

THE
IRANIAN
LANGUAGES

Edited by
Gernot Windfuhr

 **Routledge**
Taylor & Francis Group
LONDON AND NEW YORK

First published 2009

by Routledge

2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN, UK

Simultaneously published in the USA and Canada

by Routledge

270 Madison Avenue, New York, NY 10016

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2009 Gernot Windfuhr

Typeset in Times New Roman by

RefineCatch Limited, Bungay, Suffolk

Printed and bound in Great Britain by

The MPG Book Group, Bodmin, Cornwall

All rights reserved. No part of this book may be reprinted or reproduced or utilized in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data

A catalog record for this book has been requested

ISBN13: 978-0-7007-1131-4 (hbk)

ISBN13: 978-0-203-64173-6 (ebk)

BRIEF CONTENTS

<i>List of illustrations</i>	lix
<i>Notes on contributors</i>	lxv
<i>Foreword</i>	lxix
<i>List of abbreviations</i>	lxxi
1 Introduction to <i>The Iranian Languages</i>	1
<i>Gernot Windfuhr</i>	
2 Dialectology and Topics	5
<i>Gernot Windfuhr</i>	
3 Old Iranian	43
<i>Prods Oktor Skjærvø</i>	
4 Middle West Iranian	196
<i>Prods Oktor Skjærvø</i>	
5 Sogdian	279
<i>Yutaka Yoshida</i>	
6 Khwarezmian	336
<i>Desmond Durkin-Meisterernst</i>	
7 Khotanese and Tumshuqese	377
<i>Ronald E. Emmerick†</i>	
8 Persian and Tajik	416
<i>Gernot Windfuhr and John R. Perry</i>	
9 Zazaki	545
<i>Ludwig Paul</i>	
10 Kurdish	587
<i>Ernest N. McCarus</i>	
11 Balochi	634
<i>Carina Jahani and Agnes Korn</i>	
12 Parachi	693
<i>Charles M. Kieffer</i>	
13 Pashto	721
<i>Barbara Robson and Habibullah Tegey†</i>	
14a The Pamir Languages	773
<i>D. (Joy) I. Edelman and Leila R. Dodykhudoeva</i>	

14b Shughni	787
<i>D. (Joy) I. Edelman and Leila R. Dodykhudoeva</i>	
15 Wakhi	825
<i>Elena Bashir</i>	
Notes	859
Acknowledgements	859
Bibliography and references	860
Index	863

FULL CONTENTS

<i>List of illustrations</i>	lix
<i>Notes on contributors</i>	lxv
<i>Foreword</i>	lxix
<i>List of abbreviations</i>	lxxi
1 Introduction to <i>The Iranian Languages</i> Gernot Windfuhr	1
1 Overview	1
Bibliography	3
2 Dialectology and topics Gernot Windfuhr	5
1 Introduction	5
1.1 Origins: The Central Asian component	5
1.2 Ancient non-Iranian contact languages	8
1.3 Listing of Iranian languages	9
1.4 Predecessors of Modern Iranian languages	15
1.5 Modern non-Iranian contact languages	15
2 Phonology	18
2.1 Early Iranian dialects	18
2.2 West Iranian consonantal developments	18
2.3 Innovations: SW drift vs. NW	20
2.4 East and West Iranian	21
2.5 Stress	21
3 Morphology and syntax	22
3.1 Gender and animacy	22
3.2 Cases and personal enclitics	23
3.3 Deixis	23
3.4 The verbal quincunx system	24
3.5 Markers of aspect	25
3.6 Present marker <i>*-ant</i>	26
4 Syntax	27
4.1 Word order typology: adjectival noun phrase	27
4.1.1 Diachrony	28
4.1.2 Uartian and Elamite substrates	28
4.1.3 Iranian as a “buffer zone”	29
4.2 Bundling West Iranian isoglosses	29
4.3 The Iranian ergative construction	31
4.4 Differential object marking	33
4.5 Clause complementation	35
Bibliography and references	35
A Selected topical references	35
B Alphabetical	36

3	Old Iranian <i>Prods Oktor Skjærvo</i>	43
1	Introduction	43
1.1	The Old Iranian languages	43
1.1.1	The <i>Old</i> and <i>Young Avesta</i> , Old and Young Avestan	44
1.1.2	The oral background of the Avestan text	44
1.1.3	The Avestan alphabet	45
1.1.4	Stages of Avestan	46
1.1.5	Old Persian	46
1.1.6	The Old Persian script	47
1.1.7	Old Iranian grammars	47
1.2	The phonology of Indo-Iranian	48
1.2.1	Consonants	48
1.2.1.1	IIr. Velars	48
1.2.1.2	IIr. Affricates	48
1.2.1.3	IIr. <i>š ž</i>	48
1.2.1.4	IIr. The laryngeals	49
1.2.1.5	IIr. Liquids and nasals	49
1.2.2	Vowels	49
1.2.3	Proto-Indo-Iranian phoneme inventory	49
1.2.4	Ablaut	49
1.3	The phonology of Proto-Iranian	50
1.4	Early Iranian dialects	50
1.5	Writing systems	51
1.5.1	The Avestan script	51
1.5.2	The Old Persian script	52
2	Phonology	53
2.1	The phonology of Avestan	53
2.1.1	Vowel systems	54
2.1.1.1	Notes on the vowels	54
2.1.1.2	Vocalic length	54
2.1.1.3	Diphthongs	55
2.1.1.4	Hiatus	55
2.1.1.5	IIr. <i>r</i>	55
2.1.1.6	Nasalised vowels	55
2.1.2	Consonant systems	56
2.1.2.1	Notes on the consonants	56
2.1.2.2	Sibilants	57
2.1.2.3	‘Shibilants’	57
2.1.2.4	Distribution	57
2.2	The phonology of Old Persian	58
2.2.1	The vowels <i>ā, ī, ū</i>	58
2.2.2	Consonants	58
2.2.2.1	Notes on the consonants	58
2.2.3	Late Old Persian	59
2.3	The morphophonology of Avestan and Old Persian	59
2.3.1	Vowels	59
2.3.1.1	Centralising of <i>a, ā > ə (ē)</i>	59
2.3.1.2	Rounding of <i>a, ā > ā̄ or ǭ</i>	60

	2.3.1.2a Old Avestan initial $\bar{\sigma}^{\circ}\bar{a}$ - < a -/ \bar{a} -	60
	2.3.1.3 Raising of a , \bar{a} > e	60
	2.3.1.4 Combined centralising and raising and rounding of a	61
	2.3.1.5 Shortening of \bar{a} and lengthening of a	61
2.3.2	Consonants	62
	2.3.2.1 Assimilation and dissimilation	62
	2.3.2.1a Voice assimilation	62
	2.3.2.1b ‘Bartholomae’s law’	62
	2.3.2.1c Geminates	63
	2.3.2.1d Dissimilation	63
	2.3.2.2 Spirantisation	63
	2.3.2.2a Unvoiced stops and the unvoiced fricatives <i>f θ x</i>	63
	2.3.2.2b Voiced stops <i>b d g</i> and voiced spirants $\beta \delta \gamma$	63
	2.3.2.3 Consonant groups and anaptyxis	64
	2.3.2.3a Groups at morpheme boundary	65
	2.3.2.3b Anaptyxis in Old Persian	65
	2.3.2.4 Palatalisation and labialisation of consonants	65
	2.3.2.4a <i>i</i> - and <i>u</i> -epenthesis in Avestan	65
	2.3.2.4b Dissimilation of $\bar{i}\bar{i}$ and $u\bar{u}$ to $\bar{e}i$, $a\bar{i}$, and $\bar{e}u$	65
	2.3.2.4c Epenthesis in Old Persian	66
	2.3.2.5 Glides	66
	2.3.2.6 Simplification of consonant groups	66
2.3.3	Sandhi	67
	2.3.3.1 Anaptyxis in sandhi	67
	2.3.3.2 Final vowels and diphthongs in sandhi	67
2.3.4	Stress-related alternations	67
2.3.5	Metro-phonology	68
	2.3.5.1 Disyllabic long vowels and diphthongs	68
	2.3.5.2 Siever’s law	69
2.3.6	Phonological changes associated with inflection and endings	69
3	Morphology i: Nominal morphology	70
3.1	Nouns	71
	3.1.1 Gender	71
	3.1.2 Number and case	71
	3.1.3 Stem classes and declensions	71
	3.1.3.1 Derived feminine adjectives and nouns	72
	3.1.4 Protero-, hystero- and holo-kinetic declensions	72
	3.1.5 Case endings	72
	3.1.5.1 The <i>zero</i> ending	73
	3.1.5.2 Endings in <i>b</i> -	74
	3.1.5.3 Added $-\bar{a}$	75
	3.1.6 Individual declensions	75
	3.1.6.1 The <i>i</i> - and <i>u</i> -stems	75
	3.1.6.2 Monosyllabic and polysyllabic <i>ai</i> - and <i>au</i> -stems	75
	3.1.6.3 Holo-kinetic stems	76
	3.1.6.4 <i>r</i> -stems	76

3.1.6.5	Stem-formants containing <i>n</i>	77
3.1.6.5a	Stems in <i>-an-</i>	77
3.1.6.5b	Stems in <i>-ian-</i>	77
3.1.6.5c	Stems in <i>-uan-, -man-, -Han-</i>	77
3.1.6.5d	Stems in <i>-ant-</i>	78
3.1.6.5e	Stems in <i>-uant-, -mant-, -Hant-</i>	78
3.1.6.5f	Stems in <i>-Hank, -iHank-</i>	78
3.1.6.6	Stems in <i>-ah-, -iah-, -uah-</i>	78
3.1.6.7	Neuter <i>r/n</i> -stems	78
3.1.6.8	Stems in laryngeals	78
3.1.6.9	Other consonant stems	79
3.1.6.10	Suppletive stem-systems	79
3.2	Adjectives: comparative and superlative	79
3.3	Adverbs	80
3.4	Pronouns	80
3.4.1	Personal pronouns	81
3.4.2	Possessive pronouns (adjectives)	82
3.4.3	Demonstrative pronouns	82
3.4.4	Relative pronoun	82
3.4.5	Interrogative and indefinite pronouns	82
3.4.6	Reflexive pronouns and reciprocity	82
3.4.7	Pronominal case endings	83
3.5	Adpositions	83
3.6	Numerals and measure	84
3.6.1	'One' to 'four'	84
3.6.2	The remaining numerals	85
3.6.3	Derived numerals	86
3.6.3.1	Multiplicatives	86
3.6.3.2	'-th time'	86
3.6.3.3	Fractions	86
4	Morphology ii: Verbal morphology	86
4.1	Stem classes	87
4.1.1	Present stems	87
4.1.1.1	Athematic present stems	87
4.1.1.2	Thematic present stems	88
4.1.1.3	The future stem	88
4.1.1.4	Passive stem	89
4.1.1.5	Causative stem	89
4.1.1.6	Denominative stem	89
4.1.2	Aorist stems	89
4.1.3	Perfect stems	89
4.1.4	Suppletive stem systems	90
4.1.5	Periphrastic formations	90
4.2	Moods	90
4.2.1	The subjunctive	90
4.2.2	The optative	90
4.3	Voice	90
4.3.1	The passive in <i>-ia-</i>	90
4.3.2	The 3rd singular passive in <i>-i</i>	91

4.4	Person marking (endings)	91
4.5	Preverbs	93
4.6	Nominal forms of the verb	93
4.6.1	Participles	93
4.6.1.1	Present, future and aorist active participles	93
4.6.1.2	Perfect active participles	93
4.6.1.3	Middle participles	94
4.6.1.4	Past participles in <i>-ta-</i>	94
4.6.2	Verbal adjectives in <i>-0a-</i> , <i>-0βa-</i> , <i>-iia-</i>	94
4.6.3	Infinitives	94
5	Syntax i: Word order and clause structure	94
5.1	Word order	94
5.1.1	Adpositions	95
5.1.2	Preverbs	95
5.1.3	Repetition of preverb in tmesis	96
5.2	Topicalisation: 'raising' and 'lowering'	96
5.2.1	Raising of verb	97
5.2.2	Raising of direct object	97
5.2.3	Lowering of subject	97
5.2.4	Lowering of direct/indirect object	97
5.2.5	Lowering of adverbial complements	98
5.3	Special types of clauses	98
5.3.1	Questions	98
5.3.2	Exhortations and commands	99
5.3.3	Parenthetical clauses in the nominative	99
5.3.4	Reported speech	99
6	Syntax ii: Nominals	100
6.1	Specificity and reference	100
6.1.1	The Young Avestan and Old Persian connecting relative and the relative particle	100
6.1.1.1	Antecedent outside the relative phrase	101
6.1.1.1a	With invariable <i>yač</i>	101
6.1.1.2	Antecedent inside the relative phrase	101
6.2	Number	101
6.2.1	Neuter plural	101
6.2.2	Dual	102
6.3	Grammatical concord	102
6.3.1	Gender	102
6.3.2	Dual determinants	102
6.3.3	Plural subject and singular verb	103
6.4	Uses of the cases	103
6.4.1	Nominative	103
6.4.1.1	The Old Persian nominative naming phrase	104
6.4.2	Vocative	104
6.4.3	Accusative	105
6.4.3.1	Direct object and predicate of direct object	105
6.4.3.2	Direct objects with nouns and adjectives expressing actions	105
6.4.3.3	Double accusative	106

6.4.3.4	Accusative with impersonal verbs	106
6.4.3.5	Adverbial accusative	106
6.4.3.6	Accusative of length of time, space, and measure	107
6.4.3.7	Accusative 'of respect'	107
6.4.3.8	Accusative with adposition	107
6.4.3.9	Young Avestan nominative for accusative	107
6.4.4	Genitive	108
6.4.4.1	Possessive genitive	108
6.4.4.2	Subjective and objective genitive	108
6.4.4.3	Defining or descriptive genitive	109
6.4.4.5	Partitive genitive	109
6.4.4.6	Genitive with adjectives and verbs	109
6.4.4.7	Genitive with adpositions	110
6.4.5	Dative	110
6.4.5.1	Indirect object	110
6.4.5.2	Final dative	110
6.4.5.3	Dativus commodi	111
6.4.5.4	Dative with nouns and adjectives	111
6.4.5.5	Dative of agent	111
6.4.6	Ablative	112
6.4.6.1	Ablative of separation	112
6.4.6.2	Ablative of comparison, difference, and preference	112
6.4.6.3	Ablative of time or place throughout which	112
6.4.6.4	Ablative with adpositions	113
6.4.6.5	Young Avestan ablative = genitive	113
6.4.7	Instrumental	113
6.4.7.1	Instrumental of means	113
6.4.7.1a	Instrumental of space through which	114
6.4.7.1b	Instrumental of reason and cause	114
6.4.7.1c	Instrumental of respect	114
6.4.7.1d	Measure and price	114
6.4.7.2	Instrumental of accompaniment and dissociation/ deprivation	115
6.4.7.2a	Accompaniment	115
6.4.7.2b	Dissociation/deprivation	115
6.4.7.3	Instrumental with adpositions	115
6.4.7.4	Case replacements	115
6.4.8	Locative	115
6.4.8.1	Locative of place and time	116
6.4.8.2	Locative of prize won	116
6.4.8.3	Locative with adpositions	116
6.5	Uses of pronouns and pronominal forms	116
6.5.1	Personal pronouns	116
6.5.1.1	1st and 2nd persons, tonic forms	116
6.5.1.2	3rd person	117
6.5.2.2a	<i>ha-</i> , <i>ta-</i>	117
6.5.2.2b	The oblique pronominal stem <i>a-</i>	117
6.5.2.2c	<i>i-</i> , <i>hi-lši-</i> , <i>di-</i>	117
6.5.1.3	Enclitic pronouns 1st, 2nd, and 3rd person	118

	6.5.1.3a Nominative	118
	6.5.1.3b Accusative	118
	6.5.1.3c Genitive-dative	119
	6.5.1.3d Ablative	119
	6.5.2 Possessive pronouns	119
	6.5.3 Demonstrative pronouns	119
	6.5.3.1 The near-deictic <i>ima-la-</i>	120
	6.5.3.2 The near deictic <i>aēša-laēta-</i>	120
	6.5.3.3 The demonstrative pronoun <i>huuō, ana-</i>	121
	6.5.3.4 The far-deictic <i>auua-</i>	121
	6.5.4 Reflexive pronouns	122
	6.5.4.1 Use of <i>tanū-</i> ‘body’ as ‘self’	122
	6.5.5 Reciprocal expressions and ‘other’	123
	6.5.6 Interrogative pronouns	123
	6.5.7 Indefinite pronouns	123
6.6	Uses of the numerals	124
6.7	Uses of comparative and superlative	124
	6.7.1 Comparative	124
	6.7.1.1 Comparative + positive in contrasting pairs	125
	6.7.1.2 Comparative in compounds	125
	6.7.2 Superlative	125
7	Syntax iii: The verb	126
	7.1 Tenses	126
	7.1.1 Present indicative	126
	7.1.1.1 Av. present indicative with ‘before’ = past	127
	7.1.1.2 Young Avestan, Old Persian present <i>asti</i> for imperfect	128
	7.1.1.3 Old Persian present for continuing state	128
	7.1.2 Present injunctive and imperfect	128
	7.1.2.1 Present injunctive	128
	7.1.2.2 Narrative past	128
	7.1.2.3 Negation and prohibition	129
	7.1.3 Future in <i>-Sija-</i>	130
	7.1.4 Aorist	130
	7.1.4.1 Aorist indicative	130
	7.1.4.2 Aorist injunctive	130
	7.1.4.2a Aorist injunctive expressing anteriority	131
	7.1.4.2b Aorist injunctive expressing future	131
	7.1.4.2c Contrasting aspects of present and aorist injunctive	131
	7.1.4.2d Mythical events	131
	7.1.4.3 Negated aorist injunctive to express prohibition	132
	7.1.4.4 Young Avestan and Old Persian aorist	132
	7.1.5 Perfect	132
	7.1.5.1 Perfects without present	132
	7.1.5.2 Perfect expressing resulting state	133
	7.1.5.4 The past perfect/pluperfect	133
	7.1.6 The verb ‘to be’ and noun clauses	133
	7.1.6.1 The copula	134

7.1.6.2	The existential verb	134
7.1.6.3	Possession	134
7.2	Moods	134
7.2.1	Subjunctive present and aorist	135
7.2.1.1	The subjunctive in main clauses	135
7.2.1.1a	1st person	135
7.2.1.1b	2nd and 3rd persons	135
7.2.2	Optative present and aorist	136
7.2.2.1	Optative in main clauses	136
7.2.2.2	Optative of past action in Young Avestan and Old Persian	137
7.2.3	Imperative	138
7.3	Middle, passive, causative	139
7.3.1	Middle	139
7.3.2	Passive	139
7.3.2.1	Passive with agent	140
7.3.2.2	Old Persian alternatives to passive with agent	140
7.4	Participles	141
7.4.1	Present, future, aorist, and perfect participles	141
7.4.1.1	Present participles	141
7.4.1.2	Other participles	142
7.4.2	Past participles in <i>-ta-</i>	142
7.4.2.1	Medial-reflexive function of the participle in <i>-ta-</i>	143
7.4.2.2	Past participle with copula	143
7.4.3	The Old Persian perfect	144
7.4.3.1	Without agent	144
7.4.3.2	With agent	145
7.4.4	The Old Persian potential construction	145
7.5	Infinitives	145
7.5.1	Infinitive with intentional verbs	145
7.5.2	Infinitive with verbs of speaking and thinking	146
7.5.3	Independent use	146
7.6	Verbal adjectives of necessity/gerundives	146
8	Syntax iv: complex sentences	147
8.1	Coordination	147
8.1.1	Parataxis	147
8.1.2	Repetition	147
8.1.3	Conjunctions	149
8.1.3.1	Coordination	149
8.1.3.2	Disjunction	150
8.1.4	Coordinating particles	150
8.1.4.1	Sentence-introductory <i>aṭ, ā"ṭ</i>	150
8.1.4.2	Emphatic particles	152
8.1.4.3	<i>zī</i> 'for'	152
8.1.4.4	Responsive <i>bā, bōiṭ</i>	153
8.1.4.5	OAv. <i>θβaṭ</i> , YAv. <i>bāδa</i> '*sometimes, *again and again'	153
8.1.4.6	Interjections	153

8.2	Subordination	153
8.2.1	Substantial that-clauses	153
8.2.2	Relative clauses	154
8.2.2.1	The use of moods and tenses in relative clauses	155
8.2.2.2	Old Avestan relative clauses	155
	8.2.2.2a Relative clause precedes the main clause	155
	8.2.2.2b Relative clause follows the main clause	156
	8.2.2.2c Case attraction of relative pronoun	157
8.2.2.3	Young Avestan relative clauses	158
	8.2.2.3a Relative clause precedes main clause	158
	8.2.2.3b Relative clause follows the main clause	158
	8.2.2.3c <i>asti ya-</i> 'there is/some of which'	158
8.2.2.4	Old Persian relative clauses	159
8.2.2.5	Indefinite relative clauses	159
8.2.2.6	Conditional relative clauses	159
8.2.3	Adverbial clauses	160
8.2.3.1	Clauses of time	160
	8.2.3.1a Old Avestan	160
	8.2.3.1b Young Avestan	160
	8.2.3.1c Old Persian	161
8.2.3.2	Clauses of place	162
8.2.3.3	Clauses of purpose	162
8.2.3.4	Clauses of result	162
8.2.3.5	Clauses of cause	163
8.2.3.6	Clauses of manner, degree and comparison	163
8.2.4	Conditional clauses	164
8.2.4.1	Conditional clauses in Old Avestan	164
8.2.4.2	Conditional clauses in Young Avestan	165
	8.2.4.2a <i>yezi</i> -clauses: simple conditions	165
	8.2.4.2b <i>yezi</i> -clauses: hypothetical conditions	165
	8.2.4.2c Conditions contrary to fact	166
	8.2.4.2d <i>yāt</i> -clauses	166
	8.2.4.2e Periphrasis with present participle	166
8.2.4.3	Conditional clauses in Old Persian	166
9	Lexis	167
9.1	Word Formation	167
9.1.1	Derivation	167
	9.1.1.1 Suffixes	167
	9.1.1.2 Prefixation	168
	9.1.1.2a Preverbs used as prefixes	168
	9.1.1.2b Nominal prefixes	168
9.1.2	Compounding	168
	9.1.2.1 Form of first and second members of compounds	168
	9.1.2.2 Adverbial <i>āmređita</i> compounds	169
	9.1.2.3 'Dual' or 'open <i>dvandvas</i> '	169
10	Stylistic features	169
10.1	Old Avestan poetic word order	169
10.2	Parallelism and chiasmus	170
	10.2.1 Parallelism	170

10.2.2	Chiasmus	171
10.2.3	Combined parallelism and chiasmus	171
10.3	Ellipsis in Old Avestan poetry	171
10.4	Formulaic diction in Old Persian	172
10.5	Figura etymologica	173
11	Sample texts	174
11.1	Old Avestan	174
11.2	Young Avestan	179
11.3	Old Persian	185
	Abbreviations of texts	189
	Avestan texts	189
	Old Persian texts	190
	Bibliography and references	190
	Linguistic descriptions	190
	Bibliographies	190
	Further readings	195
4	Middle West Iranian <i>Prods Oktor Skjærvø</i>	196
1	Introduction	196
1.1	Overview	196
1.2	Writing Systems	197
1.2.1	Merger of letter shapes	198
2	Phonology	200
2.1	Inventory and distribution	200
2.1.1	Vowels	200
2.1.2	Consonants	201
2.1.3	Syllable structure and consonant groups	201
2.2	Non-segmental aspects	202
2.3	Relationship between orthography and phonology in Pahlavi	202
2.3.1	Vowels	202
2.3.2	Consonants	202
2.3.3	Historical and pseudo-historical orthography	203
2.3.4	Arameograms	203
2.3.5	Transliteration and transcription systems	203
3	Morphology	204
3.1	Nominal morphology	204
3.1.1	Nouns	204
3.1.1.1	Gender	204
3.1.1.2	Number and case	205
3.1.1.3	Definiteness	205
3.1.1.4	Address	205
3.1.2	Adjectives and comparison	206
3.1.3	Adverbs and particles	206
3.1.3.1	Adverbs of time, place, and manner	207
3.1.3.2	Adverbs formed by repetition	207
3.1.4	Pronouns and deixis	207
3.1.4.1	Personal pronouns	207
3.1.4.2	Possessive pronouns	208

3.1.4.3	Reciprocity and reflexivity	208
3.1.4.4	Demonstrative pronouns	209
3.1.4.5	Interrogative and indefinite pronouns and related adverbials	209
3.1.5	Adpositions	210
3.1.5.1	Nouns and adjectives as adpositions	210
3.1.6	Numerals	211
3.1.6.1	Cardinals and ordinals	211
3.1.6.2	Distributives	212
3.1.6.3	Multiplicatives	212
3.1.6.4	Fractions	212
3.1.6.5	Quantity	212
3.2	Verb morphology	213
3.2.1	Stem formation and semantic modification	213
3.2.1.1	Present and past stems	213
3.2.1.2	Causative and denominative verbs	213
3.2.1.3	Compound verb phrases	214
3.2.1.4	Preverbs	214
3.2.2	Nominal verb forms	215
3.2.2.1	Nominal forms derived from the present stem	215
3.2.2.1a	Agent noun in <i>-āg</i>	215
3.2.2.1b	Participle in <i>-andag</i>	215
3.2.2.1c	Active participle in <i>-ān</i>	215
3.2.2.1d	Participle of necessity in <i>-išn</i>	215
3.2.2.2	The verbal adjective in <i>-ag</i> from the past stem	215
3.2.2.3	Infinitives	215
3.2.3	'To be'	216
3.2.3.1	'To be'	216
3.2.3.2	Enclitic forms	216
3.2.4	On the endings	217
3.2.4.1	Notes on the present endings	217
3.2.5	System of tenses and moods	218
3.2.6	Intransitive and transitive	219
3.2.6.1	Intransitive-transitive verbs	220
3.2.6.2	Morphological passive	220
3.2.6.3	Periphrastic passive	221
4	Syntax and use of forms	221
4.1	Coordination and negation	221
4.2	Noun phrase structure	221
4.2.1	Modification of nouns by adjectives and nouns	221
4.2.2	Demonstrative pronouns and numerals	223
4.2.3	Possessive pronouns	223
4.2.4	Periphrastic adnominal constructions	223
4.2.5	The use and word order of adpositions	224
4.2.5.1	Enclitic and relative pronouns governed by — adpositions	224
4.2.5.2	Adpositions governing complex noun phrases	225
4.3	Semantics and use of verb forms	226
4.3.1	'To be'	226

4.3.1.1	'To be' and the existential verb	226
4.3.1.2	'Belonging to'	226
4.3.1.3	'To have'	226
4.3.2	Agreement between (grammatical) subject and verb	227
4.3.3	The ergative construction	227
4.3.3.1	Reflexive pronouns in ergative constructions	229
4.3.4	The use of the tenses	229
4.3.4.1	Present indicative	229
4.3.4.2	Imperfect and past imperfect	229
4.3.4.3	Preterite and past preterite	230
4.3.4.4	Present and past perfect	231
4.3.5	Passive and agency	232
4.3.5.1	Instrument	232
4.3.5.2	Passive agent	232
4.3.5.3	Unspecified agent	232
4.3.6	Direct and indirect object marking	233
4.3.6.1	Unmarked and adpositional direct object	233
4.3.6.2	Direct object marked by the preposition <i>ō</i>	233
4.3.6.3	Indirect object	233
4.3.6.4	Indirect object and possessor raising	234
4.3.7	The use of the moods	234
4.3.7.1	Subjunctive	234
	4.3.7.1a Main clauses	234
	4.3.7.1b Subordinate clauses	235
4.3.7.2	Optative	236
	4.3.7.2a Impossible conditions (irrealis)	236
	4.3.7.2b Irrealis in questions	237
	4.3.7.2c The 'parabolic' optative	237
4.3.7.3	Imperative	238
	4.3.7.3a The 'pseudo-imperative' <i>avar</i> , <i>avar-ēd</i>	239
4.3.8	Aspectual and modal particles	239
4.3.8.1	<i>hamē</i>	239
4.3.8.2	<i>ēw</i>	239
4.3.8.3	<i>bē</i>	240
4.3.8.4	<i>ōh</i>	240
4.3.8.5	<i>kāč</i>	240
4.3.9	Impersonal constructions	241
4.3.9.1	Impersonal verbs	241
4.3.9.2	Impersonal constructions with adjectives	242
4.3.10	The use of verbal nouns and adjectives	242
4.3.10.1	The participle of necessity in <i>-išn</i>	242
4.3.10.2	Infinitive constructions	243
	4.3.10.2a Infinitive as subject	243
	4.3.10.2b Infinitive as direct object	243
	4.3.10.2c Infinitive with adpositions	244
4.3.10.3	Participial constructions	244
	4.3.10.3a Present agent noun in <i>-āg</i> .	244
	4.3.10.3b Present verbal adjective (participle) in <i>-andag</i>	245

	4.3.10.3c Active participle in <i>-ān</i>	245
	4.3.10.3d Past verbal adjective in <i>-ag</i>	245
	4.3.10.3e Past participle	246
5	Complex sentences	246
5.1	Clause structure and word order	246
5.1.1	Word order in ergative constructions	247
5.1.1.1	Sequence of enclitics	247
5.1.1.2	Fronted verb	247
5.1.1.3	Agent resumption	248
5.1.1.4	Fronted agent and resumptive pronoun	249
5.1.2	Questions	249
5.2	Coordination and connectors	250
5.2.1	Sentence connectors	250
5.2.2	Coordination and disjunction	251
5.2.2.1	Coordination	251
5.2.2.2	Adversative conjunctions	252
5.2.2.3	Correlative negation	253
5.2.2.4	Coordination and negation by repetition	253
5.2.2.5	Disjunction	253
5.2.2.6	Restriction	254
5.2.2.7	Parataxis	254
5.2.2.8	Particles introducing exegetical passages	254
5.3	Subordination	255
5.3.1	Conjunctions	255
5.3.2	Relative clauses	255
5.3.2.1	Generic statements	256
5.3.2.2	Complex relative clauses	256
5.3.3	Nominal clauses	256
5.3.3.1	Reported speech	257
5.3.4	Adverbial clauses	257
5.3.4.1	Adverbial phrases	258
5.3.5	'Subordination' by means of derived nouns and adjectives	259
5.3.6	Dating	260
6	Lexis: word formation	260
6.1	Nominal derivation	260
6.1.1	Prefixation	260
6.1.2	Suffixation	260
6.1.2.1	Nouns	261
6.1.2.2	Action nouns in <i>-išn</i>	261
6.1.2.3	Agent nouns in <i>-tār/-dār</i>	261
6.1.2.4	Abstract nouns in <i>-ih</i>	261
6.1.2.4a	Quotational <i>-ih</i>	262
6.1.2.5	Nouns and adjectives	262
6.1.2.6	Adjectives	262
6.2	Compounding	263
6.2.1	Nouns	263
6.2.2	Adjectives	263

7	Lexicon and styles	263
7.1	Dialectology and loans	263
7.2	Religious terminology	263
7.3	Sociolinguistic aspects: styles and levels of speech	264
8	Sample texts	265
8.1–2	Middle Persian and Parthian Inscriptions of Šābuhr I at Hajiabad	265
8.3	Pahlavi Psalter, from <i>Psalm</i> 127	268
8.4	Pahlavi <i>Bundahišn</i> (Bd.33.0–4)	269
8.5	Pahlavi <i>Dēnkard</i> (Dk.9.45.4)	270
8.6	Manichaean Middle Persian	271
8.7	Manichaean Parthian poetic text	272
	Abbreviations of texts	273
	Linguistic descriptions	273
	Bibliography and references	274
5	Sogdian <i>Yutaka Yoshida</i>	279
1	Introduction	279
1.1	Overview	279
1.1.1	History	279
1.1.2	Materials	280
1.2	Scripts, orthography, and phonetic values	281
2	Phonology	284
2.1	Inventory and distribution	284
2.1.1	Vowels and diphthongs	284
2.1.1.1	Vowel system	284
2.1.1.2	Rhotacised vowels	285
2.1.1.3	Diphthongs	285
2.1.2	Consonants	285
2.1.2.1	Consonantal system	285
2.1.2.2	Marginal <i>l</i> , <i>h</i> , <i>ts</i> , and <i>ŋ</i>	286
2.1.3	Syllable structure and clusters	286
2.2	Non-segmental features and rhythmic law	286
2.3	Alternations and combinatory phenomena	287
3	Morphology	287
3.1	Nominal morphology	287
3.1.1	Stem classes and declensions	288
3.1.1.1	Light stems	288
3.1.1.2	Heavy stems	289
3.1.1.3	Irregular plural forms, <i>-ar-t</i> , <i>-īš-t</i>	289
3.1.1.4	<i>aka-</i> and <i>ākā-</i> stems	289
3.1.1.5	Indeclinables and archaic plural <i>-ān</i>	290
3.1.2	Adjectives	290
3.1.2.1	Degree	290
3.1.2.2	Elativ	290
3.1.3	Pronouns and deixis	290
3.1.3.1	Personal pronouns	290
3.1.3.2	Reflexive and reciprocal forms	291

3.1.3.3	Articles and demonstrative pronouns and adverbials	291
3.1.3.4	Interrogative, relative, and indefinite pronouns	293
3.1.4	Adpositions	293
3.1.4.1	Fused prepositions	293
3.1.4.2	Postpositions and circumpositions	294
3.1.5	Adverbs and interjections	294
3.1.5.1	Adverbs	294
3.1.5.2	Interjections	294
3.1.6	Numerals	295
3.1.6.1	Cardinal numbers	295
3.1.6.2	Distributives, multiples, and fractions	295
3.1.6.3	Ordinal numbers	295
3.2	Verbal morphology	295
3.2.1	Stem formation	295
3.2.1.1	Present and past stems	295
3.2.1.2	Imperfective stems	296
3.2.1.3	Derived voice stems	296
3.2.2	Nominal forms	296
3.2.2.1	Infinitives and verbal nouns	296
3.2.2.1a	Past infinitive	296
3.2.2.1b	Present infinitives	296
3.2.2.1c	Verbal noun	297
3.2.2.2	Gerund and participles	297
3.2.2.2a	Gerund	297
3.2.2.2b	Present participles	297
3.2.2.2c	Passive participles	297
3.2.3	System of tenses, aspects, and moods	298
3.2.3.1	Forms based on the present and imperfect stems	298
3.2.3.1a	Major conjugations	298
3.2.3.1b	New minor and secondary conjugations	299
3.2.3.1c	Progressive or durative particle: <i>skun</i> , archaic <i>-əṣṭan</i>	300
3.2.3.1d	Future and past prospective particle: <i>kām</i>	300
3.2.3.2	Copula and existential verb	300
3.2.3.3	Forms based on the past stem	301
3.2.3.3a	Preterite	301
3.2.3.3b	Ergative construction	302
3.2.3.3c	Potentialis system	302
3.2.3.4	Forms based on the past participle	302
3.2.3.4a	Periphrastic perfect	302
3.2.3.4b	Periphrastic passive	303
3.2.3.5	Voice and causativity	303
3.2.3.5a	Passive and middle	303
3.2.3.5b	Causativity	304
3.2.4	Negation	304
3.2.5	Synopsis of the system of tense, aspect, and mood	304
4	Syntax and use of forms	305
4.1	Use of cases	305

4.1.1	Light stems	305
4.1.2	Heavy stems	306
4.1.3	Direct object and animacy hierarchy	306
4.2	Use of verb forms	307
4.2.1	Possession	307
4.2.2	Use of tense and aspect	307
4.2.2.1	Examples	308
4.2.2.2	Imperfect and preterite	309
4.2.3	Use of modal forms	310
4.2.3.1	Functions	310
4.2.3.2	Merger of modal categories: non-indicative	310
4.2.3.3	Modal particles	310
4.2.3.3a	Hypothetical particle <i>xāt</i>	310
4.2.3.3b	Hypothetical particle <i>-(a)n</i>	311
4.2.3.4	Examples	311
4.2.4	Modal constructions	312
4.2.4.1	Necessity construction	312
4.2.4.2	Desiderative construction	312
4.3	Concord and agreement	312
4.3.1	Light and heavy stems	312
4.3.2	Group inflexion and nouns after cardinal numbers	313
4.3.2.1	Group inflexion	313
4.3.2.2	Nouns after cardinal numbers	313
4.3.3	Agreement of the subject and the predicate	313
5	Syntax of simple and complex sentences	314
5.1	Order of constituents	314
5.1.1	Noun phrase structure	314
5.1.2	Clause structure and word order	314
5.1.2.1	Complementizer	314
5.1.2.2	Enclitics	315
5.1.2.3	Word order	315
5.1.2.3a	SOV	315
5.1.2.3b	Counter-examples	316
5.2	Question	316
5.3	Complex sentences	317
5.3.1	Conjunctions	317
5.3.2	Relative clauses	318
5.3.2.1	Correlatives and gaps	318
5.3.2.2	Other relatives	318
5.3.2.3	Indefinite relative clauses	319
5.3.3	Adverbial clauses	319
5.3.4	Nominal clauses and direct quotation	320
5.3.4.1	Nominal clauses	320
5.3.4.2	Direct quotations	320
5.3.5	Use of infinitives and verbal nouns	320
5.3.5.1	Past infinitives	320
5.3.5.2	Present infinitives and verbal nouns	321
6	Lexis	321
6.1	Sogdian lexicon	321

6.2	Derivation	322
6.2.1	Suffixes	322
6.2.2	Prefixes	322
6.3	Composition	323
6.4	Collocation	323
6.4.1	Hendiadys	323
6.4.2	Phrasal verbs or idioms	323
7	Sociolinguistic aspects	324
7.1	Polite phraseology and signals	324
7.1.1	Honorific expressions	324
7.1.1.1	Verbs	324
7.1.1.2	Nouns and other expressions	324
7.1.1.3	Expressions showing the speaker's humility	325
7.2	Dialects and variations	325
7.2.1	Archaic and late features	325
7.2.1.1	Ancient letters: oldest stage	325
7.2.1.2	Christian Sogdian text C5: latest stage	326
7.2.2	Sogdian dialects?	326
7.2.3	Sogdian and Yaghnobi	327
7.3	Language contact and loanwords	327
7.3.1	Borrowings	327
7.3.1.1	Contact with Bactrian or Kushan influence	327
7.3.1.2	Loanwords from West Middle Iranian	328
7.3.1.3	Indian loanwords	328
7.3.1.4	Borrowing from other languages	328
7.3.2	Sogdian elements in other neighbouring languages	329
7.4	Language death	329
7.4.1	Death of Sogdian: Turkicization and Persianization	329
7.4.1.1	Turkicization	329
7.4.1.2	Persianization	330
8	Sample text	330
	Acknowledgement	333
	Bibliography and references	333
6	Khwarezmian <i>Desmond Durkin-Meisterernst</i>	336
1	Introduction	336
1.1	Overview	336
1.2	Writing systems	338
1.3	A note on transcription	338
2	Phonology	339
2.1	Inventory and distribution	339
2.1.1	Vowels and diphthongs	339
2.1.1.1	Vowels	339
2.1.1.2	Semivowels and diphthongs	339
2.1.1.3	Nasalization	340
2.1.2	Consonants	340
2.1.3	Syllable structure	341
2.2	Non-segmental features	341
2.2.1	Stress and unstable vowels	341

2.1.1.1	Unstable vowels	341
2.1.1.2	Position and marking of stress	341
2.1.1.3	Unstressed lengthening	341
2.1.1.4	Pre-pausal position	341
2.3	Morphophonemic change	342
3	Morphology	342
3.1	Nominal morphology	342
3.1.1	Categories	342
3.1.1.1	Gender, number, and case	342
3.1.1.2	Animate terms	343
3.1.1.3	Dual and numerative	343
3.1.1.4	Definiteness and indefiniteness	343
3.1.1.4a	Definite article	343
3.1.1.4b	Indefiniteness	343
3.1.1.5	Nouns and inflectional classes	343
3.1.2	Adjectives and comparison	344
3.1.3	Pronouns and deixis	344
3.1.3.1	Personal pronouns	344
3.1.3.2	Demonstrative pronouns	344
3.1.3.3	Interrogative and relative pronouns	344
3.1.3.4	Reflexive and reciprocal pronominals	344
3.1.4	Adpositions	344
3.1.4.1	Pre- and postpositions	344
3.1.4.2	Circumpositions	346
3.1.4.3	Nominal inflection and adpositions	346
3.1.4.3a	Masculine singular	346
3.1.4.3b	Masculine dual/numerative	347
3.1.4.3c	Masculine plural	347
3.1.4.3d	Feminine singular	347
3.1.4.3e	Feminine plural	347
3.1.5	Adverbs and adverbial phrases	348
3.1.6	Numerals	348
3.1.6.1	Cardinal numbers	348
3.1.6.2	Ordinal numbers	348
3.1.6.3	Fractions	348
3.1.6.4	Distributives	348
3.1.6.5	Multiplicatives	348
3.2	Verb morphology	349
3.2.1	Stem formation	349
3.2.1.1	Present and imperfect stems	349
3.2.1.2	Intransitive stem classes and derived causatives	349
3.2.1.3	Compound verbs	350
3.2.1.4	Directional adpositional particles	350
3.2.2	Nominal forms	351
3.2.2.1	Present stem	351
3.2.2.1a	Present participles	351
3.2.2.1b	Verbal nouns	351
3.2.2.2	Past stem	351

	3.2.2.2a Normal past participles and ‘past infinitives’	351
	3.2.2.2b Perfect participles	351
	3.2.2.3 Morphophonology of tense stems	351
	3.2.3 Person marking	352
	3.2.4 System of tense, mood, and aspect	352
	3.2.4.1 Present and imperfect forms	353
	3.2.4.1a Imperfect	353
	3.2.4.1b Injunctive	353
	3.2.4.2 Compound tenses, perfect forms	353
	3.2.4.2a Present perfect	354
	3.2.4.2b Perfect subjunctive	354
	3.2.4.2c Remote perfect	354
	3.2.4.3 Basic inflectional system	354
	3.2.4.4 Phasal and modal enclitic particles	355
	3.2.4.4a Permansive particle <i>-ina, -ī</i>	355
	3.2.4.4b Future particle <i>-kām</i>	355
	3.2.4.4c Conditional particle <i>-manc(a)</i>	356
	3.2.4.5 Overall system of tenses, moods, and aspect	356
	3.2.5 Negation	356
	3.2.5.1 Basic negation and prohibition	356
	3.2.5.2 Particle <i>fa + ne</i> and time reference	357
4	Syntax and use of forms	358
	4.1 Noun phrase structure and word order	358
	4.1.1 Noun and adjective	358
	4.1.1.1 Indefinite	358
	4.1.1.2 Definite	358
	4.1.1.3 Definite, postponed adjective	358
	4.1.2 Nouns and dependent nouns	358
	4.1.3 Nouns and demonstrative pronouns	358
	4.1.4 Nouns and pronominal suffixes	359
	4.2 Clause structure and word order	359
	4.2.1 Order of constituents	359
	4.2.1.1 Basic order	359
	4.2.1.2 Questions	359
	4.2.1.3 Initial negation	360
	4.2.2 Affixes and ‘principle of anticipation’	360
	4.2.2.1 Direct object + indirect object	360
	4.2.2.2 Adverbial suffixes	360
	4.2.2.3 Pre- and postverbal position	360
	4.2.2.4 Clause-initial particles	361
	4.2.2.5 Agency hierarchy	361
	4.3 Semantics and use of forms	362
	4.3.1 Use of number and case	362
	4.3.1.1 Impersonal plural	362
	4.3.1.2 Dative of affectee	362
	4.3.2 Use of verb forms	362
	4.3.2.1 Possession, ‘have’	362
	4.3.2.2 Phasal constructions	363

	4.3.2.2a Duration, <i>ḍāray-</i>	363
	4.3.2.2b Incipient action, <i>āyāz-</i>	363
	4.3.2.2c Ingressive action, 'about/close to', * (<i>-m-ah-</i> ; ('šxt 'β-))	363
	4.3.2.3 Modal constructions	363
	4.3.2.3a Optative particles <i>ās</i> and <i>hā(n)</i>	363
	4.3.2.3b Potential construction, <i>-k-</i>	364
	4.3.2.3c Necessity, <i>γuw-</i>	364
5	Complex sentences	364
	5.1 Conjunctions, particles, and clitics	364
	5.2 Subordination	364
	5.3 Relative clauses, <i>ki</i>	365
	5.4 Subject and object clauses	365
	5.4.1 Subject clauses, <i>cā</i>	365
	5.4.2 Object clauses, <i>ka</i>	365
	5.4.2.1 Basic pattern	365
	5.4.2.2 Direct speech	365
	5.4.2.3 Topicalized object clause	365
	5.5 Adverbial clauses	366
	5.5.1 Temporal clauses, <i>ka</i>	366
	5.5.1.1 'at time that, when'	366
	5.5.1.2 'as soon as'	366
	5.5.1.3 'while'	366
	5.5.1.4 'when(ever)'	367
	5.5.2 Conditional clauses, <i>ka . . . da</i>	367
	5.5.2.1 Basic patterns	367
	5.5.2.2 Implicit conditional, imperfect	368
	5.5.2.3 Wishes, <i>hān</i>	368
	5.5.3 Causal clauses, <i>ka</i>	368
	5.5.4 Final clauses	369
	5.5.4.1 Basic patterns, <i>ka</i>	369
	5.5.4.2 'until, so that', <i>camend</i>	369
	5.6 Nominalized clauses	370
6	Lexis	370
	6.1 Word formation	370
	6.1.1 Derivation of nouns and adjectives	370
	6.1.1.1 Suffixation	370
	6.1.1.1a Verbal nouns and infinitives	370
	6.1.1.1b Abstract nouns	370
	6.1.1.1c Adjectives	371
	6.1.1.1d Diminutives	371
	6.1.1.2 Prefixation	371
	6.1.2 Compounding	371
	6.1.2.1 Verbal components	371
	6.1.2.2 Nominal component	371
	6.2 Sociolinguistic aspects	372
	6.2.1 Standard and colloquial registers	372
	6.2.2 Kinship terms	372
7	Language development, loan component	372

8	Sample text	373
	Acknowledgements	373
	References and bibliography	374
7	Khotanese and Tumshuqese <i>Ronald E. Emmerick</i> †	377
1	Introduction	377
1.1	Overview	377
1.1.1	Historical background	377
1.1.2	Textual material	378
1.1.2a	Khotanese	378
1.1.2b	Tumshuqese	379
1.2	Writing systems	379
1.3	Orthography	381
2	Phonology	382
2.1	Inventory and distribution	382
2.1.1a	Khotanese	382
2.1.1a.1	Vowels	382
2.1.1a.2	Consonants	382
2.1.1b	Tumshuqese	383
2.1.1b.1	Vowels	383
2.1.1b.2	Consonants	383
2.1.2	Syllable and clusters	383
2.2	Non-segmental features	383
3	Morphology	384
3.1	Nominal morphology	384
3.1.1a	Nominal stem classes and declensions—Khotanese	384
3.1.1a.1	Vocalic declensions	385
3.1.1a.2	Declensions, <i>āa</i> - and <i>āā</i> - stems	385
3.1.1a.3	Consonantal declensions	386
3.1.2a	Adjectives—Khotanese	386
3.1.1/2b	Declension of nouns and adjectives—Tumshuqese	387
3.1.3	Pronouns and deixis	387
3.1.3.1a	Personal and demonstrative pronouns— Khotanese	387
3.1.3.1b	Personal and demonstrative pronouns— Tumshuqese	388
3.1.3.2	Reflexive and reciprocal pronouns	388
3.1.3.3	Indefinite and interrogative pronouns and adverbials	388
3.1.4a	Adpositions—Khotanese	389
3.1.4a.1	Postpositions	389
3.1.4a.2	Prepositions	389
3.1.4b	Adpositions—Tumshuqese	389
3.1.5	Adverbs	389
3.1.6a	Numerals—Khotanese	390
3.1.6a.1	Cardinal numbers	390
3.1.6a.2	Ordinal numbers	390
3.1.6a.3	‘Both’ and indefinite adjectives	390

3.1.6b	Numerals—Tumshuqese	391
3.2	Verb morphology	391
3.2.1	Stem formation—Khotanese	391
3.2.1.1	Present and past stems	391
3.2.1.2	Directional particles	391
3.2.2	Nominal forms	392
3.2.2.1	Present participles	392
3.2.2.2	Past participles	392
3.2.2.3	Infinitives	392
3.2.2.4	Gerundives and gerunds	393
3.2.3	Person marking and ‘to be’	393
3.2.4	Conjugation classes	394
3.2.5	Negation	395
3.2.6	System of tenses	395
3.2.6.1	Present and past	395
3.2.6.2	Periphrastic perfect and pluperfect	396
3.2.7	Modal and aspectual constructions	396
3.2.7.1	Potential-completive construction and passive	396
3.2.7.1a	Ability	397
3.2.7.1b	Completed action	397
3.2.7.1c	Passive	397
3.2.7.2	Infinitive-based constructions	398
3.2.7.3	Gerundives, gerunds, and necessity construction	398
3.2.8	Diathesis	398
4	Syntax and use of forms	399
4.1	Coordination	399
4.1.1	Coordinating conjunctions and clitics	399
4.1.2	Enclitic particles and words	399
4.2	Noun phrase structure	399
4.2.1	Concord and agreement	399
4.2.1.1	Adjectives	399
4.2.1.2	Group inflection	399
4.2.1.3	Demonstrative adjectives	400
4.2.2	Dependent nouns	400
4.2.3	Extended noun phrase	400
4.3	Clause structure and word order	400
4.3.1	Nominal sentence	401
4.3.2	Order of constituents	401
4.3.3	Enclitic pronouns	401
4.3.4	Topicalisation	402
4.3.5	Questions	402
4.4	Semantics and use of nominal forms	403
4.4.1	Use of cases	403
4.4.1.1	Genitive-dative	403
4.4.1.2	Locative	403
4.5	Semantics and use of verb forms	404
4.5.1	Possession	404
4.5.2	Tense and aspect	404
4.5.2.1	Future	404

4.5.2.2	Accomplishment of past action	404
4.5.2.3	Immediacy	404
4.5.3	Use of modal forms	404
5	Complex sentences	405
5.1	Quotative clitic	406
5.2	Relative clauses	406
5.2.1	Indicative, factual	406
5.2.2	Relative pronoun in genitive-dative	406
5.2.3	Injunctive, potential	406
5.2.4	Clauses with head noun	406
5.3	Adverbial clauses	406
5.3.1	Temporal and local clauses	406
5.3.1.1	Temporal	406
5.3.1.2	Local	407
5.3.2	Conditional clauses	407
5.3.3	Consecutive and comparative clauses	407
5.4	Nominalised clauses	407
5.4.1	Infinitival constructions	407
5.4.2	Participial-adjectival clauses	408
6	Lexis and sociolinguistic aspects	408
6.1	Word formation	408
6.1.1	Derivation	408
6.1.1.1	Suffixation	408
6.1.1.2	Prefixation	408
6.1.2	Composition	409
6.1.3a	Loanwords—Khotanese	409
6.1.3a.1	Iranian	409
6.1.3a.2	Indian	409
6.1.3b	Loanwords—Tumshuqese	410
6.2	Sociolinguistic aspects	410
7	Dialects and language development	410
7.1a	Khotanese	411
7.1b	Tumshuqese	411
7.2	Demise of the languages	411
8	Sample Khotanese text	411
	Acknowledgements	412
	Bibliography and references	412
8	Persian and Tajik <i>Gernot Windfuhr and John R. Perry</i>	416
1	Introduction	416
1.1	Overview	416
1.1.1	Historical background	416
1.1.2	Persian and related groups	417
1.1.2.1	Varieties of Persian	417
1.1.2.2	Related varieties and dialect groups	418
1.1.2.3	Number of speakers	418
1.1.3	Evolution of Standard Persian	419
1.1.3.1	Early linguistic groups	419

1.1.3.2	Social dynamics	419
1.2	Tajik	420
1.3	Writing systems	421
1.3.1	Persian	421
1.3.2	Tajik	423
2	Phonology	423
2.1	Inventory	423
2.1.1	Vowels	425
2.1.1a	Persian	425
2.1.1b	Tajik	426
2.1.2	Consonants	426
2.1.2.1	Stops and fricatives	426
2.1.2.2	Velar γ and uvular q	427
2.1.2.3	Glottal stop	427
2.1.2.4	Palatal \check{z}	427
2.1.2.5	$v \sim w$	427
2.1.2.6	Lateral l	427
2.1.2.7	Final position	427
2.1.3	Syllable structure	427
2.1.4	Clusters	428
2.1.4.1	Initial clusters	428
2.1.4.2	Final clusters	428
2.1.4.3	Gemination	428
2.1.4.4	Glottal stop, h and r	428
2.1.5	Hiatus	429
2.1.5.1	Standard Persian and Tajik	429
2.1.5.2	Colloquial absence of hiatus	429
2.1.5.3	Persian final $-e$ as a dialect feature	429
2.2	Non-segmental features	429
2.3	Morphophonemic alternations	430
3	Morphology	430
3.1	Nominal morphology	430
3.1.1	Nouns	431
3.1.1.1	Number	431
3.1.1.1a	$-h\bar{a}l$ - ho and $-\bar{a}nl$ - on	431
3.1.1.1b	Arabic loan plurals	431
3.1.1.2	Indefinite $-il$ - e and $yeklyak$	432
3.1.1.3	Tajik indefinite determiners	432
3.1.1.4	Persian referential $-e$	432
3.1.2	Adjectives	432
3.1.2.1	Subsets	433
3.1.2.1a	Substantivized adjectives	433
3.1.2.1b	Morphological subsets	433
3.1.2.2	Comparative	433
3.1.2.3	Superlative	433
3.1.2.4	Intensive reduplicated and echoic superlative	434
3.1.3	Pronouns and deixis	434
3.1.3.1	Personal pronouns and deixis	434
3.1.3.1a	Register and animacy	435

3.1.3.1b	Attachment of object marker	435
3.1.3.2	Possessive construction	435
3.1.3.3	Enclitic pronouns and oblique	435
3.1.3.4	Demonstratives and quasi-articles	436
3.1.3.4a	Emphatic <i>ham-</i>	436
3.1.3.5	Reflexive-emphatic pronoun	436
3.1.3.5a	Reflexive	436
3.1.3.5b	Emphatic	437
3.1.3.6	Reciprocals	437
3.1.3.7	Interrogatives, indefinites, and related adverbials	438
3.1.3.7a	Basic interrogatives and indefinite determiners	438
3.1.3.7b	Pronominal quantifiers and demonstratives	438
3.1.3.7c	Compound indefinite forms	439
3.1.3.7d	Other indefinites	439
3.1.3.7e	Compounds with <i>če-čā-</i>	440
3.1.3.7f	<i>digar/digar</i> 'other, else'	440
3.1.3.7g	Tajik hybrids	440
3.1.4	Adpositions and case relations	440
3.1.4.1	Prepositions	441
3.1.4.1a	Primary prepositions	441
3.1.4.1b	Secondary prepositions	442
3.1.4.2	Postposition <i>-rāl-ro</i>	442
3.1.4.2a	Specific direct object	442
3.1.4.2b	Temporal and local extension	443
3.1.4.2c	Tajik circumpositional <i>-ro</i>	443
3.1.4.2d	Northern Tajik EZ with <i>-ro</i>	443
3.1.4.3	Other Tajik postpositions	443
3.1.5	Adverbs	444
3.1.5.1	Nouns	444
3.1.5.2	Adverbs proper	444
3.1.5.3	Adjective-adverbs and noun-adverbs	444
3.1.6	Numerals	445
3.1.6.1	Cardinal numbers	445
3.1.6.2	Ordinal numbers	446
3.1.6.3	Fractions	446
3.1.6.4	Distributive adverbs	446
3.2	Verb morphology	446
3.2.1	Stem formation	446
3.2.1.1	Present and past stems	446
3.2.1.2	Denominal verbs	447
3.2.1.3	Derived causative verbs	448
3.2.1.4	Preverbs	448
3.2.2	Nominal forms	449
3.2.2.1	Forms derived from Stem I	449
3.2.2.1a	Present participle	449
3.2.2.1b	Active participle	449
3.2.2.2	Forms derived from Stem II	449

3.2.2.2a	Infinitive	449
3.2.2.2b	Short infinitive	449
3.2.2.2c	Past (perfect) participle	449
3.2.2.2d	Tajik participle	449
3.2.2.2e	Future participle or gerundive	449
3.2.3	Person marking and ‘be’	450
3.2.4	Aspectual-modal markers and negation	451
3.2.4.1	Prefix <i>mi-/me-</i>	451
3.2.4.2	Prefix <i>be-/bi-</i>	451
3.2.4.3	Negation <i>na-/na-</i>	452
3.2.5	Basic system of tense, aspect, and mood	452
3.2.5.1	Present-future	452
3.2.5.2	Imperfect	452
3.2.5.3	Preterite, or “aorist”	454
3.2.5.4	Present perfect	455
3.2.5.5	Pluperfect or distant past	455
3.2.5.6	Imperative	456
3.2.5.7	Present-future subjunctive	456
3.2.5.8	Past (perfect) subjunctive	457
3.2.5.9	Imperfect counterfactual	458
3.2.5.10	Pluperfect counterfactual	458
3.2.5.11	Defective verbs	459
3.2.5.11a	‘be’	459
3.2.5.11b	‘have’	460
3.2.6a	Extended verb system of Modern Standard Persian	461
3.2.6a.1	Evidential or non-witnessed mode	461
3.2.6a.2	Progressive	461
3.2.6b	Extended verb system of Modern Standard Tajik	462
3.2.6b.1	Evidential mode	462
3.2.6b.1a	Perfect as evidential form	463
3.2.6b.1b	Evidential pluperfect	463
3.2.6b.1c	Evidential durative	464
3.2.6b.1d	Evidential progressive	464
3.2.6b.2	Progressive forms	465
3.2.6b.2a	Present progressive	465
3.2.6b.2b	Past progressive	465
3.2.6b.3	Subjunctive forms	465
3.2.6b.3a	Durative perfect subjunctive	465
3.2.6b.3b	Present progressive subjunctive	466
3.2.6b.4	Conjectural mood	466
3.2.6b.4a	Past conjectural	466
3.2.6b.4b	Present-future conjectural	467
3.2.6b.4c	Present progressive conjectural	467
3.2.6b.4d	Pluperfect counterfactual conjectural	467
3.2.6b.5	Synopsis of Tajik verb forms	467
3.2.7	Synopsis of Persian and Tajik verb systems	468
4	Syntax and use of Forms I: noun phrase structure	468
4.1	Coordination	468
4.1.1	Single and paired conjunctions	468

4.1.2	Concatenation and group inflection	470
4.2	Constituents following the head: <i>Ezāfe</i>	470
4.2.1	Basic structure	470
4.2.1.1	NP levels	470
4.2.1.2	NP clitics	471
4.2.1.3	Alternate attachment of indefinite marker	472
4.2.2	Types of noun phrases and complements	473
4.2.2.1	Overview	473
4.2.2.2	Nominal heads and valence relationships	473
4.2.2.3	Possessor	474
	4.2.2.3a Preposed possessor	474
	4.2.2.3b Tajik Turkic-type possessor inversion	474
4.2.2.4	Adjectives and ordinals	475
	4.2.2.4a Superlatives and ordinals	475
	4.2.2.4b Preposed affective adjectives	475
	4.2.2.4c Pronouns as heads	475
4.2.2.5	Nominalized complements	475
	4.2.2.5a Verbal nouns as head	475
	4.2.2.5b Compound participles and adjectives	476
	4.2.2.5c Participial phrases as complements	476
	4.2.2.5d Tajik nominalized relative clauses	477
	4.2.2.5e Tajik constructions with participle in <i>-agī</i>	477
4.3	Constituents preceding the head noun	477
4.3.1	Overview	477
4.3.2	Classifiers	478
4.3.2.1	Numerative classifiers	478
4.3.2.2	Mensurative classifiers	479
4.3.2.3	Kind and manner classifiers	479
5	Syntax and use of forms II: clause structure	479
5.1	Order of constituents	479
5.1.1	OSV and OVS	480
	5.1.1.1 Object raising	480
	5.1.1.2 Concomitant subject lowering	480
5.1.2	Adverbials	480
5.1.3	Questions	481
5.1.4	Exclamations	482
5.2	Definiteness and subject and object marking	482
5.2.1	Marking of definiteness and specificity	482
	5.2.1.1 Definiteness	483
	5.2.1.2 Specificity	483
5.2.2	Marking of subjects	483
	5.2.2.1 Verbal agreement	483
	5.2.2.2 Copular predicates and ellipsis	484
	5.2.2.3 Specified and unspecified subjects	485
5.2.3	Marking of direct objects	485
	5.2.3.1 Definite direct objects	485
	5.2.3.2 Indefinite specific and non-specific direct objects	485
	5.2.3.3 Scale of specificity	485

5.3	Use of pronominal clitics	486
5.3.1	Pronominal direct and indirect objects	486
5.3.2	Indirect experience	487
5.4	Subject and object complements	488
5.5	Complex predicates	488
5.5.1	Possession	489
5.5.2	Modal constructions, verb + verb	489
5.5.2.1	Definite future	489
5.5.2.2	Volition	489
5.5.2.3	Intention and expectation	489
5.5.2.4	Obligation and presumption	490
5.5.2.4a	Obligation and necessity	490
5.5.2.4b	Presumption	491
5.5.2.5	Ability	492
5.5.2.6	Possibility	492
5.5.3	Aktionsart constructions, V + V	493
5.5.3.1	Inception, non-intentional	493
5.5.3.2	Inception	494
5.5.3.3	Inception, completive	494
5.5.3.4	Inception, ingressive	494
5.5.4	Tajik conjunct verb constructions	495
5.5.5	Light verb constructions	496
5.5.5.1	Reflexivity	498
5.5.5.2	Experiencer-subject LVCs	498
5.5.6	Voice and agency	498
5.5.6.1	Passive voice	498
5.5.6.2	Scale of agency and causation	499
6	Syntax and use of forms III: Complex sentences	500
6.1	Clause linking	500
6.1.1	Overview	500
6.1.2	Coordinating conjunctions and clitics	501
6.2	Subordinate clauses and linking	502
6.3	Relative clauses	503
6.3.1	Restrictivity and specificity	503
6.3.2	Gapping and anaphora	503
6.3.3	Ordering of restrictive RCs	504
6.3.4	Use of moods	505
6.3.5	General-indefinite RCs	506
6.3.6	Head incorporation	508
6.3.7	Tajik nominalized relative clauses	508
6.3.7.1	Levels of nominalizations	509
6.4	Complement clauses	511
6.4.1	Subject and object clauses	511
6.4.1.1	Use of TMA	512
6.4.2	Embedded speech	513
6.4.2.1	Levels of integration	513
6.4.3	Adpositional and EZ complement clauses	513
6.4.4	Tajik Turkic-type nominal clauses	514
6.4.5	“Idiomatic” clitic <i>ke</i>	515

6.5	Adverbial clauses	516
6.5.1	Position and use of TMA	517
6.5.2	Pre-matrix clauses	518
6.5.2.1	Locative clauses	519
6.5.2.2	Temporal clauses	519
6.5.2.2a	Since-clauses	519
6.5.2.3	Conditional clauses	519
6.5.2.4	Concessive clauses	521
6.5.2.5	Causal clauses	521
6.5.3	Post-matrix clauses	521
6.5.3.1	Causal-explanatory clauses	521
6.5.3.2	Temporal clauses	522
6.5.3.2a	Subordinator switch	522
6.5.3.3	Result clauses	522
6.5.3.4	Purpose clauses	522
6.5.3.5	Consecutive clauses	523
6.5.3.6	Comparative clauses	523
6.5.3.7	Conditional exception clauses	523
6.5.4	Tajik Turkic-type nominal clauses	523
6.5.4.1	Participial relative clauses in <i>-agi</i>	524
6.5.4.2	Infinitival clauses	524
6.5.4.2a	Object clauses	524
6.5.4.2b	Adverbial clauses	525
6.5.4.3	Embedded infinitival clauses	525
6.5.4.4	Persian conjunct participle	525
7	Lexis and sociolinguistics	526
7.1	Word formation	526
7.1.1	Homonymy and conversion	526
7.1.2	Suffixation	526
7.1.3	Prefixation	527
7.1.4	Compounding	528
7.2	Loanwords	529
7.2.1	Arabic	529
7.2.2	Turkic	530
7.2.3	Russian	530
7.3	Lexical distribution, Persian–Tajik	530
7.4	Sociolinguistic aspects	531
7.4.1	Registers and <i>ta'ārof</i>	531
7.4.2	Kinship terms	532
7.4.3	Modes of address	532
8	Diachrony and dialectology	533
8.1	Phases	533
8.2	Comparative dialectology	533
8.3	Innovative constructions, summary	534
8.3.1	Progressive	534
8.3.2	Uncertainty constructions	534
8.3.3	Evidentiality	534
8.4	Isoglosses West vs. East	534
8.4.1	Conjunct constructions and complementizer	534

8.4.2	Persian homoglossia	535
8.4.3	Topical references to Chapter 2	535
9	Sample texts	536
9.1	Persian	536
9.2	Tajik	538
9	Zazaki <i>Ludwig Paul</i>	545
1	Introduction	545
1.1	Overview	545
2	Phonology	546
2.1	Inventory and distribution	546
2.1.1	Vowels and diphthongs	547
2.1.1.1	Vowels	547
2.1.1.2	Diphthongs	547
2.1.2	Consonants	547
2.2	Non-segmental features	547
3	Morphology	548
3.1	Nominal morphology	548
3.1.1	Nouns	548
3.1.1.1	Gender	548
3.1.1.2	Number and case	548
3.1.1.3	Vocative	549
3.1.1.4	Definiteness and unity	549
3.1.2	Adjectives	549
3.1.2.1	Gender, case and number	550
3.1.2.2	Comparison	550
3.1.3	Pronouns and deixis	551
3.1.3.1	Personal pronouns	551
3.1.3.2	Demonstrative pronouns	551
3.1.3.3	Reflexive and possessive pronoun	551
3.1.3.4	Reciprocal pronoun	552
3.1.4	Adpositions	552
3.1.4.1	Primary adpositions	552
3.1.4.1a	Simple preposition	552
3.1.4.1b	Simple postpositions	553
3.1.4.2	Secondary adpositions	553
3.1.4.2a	Secondary prepositions	553
3.1.4.2b	Secondary postpositions	553
3.1.4.3	Circumpositions	553
3.1.5	Adverbs	553
3.1.6	Numerals	554
3.1.6.1	Cardinal numbers	554
3.1.6.2	Ordinal numbers	555
3.1.6.3	Quantifiers	555
3.2	Verb morphology	555
3.2.1	Stem formation	555
3.2.1.1	Subjunctive, past, and present stems	555
3.2.1.2	Causative and passive stems	555

3.2.1.3	Preverbs	556
3.2.1.4	Compound verbs	556
3.2.2	Nominal forms	557
3.2.2.1	Infinitives	557
3.2.2.2	Past participle	557
3.2.2.3	<i>miš</i> -participle	557
3.2.3	Person marking and ‘be, become’	557
3.2.3.1	Person marking	557
3.2.3.2	Existential verb	557
3.2.4	Aspectual-modal and negative affixes and particles	558
3.2.4.1	Modal prefix <i>bi-</i>	558
3.2.4.2	Negative prefixes <i>ne-</i> and <i>me-</i>	558
3.2.4.3	Imperfective particle <i>-ē</i>	558
3.2.4.4	Future particle <i>do</i>	558
3.2.4.5	Optative particle <i>wā</i>	558
3.2.5	System of tense, mood, and aspect	559
3.2.5.1	Forms based on present stem	559
3.2.5.1a	Present indicative	559
3.2.5.1b	Continuous present	559
3.2.5.1c	Present subjunctive	559
3.2.5.1d	Future	560
3.2.5.1e	Optative	560
3.2.5.1f	Imperative	560
3.2.5.2	Forms based on past stem	560
3.2.5.2a	Preterite	560
3.2.5.2b	Imperfect	560
3.2.5.2c	Past subjunctive	560
3.2.5.2d	Past conditional	561
3.2.5.2e	Past optative	561
3.2.5.3	Perfect forms	561
3.2.5.3a	Present perfect	561
3.2.5.3b	Non-witnessed function	561
3.2.5.3c	Perfect II	562
3.2.5.3d	Pluperfect	562
3.2.5.4	Overview of system of tenses	562
4	Syntax and use of forms	563
4.1	Coordinating conjunctions and clitics	563
4.2	Noun phrase structure	563
4.2.1	<i>Ezāfe</i> constructions and case marking	563
4.2.1.1	Adjectival <i>ezāfe</i> constructions	563
4.2.1.2	Nominal <i>ezāfe</i> constructions	564
4.2.1.3	Group inflection	565
4.2.1.4	Multiple <i>ezāfe</i> constructions	565
4.2.1.5	Indefinite head noun	566
4.2.1.6	Inalienable possession	566
4.2.1.7	Elliptic <i>ezāfe</i> forms	567
4.3	Clause structure and word order	567
4.3.1	Pre- and postverbal placement	567
4.3.1.1	Preverbal placement	567

4.3.1.2	Postverbal placement	567
4.3.2	Other conditions for word order	568
4.3.2.1	Omission of <i>bi-</i>	568
4.3.2.2	Adverbial complements	568
4.3.2.3	Topicalization	568
4.4	Semantics and use of nominal forms	568
4.4.1	Use of cases	568
4.4.1.1	Ergative	568
4.4.1.2	Verbal rection and transitivity	569
4.4.2	Subject marking	569
4.4.2.1	Impersonal subjects and expressions	569
4.4.2.2	Non-specific past agents	570
4.4.3	Object marking	570
4.4.3.1	Implicit direct object	570
4.4.3.2	Compound verbs	570
4.5	Semantics and use of verb forms	570
4.5.1	Possession, 'have'	570
4.5.2	Aspectual construction	571
4.5.2.1	Incipient action	571
4.5.2.2	Instantaneous action	571
4.5.3	Modal constructions	571
4.6	Sentence types and word order	571
4.6.1	Nominal sentences	571
4.6.2	Other types	572
4.6.2.1	Interrogative sentences	572
4.6.2.2	Imperative sentences	572
4.6.2.3	Optative sentences	572
4.6.2.4	Exclamative sentences	573
5	Complex sentences	573
5.1	Coordinate clauses	573
5.2	Subordinate clauses	574
5.3	Relative clauses	574
5.3.1	Functions of RC in the main clause	574
5.3.2	Resumptive pronouns	575
5.3.3	Topicalized relative clauses	575
5.3.4	Head incorporation	575
5.4	Nominal clauses	576
5.4.1	Subject clauses	576
5.4.2	Object clauses	576
5.4.3	Quotations	576
5.5	Adverbial clauses	576
5.5.1	Temporal and local clauses	576
5.5.1.1	Anterior temporal clauses	577
5.5.1.2	Simultaneous temporal clauses	577
5.5.1.3	Posterior temporal clauses	577
5.5.1.4	Local clauses	578
5.5.2	Conditional clauses	578
5.5.2.1	Possible conditions	578
5.5.2.1a	Present and subjunctive	578

5.5.2.1b	Preterite	579
5.5.2.1c	Perfect	579
5.5.2.2	Impossible conditions	579
5.5.3	Concessive clauses	579
5.5.4	Causal clauses	580
5.5.5	Final clauses	580
5.5.6	Consecutive clauses	581
5.5.7	Comparative clauses	581
6	Lexis and registers	582
6.1	Word formation	582
6.1.1	Derivation of nouns and adjectives	582
6.1.2	Compounding	582
6.2	Sociolinguistic features	582
6.2.1	Standard and colloquial registers	582
6.2.2	Kinship terminology	583
6.2.3	Terms of address	583
7	Sample text	583
	Bibliography and references	586
10	Kurdish <i>Ernest N. McCaus</i>	587
1	Introduction	587
1.1	Overview	587
1.2	Writing systems	589
2	Phonology	591
2.1	Inventory	591
2.1.1	Vowels	591
2.1.2	Consonants	591
2.1.2.1	Overview	591
2.1.2.2	Pharyngeals	592
2.1.2.3	Labio-dental ν	592
2.1.2.4	l and l'	592
2.1.2.5	r and r'	592
2.1.2.6	Nasals	593
2.1.2.7	Gemination	593
2.1.2.8	Diphthongs and glides	593
2.1.2.9	Distribution	593
2.1.3	Syllable structure and consonant clusters	593
2.2	Non-segmental features	594
2.3	Conditioned and morphophonemic changes	595
2.3.1	Vowels	595
2.3.1.1	Glide insertion, V - G - V	595
2.3.1.2	Conditioned shortening	595
2.3.1.3	Contraction	595
2.3.1.4	Unstressed i	596
2.3.1.5	Clusters and homorganic glides	596
2.3.2	Consonants	596
2.3.2.1	Voice assimilation	596
2.3.2.2	Palatalization and fronting	597

2.3.2.3	Fronting of <i>w</i>	597
2.3.2.4	Dentals	597
2.3.2.5	The “Zagros <i>d</i> ”	597
2.3.2.6	Nasal assimilation	597
2.3.2.7	Intrusive <i>d</i>	597
2.3.2.8	Deletion of <i>h</i>	598
3	Morphology	598
3.1	Nominal morphology	598
3.1.1	Nouns	598
3.1.1.1	Gender, number, and definiteness	598
3.1.1.2	Vocative	598
3.1.1.3	Locative	599
3.1.2	Adjectives and degree	599
3.1.3	Pronouns and deixis	599
3.1.3.1	Personal and demonstrative pronouns	599
3.1.3.2	Possessive forms	600
3.1.3.3	Reflexive-emphatic pronouns	600
3.1.3.4	Interrogatives and related adverbials	600
3.1.4	Adpositions	600
3.1.4.1	Pre- and postpositions	600
3.1.4.2	Absolute forms	601
3.1.4.3	Circumpositions	601
3.1.4.4	Compound absolute forms	601
3.1.5	Adverbs	602
3.1.5.1	Lexical adverbs	602
3.1.5.2	Adverbial phrases	602
3.1.5.3	<i>Izāfa</i> construction	602
3.1.6	Numerals	603
3.1.6.1	Cardinal numbers	603
3.1.6.2	Ordinal numbers	603
3.1.6.3	Quantifiers	603
3.2	Verb morphology	604
3.2.1	Stem formation and modifiers	604
3.2.1.1	Present and past stems	604
3.2.1.2	Morphological passive and causative	604
3.2.1.3	Preverbs and postverbs	604
3.2.1.4	Adpositions	605
3.2.1.5	Compound verbs	606
3.2.2	Nominal forms	606
3.2.2.1	Infinitive	606
3.2.2.2	Past participles	606
3.2.3	Person marking and ‘to be, become’	606
3.2.4	Aspectual-modal and negative prefixes	607
3.2.5	System of aspect, mood, and tense	608
3.2.6	Transitivity and ergativity	608
3.2.7	Individual verb forms	608
3.2.7.1	Present tense	609
3.2.7.2	Preterite	609
3.2.7.2a	“Aorist” function	610

	3.2.7.3 Imperfect	610
	3.2.7.4 Present perfect	610
	3.2.7.5 Past perfect	610
	3.2.7.6 Imperative	611
	3.2.7.7–8 Present and perfect subjunctives	611
	3.2.7.9–10 Imperfective and perfective counterfactual	611
	3.2.7.11 Overview of system of aspect, tense, and mood	611
4	Syntax and use of forms	612
	4.1 Coordinating conjunctions and clitics	612
	4.2 Noun phrase structure	612
	4.2.1 Post-head modifiers	612
	4.2.2 Dependent nominals	612
	4.2.2.1 Definite nominal phrases	612
	4.2.2.2 Indefinite nominal phrases	612
	4.2.3 <i>Izāfa</i> -construction	612
	4.2.4 Specific dependents	613
	4.2.5 Naming, apposition	613
	4.2.6 Infinitival head or dependent	613
	4.3 Clause structure	613
	4.3.1 Word order	613
	4.3.2 Verbal modifiers	614
	4.3.3 Questions	615
	4.3.4 Topicalization	615
	4.4 Syntax and use of nominal forms	615
	4.4.1 Pronouns	615
	4.4.2 Personal enclitics	615
	4.4.3 Subject and object marking	615
	4.4.4 Past agent suffixes	616
	4.4.5 Particles <i>čōn</i> and <i>wā</i>	616
	4.4.6 Verb-only phrases	616
	4.4.7 Preposing of adpositional suffixes	617
	4.4.8 Past tense object and possessor raising	618
	4.4.9 Three-place verb ‘to give to’	618
	4.5 Semantics and use of verb forms	619
	4.5.1 Possession	619
	4.5.2 Aspectual constructions	619
	4.5.2.1 Virtual accomplishment	619
	4.5.2.2 Progressive action	619
	4.5.2.3 Continued intensive action	619
	4.5.3 Modal constructions	619
	4.5.3.1 Ability	619
	4.5.3.2 Desire, necessity, and assumption	620
5	Complex sentences	620
	5.1 Coordinate clauses	620
	5.2 Subordinate clauses	620
	5.3 Relative clauses	621
	5.3.1 Subject	621
	5.3.2 Pronoun other than subject	621

5.4	Subject and object clauses	622
5.4.1	Subject clauses	622
5.4.2	Object clauses	622
5.4.3	Quotations	622
5.4.4	Predicative clauses	623
5.4.5	Appositive complement clauses	623
5.5	Adverbial clauses	623
5.5.1	Temporal clauses	623
5.5.2	Conditional clauses	623
5.5.2.1	Real, possible, unreal conditions	623
5.5.2.2	<i>agīna</i> 'if not, otherwise'	625
5.5.2.3	Wishes	625
5.5.3	Causal clauses	625
5.5.4	Purpose clauses	625
5.5.4.1	Potential action	625
5.5.4.2	Non-factive phrases	626
5.5.4.3	Suggestions	626
6	Lexis	626
6.1	Word formation	626
6.1.1	Derivation	626
6.1.1.1	Nouns	626
6.1.1.2	Nouns or adjectives	626
6.1.1.3	Adjectives	627
6.1.1.4	Negation	627
6.1.2	Compounding	627
6.1.2.1	With or without a composition vowel	627
6.1.2.2	Verbal component	627
6.1.2.3	Reduplication	627
6.2	Sociolinguistic aspects	627
6.2.1	Registers	627
6.2.2	Modes of address	628
7	Dialectology	628
7.1	Phonology	628
7.2	Nominal inflection	629
7.3	Verb system	630
8	Sample Sōrānī text	630
	Acknowledgment	632
	Bibliography and references	632
11	Balochi <i>Carina Jahani and Agnes Korn</i>	634
1	Introduction	634
1.1	Overview	634
1.2	Dialects and other variations	636
1.2.1	Western Balochi	637
1.2.2	Southern Balochi	637
1.2.3	Eastern Balochi	637
1.3	Writing systems	638
1.4	Descriptive approach	640

2	Phonology	641
2.1	Inventory	641
2.1.1	Vowels	641
2.1.1.1	Common Balochi vowel phonemes	641
2.1.1.2	Vowel systems of Balochi dialects	641
2.1.2	Consonants	642
2.1.2.1	Specific consonant classes	643
2.1.2.2	Pronunciation of loanwords	645
2.1.2.3	Consonant systems of Balochi dialects	645
2.1.3	Syllable structure and consonant clusters	647
2.1.3.1	Syllable patterns	647
2.1.3.2	Diphthongs	648
2.1.3.3	Geminated consonants	648
2.2	Non-segmental features	648
2.2.1	Word stress	648
2.2.2	Intonation	649
2.3	Conditioned morphophonemic changes	649
2.3.1	Nasal vowels	649
2.3.2	Consonants	650
2.3.2.1	Hiatus	650
2.3.2.2	The modal prefix <i>bi-</i>	650
2.3.2.3	Verbal suffixes	650
2.3.2.4	Metathesis	650
3	Morphology	651
3.1	Nominal morphology	651
3.1.1	Nouns	651
3.1.1.1	Nominal categories	651
3.1.1.2	Case system	651
3.1.2	Pronouns and deixis	653
3.1.2.1	Personal pronouns	653
3.1.2.2	Pronominal suffixes	654
3.1.2.3	Demonstrative pronouns	655
3.1.2.4	Other pronouns	655
3.1.3	Adjectives	655
3.1.4	Adverbs	657
3.1.5	Adpositions	657
3.1.6	Numerals	658
3.2	Verb morphology	658
3.2.1	Simple and complex verbs	658
3.2.2	Stem formation	659
3.2.2.1	Present and past stems	659
3.2.2.2	Causatives and double causatives	659
3.2.3	Non-finite forms	659
3.2.3.1	Infinitive	659
3.2.3.2	Participles	659
3.2.3.3	Gerundive	659
3.2.3.4	Agent noun	660
3.2.4	Person marking and auxiliaries	660
3.2.5	Aspectual-modal affixes and negation	661

3.2.5.1	Imperfective markers	661
3.2.5.2	Modal markers	661
3.2.5.3	Negation	661
3.2.6	System of tense, aspect, mood and voice	662
3.2.6.1	Tense	662
3.2.6.2	Aspect	662
3.2.6.3	Mood	662
3.2.6.4	Voice	662
3.2.7	Transitivity and ergativity	663
3.2.8	Individual verb forms	664
3.2.8.1	Simple tenses	664
3.2.8.2	Compound tenses	664
3.2.8.3	Basic Balochi TAM-forms	665
3.2.8.4	Possession	666
4	Syntax	667
4.1	Noun phrase structure	667
4.2	Definiteness	667
4.3	Clause structure and word order	668
4.3.1	Word order	668
4.3.2	Focus and prominence	668
4.3.3	Nominal clauses	669
4.4	Case marking	669
4.4.1	Subject marking	669
4.4.2	Object marking	669
4.4.2.1	Direct object	669
4.4.2.2	Indirect object	670
4.5	Semantics and use of verb forms	670
4.5.1	Concord	670
4.5.2	Use of TAM-forms	671
4.5.2.1	Present indicative	671
4.5.2.2	Present subjunctive	671
4.5.2.3	Imperative	671
4.5.2.4	Optative	672
4.5.2.5	Preterite	672
4.5.2.6	Imperfect	673
4.5.2.7	Present perfect	673
4.5.2.8	Past perfect	673
4.5.2.9	Past subjunctive	674
4.5.3	Aktionsart and modal constructions	675
4.5.3.1	Aktionsart constructions	675
4.5.3.2	Modal constructions	677
5	Complex sentences	678
5.1	Coordinate clauses	678
5.2	Subordination	678
5.2.1	Complement clauses	679
5.2.2	Relative clauses	680
5.2.3	Adverbial clauses	681
5.2.3.1	Temporal clauses	681
5.2.3.2	Locational clauses	681

5.2.3.3	Clauses of manner	682
5.2.3.4	Causal clauses	682
5.2.3.5	Consecutive and final clauses	683
5.2.3.6	Conditional clauses	683
5.2.3.7	Concessive clauses	684
6	Word formation	684
6.1	Nominal word formation	684
6.2	Verbal word formation	684
7	Sample texts	685
	Acknowledgements	688
	List of abbreviations	689
	Bibliography and references	689
12	Parachi <i>Charles M. Kieffer</i>	693
1	Introduction	693
2	Phonology	694
2.1	Inventory and distribution	694
2.1.1	Vowels	694
2.1.1.1	Diphthongs	694
2.1.2	Consonants	694
2.2	Non-segmental features: stress	695
2.3	Morphophonemic changes	695
3	Morphology	696
3.1	Nominal morphology	696
3.1.1	Nouns	696
3.1.1.1	Gender	696
3.1.1.2	Number	696
3.1.1.3	Declension	696
3.1.2	Adjectives	697
3.1.3	Pronouns and deixis	697
3.1.3.1	Personal pronouns	697
3.1.3.1a	Exclusive forms	697
3.1.3.2	Demonstrative pronouns	697
3.1.3.3	Reflexive pronouns	698
3.1.3.4	Interrogative-indefinite pronouns and related adverbials	698
3.1.4	Adpositions	699
3.1.4.1	Prepositions	699
3.1.4.2	Ablative - <i>ā</i> and postpositions	699
3.1.5	Adverbs	699
3.1.6	Numerals	699
3.1.6.1	Cardinal numbers	699
3.1.6.2	Ordinal numbers	700
3.1.6.3	Distribution	700
3.1.6.4	Other numerals and measure	700
3.1.6.5	Numerals and case marking	700
3.1.6.5a	Numerative - <i>a</i>	700
3.1.6.5b	Genitive of measure	700

	3.1.6.5c Plural of specificity	700
3.2	Verb morphology	701
	3.2.1 Stem formation	701
	3.2.2 Nominal forms	701
	3.2.2.1 Forms based on the present stem	701
	3.2.2.2 Forms based on the past stem	701
	3.2.3 Person marking and 'to be'	701
	3.2.4 Negation	702
	3.2.5 System of tenses and moods	702
	3.2.5.1 General present and past	703
	3.2.5.1a Past immediacy marker	703
	3.2.5.1b Continuous present and past	703
	3.2.5.1c Temporary state present and past	703
	3.2.5.2 Present perfect and past perfect	703
	3.2.5.3 Dubitative	704
	3.2.5.3a Dubitative-future	704
	3.2.5.3b Dubitative future perfect	704
	3.2.5.4 Subjunctive and conditional	704
	3.2.5.4a Present subjunctive and conditional	704
	3.2.5.4b Continuous subjunctive	704
	3.2.5.4c Perfect subjunctive and conditional	704
	3.2.5.5 Note on contracted past and perfect forms	705
	3.2.5.6 Overview of past and perfect forms	705
	3.2.5.7 Overview of tenses, moods, and aspect	705
	3.2.6 Aktionsart and modal constructions	706
	3.2.6.1 Incipient action	706
	3.2.6.2 Intention	706
	3.2.6.3 Ability and possibility	706
	3.2.6.4 Necessity and obligation	707
	3.2.7 Passive and causative formations	707
	3.2.7.1 Passive and gerundive	707
	3.2.7.1a Passive	707
	3.2.7.1b Potential passive	707
	3.2.7.1c Adjectival gerundive	707
	3.2.7.2 Causation	707
4	Syntax and use of forms	708
	4.1 Conjunctions and clitics	708
	4.2 Noun phrase structure	708
	4.2.1 Noun and adjective	708
	4.2.2 Noun and dependent noun	708
	4.2.3 Functions of the oblique	709
	4.3 Clause structure and word order	709
	4.3.1 Order of constituents	709
	4.3.2 Semantics and use of nominal forms	709
	4.3.2.1 Animacy	709
	4.3.2.2 Plurality and agreement	709
	4.3.2.3 Definiteness and specificity	710
	4.3.3 Subject and object marking: ergative	710
	4.3.3.1 Patient ending	710

4.3.3.2	Indirect object	711
4.3.3.3	Indirect object raising	711
4.3.3.4	Attachment of personal enclitics	711
4.3.4	Types of possession	712
4.3.5	Future and modal clause clitics	712
4.3.5.1	Future clitic <i>te</i>	712
4.3.5.2	Assertive clitic <i>e</i>	712
5	Complex sentences	713
5.1	Coordinate clauses	713
5.2	Subordinate clauses	713
5.3	Relative clauses	713
5.3.1	Generalized clauses	713
5.3.2	Anaphoric pronouns	713
5.3.3	Case attraction	714
5.4	Object clauses	714
5.5	Adverbial clauses	714
5.5.1	Temporal clauses	714
5.5.2	Conditional clauses	714
5.5.3	Causal clauses	715
5.5.4	Final clauses	715
5.5.5	Other clauses and conjunctions	715
5.6	Nominalized clauses	715
5.6.1	Coordinating participle in <i>-en</i>	715
5.6.2	Subordinating participle in <i>-amān</i>	716
5.6.3	Verbal noun + <i>tar</i> 'from'	716
6	Lexis and sociolinguistic aspects	716
6.1	Word formation	716
6.2	Sociolinguistic aspects and loan component	716
7	Dialectology and language development	717
8	Sample text	717
	Bibliography and references	719
13	Pashto <i>Barbara Robson and Habibullah Tegey†</i>	721
1	Introduction	721
1.1	Overview	721
1.2	Writing systems	722
2	Phonology	723
2.1	Inventory	723
2.1.1	Vowels	723
2.1.1.1	Conditioned variation	723
2.1.1.2	Distribution	723
2.1.2	Consonants	723
2.1.2.1	Inventory	723
2.1.2.2	Distribution	725
2.1.3	Clusters and syllable structure	725
2.2	Stress	725
3	Morphology	726
3.1	Nominal morphology	726

3.1.1	Nouns	726
3.1.1.1	Gender	726
3.1.1.2	Inflectional classes	727
3.1.1.2a	Irregular nouns and ablaut	728
3.1.1.2b	Dialectal variation	729
3.1.1.2c	Singular-only and plural-only nouns	729
3.1.1.2d	Inflection of infinitives and of nominal clauses	730
3.1.1.3	Use of cases	730
3.1.1.3a	Direct case, subject in present and intransitive past tense	730
3.1.1.3b	Direct case, direct object in all tenses	730
3.1.1.3c	Oblique case, subject in transitive past tenses	730
3.1.1.3d	Oblique case, object of adpositions	731
3.1.1.3e	Oblique case, vocatives	731
3.1.2	Adjectives	731
3.1.2.1	Classes	731
3.1.2.2	Irregular adjectives	732
3.1.2.3	Comparison	732
3.1.3	Pronouns and deixis	733
3.1.3.1	Personal pronouns	733
3.1.3.1a	Enclitic pronouns	733
3.1.3.1b	Independent pronouns	734
3.1.3.2	Demonstrative pronouns	734
3.1.3.3	Reflexive and possessive pronouns	735
3.1.3.4	Interrogative and indefinite pronouns and related adverbials	735
3.1.4	Adpositions	735
3.1.4.1	Prepositions and postpositions	736
3.1.4.2	Ambipositions	736
3.1.5	Adverbs	736
3.1.6	Numerals	736
3.1.6.1	Cardinal numbers	736
3.1.6.2	Ordinal numbers	737
3.2	Verb morphology	738
3.2.1	Person marking and 'be, become'	738
3.2.1.1	Basic forms	738
3.2.1.2	Third person existential and modal forms	738
3.2.2	Stem formation and tense/aspect	739
3.2.2.1	Simple verbs	739
3.2.2.2	Derivative verbs	740
3.2.2.3	Prefixed stress-shifting verbs	740
3.2.3	Nominal forms	742
3.2.3.1	Present participle	742
3.2.3.2	Perfect participle	742
3.2.3.3	Conditional participle	742
3.2.3.4	Infinitive	742
3.2.4	Aspectual-modal particles and negation	743

3.2.4.1	Future, necessity, and obligation	743
3.2.4.2	Negation	743
3.2.5	Overview of tense, mood, and aspect	743
3.2.6	Preliminary notes on inflectional pattern	744
3.2.6.1	Inflection of verb classes	744
3.2.6.2	Transitivity and ergative construction	745
3.2.6.3	Order of presentation	745
3.2.7	Basic tense constructions	746
3.2.7.1	Present time constructions	746
3.2.7.1a	Imperative	746
3.2.7.1b	Imperfective present	746
3.2.7.1c	Perfective present	747
3.2.7.1d	Present future	747
3.2.7.2	Past time constructions	748
3.2.7.2a	Imperfective past	748
3.2.7.2b	Perfective past	749
3.2.7.3	Perfect constructions	749
3.2.7.3a	Present perfect	749
3.2.7.3b	Past perfect	750
3.2.7.3c	Future perfect	750
3.2.7.3d	Conjectural perfect	750
3.2.8	Modal constructions	751
3.2.8.1	Ability	751
3.2.8.1a	Present ability	751
3.2.8.1b	Past ability	751
3.2.8.1c	Future ability	751
3.2.8.2	Necessity and assumption	752
3.2.8.3	Obligation	753
3.2.9	Passive constructions	753
3.2.10	Infinitives as subject and object	754
4	Syntax and use of forms	754
4.1	Noun phrase structure	754
4.1.1	Agreement	754
4.1.2	Dependent noun phrases	754
4.1.3	Topcialization	755
4.2	Clause structure and word order	755
4.2.1	Basic word order	755
4.2.2	Split-ergativity and agreement	756
4.3	Rules of clitization	756
4.3.1	Sentence	757
4.3.2	Order of particles	757
4.4	Conjunctions	758
4.4.1	Agreement	758
4.4.2	Reduplicative conjunctions	758
4.5	Questions	759
5	Complex sentences	759
5.1	Relative clauses	759
5.2	Noun clauses	760
5.2.1	Subject clauses	760

5.2.2	Object clauses	760
5.2.2.1	Quotations as object clauses	761
5.2.2.2	Infinitival direct object clauses	761
5.3	Complement clauses	762
5.3.1	Clauses as objects of adpositions	762
5.4	Adverbial clauses	762
5.4.1	Causal clauses	762
5.4.2	Temporal clauses	763
5.4.3	Conditional clauses	764
5.4.3.1	Overview	764
5.4.3.2	Examples of conditional clauses	765
5.5.4	Consecutive clauses	766
6	Dialectology, lexis and sociolinguistic aspects	766
6.1	Sub-groupings of dialects	766
6.2	Contacts	767
6.3	Lexis	767
7	Sample text	767
	Acknowledgement	770
	Bibliography and references	771
14a	The Pamir Languages <i>D. (Joy) I. Edelman and Leila R. Dodykhudoeva</i>	773
1	Introduction	773
1.1	Overview	773
1.2	Social status and survival	773
2	Linguistic setting and dynamics	775
2.1	Substrate and superstate	775
2.2	Genetic relationships and Sprachbund	776
3	Main linguistic characteristics and typology	777
3.1	Phonology	777
3.1.1	Vocalic systems	777
3.1.2	Consonantal systems	777
3.1.2.1	Velar and uvular fricatives	777
3.1.2.2	Palatalized affricates and fricatives	779
3.1.2.3	Labialized velars and uvulars	779
3.1.2.4	Cerebrals	779
3.1.3	Suprasegmental features	780
3.2	Morphology	780
3.2.1	Nominal morphology and categories	780
3.2.1.1	Gender	780
3.2.1.2	Number	781
3.2.1.3	Case marking	781
3.2.1.4	Definiteness	781
3.2.2	Adjectives	781
3.2.3	Pronouns	781
3.2.4	Numerals	782
3.3	Verb morphology and system	782
3.3.1	Past tenses	782

4	Lexis	783
4.1	Word formation and semantic sets	783
4.2	Loans	784
	Bibliography and references	784
14b	Shughni <i>D. (Joy) I. Edelman and Leila R. Dodykhudoeva</i>	787
1	Introduction	787
1.1	Overview	787
1.2	Writing systems	788
2	Phonology	789
2.1	Inventory and distribution	789
2.1.1	Vowels	789
2.1.2	Consonants	790
2.1.2.1	Back-lingual <i>k g</i> and fronting	790
2.1.2.2	Uvular <i>x ɣ</i> vs. velar <i>ɣ̣ ʝ</i>	790
2.1.2.3	Palatal <i>č ʝ</i> vs. dental <i>c j</i>	790
2.1.2.4	Labio-dental <i>v</i> vs. bilabial <i>w</i>	790
2.1.2.5	Non-phonemic <i>h</i>	790
2.1.3	Consonant clusters	791
2.1.3.1	Clusters and structural assimilation	791
2.1.3.2	Metathesis	791
2.1.3.3	Voice assimilation	791
2.1.3.4	Assimilation of <i>n</i>	791
2.1.3.5	<i>v > δ, θ > f</i>	791
2.2	Non-segmental features	791
2.2.1	Word stress	791
2.2.2	Phrasal stress	791
2.2.3	Sentence stress and intonation	792
2.3	Morphophonemic alternations	792
3	Morphology	792
3.1	Nominal morphology	792
3.1.1	Nouns	792
3.1.1.1	Gender	792
3.1.1.1a	Morphological gender	792
3.1.1.1b	Gender marking by attributives and verb forms	793
3.1.1.1c	Classificatory gender	793
3.1.1.2	Number	793
3.1.1.3	Case marking by demonstratives	794
3.1.1.4	Definiteness	794
3.1.2	Adjectives and degree	794
3.1.3	Pronouns and deixis	794
3.1.3.1	Personal pronouns	794
3.1.3.2	Demonstrative pronouns and deixis	794
3.1.3.3	Reflexive and possessive pronouns	795
3.1.3.4	Interrogatives and indefinites and related adverbials	795
3.1.4	Adpositions	796

3.1.4.1	Prepositions	796
3.1.4.2	Postpositions	796
3.1.4.3	Case functions	796
3.1.5	Adverbial demonstratives and adverbs	797
3.1.6	Numerals	797
3.1.6.1	Cardinal numbers	797
3.1.6.2	Ordinal numbers	797
3.1.6.3	Distributives	797
3.2	Verb morphology	797
3.2.1	Stem formation	797
3.2.1.1	Present, past, and perfect stems	797
3.2.1.2	Causative stems	798
3.2.1.3	Compound verbs	798
3.2.2	Nominal forms	798
3.2.2.1	Agent noun	798
3.2.2.2	Infinitive	798
3.2.2.3	Perfect participle	798
3.2.2.4	Adjectival participle	799
3.2.2.5	Future participle	799
3.2.3	Person marking and ‘to be’	799
3.2.4	Negation	799
3.2.5	System of tenses	800
3.2.6	Transitivity and voice	800
3.2.6.1	Transitivity	800
3.2.6.2	Passive	800
4	Syntax and use of forms	800
4.1	Coordination	800
4.2	Noun phrase structure	801
4.2.1	Attributive adjectives	801
4.2.2	Pronominal adjectives	802
4.2.3	Dependent nouns and noun phrases	802
4.2.4	<i>Izāfa</i>	803
4.3	Clause structure and word order	803
4.4	Semantics and use of nominal forms	803
4.4.1	Person, animacy, and alienability	803
4.4.2	Syntactic cases	804
4.4.2.1	Subject marking	804
4.4.2.2	Direct object marking	804
4.4.2.3	Indirect object marking	804
4.4.2.4	Adverbial phrases	804
4.5	Semantics and use of verb forms	804
4.5.1	Possession	804
4.5.2	Person, number, and agreement	805
4.5.2.1	Person and number	805
4.5.2.2	Subject-verb agreement	805
4.5.2.3	Predicates	805
4.5.3	Use of “tenses”	806
4.5.3.1	Present-future, modal function	806
4.5.3.2	Factual enclitic <i>ta</i>	806

4.5.3.3	Modal function of the perfect	807
4.5.3.4	Evidential function of the perfect	807
4.6	Aktionsart constructions	807
4.6.1	Anticipated action	807
4.6.2	Inception	808
4.6.3	Habitual action	808
4.7	Modal constructions	808
4.7.1	Ability	808
4.7.2	Wishes	809
4.7.3	Intention	809
4.7.4	Necessity and obligation	809
4.7.5	Supposition	810
4.7.6	Assumption	810
5	Complex sentences	810
5.1	Coordinate clauses	810
5.2	Subordinate clauses	811
5.3	Relative clauses	811
5.3.1	Basic structure	811
5.3.2	Generalized relative clauses, <i>ar-</i> 'each, all'	812
5.3.3	Head noun incorporation	812
5.3.4	Focused restrictive clauses	812
5.4	Object clauses	812
5.5	Adverbial clauses	813
5.5.1	Temporal clauses	813
5.5.2	Conditional clauses	813
5.5.2.1	Real or likely conditions	813
5.5.2.1a	Present + imperative	813
5.5.2.1b	Present + present	813
5.5.2.2	Possible conditions	813
5.5.2.2a	Present + perfect	813
5.5.2.2b	Perfect + perfect	814
5.5.2.3	Irreal conditions	814
5.5.2.3a	Past perfect + past	814
5.5.2.3b	Past perfect + past perfect	814
5.5.3	Concessive clauses	814
5.5.4	Consecutive clauses	815
5.5.5	Causal clauses	815
5.5.6	Final clauses	815
5.5.7	Wish clauses	815
5.5.8	Request clauses	816
5.5.9	Narration and dependent speech	816
6	Lexis	816
6.1	Word formation	816
6.1.1	Derivation	816
6.1.1.1	Suffixation	816
6.1.1.2	Prefixation	817
6.1.2	Compounding	817
6.1.2.1	Basic types	817
6.1.2.2	Determinative compounds	817

6.2	Sociolinguistic aspects	817
6.2.1	Naming	817
6.2.2	Kinship terms	818
6.2.3	Familiar address and phraseology	818
7	Dialectology and contact	819
7.1	Internal dialectology of Shughni	819
7.2	Adjacent languages and dialects	819
7.3	Loan component	820
7.4	Pamir languages and Indo-Aryan	820
7.5	Turkic Kyrghyz	820
7.6	Substrate features	821
8	Sample texts	821
8.1	The Spring Equinox Festival	821
8.2	The Prodigal Son	822
	Bibliography and references	824
15	Wakhi <i>Elena Bashir</i>	825
1	Introduction	825
1.1	Overview	825
1.2	Writing systems	826
2	Phonology	826
2.1	Inventory and distribution	826
2.1.1	Vowels	826
2.1.2	Consonants	827
2.1.2.1	Overview	827
2.1.2.2	Fricatives and affricates	827
2.1.2.3	Interdental δ	828
2.1.2.4	<i>q</i> and <i>k</i>	828
2.1.3	Syllable structure and consonant clusters	828
2.2	Non-segmental features	828
3	Morphology	828
3.1	Nominal morphology	828
3.1.1	Nouns	828
3.1.2	Adjectives	829
3.1.3	Pronouns and deixis	829
3.1.3.1	Personal and demonstrative pronouns	829
3.1.3.2	Possessive forms	831
3.1.3.3	Reflexive-emphatic and reciprocal pronouns	831
3.1.3.4	Interrogative and indefinite pronouns and related adverbials	831
3.1.4	Adpositions	831
3.1.4.1	Prepositions	831
3.1.4.2	Postpositions	832
3.1.4.3	Circumpositions	832
3.1.5	Adverbial demonstratives and adverbs	832
3.1.6	Numerals	832
3.1.6.1	Cardinal numbers	832
3.1.6.2	Ordinal numbers and distributives	833

3.1.6.3	Classifiers	833
3.2	Verbal morphology	833
3.2.1	Stem formation	833
3.2.1.1	Overview	833
3.2.1.2	Present and past stems	834
3.2.1.3	Perfect stems	834
3.2.1.4	Causative stems	834
3.2.2	Non-finite forms	834
3.2.2.1	Verbal nouns or infinitives	834
3.2.2.2	Participles	834
3.2.2.2a	Present participles	834
3.2.2.2b	Perfect participles	835
3.2.3	Person marking and the verb ‘to be’	835
3.2.3.1	Overview	835
3.2.3.2	Existential verb ‘to be’	835
3.2.4	Aspectual-modal affixes and negation	836
3.2.4.1	Aspectual clitic - <i>ǰ</i>	836
3.2.4.2	Particle -(<i>V</i>) <i>p</i>	836
3.2.4.3	Negative particles	836
3.2.5	System of tense, mood, and aspect	837
3.2.5.1	Overview	837
3.2.5.2	Present-future	837
3.2.5.2a	Subjunctive function without particle	837
3.2.5.2b	Specificity and immediacy with - <i>ǰ</i> ,	837
3.2.5.2c	Future and tenseless/generic function with -(<i>V</i>) <i>p</i>	837
3.2.5.3	Past	838
3.2.5.3a	Anterior to present and conditional perfective	838
3.2.5.3b	Imperfective meanings and conditional irrealis with - <i>ǰ</i>	838
3.2.5.4	Distant past	838
3.2.5.4a	Anterior to past event and distant past	839
3.2.5.4b	Unrealized actions and conditional irrealis with - <i>ǰ</i>	839
3.2.5.5	Perfect	839
3.2.5.6	Perfect subjunctive	839
3.2.5.7	Distant perfect	840
3.2.5.7a	Unwitnessed/inferred action	840
3.2.5.7b	Irrealis wishes in past time	840
4	Syntax and use of forms	840
4.1	Coordination	840
4.2	Noun phrase structure	840
4.2.1	Genitive phrases	840
4.2.2	Definiteness marking	841
4.3	Clause structure and word order	841
4.3.1	Clitics	841
4.3.2	Questions	841
4.3.3	Nominal sentences	841

4.4	Case marking	842
4.4.1	Subject marking	842
4.4.1.1	Non-past tenses	842
4.4.1.2	Past tense and perfect	842
4.4.1.2a	Lower dialect of Wakhan W (Namadgut)	842
4.4.1.2b	Upper dialect of Wakhan W	842
4.4.1.2c	Hunza W	843
4.4.1.3	OBL ₁ and OBL ₂	843
4.4.1.4	Oblique intransitive subjects	844
4.4.2	Direct object marking	844
4.4.2.1	Wakhan Wakhi	844
4.4.2.2	Hunza Wakhi	845
4.4.3	Causee marking	845
4.4.3.1	Pronominal causees	845
4.4.3.2	Causees with analytical causatives	846
4.5	Semantics and use of verb forms	846
4.5.1	Possession	846
4.5.2	Aspectual-phasal constructions	846
4.5.2.1	Prospective action	846
4.5.2.2	Intended action	847
4.5.2.3	Inception	847
4.5.3	Modal constructions	847
4.5.3.1	Deontic modality	847
4.5.3.2	Presumption	848
4.5.3.3	Ability	848
4.5.3.3a	Wakhan Wakhi	848
4.5.3.3b	Hunza Wakhi	848
5	Complex sentences	849
5.1	Coordinate clauses	849
5.2	Subordination	849
5.3	Relative clauses	849
5.3.1	Finite relative clauses	849
5.3.1.1	Relative noun in matrix clause	850
5.3.1.2	Relative noun in relative clause	850
5.3.1.3	Demonstrative + head noun in matrix clause	850
5.3.1.4	<i>ki</i> and <i>tsə</i> in relative clause	850
5.3.1.5	Interrogative/indefinite pronouns	850
5.3.1.6	Headless relative clauses	850
5.3.2	Participial relative clauses	851
5.3.2.1	Present and perfect participles	851
5.3.2.2	Substantival perfect participle	851
5.4	Temporal clauses	851
5.4.1	Finite temporal clauses	851
5.4.1.1	Simultaneous actions/events	851
5.4.1.2	Sequential actions	851
5.4.1.3	Punctual event	852
5.4.1.4	Time after which	852
5.4.1.5	Time up to which	852
5.4.2	Infinitival temporal constructions	852

5.4.2.1	Inception of an event	852
5.4.2.2	Sequence of two events	852
5.5	Reason and purpose clauses	853
5.5.1	Reason clauses	853
5.5.2	Purpose clauses	853
5.5.2.1	Infinitival construction	853
5.5.2.2	Finite-verb clause	853
5.6	Complement structures	853
5.7	Conditional clauses	854
5.7.1	Realis	854
5.7.1.1	Present	854
5.7.1.2	Simple past	854
5.7.1.3	Present subjunctive	854
5.7.1.4	Perfect subjunctive	855
5.7.2	Irrealis	855
5.7.2.1	Simple past	855
5.7.2.2	Distant past	855
6	Lexis and word formation	855
7	Dialectology and contact phenomena	855
7.1	Wakhi dialects	855
7.2	Adjacent languages and dialects	856
7.3	Regional contact languages	856
7.3.1	The Pamir linguistic area	857
7.3.2	Dardic	857
8	Sample text	858
	Notes	859
	Acknowledgements	859
	Bibliography and references	860
	Index	863

ILLUSTRATIONS

MAPS

2.1	Andronovo, BMAC and Yaz cultures	6
2.2	Median Empire (ca. 700 – ca. 558 BCE)	7
2.3	Achaemenid Empire (ca. 558–330 BCE)	7
2.4	Scythia during Parthian period (ca. 247 BCE – 224 CE)	8
2.5	Locator map	10
2.6	The Iranian languages	11
2.7	Modern non-Iranian languages	16
2.8	Loss of personal enclitics and present in <i>*-ant</i>	27
5.1	Sogdia and its neighbours	280
6.1	Ancient Khwarezmia	337
7.1	Khotan and Tumshuq	378
9.1	Zazaki	546
10.1	Kurdish	588
11.1	Balochi	638
14a.1	Pamir languages, Parachi and Ormuṛi	774

TABLES

2.1	Timelines of Iranian and intersecting languages	9
2.2	The four Old Iranian dialects	18
2.3	Selected West Iranian isoglosses	19
2.4	Some SW vs. NW Iranian cluster reductions	20
2.5	Some East vs. West Iranian cluster reductions	21
2.6	Synopsis of the Khwarezmian verb system	25
2.7	The three Iranian noun + adjective typologies	28
2.8	Bundled West Iranian isoglosses	29
2.9	Earlier locations of West Iranian groups	30
2.10	Ergative subject marking in Iranian languages	32
2.11	Select synopsis of differential case marking	34
3.1.1	Late Indo-European phonemes	48
3.1.2	Proto-Indo-Iranian phonemes	50
3.1.3	Proto-Indo-Iranian ablaut	50
3.1.4	Early Iranian consonant phonemes	51
3.1.5	The Avestan alphabet	52
3.1.6	The Old Persian syllabary	53
3.2.1	Avestan vowels	54
3.2.2	Avestan consonants	56
3.2.3	Old Persian consonants	58

3.3.1	Case endings: consonant stems	73
3.3.2	Case endings: vowel stems	74
3.3.3	<i>ai-</i> and <i>au-</i> stems, proterokinetic <i>i-</i> and <i>u-</i> stems	76
3.3.4	Examples of stems in stops	79
3.3.5	1st and 2nd person personal pronoun	81
3.3.6	3rd person pronouns	81
3.3.7	Demonstrative pronouns	83
3.3.8	Relative and interrogative pronouns	84
3.3.9	Declension of numerals	85
3.4.1	Verbal endings 1. present indicative; present and aorist injunctive and imperative	91
3.4.2	Verbal endings 2. perfect, pluperfect	92
3.4.3	Verbal endings 3. subjunctive and optative	92
3.4.4	'to be'	92
4.1	Alphabets. Parthian and Middle Persian scripts derived from Imperial Aramaic	198
4.2	Alphabets. The Pahlavi script	199
4.3	Alphabets. The Manichean script	199
4.4	Phonology. Vowels	200
4.5	Phonology. Consonants	200
4.6	Nominal number and case marking in the third-century inscriptions	205
4.7	Personal pronouns	208
4.8	Demonstrative pronouns	209
4.9	Cardinals	211
4.10	Ordinals	211
4.11	Middle Persian 'to be'	216
4.12	Parthian 'to be'	216
4.13	Middle Persian and Parthian 'to be': Arameograms	216
4.14	Basic verb system	218
4.15	Middle Persian present: 'to do'	219
4.16	Middle Persian past: 'to go'	219
4.17	Middle Persian perfect: 'to go'	219
4.18	Parthian present: 'to do'	219
4.19	Intransitive and transitive constructions	220
5.1	Sogdian scripts	282
5.2	Vowels and diphthongs	284
5.3	Consonants	285
5.4	Declension of light stems	288
5.5	Declension of heavy stems	289
5.6	Declension of <i>aka-</i> and <i>ākā-</i> stems	289
5.7	Personal pronouns	290
5.8	Articles and personal pronouns	291
5.9	Demonstratives	292
5.10	Demonstrative adverbs	292
5.11	Fused prepositions	294
5.12	Cardinal numbers	295
5.13	Light stem conjugation, <i>wan-</i> 'do'	298
5.14	Heavy stem conjugation, <i>patyōš</i> 'hear'	298
5.15	Preterite conjugations	301

5.16	Sogdian system of tenses and moods	305
6.1	Vowels	339
6.2	Consonants	340
6.3	Nominal inflection	343
6.4	Personal pronouns	345
6.5	Demonstrative pronouns	345
6.6	Interrogative and relative pronouns	345
6.7	Adpositions and cases	346
6.8	Cardinal numbers	348
6.9	Stem derivation	350
6.10	Morphophonology of tense stems	352
6.11	Person marking	352
6.12	Basic inflectional system, 'do, make'	354
6.13	Basic inflectional system, 'become, be'	355
6.14	Basic inflectional system, existential verb	355
6.15	System of tenses and moods	356
6.16	Order of personal suffixes	361
6.17	Kinship terminology	372
7.1	The Early South Turkestan Brāhmī script of Khotan (from Leumann 1934)	380
7.2	Comparative arrangements of Khotanese and Tumshuqese alphabets	381
7.3	Vowels – Khotanese	382
7.4	Consonants – Khotanese	382
7.5	Vocalic declensions – Khotanese	385
7.6	Declensions, <i>āa</i> - and <i>āā</i> -stems – Khotanese	386
7.7	Consonantal declensions – Khotanese	386
7.8	Pronouns – Khotanese	387
7.9	Adverbial demonstrative – Khotanese	389
7.10	Cardinal numbers – Khotanese	390
7.11	Present and past person marking – Khotanese	393
7.12	Copula and 'be, become' – Khotanese	394
7.13	Conjugation classes – Khotanese	394
7.14	Periphrastic perfect and pluperfect – Khotanese	396
7.15	Overview of tenses and moods – Khotanese	396
8.1a	Perso-Arabic alphabet	422
8.1b	Tajik alphabets	424
8.2	Vowel systems	425
8.3	Consonants	426
8.4	Personal and deictic pronouns	434
8.5	Cardinal numbers	445
8.6	'Be' and personal endings	450
8.7	Persian and Tajik conjugations	453
8.8	Basic Persian and Tajik verb system, 3s	459
8.9	Persian evidential forms, 3s	461
8.10	Persian progressive, 3s	462
8.11	Tajik innovated verb forms, 3s	468
8.12	Synopsis of verb systems, 3s	469
8.13	Clitic attachments	472
8.14	Order of relative clauses	503

8.15	Tajik finite and participial forms	509
8.16	Focused and topical order	512
8.17	Patterns of conjunctions and moods	518
8.18	Persian kinship terms	532
9.1	Zazaki vowels	547
9.2	Consonants	547
9.3	Basic nominal inflection	548
9.4	Adjectival inflection	550
9.5	Personal pronouns	551
9.6	Demonstratives	551
9.7	Cardinal numbers	554
9.8	Person marking	557
9.9	Present indicative	559
9.10	Continuous present	559
9.11	Present subjunctive	559
9.12	Intransitive preterite	560
9.13	Intransitive perfect	561
9.14	Intransitive pluperfect	562
9.15	System of tenses and moods	562
9.16	<i>Ezāfe</i> patterns	563
9.17	Scale of singular oblique marking	569
10.1	Kurdish alphabet	590
10.2	Vowels	591
10.3	Consonants	592
10.4	Hierarchy of vocalic sets	595
10.5	Specificity and definiteness	598
10.6	Independent and demonstrative pronouns	599
10.7	Cardinal numbers	603
10.8	Preverbs and prefixes	605
10.9	Preverbs: extended prepositions	605
10.10	Person marking and 'be, become'	606
10.11	Present-future	609
10.12	Preterite	609
10.13	System of aspect, mood, and tense	611
10.14	Preterite agent and patient	617
10.15	Combined Kurdish vowel phonemes	628
10.16	Combined Kurdish consonant phonemes	629
10.17	Personal pronouns in Cizre-Botan	629
10.18	Nominal inflection in Cizre-Botan	629
11.1	Alphabet commonly used for Balochi in Pakistan	639
11.2	Vowel phonemes of Common Balochi	641
11.3	Vowel systems occurring in Balochi dialects in Iran	642
11.4	Consonant phonemes of Common Balochi	642
11.5	Eastern Balochi consonant shift	642
11.6	Consonants in Southern and Western Balochi	645
11.7	Consonants in Eastern Balochi	646
11.8	Consonants in Eastern Balochi dialects of Sindh	646
11.9	Syllable patterns in Balochi	647
11.10	Case system of Pakistani Balochi	652

11.11 Case system of Turkmenistan and Afghanistan Balochi	652
11.12 Case system of Iranian Balochi	652
11.13 Case system of Central Sarawani	653
11.14 Pronouns for the 1st and 2nd singular	653
11.15 Pronouns for the 1st and 2nd plural	654
11.16 Pronominal suffixes	654
11.17 Demonstrative pronominal stems	655
11.18 Balochi adpositions (selection)	657
11.19 Cardinal numbers	658
11.20 Verbal endings and copula	660
11.21 Basic Balochi TAM-forms: system I	665
11.22 Basic Balochi TAM-forms: system II	666
11.23 Aktionsart constructions	675
11.24 Modal constructions	677
12.1 Parachi vowels	694
12.2 Consonants	695
12.3 Nominal declension	696
12.4 Personal pronouns	697
12.5 Demonstrative pronouns	698
12.6 Numerals	700
12.7 Person marking	702
12.8 Past and perfect forms	705
12.9 System of tense, mood, and aspect	706
13.1 Pashto alphabet	722
13.2 Vowels	723
13.3 Consonants	724
13.4 Nominal endings	728
13.5 Examples of declensions	728
13.6 Adjectival classes	731
13.7 Irregular adjectives	732
13.8 Clitic pronouns	733
13.9 Independent pronouns	734
13.10 Demonstrative pronouns	734
13.11 Basic cardinal numbers	737
13.12 Decades + single digits	737
13.13 Person marking and 'be, become'	738
13.14 Regular simple verbs	739
13.15 Irregular simple and suppletive verbs	739
13.16 Derivative verbs: intransitive-transitive	740
13.17 Prefixed verbs, <i>rā-</i> , <i>dar-</i> , <i>war-</i>	741
13.18 Prefixed verbs, <i>ke-</i> , <i>pre-</i> , <i>nāna-</i> ; noun + verb	741
13.19 Verbs with fused prefixes	742
13.20 Pashto basic verb system	744
13.21 Overview of verb classes and inflection	745
13.22 Protasis and apodosis in conditional clauses	764
14a.1 Transcriptions for the Pamir languages	778
14a.2 Pamir vocalic systems	778
14a.3 Pamir consonantal systems	779
14a.4 Pamir person marking in present and past	783

14b.1	Shughn(an)i vowels	789
14b.2	Consonants	790
14b.3	Pronouns and deictic adverbials	795
14b.4	Copula and personal endings	799
14b.5	Intransitive and transitive stems and imperative	800
14b.6	Intransitive conjugation	801
14b.7	Transitive conjugation	802
14b.8	Kinship terms	818
15.1	Vowels	827
15.2	Consonants	827
15.3	Case endings	829
15.4	Personal and demonstrative pronouns, Hunza	830
15.5	Personal and demonstrative pronouns, Wakhan	830
15.6	Adverbial demonstratives	832
15.7	Cardinal numbers	833
15.8	Copula, personal endings and clitics	835
15.9	Existential verb 'to be'	836

CONTRIBUTORS

Elena Bashir specializes in the languages of northwestern Pakistan. She has conducted fieldwork on a number of these languages, including Kalasha, Khowar, Wakhi, Burushaski, Shina and Balti. Additionally, she has worked on Balochi, Pashto and Brahui. Her PhD dissertation from the University of Michigan is entitled “Topics in Kalasha syntax: An areal and typological perspective”. At present her main research concentration is on Khowar, and she has been instrumental in the authorship and acceptance of new Unicode characters needed for Khowar, Burushaski and Khowar. Currently she teaches Urdu at the University of Chicago.

Leila R. Dodykhudoeva, PhD, is Senior Research Fellow at the Unit of Iranian Languages of the Department of Indo-European Languages at the Institute of Linguistics, Russian Academy of Sciences, Moscow. She is an expert in minor Iranian languages. She has worked extensively in the field with speakers of Pamiri languages. She has numerous publications on Pamiri languages, especially of the Shughnani-Rushani group.

D. (Joy) I. Edelman is Professor PhD of Linguistics at the Unit of Iranian Languages of the Department of Indo-European Languages at the Institute of Linguistics, Russian Academy of Sciences, Moscow. She is an expert in Iranian, Dardic, Nuristani and Burushaski languages. Professor Edelman has published extensively on broader issues of these languages in descriptive, typological, regional and comparative perspective.

Ronald E. Emmerick was Professor of Iranian at the University of Hamburg, Germany, and Projektleiter of the Akademienvorhabens Turfanforschung, Berlin. He was an internationally renowned expert in Khotanese, a Middle-Iranian language spoken in Xinjiang, as well as Tibetan and Sanskrit medical literature. His publications include an edition and translation of a Khotanese poem, the Book of Zambasta, and editions of the Sanskrit and Tibetan texts of the medical text Siddhisāra. He died in 2001.

Desmond Durkin-Meisterernst is a researcher at the Berlin Brandenburg Academy of Sciences and Humanities; Berlin, Germany where he works on the Iranian fragments in the Berlin Turfan Collection.

Carina Jahani is Professor of Iranian Studies at Uppsala University, Sweden. She is an internationally renowned scholar on the Balochi language and has published books and articles both about structural and sociolinguistic aspects of this language as well as about its literature. She has also organized one workshop and two conferences on Balochistan studies in Uppsala and edited proceedings from these events (alone and together with Agnes Korn and Paul Titus). She has made several field journeys to Balochi speaking parts of Pakistan and Iran from 1986 up to the present.

Dr. Charles M. Kieffer is an expert in the Iranian languages of Afghanistan and Pakistan. He has devoted his life's work to the documentation of endangered languages in the region and has published widely in this field, in particular on Parachi and Ormuri, and the increasing progression of Persian superstrates, culminating in his most recent work, *Grammaire de l'ōrmuṛī de Baraki-Barak (Lōgar, Afghanistan)* (2003).

Agnes Korn is Associate Professor for Comparative Linguistics at Frankfurt a.M. University, Germany. She holds an MA in Indo-European linguistics (Vienna University, Austria) and a PhD in comparative linguistics (Frankfurt a. M. University). She has published widely on Balochi, specifically its historical grammar of Balochi. Currently her research interests focus on Western Iranian languages, including issues of diachronic and synchronic grammar and typology.

Ernest N. McCarus is Professor Emeritus of Arabic and Kurdish Linguistics at the University of Michigan in Ann Arbor, Michigan. His publications on Kurdish include a grammar of and textbooks for the Kurdish of Sulaimania, Iraq, and essays on the phonology, morphology, and syntax of Sorani Kurdish.

Ludwig Paul is Professor of Iranian Studies at the University of Hamburg. He is an expert in Iranian languages, specializing on Middle Persian, New Persian, and West Iranian languages and dialects. His approach is both historical-linguistic and typological. He has published a reference grammar of Zazaki (NW-Iranian) and is currently working on a grammar of Early Judaeo-Persian.

John R. Perry is Professor of Persian (emeritus) at the University of Chicago. He has conducted fieldwork in linguistics and folk literature in Iran, Iraq, Afghanistan, and Tajikistan. His publications include *Form and Meaning in Persian Vocabulary: The Arabic Feminine Ending* (Mazda Publishers, 1991), *A Tajik Persian Reference Grammar* (Brill, 2005), and a broad range of articles on the history and linguistics of Persian.

Barbara Robson is the co-author, with Habibullah Tegey, of a series of Pashto language textbooks funded by the US Department of Education. She holds a PhD in Linguistics from the University of Texas/Austin. As a senior researcher at the Center for Applied Linguistics in Washington, DC, she wrote a number of publications for general audiences on languages and cultures of the Middle East and Indochina.

Prods Oktor Skjærvø is Aga Khan Professor of Iranian at Harvard University, Cambridge, Massachusetts, USA. He is an internationally recognized expert in Iranian languages and Zoroastrianism and has published extensively in these fields. His work includes an edition of the Middle Persian inscription of Narseh at Paikuli and a catalogue of the Khotanese manuscripts in the British Library, as well as a translation into Norwegian of Zoroastrian texts.

Habibullah Tegey was arguably the world's foremost authority on the Pashto language. He was a native Pashtun and holder of a PhD in linguistics from the University of Illinois. As a Professor of Pashto at Kabul University, he published numerous articles and papers on the grammars and literatures of both Pashto and Dari. In the United States, he was for many years a senior editor of the Voice of America's Pashto service, where he received a number of accolades and awards for his programs about and in Pashto. He died in 2005.

Gernot Windfuhr is Professor of Iranian Studies (emeritus) at the University of Michigan, USA. He is an internationally recognized expert in Persian linguistics, Iranian languages and related areal linguistics as well as pre-Islamic Iranian religions. His publications include a monograph of the state and history of Persian grammatical studies, a grammar and analytical dictionary of Sangesari, a Persian textbooks series, numerous theoretical and descriptive studies of verb systems, Iranian dialects, as well as studies on the cosmology, philosophy, and ritual of Zoroastrianism and other pre-Islamic Iranian religions.

Yutaka Yoshida is Professor of Linguistics at Graduate School of Letters, Kyoto University, Japan. His main interest is Sogdian and other Iranian languages once spoken along the Silk Road. He is well known as a decipherer of the Sogdian materials recently discovered in China. He has also published all the Sogdian texts preserved in Japan, among which is one brand on two pieces of sandalwood imported to Japan as long ago as 1300 years before. He is unique among Iranists in studying Iranian words transcribed in Chinese characters.

FOREWORD

Sixteen colleagues have contributed to this volume: In addition to this editor, these include Prods Oktor Skjærvø, Yutaka Yoshida, Desmond Durkin-Meisterernst, and Ronald E. Emmerick for the Old and Middle Iranian languages; John R. Perry, Ludwig Paul, Ernest N. McCarus, Carina Jahani and Agnes Korn for the West Iranian languages; Barbara Robson, Habibullah Tegey, D. (Joy) I. Edelman, Leila Dodykhudoeva, and Elena Bashir for the East Iranian languages; and Charles M. Kieffer for the Southeast Iranian languages. In collaborating with these colleagues, this editor did aim to attain as much equal sequence and coverage of the topics as possible in each chapter. Admittedly, that hands-on approach sometimes proved challenging for both parties, but was appreciated by others, and gratitude is due to all contributors for their patience.

The present volume could only appear because of the generous help in the editorial tasks by a number of these colleagues. Greatest gratitude is due to Prods Oktor Skjærvø, both for his advice and assistance in editing several chapters, and for his hands-on data-processing expertise and the availability of his data bases.

Work on this volume began some twelve years ago. Sadly two dear colleagues passed away, Ronald E. Emmerick in 2001 and Habibullah Tegey in 2005. In addition, the project encountered repeated critical challenges that threatened its completion. Most recently, it was only by the extraordinary gracious collaboration of Carina Jahani and Agnes Korn who, in spite of numerous other commitments, accepted the challenge of composing the chapter on Balochi in just a few months, that this volume could finally appear.

My fond appreciation goes to a good number of my former students at the University of Michigan, including Donald L. Stilo, Iraj Bashiri, David J. Peterson, and Karl J. Krahnke, and their enthusiasm and inspiration during work on earlier dialectological projects. For the present project doctoral student Rev. Harry Weeks assisted the project with his sophisticated editing and typesetting of earlier drafts, until his untimely death in 2005. I am likewise grateful for those students who diligently worked on the multiple aspects of the project during its various phases, most recently doctoral student Robert Haug who produced the excellent four essential maps for Chapter 2 on the typology of the Iranian languages.

The original impetus for this volume came from a letter in December of 1996 sent by Jonathan Price, then Commissioning Editor at Curzon Press (since merged with Routledge/Taylor & Francis). Once the project got started a couple of years later, he also worked with me editorially until 2005 and earned my lasting gratitude. I am equally grateful for the cordial cooperation with successive editorial teams who took his place following the commercial mergers. These include most recently Senior Production Editor Geraldine Martin, Editorial Assistant Samantha Vale Noya, and, at Swales & Willis, Production Managers Colin Morgan and Richard Willis, and copy-editor Thérèse Wassily Saba for their stellar efforts during the production phase of this volume. I would also like to thank all editorial teams who have collaborated with me on this project for

their patience during delays that occurred over the years. Finally, I am grateful to Routledge/Taylor & Francis for assuming the editorial costs involved in preparing this volume.

Geriot Windfuhr
Ann Arbor,
April 27, 2009

ABBREVIATIONS

ABBREVIATIONS I: JOURNALS, COLLECTIVE WORKS, ENCYCLOPEDIAS

<i>AcIr</i>	<i>Acta Iranica. Encyclopédie permanente des études iraniennes.</i> Téhéran/Liège/Leiden.
<i>AoF</i>	<i>Altorientalische Forschungen.</i> Berlin.
<i>AAntH</i>	<i>Acta Antiqua Academiae Scientiarum Hungaricae.</i> Budapest.
<i>AION</i>	<i>Annali. Istituto Orientale di Napoli.</i> Napoli.
<i>AION-L</i>	<i>Annali. Istituto Universitario Orientale. Sezione Linguistica.</i> Napoli.
<i>AKM</i>	<i>Abhandlungen für die Kunde des Morgenlandes.</i> Wiesbaden.
<i>AO</i>	<i>Archiv Orientální.</i> Praha.
<i>APAW</i>	<i>Abhandlungen der Preußischen Akademie der Wissenschaften. Philosophisch-historische Klasse.</i> Berlin.
<i>BAI</i>	<i>Bulletin of the Asia Institute.</i> Dearborn, Michigan.
<i>BSL</i>	<i>Bulletin de la Société Linguistique.</i> Paris.
<i>BSOAS</i>	<i>Bulletin of the School of Oriental and African Studies.</i> University of London, London.
<i>CIIran</i>	<i>Corpus Inscriptionum Iranicarum.</i> London, 1955 ff.
<i>CLI</i>	<i>Compendium Linguarum Iranicarum,</i> Schmitt, R. (ed.). Wiesbaden, 1989.
<i>CSI</i>	<i>Cahiers de Studia Iranica,</i> Paris.
<i>EIr</i>	<i>Encyclopædia Iranica,</i> London. 1982 ff.
<i>GGA</i>	<i>Göttingische Gelehrte Anzeigen,</i> Göttingen.
<i>Grundriß</i>	<i>Grundriß der Iranischen Philologie.</i> I Abteilung 1–2; Anhang II. Strassburg 1895–1901, 1898–1901, 1903, 1896–1904. Reprint Berlin/New York 1974.
<i>HbO, Iranistik I</i>	<i>Handbuch der Orientalistik.</i> I, 4: Iranistik. I: Linguistik. Leiden/Köln 1958.
<i>HbO, Iranistik II</i>	<i>Handbuch der Orientalistik.</i> I, 4: Iranistik. II: Literatur, Lieferung 1–2. Leiden/Köln 1968–1981.
<i>Iazyki mira</i>	<i>Iazyki mira. Iranskije iazyki</i> (Languages of the world. Iranian languages), 3 vols. Moskva 1997–2000.
<i>Iazyki mira I</i>	<i>Iugo-zapadnye iranskije iazyki</i> (Southwestern Iranian languages), 1997.
<i>Iazyki mira II</i>	<i>Severo-zapadnye iazyki</i> (Northwestern Iranian languages), 1999.
<i>Iazyki mira III</i>	<i>Vostochnoiranskije iazyki</i> (Eastern Iranian languages), 2000.
<i>IF</i>	<i>Indogermanische Forschungen. Zeitschrift für Indogermanistik und allgemeine Sprachwissenschaft.</i> Berlin.
<i>IJ</i>	<i>Indo-Iranian Journal.</i> Dordrecht/Boston.
<i>JA</i>	<i>Journal Asiatique.</i> Paris.

<i>JRAS</i>	<i>Journal of the Royal Asiatic Society of Great Britain and Ireland.</i> London.
<i>LSI X</i>	<i>Linguistic Survey of India.</i> Compiled and edited by G.A. Grierson. X: Specimens of Languages of the Eranian Family. Calcutta 1921.
<i>MO</i>	<i>Le Monde Oriental.</i> Uppsala.
<i>MSS</i>	<i>Münchener Studien zur Sprachwissenschaft.</i> München.
<i>NTS</i>	<i>Norsk Tidsskrift for Sprogvidenskap.</i> Oslo.
<i>Opyt</i>	<i>Opyt istoriko-tipologicheskogo issledovaniia iranskikh iazykov</i> (Essays on the historical-typological study of Iranian languages), general editor V. S. Rastorgueva, 2 vols. Moskva, 1975.
<i>Opyt I</i>	<i>Fonologiiia, evoliutsiia morfologicheskogo tipa</i> (Phonology, evolution of morphological type).
<i>Opyt II</i>	<i>Evoliutsiia grammaticheskikh kategoriï</i> (Evolution of grammatical categories).
<i>OS</i>	<i>Orientalia Suecana.</i> Stockholm.
<i>Osnovy</i>	<i>Osnovy iranskogo iazykoznaniiia iazyki</i> (Fundamentals of Iranian languages), general editor V. S. Rastorgueva, 6 vols. Moskva, 1979–97.
<i>Osnovy I</i>	<i>Drevneiranskii iazyki</i> (Old Iranian languages), 1979.
<i>Osnovy II</i>	<i>Sredneiranskii iazyki</i> (Middle Iranian languages), 1981.
<i>Osnovy III</i>	<i>Novoiranskii iazyki: zapadnaia grupa, prikaspiiskii iazyki</i> (New Iranian languages: western group, Caspian languages), 1982.
<i>Osnovy III/2</i>	<i>Novoiranskii iazyki: vostochnaia grupa</i> (New Iranian languages, eastern group), 1987.
<i>Osnovy IV/1</i>	<i>Novoiranskii iazyki: severo-zapadnaia grupa, I</i> (New Iranian languages: northwestern group, I), 1991.
<i>Osnovy IV/2</i>	<i>Novoiranskii iazyki: severo-zapadnaia grupa, II</i> (New Iranian languages: northwestern group, II), 1997.
<i>SbBAW</i>	<i>Sitzungsberichte der Bayerischen Akademie der Wissenschaften. Philosophisch-historische Klasse.</i> München.
<i>SbHAW</i>	<i>Sitzungsberichte der Heidelberger Akademie der Wissenschaften. Philosophisch-historische Klasse.</i> Heidelberg.
<i>SbPAW</i>	<i>Sitzungsberichte der Preussischen Akademie der Wissenschaften. Philosophisch-historische Klasse.</i> Berlin.
<i>StII</i>	<i>Studien zur Indologie und Iranistik.</i> Reinbek bei Hamburg.
<i>StIr</i>	<i>Studia Iranica.</i> Paris.
<i>STUF</i>	<i>Sprachtypologie und Universalienforschung.</i> Berlin.
<i>TPhS</i>	<i>Transactions of the Philological Society.</i> Oxford.
<i>ZDMG</i>	<i>Zeitschrift der Deutschen Morgenländischen Gesellschaft.</i> Wiesbaden.
<i>ZII</i>	<i>Zeitschrift für Indologie und Iranistik,</i> Leipzig.

ABBREVIATIONS II: LANGUAGES

Ar.	Arabic	Bal.	Balochi
Aram.	Aramaic	Br.	Brahui
Arm.	Armenian	Casp.	Caspian dialects
Az.	Old Azeri dialects, Iranian	Centr.	Central (Plateau) dialects
Bact.	Bactrian	ComBal.	Common Balochi
Baj.	Bajelani (Gorani dialect)	D.	Digor

Drd	Dardic	Psht.	Pashto
EBal.	Eastern Balochi dialects	Pers.	Persian, specifically Modern Standard Persian
EIr.	East Iranian		
ENP	Early New Persian	Pers.-Ar.	Perso-Arabic
Gath.	Gathic Avestan	Russ.	Russian
Ge.	Georgian	Sang.	Sangesari
Gil.	Gilaki	SBal.	Southern Balochi dialects
Gor.	Gorani	SBalKar.	Southern Balochi of Karachi
HW	Hunza Wakhi		
I.	Ishkashmi	Sem.	Semnani
I.	Iron	Sh.	Shughni
I.-A.	Indo-Aryan	Siv.	Sivandi
Ind.	Indic	Skt.	Sanskrit
Ir.	Iranian	Sogd.	Sogdian
IrBal.	Balochi spoken in Iran	SWBal.	Southern and Western Balochi
IrBalSarh.	Iranian Balochi of Sarhadd		
IrBalSarw.	Iranian Balochi of Sarawan	SWIr.	Southwest Iranian
Kb	Kaboli Persian	Taj.	Tajik Persian
Kd.	Kurdish	TBal.	Balochi spoken in Turkmenistan
Khot.	Khotanese		
Khwar.	Khwarezmian	Tumsh.	Tumshuqese
Kumz.	Kumzari	Tk.	Turkic
Maz.	Mazandarani	Tu.	Turkmen
MIr.	Middle Iranian	Uzb.	Uzbek
NIr.	New Iranian	Ved.	Vedic
NP	New Persian	W.	Wakhi
Nur.	Nuristani	WBal.	Western Balochi dialects
NWIr.	Northwest Iranian	WBalNosh.	Western Balochi of Noshke, Pakistan
OAv.	Old Avestan		
OIr.	Old Iranian	WIr.	West Iranian
OP., OPers.	Old Persian	WW	Wakhan Wakhi
Orm.	Ormuri	Y.	Yazghulami
Oss.	Ossetic	Yagh.	Yaghnobi
Pam.	Pamir languages	YAv.	Young(er) Avestan
Par.	Parachi	Zaz.	Zaza
Parth.	Parthian		

ABBREVIATIONS III: GRAMMATICAL TERMS

Note: These abbreviations may appear in capital letters or in lower case with period.

1, 2, 3	person (with s and p)	ABS	absolute
1st pers.	first person	ACC	accusative
2nd pers.	second person	ACT	active voice
3rd pers.	third person	AD(ESS)	adessive
		ADJ	adjective
A	accusative	ADDR	address
ABL	ablative	ADP	adposition

ADV	adverb	DIM(IN)	diminutive
AG	agent	DIR	direct case
AGR	agreement	DISTPT	distant/remote past
ALL	allative, allative function	DISTR	distributive
ANIM	animate	DO	direct object
AOR	aorist	DU	dual
ART	article	DUR	durative
ASP	aspect		
ATTR	attribute	EMPH	emphatic
ATTRV	attributive suffix	ENC(L)	enclitic
AUX	auxiliary	EQ	equative
		ERG	ergative
BEN	beneficiary, benefactive verb (Tajik)	EX	existential verb
		EV(ID)	evidential
		EZ	ezāfe/izāfa
C	any consonant		
CAUS	causative	F	feminine
CL	clause, clause linker	FEM	feminine
CLC	clitic	FOC	focus
CLSS	classifier	FUT/FT	future
CMPC	comparative clause		
CMPV	comparative	G-D	genitive-dative
CNCC	concessive clause	GEN	genitive
CND	conditional	GENRC	generic
CNDC	conditional clause	GL	goal
CNJECT	conjunctural	GLD	glide
CNJV	conjunction	GER	gerund
COLL.	collective	GERV	gerundive
COM	comitative	GLOTT	glottal
COMP	complementizer		
COMPD	compound	H	hiatus filler
COMPV	comparative (degree)	HAB	habitual
COND	conditional	HI	hiatus filler (glide)
CONJ	conjunct verb	HON	honorific
CONN	connective	HORT	hortative
CONV	converb	HS	heavy stem
CONT	continuative marker	HUM	human
COP/COPU	copula		
CP	complement (clause, predicate)	I-A	instrumental-ablative
		IDEF	indefinite
		IMP	imperative
D	direct case	IMPF	imperfect
DAT	dative	IMPFV	imperfective
DBL	double	INAN	inanimate
DEF	definite	INCL	inclusive
DEM	demonstrative	IND	indicative
DENOM	denominative	I(N)DEF	indefinite
DENT	dental	INDET	indeterminate
DET	determinate, determiner	INESS	inessive

INF	infinitive	QUOT	quotative (particle; verb forms)
INF2	short infinitive (= PT stem)		
INGR	ingressive		
INJ	injunctive	P	plural
INST	instrumental	PAL	palatal
INT	intentional (construction)	PART	participle; particle
INTERJ	interjection	PARTV	partitive
INTR	intransitive	PASS	passive
IO	indirect object	PAST	past
IPFV	imperfective	PAT	patient
IPR	imperative	PC	personal clitic
IRR	irreal, irrealis	PERF	perfect
ITR	intransitive	PERM	permansive
ITER	iterative	PERS	person
IZ	izāfa	PF	perfect
		PFV	perfective
L	linker	PGR	group plurality
LAB	labial	PHAR	pharyngeal
LINK	linker	PL(UR)	plural
LOC	locative	PN	proper name
LS	light stem	POSS	possessive
LW	loanword	POST	postposition
		POT	potentialis, potential construct
M	masculine	PP	past perfect, pluperfect; prepositional phrase
MASC	masculine		
MDL	middle voice	PR	present
MID	middle voice	PREF	prefix
MOD	modal	PREP	preposition
		PRES	present
N	any noun; nominative	PR.PRT	present participle
NEUT	neuter	PRET	preterite
NEG	negative	PREV	preverb
NOM, N	nominative	PRO	pronoun
NT	neuter	PROG	progressive
NOM	nominative	PROH	prohibitive
NOMZ	nominalizer	PRO(N)	pronoun
NOMCL	nominal clause	PT	past
NP	noun phrase	PTCP	participle
NUM	number, numeral	PTPART	passive participle
NUMV	numerative	PTT	partitive
NVE	non-verbal element	PRPC	purpose clause
O	oblique case; object	RC	relative clause
OBJ	object; objective	REF	referential
OBJC	object clause	REFL	reflexive
OBL	oblique case	REL	relative pronoun/adverb
OPT	optative	REM	remote
		RETR	retroflex

RSLT	resultative	V.EL	verbal element
S	singular; subject	V.EX	verbum existentiae
SG/SING	singular	V.I.	intransitive verb
SRC	source	VN	verbal noun
SUBJ/SBJ	subjunctive; subject	VOC	vocative
SUFF	suffix	VP	verb phrase
SUPER	superlative	V.PREF	verbal prefix
TAM	tense, aspect, and mood	V.T.	transitive verb
TMPC	temporal clause	W	any semivowel
TOP	topic (particle)	X	any lexical item
TR/TRANS	transitive		
UV	uvular	SIGLA	
V	any vowel; verb, verbal	*	reconstructed form
VA	verbal adjective	°	form cited without prefix, suffix, or enclitic
VB	verb	<	from
Vel.	velar	>	becomes

CHAPTER ONE

INTRODUCTION TO *THE IRANIAN LANGUAGES*

Gernot Windfuhr

1 OVERVIEW

The Iranian languages constitute the western group of the larger Indo-Iranian family which represents a major eastern branch of the Indo-European languages. With an estimated 150 to 200 million native speakers, the Iranian languages are one of the world's major language families. The present volume thus relates linguistically most closely to four other volumes in the Language Family Series: genetically to *The Indo-European Languages* and *The Indo-Aryan Languages*, areally to the latter as well as to *The Turkic Languages* and *The Semitic Languages*, and typologically to all four of them due to adjacency and partial symbiosis.

Following an overview of the typology of the Iranian languages and selected topics, this volume provides detailed descriptions of principle Iranian languages from Old Iranian to New Iranian. In terms of descriptive orientation, it aims to present the typological dynamics of the Iranian languages through time and space. In terms of coverage, each chapter addresses issues on all linguistic levels including not only an overview, writing systems, phonology and morphology, but also phrase, clause, and sentence level syntax, and pragmatic aspects, which are all documented by examples with close interlinear translations and comments. That is, the overriding focus is on how these languages “work”, highlighting on each level significant typological features. As such, the volume is complementary to the *Compendium Linguarum Iranicarum* (1989), edited by Rüdiger Schmitt (see *CLI* in the List of Abbreviations, I.), which, with its focus on the phonological and morphological levels, will stand as the standard reference work for many years to come. In fact, several contributors to the present volume also contributed to the *Compendium* (Skjærvø, Emmerick, Windfuhr, Kieffer).

The orientation towards typology reflects the appearance of an increasing number of publications on Iranian typology and linguistic universals on all levels, which necessarily encompass diachrony and diatopy, by a growing group of specialists in Iranian linguistics and of general linguistics working on Iranian languages, an orientation originally spearheaded, among others, by Joy I. Edelman (e.g. 1968). Also, comprehensive studies of individual Iranian languages and language groups have been published, a good number of them by contributors to this volume, and comprehensive series on Iranian languages have appeared, such as the two volumes of the *Opyt istoriko-tipologicheskogo issledovaniia iranskikh iazykov* (1975), the six volumes of the *Osnovy iranskogo iazykoznaniiia* (1979–1997), both edited by Vera S. Rastorgueva, and the three volumes dedicated to Iranian languages of the *Iazyki mira. Iranskie iazyki* (1997–2000), edited by Andrei Kibrik (for these see *Opyt*, *Osnovy*, *Iazyki mira* in the List of Abbreviations, I.). Two

recent overviews of the Iranian languages and the symbiotic non-Iranian languages are Skjærvø (2006) and Windfuhr (2006), respectively.

This volume does not include descriptions of all Iranian languages, but only of 16 languages of the many, as representatives for the following characteristics: (1) the three historical stages of documentation, Old, Middle, and New Iranian; (2) the four main dialectological groups, North-West Iranian, South-West Iranian, East Iranian, and South-East Iranian; (3) geographical location. Specifically, modern South-West Iranian is represented by Persian and Tajik; North-West Iranian is represented by Zazaki, Kurdish, and Balochi; East Iranian is represented by Pashto, Shughni, and Wakhi; South-East Iranian is represented by Parachi. Geographically, Kurdish represents the (north)-westernmost expansion of the Iranian languages just as diametrically opposed Balochi represents their south-easternmost expansion, and the Pamir languages with Shughni and Wakhi represent the north-easternmost Iranian languages. Regrettably, Ossetic could not be included.

Overall, the coverage of languages in this volume can be seen as representing the center and the outer circle of Iranian, the latter in contact with non-Iranian languages and thus marked by Randsprachen as well as interference phenomena, anchored on two chapters: Chapter 3 on Old Iranian which represents the foundation of the Iranian languages, and Chapter 8 on Persian and Tajik which represents the superstrate language over the Iranian expanse and beyond, both as the literary language and through its regional varieties and vernaculars. Finally, in terms of morphological complexity, Persian represents the least inflectional language, while two languages represent the most highly inflectional languages: Zazaki among the West Iranian languages and morphologically even more complex Pashto among the East Iranian languages.

In terms of descriptive strategy, a number of chapters discuss closely related languages jointly and thereby highlight their comparative dynamics. These include: Avestan and Old Persian in Chapter 3 (Skjærvø); Middle Persian and Parthian in Chapter 4 (Skjærvø); Khotanese with Tumshuqese in Chapter 7 (Emmerick); Persian (Windfuhr) and Tajik (Perry) in Chapter 8. The dynamics of “Common” Balochi in relation to a great number of varieties, rather than an arbitrarily selected “Standard” one, is discussed in Chapter 11 (Jahani and Korn). Chapter 14a on the complex Sprachbund of the Pamir languages (Edelman and Dodykhudoeva) precedes the description of its dominant language, Shughni, in Chapter 14b (Edelman and Dodykhudoeva); Wakhan Wakhi and Hunza Wakhi are contrasted in Chapter 15 (Bashir). The other chapters, while focusing on individual languages, likewise provide notes on dialectology: Middle Iranian Sogdian in Chapter 5 (Yoshida) and Khwarezmian in Chapter 7 (Durkin-Meisterernst); Modern Iranian Zazaki in Chapter 9 (Paul); Kurdish in Chapter 10 (McCarus); Parachi in Chapter 12 (Kiefler); Pashto in Chapter 13 (Robson and Tegey).

Like the chapters on the Old and Middle Iranian languages, the chapters on the modern languages reflect the present state of research and are innovative in their detail, and specifically in their syntactic and typological coverage within the framework of a book chapter. A particular kind of innovation is the daring decision by the authors of the chapters on Sogdian and Khwarezmian to use phonemic transcription throughout in addition to transliteration.

Finally, not discussed in this volume are challenging recent studies that, with due caution, investigate the correlation of regional gene pools of contemporary populations with language groups and language shifts (rather than populations shifts) for which small elites are often sufficient, as shown for example in the study of the origin of the Kurdish- and Zazaki-speaking populations by Hennerbichler (forthc.).

BIBLIOGRAPHY

- Edelman, D. I. (1968) *Osnovnye voprosy lingvisticheskoi geografii. Na materiale indo-iranskikh iazykov* (Fundamental problems of linguistic geography. Based on data of the Indo-Iranian languages), Moskva: Nauka.
- Hennerbichler, F. (forthc.) *Die Herkunft der Kurden*, fully revised edition.
- Skjærvø, P. O. (2006) 'Iran vi. Iranian languages and scripts', in *EnIr* 13.4, pp. 344–377 (additional parts forthcoming)
- Windfuhr, G. (2006) 'Iran vii. Non-Iranian languages,' in *EnIr* 13.4, pp. 377–410.

CHAPTER TWO

DIALECTOLOGY AND TOPICS

Gernot Windfuhr

1 INTRODUCTION

Today the Iranian languages are spoken from Central Turkey, Syria and Iraq in the west to Pakistan and the western edge of Xinjiang Uygur Autonomous Region of China in the east. In the North, its outposts are Ossetic in the central Caucasus and Yaghnobi and Tajik Persian in Tajikistan in Central Asia, while in the South they are bounded by the Persian Gulf, except for the Kumzari enclave on the Masandam peninsula in Oman.

Historically, the New Iranian stage overlaps with the Islamization of Iranian-speaking lands in the seventh century CE. The Middle Iranian stage began in the third century BCE. The oldest stages go back to the beginning of the second millennium BCE. The oldest physical document of Iranian is the Old Persian inscription by Darius I. of 522 BCE on the rock face of Mt. Behistun near Kermanshah along the highway that leads down from the Iranian plateau into Mesopotamia.

1.1 Origins: The Central Asian component

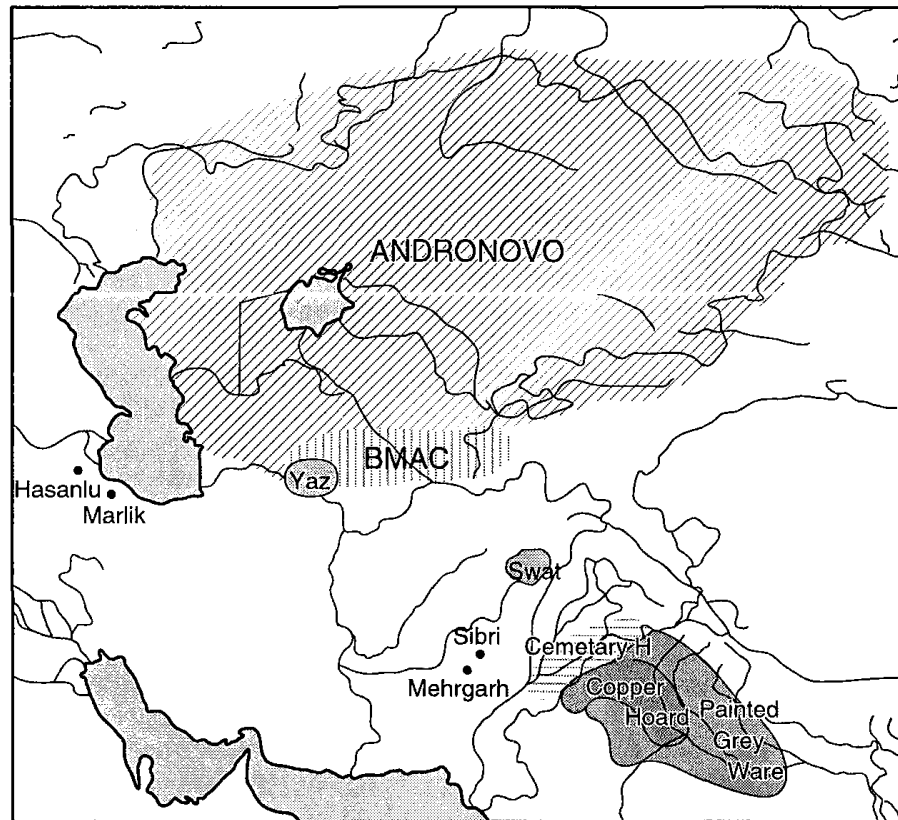
For the following section, cf. also Windfuhr (2006b) and the Introduction to Chapter 3.

Research during recent decades suggests that the Proto-Indo-Iranians originated in the eastern European steppes (Pit-Grave culture, ca. 3500–2500 BCE). From there they apparently moved eastward to the southern Ural steppes and the Volga (Potapovo culture, 2500–1900 BCE), then further on to Central Asia (Andronovo culture, from 2200 BCE onwards). At that stage they appear to have already formed two groups: the Proto-Iranians in the north, and the Proto-Indo-Aryans in the south. They came into contact with the proto-urban population of the Bactria-Margiana Archaeological Complex (BMAC), also known as the Oxus Culture), which had ancient connections to northwest India, Elam and northern Mesopotamia. They assimilated, gained prominence, and transformed it, thereby attracting non-Indo-Iranian elements. In the process they had developed a new type of social structure, called the *khanate*, which was ruled by a landlord (*khān*) residing in fortified farmsteads (*qala*).

After 2000 BCE, most of the later Indo-Aryans moved southeast probably via Afghanistan into the Indian subcontinent (Panjab), and also southwest via the Iranian plateau into northern Mesopotamia (Mitanni kingdom), probably under pressure from the Iranians to their north. The later Iranians moved into and across the Iranian plateau, both carrying the new social structure with their languages, with a lasting impact on the socio-political structures of Iran and Afghanistan, and the subcontinent.

Linguistically, these cultural contacts with the non-Indo-European languages of the proto-urban civilization in lower Central Asia left distinct shared layers of loanwords in the lexicon of Indo-Aryan and Iranian (Lubotzky 2001).

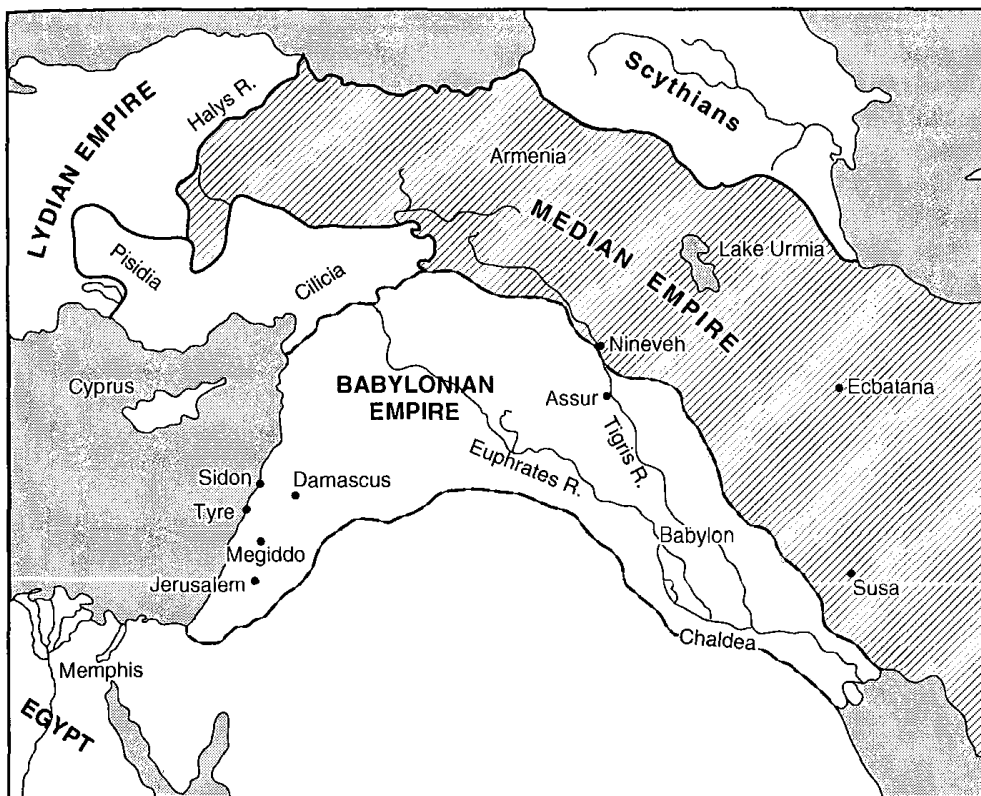
The Iranians on their part can probably be correlated with the subsequent so-called Yaz I culture in the BMAC complex, which reflects major cultural changes towards a more rural society after 1500 BCE. They apparently remained in Central Asia, and only by the end of the second millennium BCE began to spread over the Iranian plateau.



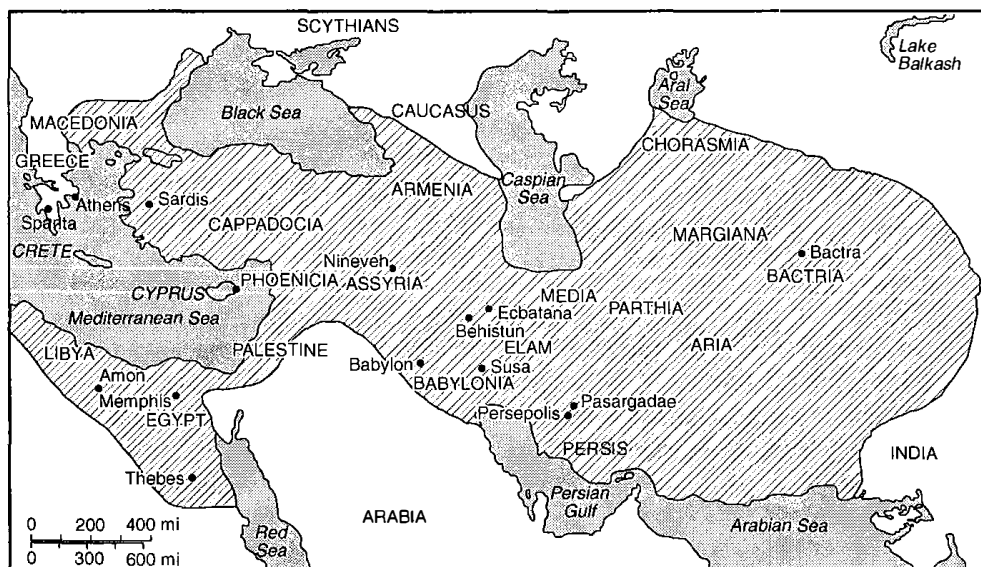
MAP 2.1 ANDRONOVO, BMAC AND YAZ CULTURES

By the second half of the eighth century BCE, Iranian Median and Persian tribes (Mada and Parswa) had already been long established among the original non-Iranian speakers of the Zagros mountain ranges of Iranian Kurdistan, according to the records of the Assyrian ruler Shalmaneser III (r. 858–824 BCE). Minorsky (1957: 78) recognized that the name of the tiny village *Qal'a Paswē* near Solduz in Kurdistan retains the memory of the Iranian settlements (cf. also Zadok 2001; 2002). The successors of the Parswa tribes who settled in the southwest of the Iranian plateau created the Achaemenid Empire (ca. 558–330 BCE) which, beginning with the Sasanian period (224–651 CE) and thereafter, ultimately resulted in the dominance of Persian over the Iranian expanse.

While the Iranian plateau was increasingly Iranianized, the Iranian tribes known as Scythians by the Greek and as Saka by the Achaemenids remained in the wide expanses

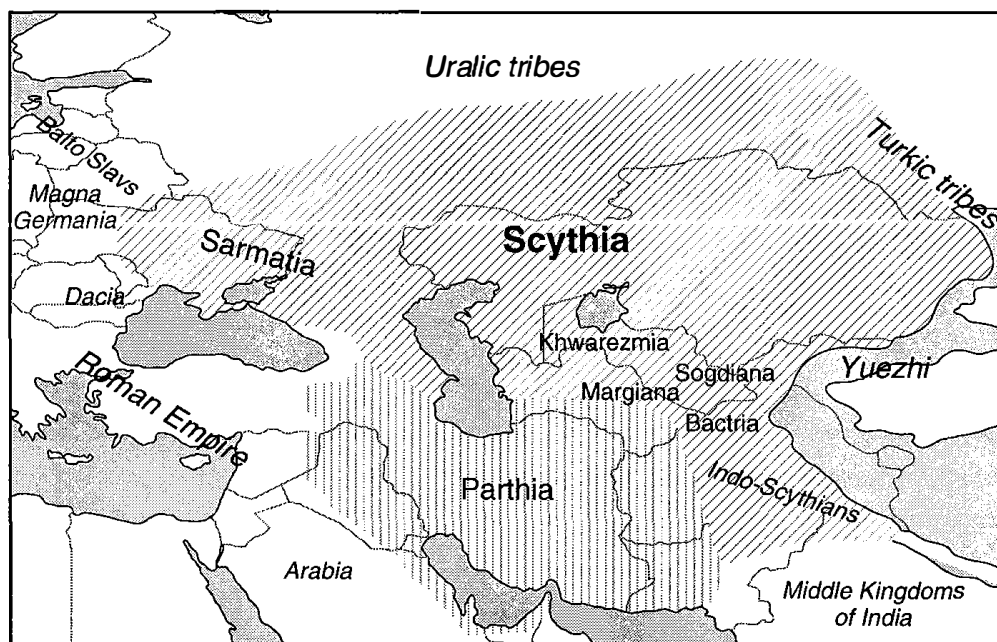


MAP 2.2 MEDIAN EMPIRE (CA. 700 – CA. 558 BCE)



MAP 2.3 ACHAEMENID EMPIRE (CA. 558–330 BCE)

to the north, ranging from the Southern Russia to Central Asia and beyond, and became the predecessors of the Middle East Iranian languages, and of the surviving modern languages, including Ossetic. However, Scythian and Saka groups also invaded the Median Kingdom from southern Russia as early as the later 8th century. Numerous later incursions are known, to which belong the subgroups who settled in Zranka, later Sistan, in the two centuries straddling the common era during Parthian rule. The name Saka is still reflected in a good number of locations, including the province of Sistan < **Saka-stān*, and probably Sangesar < ?**Saka-sar(a)*- (for Scythians invasions of NW Iran and Media in the early 8th century BCE, and Scythian loans in Old Persian, cf. Lubotsky 2002).



MAP 2.4 SCYTHIA DURING PARTHIAN PERIOD (CA. 247 BCE – 224 CE)

1.2 Ancient non-Iranian contact languages

For western Iran, the main identifiable non-Indo-European languages are located along the Zagros Mountains: (1) Hurro-Urartian, which is a linguistic isolate, but may be remotely related to the Northeast Caucasian languages, in Azerbaijan and Iranian Kurdistan; (2) the linguistic isolate Kassite in the central Zagros; and (3) Elamite further south, possibly remotely related to Dravidian as mentioned. In terms of political and tribal entities, these correspond to the kingdoms/regions of Mannea, Parsua, Ellipi, Elam, and Anshan.

In the North, the languages along the western Caspian littoral probably included speakers of South Caucasian languages, and those along the eastern littoral may have belonged to an assumed so-called Central Asian linguistic continuum. On the central plateau and probably in Kerman province, at least some languages may have been related to Elamite and to the Dravidian languages towards the east.

Little is known about the non-Iranian speakers whom the Iranian speakers

encountered on their way into and across Iran. In the east, in Baluchestan and further east in the Gandhara/Kabul region and Arachosia/Kandahār, the Iranians were most likely still in contact not only with speakers of Indo-Aryan, but also with speakers of Dravidian during the second half of the second millennium BCE.

The following synoptic table correlates the diachrony of the Iranian languages up to the New Persian period with those of Indo-Aryan and the major non-Indo-European languages with which they intersected at various stages. Omitted from the tabulation is Dravidian because of the uncertain information, and Greek.

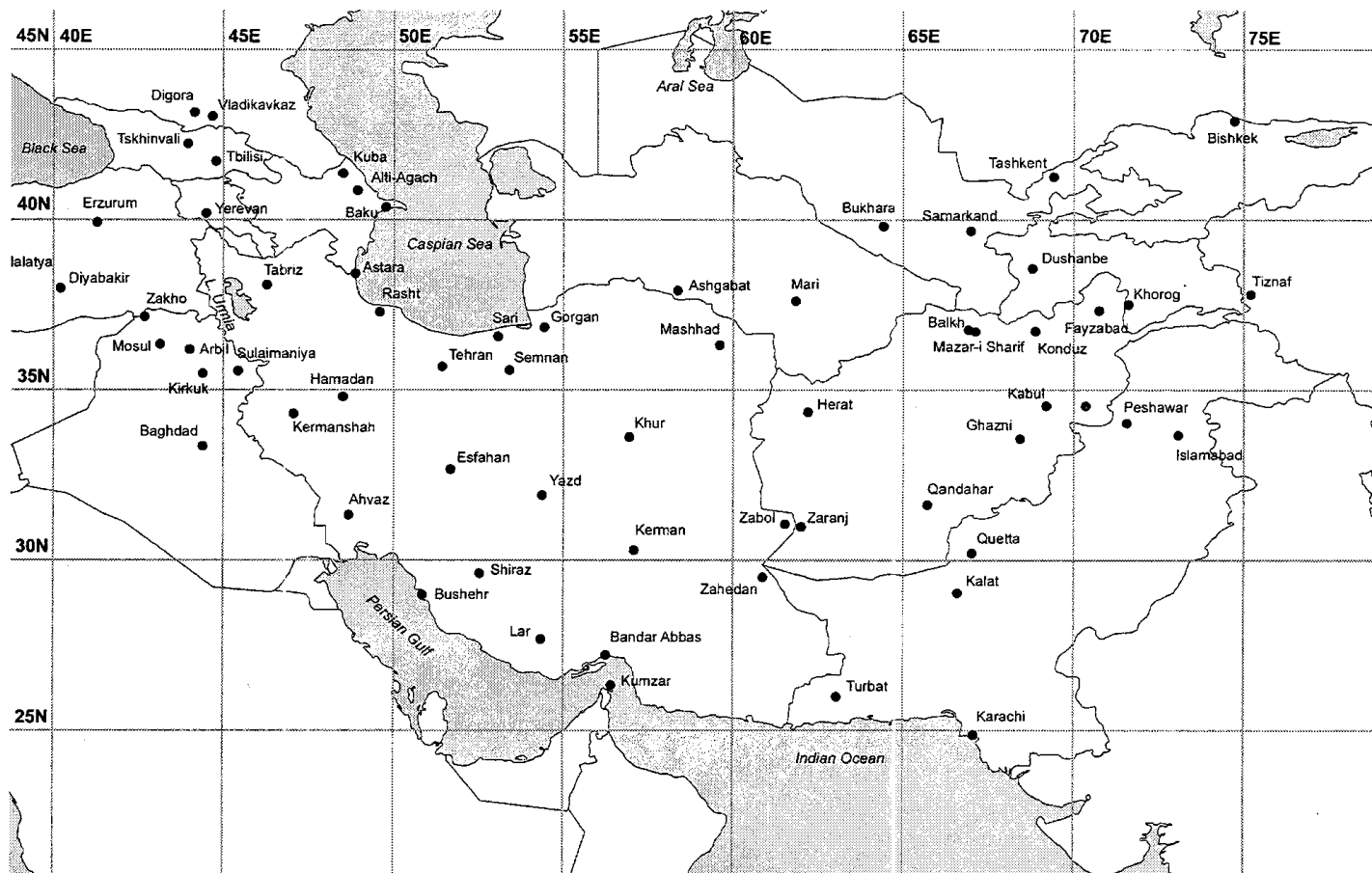
TABLE 2.1: TIMELINES OF IRANIAN AND INTERSECTING LANGUAGES

BMAC					
INDO-ARYAN					
Early I-A 2000–1500	Vedic Stage 1500–1000 Mitanni/Panjāb	Late Vedic 1000–500	Middle Indic 500 BCE	Middle I-A	New I-A
IRANIAN					
Early Iranian 2000–1500	Old Avestan 1500–1000	Young Av 1000–600 > Parsua/Medes Old Persian	Early Middle 600–300 Achaemenid	Middle Iranian 300 ^{BC} –700 CE Middle Persian	New Iranian 700 CE New Persian
ELAMITE					
Old Elamite 2600–1500	Middle Elamite 1500–1000	Neo-Elamite 1000–550	Achaemenid 550–330	Late Elamite?	Khuzi ? 1000 CE
ARAMAIC					
NW Semitic	Proto-Aramaic 1500–950	Old Aramaic 950–600 Replacing Akkadian	Achaemenid 600–200	Mid & Late Aramaic 200 BCE–700 CE	New Aramaic 700 CE
AKKADIAN East Semitic (Babylonian & Assyrian)					
Old Bab/Ass 2000–1500	Mid Bab/Ass 1500–1000	Neo-Bab/Ass 1000–600	Late Bab [600 BCE	Latest Texts 100 CE]	
HURRIAN-URARTIAN					
Old Bronze 2200–2000	Mid Bronze Late Hurrian 1500–1000	Late Bronze Urartian 850–600			

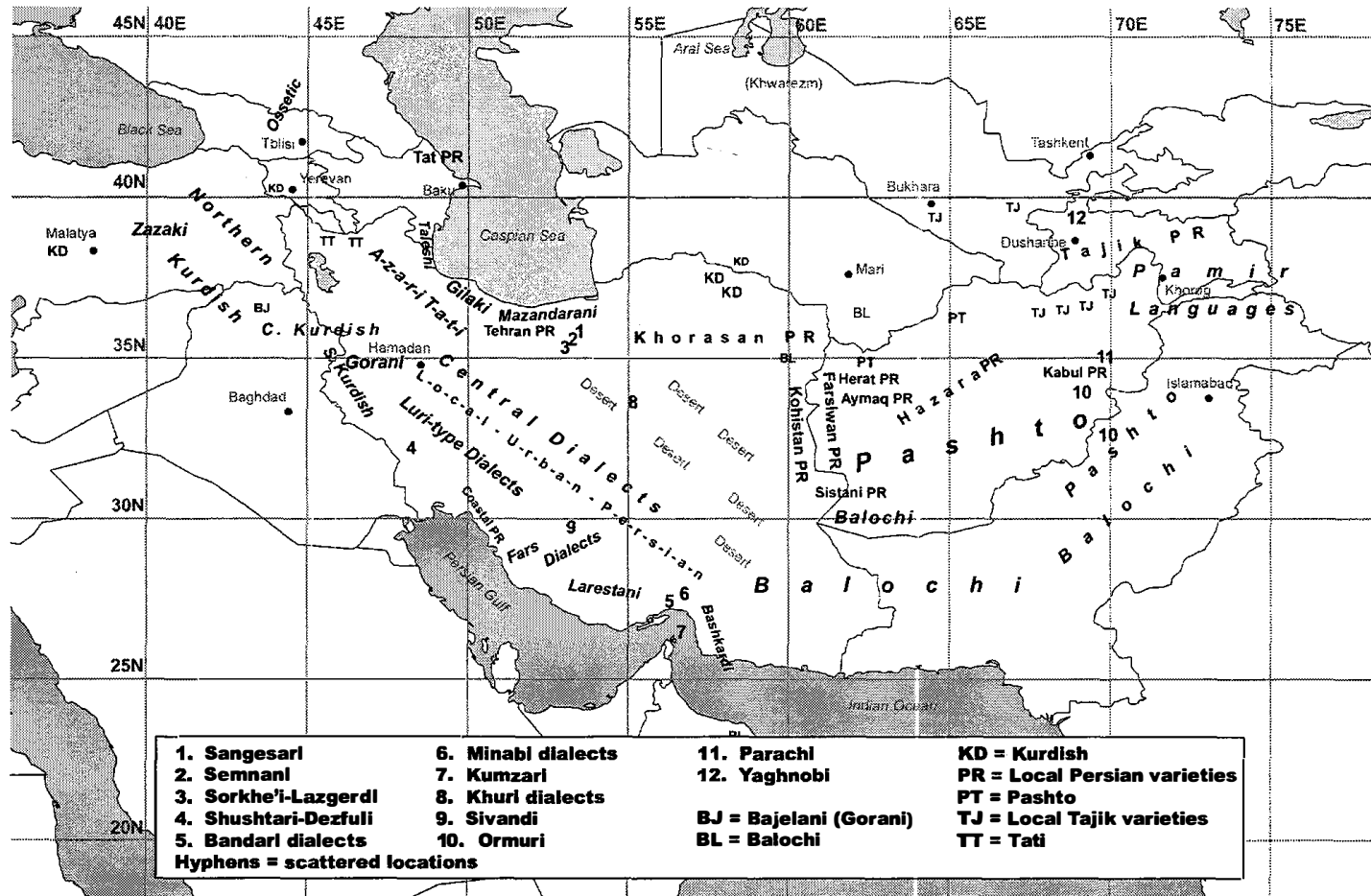
1.3 Listing of Iranian languages

In overview, the Pre-Islamic languages include well-documented Avestan and Old Persian and two language groups, Median and Scythian/Saka, with only onomastic documentation, while Middle Iranian includes six well documented languages.

There are a multitude of New Iranian languages. While the overall grouping of these languages has been well established, the internal dialectal divisions of the more extensive language groups in Iran are increasingly better understood due to renewed extensive field work, and data bases. That is the case in particular in regard to the relationship between Northern vs. Central and Southern Kurdish, and the re-assessment of earlier division of Balochi (Jahani and Korn, cf. Chapter 11). More recent yet is the recognition of the distinction between Northern Talyshi vs. Central and Southern Talyshi; the separate



MAP 2.5 LOCATOR MAP



MAP 2.6 THE IRANIAN LANGUAGES

status of Eastern Gilaki vs. Western Gilaki; the distinction among the dialects of Mazandarani. (cf. Stilo 1981; 2001; 2005; 2006; ms.), and the four-fold division of the Central Plateau dialects (for the latter cf. also LeCoq 1989b; Windfuhr 1992a). Sharper dialect boundaries have been highlighted, such as those between Talyshi and Gilaki (Stilo), and between the Larestan dialects, which have a number of Tatic features, and the Fars dialects (Windfuhr 1999; 2004).

The following list of the Iranian languages and their sub-groups is arranged according to diachronic, geographic, and dialectological parameters. The sub-groupings of the modern Tatic, Caspian, and Central (Plateau) languages mostly follow Stilo.

Notes on terminology used here:

Northwest = conventional cover term for all of West Iranian languages other than Southwest Iranian
Northeast = conventional cover term for Iranian languages north of Iran and Afghanistan
distinct = locally confined, largely or partially distinct from surrounding group
isolate = remnant in non-Iranian language group
onomastic = only onomastic documentation

OLD IRANIAN

Southwest: Old Persian

Northwest: Median, onomastic

Central: Avestan

Northeast: Scythian/Saka, onomastic

MIDDLE IRANIAN

Middle West Iranian

Southwest: Middle Persian

Northwest: Parthian

Middle East Iranian

Bactrian

Sogdian

Khwarezmian (Khwarezmian), also Choresmian

Khotanese and Tumshuqese (Khotānese)

Middle Scythian and Sarmatian, onomastic

NEW IRANIAN

WEST IRANIAN

NORTHWEST IRANIAN, TIER 1

Upper Zagros and Central Plateau Group

Zazaki, endonym Dim(i)li (Zāzāki)

Kurdish

Northern Kurdish, called Kurmanji (Kurmānji)

Central Kurdish, called Sorani (Sōrāni)

Southern and Southeastern Kurdish

Gorani (Gōrāni), also Gurani

Auroman group, in Kermanshah area, Iran (Hawrāmān)

Bajalani, in Mosul area, Iraq (Bājalāni)

SOUTHWEST IRANIAN

Persian and Varieties

Persian

Tat Persian, in SE Caucasus (Tāt)

Khorasani Persian (Khorāsāni)

Varieties on both sides of the Iran-Afghan border

Afghan varieties, incl. Dari and Kaboli (Kāboli), etc.

Tajik (Tājik ~ Tājiki)

PERSIDE GROUPS, SOUTHERN ZAGROS AND FARS

Shushtari and Dezfuli, distinct

Lori-type languages

Fars dialects (Fārs)

Sivandī, distinct

Davani, distinct (Davāni)

NON-PERSIDE GROUPS, LARESTAN AND GULF

Larestani (Lārestāni)

Gulf Group

Bandari

Minabi (Minābi)

Bashkardi

Kumzari, on Masandam pensinsula in Oman (Kumzāri)

*NORTHWEST IRANIAN, TIER 2**Tatic Group* (Tātic)Talyshi (Tālyshi, also Taleshi)*Northern**Central**Southern*Tati, also called Azari (Tāti, Āzari)*Northern*

Harzani, isolate in NE Azerbaijan

Keringani, isolate in N Azerbaijan (Keringāni)

Western

Khoi dialects, Zanzan province

Southeastern

Khalkhal dialects, in Eastern Azerbaijan

Upper Tarom dialects in Zanzan province (Tārom)

Rudbar dialects in Sefid Rud Valley, transitional to the Gilaki (Rudbār)

Eastern

“Southern” Tati, in Ramand district of Qazvin

Eshtehardi, in Karaj district of the Central [Tehran] province (Eshtehārdi)

Dialects N and NE of Qazvin

Transitional clusters, *Tafresh region east of Hamadan and south of Save*

Vafsi-cluster

Ashtiani (Āshtiāni), Kahaki, Amore’i (Āmore’i)

Alviri-Vidari

Central (Plateau) Dialects

Northwestern
 Northeastern
 Southwestern
 Southeastern

Eastern Iran, Afghanistan, Pakistan

Khuri dialects (Kurdish features), in Kavir Desert

Balochi (Balōchi, also Baluchi)

Western
 Southern
 Eastern

CASPIAN DIALECTS

Galeshi, spoken by mountain herdsmen along the Alborz (Gāleshi)

Gilaki

Western

Eastern

Taleqani-Tonkaboni, etc., transitional (Tāleqāni, Ton(e)kāboni)

Mazandarani, former name Tabari (Māzandarāni)

Gorgani, extinct since 16th century (Gorgāni)

SEM NAN AREA

Sangesari, distinct

Semnani, distinct (Semnāni)

Sorkhe'i and Lasgerdi (Lāsgerdi) with Aftari, distinct

EAST IRANIAN

Pashto

Pashto proper (Ch. 13)

Northwestern

Northeastern

Southwestern

Southeastern

Waḡetsi Pashto, in SE

Pamir Area

North Pamir Group

Yazghulami (Yazghulāmi) with Wanji, extinct (Wānji)

Shughni-Rushani group, fuller term Shughnani (Shughnāni) including

Shughn(an)i with Bajuwi, Barwozi, Rushani, Khufi, Bartangi, Roshorvi, and Sarikoli, in Chinese Xinjiang-Uygur province

Ishkashimi group

Ishkashimi proper (Ishkāshemi), Sanglich and Zebaki (extinct)

Wakhi Varieties

Wakhi, distinct

Related to the other three

Yidgha and Munji

SOUTHEAST IRANIAN

Parachi (Ch. 12) (Parāchi)

Ormuṛi ~ Ormulji (Ōrmuṛi)

NORTH IRANIAN

Yaghnobi, in NW Tajikistan (Yaghnābi)Ossetic, in central Caucasus**1.4 Predecessors of Modern Iranian languages**

Evidence for medieval stages mainly of numerous Western Iranian languages and dialects in particular are found in relatively brief, often poetic, citations in Classical and Early Modern Persian sources.

In general, even without any pre-modern evidence, the modern languages can safely be assumed to continue lost local and regional languages. For example, the modern languages of Azarbaijan and Central Iran, located in ancient Media Atropatene and Media proper, are “Median” dialects, even though Old Median is known mainly from Medisms in Old Persian. The traditional term Pahlav/Fahlav for dialect poetry and other samples of locales in western Iran found in medieval Persian sources reflects the Parthian period in these regions.

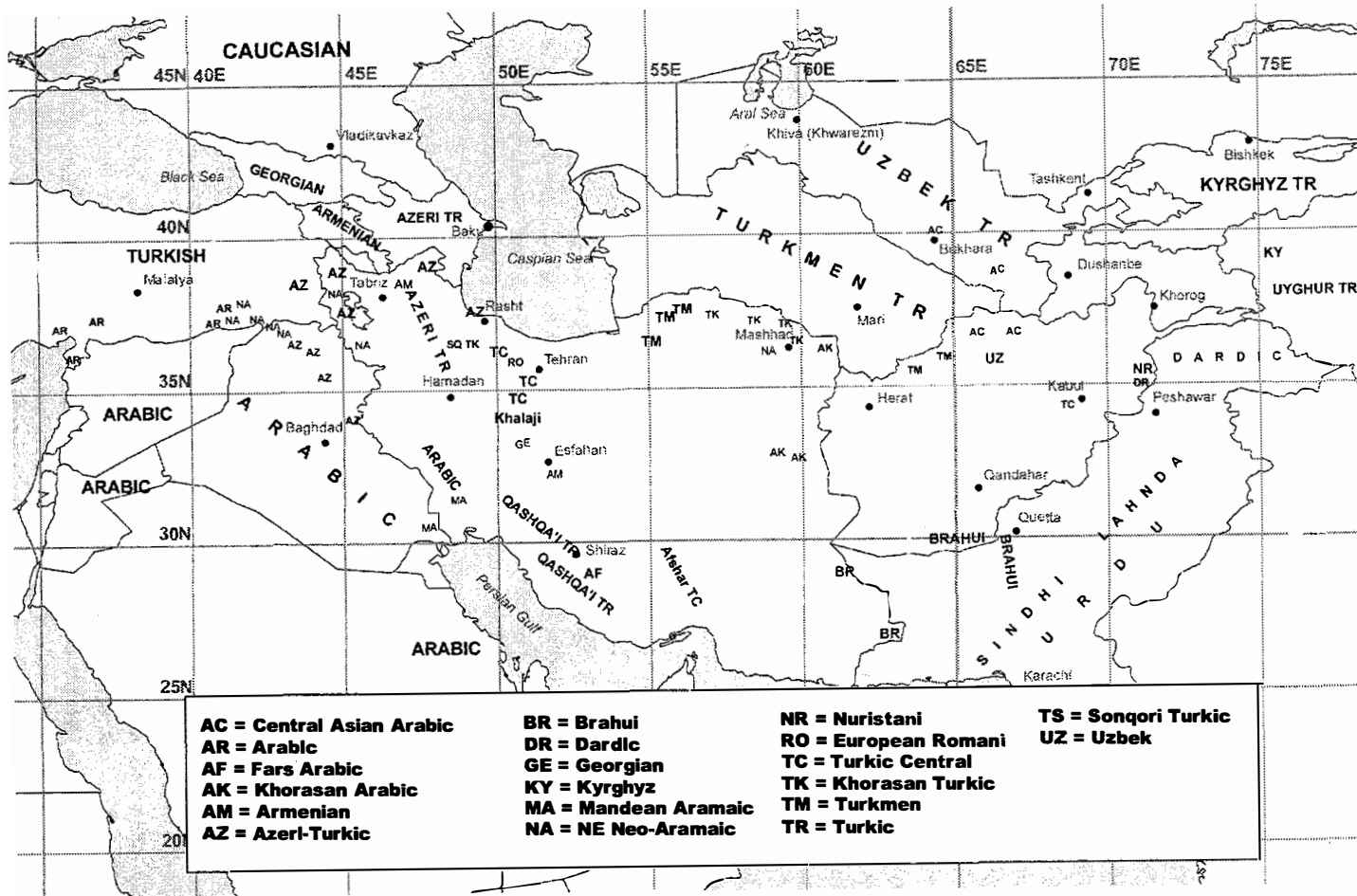
In addition to the continuation of Middle Persian in New Persian, three small modern languages show significant grammatical and lexical reflexes of other documented Middle Iranian languages: In Iran, Sangesari of the Semnan group shares a distinct set of features with Khwarezmian. In the east, Yaghnobi in Tajikistan continues a dialect of Sogdian, and Wakhi in the Pamirs shows distinct reflexes of Khotanese and Tumshuqese Saka. In fact, Wakhi is an example of the repeated invasions of Saka since antiquity.

Little is known about the languages that were erased by the progression of varieties of Persian into eastern Iran and much of western and central Afghanistan, and the expansion of Pashto inside Afghanistan (cf. Kieffer 1989). Only Southeast Iranian Parachi and Ormuṛi still survive in a few islands, but are threatened by extinction (see Map 2.6 and Chapter 12 Parachi). That they represent remnants of once continuous language areas, Parachi in the north and Ormuṛi in the south, is suggested not only by toponymy, but also by linguistic features. In particular, the retention of initial voiced stops (unlike their fricativation in other Eastern Iranian groups) suggests earlier transitional contact somewhere in central Afghanistan with non-Perside, i.e. “Northwest”, Iranian language groups, now also erased. At the same time, morphology and morpho-syntax reflect successive contacts with Pashto for Ormuṛi, and for both contact with Indo-Aryan Dardic languages, and later superstrate interference from Afghan Persian varieties.

1.5 Modern non-Iranian contact languages

Bilingualism and even multilingualism are the norm in many regions, given that Persian is the superstrate language, and that varieties of Azeri Turkic are widely used as the language of the bazaar. Identity therefore is determined by complex intersecting layered patterns of cultural, ethnic, and linguistic affiliations.

There were two successive periods of major overlays of speakers of non-Iranian languages over most of the Iranian linguistic ground. Both covered the entirety of Greater Iran from the western to the eastern regions, including Greater Khorasan, Afghanistan and Central Asia. The first was the Arabic overlay which began before Islam and continued for the next few centuries. The second was the Turkic overlay which began in the eleventh century.



MAP 2.7 MODERN NON-IRANIAN LANGUAGES

Arabic

Literary Arabic, as the administrative and literary superstrate during the early centuries of Islam, and as the language of the Qor'an, left an indelible impact on Persian and, through it, on virtually all language of Greater Iran, prominently in their lexicon, but also in morpho-morphology. However, due to a millennium of assimilation, the local Arabic overlay in the heartland and hinterland of Iran is now reduced to a few pockets in eastern Iran (as well as northern Afghanistan and Central Asia). Quite distinct is the situation of the southwestern regions and along the littoral of the Persian Gulf in direct contact with, or close to, the Arabic-speaking regions of Iraq and the Gulf states.

Turkic

In terms of numbers, of a present population of ca. 70 million, Turkic speakers constitute an estimated nearly third of the total population of Iran (compared to Persian with ca. 50 percent) Once Turkic speakers had settled first in Greater Khorasan in the northeast, then in Azerbaijan in the northwest, and then mostly from there in Fars in the southwest, they began to overlay the Iranian languages. Historically and geo-linguistically, this triangular pattern of Turkic settlement repeated the pattern of the original Iranian settlement some two thousand years earlier.

These expansions during the last millennium resulted in the Turkicization of most of the Iranian-speaking areas in Azerbaijan, leaving a few Iranian-speaking pockets, as was the case in many of the Tati-speaking areas in western Iran, including the areas along the Alborz up to Qazvin and the areas east of the Zagros mountains. Moreover, dynasties of original Turkic speakers ruled the country for most of the second millennium, most importantly the Safavids. As such, Turkic has had a distinct impact on Persian and other Iranian languages, both in the lexicon and grammar, particularly in the border provinces in contact with Turkic in the northeast and the northwest of Iran and beyond. Here may also be mentioned groups of Altaic Mongols who settled in Iran, mainly on Iranian Kurdistan, and throughout Afghanistan, including the now Persian-speaking Hazara in Central Afghanistan and the groups called Mogholi (cf. Kieffer 1983 and 2003), and linguistically left some traces in the Persian lexicon.

Indo-Aryan

In the east, Iranian languages have been in continued contact with speakers of Indo-Aryan. All of these show various degrees of Indo-Aryan features on all linguistic levels and the lexicon and, significantly, typological shifts. In particular, Dardic and Nuristani languages have been in long contact with the Pamir languages, in a multilingual Sprachbund.

Dravidian

Brahui is spoken in Iranian Baluchestan and in the adjacent areas of Afghanistan and Pakistan. The language is the northernmost remnant of the Dravidian languages which are now found mainly in the southern Indian subcontinent, but prehistorically may once have been more widely found in Iran. This is suggested in particular by the possible

distant relationship between Proto-Elamite and Proto-Dravidian. It is therefore not unlikely that the unknown ancient non-Iranian languages in the South of Iran from Elam to India may have included Dravidian speakers.

Minor linguistic groups

These include small communities of Neo-Aramaic speakers, both Jewish and Christians, in Iranian, Turkish and Iraqi Kurdistan, and Neo-Aramaic Mandeans in Khuzestan, now largely reduced by emigration, and small communities of Armenians in northern Azarbaijan and in Esfahan, close to which exist small rural Georgian communities. European Romanis, deported some 200 years ago from Bulgaria, then part of the Ottoman Empire, are found in two villages near Qazvin and Khorasan.

2 PHONOLOGY

2.1 Early Iranian dialects

The basic dialect division of the Iranian languages today goes back to the earliest stages of Iranian (cf. also Chapter 3 Old Iranian, sections 1.2 to 1.4). A primary development in the shift from Indo-Iranian to Proto-Iranian included the shift of the palatal affricate pair $\acute{c} \acute{j} > ts dz$. Proto-Iranian further split into at least four distinct Iranian dialect groups, defined by the divergent developments of the palatal clusters $*\acute{c}w/*\acute{j}w$, and further by the absorption of the semivowel w in the diametrically opposite NE and SW groups. The two “upper” dialects are correlated with the regions/languages of the Black Sea/Ossetic and the Tarim Basin/Khotanese-Tumshuqese, and two “lower” dialects with central Old Iranian regions/Avestan, and Fars in the southwest/Old Persian. The Scythian development here is that reflected in Ossetic.

TABLE 2.2: THE FOUR OLD IRANIAN DIALECTS

	Upper Tier				Lower Tier			
	NW Ossetic		NE Khotanese		SW Old Persian		Central/Avestan	
1	$*\acute{c}w$	$*\acute{j}w$	$*\acute{c}w$	$*\acute{j}w$	$*\acute{c}w$	$*\acute{j}w$	$*\acute{c}w$	$*\acute{j}w$
2	$t\acute{s}w$	$\acute{z}w$	$t\acute{s}w$	$d\acute{z}w$	$t\acute{s}w$	$d\acute{z}w$	$t\acute{s}w$	$d\acute{z}w$
3	$\acute{s}f$	$\acute{z}w$	$\acute{s}\acute{s}$	$\acute{z}\acute{z}$	ss	zz	sw	zw
4	fs	vz	\acute{s}	\acute{z}	s	z	sp	zb

2.2 West Iranian consonantal developments

The basic differentiation of the West Iranian languages into Northwestern and Southwestern, i.e. Perside, dialects was pioneered by Tedesco (1921) and Lentz (1927). Following a number of other studies, including Hadank (edited posthumously by Selcan 1992), MacKenzie (1961) outlined basic internal relationships between the dialect groups, followed by Windfuhr (1975). Among more recent studies with focus on phonology, LeCoq (1989) suggested a detailed modified classification. Paul (1998: 170) showed major phonemic isoglosses of the main West Iranian dialect groups. Korn (2003: 59) showed the

major isoglosses that determine the position of Balochi, followed in the same volume by Paul's (2003) pendent on major lexical isoglosses. All authors duly include caveats, emphasizing the provisional nature such studies.

The following table of selected West Iranian phonological isoglosses is arranged from Northwest Iranian to Southwest Iranian. For ease of comparison, the table is doubly framed by their Middle and Old Iranian pre-forms. The sequence of dialects and features is arranged to best visualize the patterns, and does not follow the probable historical sequence of the innovations, such as the one suggested by Korn (2003) which is: Old Iranian, (1) *ts*; (2) *dz*; (3) *tsw* (and *dzw*, not included here) (4) *θr*; late Old Iranian and Middle West Iranian (5) *rdz*; (6) *rd*; (7) *dw*; (8) *hw*; (9) *θw*. Compare also Map 2.6 Iranian Languages. Shifts towards Southwest Iranian are marked by underlined capital letters.

In overview, except for *rz*, *rd*, this table shows that by and large the sets of innovations form a continuum from Parthian-type to Middle Persian-type changes:

- (1) Zazaki and Gorani are the least innovative in relation to Parthian.
- (2) The middle group constitutes a fairly unified set, except for the late lateralization of *rz* and *rd* to *l* in the latter three groups.
- (3) Balochi, and more so Kurdish, are transitional to Persian:
- (4) *hw*: Zazaki and Balochi have *w*, and the remainder mostly has reflexes of *hu*, with variation in the Central dialects and Balochi.

TABLE 2.3: SELECTED WEST IRANIAN ISOGLOSSES

*Indo-Iranian	<i>j</i>	<i>čw</i>	<i>tr</i>	<i>ś</i>	<i>-tw-</i>	<i>sw</i>	<i>dw</i>	<i>rj</i>	<i>rd</i>
*Proto-W Iranian	<i>dz</i>	<i>tsw</i>	<i>θr</i>	<i>ts</i>	<i>-θw-</i>	<i>hw</i>	<i>dw</i>	<i>rdz</i>	<i>rd</i>
NORTH-WEST									
Avestan	<i>z</i>	<i>sp</i>	<i>hr</i>	<i>s</i>	<i>θw</i>	<i>hw</i>	<i>dw</i>	<i>rz</i>	<i>rd</i>
(Middle) Parthian	<i>z</i>	<i>sp</i>	<i>hr</i>	<i>s</i>	<i>f</i>	<i>w^h</i>	<i>b</i>	<i>rz</i>	<i>rd</i>
Zazaki	<i>z</i>	<i>sp</i>	<i>(h)r</i>	<i>s</i>	<i>w</i>	<i>w</i>	<i>b</i>	<i>rz</i>	<i>ř</i>
Gorani	<i>z</i>	<i>sp</i>	<i>hr</i>	<i>s</i>	<i>u</i>	<i>w</i>	<i>b</i>	<i>rz</i>	<i>r/L</i>
MIDDLE GROUP									
Talyshi	<i>z</i>	<i>sb</i>	<i>hr</i>	<i>s</i>	<i>h</i>	<i>x(u)</i>	<i>b</i>	<i>rz</i>	<i>r</i>
Tati (Azeri)	<i>z</i>	<i>sb</i>	<i>hr</i>	<i>s</i>	<i>h</i>	<i>x(u)</i>	<i>b</i>	<i>rz</i>	<i>r</i>
Central Dialects	<i>z</i>	<i>sb</i>	<i>hr</i>	<i>s</i>	<i>h</i>	<i>xu,h,f</i>	<i>b</i>	<u><i>L(rz)</i></u>	<u><i>L</i></u>
Caspian	<i>z</i>	<i>sb</i>	<i>r</i>	<i>s</i>	<i>h</i>	<i>x</i>	<i>b</i>	<i>rz</i>	<i>rd</i>
TRANSITION									
Balochi	<i>z</i>	<u><i>S</i></u>	<u><i>S</i></u>	<i>s</i>	<i>h</i>	<i>w, h</i>	<i>b</i>	<i>rz</i>	<i>rd</i>
Kurdish	<i>z</i>	<u><i>S</i></u>	<u><i>S</i></u>	<i>s</i>	<i>h</i>	<i>xw</i>	<u><i>D</i></u>	<i>l</i>	<i>l</i>
SOUTH-WEST									
Modern Persian	<u><i>D</i></u>	<u><i>S</i></u>	<u><i>S</i></u>	<u><i>H</i></u>	<i>h</i>	<i>x</i>	<u><i>D</i></u>	<i>l</i>	<i>l</i>
Middle Persian	<i>d</i>	<i>s</i>	<i>s</i>	<i>h</i>	<i>h</i>	<i>xw</i>	<i>d</i>	<i>l</i>	<i>l</i>
Old Persian	<i>d</i>	<i>s</i>	<i>s'</i>	<i>θ</i>	<i><θuv>*</i>	<i><uv>*</i>	<i><duv>*</i>	<i>rd</i>	<i>rd</i>
*Proto-W Iranian	<i>dz</i>	<i>tsw</i>	<i>θr</i>	<i>ts</i>	<i>-θw-</i>	<i>hw</i>	<i>dw</i>	<i>rdz</i>	<i>rd</i>
*Indo-Iranian	<i>j</i>	<i>čw</i>	<i>tr</i>	<i>ś</i>	<i>-tw-</i>	<i>sw</i>	<i>dw</i>	<i>rj</i>	<i>rd</i>

* Old Persian spelling.

- (5) *rʒ* and *rd*: The contraction of these clusters to laterals does not correlate as well as the other columns with the sequence of innovations, suggesting intersecting isoglosses or typologically independent change.

Overall then, from Zazaki and Gorani downward to Persian, the dialects are successively less “archaic”, each subgroup accumulating additional innovations. At the same time, Persian has been increasingly the superstrate language, and as such may have most directly affected dialect groups in contact with it at various historical periods. This was probably the case with the “transitional” Northwestern pair, Balochi and Kurdish, which have the Southwestern features *tsw* and *θr > s*, *θw > h*. Kurdish, in addition, also shares *dw > d* and *hw > xw* with Persian (where later *xw > x-*). As convincingly argued by Korn (2003), the two must have acquired those features when in contact with Persian for a considerable period of time, longer for Kurdish because of the additional two features (for such contact cf. also MacKenzie 1961 and Windfuhr 1975).

2.3 Innovations: SW drift vs. NW

The successive innovations are essentially cluster reductions. Of the nine clusters discussed, the contraction of *r*-clusters may be less significant given that lateralization may occur as independent innovations.

Of the remaining seven clusters, more than half, i.e. four, are *Cw*-clusters. In all of these *w* is lost in the Southwestern dialects. That development is opposed to the Northwestern innovations where (1) the labial is retained with partial devoicing, and (2) initial consonants tends to be lost. These innovations can be seen as divergent drifts that began with *tsw* and ultimately extended to the other three. Windfuhr (1972: 58–59) formulated a “tendency” rule, which also included loss of *w* in the clusters *pw/bw* and *gw* (not included in the table here).

The phonotactic conditions for the SW-NW divergence appear to be the position of the articulatory focus in these clusters: in the SW on the initial component, but in the NW on the final component of the cluster. This conclusion is also suggested by the Old Persian and Parthian “orthographies”, that is, their representation of the cluster **h-w*: It is spelled <uv> in OP. (similar OP. <θuv>, <duv> for **θ-w*, **d-w*) which suggests /hu/. But in Parthian it is spelled <wx> which points to /w^h/ as suggested by MacKenzie (1967: 26 n. 29).

In fact, the same articulatory conditions may account for the NW–SW divergence in other clusters as well, beginning with the divergent four-fold Early Iranian changes discussed above (see Table 2.2).

TABLE 2.4: SW vs. NW IRANIAN CLUSTER REDUCTIONS

SW	NW	SW	NW	SW	NW
	<i>*dz</i>		<i>*dzw</i>		<i>*dw</i>
<u>d</u>	<u>z</u>	<u>d</u>	<u>zb</u>	<u>d</u>	<u>b</u>
SW	NW	SW	NW	SW	NW
	<i>*ts</i>		<i>*tsw</i>		<i>*hw</i>
<u>h</u>	<u>s</u>	<u>s</u>	<u>sq</u>	<u>xw</u>	<u>w^h</u>

2.4 East and West Iranian

The developments in the East Iranian languages shows a subset of changes identical to those of the Northwest languages. But the dental *w*-clusters and *tr* do not, and may presuppose splits prior to the Iranian fricativation of preconsonantal stops. In turn, Sangesari and the two Southeast languages Parachi and Ormuṛi diverge in their palatalization of the *tr*-cluster, as shown schematically in Table 2.5 (cf. Skærvø 1989: 377):

TABLE 2.5: SOME EAST vs. WEST IRANIAN CLUSTER REDUCTIONS

Pt.-Ir.	* <i>ts</i>	* <i>dz</i>	* <i>tsw</i>	*- <i>tr</i> -
East	<i>s</i>	<i>z</i>	<i>spfs</i> ¹	<i>ārt</i> ³ * <i>θp</i> ⁴ <i>tr</i> ⁵ <i>dr</i> ⁶ <i>ts</i> ⁷ (<i>h</i>) <i>ar</i> ⁸ <i>r</i> ⁹ <i>čr</i> ¹⁰ <i>š</i> ^{r11} <i>š</i> ¹²
NW Ir.	<i>s</i>	<i>z</i>	<i>sp</i>	<i>hr</i>
SW Ir.	<i>h</i>	<i>d</i>	<i>s</i>	<i>s</i>

(1) Ossetic; (2) Wakhi; (3) Ossetic; (4) Yaghnobi; (5) Wakhi; (6) Pashto; (7) Yazghulami; (8) Shughni group; (9) Sanglichi-Ishkashimi; (10) Yidgha-Munji; (11) Ormuṛi; (12) Parachi.

Sims-Williams (1989: 168–169) and Steblin-Kamenskii (1992) highlight various major phonological developments from Old Iranian to common Middle and New East Iranian. These mostly involve word boundaries and clusters and may be arranged as follows (note that they do not apply to Southeast Iranian Parachi and Ormuṛi).

- (1) In initial position the voiced stops become fricative: OIr. #*b dg* > *β δ γ* (with various further changes).
- (2) There is widespread, though widely varying, palatalization of consonants.
- (3) Contraction of clusters is distinctly more prominent than in WIr (and similarly, palatalization of vowels).
- (4) The phonetic clusters of the palatal affricates become dental, but not in Sogdian: OIr. *č ĵ* > *ts dz* (further > *s z*).
- (5) The voiceless fricatives *f x* become voiced before *t*: OIr. *ft xt* > *βd (> ud) yd*.

In West Iranian, similar isoglottic changes occurred under various conditions across large regions (cf. Windfuhr 1989b: 253–254).

Cluster lenization: (1) Throughout West Iranian both **ft, *xt* > *ht* > *t*, and further changes, e.g. Bakhtiari (Lori-type) *raft-am* > *rahd-om* ~ *rε:-m* ‘I went’. Moreover, initial **fr, *xr* > (*h*)*r* in most of West Iranian, except Kurdish, Gilaki and Mazandarani, and Persian and Perside in the SW (for a map, cf. Azami and Windfuhr 1972: 199).

Fronting of palatals *č ĵ* > *ts dz*: More recent, it is found in the central NW in contact with Turkic (Stilo 1994, with map).

Initial clusters: The modern reflexes of Middle and Early New Iranian #CC-clusters show a distinct areal distribution: insertion of a short vowel, CVC-, along the Zagros, including the NW tier 1 from Kurdish, Zazaki to the SW Fars and Larestan dialects, as opposed to initial vowel, VCC-, elsewhere, e.g. *šVmā* vs. *Všmā* ‘you-p’. In the east, Balochi does allow initial clusters, as do most East Iranian languages.

2.5 Stress

The only modern Iranian language to have retained phonemic stress is Pashto. In West Iranian, a major change from Old to Middle Iranian was the loss of phonemic stress due

to the shift of stress to the penult or antepenult syllable. Thereby, derivational as well as inflectional formants and endings were contracted, and inflectional classes in the nominal and verbal systems were reduced to the nominal **a*-declension and the thematic **a*-conjugation, or were lost. Similarly, contractions occurred in East Iranian, notably in nominal and verbal forms with the formant of OIr. **-aka*.

Initial vocative stress is found throughout Iranian and inherited from Indo-European, but is also found in symbiotic Turkic and Semitic.

3 MORPHOLOGY AND SYNTAX

3.1 Gender and animacy

Gender retention is found in most Northwest Iranian and in East Iranian languages. In West Iranian, these include the languages along NW Iranian tier 1: Zazaki, N Kurdish, Gorani, and along the NW tier 2: Talyshi, Tati, the northwestern Central dialects, and Semnani and the enclave Sangesari. In East Iranian they include Pashto, the North and South Pamir groups, except the exclave Sarikoli, Wakhi, and Ormuři. Gender is lost in all SW Iranian, that is, all Persian varieties, Central and Southern Kurdish, Caspian Gilaki and Mazandarani, and Balochi, and East Iranian Ossetic and Yaghnobi, and Southeastern Parachi. The degree of gender retention varies. For example, plural gender is retained only in Pashto and Ormuři.

Noteworthy is the masculine plural marking of complement clauses in Pashto (see Chapter 11, section 3.1), and in Zazaki the feminine marking in both 3s and 2s feminine endings which reflects the origin of the present stem in the OIr. present participle in **-ant* + copula (see 3.5 below).

In addition to natural gender marking, classificatory gender marking contrasts masculine = abstract vs. feminine = concrete in Talyshi-Tati and Sangesari, while feminine includes collectives and animals in Khwarezmian and Shughni (for the complex system in the latter see Chapter 14b, section 3.1.1.1c).

Kinship and a few related terms have the grammaticalized oblique marker *-r* in Zazaki, Tati, Talyshi, Sangesari, e.g. Zaz. *mālmā-r* ‘mother’.

Common Iranian is the human-animacy distinction in interrogative and related pronouns, e.g. Persian *kilči* ‘who/what’ and *ān-kelān-če* ‘he who, that which’, inherited from Indo-Iranian.

Inherited from Indo-European is plural agreement in the verb for animates, but singular marking for inanimates. This rule is optional in Persian, and other modern languages, which allow for plural marking for inanimate individuation (already Middle Persian).

Semantic classification is found in spite of loss of gender marking, expressed by specializations of alternate plural markers. Thus, in Persian, human may be morphologically distinguished by pl. *-ān* (< OIr. gen. pl. **-ānām*) vs. the generalized plural marker *-hā*. It may optionally be the literary plural marker of humans, larger animals, birds, and trees, and dual marker for body parts, e.g. *može-g-ān-e u* ‘her eyelashes’. In the Pamir languages, human groups and a few other animate sets have differential plural markers (see Chapter 14b). In Parachi, human is distinguished in the genitive (oblique) singular *-ān* vs. *-eka*, and the existential verb. In Sogdian, human distinguishes animate pl. *-št* vs. *-t*. Similar marking is found elsewhere.

3.2 Cases and personal enclitics

For the following cf. also Windfuhr (1990).

The eight-case system of the Old Iranian *a*-declension was generalized. By or during the Middle Iranian period, a binary direct vs. oblique case system emerged. The latter marker is the outcome of the fusion of the case endings into a single case marker, mostly *-ē* or *-ī* in the singular, phonetically from the masculine genitive-dative, and plural *-ān* from the masculine genitive plural **-ānām* in a large number of languages. As a fused case opposed to the unmarked direct case, the oblique continued the functions not only of the genitive-dative, but also those of other non-nominative cases, including the function as marker of the innovative ergative agent (for which see section 4.3).

In a number of languages, the singular oblique marker was generalized and attached to the plural marker, here shown with the examples of languages with plural markers other than **-ān* (for Balochi, see Chapter 11, Table 11.11, for Hunza Wakhi see Chapter 15, Table 15.3, and for Sogdian Chapter 5, Table 5.5):

	Balochi		Hunza Wakhi		Late Sogdian	
	Sg.	Pl.	Sg.	Pl.	Sg.	Pl.
DIR	-∅	-∅	-∅	-išt	-∅	-t
OBL ₁	-ā	-ān	-∅	-ev	-ī	-t-ī
OBL ₂	-ā	-ān-ā	-e	-ev-e		
GEN	ay, -ī	-ān-ī				

A distinct feature of the Iranian languages is the set of personal enclitics inherited from Old Iranian. These function as person markers in all oblique cases, including possessor, indirect object, direct object, and the ergative agent, just like the generalized oblique case (see section 4.3 below).

Therefore the loss, or absence, of these enclitics is both historically and typology of major significance, and can best be accounted for only by the assumption of strong and lasting substrate typology. As shown in Map 2.8, it is found in northern West Iranian, beginning with Zazaki and Northern Kurdish, the Northern Tati isolate Harzani, then further east along the Caspian shore in Talyshi, Gilaki, Mazandarani, also in Semnani and Sangesari. It is likewise found in Tat Persian, isolated in the Republic of Azerbaijan and Daghestan (see also Map 3, “Pronominal possessive forms in Iranian and non-Iranian languages” in Stilo 1981: 163).

These Iranian groups are typologically quite distinct from each other; in particular they include a variety of Persian. Given the fact that the same absence of pronominal enclitics is found in South Caucasian, it is apparent that this Iranian isogloss is a substrate feature that continues the effects of the pre-Iranian Caucasian (for possible Caucasian substrates in earlier Iranian languages on the eastern Caspian shores, now mostly Turkic, cf. Windfuhr 2006: 383–384).

3.3 Deixis

A distinctly eastern feature is triple vs. binary western deixis, with the exception of Yaghnobi and Ossetic, and Yazghulami adjacent to Tajik. Speaker orientation is copied into a triple quasi-pronominal system of *ich-*, *du-*, *er-*deixis in Pashto (Chapter 13, section 3.1.3.1), Ormuṛi, and is found already in Late Middle Iranian Khotanese (Chapter 7, section 3.2.1.2).

3.4 The verbal quincunx system

The history of the parameters and axes of the verb systems from Old Iranian to Modern Iranian shows a cycle from a five-member quincunx to varying Middle Iranian systems back to a quincunx (Windfuhr 1980, 1986, 1987). The development is shown here with the example of Persian.

The inherited fundamental and primary verbal parameter of the Early Old Iranian system is triple aspect which intersects with the binary tense parameter of present and past (marked by the augment *a-*). It is centered on the perfective aorist:

Early Old Iranian	Present	Past	
Imperfective	PR	<i>a</i> -PR	“Present system”
Perfective		AOR	“Aorist system”
Resultative-stative	PF	(<i>a</i> -PF)	“Perfect system”

In time, this triple aspect system was reduced to forms of the “present” system, i.e. imperfect present and imperfective past, leaving only a few forms of the aorist and the perfect. With their loss, the highly complex inherited system was reduced to a single imperfective stem, distinguishing present vs. augmented imperfect: PR vs. *a*-PR.

Concomitantly, however, the vacated aorist and perfect ranges of the system were partially filled by the innovation of a new perfective system based on the adjectival completive participle in *-tá* plus the present and past copula, with both intransitive and transitive verbs (see Chapter 3, Old Iranian, sections 7.4.2.2 and 7.4.3.2).

In Middle Persian, the resulting four-member system of two imperfective and two perfective forms was extended by replacing the copula with the stative verb *ēst-* ‘to stand’. The outcome was a six-member system with a triple aspect axis and a binary tense axis:

Middle Persian	Present	Past		
Imperfective	<i>raw-</i>	(<i>a</i> - <i>raw-</i>)	present	imperfect (later lost)
Perfective	<i>raft</i> COP	<i>raft būd</i> COP	preterit	past preterit
Resultative-st.	<i>raft ēst-</i>	<i>raft ēstād</i> COP	perfect	pluperfect

In addition, the adverb *hamē*, lit. ‘forever’ (see 3.4 below) expressed ongoing and progressive action as well as continuing state, while its pendant *bē* (homophonous with the adverb *bē* ‘out, away’) expressed the singularity of an event in present and past, and assumed inchoative or future connotation with the present stem (see Chapter 4 Middle West Iranian, sections 4.4.8.1 and 4.4.8.3).

In Early New Persian (*ha*)*mē-* and *bē-* were continued, but the periphrastic resultative *ēst-* forms were replaced by extended forms based on the verbal adjective in *-tag* (< **-ta-ka-*). *bi-* and *mē-* could still occur with these verb form, and neither was obligatory. The core system in terms of frequency was the following:

ENP	Present	Past		
Imperfective	<i>mē-raw-</i>	<i>mē-raft-</i>		
Perfective	<i>bi-raw-</i>	<i>bi-raft-</i>	inchoat.-fut.	singularity
Unmarked	<i>raw-</i>	<i>raft-</i>	gen. present	gen. past
Resultative-st.	<i>raft-a</i> COP	<i>raft-a bud-</i>		

Subsequently the system was restructured by the coalescence of the unmarked forms with the perfective forms by the fifteenth century.

- (1) In the present, the perfective *bi*-form assumed distinct subjunctive function, alternating with the unmarked general present form, now opposed to the indicative present-future *mē*-form.
- (2) In the past, the general unmarked form subsumed the function of the *bi*-form to express both general and perfective events, now opposed to the imperfective *mē*-past form. It thereby assumed the central role of an aorist in the resulting five-member system.

This core of the system became thus as follows, and has not changed since:

INDICATIVE

Pre-Modern	Present		Past
Imperfective	<i>mē-rav-</i>		<i>mē-raft-</i>
Perfective		<i>raft-</i>	
Resultative-st	<i>raft-a COP</i>		<i>raft-a bud-</i>

The non-indicative sub-system developed in parallel to the indicative core, using the imperfect and past perfect forms for irreal function, and using the present subjunctive of 'to be' for the perfect subjunctive:

NON-INDICATIVE

Pre-Modern	Present		Past
	Subjunctive		Irreal
Imperfective	<i>bi-rav-</i>		<i>mē-raft-</i>
Perfective		<i>raft-</i>	
Resultative-st.	<i>raft-a bāš-</i>		<i>raft-a bud-</i>

Most languages developed similar periphrastic forms already during the Middle Iranian stage. An example is the system of Middle East Iranian Khwarezmian. Note that the following synoptic table (see also Chapter 6, Table 6.15) includes typical grammaticalized particles that mark further aktionsart and modal distinctions.

TABLE 2.6: SYNOPSIS OF KHWAREZMIAN VERB SYSTEM

					New Periphrastic Perfect Forms		
PRS	perm.	FUT perm.	PAST	conditional	PF	perm.	RemotePF
IND	+	<i>-kām +</i>	Impf.	<i>-manc-</i>	Ind.	+	
SBJ	+	<i>-kām</i>	Injunc.		Subj.		
OPT	–						
IMP	–						

+ = documented present and future forms which add the permansive particle; – = none of these particles.

Verb systems vary considerably. A most rudimentary is that of the Pamir language Shughni, which consists of only four forms: present-future and past, present perfect and past perfect (see Chapter 14b, section 3.2.5). The most complex systems are found in contact with Turkic, including the system of Tajik Persian (see Chapter 8, Table 8.12).

3.5 Markers of aspect

Grammaticalizations such as the Persian prefix *mi-* reflect the re-emergence of aspect as a primary parameter of the verb system. In ENP *mē-* was the marker not only of ongoing

and progressive events, but also of continuing state. It originates in OIr. **hama-aiva-* ‘same duration, time’. **ham-* alone is found in the Khorasani Persian dialect of Shahrud, apparently unique: *haŋ-ger-om* ‘I am taking’; *ham* is also reflected in the Sogdian and Khwarezmian pre-vocalic imperfective marker *m(a)-* < *ham* + augment *a-* (see Chapter 5, section 3.2.1.2; Chapter 6, section 3.2.5.1a).

Widely found also is the dental marker *at-*. It occurs in Kurdish as *di- ~ a(t)-*, and as *e(t)-* in the Central Plateau languages. This *at-* isogloss further extends down to the Southwest Iranian languages along the Gulf. However, the immediate area of Isfahan has the enclitic *-e* (< earlier *hait*, opt. of ‘to be’; a detailed discussion these markers is found in Stilo (2007:106–108).

In past forms of many West Iranian languages, the imperfective marker contrasts with the perfectivity *ba-*. However, in the Caspian languages, but not in the Semnan group, the past imperfective is marked by the *absence* of *ba-*, and similarly in Pashto, where *perfectivity* is marked by *wə-* and by stress-shift with directional preverbs (see Chapter 13, Table 13.20 and Section 3.2.2.3).

A number of Tati varieties such as Harzani, where a Turkic-type locative progressive constructions based on the infinitive, developed into the general present, the original present was downshifted to subjunctive function (see also section 3.5).

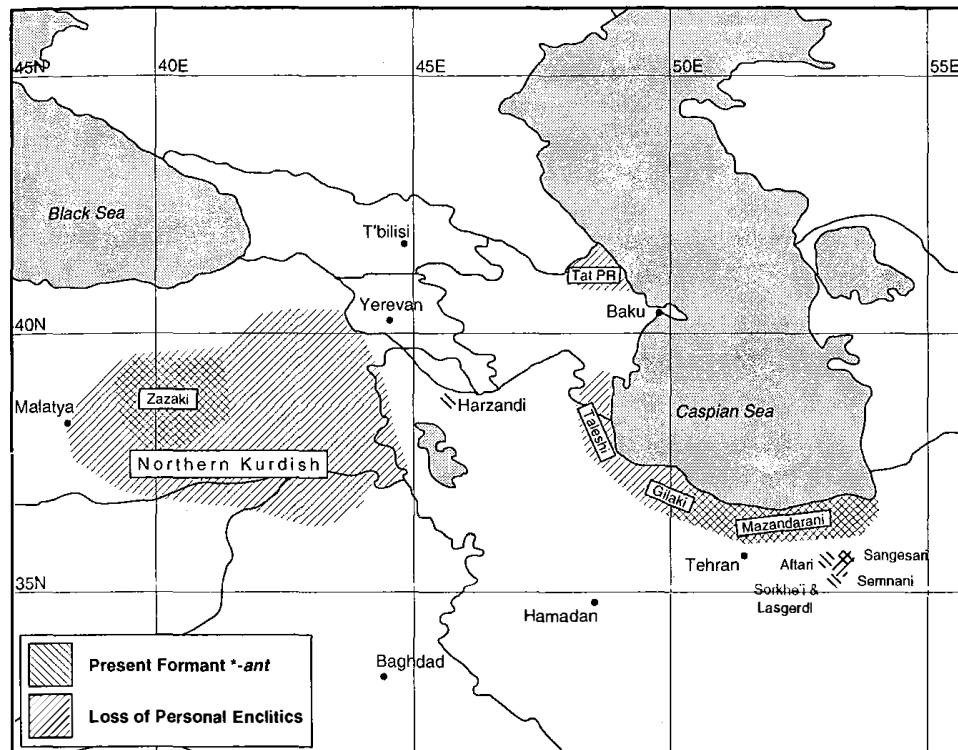
By contrast, a distinct archaism from Old Iranian is found in some languages where even the past imperfective continues to be based on the *present stem* of the verb. These include even heavily Turkicized Talyshi and several member of the Tati group, as well as Gorani, and Yaghnobi in the East. The latter is unique in also retaining the Old Iranian imperfective augment: *a-PR-*.

3.6 Present marker **-ant*

A number of language groups differ from all others in their formation of the present indicative. They include: (1) Zazaki (surrounded by Northern Kurdish), (2) the Northern Tati isolate Harzani within Turkic Azerbaijani, (3) Eastern Gilaki and Mazandarani, and (4) several small communities around Semnan, but not the city itself, including Sangesari, Aftari, Sorkhe’i-Lasgerdi (see Map 2.6). The formation originates in an innovative OIr. progressive construction based on the active present participle in **-ant*, which later took the place of the regular present based on the present stem, e.g. non-progressive Sang. *a vin-ænd-i* ‘I see’ < **azam wain-ant-ah ahmi*. In spite of their separation of the groups today, their original close links, and subgrouping, are shown by shared features, such as contractions, e.g. both Zazaki and Sangesari near Semnan have *wā-nd-/vā-nd-* ‘be saying’ < **wāč-ant-* from **wāč-* ‘to say’, while East Gilaki and Mazandarani have *gum-me* ‘I am saying’ < **gaub-ant-* from the different root **gaub-* ‘to say’. In Harzandi, the participial **-ant*-forms have been downshifted to subjunctive function, and replaced by a new Turkic-type locative progressive construction based on the infinitive + *-(n)d-* < **andar* ‘in’.

These observations suggest a former continuum that was eroded by north-migrating Iranian groups, such as the predecessors of the Talyshi-Tati groups, and yet again by the wave of the Turkic overlay.

As shown on Map 2.8, the **-ant-* isogloss overlaps with the areas showing loss of the personal enclitics. While the latter is probably due to Caucasian, the **-ant* isogloss appears to reflect an earlier Iranian isogloss, for which likely candidates are Scythian/Saka groups incoming through the Caucasus, or Saka incoming from east of the Caspian.



MAP 2.8 LOSS OF PERSONAL ENCLITICS AND PRESENT IN *-ANT

4 SYNTAX

4.1 Word order typology: adjectival noun phrase

Right-branching vs. left-branching typologies of HEAD and MODIFIER intersect the genetic NW vs. SW grouping of Western Iranian. The relative position of adjectives, which tends to be closely aligned with that of dependent nouns and noun phrases is a case in point.

Specifically, right-branching nominal typology is marked by the insertion of a linker between noun and adjective, which may be inflected for number and gender, and even for case as in Zazaki and Northern Kurdish (see Chapter 9, section 4.2.1, and Chapter 10, section 7.2). Left-branching typology is marked by an “inverse *ezāfe*”, i.e. by the insertion of a linker between adjective and noun. As shown in Table 2.7, there are two tiers: one is shared by both SW Iranian and a number of western NW Iranian groups, the other is confined to the remainder of NW Iranian:

- (1) the “Zagros” tier has N- EZ₁ ADJ, which stretches from the Fertile Crescent downward along the Zagros range into Fars and further into southern Iran, and also extends into the center of Iran.
- (2) the “Alborz” tier has ADJ-EZ₂ N, which stretches from Azerbaijan along both sides of the Alborz range and, mostly represented by Balochi varieties, further extends into southeastern Iran, Afghanistan and Pakistan.

Both tiers contrast with the East Iranian “Hindukush-Pamir” tier plus Ossetic which is marked by the absence of a linker, ADJ N.

TABLE 2.7: THE THREE IRANIAN NOUN + ADJECTIVE TYPOLOGIES

“Zagros” Tier N-EZ ₁ ADJ			“Alborz” Tier ADJ-EZ ₂ N			“Hindukush-Pamir” Tier/Ossetic ADJ N		
Zazaki	N- <i>o</i> ¹	ADJ	Tat-Persian	ADJ- <i>e</i>	N	Pashto	ADJ	N
Kurdish	N- <i>ê</i> ²	ADJ	Tatic	ADJ- <i>æ</i>	N	Yazghulami	ADJ- <i>i</i>	N
Gorani	N- <i>i</i>	ADJ	Gilaki	ADJ- <i>ə</i>	N	Pamir	ADJ	N
Central	N- <i>e</i>	ADJ	Mazandarani	ADJ- <i>e</i>	N	Parachi	ADJ	N
SW	N- <i>e</i>	ADJ	Balochi	ADJ- <i>ēn</i> ³	N	Ormuri	ADJ	N
						Yaghnobi	ADJ	N
						Ossetic	ADJ	N

1 = inflected for gender, number, case; 2 = inflected in N Kurdish; 3 = derivational suffix.

Exceptions in West Iranian include Sangesari of the “Alborz” tier. It has ADJ N, which is one of the numerous features reflecting the Eastern origin of this language. However, this construction has assumed indefinite function and is paired with a definite EZ construction which aligns with the “Alborz” isogloss: ADJ-*ə* N. In East Iranian, the exception is *-i* in Yazghulami, close to Tajik-speaking areas.

4.1.1 Diachrony

The two exceptions to the areal typology just cited, as well as the membership of Tat Persian in the “Alborz” tier, exemplify the well-known fact that word order typologies require no great time depth, and are initially partial. Thus, while in Tat Persian the adjectival EZ has aligned with the “Alborz” isogloss, the nominal EZ has remained unassimilated: N₁-*i* N₂. Nevertheless, the contiguous three tiers with their synchronic triple typology of noun phrase branching imply a long-standing process in West Iranian involving the adaptation of the relative pronoun to a linker in both the “Alborz and “Zagros” tiers, and order inversion in the latter.

4.1.2 Urartian and Elamite substrates

The two documented non-Iranian languages which the earliest immigrating Persian and Median tribes encountered in the west sometime before the 8th cent. B.C.E. are Urartian at the northern end of the “Zagros” tier, and Elamite at its southern end. Both Urartian and Elamite were right-branching languages. Therefore the Elamite typology has been recognized as the likely source of the innovative Old Persian relative construction N *hay*-X ‘N who (is) X’, which in turn is the source for the *ezāfe* in the Perside languages. Similarly, the likely source at the northern end of the “Zagros” tier was the equally right-branching Urartian typology. The process was clearly quite gradual. In both Early Parthian and Middle Persian the descriptive adjective still followed the noun without linker, and only in Middle Persian did the linker *-ī* become generalized. Moreover, in Middle Persian even the left-branching ADJ N is still found side by side the right-branching constructions (cf. chpt. 4 Middle Iranian, sect. 4.2.1):

[*weh*] *dēn* *šāh* [*ī wazurg*]
 ‘[good] religion’ ‘king [EZ great]’ = ‘a/the great king’

4.1.3 Iranian as a “buffer zone”

The mixed right-branching vs. left-branching adjectival NP typology highlighted above is only part of an extensive set of implicational universals as expressed in Iranian. It is number six of eight isoglosses investigated in a succinct pioneering article by Stilo (2005), who also includes the relative position of demonstrative adjectives, numerals, adverb + adjective, object + verb, relative clause + noun, and object + adposition. Not only is the detailed distribution of this extensive set of isoglosses mapped within the Iranian-speaking areas, but the Iranian isoglosses are also embedded in the wider context of the strictly right-branching typology of the languages to the west, represented by Semitic, and the strictly left-branching typology of the languages to the east of Iranian, represented by Turkic. Stilo could thus show how, overall, the multi-faceted internal dynamics of the Iranian languages reflects the mixed typologies distinctive for a linguistic “buffer zone”.

4.2 Bundling West Iranian isoglosses

Word order typology is thus clearly contiguous, which implies a longstanding process. At the same time, there are a number of other well known isoglosses that are discontinuous, and thus suggest displacements. For Western Iranian most salient is the imperfective and irreal marker *-EN* (< *ahēndē*, the Parthian optative of ‘to be’), which is found in Zazaki, Gorani, East Gilaki (but not West Gilaki and Mazandarani), and in Balochi. A lexically salient inventory item with discontinuous distribution is **gaub-* in the sense of ‘to say, speak’ in Persian, in both Caspian languages, and in past forms of ‘to say’ in Kurdish, as opposed to **wāč-* elsewhere.

These features suggest earlier contiguity of the diverse language groups involved, which can be deduced from the bundling of the phonological, morphological and syntactic features discussed so far (*ND* = **-ant*; *EN* = imperfective/irreal marker; **gaub* = ‘to say, speak’; capital letters plus underline = South-West features).

TABLE 2.8: BUNDLED WEST IRANIAN ISOGLOSSES

Avestan	<i>z</i>	<i>sp</i>	<i>hr</i>	<i>s</i>	<i>θw</i>	<i>hw</i>	<i>dw</i>	<i>rz</i>	<i>rd</i>
(Middle) Parthian	<i>z</i>	<i>sp</i>	<i>hr</i>	<i>s</i>	<i>f</i>	<i>w^h</i>	<i>b</i>	<i>rz</i>	<i>rd</i>
N-EZ₁ ADJ									
Zazaki <i>EN/ND</i>	<i>z</i>	<i>sp</i>	<i>(h)r</i>	<i>s</i>	<i>w</i>	<i>w</i>	<i>b</i>	<i>rz</i>	<i>r</i>
Gorani <i>EN</i>	<i>z</i>	<i>sp</i>	<i>hr</i>	<i>s</i>	<i>u</i>	<i>w</i>	<i>b</i>	<i>rz</i>	<i>r/L</i>
Central	<i>z</i>	<i>sb</i>	<i>hr</i>	<i>s</i>	<i>h</i>	<i>flxulh</i>	<i>b</i>	<u><i>L</i>(rz)</u>	<u><i>L</i></u>
ADJ-EZ₂ N									
Tatic	<i>z</i>	<i>sb</i>	<i>hr</i>	<i>s</i>	<i>h</i>	<i>x(u)</i>	<i>b</i>	<i>rz</i>	<i>r</i>
Caspian <i>ND/gaub</i>	<i>z</i>	<i>sb</i>	<i>r</i>	<i>s</i>	<i>h</i>	<i>x</i>	<i>b</i>	<i>rz</i>	<i>rd</i>
E Gil. <i>EN</i>									
Balochi <i>EN</i>	<i>z</i>	<u><i>S</i></u>	<u><i>S</i></u>	<i>s</i>	<i>h</i>	<i>wlh</i>	<i>b</i>	<i>rz</i>	<i>rd</i>
N-EZ₁ ADJ									
Kurdish <i>wāč/gaub</i>	<i>z</i>	<u><i>S</i></u>	<u><i>S</i></u>	<i>s</i>	<i>h</i>	<i>xw</i>	<u><i>D</i></u>	<i>l</i>	<i>l</i>
Persian <i>gaub</i>	<u><i>D</i></u>	<u><i>S</i></u>	<u><i>S</i></u>	<u><i>H</i></u>	<i>h</i>	<i>x</i>	<u><i>D</i></u>	<i>l</i>	<i>l</i>
Middle Persian	<i>d</i>	<i>s</i>	<i>s</i>	<i>h</i>	<i>h</i>	<i>xw</i>	<i>d</i>	<i>l</i>	<i>l</i>
Old Persian	<i>d</i>	<i>s</i>	<i>s'</i>	<i>θ</i>	< <i>θuv</i> >	< <i>uv</i> >	< <i>duv</i> >	<i>rd</i>	<i>rd</i>

In terms of word order, the grouping is in sets of three:

- (1) N-EZ₁ADJ: (a) Zazaki; (b) Gorani; (c) Central Plateau. These three are closest to Parthian; (2) ADJ-EZ₂ N: (a) Tatic; (b) E Gilaki (Caspian); (c) Balochi; (3) N-EZ₁ADJ: (a) NW Kurdish; (b) SW Persian.

In terms of the features EN, ND, and their phonological features, the eight groups are paired:

- (1) Tatic and Central; (2) E Gilaki (Caspian) and Zazaki; (3) Balochi and Gorani; (4) Kurdish and Persian.

Assuming that shared features imply earlier adjacency, the following relative positions can tentatively be reconstructed:

TABLE 2.9: EARLIER LOCATIONS OF WEST IRANIAN GROUPS

ADJ-EZ ₂ N:	Tatic	E Gil. <i>ENIND</i> Gil., Maz. <i>gaub</i>	Balochi	<i>EN</i>	N-EZ ₁	Kurdish <i>wač~gaub</i>
N-EZ ₁ ADJ:	Central	Zazaki <i>ENIND</i>	Gorani	<i>EN</i>	N-EZ ₁	Persian <i>gaub</i>

This rather mechanical re-arrangement of the eight groups reveals a reasonably ordered continuum of isoglosses on both sides of the Alborz range. It may well reconstruct the linguistic situation during Parthian times, and supports suggestions made by various scholars.

Zazaki

Its earlier location in the mountainous region of Gilan known as Deylam has long been suggested, reinforced by the endonym Dim(i)li. The Deylamites (Gk. *Dolomītai*) are first mentioned in Classical sources in the late second century BCE (cf. Felix 1995).

Balochi

For historical and linguistic reasons, the origin of this group has long been assumed to have been somewhere “just east or southeast of the central Caspian region, the meeting point of Middle Persian and Parthian” (cf. Elfenbein 1989; a more general North-West Iranian area is suggested in Chapter 11, section 1.1).

Gorani

According to MacKenzie (2003) the origin of Gorani was in the Caspian provinces, from where the group that is now in the southern Zagros around Awroman moved first at an unknown early date, while the Bajelan group near Mosul followed later. Later, many Gorani-speaking areas were overrun by Kurdish speakers, leading to a merging of the two languages, as is evident from the differences between the “archaic” Northern Kurdish and the Gorani-influenced Central Kurdish dialects.

Kurdish and Persian

MacKenzie (1961) and Windfuhr (1975) suggested locating Kurdish closest to Persian somewhere in the northeast as well. The comparative phonology in Table 2.11 indicates that Kurdish must have been in contact longer than Balochi.

It remains difficult to assign a more circumscribed northeastern location of earlier Kurdish, and earlier Persian varieties, in contact with the six groups. Kurdish may have been located south of the other Northwest Iranian groups, and north of the Fars languages in the south.

gaub- ‘to say, speak’

Even in the spatial reconstruction above, the distribution of this root is discontinuous, which suggests an earlier configuration with pre-Balochi in contact with pre-Kurdish and pre-Persian. – The non-Indo-European Iranian root **gaub-* is well attested in Central Asian Iranian, but not in the general sense of ‘to say, speak’: Sarikoli and Wakhi ‘to bark’; Khotan. ‘to dispraise, abuse’, Bactrian ‘to invoke (legally)’; Sogd., Khwar. ‘to praise’ (cf. Cheung 2007: 113–114). Accordingly, the shift to neutral ‘to say’ could indeed have evolved in a pocket in the area between Sogdia and the southeastern Caspian coast.

Any discussion of people movements must recognize the fact that the movement of small groups and elites are sufficient to superimpose their language, even over large regions, while those remaining become subject to language shifts.

4.3 The Iranian ergative construction

One of the typologically distinct features of the Iranian languages is the so-called tense-split ergative. It is morphologically restricted to forms derived from the past stems of transitive verbs. This construction is first documented occasionally in Young Avestan and extensively in Old Persian (cf. Chapter 3, sections 7.4.2.2 and 7.4.3.2). In its “classical” form, the subject/agent (A) is expressed by an oblique case, or by a personal enclitic, while the direct object/patient (P) is expressed by the direct case and person agreement in the verb.

Accordingly, verb forms in the past domain that are based on the present stem were never affected by this construction: (1) One is the past imperfect form which to this day is derived from the present stem, as in Gorani, several Tati-Azari dialects and Yaghnobi. (2) The other is the subjunctive of the perfective-resultative which in all languages has as its auxiliary the present subjunctive form of the verb ‘to be’.

The emergence and decay of the Iranian ergative is well documented. Its emergence is due to the decay and ultimate loss of the OIr. aorist (see section 3.3). Morphologically, it was replaced by one of the innovative periphrastic constructions in Old Iranian, specifically one based on the completive participle in **-tá* followed by the inflected copula, to express perfective aspect with both intransitive and transitive verbs. The subject/agent was generally marked by the genitive-dative or the personal enclitics (here N = nominative neuter):

Old Persian

<i>(adam)</i>	<i>āgma-ta</i>	<i>ah-mi</i>	‘(I)	come-NOM.3sm	am’
<i>N mana</i>	<i>kār-ta-m</i>	<i>(as-ti)</i>	‘N my.GEN	done-NOM.3sn	(is)’
<i>N-mai</i>	<i>kār-ta-m</i>	<i>(as-ti)</i>	‘N-my	done-NOM.3sn	(is)’

The use of the genitive *mana* ‘my’ (also ‘me’ with gen.-dat. function, see Chapter 3, section 11.3) as well as the use of the enclitic *-mai* with gen.-dat. function are earliest evidence that the genitive(-dative) is the original agent case, and source of the generalized innovative oblique case in Iranian.

The decay of this construction to morphologically parallel the present tense direct-oblique (nominative-accusative construction) involves various processes: (1) loss of the patient ending; (2) loss of the mobility of the enclitics, with increasing obligatory movement towards the components of the predicate: attachment to the direct object, to the preverbal nominal or preverb, to modal-aspectual-negative prefixes, and finally to post-verbal position; (3) loss of agent marking; (4) loss of enclitic marking. Intermediate stages also show differential object marking, which may result in double-oblique systems. At each stage, there are certain restrictions of occurrence and co-occurrence. Also, the co-occurrence of the enclitics with agentive function together with enclitics in possessive function may result in ambiguity.

It is quite certain that all Iranian language developed the ergative construction: Thus, the Middle Iranian languages still had retained it into their early stages. In Sogdian, the ergative construction was replaced by a periphrastic construction with the auxiliary *dār-* ‘to hold, keep’.(cf. Chapter 5, section 3.2.3.3b). Khwarezmian must similarly have gone through an ergative stage, although the documented texts only have *dār-* perfects which is generalized to include the intransitive. Among the modern languages, Ossetic contrasts transitive and intransitive endings: *kod-t-on* ‘I did’ vs. *cyd-t-æn* ‘I went’.

The following table extends the table of West Iranian languages with ergative in Scheucher (2006: 188) to include both the Old and Middle Iranian languages and the East Iranian languages, as well as the languages that have lost such marking. The columns show the three major ways of agent marking: nominal or pronominal subject marked by case ending (OBL), mobile personal enclitics (AS), and fixed enclitics in post-verbal position (VB-ENCL). The agential oblique case is still quite prominent, being lost towards the south; the majority has enclitics, some have both. Note that in three separate areas, Zazaki and Northern Kurdish, northernmost Azari, and in Caspian Gilaki and Mazandarani, the personal enclitics are absent (KD = Kurdish; CPD = Central Plateau languages).

TABLE 2.10: ERGATIVE SUBJECT MARKING IN IRANIAN LANGUAGES

Agent	OBL	ENCL	VB-ENCL (only)
Avestan, Old Persian	+	+	
Early Middle Persian	+	+	
Early Parthian	+	+	
Bactrian	+	+	
Early Sogdian	+	+	
Khwarezmian			later: Past stem + <i>dār-</i> ‘hold’, tr. & intr. Past stem + <i>dār-</i> ‘hold’, tr. & intr.
Zazaki	+	absent	
Kurmanji N KD	+	absent	
Group C KD ¹	+	+	
Mukri C KD ²		+	
Sul./Warm C KD		+	
Gorani	+	+	
Harzani, N Azari	+	absent	

Agent	OBL	ENCL	VB-ENCL (only)
Keringani N Azari	+	absent	
Talyshi	+	+	
Azari	+	+	
Sangesari	+	absent	
Semnani	+		+ transitive past endings
Semnan area			+ transitive past endings
Tafresh-CPD	+	+	
Central Plateau		+	
Sivandi		+	
Fars Languages		+	
Larestani		+	
Bashkardi		+	
Balochi		+	
Pashto	+	+	
Pamir Languages	+	+	“moving” enclitics
Parachi, Ormuḡi	+	+	
Yaghnobi	+	+	
Ossetic			+ transitive past endings
Mazandarani	no ergative		
Gilaki	no ergative		
Persian, Perside	no ergative		

(1) Bingerd, Pishdar, Arbil, Rewandiz, Khoshnaw; (2) Sulaimaniya, Warmawa

4.4 Differential object marking

Given that in those past constructions the direct object/patient was in the nominative and later in the direct case, many languages disambiguated the role of the direct object by the strategy of grammaticalizing selected adpositions. Most of those already functioned to mark other cases so that their use for the direct object was in fact an extension of their functions to differential, or differentiated, object marking (DOM) discussed in detail for the modern languages by Bossong (1985). This holds for adpositional case marking in languages which have lost case inflection.

Differential object marking in Iranian is a response to the loss of inflectional case marking, and grammaticalizes specificity: specific (incl. definite) vs. non-specific (definite or non-specific) direct objects. Languages with a distinct oblique case use this marker, and/or the new markers. Differential marking depends on three interlocking variables: semantic, including the person scale and animacy scale for specificity; syntactic, including tense marking, nominative-accusative and past ergative; and pragmatic discourse prominence (cf. Bashir 2008: 49 on transitional areal and diachronic shifts in Eastern Balochi). Even neighboring and closely related dialects may differ widely in the progression of their ergative decay and their direct object marking.

One of the markers is the postposition *rā*. The three stages of Persian may serve as an example for its development. This marker originates in the Old Persian postposition *rādi* ‘on account of, for the sake of, concerning (cf. Latin *ratione*). In Middle Persian, *rāy* assumed a wide range of dative functions: (1) possessor with the copula, *ō-rā pīsar-ē būd* ‘to him was a son’ (in the absence of a verb for “have”); (2) beneficiary ‘for’; (3) regular marker of the indirect object, in which function it alternated with the preposition *ō* ‘to’;

and, as mentioned, (4) occasional marker of the direct object. However, after the loss of the oblique case, the direct object remained regularly unmarked, and only in late texts was it occasionally marked by the postposition *rāy*, clearly under influence of Early New Persian. All of these functions were continued in ENP. Today, *rā* has become the obligatory marker of the specific direct object, both definite and indefinite. This topicalization of the direct object further extended to include that of a pragmatic anaphoric marker, and of a topical marker of temporal and spatial extension: *šab-rā injā bāš* ‘stay here (for) tonight!’ The indirect object function of *rā*, while retained in Tajik Persian, is lost in Iranian Persian. In both, its beneficiary function is preserved in the prepositional *barāyelbaro-i* < MP *pad rāy ī* ‘for the sake of’.

The following table shows the marking of the main cases in select Iranian languages (cf. Windfuhr 1990: 36) (PR = verbal form of present stem; PT = verbal form from past stem; A = accusative where direct object is marked by adpositions and/or oblique; N = unmarked; O = oblique; P = oblique only in pronouns; (1) S Tati Eshtehardi has DO/DD; (2) Some Balochi dialects have retained the ergative; (3) only marked in plural.

TABLE 2.11: SELECT SYNOPSIS OF DIFFERENTIAL CASE MARKING

	OBL	ERG	GEN	DAT	ACC	PR/PT
Cl. Persian	—	—	-i EZ	-rā	-rā	DA/DA
ModPersian	—	—	-e EZ	be N	-rā	DA/DA
Lori	—	—	(-i-) EZ	va N	-(n)a	DA/DA
Gilaki	-ə	—	ə	-ra	-ra	DO/DO
Mazand.	-e	—	-e	-re	-re	DO/DO
Sangesari	-i	-i	-i	-i/O-rē	-də	DA/OA
Zazaki	-i	-i	-i	-O-rē	-i	DO/OD
Gorani	-i	-i	-i	ba O	-i	DO/OD
N Kurdish	-i	-i	-i	-i	-i	DO/OD
Talyshi	-i	-i	-i	-i	-i	DO/OO
S Tati	-e	-e	-e	-e	-e	DO/OO ¹
Semnani	-i	-i	-i	-i	-i	DO/OO
Balochi	-ā	-ā	-āi	-ā(-rā)	-ā(-rā)	DO/DO ²
Parachi	—	—	-an	kun-N-i	ma-	DA/DA
Ormuri	—	—	-tar	ku-N-ki	ku-	DA/DA
Pashto	-l-ə	-l-ə	da O	Ota	—	DD~P/OD
Yazghulami	P	-i	-i	P-ra	na P	DA/OA
Rushani	P	—	—	-tar	as P	DA/OA
Sanglichi	P	—	—	-be	va-	DA/PA
Ishkashmi	P	—	-no-i	P bo	(-i)	DO/DO
Munji	-an	-an	-an	na-	va-O	DA/OA
Wakhi	-əy	-əy	-əy	-ərk	-əy	DA/OA
Sarikoli	(-ef) ³	—	-an	-ri	a-(ef)	DA/DA
Yaghnobi	-i	-i	-i	ba O	-i	DO/OO
Ossetic	-ī	—	-ī	O-ān	-ī	DO/DO

Clitics

Middle Iranian Khwarezmian and Sogdian provide clear evidence for the universal person scale. Thus, as established by MacKenzie (1993: 141), it determines the syntactic

sequence of the personal enclitic chains in Khwarezmian, irrespective of their case function: 1st before 2nd before 3rd person, and, on a higher interlocutionary level, the speaker is first and the addressee in the dative is last; in addition, singular appears to come before plural (see Chapter 6, section 4.2.2.5). Similar 1st and 2nd person hierarchy is found in other languages, such as Eastern Balochi (cf. Bashir 2008: 54–55).

For the syntax of complex predicates, see 4.4.6 ‘Light verb constructions’ in Chapter 8, and similar sections in other chapters.

4.5 Clause complementation

Subordination is either paratactic, or marked by a small set of semantically distinct conjunctions. Many of the lexical and morpho-syntactical features and combinatory options had already developed in Middle Iranian. Typologically significant was the step-wise emergence of a universal complementizer (cf. also Öhl and Korn 2006).

Middle Persian had the interrogative-relative pronouns *kē* ‘who’ and *čē* ‘what’, and five simple adverbial conjunctions: *ka* ‘when’, *tā* ‘till’, *kū* ‘where’ and ‘that’, *čiyōn* ‘as, in a way that’, *agar* ‘if’, *čē* ‘because, for’. These could be combined: *tā ka* ‘until ~ before (that . . .)’, *kū tā* ‘so that’, *čiyōn ka* ‘as if’, and could also form conjunctive phrases: *tā pēš ka* ‘before, until’, *az ān čiyōn* ‘from that way that’. They could also be introduced by the relative particle *ī*: *ī ka* ‘(which) when’, *tā ān ī ka* ‘(till that which) when’. Two of these had become generalized complementizers: temporal *ka* included condition ‘if’ and cause: ‘since/because’, while locative *kū* also functioned as the complement marker ‘that’. In New Persian, *kē* and *kū* merged with *ka*, which became the general complementizer, modern Pers./Taj. *kelki*, besides *agar* ‘if’, *tā* ‘till’, and *čōn* ‘as; because; when’; and *če ~ čerā* ‘(that is) because’.

The evolution of the complementizer *kelki* shows that common Persian participates in another vast isogloss, one that stretches westward from Iran up to the Balkans (see Matras 2002, in reference to the generalized complementizer *kū* in Kurmanji Kurdish). Finally, a distinctive ‘eastern’ feature, including Tajik Persian, is the prominence of participial and infinitive nominalization.

BIBLIOGRAPHY AND REFERENCES

A. Selected topical references

Phonology: Stilo 1994 phonological system, Iran & Transcaucasus. – *Gender and animacy*: Haig 2004; Ivanov 2003. – *Deixis*: Sims-Williams 1994. – *Nominal system*: Korn 2005a, 2006. – *Pronominal clitics*: Korn 2008. – *Adpositions*: Heston 1987, ambipositions Pashto; Stilo 1987, ambiposition as areal response; 2006, circumpositions. – *Noun phrase*: Boeder 1998, attribution and secondary predicates, includes Kurdish; Stilo, 2004, coordination. – *Ezāfe*: Haider and Zwanzinger, 1984, diachronic; Paul 2002, nominal compounds. – *Verb system*: Bubenik 1977; Korn 2000b, 2001; Windfuhr 1971, 1986, 1992b, 2006a. – *Preverbs*: Moshkalo 1988. – *Potentialis construction*: Sims-Williams 2007, Zaršenās 2002. – *Evidential modus*: Haarmann 1970. – *Progressive constructions*: Jeremiás 1993. – *Ergative and voice*: Bubenik 1989; Haig 1998, 2006, 2008, (forthc.); Korn 2008a; (forthc.); Lazard 1984; Matras 1997; Payne 1979; 1980, 1988. – *Case marking*: Bashir 2008; Bossong 1985; Bubenik 1989; Jahani 2003. – *Word order*: Haider 1985. – *Complex predicates*: Haig 2002; Korn 2000b. – *Clause linking*: Matras 2002; Öhl and Korn 2006; Payne 1982. – *Code copying*: Johanson 2005.

B. Alphabetical

- Azami, Č. A. and Windfuhr, G. (1972) *A Dictionary of Sangesari, with a Grammatical Outline*, Tehrān: Serkat-e Sahāmi-ye Ketābhā-ye Jibi, Franklin Book Programs.
- Bashir, E. (2008) 'Some transitional features of Eastern Balochi: An areal and diachronic perspective', in Jahani, C., Korn, A. and Titus, P. (eds) *The Baloch and Others*, Wiesbaden: Reichert, pp. 45–82.
- Bednarczuk, L. (2003) 'Consonant lenition in Iranian', in van Tongerloo, A. (ed.) *Silk Road Studies VIII. Iranica Selecta. Studies in Honour of Professor Wojciech Skalmowski on the Occasion of his Seventieth Birthday*, Turnhout: Brepols, pp. 1–13.
- Benveniste, É. (1954) 'Expression de "pouvoir" en iranien', *BSL* 50: 56–57 (= *Études sogdiennes*), Wiesbaden 1979, pp. 287–298.
- (1966) 'Le verbe *stā-* comme auxiliaire en iranien', *Acta Orientalia* 30: 45–49.
- Bielmeier, R. (1989) 'Sarmatisch, Alanisch, Jassisch und Altossetisch', in *CLI*, pp. 236–245.
- Boeder, W. (1998) 'Attribution und sekundäre Prädikate' im Sprachvergleich: Deutsch, Englisch, Kurdisch, Georgisch, Türkisch', *STUF* 51.3: 207–227.
- Bossong, G. (1985) *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*, Tübingen: Narr.
- Bubenik, V. (1989) 'An interpretation of split ergativity in Indo-Iranian languages', *Diachronica* 6.2: 181–212.
- (1997) 'Development of Modern Iranian tense-aspect system', in Hewson, J. (ed.) *Tense and Aspect in Indo-European*, Amsterdam Philadelphia: Benjamins, pp. 304–313.
- Cheung, J. (2007) *Etymological Dictionary of the Iranian Verb*, Leiden: Brill.
- Csató, É. Á., Isaaksson, B. and Jahani, C. (eds) (2005) *Linguistic Convergence and Areal Diffusion: Case Studies from Iranian, Semitic and Turkic*, London, New York: RoutledgeCurzon.
- Doerfer, G. (1963) *Türkische und mongolische Elemente im Neupersischen; unter besonderer Berücksichtigung älterer neupersischer Geschichtsquellen, vor allem der Mongolen- und Timuridenzeit*, Wiesbaden: F. Steiner.
- Durkin-Meisterernst, D. (2002) 'The uses of *baw-* in Parthian', *CSI* 25: 47–62.
- Edelman, D. I. (1968) *Osnovnye voprosy lingvisticheskoi geografii. Na materiale indo-iranskikh iazykov* (Fundamental problems of linguistic geography. Based on data of the Indo-Iranian languages), Moskva: Nauka.
- (1980) 'K substratnomu naslediiu tsentral'noaziatskogo iazykovogo soiuzu' (On the substratal heritage of the Central-Asian Sprachbund), *VJa* 5: 21–32.
- (1986) *Sravnitel'naia grammatika vostochnoiranskikh iazykov: fonologii* (Comparative grammar of the East Iranian languages: Phonology), Moskva: Nauka.
- (1988) 'Nekotorye problemy sravnitel'no-istoricheskoi morfologii iranskikh iazykov' (Some problems of the comparative-historical morphology of the Iranian languages), *VJa* 6: 45–62.
- Elfenbein, J. (1989) 'Baluchistan 3: Baluchi language and literature', in *EnIr* 3, pp. 633–644.
- Felix, W. (1995) 'Deylamites i. In the Pre-Islamic period', in *EnIr* 7.4, pp. 342–343.
- Gippert, J. (1996/1998) 'Die historische Entwicklung der Zaza-Sprache', special internet edition of paper, first published in *Ware. Pêseroka Zon u Kulturê ma: Dimili-Kurmanç-Zaza* 10, pp. 148–154 – <http://titus.fkidgl.uni-frankfurt.de/personal/jg/pdf/jg1996f.pdf>

- Haarmann, H. (1970) *Die indirekte Erlebnisform als grammatische Kategorie. Eine Eurasische Isoglosse*, Wiesbaden: Harrassowitz.
- [Hadank] Selcan, Z. (1992) 'Zur Klassifizierung westiranischer Sprachen', *Acta Orientalia* 53: 28–75 (Herausgabe aus dem Nachlaß Karl Kadanks mit einer Vorbemerkung).
- Haider, H. (1985) 'Chance and necessity in diachronic syntax – Word order typologies and the position of Modern Persian relative clauses', in Fisiak, J. (ed.) *Papers from the 6th International Conference on Historical Linguistics*, Amsterdam: Benjamins, pp. 199–216.
- Haider, H. and Zwanziger, R. (1984) 'Relatively attributive: The “ezāfe”-construction from Old Iranian to Modern Persian', in Fisiak, J. (ed.) *Historical Syntax*, Berlin, New York, Amsterdam: Mouton Publishers, pp. 132–172.
- Haig, G. (1998) 'On the interaction of morphological and syntactic ergativity: Lessons from Kurdish', *Lingua* 105: 149–173.
- (2001) 'Linguistic diffusion in modern East Anatolia: from top to bottom', in Aikhenvald, A. and Dixon, R. (eds) *Areal diffusion and genetic inheritance: Problems in comparative linguistics*, Oxford: Oxford University Press, pp. 195–224.
- (2002) 'Complex predicates in Kurdish: Argument sharing, incorporation, or what?', *STUF* 55.1: 25–48.
- (2004) 'Das Genusssystem in der kurdischen Sprache: strukturelle und sozio-linguistische Aspekte', in Hajo, S., Borck, C., Savelsberg, E. and Sukriye, D. (eds) *Gender in Kurdistan und der Diaspora*, Münster: Lit, pp. 33–58.
- (2006) 'The emergence of ergativity in Iranian: reanalysis or extension?', in Karimi, S. et al. (eds) *Aspects of Iranian Linguistics*, pp. 113–127.
- (2008). *Alignment Change in Iranian Languages. A Construction Grammar Approach*, Berlin, New York: Mouton de Gruyter.
- (forthc.) 'Why ergativity in Iranian did not originate from an agentive passive', in Lehmann, C. and Skopeteas, S. (eds) *The Evolution of Syntactic Relations*, Berlin: Mouton.
- Hennerbichler, F. (forthc.) *Die Herkunft der Kurden. Interdisziplinäre Studie*, Berlin: Institut für Kurdische Studien und Institut für Iranistik der Freien Universität Berlin.
- Henning, W. B. (1954) 'The ancient language of Azerbaijan', *TPS* 1954: 157–77.
- Heston, W. (1987) 'Pashto ambipositions and historical antecedents', in Bashir, E., Deshpande, M. M. and Hook, P. Edwin (eds) *Select Papers from Sala-7, South Asian Languages Roundtable Conference, Ann Arbor, Michigan, May 1985*, Bloomington, Indiana: Indiana University Linguistics Club, pp. 163–181.
- Hintze, A. (1998) 'The migrations of the Indo-Iranians and the Iranian sound change $s > h$ ', in Meid, W. (ed.) *Sprache und Kultur der Indogermanen. Akten der X. Fachtagung der Indogermanischen Gesellschaft*, 22.–28. September 1996 (Innsbrucker Beiträge zur Sprachwissenschaft), Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck, pp. 139–153.
- Ivanov, V. (2003 [2004]) 'Traces of animacy in Persian', in Hourcade, B. (ed.) *Iran et Connaissances. Actes du IVe congrès européen des études iraniennes, organisé par la Societas Iranologica Europaea, Paris, 6–10 septembre 1999, III, Cultures et sociétés contemporaines*, Paris: Association pour l'Avancement des Études Iraniennes, pp. 57–63.
- Jahani, C. (2003) 'The case system in Iranian Balochi in a contact linguistic perspective', in Jahani, C. and Korn, A. (eds) *The Baloch and Their Neighbours: Ethnic and Linguistic Contact in Balochistan in Historical and Modern Times*, Wiesbaden: Reichert, pp. 113–131.
- Jahani, C. and Korn, A. (eds) (2003) *The Baloch and Their Neighbours: Ethnic and Linguistic Contact in Balochistan in Historical and Modern Times*, Wiesbaden: Reichert.

- Jahani, C., Korn, A. and Titus, P. (eds) (2008) *The Baloch and Others: Linguistic, Historical and Socio-Political Perspectives on Pluralism in Balochistan*, Wiesbaden: Reichert.
- Jeremiás, E. (1993) 'On the genesis of the periphrastic progressive in Iranian languages', in Skalmowski, W. and van Tanagerloo, A. (eds) *Medioiranica: Proceedings of the International Colloquium organized by the Katholieke Universiteit Leuven from the 21st to the 23rd of May 1990*, Leuven: Peeters and Department Orientalistiek, pp. 99–116.
- Johanson, L. (2005) 'Converging codes in Iranian, Semitic and Turkic', in Csató, É. Á., Isaaksson, B. and Jahani, C. (eds) *Linguistic Convergence and Areal Diffusion*, pp. 3–31.
- (2005) 'Bilateral code copying in Eastern Persian and South-Eastern Turkic', in Csató, É. Á., Isaaksson, B. and Jahani, C. (eds) *Linguistic Convergence and Areal Diffusion*, pp. 205–214
- Johanson, L. and Bulut, C. (eds) (2006) *Turkic-Iranian Contact Areas. Historical and Linguistic Aspects (Turcologica 62)*, Wiesbaden: Harrassowitz.
- Johanson, J. and Utas, B. (eds) (2000) *Evidentials: Turkic, Iranian and Neighbouring Languages (Empirical Approaches to Language Typology 24)*, Berlin, New York: Mouton de Gruyter.
- Kieffer, C. (1974) 'L'établissement des cartes phonétiques: premiers résultats', in *L'atlas linguistique des parlers iraniens: Atlas d'Afghanistan*, Bern: Universität Bern. Institut für Sprachwissenschaft. *Arbeitspapiere* 13, pp. 21–91.
- (1983) 'Afghanistan v. Languages', in *EnIr* 1.5, pp. 501–516.
- (1989) 'L'avancée du persan vers l'Est: le cas des isolats persanophones les plus orientaux en Afghanistan', *StIr* 18.2: 221–236 [dialects of Laghmān, Sorkhrōd, Kandābāgh, Ūrgūn].
- (2003) 'Hazāra iv. Hazāragi dialect', in *EnIr* 12.1, pp. 90–93.
- Korn, A. (2001) 'Archaismus und Innovation im Verbalsystem des Balochi', Vortrag beim Symposium "Iranistik 2000" (Bamberg, Oktober 2000) – <http://titus.uni-frankfurt.de/personal/agnes/bamberg.htm>
- (2003) 'Balochi and the concept of North-Western Iranian', in Jahani, C. and Korn, A. (eds) *The Baloch and Their Neighbours*, pp. 49–60 – <http://titus.uni-frankfurt.de/personal/agnes/upsala00.pdf>
- (2005a) 'Das Nominalsystem des Balochi, mitteliranisch betrachtet', in Schweiger, G. (ed.) *Indogermanica: Festschrift Gert Klingenschmitt. Indische, iranische und indogermanische Studien dem verehrten Jubilar dargebracht zu seinem fünfundsiebzehnten Geburtstag*, Tübingen: VWT-Verlag, pp. 289–302.
- (2005b) *Towards a Historical Grammar of Balochi. Studies in Balochi Historical Phonology and Vocabulary* (Beiträge zur Iranistik 26), Wiesbaden: Reichert.
- (2005c) 'Das Balochi im Kontakt mit den anderen iranischen Sprachen', in Hackstein, O. and Meiser, G. (eds) *Sprachkontakt und Sprachwandel. Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17.–23. September 2000, Halle an der Saale*, Wiesbaden: Reichert, pp. 285–294.
- (2006) 'Historical morphology groupwise: Issues in Western Iranian nominal systems', paper given at the *6th Biennial Conference of Iranian Studies* (August 2006, SOAS London) – <http://titus.uni-frankfurt.de/personal/korn.htm>
- (2008a) 'Marking of arguments in Balochi ergative and mixed constructions', in Karimi, S., Samii, V. and Stilo, D. (eds) *Aspects of Iranian Linguistics*, Newcastle: Cambridge Scholars Publishing, pp. 249–276.
- (2008b) 'A case for the accusative: Western Iranian pronominal clitics', in Brunner, R., Laut, J. P., Reinkowski, M. (eds): *XXX. Deutscher Orientalistentag, Freiburg, 24.–28. September 2007. Ausgewählte Vorträge*, <http://titus.uni-frankfurt.de/personal/agnes/halle.htm>

- (2008c) ‘Voice and transitivity in complex predicates’, paper given at the 1st International Conference on Languages and Dialects in Iran, Tuesday 28–31 October, 2008, University of Sistan and Baluchestan, Zahedan, Iran, 14 pp – <http://www.iranianlinguistics.org/complexpredicates/presentations/korn-2008.pdf>
- (forthc.) ‘The ergative system in Balochi from a typological perspective. in Mahmoodi Bakhtiari, B. (ed.) *Studies on the Typology of the Iranian Languages*, Fremont, CA.: Jain Publishing – <http://titus.uni-frankfurt.de/personal/agnes/ergativ.pdf>
- Krahnke, K. (1976) *Linguistic relationships in Central Iran*, Ph.D. diss., University of Michigan.
- Lazard, G. (1984) ‘Deux questions de linguistique Iranienne: La construction passive du parfait transitif. La versification du moyen-iranien occidental’, in Taillardat, J., Lazard, G. and Serbat, G. (eds) *E. Benveniste aujourd'hui*, Paris: Bibliothèque de l'Information Grammaticale, diffusion, Louvain: Editions Peeters, pp. 239–48.
- (1986) ‘Formes et fonctions du passif et de l'antipassif’, *Actances* 2: 7–57.
- Lecoq, P. (1989a) ‘Les dialectes caspiens et du nord-ouest’, in *CLI*, pp. 313–326.
- (1989b) ‘La classement des langues irano-aryennes occidentales’, in de Fouchécour, C.-H. and Gignoux, P. (eds) *Études Irano-Aryennes offertes à Gilbert Lazard*, Paris: Association pour l'Avancement des Études Iraniennes, pp. 247–64.
- Leezenberg, M. (1993) ‘Gorani influence on Central Kurdish: Substratum or prestige borrowing?’, X-93–03, ILLC Research and Technical Notes Series, The Netherlands – <http://www.cogsci.ed.ac.uk/~siamakr/Kurdish/Papers/Leezenberg93/text.html>
- Lentz, W. (1927) ‘Die nordiranischen Elemente in der neupersischen Literatursprache bei Firdosi’, *ZII* 4: 251–316.
- Lubotsky, A. (2001) ‘The Indo-Iranian substratum’, in Carpelan, C., Parpola, A. and Koskikallio, P. (eds) *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations*, Helsinki: Société Finno-ougrienne, pp. 301–17.
- (2002) ‘Scythian elements in Old Iranian’, in Sims-Williams, N. (ed.) *Indo-Iranian Languages and Peoples (Centennial Bailey)*, (= Proceedings of the British Academy 116), Oxford: For the British Academy by Oxford University Press, pp. 189–202.
- MacKenzie, D. N. (1961) ‘The origins of Kurdish’, *TPS* 1961: 68–86.
- (1961–1962) *Kurdish Dialect Studies*, 2 vols, London and New York: Oxford University Press.
- (1966) *The Dialect of Awroman (Hawramān-i Luhon)*, København: Munksgaard.
- (1967) ‘Notes on the transcription of Pahlavi’, *Transactions of the Philological Society*, 17–29.
- (1993) ‘Clitics in Khwarezmian’, in Skalmowski, W. and van Tongerloo, A. (eds) *Medioiranica: Proceedings of the International Colloquium organized by the Katholieke Universiteit Leven from the 21st to the 23rd May 1990*, Leuven: Peeters, pp. 135–145 [*Iranica Diversa* II, pp. 595–605]
- (2003) ‘Gūrānī’, in *EnIr* 11.4, pp. 401–403.
- Matras, Y. (1997) ‘Clause combining, ergativity, and coreferent deletion in Kurmanji’, *Studies in Language* 21.3: 613–663.
- (2000) ‘Kurmanji complementation: Semantic-typological aspects in an areal perspective’, *STUF* 55: 49–63.
- Minorsky, V. (1957) ‘Mongol Place-Names in Mukri Kurdistan’, *BSOAS* 19: 58–81.
- Morgenstierne, G. (1973) ‘Traces of Indo-European accentuation in Pashto?’, *NTS* 27: 61–65.

- Moshkalo, V. V. (1988) 'The reflections of the Old Iranian preverbs on the Baluchi verbs', *Newsletter of Baluchistan Studies*, 5: 71–74.
- Öhl, P. and Korn, A. (2006) 'Performanzbasierte und parametrische Wandel in der linken Satzperipherie des Persischen. Der Subordinationsmarker *ke* und die Interrogativpartikel *āyā*', *Die Sprache* 46: 137–202.
- Paul, L. (1998) 'The position of Zazaki among West Iranian languages', in Sims-Williams, N. (ed.) *Proceedings of the Third European Conference of Iranian Studies held in Cambridge, 11th to 15th September 1995, Part 1, Old and Middle Iranian Studies*, Wiesbaden: Reichert, pp. 163–177.
- (2002) 'The constituent order of nominal compounds in Zazaki and other Iranian languages', *STUF* 55.1: 64–79.
- (2003) 'The position of Balochi among Western Iranian languages: The verbal system', in Jahani, C. and Korn, A. (eds) *The Baloch and Their Neighbours*, Wiesbaden: Reichert, pp. 61–71.
- Payne J. R. (1979) 'Transitivity and intransitivity in the Iranian languages of the U.S.S.R.', in *Papers from the Conference on Non-Slavic Languages of the U.S.S.R.*, Chicago: Chicago Linguistic Society, pp. 436–447.
- (1980) 'The decay of ergativity in Pamir languages', *Lingua* 51:147–186.
- (1982) 'Relativization in the Iranian languages of the U.S.S.R.', *Folia Slavica* 5.1–3: 344–363.
- (1991) 'Transitivity and intransitivity in the Iranian Languages of the U.S.S.R.', in *Issues in Diachronic Syntax*. Material for 1991 Linguistic Institute, UC Santa Cruz, pp. 436–447.
- (1998) 'Ergative construction', in *EnIr* 9, pp. 555–558.
- Perry, J. R. (1991) *Form and Meaning in Persian Vocabulary: The Arabic Feminine Ending*, Costa Mesa: Mazda Publishers.
- (2005) 'Lexical areas and semantic fields of Arabic loanwords in Persian and beyond', in Csató, É. Á., Isaksson, B. and Jahani, C. (eds), *Linguistic Convergence and Areal Diffusion*, pp. 97–109.
- Redard, G. (1974) 'État des travaux et publications. Quelques cartes onomasiologiques', *L'atlas linguistique des parlers iraniens: Atlas d'Afghanistan*, Bern: Universität Bern. Institut für Sprachwissenschaft. Arbeitspapiere 13, pp. 7–19.
- Scheucher, B. (2006) 'Teilergativität in den modernen westiranischen Sprachen', in Eichner, H., Fragner, B. G., Sadovski, V. and Schmitt, R. (eds) *Iranistik in Europa – Gestern, heute, morgen*, Wien: Österreichische Akademie der Wissenschaften, pp. 169–193.
- Schmitt, R. (1990) 'Caspian', in *EnIr* 5.1, pp. 62–63.
- (2000) *Die iranischen Sprachen in Geschichte und Gegenwart*, Wiesbaden: Reichert.
- Sims-Williams, N. (1989) 'Eastern Middle Iranian', in *CLI*, pp. 166–172.
- (1996) 'Eastern Iranian Languages', in *EnIr* 7.6, pp. 649–652.
- (1994) 'The triple system of deixis', *TPS* 92: 41–53
- (2007) 'The Sogdian potentialis', in Macuch, M., Maggi, M. and Sundermann, W. (eds) *Iranian Languages and Texts from Iran and Turfan. Ronald E. Emmerick Memorial Volume* (Iranica 13), Wiesbaden: Harrassowitz, pp. 377–386.
- Skjærvø, P. O. (1989) 'Modern East Iranian languages', in *CLI*, pp. 370–383.
- (2006) 'Iran vi. Iranian languages and scripts', in *EnIr* 13.4, pp. 344–377 (additional parts forthc.)
- Steblin-Kamenskii, I. M. (1992) 'Central Asia xiii. Iranian languages', in *EnIr* 5, pp. 223–226.
- Stilo, D. (1971) *A Grammar of Vafsi-Tati: An Application of a Transformational Computer Model*, Ph.D. diss., University of Michigan.

- (1981) 'The Tati language group in the sociolinguistic context of Northwestern Iran and Transcaucasia', *Iranian Studies* 14.3-4: 137-187.
- (1987) 'Ambipositions as an areal response: The case-study of the Iranian zone', in Bashir, E., Deshpande, M. M., and Hook, P. E. (eds) *Select Papers from Sala-7, South Asian Languages Roundtable Conference, Ann Arbor, Michigan, May 1985*, Bloomington, Indiana: Indiana University Linguistics Club, pp. 308-336.
- (1988) 'Gender loss and gender differentiation in Northwestern Iranian', paper presented at the 1988 Meeting of the Middle East Studies Association of North America.
- (1994) 'Phonological systems in contact in Iran and Transcaucasia', in M. Marashi (ed.) *Persian Studies in North America. Studies in Honor of Mohammad Ali Jazayeri*, Bethesda, Maryland: Iranbooks, pp. 75-94.
- (2001) 'Gīlān x. Languages', in *EnIr* 10.6, pp. 660-668.
- (2003) 'The polygenetic origins of the Northern Talyshi language', paper given at the International Conference on The Sociolinguistic Situation of Present-Day Caucasia, Frankfurt am Main, 12-14 December 2003.
- (2004) 'Coordination in three Western Iranian languages: Vafsi, Persian and Gilaki', in Haspelmath, M. (ed.) *Coordinating Constructions*, Amsterdam: Benjamins, pp. 269-330.
- (2005) 'Iranian as buffer zone between the universal typologies of Turkic and Semitic', in Csató, É. Á., Isaaksson, B. and Jahani, C. (eds) *Linguistic Convergence and Areal Diffusion*, pp. 35-63.
- (2006) 'Circumpositions as an areal response: The case study of the Iranian zone', in Johanson and Bulut (eds), pp. 310-31.
- (2007) 'Isfahan xx. - Provincial dialects', in *EnIr* 14.2, pp. 93-112.
- Tedesco, P. (1921) 'Dialektologie der mittelliranischen Turfantexte', in *Le Monde Oriental* 15: 184-258.
- Thordarson, F. (1989) 'Ossetic', in *CLI*, pp. 456-479.
- Tremblay, X. (2006) 'Ostiran vs. Westiran: Ein oder zwei Iran vor der islamischen Eroberung?', in Eichner, H., Fragner, B. G., Sadovski, V. and Schmitt, R. (eds) *Iranistik in Europa - Gestern, Heute, Morgen*, Wien: Österreichische Akademie der Wissenschaften, pp. 217-239.
- Wendtland, A. (2006) 'On the ergativity in the Pamir languages', in Karimi, S. et al. (eds) *Aspects*, pp. 419-433.
- Windfuhr, G. (1971) Review of Yarshater, E. (1969) *A Grammar of Southern Tati Dialects* (Median Dialect Studies, I), *JAOS* 91: 370-374.
- (1972) 'Some Avestan rules and their signs', *JAOS* 92: 52-59.
- (1975) 'Isoglosses: A sketch on Persians and Parthians, Kurds and Medes', in *Hommages et Opera Minora, Monumentum H. S. Nyberg*, vol. 2 (*Acta Iranica* 5). Téhéran-Liège: Bibliothèque Pahlavi, pp. 457-472.
- (1979) *Persian Grammar. History and State of Research* (Janua Linguarum. Trends in Linguistics. State of the Art Series, 12). The Hague: Mouton.
- (1986) 'A spatial model for tense, aspect and mood', *Folia Linguistica* 19.3: 415-461 (verb systems of Persian, French, Russian, Arabic, Turkish, Creole).
- (1987) 'Convergence: Iranian Talyshi and Turkic Azari', in Bashir, E., Deshpande, M. M. and Hook, P. H. (eds) *Select Papers from Sala-7, South Asian Languages Roundtable Conference, Ann Arbor, Michigan, May 1985*, edited by Bloomington, Indiana: Indiana University Linguistics Club, pp. 385-405, 7 diagrams.
- (1989a) 'New Iranian languages', in *CLI*, pp. 246-250.
- (1989b) 'New West Iranian,' in *CLI*, 251-262.
- (1989c) 'West Iranian dialects', in *CLI*, pp. 294-295.

- (1989d) ‘Lārestān dialects, in *EnIr* (electronic version; hard copy of fascicle not yet published)
- (1990) ‘Typological notes on pronominal cases in Iranian Tati’, in *Aspects of Iranian Culture in Honor of Richard Nelson Frye. Bulletin of the Asia Institute, N.S.*, vol. 4, pp. 313–324.
- (1991) ‘Cases, forms and uses, in Iranian languages and dialects, in *EnIr* 5.1, pp. 25–37.
- (1992a) ‘Central dialects’, in *EnIr* 5.3, pp. 242–252.
- (1992b) ‘A spatial model applied to West Iranian verb systems’, in Wetzler, A. and Hammerschmidt, E. (eds) *Proceedings of the XXXII International Congress for Asian and North African Studies, Hamburg, Germany, August 1986, Section 6*, Wiesbaden: Steiner Verlag, pp. 394–395.
- (1995) ‘Dialectology’, in *EnIr* 7.4, pp. 362–370.
- (1999) ‘Fārs viii. Dialects, in *EnIr* 9.4, pp. 362–73.
- (2006a) ‘Language change and modeling modal axes: Irano-Turkic convergence’, in Johanson, L. and Bulut, C. (eds) *Turkic-Iranian Contact Areas. Historical and Linguistic Aspects*, Wiesbaden: Harrassowitz, pp. 252–83.
- (2006b) ‘Iran. Non-Iranian languages’, *EnIr* 13.4, pp. 377–410.
- Yarshater, E. (1987) ‘Azari, the Old Iranian language of Azerbaijan’, in *EnIr* 3.3, pp. 238–245.
- Zadok, R. (2001) ‘On the location of NA Parsua’ in *Nouvelles assyriologiques brèves et utilitaires (NABU)* 2001–28 at <http://www.achemnet.com>
- (2002) *The Ethno-Linguistic Character of Northwestern Iran and Kurdistan in the Neo-Assyrian Period*, Jaffa, Israel, 2002; also published in *Iran* 40 (2002): 89–151.
- Zaršenās, Z. (2002/1381) ‘Šavāhedi az sāxt-e emkāni dar zabān-e baluči’ (Some examples of the potential construction in Balochi), in Rezā’i Bāgbidi, H. (ed.): *Majmū’e-ye maqālāt-e naxostīn ham-andiše-ye gūyeš-šenāsi-ye Irān*, Tehrān, pp. 259–270.

CHAPTER THREE

OLD IRANIAN

Prods Oktor Skjærvø

1 INTRODUCTION *

1.1 The Old Iranian languages

The Old Iranian languages known from texts are Old and Young Avestan and Old Persian. Comparative study of these three languages permits the reconstruction of proto-Iranian as a branch of the Indo-Iranian languages, an eastern branch of the Indo-European languages that may have been spoken in the area south and south-east of the Aral sea in the third millennium BCE, splitting into Iranian and Indo-Aryan some time before 2000 BCE. Geographical names contained in the *Young Avesta* confirm this location and also show that, by the time of the Achaemenids, the Avestan speakers had moved farther south as far as the Helmand valley in modern south Afghanistan (Skjærvø 1995: 163–66). See also section 1.4.

The Avestan texts contain no historical allusions, so they cannot be dated on such grounds, but Old Avestan (see below) is closely similar in grammar and vocabulary to the oldest Indic language as seen in the oldest parts of the Rigveda and should therefore probably be dated to about the same time. Similarly, Young Avestan must have been quite similar to Old Persian, which suggests it was spoken in the first half of the first millennium BCE.

Iranian tribes calling themselves Parswa and Mada are found in (north)western Iran from the ninth century BCE onward (see Waters 1999), but the extant Old Persian texts, written in a cuneiform script, are from the Achaemenid period (ca. 558–330 BCE; the texts date from between 522 and ca. 350 BCE). Thus, while the language of the Avesta probably belonged to tribes from north-eastern Iran, both history and linguistics indicate that Old Persian was the language spoken in southern (south-western) Iran.

Other Iranian languages, beside Old Persian and Avestan, were spoken in the first millennium before our era, about two of which we have some information:

Median was spoken in Western and Central Iran by the Medes, who ruled Western Iran ca. 700–ca. 558 BCE, and presumably used their language in official proclamations. Numerous non-Persian words in the Old Persian inscriptions are from a dialect that shared important phonological isoglosses with Avestan, rather than Old Persian, and are assumed to be from Median. Sometimes, both the Median and Old Persian forms of words are found.

Scythian (Saka) dialects were spoken by tribes in Central Asia.

* Marking reconstructed older forms with an asterisk (*) has been dispensed with, as they will be easily recognisable as such by their orthography and by '<' (coming from) and non-reconstructed forms will be marked throughout (OAv., etc.). The asterisk has therefore been reserved for restored (and uncertain) forms. Note the convention '-S-' for the alternating *-h/-š/-s-*.

1.1.1 *The Old and Young Avesta, Old and Young Avestan*

The Avestan languages are known from the *Avesta*, the sacred book of the Zoroastrians (Mazdeans, Mazdayasnians), a collection of mostly ritual texts assumed to have been composed in the second and first millennia BCE. The texts were orally transmitted until committed to writing some time in the late Sasanian period (ca. 224–651 CE; see Skjærvø 1995, 2006; Kellens 1998). The extant Avestan texts represent only a part of the *Avesta* known at that time.

Avestan falls into two chronologically distinct layers: Old Avestan (OAv.) and Young Avestan (YAv.). Young Avestan represents a more developed form of the language than Old Avestan, especially in phonology, but also in morphology and syntax. Thus, Young Avestan is typologically closer to Old Persian (OPers.), and the simplest assumption is that these common developments had taken place before the Old Persian-speaking tribes migrated out of Central Asia at about the turn of the millennium (Skjærvø 2003–2004). Old Avestan and Young Avestan are apparently not simply different stages of the same language, as there are phonological and morphological isoglosses separating them (ibid.; de Vaan 2003: 5–10; Panaino 2007).

1.1.2 *The oral background of the Avestan text*

The extant text of the Avesta does not represent a text composed in writing in ancient times. Rather, it is a compilation of mainly ritual and a few didactic texts that had been transmitted orally for centuries and even millennia before being written down some time after ca. 600 CE, though not all, necessarily, at the same time. During this time, the texts had been, presumably, judging from what we know of oral poetry, first been composed and recomposed in memory and performance; then, at some time committed to memory so as not to be recomposed or changed in any way ('crystallised'). This happened, at different times, to both the Old Avestan and the Young Avestan corpora. Nevertheless, over time, the priests by necessity interfered with the text, both intentionally and unintentionally, as it was passed down through the generations and as the Iranians moved into new areas and maybe, even, changed dialects. What the situation was in the late sixth century, therefore, when it was decided to write the tradition down, we can only speculate about. It seems reasonable, however, to assume that the two principal texts, the liturgies of the *yasna* and the *videvdad sade* ceremonies, were well known by the priests among whom the alphabet was devised and so represent 'official' texts. The same may be the case of the principal texts recited at festivals to individual deities (the *yašts*) and some others, but it is not reasonable to think that all the texts were present at the religious centres; some were probably remembered only in some places by some priests, who may then have been called in to perform what they remembered to the newly educated scribes.

The oral background of the Avestan texts suffices to explain the grammatical inconsistencies and (from our point of view) erroneous forms and uses. The ideas, common throughout the twentieth century, that the 'errors' were due to the 'authors' and 'late' composition, was based on the notion that oral literature could be evaluated in this respect like written literature (see Skjærvø 2005–2006, 2006b: 112–15). Note also that the suggestion that Old Avestan might in fact be later than Young Avestan, adducing the comparison with the use of Latin in Europe long after it was no longer a spoken language, is faulty, as Latin was learned from existing manuscripts and grammars.

The most important fact to keep in mind is that the priests who performed the texts no longer spoke the languages and that their understanding of them was that of the secondary traditions as recorded later in the Pahlavi texts.

As a result, on the one hand, the Old Avestan texts contain many elements that are clearly borrowed from or influenced by Young Avestan and, on the other hand, Young Avestan texts contain both elements that are imitations of Old Avestan ('pseudo-Old Avestan') and later features introduced by the scribes (including from local spoken languages). This makes it a challenge to determine which of the sound changes we observe in our extant manuscripts already belonged to the 'original' two languages and which ones were introduced during the oral and written transmission of the texts. It renders even more problematic attempts to identify additional linguistic stages between Old and Young Avestan (see Tremblay 2006).

Nevertheless, although it is not likely that the texts in every detail reflect a genuine spoken language, research has shown that, in spite of all the inconsistencies of the extant text and contrary to the common opinion of pre- Second World War European scholarship, it reflects a real linguistic system (Morgenstierne 1942).

1.1.3 *The Avestan alphabet*

Some time during the Sasanian period a phonetic alphabet was invented, which was used to write down the known Avestan texts to ensure their correct recitation, crucial to the success of the rituals in which they were used. The alphabet was based on the Middle Persian (Pahlavi) script, of which various stages and styles are known. It is usually assumed that the Avestan script was primarily based upon the common Pahlavi script as known from the extant manuscripts, with the addition of forms taken from the Psalter script (see, e.g. Hoffmann 1988). We have no guarantee, however, that the script seen in the extant manuscripts is identical with that of the first manuscripts. It is just as likely that the shape of the Avestan letters developed together with those of the Pahlavi letters, for instance.

What the principles were that governed the phonetic analysis of the spoken text we do not know for certain, but they were probably the same that applied to the learning of the text. Thus, it is possible that the priests simply assigned a sign to each of the sounds they had been taught to utter during recitation, e.g.: 'to speak this word you say the sound X' > 'to write this word you use the sign X'.

The oldest manuscripts of the *Avesta* are from the thirteenth–fourteenth centuries (*Vispered* ms. K7a: 1278?; *Yasna* mss. J2 and K5: 1323; *Videvdad* mss. K1: 1324, L4: 1323; *Xorde Avesta* ms. Jm4: 1352), and, from the evidence of the manuscript colophons and mistakes common to all the manuscripts of a text, all go back to single manuscripts for each part of the *Avesta* that were in existence around 1000 CE.

Most manuscripts of the *Avesta* are much later, however, and, for some parts of the text, the manuscript tradition does not go back beyond the sixteenth–eighteenth centuries. This means that text criticism in the Classical sense can only restore the readings of manuscripts no older than 1000 CE, but mostly later.

A serious desideratum is a study of individual manuscripts in order to determine the scope of scribal variations in paleography and orthography. Until this has been done, no definitive descriptive orthography, hence also phonology and morphology, of Avestan can be written. Current and earlier descriptions are all, to a large extent, based upon the orthography of K. F. Geldner's critical edition, which is the one most often referred to for

grammatical purposes. It is not a completely ‘critical’ edition, however, since the relative importance of the manuscripts was not clear to him during the publication (see his *Prolegomena*). The edition can therefore not (as is now widely recognized) be used directly as the basis for grammatical analysis.

This situation also renders theories about dialect features in Young Avestan doubtful, such as that of Schindler (1982), who interpreted the different treatments of final *-aŋh* (< Iir. *-ans*) in terms of dialects, and Hoffmann, who ascribed features such as *h_u > x^v* instead of *h_{uu}-* (Hoffmann and Forssmann 1996, §6.2) and *VβV > V_uV* to an Arachotian dialect (ibid., §63cg).

1.1.4 Stages of Avestan

We may represent the history of the Avestan texts after the proto-Iranian stage schematically as follows (cf. Hoffmann 1970, Skjærvø 1994, Kellens 1998, de Vaan 2002: 11–15; Tremblay 2006: 276–77):

- Mid-second millennium BCE: Composition of ritual texts constantly recomposed and linguistically updated in performance, the last direct evidence of which are the extant Old Avestan texts.
- End of second / early first millennium: Composition of texts, constantly linguistically updated, etc. which resulted in the Young Avestan texts.
- Late Old Avestan period: Crystallisation of a set of Old Avestan texts as unchangeable, but, probably, with the introduction of editorial changes then and/or later. These were the ‘five *Gāōās* of Zarathustra’, as they are called in the Young Avesta, plus the *Yasna Haptaŋhāti*, the sacrifice in seven sections.
- First half of first millennium(?): Crystallisation of the Young Avestan text (containing the *Old Avesta*) as unchangeable.
- Up to ca. 500 CE: Transmission of the entire immutable text with introduction of linguistic novelties and changes by the oral transmitters.
- 600+: Creation of a phonetically unambiguous alphabet, in which the entire known corpus was written down (to the extent it was deemed worthy?).
- Ca. 600 to thirteenth century: Written transmission of the text (copying of manuscripts), still probably influenced by the oral tradition, and oppression of the religion and its texts by the Arab conquerors, all of which contributed to deterioration of the text. Ca. 1000 CE there was only single manuscripts in existence of each part of the extant Avesta.

Another problem is the ‘canonisation’ of the corpus, that is, the process of selecting which texts were to be part of the sacred corpus. The concept, however, is based on the canonisation process of the Bible and may not have the same relevance for the Avestan texts.

1.1.5 Old Persian

We do not know when or how the Old Persian-speaking tribes came from Central Asia to south-western Iran, where they are found in the historical period from the ninth century onward. The extant Old Persian texts all date from the sixth to the fourth century BCE. They are written in a cuneiform script, probably invented under Darius for the purpose of recording his deeds. The Old Persian language as we know it from the inscriptions

(sixth–fourth centuries) was already about to change to ‘proto-Middle Persian’, the predecessor of Middle Persian as known from the first century BCE on, as we can see from the late inscriptions, in which ‘wrong’ orthography, especially endings, are common (Skjærvø 1999: 158–61). It is therefore probable that Old Persian had already been spoken throughout most of the first half of the first millennium BCE and had been more or less contemporary with Young Avestan.

1.1.6 The Old Persian script

The Old Persian script is a cuneiform script, but differs from all the neighboring script of the time in having a small set of signs.

Opinions vary about who invented and first used this script, but strong arguments have been adduced that Darius invented it for his Behistun (Bisotun) inscription (520 BCE). On the one hand, it has been shown that the Old Persian version of his ancestor Cyrus’s inscription is a later addition to the Akkadian and Elamite versions, and those attributed to his grand- and great-grandfathers Arsames and Ariaramnes are probably modern, less likely antique, forgeries (Schmitt 2007: 25–31). On the other hand, in §70 of the Behistun inscription, Darius appears to say that he was the first to write ‘in Aryan’.

Although the orthography is relatively consistent, there is no particular reason to think that an orthographic standard had been established (e.g. that of the Behistun inscription) that was supposed to be followed both under Darius and after him and against which modern scholars are entitled to judge diverging spellings to be errors. This is all the more true for late inscriptions, which were obviously written by scribes who no longer spoke ‘the King’s Old Persian’ (see Skjærvø 1999b: 158–61).

1.1.7 Old Iranian grammars

As the extant Avestan text cannot be assumed to represent actually spoken languages, any description of the two Avestan languages based upon this text will also not be of actually spoken languages. Let us sum up:

- the ‘crystallized’ text probably represents a language no longer spoken by the current generation;
- the oral transmission took place over a large territory, and we do not know from which local traditions the extant texts have come to us, which may all have left their imprints on the text in the form of dialect features;
- the oral transmission went on for centuries, and we do not know to what extent phonological and grammatical features may reflect the languages of the transmitters, rather than the original languages;
- the oral transmitters, at some stage, grew increasingly unfamiliar with the (whole) ‘correct’ text and would substitute passages they knew in places they did not belong, upsetting the grammar and the context (and metre);
- inferior oral traditions influenced the written tradition and, probably, vice versa. The most serious consequence of this situation is, of course, that no complete phonemic analysis of the two Avestan languages is possible, since it is a concept presupposing the possibility of capturing an actual linguistic state.

It must always be kept in mind that ‘Avestan’ always means ‘the Avestan text as presented by the manuscripts’. One consequence is that ‘Old Avestan’ does not

necessarily imply that a form is thought to have been spoken by the original speakers; it can also be a form modified by Young Avestan speakers. For instance, it is not likely that Old Avestan had OAv. *ha'θiiā̄m* beside *ha'θīm* = YAv. 'true, real'; it is, in fact, more likely that the 'original' form, that spoken by the composers, was **haθjam*.

As for Old Persian, two features need to be kept in mind:

1. the lateness of the language of the inscriptions in the history of Old Persian warns against forcing phonetic and grammatical forms too much into an Old Iranian mould; instead one should consider Old Persian as suspended, as it were, between the Old Iranian and the Middle Iranian language types;
2. the mixed-language type prevents us from grasping the genuine South-West Iranian phonological system.

1.2 The phonology of Indo-Iranian

The traditional reconstruction of late Indo-European phonemes, after the laryngeals H_2 and H_3 had colored e to a and o (etc.), is set out in Table 3.1.1. A number of changes distinguish Indo-Iranian from this reconstruction.

TABLE 3.1.1: LATE INDO-EUROPEAN PHONEMES

Vowels and diphthongs				Syllabic liquids, nasals			
<i>i</i>	<i>e ē</i>	<i>a ā</i>	<i>o ō</i>	<i>u</i>	(ə)	<i>ɾ ʎ</i>	<i>n m</i>
	<i>ei ēi</i>	<i>ai āi</i>	<i>oi ōi</i>				
	<i>eu ēu</i>	<i>au āu</i>	<i>ou ōu</i>				
Consonants							
Labials	<i>p</i>	<i>b</i>	<i>b^h</i>	<i>ɸ</i>	<i>m</i>		
Dentals	<i>ṭ</i>	<i>d</i>	<i>d^h</i>	<i>l r</i>	<i>n</i>	<i>s (z)</i>	
Palatals	<i>k</i>	<i>ǵ</i>	<i>ǵ^h</i>	<i>j</i>			
Velars	<i>k</i>	<i>g</i>	<i>g^h</i>				
Labiovelars	<i>kɸ</i>	<i>g^ɸ</i>	<i>g^{ɸh}</i>				
Laryngeals	<i>H₁</i>	<i>H₂</i>	<i>H₃</i>				

1.2.1 Consonants

1.2.1.1 IIr. Velars

The IE. labio-velars *k^ɸ g^ɸ g^{ɸh}* merged with the velars into one series, *k g g^h*.

1.2.1.2 IIr. Affricates

The IE. palatals *ǵ ǵ^h* became the palatal affricates *č j^h* [tʃ dʒ dʒ^h].

The velars *k g g^h* produced the allophones *k^y g^y g^{yh}* before *i, i, e*, which developed into (post-)alveolar affricates *č j^h* [tʃ dʒ dʒ^h]; when IE. *e, a, o* merged into IIr. *a*, the conditioned variants became phonemes.

1.2.1.3 IIr. š ž

IIr. *š* and *ž* developed from various sources:

- IE. *s* (*z*) became *š* (*ž*) after *i* and *u*, *r* and *ʀ* (<*r* and *l* and their syllabic variants), *k* and *g*^(h), and *p* and *b*^(h) (the ‘*ru(p)ki*’ rule); this rule also worked after an intervening *n*, e.g. acc. plur. *-inš*, *-unš*, *-rnš* (Av. *-īš*, *-ūš*, *-əʀ^(h)š*);
- The resulting *š* became voiced *ž* before voiced stops, but also before vowels, notably in prefixes and before enclitic particles (Ir. *duž-*, *niž-*; *yūž-am*).
- The IIr. palatal affricates *č j j^h* [*tš*, *dž*, *dž^h*] became *š* and *ž^(h)* before dentals and, probably, after labials.
- *š* and *ž* developed in the IE. ‘thorn’ groups, ‘*k̂p*, *kp*’, etc. > IIr. *čš*, *jž*, *kš*, *g^hž* > Ir. *š*, *ž*, *xš*, *gž* (all Ind. *kš*).

1.2.1.4 IIr. The laryngeals

The IE. laryngeal *H*₂ aspirated the (voiceless) stops before vowels (*pH*₂ > *p^h*, *tH*₂ > *t^h*, *kH*₂ > *k^h*).

Between vowels, laryngeals left a hiatus (or some kind of glide) and, between consonants, it is thought, a *schwa*-like central vowel *ə*. They were lost after vowel before consonant with lengthening of the vowel (e.g. *eH* > *ē*).

1.2.1.5 IIr. Liquids and nasals

r and *l* (and syllabic *l* and *r*) merged, though *l* was preserved sporadically in dialects, both Iranian and Indic.

The syllabic nasals *ŋ* and *ŋl* merged with *a* and *ŋH* and *ŋlH* before consonants with *ā*.

1.2.2 Vowels

o in open syllable became *ā* (Brugmann’s Law). IIr. forms with *a* instead of *ā* are often caused by a laryngeal following the syllable-closing consonant, e.g. *sauu-aiia-* ‘revitalize’ < *čauHaja-* vs. *srāuu-aiia-* ‘recite, sing’ < *črāuaja-*.

ē ā ō and the corresponding diphthongs *ēi*, etc. merged into *ā* and *āi*, etc.

The IE. qualitative ablaut *e* ~ *o* was lost by this merger and only partly replaced by the quantitative ablaut *a* ~ *ā*. See also de Vaan 2003, §§30.2–3.

1.2.3 Proto-Indo-Iranian phoneme inventory

In overview, the reconstructed phoneme inventory of proto-Indo-Iranian was as set out in Table 3.1.2:

1.2.4 Ablaut

A distinctive feature of IE. inflection was the qualitative and quantitative ablaut, correlated with stress patterns. While in Indo-Iranian the qualitative ablaut was lost after the merger of *ā ē ō*, the quantitative ablaut was retained albeit modified by these various developments, and ensuing analogies. The basic vocalic ablaut grades were thus as in Table 3.1.3, but other forms are found, as well, especially in connection with lost laryngeals (examples below).

TABLE 3.1.2: PROTO-INDO-IRANIAN PHONEMES

Vowels and diphthongs		Syllabic liquids				
<i>i</i>	<i>ə</i>	<i>u</i>	<i>r</i>	<i>r̄</i>		
	<i>a</i>	<i>ā</i>				
	<i>ai</i>	<i>āi</i>				
	<i>au</i>	<i>āu</i>				
Consonants						
Labials	<i>p</i>	<i>b</i>	<i>b^h</i>	<i>u</i>	<i>m</i>	
Dentals	<i>t</i>	<i>d</i>	<i>d^h</i>	<i>r (l)</i>	<i>n</i>	<i>s (z)</i>
Palatals	<i>č [tš]</i>	<i>ǰ [dž]</i>	<i>ǰ^h</i>			
Palato-Alveolars	<i>č̣ [tṣ̌]</i>	<i>ǰ̣ [dẓ̌]</i>	<i>ǰ̣^h</i>	<i>i</i>		<i>š (ž)</i>
Velars	<i>k</i>	<i>g</i>	<i>g^h</i>			
Laryngeal	<i>H</i>					

TABLE 3.1.3: PROTO-INDO-IRANIAN ABLAUT

long:	<i>ā</i>	<i>āi</i>	<i>āu</i>	<i>ār</i>	<i>ān</i>	<i>ām</i>
full:	<i>a</i>	<i>ai</i>	<i>au</i>	<i>ar</i>	<i>an</i>	<i>am</i>
zero:		<i>i / i</i>	<i>i / u</i>	<i>r / r̄</i>	<i>n / a</i>	<i>m / a</i>

1.3 The phonology of Proto-Iranian

Proto-Iranian shows the following typical changes separating it from proto-Indic:

- Ilr. interconsonantal *ə* was lost in Iranian in all positions (with sporadic development, helped by analogy, of anaptyctic vowels in initial consonant groups; see Beekes 1981, Ravnæs 1981, Mayrhofer 1981, Pirart 1988);
- Ilr. *r̄H* became Ir. *ar* in most contexts, but sometimes *r* (Av. *ər*) (acc. to Cantera 2001, in labial context when unstressed);
- The voiced and aspirated voiced consonants merged: *b, b^h > b; ǰ, ǰ^h > ǰ*, etc;
- The new aspirated stops *p^h t^h k^h* and unaspirated stops *p t k* before consonants were spirantised to *f θ x* (e.g. *kt > xt*);
- A sibilant developed in the clusters dental + dental: *t-t, d-d, d-dh > t^st, d^sd*;
- Dentals were lost before *s/z, ś/ž, š/ẓ̌*, including in the sequences $T_1^{s/\acute{s}}T_2 > S/\acute{S}T_2$: *t^st, d^sd > st, zd*, and *čt, ǰd [tšt, džd] > št, žd*. This rule also affected voiced aspirated clusters that resulted from ‘Bartholomae’s Law’ (see section 2.3.2.1 b);
- Geminates were simplified, including those resulting from assimilation (*s-s > s, z-z > z, s-č > s, ǰž > ž, d-n > n, p-b > b*, etc.);
- Most significantly, also for later vocalic modifications and changes, *s > h* except before stops and *n* (see section 2.1.2.2).

1.4 Early Iranian dialects

Proto-Iranian split into at least four distinct proto-Iranian dialect groups, characterised, among other things, by the developments of the palatal affricates *č, ǰ* and the groups *č, ǰ* and the groups *č^u, ǰ^u* (see Schmitt 1989: 27–28).

The two dialect groups unattested by texts are:

- *Old Northwest Iranian*, represented by the later Alanic dialects and modern Ossetic, in which initial $p > f$ and internal $r_i > l$;
- *Old Northeast Iranian*, represented by Middle Iranian Khotanese and modern Wakhi, in which $č^u$ and $ǰ^u$ were assimilated to $ś$ and $ž$.

The two attested groups are:

- *Old Central Iranian*, represented by most of the remaining dialects, including Avestan and Median, in which $č$ and $ǰ$ merged with Ir. s and z , respectively, but $č^u$ and $ǰ^u$ became sp and zb ;
- *Old South-West (Perside) Iranian*, represented in historical times by the dialects of Pārsa/Fārs, including Old Persian, in which $č$ and $ǰ$ merged with Ir. $θ$ and d , but $č^u$ and $ǰ^u$ with s and z . Other Perside developments: Ir. $θ > š$ before i and n (Av. *hdθiia-*, OPers. *hašiya-* ‘real, true’; Av. *araθni-*, OPers. *arašni-* ‘ell’);
- Ir. $θr$ and $čr > a$ sibilant $<ç>$ of uncertain nature that later merged with s (Av. *puθra-*, OPers. *puça-*, MPers. *pus* ‘son’; Av. *sraiiia-* ‘to lean’, OPers. *ničāraya-* ‘give back’).

There are numerous ‘Median’ forms in OPers., e.g. *asan-* ‘stone’ vs. OPers. *aθd'ga* ‘stone’, *aspa-* ‘horse’ vs. OPers. *asa-*. In some instances, the Elamite transcriptions reflect Perside forms not found in the inscriptions (see, e.g. Hinz 1973; cf. section 3.6.3.3). As a rule, the Elamite transcriptions of Old Persian names show the Perside form, while the Akkadian ones show the Median forms, e.g. *ciça'taxma-*, but Elam. *ti-iš-šá-an-tam-ma = *tiçantama*, with $č - ç > t - ç$ (cf. Greek *Tissaphernēs* < OPers. **čiça-farnah-*).

At this stage, the consonant phonemes of the four groups were presumably identical except for the palatal sibilants in proto-North-East Iranian, see Table 3.1.4..

TABLE 3.1.4: EARLY IRANIAN CONSONANT PHONEMES

Consonants						
Labials	<i>p</i>	<i>b</i>	<i>f</i>	<i>u</i>	<i>m</i>	
Dentals	<i>t</i>	<i>d</i>	<i>θ</i>	<i>r (l)</i>	<i>n</i>	<i>s z</i>
Palatals				<i>i</i>		[<i>ś ž</i> NE-Ir.]
Palato-Alveolars	<i>č</i>	<i>ǰ</i>				<i>š ž</i>
Velars	<i>k</i>	<i>g</i>	<i>x</i>			
Pharyng.-Glott.		<i>'</i>	<i>h</i>			

1.5 Writing systems

1.5.1 The Avestan script

The Avestan alphabet contains ca. 55 letters (Table 3.1.5), some being, apparently, only scribal and/or local and chronological variants. Each letter is usually written separately, although ligatures like those of Pahlavi are also found. There are fairly distinct differences in ductus between older and later manuscripts and from scribe to scribe, but no paleographic study has yet been made to investigate chronological and geographical trends. See, for instance, the three different handwritings in manuscript L4 in the British Library, London, the original manuscript from 1323 and the fairly recent additions (L4a, b) currently available online at avesta.ana.usal.es/cataloges.htm.

TABLE 3.1.5: THE AVESTAN ALPHABET

𐬀	𐬁	𐬂	𐬃	𐬄	𐬅	𐬆	𐬇	𐬈	𐬉	𐬊	𐬋	𐬌	𐬍	𐬎	𐬏
a	ā	ā (ā)	ā (ā)	i	ī	u	ū	e	ē	o	ō	ə	ō	𐬎 (ā), *𐬎 (ā)	
𐬐	𐬑	𐬒	𐬓	𐬔	𐬕	𐬖	𐬗	𐬘	𐬙	𐬚	𐬛	𐬜	𐬝	𐬞	𐬟
p	b	β		f		m		𐬞 (hm)							
𐬠	𐬡	𐬢	𐬣	𐬤	𐬥	𐬦	𐬧	𐬨	𐬩	𐬪	𐬫	𐬬	𐬭	𐬮	𐬯
t	d	ð	ð	θ		𐬩		n		𐬮					
𐬰	𐬱	𐬲	𐬳	𐬴	𐬵	𐬶	𐬷	𐬸	𐬹	𐬺	𐬻	𐬼	𐬽	𐬾	𐬿
k	g	g ²	γ	x		x ^v	𐬶	𐬹		𐬻		𐬼	𐬽	𐬾	𐬿
𐬠	𐬡									𐬠	𐬡				
c	j									𐬠	𐬡				
𐬠	𐬡	𐬢	𐬣	𐬤	𐬥	𐬦	𐬧	𐬨	𐬩	𐬪	𐬫	𐬬	𐬭	𐬮	𐬯
y	Y	ii	V	uu	l		h								
𐬰	𐬱	𐬲	𐬳	𐬴	𐬵	𐬶	𐬷	𐬸	𐬹	𐬺	𐬻	𐬼	𐬽	𐬾	𐬿
s	z	š	ž	š	š										

Words are separated by a dot (with or without spaces) and sentences often by more elaborate punctuation, e.g. multicoloured floral designs. Individual words and components of compounds are not distinguished. In Western academia, the dot is kept in compounds, but usually replaced by a space between words. The common practice of asterisking words reinterpreted as compounds (or compounds reinterpreted as individual words) is based on unfamiliarity with the script.

Not infrequently, especially in the *Old Avesta*, punctuation is used, incorrectly from our point of view, to delimit morphemes (e.g. plural endings in *b-*), in which case sandhi forms are applied (e.g. *-at.biš* beside *-odbiš*), and even to split up consonants that the priests pronounced as double, e.g. *aēšəm.mahiiā* for *aēš^omahiiā* 'of Wrath' (the Pahlavi translation has the correct interpretation) and *həm.miiāsa-* for *həm.miiāsa-* (beside *həm^omiiāsa-*) 'harness' <*ham-iasa-*.

For notes on individual letters, see section 2.1.

1.5.2 The Old Persian script

The Old Persian script has 3 vowel signs <a, i, u>, 33 consonant signs <C(V)>, 8 ideograms/logograms, numerals, and a word separator (Table 3.1.6). There are minor differences in ductus throughout the corpus.

The consonant signs are consonantic or syllabic with inherent *-a*, *-i*, *-u*. There are only four <Ci> signs and seven <Cu> signs; the <Ca> signs are used for the missing <Ci> and <Cu> signs. Consonants not followed by vowels are written with <Ca> signs.

Short and long *ī*, *ū* are spelled <i, u>, preceded by <Ci> and <Cu> signs when they

TABLE 3.1.6: THE OLD PERSIAN SYLLABARY

Ca		Ci		Cu		Ca		Ci		Cu	
𐎠	<a>	𐎡	<i>	𐎢	<u>	𐎣	<na>			𐎥	<nu>
𐎧	<ā>					𐎦	<pa>				
𐎨	<ba>					𐎩	<ra>			𐎫	<ru>
𐎬	<ca>					𐎭	<sa>				
𐎮	<ca>					𐎯	<ša>				
𐎰	<ca>					𐎱	<ta>			𐎳	<tu>
𐎲	<da>	𐎴	<di>	𐎶	<du>	𐎷	<ta>				
𐎸	<fu>					𐎹	<θa>				
𐎺	<ga>					𐎻	<va>	𐎽	<vi>		
𐎼	<ha>					𐎿	<xa>				
𐎾	<ja>	𐎿	<ji>			𐏀	<ya>				
𐏁	<ka>			𐏃	<ku>	𐏄	<za>				
𐏅	<la>					𐏆	< : >				
𐏇	<ma>	𐏈	<mi>	𐏉	<mu>						

𐎫	<AM ₁ > = Auramazdā	𐎴	<DH ₁ > = dahayāuš	𐎫	<BG> = бага
𐎴	<AM ₂ > = Auramazdā	𐎴	<DH ₂ > = dahayāuš	𐎫	<BU> = būmiš
𐎫	<AMha> = Auramazdāha	𐎫	<XŠ> = xšāyaθiya		

exist: <Ca-i, Ci-i> and <Ca-u, Cu-u> (exceptionally <Ci> = *Ci*). Frequently, we find <-u-va> for *ū* (<pa-ru-u->, <pa-ru-u-va-> = *parū* ‘much’), <i-ya> beside <i> before consonant in *nīyašādayam* <na-i-ya-ša°> ‘I set down’ beside *nīšādayam* <na-i-ša°>, presumably reflecting the development of *-iya-* > *-ī-*.

Final *ī*, *-ai* and *-ū*, *-au* were written <-i-ya> and <-u-va>, but the final *-v* or *-y* could be omitted when followed by an enclitic word (*hau-maiy*, beside *hauv-maiy* ‘he . . . me’, *tayai-šaiy* ‘who . . . his’, etc.). Beside *-ava-* and *-aya-*, we also find *-auva* (*bava-* and *bauva-* ‘become’) and *-aiya* (*dāraya-* and *dāraiya-* ‘hold’).

After *h*, an *i* was usually not written (transcribed as *hq-*).

The syllabic *r* (transcribed as *ʀr*) was also spelled <ra>, e.g. *vazarka-* ‘big’ > Mod. Pers. *bozorg*, vs. *martiya-* > Mod. Pers. *mard*.

Logograms can take case endings (e.g. < XŠ-ha-ya-a> = *xšāyaθiya-hayā* ‘the king’s’, <DH₁-na-a-ma> = *dahayū-nām* ‘of the lands’, <BU-ya-a> = *būmī-yā* ‘of the earth’). See also section 2.2.

2 PHONOLOGY

2.1 The phonology of Avestan

The Avestan languages share some important features that give them their characteristic look, different from Old Indic and Old Persian: raising, lowering, rounding, and nasalisation of vowels, anaptyxis; spirants, palatalisation and labialisation of consonants, nasalisation of intervocalic *h*.

Thus, Avestan phonology is very complex. The exact inventory of phonemes is unclear, as the number of phonetic realisations and morphophonological variants is very large, though all part of an internally consistent phonological system, in many respects similar to those of later East Iranian languages (Morgenstierne 1942).

In the following, the phonologies of Old and Young Avestan are described as they had evolved through many layers of historical and redactory processes down to the end of the Sasanian period, when at least some of the texts were first written down, and beyond, during the written transmission. They are therefore the artifacts of a learned tradition, and can not be assumed to reflect the actual phonological systems of the languages when they were spoken.

In view of the uncertainties regarding the shaping of the Avestan phonology, I have refrained from providing a reconstruction of their phonemic systems. For a recent attempt, see de Vaan 2003: 615–29. Instead, Tables 3.2.1–2 contain the basic vowels and consonants, some of which are conditioned variants, differing in Old and Young Avestan.

2.1.1 Vowel systems

The letters *ā ā̇, ə ē, o ō, e ē*, and nasalised *q*, reflect the development and partial phonemisation of allophonic variants mainly from the basic low pair *a ā*, the diphthongs *ai, āi* and *au, āu*, and *r*. See Table 3.2.1.

TABLE 3.2.1: AVESTAN VOWELS

	Front	Central	Back, rounded	Nasal
High	<i>i ī</i>		<i>u ū</i>	<i>ī <iiā, ī> ū <uuā></i>
Mid	<i>e ē</i>	<i>ə ē</i>	<i>o ō</i>	<i>ē <ēā></i>
Low		<i>a ā</i>	<i>ā̇</i>	<i>q, q̄ <q></i>
Syllabic <i>r</i>		<i>r <ərə></i>		<i>r̄ <ər̄q></i>
Diphthongs:				
	<i>aē</i>	<i>āi</i>	<i>ōi</i>	
	<i>ao, aō</i>	<i>āu</i>	<i>āu</i>	

2.1.1.1 Notes on the vowels

For details on the Avestan vowels, see now de Vaan 2003.

All Avestan vowels except *e* and *ē* are found in initial position, but *o* and *ō* only in YAv. *oīm*, *ōiium*, etc. <*aiyam* ‘one’ and in a few words where *ōi-* is from *ui-* (e.g. *ōifra-* for *vifra-*?).

All vowels are also found in anaptyxis and epenthesis, particularly *ə*, *i*, and *u*, and may form secondary diphthongs.

2.1.1.2 Vocalic length

Old Avestan probably maintained the length opposition longest in the high row, at least in part: *-iš*, *-uš* vs. *-īš*, *-ūš* (see de Vaan 2003: §§9, 13).

In Young Avestan (and to some extent in OAv.), the distribution of short and long *iī* and *uū* may be conditioned by phonetic context (not etymology); it may also be a matter of scribal preferences and local practice at different times. There is therefore no distinction between proto-Iranian short and long *i* and *u* and Avestan contraction products, *-im*, *-ūm*, and *-iām* all > *-iml-ūm* and *-um*, *-ūm*, and *-uām* all > *-uml-ūm* (ibid., §§8, 12).

The exchange of vowel length, *a/ā*, *i/ī*, *u/ū* is common and may, at least in some instances, be caused by shifts of accent (ibid., §30.1.1).

In final position, length opposition is neutralised in all three Old Iranian languages. Old Avestan final vowels are long; Young Avestan final vowels in monosyllables are long; in polysyllabic words, we have *-a*, *-i*, *-u*, *-e* and *-ə* (epenthetic), but *-ō*, *-ā*, and *-ō̄* (= OAv. *-ō̄ηg*) and, rarely, *-ū* (< *-uuō̄*).

Short *ə* (in OAv. also written *ə̄*) is an allophone of *a* before nasals (*hə̄ηti* ‘they are’) and before *uu* followed by *il̄*. Long *ā* is a phoneme in Young Avestan, since it is a distinctive ending: *-ā* with allomorph *-a*.

It is not known whether the length in *ā̇* was originally phonemic. The short < *ā̇* > is used in only one known manuscript to spell *a* before *η* (see Hoffmann and Narten 1989: 31; de Vaan 2003, § 18). Here *ā̇* will be used for *ā̇*.

2.1.1.3 Diphthongs

ai. The diphthongs *aē* and *ōi* are partly in complementary distribution. YAv. *ōi* is preferred before consonant clusters, except *s* or *š* plus one consonant (Fortson 1996), whereas OAv. *ōi* is apparently also found before *s* and *š* < *ss* and *šš*: *dōišī* ‘you show’ < *daič-ši* (strongly doubted by de Vaan 2003: 352 n. 436). For Ir. *aij*, OAv. has *ōii* (*ōi*.) and *aii*; YAv. *aii* (e.g. OAv. *x^vāθrōi.ā*, YAv. *x^vāθrāiia* ‘in comfort’). Final *ai* in monosyllables became OAv. *ōi*, YAv. *-ē* (but *γōi*); in polysyllables, it became OAv. *ōi* and *-īē* (= *-ē* with preceding palatalisation), YAv. *-īe*. Before enclitics, both OAv. and YAv. have *-aē°*.

au-. OAv. prefers *āu*, YAv. *ao* (many mss. *aō*), but also *āu* in imitation(?) of OAv. (de Vaan 2003, § 16.1). In final position, *-au* became **-aō̄*, YAv. *-uuō̄* (also OAv.) but *-ao°* before enclitic (e.g. YAv. *drao-ca* ‘and in tree’; see Skjærvø, 2005a). The apparent diphthongs *au*, *ou*, and *aou* are the result of labialisation of *a* (see section 2.3.2.4a).

The Ilr. and Av. diphthongs *āi*, *āu* are indistinguishable in the script from Av. *ā* + epenthetic *i*, *u*.

2.1.1.4 Hiatus

In Old Avestan, long vowels and diphthongs resulting from laryngeal loss remain disyllabic: *ā̇*, *ā̇* < *a^hā̇*; *ā̇* < *a^hā̇*, *a^hō̄*; *-qm* < *-a^ham* (e.g. plur. gen., opt. 1s); *aē̇*, *ōi̇* < *a^hi̇*; *āi̇* < *a^hi̇*.

2.1.1.5 Ilr. *r*

The Ilr. syllabic *r* shows up in the later Ir. languages as *r* preceded by a vowel that varies according to the phonetic contexts, e.g. *ir*, *ur* (de Vaan 2003, §24). Avestan usually has *ər* followed by an anaptyctic *ə*, but *rš* after *t* (OAv. *ātršm*), and sporadically elsewhere. Young Avestan has *arš* for OAv. *ər^vš*.

2.1.1.6 Nasalised vowels

All vowels could at some stage be nasalised. Thus, *q* is a realisation of *an* and *ān* before spirants: *zqθβa-*, and sibilants: *vqs* < *vāns-t*, and of long *ā* before *n* or *m*: *nqma*, *nāma*.

The two signs for nasalised *q* may originally have denoted nasalised *ā̇* vs. nasalised *ā̇̄*; thus final *-ā̇* < *-aij^h* may originally have alternated with *-ā̇̄* rather than with *-q*.

In Old Avestan, the spellings *iiq(n)* and *uuq* before *m* may represent nasalised *i̇*: *friiq(n)mahī* ‘we make (you) friendly’[?] < *fri-n-mahi*; and *uq*: *huuqmahi°* ‘we press’ < *hu-n-mahi*. Similarly, the nasalised *ā̇̄* is written *ā̇q* in OAv. *māq* (Y. 28.4). Nasalised *i̇* and *q* are written (or became) *ī* and *ū* before sibilant. In Old and Young Avestan, nasalised *r* is written *ərq* before *š* and *ž*.

2.1.2 Consonant systems

Old and Young Avestan, judging from the orthography, had the consonants set out in Table 3.2.2, which include phonemes and allophones with differing distribution in Old and Young Avestan. Note also that the convention ‘X became Old Avestan/Young Avestan Y’ is subject to the above caveats.

TABLE 3.2.2: AVESTAN CONSONANTS

	Stops/Affr.	Fricatives	Contin.	Nasals	Sibilants
Bi-labials	<i>p b</i>	<i>β</i>	<i>u</i>	<i>(^hm) m</i>	
Labio-dentals		<i>f v</i>			
Dentals	<i>t d</i>	<i>θ δ</i>		<i>n</i>	<i>s z</i>
Alveolar			<i>(^hr) r</i>		
Alveo-pal.	<i>č j</i>				<i>š ž</i>
Retroflex (?)					<i>ṣ̌ ẓ̌</i>
Palatals		<i>y(?)</i>	<i>i</i>	<i>(^hi)</i>	<i>ṣ̌</i>
Velars	<i>k g</i>	<i>x γ</i>		<i>ŋ^h ŋ</i>	
Palato-velars		<i>č</i>		<i>ŋ^h ŋ</i>	
Labio-velars		<i>x^v</i>		<i>ŋ^h ŋ^v</i>	
Glottals	<i>(ʔ)</i>	<i>h</i>			

2.1.2.1 Notes on the consonants

ɮ. The letter *ɮ* probably represented an unreleased stop and was an allophone of /t/ and /d/ in final and pre-consonantal position: YAv. *janaɮ* ‘he killed’, OYAv. *ɮkaēša-* ‘*guidance’, YAv. *ɮbaēšah-* ‘hostility’. In the manuscripts, *ɮ* alternates with *δ*: *aɮka-* and *aδka-* ‘coat’, *druuaɮbiiō* and *druuaδbiiō* < *druuant-* ‘possessed by the Lie’. – The rare final *-g^ɮ* is etymological in OAv. *pa^ttiiaog^ɮ* ‘answering’, *yaog^ɮ* < *yuj-* ‘to harness’, but may be a graphic representation of an unreleased final *-g* in YAv. *-hāg^ɮ* ‘following’; *paraɮg^ɮ* ‘away from’ (Hoffmann and Forssman 1996: 99).

r had a (pre)aspirated/unvoiced(?) allophone before *p*, *k*, *t*, apparently limited to stressed syllables in proto-Avestan (see section 2.3.4). Before *k*, *p*, it is written *hr*, while proto-Av. **hrt* became *ṣ̌* (see section 2.1.2.3). In final position *-r* takes a supporting vowel (*-r^h*).

n was realised as a nasal of uncertain nature written *ni* or *n* before consonants (in editions, usually normalised as *ñ*), except *i* and *u*. Before *i*, possibly also *u*, *n* was palatalised to *ñ* (*ñ*), but the letter <*ñ*> is not consistently used in the manuscripts, e.g. *añiiō* ‘other’, *aniiō*, or *ainiiō*, *nitəmō* ‘nethermost’ or *ñitəmō*; in this description I am adopting the convention *anii-* but *aini-*.

h > *ŋ* (*h*). Intervocalic *h* and its palatalised and labialised allophones are usually preceded by their class nasal: *-āhā-* > *-āŋhā-*, *-āhiā-* > *-āñhā-*, and *-āhuā-* > *-āŋ^vhā-*; similarly, *-hr* > *-aŋ(h)r-* (usually normalised as *-aŋr-*, but mss. often *-aŋhr-*). These clusters are indistinguishable from IIr. *ns* clusters: OAv. *māñhī* < *mansi* ‘I have thought’. The nasalisation is omitted in a not well-defined subset of words with internal *h*, often before *i*: loc. sing. *manahi* ‘in thought’, less often before *u*: neut. *vohu* (see section 2.3.1.2). – A similar nasalisation in Old Persian is suggested by Elam. < pat-ti-ya-man-ya-a > for OPers. *patiyāvahayaiy* (DB 1.55), but other explanations of the Elamite form are possible, and, if the nasalisation is genuine, the form may be ‘Avestan’ (cf. Av. *auuañhe* ‘to (my) aid’; see Skjærvø 1999a: 18–19).

x^v vs. *huu* and $-ŋ^v h-$. The Avestan labialised velar fricative x^v is also an allophone of *h_v*, alternating with *huu* and the labialised velar aspirated nasal $ŋ^v h$.

Initial *h_v*- and *huu*- both became x^v - or *huu*-: Av. $x^v afna-$ ‘sleep’ (<*h_va-*) and Av. $x^v ā0ra-$ ‘good breathing space, comfort’ (<*hu-ā°*), but *huuaršta-* ‘well-done’ (<*hu-_va°*); OAv. *huuar^s* ‘sun’, gen. $x^v ēŋg$, YAv. *hū* (all disyllabic).

Medial $-h(u)_{v-}$ became OAv. $-x^v-$, but YAv. $-ŋ^v h-$: OAv. *nəmax^v d^ttī-* ‘containing homage’, YAv. *baodaŋ^v h^v tī-* ‘conscious’; remained: OAv. dat. sing. *ahuiiē* ‘for the (new) existence’ (<*ahuuai*), YAv. *aŋ^v he*; or became OAv. $-ŋhuu-$, YAv. $-ŋ^v h-$: OAv. *aojōŋhuuąnt-*, YAv. *aojaŋ^v hant-* ‘strong’. – In Young Avestan, the only examples of intervocalic x^v are *kax^v arəda-* ‘sorcerer’ (cf. OInd. *kākhorda*), fem. *kax^v ar^v dī-*, and the country name *harax^v tī-* ‘Arachosia’.

x^i vs. *hii* and *ijh*. The Avestan palatalised velar fricative x^i is an allophone of *h_i*, alternating with *hii* and the palatalised velar aspirated nasal *ijh*: initially and medially, x^i is common in Old Avestan, but rare in Young Avestan: OAv. 3s opt. *xiiāt*, YAv. *hiiāt* ‘may he be’, OAv. gen. sing. fem. *axiiā*, YAv. *ajhā* ‘her’; OAv. *vaxiiā*, YAv. *vajhā* ‘better things’; YAv. *xiiāona-* (ethnic); *daxiiuąm*, gen. plur. of *dajhu-* ‘land’. OAv. $-xii-$ alternates with $-hii-$ (perhaps a pre-stress variants, see below): OAv. gen. sing. *ašahiā* vs. *ašaxiiā-cā* (cf. YAv. *ašajhā-cā*).

2.1.2.2 Sibilants

The sibilants *s* and *z* are common before stops: *spaiia-* ‘throw’, *zbaia-* ‘invoke’ (<*é^v, j^v*), *daste* ‘is given’, *dazde* ‘is placed’, and nasals: *snaēža-* ‘to snow’, *vasna-* ‘exchange value, price’, *asman-* ‘sky, stone’; *s* is occasionally found before other consonants and in final position (Tremblay 1999), where it is the result of dental assimilation: OAv. *dasuuā* < *dad-sua*, nom. of *t*-stems: $^o tās$ < $^o tāt-s$, OAv. *stavas* ‘praising’ < *stauąnt-s*; OYAv. *šs* ‘was’ < *šs-t*.

2.1.2.3 ‘Shibilants’

The three sibilants *š* *šs* *š* had merged into one sound [š] by the time of our earliest manuscripts, but must originally have been separate:

- *š* = [š] with the voiced allophone *ž* (*duš-* vs. *duž-*);
- *šs*, written *šš*, *ššii* must have been a palatal(ised) sibilant resulting from the palatalisation of *č* [tš] before *i*: *čjāti-* > Av. *š(ii)ā^ti-* ‘joy’, OPers. *šiyāti-* (cf. Sogd. *šāt*, but Khot. *tsāta-* [ts^hāda-] ‘rich’). The corresponding voiced sound has no letter in the alphabet and probably merged with *ž*: YAv. *družā-* (OAv. *drujiiā-* ‘to lie’); *aži-* ‘dragon’ for *ajī-* (OInd. *ahi-*);
- *š* (< ^h*rt*, see section 2.1.2.1) may originally have been a rhotacised alveolar sibilant, e.g. Av. *mašiiā-* ‘(mortal) man’ (rendered in Pahl. as *mašī*) or, perhaps, a retroflex or lateral affricate or flap (cf. Pahl. *mahlī*), but later a retroflex sibilant.

2.1.2.4 Distribution

Most consonants are found in initial position before vowel; exceptions include the voiced fricatives (and probably *f*), the velar nasals (*ŋ* *ij* *ŋ^v*), and *š*.

In final position, we find *m* and *n*, *r*, the dental *t*, and the sibilants *s* and *š*. Final *-s* is also found in sandhi (see section 2.3.3). The consonants *r* and *s* (in *sandhi*) take *ə* as a supporting vowel as finals (e.g. *dātar^v* ‘O creator!’, *kas^v tē* ‘who for you?’).

2.2 The phonology of Old Persian

The Old Persian orthography does not express all the phonemic features of the language, as evidenced by transcriptions into other languages of the time and by comparison with Avestan and Middle and Modern Persian.

2.2.1 The vowels \bar{a} , \bar{i} , \bar{u}

In Old Persian, vowel length is expressed explicitly only in the case of non-initial \bar{a} < Ca-a>; in initial position, <a> spells *a-* and \bar{a} -. In final position, the distribution of *-a* and \bar{a} is historically based; OPers. *-a* < Ir. *-aC* (*-ah*, *-at*, *-an*); OPers. \bar{a} < Ir. *-a*, \bar{a} , $\bar{a}C$. The quantity of final \bar{i} and \bar{u} cannot be determined (see section 1.1.6).

The Iranian diphthongs *ai* and *au* were probably monophthongised to \bar{e} and \bar{o} sometime during the Achaemenid period. Thus, the Akkadian, Elamite, and Greek transcriptions show little if any trace of diphthongs (e.g. *haumavarga-*, Elam. <u-mu-mar-ka>, Akk. <ú-mu-ur-ga-?>, Greek *amiúrgioi*, a tribe of Sakas/Scythians).

2.2.2 Consonants

The OPers. consonant system is set out in Table 3.2.3.

TABLE 3.2.3: OLD PERSIAN CONSONANTS

	Stops/Affric.	Fricat.	Contin.	Nasals	Sibilants
Labials	<i>p b</i>	<i>f</i>	<i>v(u)</i>	<i>m</i>	
Dentals	<i>t d</i>	\emptyset	<i>r, l</i>	<i>n</i>	<i>s, ç z</i>
Palatals	<i>č ĵ</i>	<i>y</i>	<i>(i)</i>		<i>š (ž?)</i>
Velars	<i>k g</i>	<i>x</i>			
Pharyngeal	<i>h</i>				

2.2.2.1 Notes on the consonants

On $\ç$ < Ir. *θr*, *čr*, see section 1.4.

Old Persian may have had a phoneme \check{z} , judging from *nijāyam* [nižāyam?] < **niš-āyam* ‘I went out’, though it may be simpler to assign [ž] to the phoneme /j/ and assume that it was pronounced [dž] as written. Alternatively, OPers. <j> was actually pronounced [ž], and there may have been no phoneme /j/.

Elamite, Akkadian, and other transcriptions attest to etymological, but unwritten sounds, e.g. preconsonantic *n*: < ba-da-ka-> for *ba'daka* ‘bondsman’, cf. MPers. *bandag*; *ciča'taxma-*, proper name, Elam. <zi-iš-šá-an-tak-ma>, Akk. <ši-it-ra-an-tah-ma> (for Median *ciOrantaxma*); sequence ^h*u*: ^h*uwaipašiya* ‘self’, Av. *x'aēpa'θiia-*, MPers. *xwēbaš*. In some cases, the various transcriptions are inconclusive, e.g. *a^huramazdā*, Av. *ahura-mazdā-* (YH.), Elam. <u-ra-maš-da>, Akk. <ú-ra-ma-az-da, a-hu-ru-ma-az-da-?>, Greek *ōromasdēs*, MPers. *Ohrmazd*, *Hormezd*. Before *m*, *h* could be written or omitted (e.g. *a^hmiy* and *ahmiy* ‘I am’). Note also morphophonological variation such as pres. *ha^h-taxša-* ‘be diligent’, imperf. *ham-a-taxša-*.

In final position, the only allowed (written) consonants (other than *y* and *v*) are *m* and \check{s} , perhaps also *s*. The spread of \check{s} as the ending of the 3s and 3p in *akunau-š* ‘he

did/made' (Av. *akər'nao-t*) and similar forms, however, indicates that the corresponding forms of *a*-stems in fact had no consonantal ending (see Allegri and Panaino 1995).

There appear to be the same kind of restrictions on initial and internal consonants as in Avestan, and the same kind of groups are found (e.g. *xšn*: initial and intervocalic). There are no examples of two-consonant groups in final position, and forms such as Av. *āfš* 'water' nom. sing. may have been transferred to the *i*-declension (*āpī[š]-šim* (?) 'the water . . . him', beside athematic inst.-abl. plur. *abiš* < *ap-biš*).

2.2.3 Late Old Persian

The inscriptions from Artaxerxes II (404–359) on are written in what is clearly a post-Old Persian/proto-Middle-Persian stage of the language (see Schmitt 1999: 59–118; Skjærvø 1999b [2002]: 158–61). Those from Artaxerxes I and Darius II are less clear; they are written in a late-Old Persian form, as seen from the thematic forms such as *dārayavaušahayā* and some grammatical constructions, but they are too short and formulaic to tell us much.

Among the (orthographic-)phonetic peculiarities of the post-Old Persian stage note (see Schmitt 1999): <Cy> for <Ciy>: *nayāka-*, *abayapara*; use of *y* or *iy* to write long *ē*: *paradaydām* for *pardēd*^(?) (MPers. *pālēz*); contraction of *iya* > *ī* in *martihayā* (A^{3?}); *st* for *št* in *nistāya*; loss of final consonants and their preceding vowels in endings, as evidenced by the indiscriminate use of short and long vowels and omission of final *m*.

2.3 The morphophonology of Avestan and Old Persian

There are several kinds of synchronic alternations in Old Iranian: those inherited from earlier stages of the languages; those due to historical developments; those due to analogy; and, for Old Persian, those due to the existence of Median beside Old Persian forms.

2.3.1 Vowels

Most of the alternations in the vowels derive from the conditioned variants of *a* and *ā* before *h* and nasals and *ī*/*i* and *ū*/*u*. For details, see de Vaan 2003, Chapter 6.

2.3.1.1 Centralising of *a*, *ā* > *ə* (*ē*)

ah > *ə* (*ē*). Ir. *h* caused centralising of a preceding *a*: OAv. *amēhmaidī* 'we have thought'; *vacē* 'speech' (but mostly *-ō* reintroduced from YA^v?).

In Young Avestan, the *a* remained in initial and medial position (reintroduced into the OAv. text: *ahmat* 'us'), and final *-ah* became *-ō* (occasionally *ē*: *nəmē* 'homage'); in sandhi, the *a* reappears: *-a-s*^o.

aN > *əN*. Short *a* was centralised before nasal, in more positions in Old than in Young Avestan; – initial: OAv. *ēnə'tī* 'no going', *ēmauuant-* 'powerful' (YA^v., *a'niti*, *amauuant-*); in final: OYA^v. *-ām*, *-ān*.

Final *-an-h* > Ir. *-aṅh*. In Old Avestan, this became *-ēṅg*, notably in the thematic acc. plur. *daēuu-ēṅg* 'old gods' and the *n*-stem gen. sing. *x^v-ēṅg* 'sun' (< *huu-aṅh*). In Young Avestan, it probably became nasalised *-ē* which developed variously to *-ē* and *-q*: acc. pl. *yazat-ē* 'gods', *aēsm-ṇ* 'firewood', gen. sing. *bar^vsm-q* 'barsom'; final *-uē* in turn

became (*uu*)*ū*; acc. plur. *daēuu-ū* (*daēū*), gen. sing. *hū* (<*huu-ū*), etc. (see also Hoffmann 1970; de Vaan 2003: 492–93).

Final *-āh* > *ā*. Long *ā* became rounded long *ā* before *h* in final: *-āh* > *-ā*; the *ā* remains in sandhi: *-ā-s°*.

āN > *āN*. Long *ā* was centralised to *ā* before nasal in a few Old Avestan words: *xiiēm* ‘may I be’ and *strēm*, gen. pl. of *star-* ‘star’ (both disyllabic < *-a’am*). In both Old and Young Avestan, *āN* alternates with *-qN* in the manuscripts.

au- > *əuu-*. Ir. *au* became *əuu* sporadically in initial position: OAv. *ə.vaocaŋ* ‘he has said’, YAv. *əuuər°z°* ‘not producing’; and when followed by *ilī*, OAv. *kəuuitāt-* ‘the word/title of *kauui*’, YAv. *səuišta-* ‘richest in life-giving strength’, *huuəŋhəuuīm* ‘*comfort’ (<*huuəŋhəuuīam*).

2.3.1.2 Rounding of *a*, *ā* > *ā* or *ō*

a + u. YAv. *a* became *o* before (primary or secondary) (*u*)*ū* (*uu*) in the following syllable, regularly before *r*: *po^oru-* ‘much’ < *paru*, *po^oruua-* ‘earlier’ < *paruuā-*, also when the *u* was the result of contraction: *po^orum* < *paruuam* (see Kellens 1986), sporadically elsewhere: OYAv. *vohū* ‘good’ (but *vanhu-*), *vohunī-* ‘blood’, YAv. *moγu°* ‘*Magian’.

a in labial context. Change of *a* > *ā* or *ō* (also *ā* > *ō*) is found in Old Avestan, sporadically in Young Avestan, when *a* is in multiple labial context: *apō mā* vs. *apā*; YAv. dat. plur. *“ruuōⁱbiiō* < *uruuan-* ‘soul’; dat. sing. *māuuⁱtia* ‘for me’ (OAv. *maⁱbiiā*); note also *dəb^onao-* ‘deceive’ (2 syllables; OInd. *dabhno-*), *duž^oz^obā* ‘making bad invocations’ < *duž-zbāh-* (<*-juā-*). In Young Avestan, final *-uuō* for *-uua* is found in a few cases, e.g. instr. sing. *bāzuuō* ‘arm’. OAv. *hōiθōi* < *haēθa-* ‘*cordwork’, with *ōi* for *aē*, may be the result of assimilation.

ā + u. Labial umlaut of *ā* is found especially in Old Avestan when followed by *ū* (*u*) in the next syllable: *u-* stems: *jiūōtūm* < *jiūtu-* ‘*livelihood’.

a + r. The sequence *-a + r-* (*-a + r°*) in derivatives of *fra-ər°* ‘move forth’ gave OYAv. *-ər°*: *frēr^otōi* ‘moving forth’, *frēr^onao-* ‘send on its way’; the OAv. abl. sing. *frēr^otōis* ‘from moving forth’ may show assimilation.

2.3.1.2a Old Avestan initial *ā* > *a*-*lā*-

In several words with initial *ā* caused by one of the above processes, the original *a-lā* were reintroduced: *ā^oādū* for **ādū* < **ādu* ‘?’ by *u*-umlaut; *ā^onū* for *anu* ‘along’, centralised before nasal or by *u*-umlaut; *ā^oŋhā* < *āhā* instr. sing. of *āh-* ‘mouth’, centralised before nasal or before *h*; *ā^o* < *auua* ‘down’, by labial assimilation.

2.3.1.3 Raising of *a*, *ā* > *e*

a and *ā* can be raised (fronted) when preceded by *i* and followed by palatal or palatalised sounds.

ia became *ye-*, *-iie-*, *-iē* (< *-ai*):

- before palatal consonant: OYAv. *iōiejah-* < *ōijajah-* ‘danger’;
- before palatalised consonant: OYAv. *yezī* < *iazī* ‘for if’; – OAv. fem. sing. *yezuū* ‘young’ < *iazīyī-*; – YAv. inf. ending *-iieⁱδiiāi* < *-ia-dīāi*; – *frāiieze* < *-iazīē* < *-iazai* ‘I sacrifice’; – OAv. loc. sing. *yesnē* < *iasnⁱe* (<*iasn^oi*), Av. *yesniia-* < *iasnⁱiia-* ‘worthy of sacrifices (*yasna-*)’; – comp. fem. sing. *-iiehī* < *-iiahī-*; – 2nd and 3d sing. pres. ind. act./mid. *-iiehīle* < *-iiahīlē* and *-iieⁱtil-iiete* < *-iatilīē*, *-iēj^hhē* < *iahīa* and *iahai*: Av. gen. sing. *yej^hhē* < *iahīa* ‘whose’, OAv. inf. *-iiej^hhē* < *-iahai*;

- in final after *h, r, s* in some words: *ahe* ‘his’ < *ahīa* (also gen. sing. ending *-ahe*); *dre* ‘Aryans’ < *arīa*; *nase* ‘perish!’ < *nasīa*;
- *īāN* became *īeN* when followed by *i, ī, or -īē*: 1s pres. ind. act. *-iiemi* < *īāmi* subj. *-iieni* < *-īāni*, mid. *-iiene* < *-īānīē* (< *-īānai*).

2.3.1.4 Combined centralising and raising and rounding of *a*

a > ə > i. After palatals, *ə* before nasal further became *i* in Young Avestan, sporadically in Old Avestan:

- after the palatals *c j* (but *a* was sometimes restored): OAv. *hacēnā* ‘company’ (instr. sing.), beside OYAv. *hacīntē* ‘they follow’, YAv. *raocana-* ‘window’ vs. *raocinauuant-* ‘*bright’;
- *īaN* and *ūaN* regularly became OAv. *īēN*, *ūēN* or *īiN*, *ūīN* (the *a* occasionally remains), YAv. *īaN* or *īiN*, *ūuN*: OAv. *yām*, YAv. *yim* ‘whom’; OAv. nom.-acc. dual *yāmā* ‘twins’, YAv. *yima-* ‘Yima’ (but OAv. *yimas-cī* ‘also Yima’); OAv. *driiēmā*, YAv. *driiaman-* ‘Airriaman’; YAv. *po^orum* < *parūyam* ‘before, in front’; – in final syllable, Old Avestan often preserves the *ā* (*yām*, *-iiēm*; always *-iiēn*, *yēng*, *-iiēng*; *-uuēng*), but contraction of *īam > īm* is frequent: OAv. *aniiēm* ‘other’, YAv. *a^onim*; YAv. *haxāim* (< *-ājam*) ‘companion’; OAv. *hā^oθiiēm* and *hā^oθīm* ‘true, real’, YAv. *hā^oθim*. Contraction to *-ūm* is found only in the 2nd plur. mid. ending *-dūm*.

īaN and *ūaN* after vowels: Young Avestan regularly has contraction, Old Avestan sporadically: *aīiN > aēN*, *āīiN > āiN*, *ūīiN > īN* (*iN*): OAv. *aiiēm* ‘this one’, YAv. *aēm*, 3rd plur. pres. inj. *-aēn* < *-aijan*. In the 3p opt. of *aiiā*-stems, *-aijājan* became *-aiiāēn*, occasionally preserved in the manuscripts, but most often replaced by *-aiiēn* (Skjærøen 1998: 191).

aūuN > aoN (*āuN*), *āūuN > āuN*, *uūuN > ūN* (*uN*): OAv. *kər^onaon* (< *-aūan*) ‘they shall make’, YAv. *kərənāun*; YAv. *baon* ‘they became’; YAv. *adāuṇta* (< *dauua-*) ‘they spoke deceiving words’; OAv. *tuuēm* < *tuūam* ‘you’, YAv. *tum*; YAv. *bun* < *buūan* ‘they shall become’; YAv. *aū* < *aūaṇti* ‘them’.

The sequences *-aiūm*, *-aiūam*, and *-aiūam* were simplified to YAv. *-aēum*, *-ōiium* (*-aoiium*), or *-aēm*: *vaiiu-* ‘Vaiiu’: acc. *vaēm*; *daēuua-*: acc. *daēum*, *vīdaēuua-* ‘discarding the old gods’: acc. *vīdōiium*; *haoiia-* ‘left’: *haoiium*, *hōiium*, *hōim*.

Combined labialisation and palatalisation is seen in YAv. *paō^orim* ‘first’ < *parūijam* (OAv. *pa^oruuīm*, trisyllabic).

The forms with *a* are frequently restored, especially in certain morphological categories; thus, YAv. 3p *-aiiēn* is more frequent than *-aēn*. OAv. *īant* before *ī, ē* (< *-ai*) remained or became *iiēnt*, *uuant* remained. In Young Avestan, *īant* became *-iēt-* (usually after *c, j*, sibilants, and *θ*), but was often retained, and its *a* could then be raised to *e*: *-iie^onti*, *-iiēnte*. YAv. *uant* became *-uēt-*, but adjectives in *-ŋ^ohant-* have analogical *-ŋ^ohant-* beside regular *-ŋhūnt-*.

In both Old and Young Avestan, diphthongs resulting from contraction of two syllables remain metrically disyllabic.

2.3.1.5 Shortening of *ā* and lengthening of *a*

Ir. *ā* is shortened in Old Avestan under phonetic conditions, for rhythmic reasons, or as a result of stress movements: before *ii* or *uu*, in the gen. plur. ending *-anant*; in the preverb *ā-*, in initial syllable when an enclitic was attached to the word, in antepenultimate or

earlier syllables or when the word was unaccented (?): *a-iiaghā* ‘you shall request’ < *ā-yā-*; *a-xštaŋ* ‘he will stand by’ < *ā-°*; *abiias-cā* ‘and to them’ < *ābiias°*; *uštanəm* vs. *uštānāi* < *uštāna-* ‘life breath’. It is shortened in Young Avestan in similar environments.

Ir. *a* is frequently lengthened in Old Avestan in initial syllables, occasionally in non-initial syllables: nom. plur. *kāuuaiias-cā* < *kauui-* ‘poet’; – after *ii:* 3rd sing. inj. mid. *maniīātā* ‘he thinks’ < *-jata-*; – after *ui:* instr. sing. *x’ēnuuātā* < *x’ēnuuaŋt-* ‘containing the sun’; – sporadically elsewhere: gen. plur. *hātqm* < *haŋt-* ‘being’.

Ir. *a* is lengthened in Young Avestan in initial syllables: *āiia* < *aia:* *xštāuuaiiō* vs. *xštāuuīfiīō* < *xštāuui-* (ethnic name) and sporadically elsewhere.

For details, see de Vaan 2003, Chapter 2.

2.3.2 Consonants

2.3.2.1 Assimilation and dissimilation

2.3.2.1a Voice assimilation

The results of voice assimilation (voicing and devoicing) seen in Iranian occurred at various times and in various situations.

Devoicing affected unaspirated voiced stops before unvoiced sounds, mainly *t* and *s*, *š*: Av. *baxta-* ‘shared’ (< IIr. *b^hak-ta* < *b^hag-*); Av. *vista-* ‘found’ (< IIr. *uit^h-ta-* < *vid-*), *baxša-* ‘distribute’ (< IIr. *b^hakš-a* < *b^hag-*). On the analogical replacement of *bd*, etc. with *pt*, see section 2.3.2.1b.

Voicing was more common, unvoiced stops, affricates, and sibilants being assimilated to a following voiced stop or sibilant:

- stops: YAv. *upa-bdi* ‘on the foot-hill (of)’ (< *upa-* + *pad-*); YAv. dat.-abl. plur. *d^hīiō*, OPers. instr.-abl. plur. *abiš* < *āp-* ‘water’ + *-b₁ah*, *-biš*; OAv. dat.-abl. plur. *drəguuō*. *d^hīiō* YAv. *druuāŋbīiō* < *druguat-* ‘possessed by the lie’ + *-b₁ah*; instr. plur. Av. *azd^hbīš* < *ast-* ‘bone’ + *-biš*;
- affricates: OAv. 2p inj. *ōβar^hš-dūm* < *ōβar^hs-* ‘to fashion’ (< IIr. *θuaré-*);
- sibilants: OAv. *zdī* ‘be!’ < *ah-ls-*; *mazdā-* ‘all-knowing < placing (all) in (his) mind’ < *mas-dā-* (< *m₁s-*); – *niš-bər^hta-* ‘removed’ (< *niš-* + *bər^hta-*); *duš-zaotar-* ‘evil libator’ (< *duš-* + *zaotar-*).

Final *š* was voiced before vowels and voiced consonants in compounds and before suffixes (see also section 1.2.1.3: OAv. *ər^hš-uxōa-* ‘straight utterance’ (but YAv. *arš-uxōa-*); *duš-āōra-* ‘with bad (constricted) breathing space’; *aβ^hš-dāna-* ‘water container’ (< *āfš-* + *dāna-*).

2.3.2.1b ‘Bartholomae’s law’

A special case of assimilation is ‘Bartholomae’s law’, according to which an IIr. final voiced aspirate of the stem passed both its aspiration and voice on to an initial *t* or *s* of an ending, so as to produce clusters of the type, e.g., *-b^h-t-* > *-bd^h-*, *g^h-z* > *gž^h* (with *ruki*, see section 1.2.1.3). The resulting clusters remain in Old Avestan: IIr. *aug^h-* ‘to present oneself (as), say’ had 3s *aug^h-ta* > *aog-d^ha* > OAv. *aog^hdā* and 2s *augž^h-sa* > *aug-ž^ha* > OAv. *aogžā*; from IIr. *uaj^h-* ‘to convey’ we have *uaj^h-tra-* ‘conveyor, draught animal’ > *uaj^h-d^hra-* > OAv. *važdra-*; IIr. *dad^h-* ‘place’ had 3s mid. *dad^h-tai* > *dad^h-d^hai* (see section 1.3 > OAv. *dazdē* ‘it is placed’ (vs. IIr. *dad-tai* ‘it is given’ > *dat^htai* > OAv. *dastē*, YAv. *daste*);

IIr. *d^hab^h-* (> *dab^h-*) ‘deceive’ had the desiderative stem *di-db^h-sa-* > *di(d)bž^ha-* > OYAv. *diβžā-* ‘seek to deceive’ (but OInd. *dipsa-* with elimination of ž).

In Young Avestan and Old Persian, these groups remain only in isolated words: YAv. *ubda-* ‘woven’ < *ub^h-ta-* (< *uab^h-*); OAv., OPers. *azdā* ‘known’ (OInd. *addhā*), but most often they are analogically replaced by unvoiced groups: YAv. *aoxta*, *vaštar-* ‘draught animal’, *dapta-* ‘deceived’ (for *dabda-*), *gər^ofšā-* ‘grasp’ (for *gr^ob-ž^ha-* < *gr^ob^h-* ‘seize’); YAv., OPers. *basta-* ‘bound’ (OInd. *baddhā-*).

2.3.2.1c Geminates

Geminates resulting from internal sandhi and assimilation were simplified: YAv. *usnā-* ‘wash (up)’ < *us-snā-*; *višān-* < *viš-šān-* (< *vis-* + *hān-* ‘who gains a village’, with *ć-š* > *šš* > *š*).

Analogically restored forms are frequent, e.g. YAv. *viš.har^ozana-* ‘abandonment of the villages’ for *višar^o* (< *vis-* + *har^o*).

2.3.2.1d Dissimilation

Dissimilation in point of articulation is perhaps to be seen before *m* in *vahma-* ‘hymn’ if from *vaf-* (pres. *ufiia-*) ‘weave’, and *daxma-* ‘burial mound’, if from *dafma-* < *d^hab^h-* ‘construct’ (cf. Gk. *taphos*, Hoffmann 1975: 338; Skjærvø 2005b).

Voice dissimilation is found in the groups *fθr* > *fδr* and *xθ(r)* > *xδ(r)* (no examples of *fθV*): *naf^oδrō* < *naptar-* ‘grandson, scion’; *uxδa-* ‘utterance’, *apāx^oδra-* vs. *apāxtar-* ‘northern’.

2.3.2.2 Spirantisation

2.3.2.2a Unvoiced stops and the unvoiced fricatives f θ x

The unvoiced fricatives *f θ x* are independent phonemes before vowels (see section 1.2.1.4), but allophones of *p t k* before consonants, e.g. IIr. *ky* > Ir. *xu* in YAv. perf. part. *vaox^oah-* ‘having spoken’ < *uā-uk-uāh-* (see Skjærvø 1997a). The stops remain after sibilant: OAv. *āsk^oti-* ‘following’ (< *sak-* ‘follow’), gen. plur. *strēm* ‘stars’, *vāstriia-* ‘forager’.

In Avestan, Ir. *ft* appears to have reverted to *pt*: OAv. sing. nom. *ptā* ‘father’ vs. dat. *f^oδrōi* < *fθrai*, YAv. *dapta-* ‘deceived’.

2.3.2.2b Voiced stops b d g and voiced spirants β δ γ

The voiced stops *b d g* and the voiced fricatives *β δ γ* are in complementary distribution in both Old and Young Avestan.

In Old Avestan, the Iranian voiced stops *β* and *γ* are allophones of *b* and *d* before ž: *diβžā-*, *aγžāonuuamna-* ‘?’; otherwise the voiced stops remain unchanged; *β* replaces *u* after *θ*, and *δ* replaces *θ* after *x* and *f* (*xδ*, *x^oδ*; *fδ*, *f^oδ*).

In Young Avestan, the Iranian voiced stops remain regularly only initially and after nasals and sibilants (*z ž*), sporadically also in other words. Elsewhere they become the voiced fricatives *β δ γ*: *a^oβi* preverb (OAv. *a^obī*), *duγda* ‘daughter’ (OAv. *dug^odā*), *driγu-* ‘poor’ (OAv. *driγu-*); *γ* is also found initially before nasals and sibilants: *γ^onā-* ‘woman’ (OAv. *g^onā-*), *γ^omata-* ‘gone’, *γžar^o.γžar-* ‘flow’. Finally, *f* and *x* are voiced before ž: *aβ^oždāna-* ‘water container’, *vāγž(əli)biiō*, and *δ* is an allophone of *θ* in the groups *fδr* and *xδ*, see section 2.3.2.1d.

Later changes. The resulting Young Avestan voiced fricatives are all subject to further

contextual and chronological, possibly also dialectal, changes and variations: β tends to become \underline{u} > uu : *viiāuua'tī-* 'shining' < *vi-ā-βā-* < *bā-* 'to shine', *auuara-* 'to bring' (< *ā-βara-*). In turn, \underline{u} may combine with preceding *a* to form the diphthong *ao*: *a'βi* 'to' > *auui* > *aoi* (mss. *aōi*, *aōui*, etc.; *ada'βiia-* 'not to be deceived': sing. nom. masc. *adaoiiō*, but acc. *adauim*).

δ vs. θ : Sporadically, δ alternates with θ , notably in forms of *vaēθ-* (*vaēδ-*) 'to know' and in pres. *dadā-* (*daθā-*) < *dā-* 'to give, place', sporadically elsewhere.

γ is lost before \underline{u} and sporadically before *u*: YAv. *druuaṇt-* < *druγuaṇt-* *drəguuaṇt-*); *rəuuī-* 'fast' < *rayūi-* (fem. of *rāγu-*); *raom* < *rayum*; *mo^urum* < *marγum* 'Merv', but *driyum*, *moyu^o* (see Skjærvø 1997b).

2.3.2.3 Consonant groups and anaptyxis

Most Indo-Iranian consonant groups survived in Avestan, and anaptyxis does not create additional syllables as shown by the meter. The anaptyctic vowels is typically ə , but also *a*, *i*, *u* (often as conditioned variants of ə). Anaptyxis is found in more contexts in Old than in Young Avestan and varies among the manuscripts. Following are some of the more unusual groups (others are common and trivial):

Stop + stop:

- initial: OAv. *pt-* in *ptar-* 'father' (nom. *ptā*, *p^utā*; acc. *p^utarēm*, *ptarēm*); *db-* in *d^{ai}bitā* '??' (< *dūtā*); *tk-* in *tkaēša-* '*guidance'; YAv. *p^(a)tar^ata-* 'winged'; – in Young Avestan, *pt-* in *ptar-* has been replaced by *pit-* and *db-* by *tb-*; – other groups have been simplified: *db-* > *b-*: *bitīm* 'second(ly)' (< *dbitiīam*, but *ā-tbitimlā-δbitim* 'a second time'); Ir. *pt-*, *kt-* > *t-*: YAv.: *tūrīia-* < *ptūrīia-* 'uncle', *kturīia-* (< *ktūrīia-*, cf. *caθrušī* 'fourth'; but *ā-xtūrīm* 'a fourth time');
- medial: OAv. *āsk^{ti}-* 'following*' (< *ā-skti-*); gen. plur. *dug³draṃ* < *dug³dar-* 'daughter'; YAv. *hapta* 'seven', *aṭka-* (*aḍka-*) 'coat'; *ubda-* 'woven'; YAv. *ā-tbitimlā-δbitim* 'a second time'.

Stop + non-stop:

- initial, common: + continuants: OYAv. *br-*, etc.; + glides: *dii-*, *duu-*; OAv. also + fricative: *d³j-*; + nasal: OAv. *g³n*, *g³m-*, *d³m-*; – in Young Avestan less common: *dj-* > *j-*, *gn*, *gm-* > *γ³n*, *γ³m-*, *dm-* > *nm-*.

Fricative/sibilant + continuant, common:

- initial and medial: *fr-*, *fi-*, *fš-*, *sr*, *zr-*, *γž-*, *f³δr* (< *fθr*) and medial *x³δr* (< *xθr*: OAv. *f³raša-*, YAv. *fraša-* 'filled with (life-giving) juices'; YAv. *fii^uṇ^ha-* 'to hail', OYAv. *fšūiiaṇt-* 'cattle tender'; OAv. *s³raoša-*, YAv. *sraoša-* 'readiness to listen, Sraoša'; OYAv. *z^(e)razdāiti-* 'reliance'; YAv. *γžara-* 'flow');
- medial: OYAv. *vax³δra-* 'speech organ', OAv. *raf³δra-* 'support'; YAv. gen. sing. *naf³δrō* < *naptar* 'grandchild, scion'.

Two fricatives + consonant. Initial groups of two fricatives (fricative + sibilant) + consonant are found occasionally: OAv. dat. sing. *f³δrōi* < *ptar-* 'father', *fš³ratū-* '??'; YAv. *fštāna-* '(woman's) breast', *xštuaa-* 'sixth', *xšnūman-* 'favour', *xšmākəm* 'your (plur.)'.

Final groups with *š* or *t*: *kə³fš* 'body', *vāxš* 'word'; *vaxšt* 'he has grown'.

Final *-t* is, from a synchronic perspective, sometimes dropped after a sibilant.

Historically, we must distinguish between two cases: 1. *-t-t* > *-t³-t* > Av. *-s-t*: OAv. *“rūraost* 'howled' < *“raud-t*; YAv. *nāist* 'scorned' < *nāid-*; 2. *-s/š-t* > Av. *-s/š*: OAv. *ās*

‘was’, *vaxš* ‘he has overcome’ < *vān-s-t*; *xšnāuš* ‘he has favoured’ < *-š-t*. The *-t* was then reintroduced by analogy: OAv. *cōiš-t* ‘he has pointed out’, *vaxš-t*, YAv. *tāš-t* ‘he fashioned’ (Tremblay 1999).

2.3.2.3a Groups at morpheme boundary

Groups found only at the morpheme boundary include in principle all combinations of any final group + *C* or any initial group: dat.-abl. *vāyž^hbiiō* < *vak-ivac-* ‘word’, *frādaṭ.fšau-* ‘cattle-furthering’, *Orafs-ca* ‘and satisfaction’, *afš-tacin-* ‘flowing with waters’, *afš-ciOra-* ‘containing the seed of water’, *kə^rfš-x^rar-* ‘body-eater’; *fraor^tfraxšnini-* ‘(a mind) with foreknowledge of and turned toward the reward(?)’.

2.3.2.3b Anaptyxis in Old Persian

Anaptyxis is found in OPers. only in groups containing *d* in the vicinity of *u*: *d^uruwa-* ‘healthy, whole’, *sug^uda-* vs. *sugda-* ‘Sogdian’.

2.3.2.4 Palatalisation and labialisation of consonants

In the Avestan text, palatalised and labialised consonants are indicated either by special letters (*ń, ř, ́, ́́* and *ř^u, x^u*) or by writing *i* and *u* before the consonants (*i-* and *u-*epenthesis). The exact phonetic nature (and age) of the phenomenon is not known. The sibilants are not marked for palatalisation (*s, z, š, ž*), but permitted palatalisation of preceding *a* (see section 2.3.1.3), and *m* was not affected. Labialisation affected only *r* and proto-Ir. *h*. For details, see de Vaan 2003, Chapter 7.

Palatals + *i*: The alveo-palatal consonants *c, j* usually remained before *i*, e.g. *ci-* ‘who, what?’, but were palatalized before vowel:

- *cj* > *š(ii)*: YAv. *š(ii)āiti-* ‘happiness’ < *čjāiti-* (also OPers. *šiyāti-*); fem. *apaši* ‘backwards’ < *apāčī*;
- *jš, ji* > *ž, ži* (rare): YAv. *družā-*: OAv. *drujiia-* ‘to lie’; *aži-* ‘dragon’ < *aji-* (OInd. *ahi-*).

2.3.2.4a i- and u-epenthesis in Avestan

i-epenthesis occurs regularly before *r*, dentals, and labials and in consonant groups: Av. *rixta-* ‘left’, OAv. *gādī* ‘come!’, 1p mid. ending *madē*; YAv. *staoⁱti* ‘praises’, *vand^unī-* ‘victorious’ (fem.); *oiiejah-* ‘danger’; *aⁱpī, aⁱbī, a^ebiiō*; – between compounds: *tarōⁱdīti-* for *tarō.dīti-* ‘scorn’; – in consonant groups: dat.-abl. plur. *vāyžⁱbiiō* (< *vāk-* ‘word’); *a^rime* ‘in peace and quiet’ (< *ar^um^e*, cf. *armaē^o*); – vocalic *ə^r*: *miriia-* ‘die’ < *m^r-ia-*;

u-epenthesis. This occurs regularly before *r*: YAv. *rūraod-* ‘howl’, *a^uruša-* ‘rosy white’, *po^uru-* ‘much’; OAv. *pa^uruuia-* ‘first’.

Combined *i-* and *u-*epenthesis. This occurs before *r*: YAv. *pao^uriia-* < *parūia-* ‘first’; *pao^urī-* < *parūī-*, fem. of *po^uru-* ‘much’; – vocalic *ə^r*: *tūriia-* ‘father’s brother; fourth’ (< (*plk*)*trūia-*), dat.-abl. plur. *nəruuiō, nuruiō* ‘for the men’ (< *n^r-ūiō* < *n^r-bⁱah*).

2.3.2.4b Dissimilation of *iī* and *uu* to *əi*, *aī* and *əu*

In a further development, the sequences *ii* (*iī*) and *uu* (*uu*) resulting from epenthesis were dissimilated to *əi*, *aī* (*əi*, *aī*) and *əu* (*əu*): OAv. *ənəⁱti-* ‘non-going’ (< *an-ⁱiⁱti-*; cf. *x^rtīcā* ‘and good going’); *daⁱdūiaṭ* ‘shall she see’ (< *dīⁱdūiaṭ*?).

On absence of epenthesis due to stress, see section 2.3.4.

2.3.2.4c Epenthesis in Old Persian

Possible examples of epenthesis include *pa'siyā* 'before' (MPers. *pēš* if from *paθiā(k)*) and *yāu^hmani-* and *yāu^hma'ni-* 'in control(?)' if from *yau^x-man-* 'harnessing' (< *yaug-*; several other interpretations have been proposed, see, e.g. Hoffmann 1975: 56–57, 1976: 633 n. 20).

2.3.2.5 Glides

The glides–initial *y-*, *v-* and intervocalic *i*, *u*–are in complementary distribution.

i, *u* > *ī*, *ū*. Both postconsonantic and postvocalic *i* and *u* were apparently realised (by the time of our mss.) as *ij* and *uj* (including *uj* < *uβ*), written *ii* and *uu*; note OAv. *ōi-i* = YAv. *ai-i* for *ai-i*. In Old Persian, < *iy*, *uv* > is written for postconsonantic (*ī*), *i*, (*ū*), *u*.

a-uC > *aoC* or *āuC* at the morpheme boundary (reduplication, prefixes): Av. perf. stem *vaon-* < *ua-ūn-* < *van-* 'to win', OAv. *vāunu-* '*winner'; YAv. *vaox'ah-* 'having spoken' < *ua-uk-uah-* < *vak-* 'speak'.

Initial *ī-* > *ii* and *ū-* (< **uβ-*) > *uu*. These are rare and are often written *y-* (*i-*), *v-* in the manuscripts: *īiēiān* (*iēi-*, Y. 42.6), subj. of the perf. stem *ii-āii-* < *ai-* 'go' (cf. OInd. *iyāy-*), *uua* (*ūua*) < *uβa* 'both'.

ij and *uj*. These groups (also with *u* < *β* and *u*) survive occasionally in the manuscripts (only *iiuu*; no examples of *uiii*?), but sometimes were resolved by anaptyxis (or dissimilation?): initial, *viēmi* and *v^uiēmi* < *vja-* 'pursue', *y^uuua* = *yūua* (*yūua* < *yūuan-* 'youth'); – medial, *hāuu^uiia*, *hāuu^uiia-ca* 'left (hand)' (cf. fem. *haoiā-*); *driuu^uiās-ca* (< *driuu^ui-* fem. of *driyu-* 'poor'; see Skjærvø 1997b). Often they were simplified to *iu* < *iuu*>. Final *-iūō* (< *-iau*) appears as *-iiō*: voc. *maniiō-* < *maniiu-* 'spirit'; loc. *daījhō* < *daījhu-* 'land'; see de Vaan 2003, §25.10; Skjærvø 2005a: 202–3).

du. The initial cluster *du-* shows up variously (see section 2.3.2.3): OAv. *duuaēšah-* 'hostility', *d^ubišīia-* 'be hostile', YAv. *ībaēšah-*, *ībišīia-*, but perf. *diduuaēša* 'I have been hostile'; OAv. *d^ubitīia-* 'second', YAv. *bitīia-*, but *ātbitīm* 'a second time', OPers. *duviīya-*. Intervocalic *du*: 2nd plur. mid. *-duam* > OAv. *-dūm*, YAv. *-δβəm*; YAv. *ər^uδβa-* 'upright', beside *ər^uduua-*.

ou. The cluster *ou* became Av. *oβ*: *ratuš* 'model' vs. gen. *raθō*, but loc. *gātuiō* < *gātau* < *gātu-* 'place'.

In Old Persian, the *o* also spread to the rest of the paradigm: *xratu-* 'intelligence' (MPers. *xrad*) and *xraou-*, but only *gāou-* attested (MPers. *gāh*).

2.3.2.6 Simplification of consonant groups

Certain consonant groups lose one consonant. The loss may have been early (e.g. earlier than the IIr. *ru(p)ki* rule) or late: OYAv. *diβžā-* 'seek to deceive' < *di-dbžā-* (OInd. *dipsa-*); OAv. loc. plur. *naššū* < *nap(t)-šū* < *napāt-* 'grandson'; YAv. *°fšm-* 'with . . . breasts' (cf. *fštāna-* 'breast'); YAv. *sāθr-* vs. *sāstar-* '(false) teacher' (see also section 2.3.2.1a).

Final *-ant-s* regularly became Av. *-qs* (*cuuqs* 'how great?'), but appears to have been simplified to *-an-s* early enough to develop like IIr. *-an-s* in the nom. sing. of active present participles > proto-Ir. *-aŋh* (OInd. *-am* before vowels): YAv. *ja'diūq* and *framrū* (< *°mruyanh*); the YAv. nom. sing. ending *-ō* may then have been substituted for *-ō*. (Schindler 1982 proposed dialect influence in these forms.)

2.3.3 Sandhi

Stem-initial or -final consonants or vowels are regularly modified by preceding or following sounds after prefixes or before suffixes (derivational or inflectional), as well as, more generally, between members of compounds, according to the rules discussed above.

Changes to a word-final consonant or vowel are usually caused by enclitics, mostly Ir. *-ca*, *-cit*, occasionally pronouns with initial *t-*, and, exceptionally, before nouns with initial *t-*. This sandhi principally affects final vowels and final Ir. *-asl-ah*:

- before enclitic Ir. *-ca*, *-cit*: OAv. *manas-cā* ‘and thought’, *gaē0âs-cā* ‘and living beings’, *yimas-cī* ‘even/also Yima’; YAv. *mi0ras-ci* ‘even/also Mi0ra’, *haomas-ca* ‘and the haomas’, *vīspâs-ca* ‘and all’; OPers. (with *-s-c-* > *-š-c-*): *manaš-cā* ‘and thought’, *kaš-ciy* ‘anybody’;
- before enclitic pronouns and nouns: OAv. *yāngs-tū* ‘whom you’, YAv. *ōiy³matas tūrahe* (proper name).

In compounds: OAv. *rāniī0.skər³tī-* ‘joy-making’ (< *-as-k-*); YAv. *drujas-kanā-* ‘den of the Lie’; OPers. *vahāyaz-dāta-*, proper name: ‘given as the better (of the two)’.

Final *-t* was probably assimilated to following sibilants and affricates, but in Avestan it was restored as *-t*; in Old Persian, it was also analogically replaced by *-š-c-*: YAv. *aṭ-cī*, *yaṭ-cī*; OPers. *aciy*, *yacy*, but *aniyaš-ciy*, *avaš-ciy*, *ciš-ciy*.

2.3.3.1 Anaptyxis in sandhi

Anaptyxis in sandhi occurs after OAv. final *-m* before fricative or sibilant: *yām³ spašuoā* ‘whom you *regard’, *hām³.fraštā* ‘he consulted with’; – after final *-s/-š* before consonant: OAv. *vasas³.xša0ra-* ‘having command at will’, YAv. *us³.hištən* ‘they stood up’, *nəmas³ tē* ‘homage to you!’, *yas³ 0βqm* (‘half’ -sandhi: < *yas* + *0βqm* for **yas tuuqm*), *aīhās³ tanuuō* ‘of this body’.

2.3.3.2 Final vowels and diphthongs in sandhi

Before enclitic *-cā*, vowel quantities may change and, at least in Old Persian, original quantities reappear: OAv. *sauua-cā* vs. *sauuā* ‘life-giving strengths’, *x³ūti-cā* ‘and good going’ vs. *ənə³tī* ‘non-going’; – YAv. *mašiiā-ca* vs. *mašūa* ‘mortal men’, etc.; – OPers. *mana-cā* vs. *manā* ‘of/to me’; *avahāya-rādiy* ‘on account of that’ vs. *avahāyā*.

In Avestan, monophthongised final diphthongs reappear: *taē-cī* ‘they too’ vs. OAv. *tōi*, YAv. *tē*; YAv. *drao-ca* ‘and in tree’ vs. **druuō*.

2.3.4 Stress-related alternations

Changes in vowel length are found frequently throughout the Avestan corpus (see section 2.1.1.7e). At least some of these changes seem to be due to stress, which is sometimes termed ‘rhythmical shortening/lengthening’. Other cases of lengthening and shortening of vowels includes final *ā*, *ī* and *ū* before enclitics.

We do not know what the actual stress patterns of Avestan and Old Persian were, although stress must have been responsible for several morphophonological alterations, both in the case of vowels and of consonants. Vowel quantities often change when words receive an enclitic particle or in the course of declension or conjugation. Enclisis also

seems to have affected the palatalisation and labialisation of consonants and the use of anaptyxis, as well as the distribution of alternate consonants.

Epenthesis appears to be omitted in words with enclitics: OAv. *ən-ə'ti-* (< *anī'ti-*) vs. *x^v-īti-cā*; *buud'ntī* 'they shall become' vs. *buuanti-cā*; *ma'ni-madi-cā* 'and we think' vs. *vā'rōi-ma'dī* 'may we *classify'; YAv. *frādati-ca* 'and he furthers', *vīsati-ca* vs. *vīsati* 'twenty' (see de Vaan 2003: Chapter, 7).

Absence or presence of anaptyxis in Old Avestan may depend on stress patterns: *ušə'rū* (< *ušu'ru-*) vs. *ušuruiē*; *d'riiamnā* vs. *d'riiam^mnas-cā*; *ə^rž-jīš* vs. *ə^rž^ə-jiiōi*.

The (pre)aspirated/unvoiced allophone of *r* before *p*, *k*, *t* (^h*r**p*, ^h*r**k*, ^h*r**t* > *š*) is apparently limited to syllables which bore the stress in proto-Avestan: acc. *kəhrpəm* 'body' vs. *hukə^rpta-* 'having a good body'; *mahrka-* 'destruction' vs. *amə^rxti-* 'absence of destruction'; *aša-* 'Order' vs. *astuuə.ə^rta-* 'he through whom Order will have bones', acc. *ašim* 'reward' vs. *ā^rti^mca* < *artim-ca* (P. 39) beside *ašim-ca*; *bāšar-* 'rider' < *bar-tar-* vs. *bar^rtar-* 'carrier' (corresponding to *bara-* mid. 'ride' vs. act. 'carry'). For details, see de Vaan 2003, §29.

2.3.5 Metro-phonology

Certain (morpho-)phonological phenomena expected from the history of the language are illuminated by the Avestan meters.

The Old Avestan meters are syllable-counting and based on rhythmical units of (more or less) identical numbers of syllables. Stanzas contain three to four rhythmical units ('lines') composed of two smaller units ('half-lines') divided by a cesura.

The Young Avestan meter is based on regular rhythmical units of eight syllables. Next in frequency are probably units of seven and nine syllables. Set formulas tend *not* to be adjusted to the meter and frequently causes the number of syllables to differ from the standard eight (see Lazard 1984, 1990, 2002).

The principal phonological features revealed by the Old Avestan meter are disyllabic long vowels and diphthongs and 'Siever's Law' (section 2.3.5.2). In the Young Avestan octosyllabic meter, contracted vowels and Siever's Law apparently provided the poets with flexibility of syllable count, and often, apparently, they scanned these words according to the needs of the meter. Whether they actually did this cannot, of course, be verified.

2.3.5.1 Disyllabic long vowels and diphthongs

Disyllabic long vowels and diphthongs are found in laryngeal-stem nouns and verbs, in compounds, and between preverb and verb, as well as in certain morphemes which may or may not have contained laryngeals. Examples:

Nouns: neut. *h*-stem *dā* = *daō* 'gift'; *Han*-stem *maqōrā* = *maqōra^a* 'keeper of the *maqōra*' (< *manōra-Han-*, but obl. *maqōrān-* < *manōra-Hn-*); gen. plur. *-qm* (*-əm*) (always disyllabic); *-ārmāti-* 'humility' has a disyllabic initial *ā-* (cf. OInd. *arāmati-*).

Verbs: subj.: *dā'tī* = *da'dtī* 'he shall give/place' (< *daH-ati*; but *dada'tī* 'they give/place' < *da-dH-ati* < *da-dH-nti*).

Compounds: *f³rašaōstrāi* = *fraša'uštrāi*, *vīštāspa-* = *vīšta'aspa-*, *d³jāmāspa-* = *djāmā'aspa-*, all proper names, but *spitāma-* 'Spitamid' < *spita-Hma-* 'having fattened strength (*Hama-*)', *darəgāiiu-* < *darga-Hīu-* 'giving a long life span (*āiiu-*)'.

Between preverb and verb: *āiiaṭ* = *ā-aiṭ* ‘he shall come’, *āitē* = *ā-itē* ‘to come’, *āiōi* = *ā-iōi* ‘I request’.

Optional disyllabic scansion: the thematic dat. sing. ending *-āi* is most often disyllabic; – the thematic subj. endings 1s *-ā*, *āi*, 3s *-āi*, 3p *-an* are mono- or disyllabic.

2.3.5.2 Siever’s law

In Ir. *i* and *u* after consonant were realised as *i* and *u* or *ij* and *uj* according to whether the preceding syllable was light or heavy: if *i* and *u* were preceded by short vowel plus one consonant (light syllable), the phonetic realisation was *i* and *u*; if *i* and *u* were preceded by a syllable with a long vowel or diphthong plus one or more consonants or by short vowel plus more than one consonant (heavy syllable), it was *ij* and *uj*.

In Iranian, stops before consonantal [j] and [ɰ] were then spirantised, but remained before [ij] and [uj]. This is still the situation in Old Avestan:

- after light syllable, disyllabic: *ufiia-* ‘weave’ < *uf-ia-*, *mər’ōiiu-* ‘death’ < *mṛt-iu-*;
- after heavy syllable, trisyllabic: *vaēpiia-* ‘a *trembler’ < *uap-iia-*, *mašiiia-* ‘mortal man’ (< *mart-iiia-*).

The endings in *-du-*, however, apparently do not cause Siever’s Law: OAv. *vīduiē* ‘(in order) to know’ (< *viduai*, light) *mər’ngduiē* ‘(in order) to be destroyed’ (< *mṛng-duai*, heavy; there are no examples of the 2p ending *-dūm* < *-duam* after consonant in metrically unambiguous positions).

In Young Avestan, there are many exceptions: suffix *-tuua-*, *-ōβa-* (see section 4.6.2): *maōβa-* ‘that ought to be thought’ < *man-ōua* and *frāiias’ōβa-* ‘that ought to be sent forth in sacrifice’ (for *’ias-tuua-*); – suffix *-uant-* (see section 9.1.1.1): *āfənt-* < *āfuiant-* ‘rich in water’ (for *āp-uuant-*); – *tu-* stems: gen. plur. *yāōβam* < *yātu-* ‘sorcerer’ (for *iāt-uua’ am-*).

2.3.6 Phonological changes associated with inflection and endings

(Morpho)phonological changes associated with inflection and nominal and verbal endings affect vowels and consonants, among them the following.

Important vowel changes include *a*, *ā* > *e* (see section 2.3.1.3); *-an-* > *-ən* and *-in-* and *-aijan-* > *-aēn*, *-aijājan-* > *-aijāēn*, and *-auan-* > *-aon* (see section 2.3.1.4).

Consonant changes are found in stems (caused by the IE.-Irr. ablaut) and include spirantisation (see section 2.3.2.2); palatalisation (see section 1.2.1.2); dental assibilation (see section 1.3); assimilation (see sections 1.3, 2.1.2.2., 2.3.2.1); voice dissimilation (see section 2.3.2.1); and the modifications of Irr. *é* and *í* and Irr. *s(z)*.

Spirantisation of voiceless stops (*plb*, *klg*); nom. sing. *āf-š* < *ap-* ‘water’, *vāx-š* < *uak-* ‘word’ *druux-š* < *drug-* ‘(cosmic) Lie’; loc. plur. *af-šu* < *ap-*; aor. 1p *ioux-madi*, past part. *iux-ta-* < *ioug-* ‘harness’; weak perf. stem *ča-xr-* < *kar-* ‘do’, *čā-xn-* < *kan-* ‘love’; *ci-cit-* < *kait-* ‘distinguish’.

Palatalisation of velars before Irr. *e*, *i*; acc. sing. *uāc-am* < *uak-*; pres. *jan-lgn-* ‘strike down’, aor. *jam-*, past part. *gmata-* < *gam-* ‘come’.

Dental assibilation: pres. act. 3s *das-tai* < *da-d-t°*, *daz-dai* < *da-d^h-t°*, imp. act. 2s *daz-di*, mid. 2p *daz-duam* < *da-d-d°* < *dā-* ‘give’, *d^hā-* ‘place’; inj. 3s *nāis-t*, imp. 2p *nīs-ta* < *naid-* ‘scorn’ (but *nāis-mi* analogical for **nāin-mi*); past. part. *uis-ta-* < *uaid-* ‘find’, *bas-ta-* (analogical for **baz-da-*), *band^h-* ‘bind’.

Assimilation of voiceless stops before (IIr. aspirated) voiced stops: instr. plur. *abiš* < *ap-biš*; *azd-biš* < *ast-* ‘bone’; – before *ž*: *aug-ža* < *aug-*.

Assimilation of sibilants: nom. sing. *haruātās-* < *haruātāt-* ‘wholeness’; pres. 2s *vaši* < *vač-ši* (*vaš-ši*) < *vač-* ‘wish’; imp. mid. 2s *frač-šua* < *frašua* < *parč-lfrač-* ‘ask’.

Voice dissimilation of *xθ* > *xđ* is found, e.g. in *tar*-stems: *f-θr-ai* > *fādrōi* < *ptar-* ‘father’; perfect 2s: *ua-uax-θa* > *vauuax’đa* < *wak-* ‘speak’; and in the suffixes *-θa-*, *-θra-*: *vax’đra-* ‘speech organ’.

Modifications of IIr. *é* and *j*: plur. acc. *ašn-ah* < *asan-* ‘stone, heaven’; pres. 3s *vaš-ti*, 2p *uš-ta* < *vač-*; OAv. *ər’š* ‘straight’ (adv.) < *Hrj’š* (cf. *ər’zu-* ‘straight’ adj.); sing. instr. *baršn-ā* < *barj’han-* ‘height’.

Modifications of IIr. *s* (*z*): nom. sing. *vāx-š* < *uak-*, loc. plur. *af-šu* < *ap-*; loc. plur. *anj’ahu* < *anj’ah-hu* ‘in constrictions’; pres. 1s *ah-mi*, 2s *ahi*, 3s *as-ti*, 3p *h-anti*, imp. 2s *z-di* < *ah-* ‘be’; pres. inj. 2s *bar-a-h* < ‘you carry’, opt. 2s *bar-ai-š*; imp. mid. 2s *-šua*: pres. Av. *yāsa-ŋ’ha* < *yā-sa-* ‘request’, *dasua* < *da-d-sua* < *dā-* ‘give’; aor. *kṛ-šua*, pres. *kṛ-nu-šua* < *kar-* ‘do’; *s-aor.* inj. mid. 1s *maŋ-h-i*, 3s *man-s-ta* < *man-* ‘think’; imp. 2p *Orā-z-dyam*; 3s act. 3p *stā-h-at* < *stā-* ‘stand’; subj. 3s *nai-š-a-t* < *nai-* ‘lead’; perf. stem *hu-šuf-* < *hupa-* ‘sleep’.

3 MORPHOLOGY I: NOMINAL MORPHOLOGY

The morphological categories of nouns, verbs, etc. are the same in all three Old Iranian languages and are largely identical with the inherited Indo-Iranian system. They are much less well known than the Old Indic system, however, because of the limited material.

There are some modifications: in Young Avestan and Old Persian, the abl. sing. is marked in all declensions. In Old Persian, the genitive merged with the dative and the instrumental with the ablative into two cases: gen.-dat. and instr.-abl. In Young Avestan and Old Persian, a ‘preterital optative’ developed.

Note also that Old Persian shares morphological (also lexical) isoglosses with Middle and Modern Iranian languages remaining in Central Asia, so-called ‘Randsprachen’ phenomena.

Typically, the morphological categories are defined by the type and morphology of the inflection and derivation. Both nominal and verbal stems may consist of a root or root + formant, to which inflectional endings are attached: R-(f)-E.

Both nominal declension and verbal conjugation are characterised by complex patterns of quantitative ablaut in the so-called athematic classes, affecting the root, and/or the formant, and/or the ending, as opposed to the absence of ablaut in the nominal and verbal so-called thematic classes, both marked by a thematic vowel *-a* (< ablauting IE. *-o/-e*). The various ablaut patterns result in the extremely complex inflectional morphophonology described above and below.

It is useful to classify nominal and verbal stems as strong and weak (occasionally also middle). The strong stem, if possible, takes the lengthened grade of the root or stem formant, the weak one the full or zero grade. Strong cases are the nom., acc. sing., nom.-voc.-acc. dual, and the nom. plur.

Pronouns have several endings not found in nouns.

3.1 Nouns

Iranian, on the whole, maintains the inherited the system of vocalic and consonantic declensions in nouns and adjectives; the triple gender (masc., fem., neut.) and number (sing., dual, plur.) systems; and the eight cases.

3.1.1 Gender

The distribution of the genders in nouns, adjectives, pronouns, nominal forms of the verb, and numerals is that of Indo-Iranian, with some individual Iranian features, e.g. *vak-ivac-* is feminine in Old Indic (Latin *vox* fem.), but masculine in Avestan.

A few words have both feminine and neuter forms (not in complementary distribution), e.g.:

- fem.-neut.: YAv. *zaoθrā-* ‘libation’; *pəṣānā-* ‘battle’; OP. *^hunarā-* ‘skill’;
- YAv. neut. *a*-stems frequently have fem. plur., e.g. sing. *nmānəm* ‘house’, plur. *nmānā*;
- masc.-fem.: OAv. *mąθrā-* ‘poem’ (fem. *Yasna* 43.14).

See also section 3.1.3.1 Derived feminines.

3.1.2 Number and case

There are three numbers: singular, dual and plural. The few dual forms found suffice to show that this category was of the old Indo-Iranian type, but with distinct genitive and locative forms in Old Avestan. See Tables 3.3.1–2

Proto-Iranian had eight cases: nominative, vocative, accusative, genitive, dative, ablative, instrumental, and locative. All are preserved in Avestan, while in Old Persian they were reduced to six, the genitive being also used for the dative and the ablative having merged with the instrumental.

Case syncretism is common: vocative = nominative in dual and plural; genitive = ablative except in the *a*-stems (in YAv. and, partly, OPers., the abl. was distinguished from the gen.); nominative = accusative in the feminine plural; nominative = accusative (no voc.) in the neuter singular, dual, and plural; dative = ablative in the plural; dative = ablative = instrumental in the dual; Old Persian (and Young Avestan?), genitive = locative in the dual.

3.1.3 Stem classes and declensions

Synchronically, stems can be classified as vowel stems (ending in *a*, *ā*, *i*, *ī*, *u*, *ū*, *ai*, and *au*) and consonant stems (most commonly ending in *n*, *r*, *h*, but also *p*, *t*, *nt*, *d*, etc.). The main difference is between *a*-stems and all the others: only *a*-stems had distinct forms for the gen. and abl. sing. in Indo-Iranian and proto-Iranian (elsewhere, gen. = abl.) and a gen. sing. not ending in *-hl-š*. Another useful classification is declensions with and without *-hl-šl-s* (*-S* in the following) in the nom. sing.

Diachronically, several of the ‘vowel’ stems are laryngeal stems, with *ā* < *ā* or *aH*, *ī* < *iH*, and *ū* < *uH* and are still declined as consonant stems. The feminine *ī*-stems fall into two groups, the so-called *vṛkī*-type, with the invariant formant *-i-H-* (> *-īC*, *-iḡV*), and the so-called *devī*-type, with ablauting formant *-ī-l-iḡ-* < *-iH-l-iaH-*. The feminine

\bar{a} -declension apparently acquired several features from this type, including the voc. sing. in *-ai* and the element *-jā-*.

Iranian declensions are on the whole inherited from Indo-Iranian, sharing both regular paradigms and archaic singularities with Old Indic, from which they differ mainly because of the phonological developments described above.

Special Iranian features include the protero-, hystero-, and holo-kinetic declensions (see section 3.1.1.6b).

The principal innovations are the extension of the abl. sing. ending *-t* in Young Avestan and Old Persian and the syncretism of cases in Old Persian.

3.1.3.1 Derived feminine adjectives and nouns

Feminine forms of most *a*-stem adjectives are declined according to the \bar{a} -declension: Av. *sūr-a-* ‘rich in life-giving strength’, fem. *sūr-ā-*.

Some *a*-stem adjectives, notably those denoting material, have ‘*vṛkī*-type’ feminine forms: YAv. *maniiauu-a-* ‘belonging to the world of thought’, fem. *maniiauu-ī-*; *zarānaēn-a-* ‘of gold’, fem. *zarānaēn-ī-*, OPers. *aθaʹgain-a-* ‘of stone’, fem. *aθaʹgain-ī-*; female patronymics in Av. *-fʹδr-ī-* ‘daughter of’.

The *u*-stems and consonant stems have ‘*devī*-type’ feminines with zero grade of ablauting suffixes: YAv. *pour-u-* ‘much’, fem. *paoʹr-ī-* (<*par-u- parū-ī-*). In this type, the final *-ī-* probably palatalised a preceding *k > č* and further to *š*: *apaš-ī* ‘backwards’ < **apāč-ī* < *apānk-* (only example).

Non-ablauting suffixes: present participles of thematic verbs in *-aʹnt-ī-*, *-ənt-ī-*: *van-aʹnt-ī-* ‘winning’, *bar-ənt-ī-* ‘carrying’; comparatives in *-iiah-*, fem. *-iieh-ī-*: *maz-iiah-* ‘greater’, fem. *maz-iieh-ī-*;

ablauting suffixes: adjectives and present participles of athematic verbs in *-ant-*, fem. *-aʹt-ī-*: *bəʹz-ant-* ‘high’, fem. *bəʹz-aʹtī-*; perf. part. *-uuah-*, fem. *-uš-ī-*: *dad-uuah-* ‘creator’, fem. *dad-uš-ī-*; agent nouns in *-tar-*, fem. *-Ōr-ī-*: *barʹ-tar-* ‘carrier’, fem. *barʹ-Ōr-ī-* ‘womb’.

Fem. (and neut.) forms of *i*-stem adjectives are declined as *i*-stem nouns: masc., fem., neut. *āhuʹr-i-* ‘belonging to Ahura Mazdā’.

Neut. adjectives are declined like neut. nouns: nom.-acc. sing. Av. *sūr-əm*, *āhuʹr-i*, *voh-u* ‘good’, OPers. *fraš-am* ‘perfect’, *par-uv* ‘much’.

Pronouns and the numerals ‘3’ and ‘4’ have some special masc. and fem. morphemes (see section 3.6.1).

3.1.4 Protero-, hystero- and holo-kinetic declensions

In these declensions, full and zero grades in formants and endings alternate. Proterokinetic forms have stem formant in the full grade and the ending in the zero grade; hystero-kinetic forms have stem formant in the zero grade and ending in the full grade; and holo-kinetic forms have two full grades or two zero grades.

Avestan has a greater incidence than Old Indic of the proterokinetic as opposed to hystero-kinetic, including from *n*-stems (*r/n*-stems), e.g. hystero-kinetic gen. sing. *-i-ah*, *-u-ah*, *-n-ah*, but proterokinetic *-ai-š*, *-au-š*, *-aŋ-h*.

3.1.5 Case endings

The basic case endings as reconstructed for proto-Iranian are set out in Tables 3.3.1–2. Note, again, the convention *-S = -h/-š/-s*. Laryngeals are not always noted.

For actual Avestan and Old Persian endings, see on vowel and consonant changes (see sections 2.3.1–2). Note in particular assimilations between final consonants of stems and initial *S*- and *b*- of endings.

Note also that the new YAv. abl. forms are formed by replacing the proto-Av. gen. ending *-S* by *-t*: *gāri-* ‘mountain’: Ir. gen. *garai-š* → *garai-t* > *garōi-t*; *nar-* ‘man’: Ir. gen. *nr̥-š* → *nr̥-t* > *nər̥-t*; *bar²sman-* ‘barsom’: Ir. gen. *barsman-h* → abl. *barsman-t* > *bar²smən* (the ending reappears before the postposition *-a*: *vaēsmənd-a* ‘up to the entrance hall’).

TABLE 3.3.1: CASE ENDINGS: CONSONANT STEMS

	Basic	Consonant stems
Sing.		
V	-∅	-∅
N	-S, -∅	-hl-sl-š, -∅
A	-m, -am	-am
NAn	-am, -∅	-∅
G	-ah, -S	-ah, -hl-š
Ab	-t	-atl-t
D	-ai	-ai
I	-ā, -H	-ā
L	-i, -∅	-i
Plur.		
NV	-ah	-ah
A	-ah, -nS	
NAn	-i, -H, -∅	-i, -H, -∅
G	-ā ² ām	-ā ² ām
DAb	-b ² jah	-b ² jah
I	-biš	-biš
L	-Su	-hul-sul-š ^u
Dual		
NAV	-ā	-ā
NAn	-ai, -ī	-ai
G	-āh	-āh
DIAb	-b ² jā	-b ² jā
L	-ah	-ah

3.1.5.1 The zero ending

No ending is typical of several cases and declensions, sometimes accompanied by ablaut of the stem formant. Among the noteworthy instances are the following:

- nom. sing. masc. *h*-, *r*-, *n*-, and *ai*-stems, which also drop the final consonant: YAv. *na're.man-ā* (< *-āh*) ‘having manly thoughts’, OAv. *p-tā* < *p-tar-* ‘father’, *a'riiā-mā* < *a'riia-man-* ‘A'riiāman’, *kauu-ā* ‘kauui, poet’;
- nom.-acc. sing. neut. *n*- and *r*-stems have zero grade of the formant: *nā-mā* ‘name’ < *nā-mn̥*, *aii-ar²* ‘day’ < *ai-r*;
- nom.-acc. plur. neut. *h*-, *n*-, and *r*-stems have lengthened grade: YAv. *nā-mān* ‘names’, OAv. *aii-ār²* (YAv. *aii-ān*), *vaxii-ā* < *uahi-āh* ‘better things’;
- loc. sing. neut. *n/m*-stems also have (apparently) lengthened grade beside *-i*: OAv. *caš-mān* ‘in the eye’ (also *caš-mān-ī*), OYAv. *dām* ‘in the house’ (YAv. also *dām-i*).

TABLE 3.3.2: CASE ENDINGS: VOWEL STEMS

Vowel stems	<i>a</i> -stems	<i>ā</i> -stems	<i>ī</i> -stems	<i>i-lai</i> -stems	<i>u-lau</i> -stems
Sing.					
V	- <i>a</i> -∅	- <i>ai</i> -∅	- <i>i</i> -∅	- <i>ai</i> -∅	- <i>au</i> -∅
N	- <i>a</i> - <i>h</i>	- <i>ā</i> -∅	- <i>ī</i> -∅	- <i>i</i> -š, - <i>ā</i> -∅	- <i>u</i> -š, - <i>āu</i> -š
A	- <i>a</i> - <i>m</i>	- <i>ā</i> - <i>m</i>	- <i>īm</i>	- <i>i</i> - <i>m</i> , - <i>āi</i> - <i>am</i>	- <i>u</i> - <i>m</i> , - <i>āu</i> - <i>am</i>
NAn	- <i>a</i> - <i>m</i>			- <i>i</i> -∅	- <i>u</i> -∅
G	- <i>a</i> - <i>hja</i>	- <i>ā</i> - <i>iā</i> - <i>h</i>	- <i>iā</i> <i>h</i>	- <i>ai</i> -š, - <i>i</i> - <i>ah</i>	- <i>au</i> -š, - <i>u</i> - <i>ah</i>
Ab	- <i>āt</i>	[- <i>ā</i> - <i>iā</i> - <i>t</i>	- <i>iāt</i>	- <i>ai</i> - <i>t</i>	- <i>au</i> - <i>t</i> , - <i>u</i> - <i>at</i>]
D	- <i>āi</i> (- <i>a</i> ' <i>ai</i>)	- <i>āi</i> (- <i>ā</i> - <i>iā</i> - <i>i</i>)	- <i>iāi</i>	- <i>ai</i> - <i>ai</i> , - <i>i</i> - <i>ai</i>	- <i>au</i> - <i>ai</i> , - <i>u</i> - <i>ai</i>
I	- <i>ā</i>	- <i>ā</i> (- <i>a</i> <i>i</i> - <i>ā</i>)	- <i>ī</i>	- <i>ī</i>	- <i>ū</i> , - <i>uā</i>
L	- <i>a</i> - <i>i</i>	- <i>ā</i> - <i>i</i> - <i>ā</i>		- <i>ā</i> -∅	- <i>au</i> , -∅
Plur.					
NV	- <i>ā</i> (- <i>ā</i> <i>hah</i>)	- <i>ā</i> - <i>h</i>	- <i>ī</i> š	- <i>ai</i> - <i>ah</i> , - <i>āi</i> - <i>ah</i>	- <i>au</i> - <i>ah</i> , - <i>āu</i> - <i>ah</i>
A	- <i>a</i> - <i>ṇh</i>	= N	= N	- <i>i</i> - <i>n</i> š	- <i>u</i> - <i>n</i> š
NAn	- <i>ā</i>			- <i>ī</i>	- <i>ū</i>
G	- <i>ān</i> - <i>a</i> ā <i>m</i>	- <i>ān</i> - <i>a</i> ā <i>m</i>	- <i>īn</i> - <i>a</i> ā <i>m</i>	- <i>i</i> - <i>n</i> - <i>a</i> ā <i>m</i>	- <i>u</i> - <i>n</i> - <i>a</i> ā <i>m</i> , - <i>u</i> - <i>a</i> ā <i>m</i>
DAb	- <i>ai</i> - <i>biā</i> <i>h</i>	- <i>ā</i> - <i>biā</i> <i>h</i>	- <i>ī</i> - <i>biā</i> <i>h</i>	- <i>i</i> - <i>biā</i> <i>h</i>	- <i>u</i> - <i>biā</i> <i>h</i>
I	- <i>āi</i> š	- <i>ā</i> - <i>bi</i> š	- <i>ī</i> - <i>bi</i> š	- <i>i</i> - <i>bi</i> š	- <i>u</i> - <i>bi</i> š
L	- <i>ai</i> -š <i>u</i>	- <i>ā</i> - <i>ṇu</i>	- <i>ī</i> -š <i>u</i>	- <i>i</i> -š <i>u</i>	- <i>u</i> -š <i>u</i>
Dual					
NAV	- <i>ā</i>	- <i>ai</i>	- <i>ī</i>	- <i>ī</i>	- <i>ī</i>
NAn	- <i>ai</i>				
G	- <i>ai</i> - <i>ā</i> <i>h</i>	- <i>ai</i> - <i>ā</i> <i>h</i>		- <i>i</i> - <i>ā</i> <i>h</i>	- <i>u</i> - <i>ā</i> <i>h</i>
DIAb	- <i>ā</i> <i>biā</i>	- <i>ā</i> <i>biā</i>	- <i>ī</i> - <i>biā</i>	- <i>ī</i> - <i>biā</i>	- <i>u</i> - <i>biā</i>
L	- <i>ai</i> - <i>ah</i>				- <i>u</i> - <i>ah</i>

Note that the endless instr. sing. and the nom.-acc. plur. neut. probably had an original final laryngeal *-H*.

3.1.5.2 Endings in *b*-

Several endings have an original initial *b*- (see Tables 3.3.1–2; pronominal forms, see Table 3.3.5). In Young Avestan, the forms with *b* are common in all positions, but must be due to analogy, as the forms with *-b*- are phonetically expected only after consonant (-*ṇbi*š, -*i*.*bi*š, etc.). After vowels, *b* became *β* > *u* (*uu*), which is still frequently preserved (see Skjærvø 2007a): dat.-abl. plur. *xštəuui-βiiō* (<*xštəuui*-, a mythological people), *vōiyṇā-ūiiō* < *vōiyṇā*- (a kind of natural disaster, flood?), *rasma-oiō* (<*rasman*- ‘battle line’); note also the pronominal dat. sing./plurs. *māuuōiia* (<*maβiā* ‘for me’), etc.; – dat.-abl.-instr. dual *bāzu-βe* < *-biā* < *bāzu*- ‘arm’; – instr. plur. *-βi*š > *-u*iš (**-aoi*š): *vaṇuhi*š < *vaṇhu*- ‘good’. The forms *vāγž’l’biiō*, instr. *vāγžibi*š < *vak*- ‘word’ must have replaced forms such *vāγ-βiō* and **vāγ-βi*š > **vaoiiō* and **vauui*š.

The *n*-stem forms such as *nāmēni*š < *nāman*- ‘name’, *ašaonī*š < *ašauuan*- ‘Orderly’ I believe are analogical replacements for *nāməu*iš, *ašaəu*iš (see Skjærvø 2007a).

The expected forms of the *ah*-stems, **-az*-*bi*š and **-az*-*biā**h*, were replaced by *-ēbi*š, *-ēbiiō*, as if *-ah* + *bi*š.

3.1.5.3 Added -ǎ

In Young Avestan, a final -ǎ can be added to the acc. sing. (*vārəm-a* ‘*according to pleasure’) and the abl. sing. (-ǎδ-*a*, -*mənd-a*), with the specific meaning ‘all the way to, up to and including, throughout’ (see section 5.4.6.3). Old Avestan also has dat. sing. -*āi* and -*āi.ā*.

In the loc. sing./plur. we find the particle in all three languages: OAv. (only sing.) *xʷā0rōii-ā* ‘in good breathing space’; YAv. *gātauu-a*, OP. *gā0av-ā* ‘in place’; – YAv. *gātušuu-a*, OP. *dahyušuv-ā* ‘among the lands’.

This particle may be identical with OAv. *ā*, used to emphasise ‘here and now/there and then’(?), which is combined with a variety of local (temporal) cases.

3.1.6 Individual declensions

Following are some notes on special features of individual declensions. For further details, see Hoffmann and Forssman 1996, Skjærvø 2007b.

3.1.6.1 The *i*- and *u*-stems

Most *i*- and *u*-stems have ablauting (proterokinetic) stem formants -*i-l-ai*- and -*u-l-au*-, except a small set with hysterokinetic ablaut -*i-l-ia*- and -*u-l-ua*- (see section 3.1.1.4, Table 3.3.3).

The forms are distributed somewhat differently in Old and Young Avestan; thus, some masc. *u*-stems have gen. sing. from -*aus* in Old Avestan, but from -*uah* in Young Avestan: OAv. *xrat-əuš*, *pas-əuš*; YAv. *xra0β-ō*, *pas-uuō* < *xratu*- ‘wisdom’, *pasu*- ‘sheep’; OAv. instr. sing. *xratū*, OYAv. *xra0β-ǎ*. The masc. *pasu*- ‘sheep’ also has the irregular nom.-acc. plur. *pas-uuō* (OInd. acc. plur. *paśvas*).

On various aspects of *u*-stems see also Pirart 1993; Tremblay 1998; de Vaan 2003, §§16.3.1–2; Skjærvø 2005a.

A special hysterokinetic *i*-stem is *raǐ*- ‘wealth’ < *raHi*-, which has the two stems *raē*- < *raHi*- and *rāii*- (shortened *rāii*-) < *raHǐ*-.

3.1.6.2 Monosyllabic and polysyllabic *ai*- and *au*-stems

Monosyllabic *au*-stems include the well-attested *gau*- ‘cow’ and *dǐau*- ‘heaven’ (only gen. sing. *dǐiaoš*); *nau*- ‘ship’ is absent from the OIr. corpus; *gau*- is inflected as consonant stem, with standard strong and weak case forms, but acc. sing. *gqm*, plur. *gā* (OInd. *gām*, *gās*).

Old Iranian has a few polysyllabic *ai*- and *au*-stems (traditionally classified as subsets of *i*- and *u*-stems), which take the long grade of the stem formant in strong cases and proter- or hysterokinetic forms in the weak cases (the distribution of lengthened and full grade in Avestan is obscured by shortening and lengthening of *a* and *ā*), among them the proterokinetic *hax-ai*- ‘companion’ with strong stem < *sakH-ai*-, weak stem *haš*- < *hač-ǐ*- < *sakH-ǐ*-, and *kau-ai*- ‘poet’; and the hysterokinetic *dahi-au*- ‘land’. See Table 3.3.3 (only attested forms).

Old Avestan also has *hi0-au*- ‘*cord-master’ with nom. sing. *hi0-āu-š*, acc. sing. *hi0-ctm*, and proterokinetic YAv. abl. sing. *hi0-uua-tt* (?).

TABLE 3.3.3: *ai-* AND *au-* STEMS, PROTEROKINETIC *i-* AND *u-* STEMS

	<i>ai-</i> stems endings	<i>kau-ai-</i>	<i>hax-ai-</i>	<i>au-</i> stems endings	<i>dahj-ai-</i> Avestan	Old Persian
Sing. V						
N	-ā	<i>kauuā</i>	<i>haxā</i>	-āu-š	<i>dajhuš</i>	<i>dahayāuš</i>
A	-āi-am	<i>kauuaēm</i>	<i>haxāim</i>	-āu-am	<i>dajhaom</i>	<i>dahayāum</i> , <i>dahayāvam</i> (° <i>dahayum</i>)
G				-au-š	(OAv. <i>daxiiūm</i>) <i>dajhaoš</i>	<i>dahayauš</i>
Ab				-au-t	<i>dajhaoṭ</i>	
D	-i-ai		<i>haše</i>	-au-ai	<i>dajhaoe</i>	
L				-au	<i>dajhō</i>	<i>dahayauv-ā</i>
Plur. NV	-āi-ah	<i>kāuuaiiō</i>	<i>haxaiiō</i>	-āu-ah		<i>dahayāva</i>
A				-ū-š		<i>dahayāva</i>
G	-i-a'ām	<i>kaoiiaqm</i>	<i>hašqm</i>	-ū-na'ām	<i>daxiiunqm</i>	<i>dahayūnām</i>
L				-u-šu		<i>dahayušuv-ā</i>
Dual NAV				*-ā-ū-ā	<i>dajhu</i>	
	Proterokinetic <i>i-</i> stems endings	<i>pati-, kaui-</i>	<i>ra'i-</i>	Proterokinetic <i>u-</i> stems endings	<i>xratu-, pasu-</i>	
Sing. N			<i>raēš</i>			
A			<i>raēm</i>			
G	-i-ah		<i>rāiiō</i>	-u-ah	<i>xraθβō</i>	
D	-i-ai	<i>pa'oe</i>		-u-ai	<i>xraθβε</i>	
I	-i-ā		<i>raia</i>	-u-ā	<i>xraθβā</i>	
Plur. NA			<i>rāiiō</i>	-u-ah	<i>pasuuō</i>	
A			<i>raēš</i>			
G	-i-a'ām	<i>kaoiiaqm</i>	<i>raiaqm</i>	-u-a'am	<i>xraθβqm</i>	

3.1.6.3 Holo-kinetic stems

These include the following stems:

- the laryngeal stem *pantā-lpaθ-* 'path' (<*pant-aH-*, *pnt-H-*): nom. sing. *pant-ā* (<*pant-ā-h* <*pant-aH-s*) acc. sing. *pant-qm*, instr. sing. *paθ-ā* (<*pnt-H-ā*);
- a few nouns with stem formant *-Han-*: *marta-Han-* 'accompanied by death(?)': nom. plur. *mar^ot-ān-ō* (<*marta-Hān-as*), gen. sing. *mar^oθ-n-ō* (<*mart-Hn-as*);
- a few neut. *u-* stems: *dā^or-u-* 'tree, wood', *āii-u-* 'life span', *zān-u-* 'knee' (sing. nom.-acc. YAv. *dā^or-u*, *āii-u*, gen. *drao-š*, *yao-š*; dat. *yauu-e/yao-e*, instr. *yauu-a*, loc. *drao^o*; plur. dat.-abl. *žnu-biias^o*; cf. Gk. *doru* 'wood', *gonu* 'knee').

3.1.6.4 *r-* stems

The masc. *r-* stems have nom. sing. with lengthened grade of the formant and zero ending. The *r-* stems comprise the following (relatively few forms attested):

- words denoting kinship: OAv. *p^(tu)tar-*, YAv. *pitar-*, *mātar-*, OAv. *dug³dar-/YAv. duγ³dar-*, *x³aγhar-*, *brātar-*, *naptar-* ‘grandson’, as well as *nar-* ‘man’; these have full grade *-ar-* in strong cases and zero grade *-r/ṛ-* in weak cases;
- agent nouns in *-tar-*: *pātar-* ‘protector’, *dātar-* ‘maker, creator’, as well as *star-* ‘star’ have lengthened grade *-ār-* in strong cases.

Both types behave partly like vowel and partly like consonant stems: nom. sing. *-ā-*; hysterokinetic sing. gen. *-rah*, *-rai* (YAv. *piθrō*, OAv. *f³δrōi*, YAv. *dāθrō*, *dāθre*); acc. plur. *-ah*; – acc. plur. *-ṛnš* (OAv. *nəraqš*); proterokinetic gen., dat. sing. *-ṛš*: OAv. *nə³š*, YAv. *narš*; *sāstarš* < *sāstar-* ‘false teacher’ (weak stem *sāθr-*).

ātar- ‘fire’ was probably originally a neuter *r*-stem, with nom.-acc. sing. **ātṛ*, to which masc. endings were added: nom. *ātṛ-š*, acc. *ātṛ-am* > Av. *ātarš*, *ātrəm* (Hoffmann and Narten 1989: 73 n.126).

3.1.6.5 Stem-formants containing *n*

These comprise the stem formants *-an-*; *-jan-*; *-yan-*, *-man-*; *-ant-*, *-yant-*, *-mant-*; *-ank-*, *-jank-*.

In the zero grade of the formant, the *n* when between consonants is vocalised to *ṇ* > *a*, producing the alternations *-ān-l-an-l-n-l-a-*; *-ānt-l-ant-l-at-*; *-ānk-l-ānc-l-āk-l-āc-* (*-ac-*), *-i³ānk-l-i³ānc-l-īk-l-īc-*.

3.1.6.5a Stems in *-an-*

YAv. f. *xšapan-* ‘night’ and m. *asan-* ‘stone, sky’: sing. nom. *xšapa*, acc. *xšapanəm*, *asānəm*, gen. = plur. acc. *xšafnō*, *ašnō*, plur. loc. *xšapōhuua* (< *xšapahu* + *-a*).

3.1.6.5b Stems in *-jan-*

The masc. and fem. *jan*-stems have nom. sing. in *-jā* > Av. *-je*: m. *fraγrasiiān-* (proper name): sing. nom. *fraγrase*, acc. *fraγrasiiānəm* (only forms); – f. *kaniian-* ‘young woman’: sing. nom. *ka³ne*, acc. *ka³nīnəm*, gen. = plur. nom. *ka³nīnō* (sing. gen. also *kanii³*).

3.1.6.5c Stems in *-yan-*, *-man-*, *-Han-*

Example: *ašāuuān-* m. ‘righteous’: sing. nom. *ašāuuā*, acc. *ašāuuānəm*, sing. gen. = plur. acc. *ašāonō* and *ašāunō* (< *-yan-*), plur; dat.-abl. *ašāuuabiīō* (< *-ṇb-*); *manθraHan-* ‘keeper of the *māθra*’: sing. nom. *māθrā* = *māθra³* (< *manθra-Hā*), dat. *māθrāne* (< *manθra-Hn-ai*).

The neuter stems have sing. loc. and plur. nom.-acc. with lengthened grade and zero ending, beside forms with ending *-i* (see section 3.1.5.1).

Because of the morphophonological vagaries of Ir. *y*, the *yan*-stems are sometimes no longer recognizable as such, e.g. Av. *span-lspa-lsūn-* ‘dog’ (< *éyan-léyṇ-léun-*; *ruθβan-* ‘entrails’ < *ruθyan-*; – YAv. *aθaryan-*, a kind of priest: strong stem *āθrauuān-*, weak stem *aθa³run-*.

The voc. sing. of *yan*-stems ends in *-um* (*-əm*), with the final *-n* apparently assimilated to the preceding *y(ə)*: *ašāum* < *ašāyan* ‘Orderly’, *yum* < *yuyan* ‘youth’, *āθraom* < *āθrayan*, *θrizafəm* < *θri-zaf-yan* ‘three-mouthed’.

3.1.6.5d Stems in -ant-

These differ from the *n*-stems in having nom. sing. in -*S*, but are otherwise similar to the *uan*-stems. The *ant*-stems include adjectives and present and aorist participles (these have no ablaut when from thematic verbs).

The nom. sing. varies: adjectives have YAv. -*ō* and forms from -*anh*, pronouns and present participles -*qs* (see section 2.3.2.6). Acrostatic presents have the weak stem throughout (OAv. *stauuas* < *stáu-nt-s* < *stcu-* ‘praise’).

3.1.6.5e Stems in -uant-, -mant-, -Hant-

Stems in -*uant-*, -*mant-* have sing. nom. and voc. from stems in -*uah-l-mah-*: YAv. sing. nom. *druuā* < -*uāh*, voc. *druuō* < -*uah*, acc. *druuaṇtəm*, gen. *druuatō* < *drug-uant-* ‘possessed by the lie’; *maza-Hānt-* ‘big’: YAv. sing. masc. acc. *mazāṇtəm* (*mazāṇtəm*), neut. nom.-acc. *mazāt* (< *maza-Hnt*).

3.1.6.5f Stems in -Hank-, -iHank-

Most of these are derived from preverbs denoting direction, but also from other words. They have various unusual features: sing. nom. in -*š* with loss of the velar (cf. *ant*-stems); neut. sing. nom.-acc. in -*gʷt*; palatalisation in sing. acc., plur. nom.: sing. nom. *parqš* ‘away < *parān(k)-š*, neut. *parāgʷt* (see section 2.1.2.1); sing. acc. *niiāncim* (< -*cam*) ‘downwards’, plur. nom. *niiāncō*.

3.1.6.6 Stems in -ah-, -iah-, -uah-

Most of the *ah*-stems are neuter (cf. Gk. *genos*, etc.) and masculine only in compounds. They take the full grade throughout, except in the plur. nom.-acc. (*manā* < -*āh* ‘thoughts’) and masc. sing. nom. *na're.manā* ‘with manly thoughts’.

The *iah*-stems are comparatives (see section 3.2) and the *uah*-stems perfect participles (see section 4.6.1.2). They both take the lengthend grade in the strong cases, but only *uah*-stems have zero grade -*uš-*.

3.1.6.7 Neuter *r/n*-stems

The archaic heteroclitic neut. *r/n*-stems have *r*-stem nom.-acc. sing. and *r*- or *n*-stem nom.-acc.plur.: *aiiar* ‘day’, gen. sing. *aiiq* (< -*aṇh*), nom.-acc. plur. OAv. *aiiār*⁵, YAv. *aiiqn*; OAv. *luuar*⁵ ‘sun’, gen. *xʷəng*, YAv. *hū* (disyllabic < *huʷ-ar*, *huʷ-aṇh*). Note OAv. *rāzar-* ‘*straightness’, sing. instr. *rašnā*; Old Persian **vazar-* ‘greatness’ (in *vazar-ka-* ‘great’), sing. instr.-abl. *vašnā* ‘by the greatness (of Ahuramazdā)’.

3.1.6.8 Stems in laryngeals

Indo-Iranian laryngeal stems ending in vowels plus laryngeal became long vowel-stems in the attested corpora: *aH* > *ā*, *iH* > *ī*, *uH* > *ū*.

Before endings in vowels, the lost laryngeal left a hiatus, e.g. *mazdaH-* ‘omniscient’: acc. sing. *mazdaH-am* > OAv. *mazdam* (3 syllables); gen. sing. *mazdaH-as* > *mazda'ah*, OAv. *mazdā* (3 syllables) > YAv. *mazdā* (2 syllables); – OPers. nom.-acc. plur. *aθa'gainīya^h* < -*iHah* ‘of stone’; – *tanuH-* ‘body’: gen. sing. *tanuH-ah*, OAv. *tanuuas*^o. See also on *pantaH* in section 3.1.6.3.

TABLE 3.3.4: EXAMPLES OF STEMS IN STOPS

ap- ‘water’, *ʷak-* ‘voice, word’, *drug-* ‘(the cosmic) deception’, *xšnut-* ‘*satisfaction’, n. *ast-* ‘bone’, *stut-* ‘praise’, *°brt-* ‘carrying, riding’, fem. derivatives in *-tāt-*:

Sing.				
NV	<i>āf-š</i>	<i>ʷāx-š</i>	<i>druḫ-š</i>	<i>xšnu-s</i>
A	<i>āp-am</i>	<i>ʷāč-am</i>	<i>druḫ-am</i>	<i>xšnut-am</i>
NAn				<i>as</i>
G	<i>ap-ah</i>	<i>ʷač-ah</i>	<i>druḫ-ah</i>	<i>stut-ah</i>
Ab	<i>ap-at</i>		<i>druḫ-at</i>	<i>°brt-at</i>
D	<i>ap-ai</i>			
I	<i>ap-ā</i>	<i>ʷač-ā</i>		
L	<i>ap-i</i>			
Plur.				
NV	<i>āp-ah</i>	<i>ʷāč-ah</i>	<i>druḫ-ah</i>	<i>stut-ah</i>
A	<i>ap-ah</i>	<i>ʷač-ah</i>		<i>stut-ah</i>
NAn				<i>ast-i</i>
G	<i>ap-ā'am</i>	<i>ʷač-ā'am</i>	<i>°druḫ-ā'am</i>	<i>stut-ā'am</i>
DAh	<i>a(p)-hijaḥ</i>	<i>*ʷag-bijaḥ</i>		<i>°brd-bijaḥ</i>
I	<i>a(p)-biš</i>	<i>*ʷag-biš</i>		<i>azd-biš</i>
L	<i>af-šu</i>			
Dual				
NAV	<i>āp-ā</i>			<i>°tātā</i>
G				<i>°tād-ātāh</i>
DIAb		<i>*ʷag-bijā</i>		<i>°tad-bijā</i>

3.1.6.9 Other consonant stems

The stems in other stops are fairly regular. Examples are seen in Table 3.3.4 (reconstructed proto-Iranian).

In Old Persian, where final groups were apparently not allowed, some words may have been transferred to other declensions (see Section 2.2.2.1).

The two *m*-stems *zam-* ‘earth’ and *ziam-* ‘winter’ have sing. nom. *zā* and *ziā* and weak stems *z'm-* (monosyllabic, *-sm-* in compounds) and *zim-* (often confused in the mss.); *-dam-* ‘house’ has gen. sing. OAv. *dāng* (<*danḫ* < *dam-s*), YAv. loc. sing. *dam*, *damī*.

3.1.6.10 Suppletive stem-systems

There are various kinds of suppletive stem-systems, e.g.:

- alternating vowel and consonant stems: Av. *zā-/zam-* ‘earth, ground’, *ziā-/ziam-* ‘winter’; OAv. *sauua-/sauuah-* (in complementary distribution) ‘life-giving strength’, *ušā-/ušah-* ‘dawn’, YAv. *kaniīā-/ka'nīn-* ‘young woman’; OPers. *āpī-/āp-(?)* ‘water’;
- alternating consonant stems: *ntls-* stems, *tls-*, and *tn-* (*napah-/napat-/naptar-* ‘grandchild, scion’; *°carat-/°caran-*); neut. *rn-* stems.

3.2 Adjectives: comparative and superlative

The comparative and superlative of adjectives (including participles) and adverbs is made either with the suffixes *-īyah-* and *-išta-* or with the productive type *-tara-* and *-təma-*. The

suppletive type Eng. *good* vs. *better*, is attested in OPers. *vazarka-* ‘great’, *maθ-išta-* ‘greatest’.

The comparative/superlative in *-tara-l-təma-* is made from the (weak) stem of adjectives or adverbs and with appropriate sandhi before the ending: YAv. *ama-uuas-tara-l-təma-* ‘more/most powerful’ (*-uuast-* < *-uñt-t-*); *hudās-təma-* < *hudāh-* ‘giving good gifts’; *yāskər^s-tara-l^otəma-* < *yās-kər^s-t-t-* ‘performing (well) at the audition’; YAv. *tauruuāiḡs-təma-* < *^oiiañt-t-* ‘overcoming, victorious’, YAv. *hubaoⁱḡi-tara-l^otəma-* ‘more/most fragrant’; OPers. *apa-tara-* ‘farther’; YAv., OPers. *fra-tara-* ‘ahead (of)’, Av. *fra-təma-*, OPers. *fra-tama-* ‘foremost’.

The comparative/superlative in *-iih-l-išta-* is made from the root in the full grade. Adjectives with suffixes lose these: Av. *ak-a-*: *ašⁱ-iih-* (< *ašⁱñah-*), *ac-išta-* ‘more/most evil’; *būⁱ-ri-* ‘plentiful’: *baō-iih-*, *dbō-išta-* < *du-*); *driy-u-* ‘poor’: *draēj-išta-*; *spəñ-ta-* ‘life-giving’: *span-iih-*, *spəñ-išta-*; *drug-uant-* ‘possessed by the Lie’: *draoj-iih-*, *draoj-išta-*. Forms from pres. parts. (or verbal nouns) include *bār-išta-* ‘who carries most often’ < *bar-əñt-* ‘carrying’, OPers. *tauv-īyah-* ‘mightier’ < *tu-nu-va^t-* ‘who has power’.

From roots with laryngeals we have forms such as the following: Av. *dar³ya-* ‘long’ (< *d^rH-ga-*): *drāj-iih-*, *drāj-išta-* (< *draHg-*); *ād-ra-* ‘needy’ (< *ñHd-ra-*): *nā^d-iih-* (< *naHd-*); *po^rru-* ‘much’ (< *pr^rH-u-*); *frā-iih-*, *fraēšta-* (< *praH-īah-*, *praHišta-*).

A few adjectives have both kinds of superlative, but with different meanings: OAv. *po^rru-təma-* ‘in highest numbers’; *spəñtō.təma-* ‘having the name *spəñta-* in the highest degree(?)’; YAv. *aka-tara-* ‘worse (for sb.)’.

Double superlative: YAv. *draēj-ištō-təma-* ‘the poorest of the poor’.

3.3 Adverbs

Adverbs can be invariable particles or forms of adjectives (also comparative and superlative) or other words:

- underived: Av. *mošū* ‘soon, quickly’; OAv. *nū*, OYAv. *nūrəm*, OPers. *nūram* ‘now’;
- nom.-acc. neut. sing. of a noun or adjective: OYAv. *vasō* ‘at will’ (< *vasah-* ‘will’); YAv. *daršaṭ*, OPers. *daršam* ‘vigourously’; YAv. *po^rrum* ‘in front’, OPers. *paruvam* ‘before’;
- abl. and loc. sing.: Av. *dūrāṭ*, OPers. *dūradaš* ‘from afar’; Av. *dūrē* ‘in the distance’, OPers. *dūraiṯ* (*apiṯ*) ‘far (and wide)’;
- ending *-S* (Schindler 1987): OAv. *ər^s*, YAv. *arš* ‘straight’ (< *Hr^j-š*);
- compounds: YAv. *yaθā.kər^stəm* ‘as it is done’, OPers. *pasāva* (< *pasā-ava*) ‘afterward’.

Adverbs of place and manner are made from adjectives, pronouns and preverbs with suffixes: *-Ora*, *-Oa*, *-da*, *-dāt* (OPers. *-daš*), and *-tah*, e.g.: Av. *ya-Orā* ‘where’; *ka-θā* ‘how?’; OAv. *a-dā* ‘then’, OPers. *ava-dā* ‘there’; YAv. *a-dāṭ* ‘from there’, OPers. *dūra-daš* ‘from afar’; Av. *aīβi-tō* ‘all around’, OPers. *a^hma-ta^h* ‘from that (place)’.

Comparative and superlative: *bāḡa* ‘*sometimes, *again and again’, sup. *bāⁱḡ-ištəm*; *fra-* ‘forward’: comp. *fratarəm*, sup. *fratəməm*.

See also section 9.1.2.2 Adverbial *āmreḡita* compounds.

3.4 Pronouns

There are personal, possessive, demonstrative, reflexive-reciprocal, relative, interrogative and indefinite (indefinite relative) pronouns, most of them of the Indo-Iranian type.

3.4.1 Personal pronouns

The personal pronouns distinguish three persons; the 3rd person distinguishes three genders. All persons have tonic and non-tonic enclitic forms (see Tables 3.3.5–6).

TABLE 3.3.5: 1ST AND 2ND PERSON PERSONAL PRONOUN

Sing.	1st pers.		2nd pers.	
		encl.		encl.
V			<i>tū</i>	
N	<i>aJam</i>		<i>tūJam</i>	
A	<i>mām</i>	<i>mā</i>	<i>θuām</i>	<i>θuā</i>
G	<i>ma-na</i>	<i>mai</i>	<i>taJa</i>	<i>tai</i>
Ab	<i>ma-t</i>		<i>θu-a-t</i>	
D	<i>ma-bJa(h)</i>	<i>mai</i>	<i>ta-bJah</i>	<i>tai</i>
L			<i>θu(a(i))i</i>	
Plur.				
N	<i>vajam</i>		<i>jūž-am</i>	<i>jūš</i>
A	<i>ahmā</i>	<i>nāh</i>		<i>uāh</i>
G	<i>ahmākam</i>	<i>nah</i>	<i>jūšmākam</i>	<i>uāh</i>
D	<i>ahma-bJa</i>	<i>nah</i>	<i>jūšma-bJa, šma-bJa</i>	<i>uāh</i>
Ab	<i>ahma-t</i>		<i>jūšma-t, šma-t</i>	
I	<i>ahmā</i>		<i>šmā</i>	
Dual				
NAV	<i>āyā(?)</i>			
G			<i>juuākam</i>	

TABLE 3.3.6: 3RD PERSON PRONOUNS

Sing.	Tonic:			Enclitic		
	m.	n.	f.	m.	n.	f.
N	<i>*ha, *ha-h</i>	<i>ta-t</i>	<i>hā</i>	–	<i>it</i>	<i>hī</i>
A	<i>ta-m</i>	<i>ta-t</i>	<i>tām</i>			
G	<i>a-hJa</i>		<i>a-hJ-āh</i>	<i>hailšai</i>	<i>hailšai</i>	<i>hailšai</i>
Ab	<i>a-hm-āt</i>		<i>a-hJ-āt</i>			
D	<i>a-hm-āi</i>		<i>a-hJ-āi</i>	<i>hailšai</i>	<i>hailšai</i>	<i>hailšai</i>
I	<i>a-nā</i>		<i>a-jā</i>			
L	<i>a-hm-i</i>		<i>a-hJā</i>			
Plur.						
N	<i>tai</i>	<i>tā</i>	<i>tā-h</i>			
A	<i>ta-gh</i>	<i>tā</i>	<i>tā-h</i>	<i>inš (dinš)</i>	<i>ī</i>	<i>hīš*šīš</i>
G	<i>ai-š-a'am</i>		<i>ā-h-a'am</i>			
DAb	<i>ai-bJah</i>		<i>ā-bJah</i>			
I	<i>tāiš</i>		<i>ā-biš</i>			
L	<i>ai-šu</i>		<i>ā-hu</i>			
Dual						
NAV	<i>tā</i>	<i>tai</i>	–	–	<i>hī</i>	<i>hī</i>

There are three 3rd-person demonstrative pronouns: *ha-* (*ša-)/ta-*, *i-* (*hi-lši-*, *di-*), and *a-*. Some forms are poorly attested and difficult to classify.

ha-/ta- serves as 3rd-person personal pronoun and weakly deictic ‘that’, but is often difficult to distinguish from the far-deictic demonstrative pronouns; *i-* is only enclitic; and *a-* is identical with the near-deictic demonstrative pronoun.

In Old Persian, the far-deictic pronoun is used as emphatic personal pronoun.

3.4.2 Possessive pronouns (adjectives)

Only Old Avestan has genuine possessive pronouns; these are derived from the oblique stems of the 1st and 2nd sing. personal pronouns by thematisation: *m-* ‘my’, *0βa-* (< *0ua-* > ‘thy’, and from *aka-* extensions of the oblique stems of the 1st and 2nd plur. personal pronouns: *ahmāka-* ‘our’, *xšmāka-* ‘your’. For the 3rd person, the genitive of the personal pronoun is used or the possessive reflexive pronoun *hya-* ‘his, her, its, their’.

In Young Avestan and Old Persian, the genitive of the personal pronoun is used.

3.4.3 Demonstrative pronouns

The demonstrative pronouns have a three-way deixis of varying emphasis, corresponding to 1st, 2nd and 3rd person:

- 1st-person deixis (near-deixis): *ima-la-* and *aēša-laēta-* ‘this’ (Latin *hic*, Spanish *este*);
- 2nd-person deixis (unclear): OAv. *huuōlana-* ‘that near you’, also ‘derogatory’ (cf. Latin *iste*, Spanish *ese*);
- 3rd-person deixis (far-deixis): *hāulaua-* ‘that (over there)’ (Latin *ille*, Spanish *aquel*).

All three have two (or more) stems, one reserved for the nom. masc. and fem., the other for the other cases, or a more complex distribution. See Table 3.3.7.

3.4.4 Relative pronoun

The relative stem is *ja-*, with nom.-acc. neut. OAv. *hūiaŋ* (with *h-* of unclear origin), YAv. *yaŋ* (rarely *hūiaŋ*). In Old Persian, *ja-* was univerbated with the demonstrative pronoun *ha-/ta-*: *haya-ltaya-* (cf. Adiego Lajara 2000). See Table 3.3.8.

3.4.5 Interrogative and indefinite pronouns

The stems *ka-/ca-* and *ci-* serve as interrogative pronouns (see Table 3.3.8) and, with enclitic *-ca* and *-ciŋ*, repeated, or negation, as indefinite pronouns: Av. *kas-ciŋ* ‘whoever’, OPers. *kaš-ciy*, *ciš-ciy* ‘anything’; *kataras-ciŋ* ‘each (of two)’; *ka-ka-* ‘each and every’; – negative and preventive: *nōiŋ kas-ciŋ*, *naē-ci-* ‘not any’, *mā-ci-* ‘let not any’.

Indefinite relative: OAv. *yō . . . ciš-ca* (sing. m.), *yōi . . . caiias-ca* (plur. m.) ‘whoever’, *yā . . . cī-ca* (plur. n.) ‘whichever’, *yahmāi . . . kahmāicīŋ* ‘to whom-so-ever’.

3.4.6 Reflexive pronouns and reciprocity

These are OYAv. *x’a-* and YAv. *hauua-* ‘own’; YAv. *x’ aēpa’0iia-* ‘own’; OPers. *huvaipašiya-* ‘self’, *huvāipašiya-* ‘own’. The meaning of *tanū-* ‘body’ sometimes comes close to ‘own body, self’.

TABLE 3.3.7: DEMONSTRATIVE PRONOUNS

	Near-deictic (1st pers.) dem. pron <i>ima-la-</i> :			Far-deictic (3rd pers.) dem. pron <i>hau/aya-</i> :		
	m.	n.	f.	m.	n.	f.
Sing.						
N	<i>ajam</i>	<i>im-at</i>	<i>ijam</i>	<i>hāu</i>	<i>ay-a-t</i>	<i>hcu, hāu</i>
A	<i>im-am</i>	<i>im-at</i>	<i>im-ām</i>	<i>ay-a-m</i>	<i>ay-a-t</i>	<i>ay-ā-m</i>
G	<i>a-hja</i>		<i>a-h-iāh</i>	<i>ay-a-hja</i>		<i>ay-ah-iāh</i>
Ab	<i>a-hm-āt</i>		<i>a-h-iāt</i>			<i>ay-ah-iāt</i>
D	<i>a-hm-āi</i>		<i>a-h-iāi</i>			<i>ay-ah-iāi</i>
I	<i>a-nā</i>		<i>ajā</i>	<i>ay-ā, ay-a-nā</i>		
L	<i>a-hm-i</i>		<i>a-h-iā</i>			
Plur.						
N	<i>im-ai</i>	<i>im-ā</i>	<i>im-āh</i>	<i>ayai</i>	<i>ay-ā</i>	<i>ay-ā-h</i>
A	<i>im-aṛh</i>	<i>im-ā</i>	<i>im-āh</i>	<i>ay(a-ṛ)h</i>	<i>ay-ā</i>	<i>ayā-h</i>
G	<i>ai-š-a'am</i>		<i>ā-h'a'am</i>	<i>ay-aiš-a'am</i>		<i>ay-ā-bjah</i>
DAb	<i>ai-bjah</i>		<i>ā-bjah</i>			
I	<i>āiš</i>		<i>ā-biš</i>	<i>ay-āiš</i>		
L	<i>ai-šu</i>		<i>ā-hu</i>			
Dual						
NAV	<i>im-ā</i>			<i>ay-ā</i>		
G	<i>āh, ajāh</i>			<i>ay-āh</i>		
DIAb			<i>ābjā</i>			
Far-deictic (2rd pers.) dem. pron <i>hculana-</i>						
m.						
	Sing.	Plur.				
N	<i>hcu</i>					
I	<i>anā(?)</i>	<i>anāiš</i>				

Reciprocity is usually expressed by *aniia-* . . . *aniia-*, OPers. *aniya-* . . . *aniya-* ‘one . . . another, each other’ (see Jamison 1997).

3.4.7 Pronominal case endings

Pronouns have some case endings that differ from those of nouns, including the dat. sing./plur. ending *-bja(h)* in the pers. prons. 1st and 2nd (see also sections 2.3.1.2, 3.1.5.2); – neut. nom.-acc. sing. *-at* in poss. and dem. prons. (and pronominal adjs.); – masc. nom. plur. *-ai*; – masc.-neut. dat., abl. sing. formant *-hm-*; – fem. gen., dat., abl., loc. formant *-hi-*; – gen. plur. masc. formants *-š-*, fem. *-h-*. See Tables 3.3.5–8.

There is at least one OAv. fem. sing. form in *-aē*: *x'aē°* ‘own’ (cf. latin *hae-c*).

3.5 Adpositions

Old Iranian has pre- and postpositions. Some of these are only prepositions, some are also preverbs, and some are derived from nouns, adjectives, or adverbs or particles. In some cases it is difficult to distinguish between adverbs, preverbs and adpositions. For examples, see section 5.4 Uses of Cases.

TABLE 3.3.8: RELATIVE AND INTERROGATIVE PRONOUNS

Relative pronouns:						
	m.	n.	f.	m.	n.	f.
Sing.						
N	<i>ī-ah</i>	<i>ī-a-t</i>	<i>īā</i>	<i>k-ah, čī-š</i>	<i>k-at, čī-t</i>	<i>kā</i>
A	<i>ī-am</i>	<i>ī-a-t</i>	<i>īā-m</i>	<i>k-am, čī-m</i>	<i>k-at, čī-t</i>	<i>kā-m</i>
G	<i>īa-hja</i>		<i>īa-h-īāh</i>	<i>ka-hja, č-ahja</i>		<i>ka-h-īāh</i>
Ab	<i>īa-hm-āt</i>		<i>īa-h-īāt</i>	<i>ka-hm-āt</i>		
D	<i>īa-hm-āi</i>		<i>īa-h-īāi</i>	<i>ka-hm-āi</i>		<i>ka-h-īāi</i>
I	<i>īā</i>			<i>kā, ka-nā</i>		
L	<i>īa-hm-i</i>		<i>īa-h-īā</i>	<i>ka-hm-i, č-ahmi</i>		<i>ka-h-īā</i>
Plur.						
N	<i>ī-ai</i>	<i>īā</i>	<i>īā-h</i>	<i>k-ai, čajah</i>	<i>kā, čī</i>	
A	<i>ī-aṅh</i>	<i>īā</i>	<i>īā-h</i>	<i>k-aṅh</i>	<i>kā</i>	
G	<i>ī-ai-š-a'am</i>		<i>īā-h-a'am</i>	<i>k-ai-š-a'am</i>		<i>kā-h-a'am</i>
DAb	<i>ī-ai-bjah</i>		<i>īā-bjah</i>	<i>k-ai-bjah</i>		
I	<i>īāiš</i>			<i>kāiš</i>		
L	<i>ī-ai-šu</i>		<i>īā-hu</i>			
Dual						
NAV	<i>īā</i>		<i>īai</i>			
G	<i>īa-īāh</i>					
Old Persian						
	m.	n.	f.			
Sing.						
N	<i>haya</i>	<i>taya</i>	<i>hayā</i>			
A	<i>tayam</i>	<i>taya</i>	<i>tayām</i>			
IAb	<i>tayanā</i>					
Plur.						
N	<i>tayaiy</i>		<i>tayā, tayaiy</i>			
A		<i>tayā</i>	<i>tayā</i>			
GD	<i>tayaišām</i>		<i>tayaišām</i>			

3.6 Numerals and measure

A relatively large number of cardinals, ordinals and other number words are attested in Young Avestan, while Old Avestan has hardly any. Of the Old Persian number words, only *aiva* 'one', is spelled out, but several Old Persian number words are found in Elamite texts.

3.6.1 'One' to 'four'

'One' is expressed by the stems *aiya-* and *ha-* < *sṃm-* (see section 3.6.3.1):

- cardinal: Av. *aēuua-*, OPers. *aiva-*;
- ordinal: Av. *fra-təma-* 'foremost, first', OAv. *pa^uruuiia-*, YAv. *pao'riia-*, OPers. *paruiyiya-*.

'Two' and 'both': The number 'two' is expressed by the stems *d(u)ya-* and *dui-* (OAv. *d^ubi-*, YAv. *bi-*):

- cardinal: YAv. *duua*, fem., neut. *duiie*, *duuaē°* (i.e. *dúya*, *dúyai*);
- ordinal: OAv. *d^mbitiia-*, YAv. *bitiia-*, OPers. *duvitīya-*;

For ‘both’, OAv. has nom.-acc. masc. *ubā*, fem., neut. *ubē*; YAv. nom.-acc. masc. *uua*, fem., neut. *uiie* < *úya*, *úyai*); OPers. nom.-acc. masc. *ubā*.

YAv. also has the indecl. *uuaēm* ‘both’ (OInd. *ubhayam*).

‘Three’: The stem is *θri-*:

- cardinal: nom. masc. *Orāiiō*, fem. *tišrō*, neut. *Orī*;
- ordinal: Av. *Oritiia-*, OPers. *çitīya-* ‘for the third time’.

‘Four’: The stem is *čaθβar-/catur-* (YAv. *caθru°* in compounds):

- cardinal: nom. masc. *caθβārō*, fem. *catayrō*, neut. *catura*;
- ordinal: YAv. *tū’riia-* (< *ktrūja-*, cf. *ā-xtū’ rim* ‘a fourth time’).

‘One’ has pronominal declension; ‘two’ and ‘both’ are declined as dual *a*-stems (but OPers. plur. gen.-dat. *ubānām*); ‘three’ and ‘four’ have fem. forms with *-hr-/šr-* (see Table 3.3.9).

TABLE 3.3.9: DECLENSION OF NUMERALS

‘One’				
	m.	n.	f.	
N	<i>aēuuō</i>	<i>ōiium</i> , <i>ōim</i>	<i>aēuua</i>	
A	<i>ōiium</i> , <i>ōim</i>	<i>ōiium</i> , <i>ōim</i>	<i>aēuuqm</i>	
G		<i>aēuuahē</i>	<i>aēuuaijḥā</i>	
Ab		<i>aēuuahmāṭ</i>		
I		<i>aēuua</i>	<i>aēuuaia</i>	
L		<i>aēuuahmi</i>	–	
‘Two’				
	m.	n.	f.	
NA	<i>duua</i>	<i>duiie</i> , <i>duuaē°</i>	<i>duiie</i> , <i>duuaē°</i>	
DAbI		<i>duuaēibita</i>		
L		<i>duuaiiā</i>		
‘Three’				
	m.	n.	f.	
N	<i>Orāiiō</i>	<i>Orī</i>	<i>tišrō</i>	
A	<i>Oriš</i>	<i>Orī</i>	<i>tišrō</i>	
G		<i>Oraiiqm</i>	<i>tišraqm</i> , <i>tišranqm</i>	
DAb		<i>Oribiiō</i>	–	
‘Four’				
	m.	n.	f.	
N	<i>caθβārō</i>	<i>catura</i>	<i>catayrō</i>	
A	<i>caturā</i>	<i>catura</i>	<i>catayrō</i>	
G		<i>caturqm</i>		

3.6.2 The remaining numerals

The cardinals ‘five’ to ‘ten’: *pañca*, *xšuuas*, *hapta*, *ašta*, *nauua*, *dasa*; these have thematic gen. forms (*pañcanqm*, *nauuanqm*, *dasanqm*).

The ordinals (only YAv.) of ‘5’ and ‘6’ are characterised by the root vowel *u*: *puḫda-*, *xštūua-*. ‘Seventh’ has the formant *-0a-*: *hapta-0a-*; ‘eighth’ to ‘tenth’ the formant *-ma-*: *aštā-ma-*, *nao-ma-*, *dasā-ma-*.

Most of the numerals ‘11’ to ‘19’ are known only from the ordinals, which were identical with the cardinals: *aēuuān-dasa-*, *duua-dasa-*, *0ri-dasa-*, *ca0ru-dasa-*, *pañca-dasa-*, *xšuuāš-dasa-*, *hapta-dasa-*, *ašta-dasa-*, *nauua-dasa-*.

‘20’ is *vīsaiti*. ‘30’ to ‘50’ are compounds with ^o*satəm*: *0ri-satəm*, *ca0βar^o-satəm*, *pañcā-satəm*. ‘60’ to ‘90’ are fem. *ti*-stems: *xšuuāš-ti-*, *haptāⁱ-ti-*, *aštāⁱ-ti-*, *nauuaⁱ-ti-*.

‘100’ and ‘1,000’ are neut. *a*-stems: *sata-*, *hazaṅra-*, with mixed plur. forms: *tišrō sata*, *ca0βārō sata*, *nauua hazaṅra*, etc. ‘10,000’ is neut. *baēuuar/n-*.

Compounded numerals are represented by *pañcā-ca vīsāⁱti* ‘25’, *0raīias-ca 0risqs-ca* ‘33’, *pañcā-ca haptāⁱti* ‘75’, *nauua-ca nauuāⁱti* ‘99’.

Higher numerals include *duīie nauuāⁱti* ‘180’ (two 90) and *nauua-ca . . . nauuaitiš-ca nauua-ca sata nauua-ca hazaṅra nauuasōs-ca baēuuar* ‘99,999’, which is the highest Av. number, there being no word for ‘100,000’.

3.6.3 Derived numerals

3.6.3.1 Multiplicatives

Multiplicatives (‘times, -fold’) have simple and extended forms:

- ‘once’: YAv. *ha-kər²t*, OPers. *ha-karam*; cf. Av., OPers. *hama-* ‘one and the same’;
- ‘twice’: YAv. *biš*, *biž-uuaṅ*; – ‘thrice’: YAv. *0riš*, *0riž-uuaṅ*; – ‘four times’: YAv. *ca0ruš*;
- ‘six/nine times’: *xšuuāš-aiia*, *naom-aiia*;
- the tens: *vīsaiti-uuā*, *0risa0-βā*, etc.

3.6.3.2 ‘-th time’

The ‘2nd–4th time’ are formed from the ordinals with the prefix *ā-*: *ā-ḡbitīm/lā-ḡbitīm*, *ā-0ritīm*, *ā-xtūrīm*.

3.6.3.3 Fractions

Fractions are made with the formant *-Sya-*: *0ri-šuuā-*, *ca0ru-šuuā-*, *paṅta-ṅ^hha-*, *hapta-huuā-*, *ašta-huuā-*. The same formation is attested for Old Persian by Elamite texts: < ši-iš-maš > = **či-šuva-*, < ṣa-iš-šu-iš-ma > = **čaṣuva-*, < aš-du-maš > = **ašta^huva-*, etc. (Hoffmann 1965).

4 MORPHOLOGY II: VERBAL MORPHOLOGY

The Iranian verb inherited from Indo-Iranian the categories of aspect, tense, mood, voice, number, and person expressed by ‘finite’ forms of the verb, and non-finite forms, which behave like nouns (infinitives) or adjectives (participles).

Tenses and moods are expressed by endings and augment, while aspect is expressed lexically by different stems of the verb (see section 7.1).

Verb stems are classified as ‘thematic’, ending in *-a*, and ‘athematic’, ending in other vowels and diphthongs or in consonants. Athematic verbs are frequently thematised.

Most personal endings are the same in the two classes (the 2nd sing. act. is a noteworthy exception), with morphophonological changes in athematic verbs in clusters resulting from contact between the final consonant of the stem and the initial consonant of the ending.

The augment *a-*, which characterises an event as having taken place in the past, is found occasionally in Old Avestan (imperfect and aorist) and Young Avestan (imperfect), but regularly with the imperfect in Old Persian. In Young Avestan and Old Persian, it is also found with optative forms (see below).

4.1 Stem classes

Each conjugation contains a miscellany of stem formations, most of which are no longer productive in Iranian.

Athematic verbs show ablaut in the root syllable, e.g. *jan-lgn-* ‘smash’, or in the stem formant, e.g. *-nau-l-nu-*, distributed as follows (with exceptions):

- lengthened grade is found in the acrostatic ‘Narten’ presents (but see de Vaan 2004), the *s*-aor. ind. sing., and in some 3s perf. forms;
- full grade is found in the singular in the pres., aor., and perf. ind., throughout the paradigms in the subj., and in the 2p pres. ind. and 3s and 2p imper.;
- zero grade is found in the dual and plur. of the pres., aor., and perf., in the 2s and 3p imper., and throughout the opt.

Ambikinetiç ablaut, apparently conditioned by (proto-Av.) stress patterns, is found in the 3p pres. and in the 2s and 3s opt. endings:

- zero grade of the root + full grade of the ending: Av. plur. *-ənti*, *-ən*, opt. *-ii̯*, *-ii̯ā*, etc. (e.g. *h-ənti*, *h-ii̯ā* < *ah-* ‘be’);
- full grade of the root (or reduplication) + zero grade of the ending: *-d’ti*, *-at*, *-tš*, *-i̯*, etc. (e.g. *dad-d’ti* < *dad-at* < *-nti*, *-nt*, *da’d-i̯*).

4.1.1 Present stems

4.1.1.1 Athematic present stems

Athematic present stems include the following:

1. root stems: Av. *ah-las-lh-* ‘be’, *ai-li-* ‘go’, *jan-lja-lgn-* ‘smash, strike’; *stau-lstu-* ‘to praise’; OPers. *ah-las-*, *ai-*, *jan-*;
2. reduplicated stems:
 - stems with ‘light’ reduplication *CV-*: Av. *da-δā-l-da-δ-* ‘give, place’, *hišhak-lhišc-* ‘follow’ (with *hišhak-* for *hi-šak-* < *Sak-*); OPers. *da-dā-*;
 - ‘intensive’ stems with ‘heavy’ reduplication *CVC-*: Av. *zao-zao-l-zao-zu-* ‘keep calling’, *car²-kar-l-car²-kər²-* ‘keep praising’;
3. stems with *n*-infixes:
 - *na-C-l-n-C-* stems: Av. *ci-na-h-l-ciš-* (< *ci-n-š-*) ‘*assign’ (< *kaiš-l-ciš-*);
 - *nā/n-* stems (< *na-H-l-n-H-*): YAv. *zi-nā-l-zi-n-* ‘take away’, OP. *di-nā-l-* (< *ji-na-H-l-ji-n-H-*);

- *naulnu*-stems: Av. *s^hru-naolnu-* ‘hear’ (< *sṛ-na-u-lu-u-* [or *sru-nau-*] < *sru-*); Av. *kər^h-naulnu-* ‘do, make’, OPers. *kunau-* (< *kṛ-nau-lnu-* < *kar-*).

4.1.1.2 Thematic present stems

Thematic present stems include the following:

1. stems in *a-*:
 - root in zero grade: Av. *mər^hza-* ‘wipe’, OPers. *ṃārda-* (or *marda-*, MPers. *māl-*);
 - root in full grade: Av. *bara-* ‘carry’, *bauua-* ‘become’, OPers. *barā-*, *bava-*;
 - root in long grade: Av. *frāda-* ‘to further’;
2. stems in *ja* (see also section 4.1.1.4):
 - root in zero grade: Av. *miriia-*, OPers. *mṛiia-* ‘die’ (< *mṛ-ja-*);
 - root in full grade: Av. *spasiia-* ‘spy on’, OPers. *jadiya-* ‘ask for’;
3. stems in *aia* (see also section 4.1.1.5):
 - root in zero grade: Av. *saδ-aiia-* ‘appear’, OPers. *ṭad-aya-* (< *čnd-aya-* < (*s*)*čand-*);
 - root in full grade: Av. *daēs-aiia-* ‘show’, OPers. *ṭar-aya-* ‘cross (river)’;
4. stems in *-āja-* with root in zero grade: *gə^hruu-āiia-* ‘seize’, OPers. *garb-āya-* (cf. YAv. *gər^hβ-nā-*);
5. stems in *-sa-* (OInd. *-cha-*, IE. *-skē-l-sko-*)
 - root in zero grade: Av. *ja-sa-* ‘go’ (for **gasa-* < *gṃ-sa-* < *gam-*), *tər^h-sa-* ‘to become afraid’; OPers. *ā-ya-sa-* ‘take’ (< *jam-*), *tar-sa-*;
 - roots in long *ā*: Av. *yā-sa-* ‘request’, OPers. *xšnā-sa-* ‘know’ (< *jnā-*, cf. Latin, Greek *gnō-sk-*);
6. stems with *n* infix: Av. *kər^h-n-ta-* ‘cut’ (< *kart-*); YAv. *viṇda-* (< *vaid-*, athematic OAv. *vi-na-d-*);
7. reduplicated stems:
 - with ‘light’ reduplication: Av. *hi-šta-* ‘stand’, OPers. *(h)i-šta-* (< *Si-St-a-*, *stā-*); OAv. *ṃiia-* ‘implore’ (< *i-y-a-* < *yā-*);
 - with ‘heavy’ reduplication (intensive): Av. *ja-γna-* ‘smash to smithereens’ (< *gan-ljan-*);
8. desiderative stems in *-Sa-* with reduplication: OAv. *dī-dray-žā-* ‘to wish to hold firmly’ (< *drang-l-drng-*), OYAv. *diβžā-* ‘seek to deceive’ < *di-(d)b-ž^ha-* < *di-db^h-šā-* (< *d^hab^h-*); YAv. *zi-xšnā-gha-* ‘wish to know’ (< *jnā-*); *mi-mar^hx-ša-* ‘wish to destroy’ (< *mark-*).

4.1.1.3 The future stem

A future stem is formed with *-Sja-* added to the root in the full grade. Old Avestan has only two forms: *vax-šiiā-* < *vak-* ‘speak’ and *sao-šiiā-nt-* < *sau-* ‘revitalise’. Young Avestan has a few, including: *sao-šiiā-* and the irregular *bū-šiiā-* < *bau-* ‘become’. There are no Old Persian forms attested.

4.1.1.4 Passive stem

A passive stem is formed with the suffix *-iā-*, with the root in the zero or full grade: YAv. *kir-iiā-* ‘be done, be made’, OPers. *kar-iya-* (<*kr̥iā-*), YAv. *ba'r-iiā-* ‘be carried’ (<*bar-*). See also section 4.3.1.

On the 3rd sing. passive in *-i*, see section 4.3.2.

4.1.1.5 Causative stem

A causative is formed with the suffix *-aiā-* to the root in the lengthened (or full) grade. It usually corresponds to a passive or intransitive present: *s^hru-nau-* ‘hear’ vs. *sru-iiā-* ‘be heard’ vs. *srāuu-aiiā-* ‘make heard’ > ‘recite, sing’; *su-iiā-* ‘be revitalized’ vs. *sauu-aiiā-* ‘revitalize’.

4.1.1.6 Denominative stem

This group contains stems in *-iā-* and *-aiā-* and *-aniā-* (cf. Oettinger 1992): *bišaz-iiā-* and *baēšaz-iiā-* ‘to heal’ (<*bišaz-* ‘doctor’, *baēšaza-* ‘medicine’), Av. *ar^hza-iiā-* ‘to do battle (*ar^hza*)’, *pər^hs-aniiā-* ‘to discuss’, OPers. *patiy-a vah-āya-* ‘to pray for help’ (cf. Av. *auuah-* ‘help’). See also Tucker 2004.

4.1.2 Aorist stems

There are the following principal types of aorist stems:

- the root aorist: Av. *jam-lga-lgm-* ‘go, come’ (pres. *jasa-*), Av., OPers. *dā-l-d-* ‘give, place’ (pres. *da-dā-*);
- the thematic aorist: Av. *taš-a-* ‘fashion’ (pres. *tāš-*);
- the *s*-aorist: Av. *stā-h-* < *stā-* ‘stand, place’, *nai-š-* < *nai-lnī-* ‘lead’ (pres. *naiiā-*), *varš-* < *var^hz-* ‘produce’ (pres. *vər^hziiā-*), OPers. *dar-š-* < *dar-* ‘to hold’ (pres. *dāraya-*);
- the reduplicated thematic aorist: Av. *va-oca-* < *va-uc-a* < *vac-* ‘to speak’ (pres. *mrao-*);
- suppletive systems: Av. *va-oca-* (pres. *mrao-*); *dars-* ‘see’ (pres. *vaēna-*).

Some verbs have both root- and *s*-aorist forms, e.g. Av. *man-* ~ *māh-* ‘think’ (pres. *maniiā-*).

4.1.3 Perfect stems

The perfect stem is formed by reduplication, with the exception of *ād-lāδ-* (invariable) ‘say’ and *vaēd-lvaēδ-* (*vaēθ-*) ‘know’. The vowel of the reduplication syllable varies:

- *a*, *i*, or *u*, in harmony with the vowel of the root (velars are palatalized before *a* and *i*): *ca-kan-* ‘love’, *ci-kaēθ-* ‘*point out’, *ru-raoδ-* ‘howl’, OPers. *ca-xr-*, weak stem of *ca-kār-* < *kar-* ‘do’;
- rarely *ā*: *dā-dar^hs-* < *vaēn-l-dars-* ‘see’;
- roots in initial *ā-* have long *ā-*: *āp-* ‘reach’, *āṇh-* < *ah-* ‘be’ (<*a-ah-*);
- the root *ai-li-* has perf. stem *ū-ai-*: YAv. *ii-aē-* (*i-aē-*) < *aē-* ‘go’.

4.1.4 Suppletive stem systems

There are a few suppletive verb systems, e.g. Av. pres. *mrāu-*, aor., perf. *vac-* ‘speak’; pres. *vaēna-*, aor., perf. *dars-*, OPers. pres. *vaina-*, imper. *dī-* ‘see’.

4.1.5 Periphrastic formations

On the use of past participle + ‘to be’ in YAv. and OPers., see section 6.4.2.2; on the pres. part. + modal forms of ‘to be’ see section 7.2.3.7ce.

4.2 Moods

The subjunctive and optative take special formants added to the stem (present, aorist, perfect), while the imperative is characterised by a special set of endings.

4.2.1 The subjunctive

The subjunctive is formed from the present, aorist, or perfect stems by the formant *-a-*. This means that athematic verbs are thematised, while thematic verbs take the combined formant *-ā-*. The endings are primary or secondary. The primary 1s is *-ni* (vs. indicative *-mi*).

4.2.2 The optative

The optative is formed from the present, aorist, or perfect stems by the ablauting formant *-jaH-/iH-* > *-jā-/ī-*. In thematic verbs, the thematic vowel combines with *-ī-* to yield the characteristic thematic optative formant *-ai-*.

In Old Persian and, sometimes, in Young Avestan, the optative can take the augment to produce the preterital optative: YAv. *a-uuaēn-ōiš* ‘(whenever) you saw’, *nišādāi-ōiš* ‘you would seat (me)’; OPers. *avājan-iyā* ‘he would kill’, *a-kunav-ayaⁿtā* ‘they would do’ (see section 7.2.2.2).

4.3 Voice

There are two ‘voices’: active and middle. Verbs can have active and/or middle forms. Some verbs have only active forms, some only middle forms, and some both active and middle forms. Only in the third group can the middle forms have a special function, different from that of the active forms. Intransitive verbs typically have only active (e.g. YAv. *aēti* ‘he goes’) or only middle forms (YAv. *āste* ‘he sits’), while transitive verbs can take both active and middle forms. If they do, then most often the distinction between the two forms is active vs. passive (see further section 7.3.2).

4.3.1 The passive in *-ja-*

More commonly, the passive (‘it is done, he is killed’) is expressed by the special present stem in *-ja-* (YAv. *-iia-*, OPers. *-iya-*). In Avestan, this form normally takes the middle endings; in Young Avestan, active endings are also well attested (in Old Persian 3rd person *-tiy* and *-taiy* are indistinguishable <*-t-i-y*>).

4.3.2 The 3rd singular passive in -i

A special Indo-Iranian 3s passive form was made from the aorist stem with the ending -i (OAv. -ī, YAv. -i): OAv. *srāuu-ī* ‘has become renowned’. In Young Avestan, this rare form was made from the present stem (*ja'n-i* ‘was smashed’, *əʔ'nāuu-i* ‘was sent on its way’) and from the perfect stem (*ā'δ-i* ‘is/was said’).

4.4 Person marking (endings)

The endings are of the Indo-Iranian (and late Indo-European) type, those of the present and aorist indicative differing from those of the perfect indicative, and with distinct sets of ‘primary’ in the pres. and perf. ind. and ‘secondary’ endings elsewhere. The subjunctive takes endings from either set. See Tables 3.4.1-4.

In the 1s, Avestan still has the pres. ind. act. primary ending OAv. -ā, YAv. -a, beside the more common -āmi; the subj. has -ā and -āni.

The 1st person endings of the dual and plural are parallel: act. primary -yahi, -mahi, secondary -ya, -ma; middle secondary -yadi, -madi, YAv. -ma 'δel-ma'de.

On the 3s pass, ending -i, see above.

TABLE 3.4.1: VERBAL ENDINGS 1. PRESENT INDICATIVE; PRESENT AND AORIST INJUNCTIVE AND IMPERATIVE

	Present indicative		Present/aorist injunctive		Present/aorist imperative	
	Athematic	Thematic	Athematic	Thematic	Athematic	Thematic
Active						
1s	-mi	-ā, -ā-mi	-am	-a-m		
2s	-Si	-a-hi	-S	-a-h	-di	-a
3s	-ti	-a-ti	-t	-a-t	-tu	-a-tu
1p	-mahi	-ā-mahi	-ma	-ā-ma		
2p	-θa	-a-θa	-ta	-a-ta	-ta	-a-ta
3p	-anti, -ati	-a-nti	-an	-a-n	-a-ntu	-a-ntu
			aor. -an, -at			
1d	-yahi		-ya	-āya		
3d	-tah	-a-tah	-tam	-ā-tam		
Middle						
1s	-ai	-ai	-i	-ai		
2s	-Sai	-a-hai	-Sa	-a-ha	-Sya	-a-hya
3s	-tai, -ai	-a-tai	-ta	-a-ta	aor. -ām	-a-tām
			aor. -ātam			
1p	-madai	-ā-madai	-madi			
2p	-dyai	-a-dyai		-a-dyam	-dyam	-a-dyam
3p	-ntai, -atai	-a-ntai	-ata	-a-nta	pres. -rām	-a-ntām
	-rai-, -ārai					
1d			-yadi			
3d	-aitai, -aiθai	-aitai, -aiθai		-aitam		

TABLE 3.4.2: VERBAL ENDINGS 2. PERFECT, PLUPERFECT

	Perfect		Pluperfect
	Active	Middle	
Sing.			
1	-a	-ai	
2	-θa		
3	-a	-ai	-a-t
Plur.			
1	-ma		
3	-r		-rš
Dual			
3	-ātr	-ātai	

TABLE 3.4.3: VERBAL ENDINGS 3. SUBJUNCTIVE AND OPTATIVE

	Subjunctive		Optative	
	Athematic	Thematic	Athematic	Thematic
Active				
1s	-ā	-ā, -ā-ni	-jā-m	
2s	-a-h, -a-hi	-ā-h, -ā-hi	-jā-h	-ai-š
3s	-a-t, -a-ti	-ā-t, -ā-ti	-jā-t, -ī-t	-ai-t
1p	-ā-ma	-ā-ma	-jā-ma, -ī-ma	-ai-ma
2p	-a-θa, -a-ta	-ā-θa, -ā-ta	-dyai	
3p	-a-n, -a-nti	-ā-n, -ā-nti	-jā-n	
			-jā-r, -jā-rš	-aj-an
3d			-ī-tam	
Middle				
1s	-āi, -ānai	-āi, -ānai	-(ī)j-a	-aj-a
2s	-a-hai	-ā-hai	-ī-ša	-ai-ša
3s	-a-tai	-ā-tai	-ī-ta	-ai-ta
1p	-ā-madai	-ā-madai		-ai-madai
2p				-ai-dyam
3p	-a-ntai	-ā-ntai		-aj-anta

TABLE 3.4.4: 'TO BE'

	Pres.	Inj.-Ipf.	Subj.	Opt.	Imp.
1s	ah-mi		ah-ā	h-jā-m	
2s	ah-i		ah-a-h	h-jā-h	z-di
3s	as-ti	ās	ah-a-t(i)	h-jā-t	as-tu
3d	s-tah				
1p	mahi		ah-ā-ma	h-jā-ma	
2p	s-ta			h-jā-ta	
3p	h-anti	āh-an	ah-an	h-jā-r	h-antu
1d		ah-ya(?)			
2d	s-tah				

In the athematic 3rd person pres. and aor. mid., a few forms lack the *t* in the sing. and have forms with *r* in the plur.: OAv. pres. ind. sing. *is-ē* ‘desires’, aor. imper. sing. *d-qm*, *uc-qm* ‘let it be given, spoken’, pres. imper. plur. *jān²-raqm* ‘let them be smashed!’, *xrūn²-raqm* ‘let them be bled!’, YAv. pres. ind. 3rd sing. and plur. *mruii-e* ‘is said’ (< *mruu²-ē*), *mruu-ā²re* ‘they are said’ (< *mrao-lmru-*); *niγn-e*, *niγn-ā²re* ‘is/are smashed down (upon)’ (< *jan-lγn-*, see Hintze 2005); *āgh-ā²re* ‘they sit’ (but 3s *āste* < *āh-*), *sōi-re/saē-re* ‘they lie’ (but 3s *saē-te*).

4.5 Preverbs

Common preverbs include *us-* ‘up, out’ – *ni-* ‘down’; *fra-* ‘forward, forth’ – *apa-* ‘away, off’; *ati-* ‘up to’ – *niš-* ‘out, away, separate, beyond’; *ham-* ‘together’ – *ui-* ‘out of, apart’. In addition, a subset of adpositions also has a preverbal function (see section 3.5). For examples see sections 5.1.2–3.

4.6 Nominal forms of the verb

4.6.1 Participles

Active and middle participles are derived from the present, aorist, perfect, and future stems of the verb (see sections 4.6.1–3). The ‘past participle’ with the suffix *-ta-* is derived from the root, usually in the zero grade (see section 4.6.1.4), and the verbal adjectives expressing necessity from the root, usually in the full grade, with various suffixes (see section 4.6.2).

4.6.1.1 Present, future and aorist active participles

The active participles of the pres., fut., and aor. stems take the formant *-nt-*:

- thematic verbs with fixed stems in *-ant-*: pres. OAv. fem. *maēk-a²nt-ī-* ‘sparkling’; YAv. them. *bar-ənt-* ‘carrying’, *bū-šii-ant-* ‘who will be’; future: OAv. *sao-šii-ant-* ‘he who will revitalize’, fem. *bū-šii-e²nt-ī-*; – aor.: OAv. *han-ant-* ‘gaining’;
- athematic forms with strong stems in *-ant-* and weak stems in *-at-*: pres. OAv. *h-ənt-* ‘being’, fem. OAv. *šii-e²t-ī-* ‘inhabiting’; YAv. *h-a²tī-* ‘being’; – aor. YAv. *°buu-ant-* ‘becoming’; – acrostatic present stems with fixed weak stems: OAv. *stauu-at-* ‘praising’.

4.6.1.2 Perfect active participles

The perf. part. act. is made with the formant *-uah-l-uš-*: OAv. *vīd-uuah-lvīd-uš-* ‘knowing’, YAv. *vīd-uuah-lvīd-uš-* ‘knowing’ (see section 2.3.2.2b) *ja-γm-uš-* ‘having gone/come’ (< *gam-*). Note the conditioned changes (Skjærvø 1997a): *vaox²-ah-* ~ *vaokuš-* ‘having spoken’ (< *ua-uk-uah-* < *vak-*), *vauuan-uuah-* ~ *vaomuš-* ‘having won’ (< *van-*), *yōiθ-βah-* ~ *yaētūš-* ‘having taken up one’s position’ (< *yat-*), *tarš-uuah-* ‘having fashioned’ (< *ta-tš-uah-* < *taš-*).

Two types of adjectives resemble the perfect participle and have similar functions, with *-u-* and *-uan-*: OAv. *vīd-u-* ‘knowing one’, *vā-un-u-* ‘winner’, YAv. *za-z-u-* ‘winner’; OAv. *ci-ciθ-βan-* ‘who understands’, YAv. *yōiθ-βan-* ‘standing ready’.

4.6.1.3 Middle participles

The middle participles of thematic verbs end in Av. *-əmna-* (*aiia*-stems: *-aiiamna-*), OPers. *-amna-*, those of athematic verbs, including the perfect, regularly end in Av. *-āna-*.

4.6.1.4 Past participles in *-ta-*

The past participle (Eng. ‘gone, killed’) has the ending *-ta-* with the root in the zero grade if possible and with the regular sound changes: *śuta-* ‘set in motion’ (< *śau-*); *ux-ta-* ‘spoken’ (< *vak-*), *bas-ta-* ‘bound’ (< *band-*); full grade: *dā-ta-* ‘given, placed’ (< *dā-*); *aox-ta-* ‘spoken’ (< *aog-*; in *aoxtō.nāman* ‘whose name is spoken’).

On the ‘periphrastic’ perfect, see sections 7.4.2.2 and 7.4.3.

4.6.2 Verbal adjectives in *-θa-*, *-θβa-*, *-iia-*

The ‘participles of necessity’, expressing ‘that which ought to or can be done’, are formed with the suffixes *-θa-* (*-δa-* by internal sandhi): *ux-δa-* ‘that ought to be spoken’; *-θβa-* (*-tuua-* by internal sandhi): *vax³-δβa-* ‘that ought to be spoken; or *-iia-*: *karš-iia-* ‘that ought to be ploughed’. The suffix *-iia-* could also be attached to nouns: *yesn-iia-* ‘deserving of sacrifices’ < *yasna-*.

4.6.3 Infinitives

There are two sets of infinitives, which may be formed from the root or the present stem (see also Kellens 1994): those ending in *-dīāi* have medial function (cf. Lühr 1994), the others are neutral with respect to voice.

The infinitives in *-dīāi* can be made either from the pres. stem: athem. (root in zero grade) OAv. *ja¹-dīāi* ‘to strike’ (< *jan-*), YAv. *da-z-dīāi* (< *da-d²-dīāi*) ‘give, establish’, them. YAv. *vaza¹-dīāi* ‘drive’; or from the aorist stem: aor. stem. athem. OAv. *dər³-dīāi* ‘hold (up)’.

The neutral (active/middle) infinitives are the following:

- exclusively OAv.: *-ai*: *p-ōi* ‘protect’ (< *pā-*); *-uāi*: *vīd-uīē* ‘know’ (< *uīd-uāi*); *-uuanai*, *-manai*: *vīd-uuanōi* ‘know’; *-Sai*: athem. *nāšē* ‘to obtain’ (< *nās-šai*); them. *srāuuaie-ghē* ‘recite’;
- both OAv. and YAv. (rare) are infinitives in *-tai*: OAv. *gaṭ.tōi* ‘go’ (for *ga-tōi* < *gam-*), *i-tē* ‘go’, OYAv. *mrū¹-tē* ‘speak’, *s-tōi* ‘be’ (< *ah-*);
- only YAv., the infinitive in *-taiāi* (identical with the dative of an action noun in *-ti*): *ux-tēe* ‘to speak’;
- only OPers., the infinitive in *-tanaiy*, made from the root in the full grade: *bar-tanaiy* ‘to carry’, *car-tanaiy* ‘to do’ (< *kar-*).

5 SYNTAX I: WORD ORDER AND CLAUSE STRUCTURE

5.1 Word order

The basic word order is verb-final: SOV. Preverbs in tmesis are always, adverbs often, clause-initial; the indirect object usually follows the direct object. The second position in the clause is reserved for enclitics (see section 6.5.1.3).

5.1.1 Adpositions

Adpositions can precede or follow their government. The position is freer in Old Avestan than in Young Avestan and Old Persian.

Adpositions follow enclitic pronouns they govern:

YAv. *kaḍa [nō auui] ājasāṭ nmānō.pa'tiš*
 'when will our home-lord come [to us]?' (Yt. 17.10).

Enclitic postpositions will seek the second place in the clause:

[asmanəm auui] frašusāni
 'shall I go forth [to heaven]?' (Yt. 17.57),

vs.

mā [auui asmanəm] frašusa
 'do not go forth [to heaven]!' (Yt. 17.60).

Adpositions (and preverbs) are frequently placed between a noun and its modifier:

YAv. *[aēuuqm] vā [auui xšapanəm]*
 'or [for one night]' (Yt. 8.11).

5.1.2 Preverbs

In Avestan, but not in Old Persian, the preverbs may be detached from the main verb, usually preceding it, sometimes fronted to the beginning of the clause, sometimes separated by one word only (frequently the conjunction *-ca*; Hale 1993). Coordination of separated preverbs with (*-ca . . .*) *-ca* is frequent.

In Old Avestan, they are then often repeated directly before the verb, see section 5.1.3.

They may also be lowered to the end of the clause (or metrical half-line). When a verb with preverb is repeated, only the preverb may be repeated.

In Young Avestan, a verb can, apparently, have two preverbs, e.g. *paiti auua.jasa-* 'to come down *in return'.

Examples:

- joined to the verb in main and subordinate clauses:

OAv. *aiiā nōiṭ ərʷš [vī-šiiātā] . . .*
hiiṭ iš dʷbaomā . . . [upā.jasat]
 'they did not [discriminate] correctly between these two,
 because deception [came over] them' (Y. 30.6);

- separated and fronted:

YAv. *[pa'ti] maqm ərʷzuuō [pərʷsaj'ha]*
yezi maqm [pa'ti.pərʷsaj'he]
 '[consult] me [in turn], O upright (Zarathustra)! . . .
 if you [consult] me [in turn]' (V. 18.7);

OAv. *[ā]-zī dʷmānəm . . . [°dāṭ] | dušitā-cā*
 'for he [has placed] the home [in] . . . and difficulty' (Y. 31.18);

YAv. *yezī-ca hē [hqm] tafnō [jasā]*
 ‘and if fever [comes (and joins) with] her’ (V. 7.70);

- post-verbal:

OAv. *xraosəntqm upā*
 ‘let them be howled upon!’ (Y. 53.8);

- one preverb for two forms of the same verb:

OAv. *yā zī [āitī jəṅghati]-cā*
 ‘namely, (those things) that [are (now) coming] (*āiti* < *ā aēti*)
 and (those) that [shall come] (*ā . . . jəṅgha’ti*)’ (Y. 31.14);

- two preverbs for one verb:

YAv. *ā tē aṅhe fra-ca stuiie*-pres. ind.
 ‘for this I am assigning myself by my praise (*ā-stuiie?*) to you and utter your praise
 (*fra-stuiie*)’ (Y. 1.21);

- preverb minus verb:

OAv. [*apānō*] *darəgō.jiūā’tīm*
 [*ā*] *xšaθrəm vaṅhəuš manaṅhō*
 ‘[having obtained] (*ā-apānō*) long life,
 having (obtained) (*ā*) the command of good thought’ (Y. 33.5);

- with change of preverb (and verb?):

kaθā drujəm [nīš] ahmaṭ ā [°nāšāmā] /
təṅg ā [auuā] yōi asruštōiš pər³ nāṅhō
 ‘how we shall [take away] here from us the Lie,
 (and) [(bring it) down] upon those who (are) full of refusal to listen’ (Y. 44.13).

5.1.3 Repetition of preverb in tmesis

In Old Avestan (exceptionally in YAv.), preverbs separated from the verb are often repeated directly before the verb (observing sandhi rules). This phenomenon was early on recognized as ancient interference with the text, as the repeated preverbs obey sandhi rules, but are shown by the meter to be additions to the text:

OAv. [*frō*] *spəntā ārma’tē / ašā daēnā [fra-daxšaiiā]*
 ‘[launch forth], O Humility, (our) vision-souls through life-giving Order!’ (Y. 33.13);

YAv. [*auua*] *mē āziš . . . aṅhuuqm [auua-dər³nq] sadaiiēti*
 ‘that *āzi* (= ?) appears (to be) [ripping off] my life thread’ (V. 18.19).

5.2 Topicalisation: ‘raising’ and ‘lowering’

Departures from the basic word order are frequent, most often for metrical reasons, but also for emphasis of various kinds (see also section 10 Stylistic features).

‘Raising/fronting’ and ‘lowering/backing’ of the main constituents are quite common.

In addition, in the Old Avestan and Young Avestan metrical texts, word order to a large extent depends on the exigencies of the metrical units, although similar tendencies in Old Persian show that it is also a feature of elevated style (cf. Hale 1988).

In general, verbs tend to be raised to initial position, and other parts of the sentence tend to be lowered. In Old Avestan, this often happens when the sentence goes over more than one ‘half-line’. In Young Avestan and Old Persian, it frequently happens with complex direct objects. Raising is often combined with other effects, such as parallelism and chiasmus.

5.2.1 Raising of verb

OAv. [ahmī] mazdā anaēšō . . . kamnānā ahmī
‘[I am], O Mazda, weak . . . of few men I am’ (Y. 46.2) [chiasmus];

YAv. [jamiūā] vō vaŋhaoŋ vaŋhō
‘[may there come] to you (what is) better than good! (Y. 59.31);

OPers. [ōātiy] dārayavauš xšāyaθiya
‘King Darius [announces]’.

5.2.2 Raising of direct object

OAv. yaθā [ratūm] ahurō vaēdā
‘how the Lord knows [the model] (to be)’ (Y. 31.2);

YAv. mā yauue [imaŋ nmānəm] x^vāθrauuat x^var^{nō} frazahūt
‘may not ever Fortune providing comfort leave [this house]!’ (Y. 60.7);

OPers. [xšaçam] hauv agarbāyatā
‘[the royal command] he took for himself’ (DB 1.41–42).

5.2.3 Lowering of subject

OAv. āuuš . . . hēntū . . . [rātaiō]
‘let there appear . . . [gifts]!’ (Y. 33.7);

YAv. ā dim pər^{sa}ŋ [zaraθuštrō] ‘Zarathustra asked him’ (Y. 9.1).

5.2.4 Lowering of direct/indirect object

Lowering, often for metrical reasons, is frequently found with phrases (dir. obj., double dir. objs., indir. obj.) consisting of noun + determinative or sequences of nouns:

OAv. iθā āŋ yazama’dē [ahurəm mazdām]
‘thus, in this manner we are sacrificing [to Ahura Mazda]’ (Y. 37.1);

YAv. yaŋ kər^{nao}ŋ aŋhe xšāθrāda / a-maršanta [pasu vira] [8 + 8 syllables]
‘. . . that he (= Yima) made during his reign / imperishable [cattle and men]’ (Y. 9.4);

OPers. yadiy imam dipim vaināhay [imai]-vā [patikarā]
‘if you see this inscription or [these pictures] . . .’ (DB 4.72–73).

5.2.5 Lowering of adverbial complements

YAv. *mā zaṃ vaēnōiṭ [ašibiia]*
 ‘may he not see the earth [with (his evil) eyes]!’ (Y. 9.29);

OPers. *avam kāram . . . adam a-janam [vasiy]*
 ‘I smashed that army [greatly]’ (DB 1.88–89).

5.3 Special types of clauses

Regular positive statements are of the common Indo-Iranian type. Following are remarks on special clauses.

5.3.1 Questions

Explicit questions are introduced by interrogative pronouns or particles. When they are not, they are, obviously, sometimes difficult to identify. The mood in questions is often the deliberative subjunctive (see section 7.2.1.1):

- without interrogative particle:

OAv. *vīdušē zī nā mruiiāṭ*
 ‘for may a man tell the knowing one?’ (Y. 51.8);

- with interrogative pronouns:

OAv. *[kō] vā ašā āfraštā [kā] spəntā ārma’tiš*
 ‘[who] has either consulted with Order, (and) [with whom] (has) life-giving Humility (consulted)?’ (Y. 51.11);

YAv. *[kuua] hē aošō viṇdāma*
 ‘[where] shall we find death for him?’ (V. 19.46);

- with interrogative particle, including *kaṭ* and *-nā*, approximately ‘I wonder’:

OAv. *[kaṭ] mōi uruuā isē cahiiā auuaṇhō*
 ‘does my soul, I wonder, control anybody’s help (at all)?’ (Y. 50.1);

kōm[-nā] mazdā mauua’tē pāiiūm dadā
 ‘whom, [I wonder], do you, O Mazdā, give as protector to one like me?’ (Y. 46.7);

- disjunctive questions:

OAv. *[katārəm] ašauuā [vā] drəguuā [vā]*
vərauuuā’tē maziio /
vīduuā vīdušē mraotū
 ‘[whether] the sustainer of Order [or] the one possessed by the Lie will *retain for himself the greater (reward), let him who knows say to the one who knows!’ (Y. 31.17);

YAv. *[katārəm] ā0rauuu a0a^urunəm [vā] paraiiāṭ*
gaē0anəm [vā] aspə^rnō auuōiṭ
 ‘should the priest go away to (do) priestly studies, [or] should he help as *overseer of the livestock?’ (Her. 3).

5.3.2 Exhortations and commands

Exhortations (commands, suggestions, wishes) to 1st, 2nd, and 3rd person are expressed in various ways. Simple commands are expressed by the imperative. Negated commands or exhortations (also to 1st pers.) are expressed by *mā* + inj. (pres. or aor.) or *mā* + opt. (YAv., OPers.).

Wishes are expressed by the optative, negated by OAv. *nōit*, YAv., OPers. *mā*. The difference between negated commands and negated wishes is sometimes minimal.

5.3.3 Parenthetical clauses in the nominative

Parenthetical clauses in the nominative are found in Avestan:

OAv. *nū im vīspā [ciθrē zī] mazdāyḥō.dūm*
 ‘now, (you) all, – [for it (is) brilliant!] – pay attention to it (the announcement?) . . .’
 (Y. 45.1).

In Young Avestan, we occasionally find descriptive parenthetical phrases in the nominative such as the following:

upa.tacaṭ ar³duuī sūra anāhita ka'nīnō kəhr pa srīraiīā . . .
[raēuuat ciθrəm] āzātaiīā
 ‘Arəduuī Sūrā Anāhitā came running in the body of a beautiful young woman
 – [(her) seed (being) wealthy (= of wealthy seed)] – high-born’ (Yt.5.64).

On the Old Persian ‘naming phrase’, see section 6.4.1.1.

5.3.4 Reported speech

Reported speech is identical with direct speech and may be recognizable by context only: in Avestan as the object of verbs of speaking, thinking, etc., or accompanied by the particle *u'tī*; in Old Persian introduced by the conjunction *taya* (cf. Schmitt 1995). It is found in the following cases:

object of verbs of speaking, thinking, etc.:
 OAv. *ōiiā cistī [huuō ptā ašahiīā mazdā]*
 ‘by this insight: [He is the father of Order, the all-knowing one]’ (Y. 47.2);

YAv. *yaṭ *aēšō yō apər³nāiiukō . . .*
*aoxtō vā hē aoxte [*0βaṭ pairi *aṭha]*
 ‘when this child . . .
 or having been spoken to says to him: [I shall stay with you]!’ (Her. 7);

OPers. *yadi-maniyā^hi y [šiyāta ahaniy jīva utā maṛta aṛtāvā ahaniy]*
 ‘If you shall think: [Let me be happy both (while) alive, and (after I am) dead let me be blessed!] . . .’ (XPh 46–50).

In Old Persian the verb of speaking, thinking, etc., can be embedded in the reported speech:

[taya] amaniyaiy [kunavāniy] ava-maiy visam ucāram āha
 ‘[whatever] I thought [I shall do] that all was easy for me’ (DSI).

The Avestan particle *u'ti* precedes or follows reported speech:

OAv. *yaiiā spaniiā u'ti mrauuai yēm aṅgrəm /*
[nōi nā manā . . . nōi daēnā noi uruuqno hacintē]
 'of which two the life-giving one **thus** shall say (to him) whom (you know to be) the
 Evil one:
 [Neither our thoughts . . . nor vision-souls, nor souls go together]' (Y. 45.2);

YAv. *[ā'ī vō kas-ci mašiiānəm]*
u'ti mraoi ahurō mazdā
[āi ašūm zaraθuštra x'ar'nō ax'ar'təm isaēta]
 '[therefore may each of you men]
 – **thus** said Ahura Mazda –
 [O Orderly Zarathustra, seek the unseizable Fortune!]' (Yt. 19.53).

In Young Avestan, *u'ti* is often found with *aojana-*: *u'ti aojana-*, *u'tiaojana-* 'thus saying':

yō aoxta ahurāi mazdāi u'tiaojanō
[ahura mazda mainiō spēništa . . .]
 'who said to Ahura Mazda thus saying:
 [O Ahura Mazda, most Life-giving Spirit . . .]' (Yt. 8.10).

The OAv. particle *uiti*, apparently, also serves to abbreviate statements:

yōi vaṅhēuš ā manahō šiiēnti
yās-cā [u'ti]
 'the male (deities) who dwell with good thought,
 as well as the female ones who ["dwell with good thought"]' (Y. 39.3).

6 SYNTAX II: NOMINALS

6.1 Specificity and reference

There are no articles, but relative clauses and, especially, constructions with relative particles are used to express specificity, and demonstrative pronouns are used for reference to elements of the discourse (the former/latter, the aforementioned/following).

6.1.1 *The Young Avestan and Old Persian connecting relative and the relative particle*

In Young Avestan and Old Persian, the relative pronoun is frequently used to connect a noun with a nominal (adjective, genitive, etc.). Its function must originally have been 'specifying' or 'delimiting', cf. OPers. *hayā manā taumā* vs. *manā taumā* = 'my family' vs. 'my family'.

In these constructions, when the antecedent is in the nominative or accusative, the relative pronoun is usually in the same case.

In Young Avestan, the relative pronoun takes the invariable from nom.-acc. sing. neut. *yai* when the antecedent is in a case other than the nominative or accusative, occasionally also when it is in the nominative or accusative. In such clauses, the case of the relative pronoun and the predicate noun or adjective is the same as that of the antecedent.

6.1.1.1 Antecedent outside the relative phrase

Nominative-accusative:

YAv. [*vīspəm imaṭ yaṭ juiiō aṅhuš*]-**nom.**
 ‘[all this, the living existence]’ = ‘this entire living existence’ (H. 2.2);

[*0βqm*] *raṭūm āiiese yešti [yim zaraθuštrəm]*-**acc.**
 ‘[you, Zarathustra], I harness by my sacrificing as model’ (Vr. 2.4);

yō paō’riiō [gāθā] frasrāuuaiiaṭ
 [*yā pañca*]-**acc.** *spitāmahe . . . zaraθuštrahe*
 ‘(Sraoša) who (as) the first made heard [the *Gāθās*],
 [the five] of Spitāma Zarathustra’ (Y. 57.8);

haomō [təm-ciṭ yim kərəsānīm]-**acc.** *apa.xšaθrəm nišādaiiaṭ*
 ‘the haoma set [that Kərəsāni, too], down away from his command’ (Y. 9.24);

OPers. [*kāra haya manā*]-**nom.** *avam [kāram tayam hamīčiyam]*-**acc.** *ajd’ vasiy*
 ‘[my army] crushed that [conspiratorial army]’ (DB 2.25–26).

6.1.1.1a With invariable *yaṭ*

YAv.

[*puθrəm yaṭ pourušaspahe*]-**acc.**
 ‘[the son of Pourušaspa]’ (Yt. 5.18);

cuuaṭ [aētaiḡhā apō yaṭ armaēštaiiā]-**gen.**
 ‘how much [of this standing water]?’ (V. 6.30).

6.1.1.2 Antecedent inside the relative phrase

YAv. *yaṭ upaṅhacaṭ [yim yiməm xšaētəm huuaq0βəm]*-**acc.**
 ‘that he followed [splendid Yima of good herds]’ (Yt. 19.31);

OPers. *imā dahayāva [tayanā manā dātā]*-**abl.-inst.** *apariyāya*
 ‘these lands behaved [according to my law]’ (DB 1.20)

6.2 Number

Singular, plural, and dual are used in the usual ways, with a few noteworthy special features.

6.2.1 Neuter plural

The neuter plural of adjectives and pronouns is commonly used in the sense of ‘things’:

OAv. [*yā*] . . . [*acištā*]-**p/nt.** *danṭō / vaxšəntē*
 ‘giving [(those) which] (are) [the worst (things)] / to the blazing (fire)’ (Y. 32.4);

YAv. [*vaṅhu*]-**p/nt.-ca** [*vaḡhās*]-**p/nt.-ca** *āfrīnāmi*
 ‘I invite as (guest-)friends [good] and [better (things)]’ (Y.52.1).

6.2.2 Dual

The dual is used with the numeral *dya-* ‘two’ and the pronoun *uba-* ‘both’ or alone to indicate a pair of equal items (natural pairs or arbitrary items). Some words have only dual forms (OPers. *ušt* ‘hearing, consciousness’). See also section 6.3.2.

Two singular items:

OAv. *aṭ . . . [ahuuā . . . fr̥nəmnā] . . .*
mē uruuā gəuš-cā aziīā
 ‘thus, [we two were there presenting ourselves as guest friends]:
 my soul and that of the fertile cow’ (Y. 29.5).

A special form of the dual is the ‘elliptical dual’:

YAv. [*hāuuānāiās*]-**du/gen.-ca** *ašāiia [frašūtāiīā]*
 ‘and [of the mortar (and pestle) moved forth] in Orderly fashion’ (Y. 27.7);

[*duuāēbiia haca nər’biia duua nara*] *us.zaiie’nte* [plur. for dual] . . .
stri-ca na’riias-ca
 ‘[from two men (= a man and woman) two men] are born:
 female and male’ (V. 2.41).

6.3 Grammatical concord

Grammatical concord is of the usual Indo-European/Indo-Iranian type: between nouns and adjectives, pronouns, and numerals; between nouns and appositions and in comparisons introduced by *yaθa* ‘like, than’; and between subject and verb. There are some archaic and novel features:

6.3.1 Gender

Adjectives (including participles), pronouns, and verbs modifying several nouns tend to agree with the nearest one:

OAv. *vī.nənāsā [yā-s/f. kar’pō.tās-cā-s/f. kəuuītās-cā]*
 ‘ruined are [what (are) the words *karapan* and *kauuī*]’ (Y. 32.15);

[*yqm*] *dā . . . [xšnūtəm hiiaṭ ’ruuatəm . . .] taṭ nē vaocā*
 ‘[the *favour-s/f. **which** you (first) established . . . the deal-nt. **which** (you have)] . . .
 speak **that-nt.** to us’ (Y. 31.3).

6.3.2 Dual determinants

Dual forms of determinants and verbs may agree with two singular nouns forming a pair:

YAv. *staomi [maēγəm]-s/m.-ca [vārəm]-s/m.-ca*
[yā]-du/m. tē kəhrpəm [vaxšāiīatō]-3du
 ‘I praise [the fog and the rain,
 which make] your body [grow]’ (Y. 10.3).

To express ‘both X and Y’, Avestan uses the construction ‘both: (namely) X and Y’, where ‘both’ is OAv., OPers. *ubā*, YAv. *uua* or the invariable pronoun *uuaēm* (<*ub^haiam*):

OAv. *ubē-du/f. ha^rruuās-s/f.-cā . . . amər^rtatās-s/f.-cā*
 ‘both wholeness and immortality’ (Y. 34.11);

YAv. *uuaēm x^rar^rθəm-s/nt.-ca vastrəm -s/nt.-ca*
 ‘both food and clothing’ (Y. 55.2);

OPers. *utā vištāspa-s/m. utā aršāma-s/m. ubā ajīvatam-3du*
 ‘Hystaspes and Arsames were both alive’ (XPf 19–21).

6.3.3 Plural subject and singular verb

A neuter plural subject takes a singular verb both in Old and Young Avestan (where we also find plural agreement):

OAv. [*sax^rār^r]*-p/nt. . . . [*yā*]-p/nt. *zī [vā-uuər^rzōi]-3s . . . [yā]-p/nt.-cā [var^ršatē]-3s*
 ‘[the *verses], namely (those) [which have been produced] and (those) [which will be produced]’ (Y. 29.4);

YAv. [*vər^rziitatəm*]-3s-ca *iḍa [vohu vāstriia]-p/nt.*
 ‘and [let good pastures be produced] here!’ (Vr. 15.1),
 but
kuua [tā dāθra]-p pa^rti [haṅjasəṅte]-3p
 ‘where [a^re those gifts *totalled] in turn?’ (V. 19.27).

6.4 Uses of the cases

Cases fall into the generic, but not always clearcut, categories of ‘grammatical’ and ‘other’ cases. The former, nominative, accusative (DO), genitive, dative, are those required by the syntactical structure of the noun phrase and sentence, while the latter include the vocative (address, invocation) and the ‘local’ cases: (local) accusative, ablative, instrumental, locative, which are ‘free/independent’ verb- or sentence-modifying cases and are often accompanied by adpositions.

6.4.1 Nominative

The nominative is used for the subject of any verb and the predicate of ‘to be, become’ and verbs such as ‘to seem; be called, declare (oneself as); be announced, renowned, made/established (as)’, as well as appositions to such nominatives:

with copula:
 OAv. *yūžēm zəuuištiiāṅhō išō*
 ‘you (all are) the fastest invigorants’ (Y. 28.9);

YAv. *hapta həṅti . . . ratauuō*
 ‘seven are the models’ (N. 84);

OPers. *hamaranakara a^hmiy^hušhamaranakara*
 ‘(as) a fighter, I am a good fighter’ (DNb 34);

with verbs of thinking, declaration, etc.:

OAv. *aṭ hōi* [*aojī zaraθuštrō*]
‘thus, [I declare myself to him (as) Zarathustra]’ (Y. 43.8);

YAv. *nmānəm hō* [*manīaēta para.daθō*]
‘[he may think] (of himself as) [giving] a house’ (V. 18.28);

OPers. [*fratarā maniyaiy*] *afuvāyā*
‘[I consider myself superior] to fear’ (DNb 38).

6.4.1.1 The Old Persian nominative naming phrase

In Old Persian, the names of persons and places in narrative passages may be introduced by a name in the nominative + *nāmā* ‘by name’, which agrees in gender with the following noun: name-**nom.** + *nāma* + masc., neut./*nāmā* + fem. and may be regarded as an adjectival compound. The entire phrase may have any syntactic function and even be governed by a preposition; its function may then be made explicit by an anaphoric pronoun or adverb:

I martiya-s/m. [*martiya nāma*]-**s/m.** *ci'caxraiš puça-s/m.*
[kuganakā nāma]-s/nt. *vardanam-s/nt.* *pārsaiy avadā adāraya*
hauv udapataṭā ūvjaiy
‘one man [Martiya by name] son of Cincaxri,
[Kuganakā by name] a town in Persia, **there** he was staying,
he rose up in Elam’

= ‘one man called Martiya, son of Cincaxri, who was staying in the town of Kuganakā in Persia, rose up in Elam’ (DB 2.8–9);

[kāpišakāniš nāmā]-s/f *didā-s/f* *avadā hamaranam akunava*
‘Kāpišakāni by name, a fortress, **there** they fought the battle’
= ‘they fought the battle at a fortress called Kāpišakāni’ (DB 3.60–61);

hacā [*pirāva nāma rauta*]-**s/nt.**
‘from: [Nile by name, a river]’ = ‘from the river Nile’ (DZc 9).

The construction may be found in Young Avestan, but the examples are too few to be certain:

saošiiqs [*vər^oraja nqma*]
‘the Revitaliser [Obstruction-smasher by name]’ (Yt. 13.129).

6.4.2 Vocative

The vocative is used for address and is often used together with an imperative or in questions:

OAv. [*f^orašaoštrā*] *aθrā tū . . . iḏī*
‘[O Frašaoštra], you, **go** there!’ (Y. 46.16);

YAv. *kō* [*nar^o*] *ahī . . .*
azəm ahmi [*zaraθuštra*] *haomō*

‘who, [O man], are you? . . .
I am, [O Zarathustra], the *haoma*’ (Y. 9.1–2);

OPers. [*martiyā*] *hayā auramazdāhā framānā hauv-taiy gastā mā θalaya*
‘[O man], **do not let** Ahuramazdā’s command **seem** evil to you!’ (DNa 56–58).

In Young Avestan, a vocative may be introduced by the particle *āi*:

[*āi miθra vo^uru.gaoiiaō^{te}*]
‘[O Miθra, who provide wide grazing grounds], . . .’ (Yt. 10.42).

A 2nd pers. pronoun in the nominative may be followed by a vocative:

tūm [miθra vo^uru.gaoiiaō^{te}
‘you, [O Miθra, who provide wide grazing grounds], . . .’ (Yt. 10.94);

but one in the acc. by another acc.:

ā [0βā ātrəm] gāraiemī
‘I sing (of) [you, the fire = O fire], a song of praise’ (*Ātaš Niyāyišn* 2).

6.4.3 Accusative

The accusative is used for the direct object and the predicate of the direct object, as well as appositions to these. Two accusatives are found with verbs of calling, asking, taking away, etc. The accusative is used with several prepositions, most of them local with the sense of direction.

6.4.3.1 Direct object and predicate of direct object

OAv. [*ratūš*] *səṅgha^{tī} ārma^{tī}š*
‘Humility is announcing [the models]’ (Y. 43.6);

[*spəntəm*] *a^t [0βā] . . . mōj^{hī}*
‘thus, I (now) think of [you (as) life-giving]’ (Y. 43.5);

YAv. *ā^t yimō [imqm zqm] vī.š^uuuiia^t*
‘then Yima made [this earth] go to the sides’ (V. 2.11);

[*xšaiiamnəm ašauuanəm*] *dāiata^u*
place [in command the sustainer of Order]!’ (Y. 8.5);

OPers. *dārayavauš imam duvarθim . . . akunauš*
‘Darius made [this gate]’ (XPa 11–13);

[*anīyam ušabārim*] *akunavam*
‘[another] I made [camel-riding]’ = ‘I made another (cross the river) riding a camel’
(DB 1.86–87).

6.4.3.2 Direct objects with nouns and adjectives expressing actions

Nouns and adjectives expressing actions frequently take accusative objects:

OAv. *vohū xšaθrəm [va^{rīm} bāgəm] ā^{bī}.ba^{rī}štəm*
‘the good command (is) **the best bringer of** [a well-deserved portion]’ (Y. 51.1);

YAv. *sraošō . . . [driγum] Orātō.təmō*
 ‘Sraoša, **the greatest protector of** [a poor (man)]’ (Yt. 11.3);

OPers. *auramazdā [Ouvām] dauštā biyā*
 ‘may Ahuramazdā be **the one who favours** [you]!’ (DB 4.54–55).

6.4.3.3 Double accusative

Two accusative objects, one personal and one thing, is found with verbs meaning ‘ask for, take away from’:

YAv. *[yānəm vō] yāsāmi*
 ‘I request [from you a request]’ (Y. 65.11);

OAv. *duš.sastiš . . . apō [mā īstīm] °iiantā*
 ‘the one of bad announcing . . . **robs** [me of (my) wish]’ (Y. 32.9);

OPers. *[xšačam taya] gaumāta . . . adīnā [ka^mbūjīyam]*
 ‘[the royal command (of) which] Gaumāta . . . **robbed** [Cambyses]’ (DB 1.43–48).

The second direct object can be an ‘inner object’, which may etymologically related to the verb or not:

vīdāraēm zaraθuštara aom asmanəm . . .
[yim] mazdā vaste [vaγhanəm]
 ‘I held out yonder sky, O Zarathustra,
 [which] Mazdā **wears** (as) [(his) garment]’ (Yt. 13.2–3);

yō [aēšməm] . . . [x^rarəm] ja^hnti
 ‘(Sraoša), who **strikes** [Wrath a wound]’ (Y. 57.10).

6.4.3.4 Accusative with impersonal verbs

The construction of an impersonal verb (or noun + copula) + personal accusative (also known from Latin) is found in Old Persian (cf. MPers. Chapter 4, section 4.3.9.1).

with *vaγnav-* ‘believe’:

mām/Ouvām naiy vaγnavataiy
 ‘me/you it does not believe’ = ‘I/you do not believe’;

with *kāma ah-* ‘wish’:

[auramazdām] avaθā [kāma āha]
 ‘[to Ahuramazdā] thus [the wish was]’ = ‘Ahuramazdā wished it thus’ (XPf 21–22).

6.4.3.5 Adverbial accusative

The accusative is used to express direction with verbs of motion and calling. In Old Avestan, it can be personal or impersonal. In Old Persian, it is used about place whither, alternating with *abiy*, while persons take the preposition *abiy*:

OAv. *[aēš^mməm] hēnduarēntā*
 ‘they would scramble together [to Wrath]’ (Y. 30.6);

YAv. (few certain examples):

a'ime [gātūm] hē nišhiḍaēta
 'he should sit down quietly [in] his [place]' (V. 9.33);

OPers. *avam adam frāišayam [arminam]*
 'him I sent [to Armenia]' (DB 2.30)
 but

adam frāišayam . . . [abiy avam]
 'I sent . . . [to (= against) him]' (DB 3.12–14).

6.4.3.6 Accusative of length of time, space and measure

OAv. *vīspā aiiār³-p/nt.*
 'for all (his) days' (Y. 43.2);

YAv. *hqminəm-ca zaiianəm-ca*
 'throughout both summer and winter' (Y. 65.5 = Yt. 5.5);

[varci'ōim paṇtaqm] azāte
 '(she = the cow) is driven [along the road of captivity]' (Yt. 10.38);

OPers. *[XL arāšnīš] baršnā*
 '[forty cubits] in depth' (DSf 26).

6.4.3.7 Accusative 'of respect'

This accusative (like the instrumental) expresses 'with respect to' and is often used with superlatives and in *figura etymologica* (see section 6.4.7.1c).

OAv. *voḥū uxšiiā managhā . . . [tanūm]*
 'grow [in body] by good thought!' (Y. 33.10);

YAv. *yaṭ as [ašəm] ašauuastəməm*
 'as he was [in Order] the most Orderly' (Yt. 19.79).

6.4.3.8 Accusative with adposition

Numerous adpositions govern the accusative. When they govern more than one case, the function of the accusative is similar to that of the accusative in general (e.g. motion toward):

Av. *adāri* 'beneath'; – *aṇtar³*, *aṇtar³-ca*, OPers. *a^htar* 'between'; – Av. *a'bī*, *auui*, *aoi*; OPers. *abiy* 'to', etc.; – Av. *a'pī* 'after, throughout'; – OPers. *pariy* 'about'; – *paišiyā* 'before'; – Av. *pci'ti*, OPers. *patiy (patiš)* 'on, at, against' etc.; – Av. *para* 'before'; – OPers. *para^h* 'beyond'; – Av. *pasca*, OPers. *pasā* 'behind, after'; – Av. *tarō*, *tarasca* 'across, over'; – Av. *upa*, OPers. *upā* 'in, at, under (during the reign of)'; – Av. *upa'ri*, OPers. *upariy* 'over, on, above, compared to'.

6.4.3.9 Young Avestan nominative for accusative

In Young Avestan, whole phrases in the nominative, especially plural, often occupy positions in the sentence which would require them to be in the accusative. This is no doubt an effect of the oral transmission:

yaṭ bauuāni a'βi.vaniiā
[vispe daēuua mašiiā-ca vispe yātauuō pa'rikās-ca]
 for:
 **[vispē daēuuū mašiiās-ca vispē yātūš pa'rikās-ca]*
 'that I may be an overcomer of = overcome
 [all old gods and men, all sorcerers and witches]' (Yt. 15.12).

6.4.4 Genitive

The main function of the genitive is 'adnominal'. The main adnominal functions are 'possessive' and 'partitive'. The genitive is also used with verbs and adjectives.

6.4.4.1 Possessive genitive

Possession is expressed by the genitive (also in noun clauses, where gen. + 'to be' = 'to have', see section 7.1.6.3). The use of the possessive genitive ranges from material possession to genitive of definition. A possessive genitive can be coordinated with a possessive pronoun. Series of two or three dependent genitives are common:

OAv. *[vaṅhēuš] dazdā [manaṅhō]*
[šiiāoθⁿnanəm] [aṅhēuš] mazdāi
 '(the Model) [of good thought] is established –
 (and) [of the actions] [of the (new) existence] – for the All-knowing one' (Y. 27.13;
 other parsing possible);

YAv. *tum zaraθuštrō [nmānahe] [po^rrušaspae]*
 'you (are) Zarathustra [of (= belonging to) the house] [of Pourušāspa]' (Y. 9.13);

OPers. *[dār^ayavahauš XŠh^ayā] viθiyā*
 'in the house [of king Darius]' (DPi).

The genitive can be governed by the prior member of a compound:

YAv. *[kamarəδō].janō [daēuuanəm]*
 'the [head]-smasher [of the old gods]' (Y. 57.33).

See also section 10.5 Figura etymologica.

6.4.4.2 Subjective and objective genitive

Subjective and objective genitives represents 'transformations' of sentences of the structure SV or OV > N_{gen} N:

OAv. *aēšē [ašahiiā]*
 'in search [of Order]' < 'he seeks order-DO' (Y. 28.4);

YAv. *[ahe] yasna [yazatanəm]*
 'by [his]-S/gen. sacrifice [of = to the gods- DO/gen.]' < 'he-S sacrifices to the gods-DO' (Y. 57.3);

OPers. *xšāyaθiya [dah^ayūnām]*
 'king [of/over lands]' < 'king rules the lands' (DNa 8–11).

6.4.4.3 Defining or descriptive genitive

This genitive is used to define or delimit the scope of a noun. Its most common use is to define a crime of which one is guilty or accused, that for which one must pay penalty, or that which one confesses (the ‘forensic/legal’ gen.), but genitive phrases describing persons or objects are also found:

Descriptive genitive:

YAv. *[vīsō sūraiiā] Oraētaonō*
 ‘Thraetaona [of a vigorous house]’ (Y.9.7);

vazrəm . . . [zarōiš aiiāḡhō] frahixtəm
 ‘a cudgel . . . cast [in tawny bronze]’ (Yt. 10.96).

‘Forensic’ genitive:

OAv. *yadā [aēšqm] kaēnā jama’ī [aēnaḡḡm]*
 ‘when the **retribution** comes [for these sins]’ (Y. 30.8);

YAv. *hāuuāḡta [aētahe šīiaoθnahe] vər^ozīiḡn*
 ‘they should perform equal (penalties) [for this action]’ (V. 15.14).

6.4.4.5 Partitive genitive

The partitive genitive expresses that something is part of a totality or an example of group and is found with nouns (one among, etc.), superlatives or adjectives with implicit superlative meaning (often with *hātqm* ‘among/of (all) those that are’), pronouns, adverbs of place and time, numerals (from ‘twenty’ and up), and other words expressing quantity. It is often used in expressions of time (‘sometime during’), sometimes of place (‘somewhere in’):

OAv. *[hātqm] huuō aojištō*
 ‘he (is) **strongest** [of/among (all) those who are]’ (Y. 29.3);

YAv. *yātauuō [mašīiānqm]*
 ‘sorcerers [among men]’ (Yt. 8.44);

OPers. *VIII manā [taumāyā] tayaiy paruwam xšāyaθiyā āha*
 ‘(there were) eight [of my family] who had been kings before’ (DB 1.8–10).

A ‘free’ partitive genitive (French *du*, etc.) is found with verbs of giving, bringing, partaking, etc.:

YAv. *yaḡ vā mašīiō [mašīianqm xšudranqm] para.gə^oruuaiie’ti*
 ‘or when a man receives [(some) men’s semen]’ (V. 8.32).

6.4.4.6 Genitive with adjectives and verbs

The genitive with adjectives meaning ‘full of, satisfied with’ and with verbs meaning ‘rule (over), be in command of’ is probably a variant of the partitive genitive:

OAv. *yōi [a-sruštōiš] pər^onāḡhō*
 ‘who, **full** [of refusal to listen]’ (Y. 44.13);

YAv. *īm zā bauuaṭ pər³na* [*pasuuqm*]-*ca*
 ‘this earth became **full** [of sheep] and . . .’ (V. 2.8);

OPers. *adam*[-*šām*] *patiyaxšayaīy*
 ‘**I ruled** [over them]’ (DNa 6).

OAv. *kaṭ mōi uruuā isē cahiiā* [*auuaṅhō*]
 ‘does my soul, I wonder, **command** anyone’s [help]?’ (Y. 50.1);

YAv. *naēda-ca pascaēta hāu nā . . . isaēta* [*frašūtōiš*]
 ‘not **may** that man thereafter **be capable** [of going forth]’ (P. 22).

6.4.4.7 Genitive with adpositions

A few adpositions govern the genitive, some of them also other cases:

OPers. *anuv* ‘according to’; – Av. *pa’ti* ‘on, with, in return for’; – *parō* ‘before, in front of’; – *pasca*, OPers. *pasā*; – OPers. *rādiy* ‘on account of’.

6.4.5 Dative

The dative (OPers. gen.-dat.) expresses the indirect object (usually animate or personified inanimate) with transitive verbs, most often accompanying a direct object, in the sense of ‘giving sth. to sb.’, ‘doing sth. for sb.’ (*dativus commodi*), but also with intransitive verbs.

6.4.5.1 Indirect object

When accompanied by a direct object, the dative is usually (but not exclusively) a personal dative of the indirect object:

OAv. [*kahmāi*] *mā* *0βar³zdūm*
 ‘[for whom] have you carved **me**?’ (Y. 29.1);

aṭ mā . . . [*mazōi magāi.ā*] *pa’ti.zān⁴tā*
 ‘so . . . acknowledge **me** [for the great gift exchange]’ (Y. 29.11);

YAv. *yā aom pu0rəm bara’ti* [*aniiahmāi aršānāi*] *varštəm*
 [*pa’0e*] *upa.bara’ti*
 ‘(the evil woman) who bears that son **produced** [for another male]
 (and then) presents (**him**) [to (her) husband] (as his)’ (Yt. 17.58);

OPers. [*aniyahçiyā*] *asam frānayam*
 ‘[for another] I brought **a horse**’ (DB 1.87).

This dative can be governed by a compound:

YAv. *vāstrō.bər³taheca* [*gauue huδā³ḡhe*]
 ‘and of the one by whom **grass** is brought [to the cow giving good gifts]’ (Vr. 1.9).

6.4.5.2 Final dative

The dative of inanimate nouns can express purpose (final dative), often in connection with a verb of movement and together with a dative of the person for whose benefit it happens, the *dativus commodi*:

OAv. *ku0rā [mər'z̄dikāi] axštaŋ*
 'where will (someone) stand by [for (= to show) mercy]?' (Y. 51.4);

ā *mōi [raf'δrāi] zauuēṅ jasatā*
 'come to (my) calls (acc. of goal of motion) [for support] **for me!**' (Y. 28.3);

YAv. *ahmāi jasa'ti [auuajhe]*
 'he comes [for help] **for him**' = 'he comes in order to help him' (Yt. 10.87).

6.4.5.3 Dativus commodi

This dative is found especially with intransitive verbs and is often accompanied by an impersonal dative of purpose:

YAv. *ā'ŋ [yimāi] [xša0rāi] Orisatō.z'ma həṅjasənta*
 'then [for Yima] [for (his) command] three hundred winters came together' (V. 2.8);

[a'βiiō vaŋ'hibiiō] imā zao0rā
 '[for the good waters (are)] these libations' (Y. 22.2);

OPers. *imā dahayāva tayā [manā] *patiyāišd'*
 'these (are) the lands which came [to me]' (DB 1.13).

This dative is also used to refer to future time:

parō.asnāi aŋ'he
 'for (the duration of his) future existence' (Āfr. 3.7).

6.4.5.4 Dative with nouns and adjectives

The same dative is found with nouns and adjectives ('good for', etc.) and interjections (*ušta* 'wished-for-things, good luck', *nəmō* 'reverence/homage to, hail!', *sādrəm* 'bad luck for!', OAv. *auuōi*, YAv. *āuuōiia* 'woe to!'):

OAv. *huuō ["rušāē'biio] spəntō*
 'he is **life-giving** [for the meager ones]' (Y. 29.7);

YAv. *ušta buiāŋ [ahmāi na're]*
 'may there be **wished-for things/good luck** [for this man]!' (Y. 62.1).

6.4.5.5 Dative of agent

The dative is found with verbal adjectives in *-iia-*, *-0a-*, or *-0βa-* (*-tuua-*), meaning 'which should be done **by** somebody':

OAv. *aŋ tā vaxšiiā išəntō*
yā mazdāθā hiiaŋcū [vīdušē]
 'thus, I shall speak, O (you) who wish to come,
 those (words) which **should be paid attention to**,
 even when [by someone who knows] (them)' (Y. 30.1);

YAv. *mi0rō yō nōiŋ [kahmāi] aiβi.draoxdō*
 'Miθra, who (is) not **to be deceived** [by anybody]' (Yt. 10.17).

6.4.6 Ablative

The ablative expresses separation – movement away from; distance from or point of origin seen from nearby (usually with *ā*) – and comparison (‘than’). In Old Persian, these functions of the instrumental-ablative require the preposition *hacā*.

6.4.6.1 Ablative of separation

OAv. *yō īš pāt [dar^osāt] ašahiā*
 ‘which shall **keep** them [from the sight] of Order’ (Y. 32.13);

aṇtar^o vīspāng drəguuatō [haxmāng] °mruīē
 ‘**I am banning** [from (their) following] all those possessed by the Lie’ (Y. 49.3).

Note the special uses of the ablative in the following instances:

YAv. [*aētahmāt puθrāt*] *mimar^oxšaṇ^oha*
 ‘rid yourself [of this child] by destroying it!’ (V. 15.14);

ablative of separation plus *parō* ‘before, away from’:

YAv. *hō nōit . . . frānāmā^{te} [θβaēsāt {parō daēuuaēⁱbiiō}]*
 ‘he does not **yield** [{before the old gods} to (their) hostility]’ (Y. 57.18).

6.4.6.2 Ablative of comparison, difference and preference

This ablative is found with comparatives, with *aniia-* ‘other’ (tends to become a mere particle ‘other than’), and with verbs:

comparison:

YAv. *vātō . . . hubaoⁱdiš hubaoⁱ dītarō [aniiaēⁱbiiō vātaēⁱbiiō]*
 ‘a fragrant wind **more fragrant** [than other winds]’ (H. 2.7);

OPers. *fratarā maniyaiy [afuvāyā]*
 ‘I consider myself **superior** [to fear]’ (DNb 38);

with *aniia-*:

OAv. *aniiēm [θβahmāt āθras-cā manāḥas-cā]*
 ‘other [than your fire and thought]’ (Y. 46.7);

YAv. *aniiō [θβaṭ yaṭ zaraθuštrāt]*
 ‘(to anybody) other [than you, Zarathustra]’ (V. 2.2).

The ablative may be governed by the first member of a compound:

YAv. *uyra.zaošā [tbišiiānbiiō]*
 ‘those with **strong(er)** desires [than those hostile (to us)]’ (Yt. 13.31).

6.4.6.3 Ablative of time or place throughout which

In Young Avestan, the ablative can express time or place throughout which or all the way up to, often with the particle *-a* (see section 3.1.5.3):

[*asnā́t-ca xšafnā́t-ca*] . . . *āpō auua.barəntē*
 ‘[day and night] . . . the waters pour down’ (Yt. 5.15);

drujō [vaēsmənd-a] azəmnaqm
 ‘being led [(all the way) to the entrance hall] of the Lie’ (Yt. 10.86; see de Vaan 2001).

6.4.6.4 Ablative with adpositions

Several adpositions govern the ablative, some of them also other cases; the function of the ablative is then similar to that of the ablative in general (e.g. motion away):

OAv. *a’bī* ‘as far as . . . is concerned, to the extent one can’; – YAv. *adā’ri* ‘(just) under’; – OPers. *anuv* ‘along, on (river)’; – Av. *ā*, YAv. *-a* ‘all the way to, at (*chez*), out of, throughout’; – *haca* ‘from, in accordance with, on the part of’ (also with ‘fear’); – OAv. *hanar* ‘without’; – YAv. *maṭ* ‘with’; – *pa’ri* ‘from, away from, without’; – *pa’ti* ‘from on the top of, at, beside, without’; – Av. *parā* ‘before’; – YAv. *parag’t* ‘aside from’: – Av. *parō*, *parē* ‘before, in front of, from, because of’; – YAv. *pasca* ‘after’; – *upa’ri* ‘in, throughout’.

6.4.6.5 Young Avestan ablative = genitive

There are few examples in Young Avestan of genitives used as ablatives (other than of *a*-stems). Whether these are survivals of the old ablative forms or errors cannot be verified:

haca [hauuaitiās³ tanuuō]
 ‘away from [one’s own body]’ (Y. 10.5) [cf. *auui x³aēpaiθiiās³ tanuuō-p/acc.* in section. 6.5.4.1];

fraš [aiiaṅhō] frasparaṭ
 ‘he jumped forth (away) [from the metal (pot)]’ (Y.9.11).

6.4.7 Instrumental

The instrumental expresses means/instrument and accompaniment.

6.4.7.1 Instrumental of means

This is one of the most common cases in the Avestan texts, and, especially in Old Avestan, it is often difficult to decide which nuance is intended. In Old Avestan, the instrumental of means can also be used of (divine) persons.

This instrumental is also used of space through which (the road along which) on travels; to express reason or cause; measure and price; and as the instrumental of respect:

OAv. *kā [yā] mā uxšiiē’ti nər³fsa’ti θβaṭ*
 ‘who (is he) [through whom] the moon is (now) first waxing then waning’ (Y. 44.3);

YAv. *yō aēšməm [stər³θβata sna’θiša] . . . ja’nti*
 ‘he who smites Wrath [with a stunning weapon]’ (Y. 57.10);

OPers. *imā dahayāva [tayanā] manā [dātā] apariyāya*
 ‘these lands behaved [by/according to (that) which (is)] my [law]’ (DB 1.23).

6.4.7.1a *Instrumental of space through which*

OAv. *tēm aduūānəm . . . daēnā . . . [yā hū.kəʳtā] . . . "ruuāxsat*
 ‘that road . . . [along which, well made], the vision-souls have walked’ (Y. 34.13);

YAv. *huuar³ [auua paθa] aēti*
 ‘the sun goes [along yonder path]’ (Yt. 13.16);

OPers. *[viθbiš-cā] tayā-dīš gaumāta haya maguš adīnā*
adam kāram gāθavā avāštāyam
 ‘and [throughout the houses/towns] which Gaumāta the Magian had taken from them,
 I settled the people (back) in its place’ (DB 1.65–66).

6.4.7.1b *Instrumental of reason and cause*

OAv. *vaijəuš [əuuištī] manaŋhō*
 ‘[on account of not finding] a good thought’ (Y. 34.9);

YAv. *yō gāθā asrāuuaiiō [qsta vā tarō.ma’ti vā] tanum piriie’ti*
 ‘he who (goes) without reciting the Gāθās [out of evilness or scorn] forfeits his body’
 (N. 41).

6.4.7.1c *Instrumental of respect*

This instrumental (like the accusative) expresses ‘with respect to’ and is often used with superlatives and in *figura etymologica* (cf. section 6.4.3.7):

OAv. *ārmātiš vaxšt / [utaiiū’tī təuuištī]*
 ‘Humility (the earth) has grown [with respect to *texture (and) *tension]’ (Y. 34.11);

YAv. *[ama] ahmi amauuastəmō*
 ‘[in force] I am the most forceful’ (Yt. 14.3);

OPers. *utā [abiš] nāviyā āha*
 ‘and it (= the Tigris) was *deep [with waters] (= in spate)’ (DB 1.85–86).

6.4.7.1d *Measure and price*

OAv. *f³rašəm [vasnā] ha’θiiēm dā ahūm*
 ‘you have (now) made the existence “juicy” (filled it with vitalising juices) [in (exchange) value]’ (Y. 34.15);

YAv. *imqm zqm vī.šāuuaiiat [aēuua θrišuuu] ahmāt masiiehīm*
 ‘he made this earth go apart (to make it) [by one-third] larger’ (V. 2.11);

OPers. *XL arāšnīš [baršnā]*
 ‘forty cubits [in depth]’ (DSf 26).

Here may also belong the use of the instr.-abl. *raucabiš* ‘days’ in the Old Persian dating formula:

month name (X) in gen.-dat. + *māhəyā* ‘of the month’ + numeral (Y) + *raucabiš*
θakatā āha
 ‘of the month of X, **by Y days**, (the days) had passed’ = ‘on the Yth of X’,

but
 month name (X) in gen.-dat. + *māhəyā 1 rauca^h 0akatam āha*
 ‘of the month of X, one day had passed’ = ‘on the first of X’.

6.4.7.2 Instrumental of accompaniment and dissociation/deprivation

This instrumental is used with verbs of motion, notably *haca-* (mid.) in the sense of ‘be accompanied by, be together with, unite/mingle with’, as well as with other verbs and expressions for ‘togetherness, union’, etc. In Old Persian, this function requires the preposition *hadā*:

6.4.7.2a Accompaniment

OAv. *a0rā tū [ar^odrāiš] idī*
 ‘you, go there [with the (heavenly) arbiters]!’ (Y. 46.16);

YAv. [*x^rar^o0aēⁱbiiō*] *pascaēta āstaiiānta*
 ‘afterward they should approach [with foods]’ (V. 3.18).

6.4.7.2b Dissociation/deprivation

OAv. *parō vā [vīspāiš] °vaox^omā [daēuuāiš]-cā*
 ‘we have (always) said you (are) beyond [others], both [old gods] and . . .’ (Y. 34.5);

YAv. *vī [daēuuāiš] . . . sarəm mruuīē*
 ‘I forswear the company [with the old gods]’ (Y. 12.4).

6.4.7.3 Instrumental with adpositions

A few adpositions govern the instrumental, some of them also other cases: Av., OPers. *hadā/hada* ‘(together) with’; – Av. *ha0rā* ‘in the same place as, together with’; – *maṭ* ‘(together) with’; – *pa^ti*, OPers. *°patiy* ‘at, close to, in, throughout’; – *pasca* ‘after’.

6.4.7.4 Case replacements

In Young Avestan, the instrumental plural commonly takes the dat.-abl. ending *-biiō*:

yō yuⁱdiēⁱti . . . [hada daēuuāēⁱbiiō]
 ‘who fights . . . [with the old gods]’ (Y. 57.17).

In Young Avestan, the nominative-accusative plural forms of neutre (*r/*)*n*-stems commonly take attributes in the instrumental plural (a still unexplained peculiarity of YAv.):

yim [vispāiš]-p/instr. pa^tišmarənte [yāiš] spəntahe mainiiēuš [dāmaq] -p/nom.
 ‘whom [all the creatures] of the Life-giving Spirit recall’ (Yt. 8.48).

6.4.8 Locative

The locative indicates time and place in/at which (when?, where?). It also indicates motion into or inside (also ‘wishing sth. upon sb.’). A special use of the loc. is the ‘loc. of emotion’ (‘to my sorrow’; see Kellens and Pirart 1990, 43).

6.4.8.1 Locative of place and time

OAv. [0βahmī] ā [daqm] nipājēhē
‘you are (now) guarding (it) there [in your abode]’ (Y. 49.10);

YAv. x’ahmi daqm x’ahmi ciθre
‘in his own house, in his own lineage’ (Vr. 14.2);

[fraiia’re aiiqn] bauua’ti hubaδrō
‘[On a future day] he becomes lucky’ (Aog. 53);

OPers. pasāva-diš auramazdā manā [dastayā] akunauš
‘then Ahuramazdā placed them [in] my [hand]’ (DB 4.35).

6.4.8.2 Locative of prize won

The locative is used to indicate things to be won in competition, especially with the verbs *zā-* ‘to leave (the competitor behind) at = in (the race for) X’ and *yū’diia-* ‘to fight (for)’:

OAv. yōi zazənti [vaḡhāu srauuahī]
‘who shall be leaving (the others) behind = win [in (the race for) good renown]’
(Y. 30.10);

YAv. zazuš [vīspaēšu vaḡhušu]
‘having won [in (the race for) all good things]’ (P. 26);

tā yū’diie’nti pəḡanāhu [hauue asahi šōiθraēca]
‘(the fravashis) fight in battles (each) [for (her) own place and settlement]’ (Yt. 13.67).

6.4.8.3 Locative with adpositions

A few adpositions govern the locative, some of them also other cases:

OAv. a’bī ‘in, on, regarding’; – OAv. ā ‘in, on’; – Av. ākā ‘in the presence of, in view of’ (Skjærvø 2005a: 203–5); – OAv. pa’rī ‘among’; – Av. pa’tī ‘on; at, in return for’; – YAv. upa ‘in’.

6.5 Uses of pronouns and pronominal forms

6.5.1 Personal pronouns

6.5.1.1 1st and 2nd persons, tonic forms

The tonic (independent, stressed) forms of the personal pronouns are used in Old Avestan for emphasis and contrast; in Young Avestan, they are used more liberally; in Old Persian, personal pronouns are rarely omitted, and only in sequences of clauses at least one of which contains the pronoun.

OAv. [azəm] -cī . . . [0βaqm] mējēhī . . .
‘[I] for my part (now) consider [you] . . .’ (Y. 29.10);

YAv. [tūm] nō āθraom zaota-ste . . . [azəm] . . .
‘[you], O priest, (are ordered) to be our libator . . . (while) [I] . . .’ (Vr. 3.7);

OPers. [*adam*] *dārayavauš xšāyaθiya vazaraka*
 ‘[I] (am) Darius, the great king’ (DB 1).

6.5.1.2 3rd person

Of the three 3rd-person pronouns *ha-lta-*, *i-*, and *a-*, the first two have only nom. and acc. forms, the third supplies the other oblique forms. *ha-lta-* is often difficult to distinguish from the far-deictic demonstrative pronouns. The classification of the pronoun *huuō* is often unclear.

6.5.2.2a *ha-*, *ta-*

The forms of *ha-lta-* are usually emphatic and are often found in initial position followed by the emphatic *zī*, as antecedent to a relative pronoun, and as emphatic anaphoric. They are found in second position mainly in questions and after sentence-initial particles:

OAv. [*huuō*] [*tēng*] *frō.gā . . . caraṭ*
 ‘[he] shall make [them] the lead bulls’ (Y. 46.4).

When used as antecedents for relative pronouns, they may be separated from the relative or precede it immediately (see section 8.2.2):

YAv. [*hō*] *zī asti . . . yō ahurō mazdā*
 ‘for [he] is . . ., (he) **who** (is) Ahura Mazda’ (Vr. 2.7);

kō [*hō*] *aṅhaṭ . . . yō . . .* [*hō*] *bā aṅhaṭ . . . yō*
 ‘**who** may [*he*] be . . . **who** . . . will be [*he*] . . . **who**’ (V. 7.78–79).

On OPers. *hauv/ava-*, see section 6.5.3.4.

6.5.2.2b *The oblique pronominal stem a-*

This pronoun is commonly used as a simple enclitic anaphoric, also reflexive, but it also refers emphatically to an object in contrast to another. In the latter usage it may be difficult to distinguish it from its use as the oblique stem of the near-deictic demonstrative pronoun *ima-*:

OAv. [*ahmāi*]-*cā xšaθrā jasaṭ*
 ‘and [to him/this one] he shall come with command’ (Y. 30.7);

YAv. *sruuaēna* [*aēšqm*] *safūṅhō zaranīia*
 ‘benailed are [their] hooves, golden’ (Y. 57.27).

6.5.2.2c *i-*, *hi-/ši-*, *di-*

This pronoun can be used proleptically for a noun (pronoun) or for direct speech or anaphorically; the noun may be gapped:

OAv.,

- proleptic:

yōi [*īm*] *f³rašēm kər²naon ahūm*
 ‘(those) who shall make [it] perfect, (this) existence’ (Y. 30.9);

- anaphoric:

aṭ [hī] [aiiā] frauar²tā
‘but [she] chose among [those two]’ (Y. 29.7);

- both proleptic and anaphoric:

nōiṭ tā [īm] xšnāuš vaē piō . . . zaraθuštrəm . . .
hiiaṭ [hōi] [īm] . . . zōiš²nū vāzā
‘the *trembler did, for (all) that, not favour [him], Zarathustra,
when [his] two draught animals (were) *angry at [him]’ (Y. 51.12);

YAv. *pascaēta [diš] fraspaie²ti miθrō*
‘then Miθra casts [them (masc.)] away’ (Yt. 10.43);

OPers. *xšačam . . . adam[-šim] gāθavā avāstāyam*
‘the royal command . . . I set [it] down in its place’ (DB 1.61–63);

abicariš . . . tayā[-diš] gaumāta . . . adīmā
‘the pastures that Gaumāta had taken [from them]’ (DB 1.64–66).

OPers. sing. *-šim* is also used as instr. -abl.:

kāra[-šim] hacā daršam atçrsa
‘the army/people feared [him] strongly’ (DB 1.43–53).

6.5.1.3 Enclitic pronouns 1st, 2nd and 3rd person

Enclitics commonly follow the first word in a clause or metrical unit (‘verse line, half-line’), less commonly they are placed at the end of a metrical unit. In clauses with elements raised into the first position, the enclitic may remain in its place. Several enclitics may follow one another. In Old Persian, the position of enclitics is relatively free.

6.5.1.3a Nominative

OAv. *mīždəm mazdā yehiiā [tū] daθrəm*
‘the fee whose depository [you] are, O Mazdā’ (Y. 34.13);

aṭ [yūš] daēuuā vīspāṅhō akāṭ manaṅhō stā ciθrəm
‘but [you], O old gods, are all the *seed (issued) from an evil thought’ (Y. 32.3);

YAv. *ā [tū] mē aētaiiā zaoθraiiā fraṅ²harōiš*
‘then may [you] partake of this my libation’ (Yt. 5.91).

6.5.1.3b Accusative

Accusative forms are distinct from genitive-dative forms in Old Avestan, but Young Avestan uses the genitive-dative forms of the 1p and 2p, occasionally also the 3s, pronouns as accusative. In Young Avestan and Old Persian, the tonic forms of the 1st and 2nd pers. forms are also used as enclitics:

OAv. *ašā kaṭ [θβā] dar²sānī*
‘shall I see [you] through Order, I wonder’ (Y. 28.5) [*ašā* raised];

āuiš [nā] aṅtar² hēntū . . . rātaiiō
‘let gifts appear among [us]!’ (Y. 33.7)

YAv. *aoi [maqm] . . . stūdi yaθa [mā] . . . stauuq̄n*
 ‘praise [me] . . . like they will praise [me]!’ (Y. 9.2);

mā ciš paʹruuō buʹdiiaēta [nō]
 ‘may no one notice [us] first!’ (Y. 9.21) [octosyllabic verse];

OPers. *mā-taya[-mām] xšnāsātiy*
 ‘lest he recognise [me]’ (DB 1.52).

6.5.1.3c Genitive-dative

OAv. *sraotā [mōi] mərʹzđātā [mōi]*
 ‘listen to [my] . . .! be merciful for [my] . . .!’ (Y. 33.11);

YAv. *ašaiia [nō] paʹtī.jamiāi*
 ‘may he in (re)turn come [to us] in Orderly fashion!’ (Y. 7.24);

OPers. *vašnā auramazdāha utā[-mai y]*
 ‘by the greatness of Ahuramazdā and [me]’ (DB 4.46).

6.5.1.3d Ablative

OAv. *yā [0βaɪ] . . . asruštīm . . . yazāi apā*
 ‘I who shall sacrifice away [from you] lack of listening’ (Y. 33.4);

OPers. *tayā hacā[-maʹ] atarsa*
 ‘. . . which feared [me]’ (DPe 9).

6.5.2 Possessive pronouns

Only Old Avestan has genuine possessive pronouns; in Young Avestan and Old Persian, the genitive of the personal pronouns is used:

OAv. *[mōʹruuā] gēuš-cā aziiā*
 ‘[my breath-soul] and that of the fertile cow’ (Y. 29.5);

nōiɪ [nā manā] . . . hacīntē
 ‘[our (= your and my) thoughts] are not in agreement’ (Y. 45.2);

tēm [ahmākāiš azdabīš]-cā uštānāiš-cā yazamaidē
 ‘to him we are sending our sacrifice together with [our bones] and life breaths’
 (Y. 37.3);

YAv. *[yūšmākəm yasnāi]-ca . . . [ahmākəm hauuq̄ʹhāi]-ca*
 ‘for [your sacrifice], for [our] *well-being’ (Y. 14.1).

6.5.3 Demonstrative pronouns

The two demonstrative pronouns *ima-* and *aua-* denote near and far deixis, respectively, that is, what is near the speaker in space and time, both in his immediate surroundings, as opposed to farther away, and to his cosmic place on earth, as opposed to in heaven; the waters and fires on earth, as opposed to those in heaven, etc. The two are often contrasted in the texts:

OAv. [imā] raocā bar²zištəm . . . bar²zimanəqm [auuaṭ]
 '[these] lights (= the ritual fire) . . . [yonder] highest of heights (the sun)' (Y. 36.6);

YAv. [iməqm]-ca zqm yazama'de [aom]-ca asmanəm yazama'de
 'and we sacrifice to [this] earth, and we sacrifice to [yonder] sky' (Yt. 13.153);

OPers. haya [imām] būmim adā haya [avam] asmānam adā
 'who set in place [this] earth, who set in place [yonder] heaven' (DE 2-4).

6.5.3.1 The near-deictic *ima-la-*

In addition to its local-temporal deixis referring to objects in the world, this pronoun can refer to on-going, present events, for instance, those taking place in the sacrifice; it is often accompanied by 1st person references. It also refers to the just-mentioned or to the following:

YAv. [iməm] haoməm . . . āiiese yešti
 'by my sacrificing I harness [this] haoma' (Y. 22.1-2);

ušta buiāṭ [ahmāi] na're
 'may there be wished-for things [for this man]' (Y. 62.1);

aṅtar²-ca druuaṅtəm āmrūta [aiia] aṅtar².uxti
 'he forswore the Lieful One with [this] forswearing: . . . ' (Y. 19.15);

pa'ti tē . . . [imā] sruuā vaēdaiemi
 'to you I exhibit [these] nails' (V. 17.9);

OPers. mām auramazdā pātuv hacā gastā
 utā-maiy viθam utā [imām] dahayāum
 'may Ahuramazdā protect me from evil,
 both my house and [this] (my) land!' (DNa 51-55);

ava [ahayāyā] dipiyā naiy nipištam
 'that is not written in [this] inscription' (DB 4. 47-52).

6.5.3.2 The near deictic *aēša-laēta-*

In its general use in the ritual and narrative texts, this pronoun overlaps to some degree with *ima-la-*:

YAv. [aētahmi] aṅhuuō yaṭ astuua'dṅti
 'in [this] bony existence' [the existence of living beings with bones];

[aēta] vaca maḍaiiaṅ'ha
 'intoxicate yourself with [this] word!' (Vr. 8.1).

It is frequently used with 2nd-person deixis:

YAv. [aēša] zaoθra pa'ti.jamiāṭ tauua ahurāne ahurahe
 'may [this] libation (of yours) arrive, yours, O lady of the Lord' (Y. 68.1);

[aētaē]-ca rē vacō
 'and [these] words of yours' (Yt. 14.46).

The majority of occurrences of this pronoun is in the legal texts. Most of the contexts are of the type ‘how should this/these X(s) behave in the case of this Y? The meaning of the pronoun is therefore very often ‘the X in question, this X we are discussing, the aforementioned X’. The pronoun seems sometimes to be close to *ha-* in function, cf. *hō zī asti* and *aēšō zī asti*, probably with just a slight difference in deixis:

YAv. *caiiō āt [aēte] raocā aghən . . .*
yōi auuaθa ā.raocaiie'nte [aētaēšuuā] var³fšuuā
yō yimō kər³nao t
 ‘but which were [these] lights
 that shine hither in that way in [these] *enclosures
 that Yima made?’ (V. 2.39);

OPers. *mām auramazdā pātuv . . .*
[aita] adam auramazdām jadyāmiy [aita]-mai y auramazdā dadātuv
 ‘may Ahuramazdā protect me . . .
 [this] I ask Ahuramazdā for; may Ahuramazdā give me [this]!’ (DNa 51–55);

contrasting *aēta-* vs. *auua-*:

YAv. *mā mē [aētaiiā] zaoθraiiā fraṇ³harəntu . . .*
nōit [auuā] zaoθrā pā'ti.vise yā māuuōiia fraṇ³harənti
 ‘let them not partake of [this] libation of mine . . .
 I do not accept [those] libations of which they partake for my sake’ (Yt. 5.92–93).

6.5.3.3 The demonstrative pronoun *huuō*, *ana-*

The pronoun *huuō* can have 2nd-person deixis (cf. Jamison 1992, Watkins 2000):

OAv. *“ruuāzištō [huuō] nā . . . pā'ti.jamiiā*
 ‘[you there], who gladden (us) the most, may you *in return come to us’ (Y. 36.2).

The pronominal stem *ana-* may have 2nd-person and 3rd-person deixis, with contextually derogatory reference:

OAv. *[anāiš] vā nōit . . . ašəm-cā yānāiš zar⁴naēmā*
 ‘may we not anger you, as well as Order and (your) thought, the best, with [those] requests (to you)’ (Y. 28.9);

[anāiš] ā manahīm ahūm mər³ngduiē
 ‘by [those] actions/utterances [(of yours)] you are destroying (here and now) the existence of thought’ (Y. 53.6);

YAv. *aθā azēm-cīt . . . daēuuāiš sarəm vīmruuiē*
yaθā [anāiš] viāmruuītā
 ‘and thus do I too renounce union with the old gods,
 like (Zarathustra) used to renounce (union) [with them]’ (Y. 12.6).

6.5.3.4 The far-deictic *auua-*

In addition to its local-temporal deixis (see on *ima-* and *aēta-*, above), this pronoun can refer to well-known facts in the other world. In Old Persian, it also serves as emphatic or anaphoric 3rd-person pronoun (see also section 6.4.1.1):

OAv. *vīduuā [auuqm] yā īm aṅhaṭ apēmā*
 ‘knowing [yonder] (requital) that shall be the last (to reach?) him’ (Y. 44.19);

YAv. [*hāu*] . . . *huraoda jasāti*
 ‘(then) [*she*] shall come, well-shaped’ (V. 19.30);

[*aom*] *ga’rīm yazamaide yim ušadqm ušidar’nəm*
 ‘we sacrifice to [yonder] mountain: the ridge Crack of Dawn’ (Yt. 1.28).

OPers. *hauvlava-* is both fat-deictic and anaphoric, possibly with emphasis:

utā pārsam utā mādam . . . [hauv] āyasatā
 ‘(it was) [*he*] (who) took both Persia and Media’ (DB 1.46–47);

tayaīy paruvā xšāyaθiyā yātā āha
[avaīšām] avā naiy astiy kartam yaθā manā . . . kartam
 ‘the former kings, for as long as they were –
 [*they*] have not done as much as *I* have done’ (DB 4.50–52).

6.5.4 Reflexive pronouns

Examples:

OAv. *yēṅg [x’ē] "ruuā [x’aē]-cā xraodaṭ daēnā*
 ‘whom [their own] breath-soul and [their own] vision soul will make *shudder’
 (Y. 46.11);

YAv. [*x’ā*] *daēna nis’rinuiiāt acištāi aṅhe*
 ‘may [his own] vision soul give him over to the worst existence!’ (V. 5.62);

aoi nmānəm yim [x’aēpa’θim]
 ‘to [(my) own] house’ (Yt. 5.63);

OPers. [*uvaipašiyahayā*] *daršam xšayamna a^hmiy*
 ‘I am firmly in command [of myself]’ (DNb 14–15);

utā pārsam utā mādam . . . [uvāipašiyam] aku’tā
 ‘he made both Persia and Media [his own]’ (DB 1.46–47).

6.5.4.1 Use of *tanū-* ‘body’ as ‘self’

In all three languages, *tanū-* ‘body’ is used in the sense of ‘own body, self’:

OAv. *auuaēnatā . . . āuuar³nā vīciθahiiā narēm narəm*
[x’axiiāi tanuiē]
 ‘observe . . . the *preferences of discrimination (made) man-by-man
 [for his own body]!’;

YAv. *auui [x’aēpa’θiis³] tanuuō]*
 ‘upon [(their) own bodies]’ (Yt. 10.23);

OPers. *dārayavauš haya manā pitā pasā [tanūm] mām maθištam akunauš*
 ‘my father, Darius, made me the greatest after [(him)self]’ (XPf 30–32).

6.5.5 *Reciprocal expressions and 'other'*

Reciprocity is usually expressed by repeating *aniia-* 'one . . . (an-/the) other' (cf. Jamison 1997; see also section 4.6.10):

OAv. *ašā vā [aniiō d'nim] vūuəŋhatū*
 '(travelling) through Order, let [one] of you *illuminate/overcome [the other]= [one another]!' (Y. 53.5);

YAv. *yaēšqm [aniiō aniehe] uruūānəm aiβi.vāēna'ti*
 'of whom [one] looks at the soul [of the other]' (Yt. 13.84);

OPers. *[aniya aniyam] naiy ja'tiy*
 '[the one] does not kill [the other]' = 'they do not kill one another' (DSe 35–36);

[aniyam] ušabārim akunavam [aniyahçiyā] asam frānayam
 '[the one = some] I mounted on camels, [for the other = for some] I brought a horse' (DB 1.85–87).

6.5.6 *Interrogative pronouns*

Examples:

OAv. *[ciš] ahī [kahiiā] ahī*
 '[who] are you? [whose] are you?' (Y. 43.7);

YAv. *[kō] maṃ yazd'te . . . [kahmāi] tanuuō druuatātəm azəm baxšāni*
 '[who] will sacrifice to me? [to whom] shall I give health of (his) body?' (Yt. 10.108).

6.5.7 *Indefinite pronouns*

Examples:

OAv. *aṭ huuō . . . [kas-cī] ašā hužēntuš*
 'thus, he there . . . [whoever] by (his) Order (is) of good lineage' (Y. 49.5);

YAv. *[kqm-ci] vā hubao'δitəmanqm "ruuaranqm*
 'or [whichever] of most fragrant plants' (V. 8.2);

[ka]jhe ka]he] apayžā're
 'in [each and every] outlet' (Yt. 5.101–2);

[kahe kahiiā-cī] druuatqm
 '[of each and everyone] of the ones possessed by the Lie' (Y. 61.4);

pita puθras-ca . . . [kataras-ci]
 'father and son [each]' = 'both father and son' (Y. 9.5).

With relatives:

OAv. *tācā . . . ašāunē cōiš*
[yā] zī [cīcā] vahišā
 'and **those** (things) you assigned to the sustainer of Order, namely, [whichever] are the best ones' (Y. 47.5);

uštā ahmāi

[*yahmāi*] *uštā* [*kahmāicī*] . . . *mazdā dāiiā* *ahurō*

‘wished-for (things are) in the wish **for him**,
[to whomever] Mazdā Ahura shall give (them)’ (Y. 43.1).

Negated:

OAv. *aēšcəm aēnaḥqəm* [*naē-cī*] *vīduuā* *aojōi*

‘of these sins I declare I know [none] (at all)’ (Y. 32.7);

[*mā ciš*] *a* *vē* . . . *gūštā*

‘but let [no one] among you keep listening to’ (Y. 31.18);

OP. [*naiy*] *āha martiya* . . . [*naiy*] *amāxam taumāyā* [*kaš-ci*] *haya* . . .

[*kaš-ci* *naiy*] *adārsnauš* [*ciš-ci*] *θaⁿstanaiy*

‘there was no man . . . nor anybody of our family, who . . .
nobody dared say anything’ (DB 1.48–49, 53).

Without particle:

OAv. *kadā vaēdā yezi* [*cahiā*] *xšaiiaθā*

‘when shall I know whether you are in command [of anything/anybody]’ (Y. 48.9).

6.6 Uses of the numerals

The numerals ‘100’ and higher take the genitive plural of a following noun:

YAv. [*catagrō parəna*] *vīdāraiiōiš auui paθəm kataras-ci*

‘you should hold out [four feathers] toward each of the (four) roads’ (Yt. 14.44);

nauuaitīm [*upāzananəm*] *upāzōi* *aspāhe aštraia*

‘he shall administer ninety [strokes] with the horse-whip!’ (V. passim).

The nom.-acc. sing. of ordinals (in OPers. with *patiy*) is used in the meaning of ‘for the -th time’:

OAv. *nōi* [*d^mbitīm*] . . . *ahūm mərəšūiāt*

‘not [a second time] should he destroy the existence’ (Y. 45.1);

YAv. [*pa^rrīm/bitīm/θritīm*] *x^rarənō apənəmata*

‘[a first/second/third time] the Fortune turned away’ (Yt. 19.35–37);

OPers. [*patiy čitīyam/duvitīyam*] *hamiçiyā haⁿgmatā*

‘the conspirators having gathered [for the third/second time]’ (DB 2.43, 57–58).

6.7 Uses of comparative and superlative

6.7.1 Comparative

Comparison (‘than’) is expressed by comparative + ablative (see section 6.4.6.2) or comparative + *yaθa* (see section 8.2.3.6).

6.7.1.1 Comparative + positive in contrasting pairs

The comparative is used together with the positive of another adjective to form a contrasting pair, in which the first member in the comparative is depicted as superior to the second member in the positive: ‘X’er and Y’ = ‘X and, on the other hand/in contrast, (the poorer) Y’ (Humbach 1991, I § 15.2.1):

OAv. *yaiiā [spaniiā] -comp. *u’tī mrauuat yēm [aṅrəm] -pos.*
 ‘of which two, [the life-giving one] shall tell (him),
 whom (you know as) [the Evil one]’ (Y. 45.2);

OPers. *haya [tauvīyā] -comp. tayam [skauθim] -pos. naiy ja’tiy*
 ‘[the mighty one] does not kill [the poor one]’ (DSe 37-44).

6.7.1.2 Comparative in compounds

The comparative is found in an elliptic type of compounds: ‘. . . -er than (those of) X’:

YAv. *bāzauua *auruša aspō. [staoiehīš]*
 ‘arms, white (and) [thicker] (than those) of a horse’ (Yt. 5.7).

As the first member of a compound, a comparative may be replaced by the positive, but maintain its rection:

YAv. *[uyra].zaošā [tbišūaδbiiō] -abl.*
 ‘(the fravashis) with [strong(er)] desires [than those hostile] (to us)’ (Yt. 13.31);

auuā dāmaq . . . yā hānti [paōriiō].dāta [paōriiō].fraθbaršta
ašnāat-ca apāat-ca
 ‘those creations that (were) those established [first], fashioned forth [first]
 (= earlier) than the sky, the water . . .’ (Vr. 7.4) [cf. *para, parō* ‘before’ + abl., see section 6.4.6.4].

6.7.2 Superlative

The superlative is most commonly used to present something as possessing a quality in the highest degree, ‘the most A’, or to single out one thing/person from others of the same class: ‘A is the biggest of all A’s’ (see section 6.4.4.5).

The superlative is also used in contrast to an adjective in the positive or to depict something (in the superlative) as vastly superior to another thing (in the positive; Humbach 1991, I § 15.2.2):

OAv. *yā drəguuā -pos. acištā vərəziiō /*
ašəm maniuš [spəništō] -sup.
 ‘you, [who are possessed by the Lie], would perform the worst (words/actions);
 [the most life-giving] spirit . . .’ (Y. 30.5);

yā vā [kasōuš] -pos. aēnaghō ā [mazištəm] -sup. °iiamaitē būjim
 ‘or (someone) who for a [small] sin shall incur the [greatest] expiation’ (Y. 31.13).

7 SYNTAX III: THE VERB

The Iranian verb has the categories of tense and aspect, mood, and voice. Finite forms have the categories of number and person, while non-finite forms behave like nouns (infinitives) or adjectives (participles).

The negations, Av. *nōiŋ*, OAv. *naēdā*, YAv. *naēda*, *nauuu*, OPers. *naiy*, mainly negate statements and *mā* (*māda*) commands and exhortations.

7.1 Tenses

The main difference in syntax between Old Avestan, on the one hand, and Young Avestan and Old Persian, on the other, is in the use of the tenses, while that of the moods is largely the same.

The Old Avestan verbal system is based upon the opposition of the present (imperfective) aspect, which is indefinite (durative, repetitive), vs. the aorist (perfective) aspect, which is definite (punctual, ingressive, terminated); that of Young Avestan and Old Persian is based mainly upon the opposition present vs. past.

The common past narrative tense is the present injunctive (augment-less) in Young Avestan and the augmented imperfect in Old Persian.

In Avestan, an action can be characterised as definitely having taken place in the past(?) by the optional use of the augment. In Old Persian, the augment is an intrinsic part of past tenses.

The state reached after a past event is expressed in Avestan by the perfect in its old function of resultative-stative.

In Old Avestan, the aorist expresses, on the one hand, anteriority in relation to the present or preterite, on the other punctual or 'immediate', both incipient and concluding, action vs. ongoing or unfinished action, expressed by the present.

In both Young Avestan and Old Persian, repeated or habitual past action (cf. English 'he would go') can be expressed by the optative. It then takes the augment, regularly in Old Persian and occasionally in Young Avestan (see section 7.2.2.2).

Thus, all three Old Iranian languages possess the following basic tenses:

- *present*: 'he does, he is doing';
- *preterite*: 'he did, he was doing';
- *perfect*: 'he has (always, never) done' or 'he did/it was done/it happened and now is' (*present perfect*);
- *pluperfect* in the function of preterite of the *present perfect* 'he had done and now was,' etc.

The future is usually expressed by the subjunctive (in Old Avestan by the subjunctive of the present or aorist depending on the aspect), but also by the special future stem in *-hiia-l-šii-*.

7.1.1 Present indicative

The present indicative is the tense used to describe events that are currently taking place, whether they always take place or only now. It is also the 'performative' tense (especially in the 1s and 1p), used, notably, in ritual contexts to describe the actions in the process of being performed by the speaker or someone referred to by the speaker.

It is used in main (declarative, interrogative) clauses and subordinate clauses of actions or situations that obtain in general (also eternal truths), but with focus on their current or incipient validity; actions or situations that obtain at the moment of utterance, e.g. during the performance of the ritual; and other actions or situations that are concurrently taking place:

OAv. *kē yā mā [uxšiiē'tī nər'fsa'tī] 0βaṭ*
 'who (is he) through whom the moon [is (now)] first [waxing] then [waning]' (Y. 44.3);

aē'biio yōi "ruuātāiš drūjō ašahiā gaē0ā [vīmər'ncā'tē]
 '... to those who by (their) deals with the Lie [are (now) destroying] the living beings of Order' (Y. 31.1);

ahiā [yāsā] nəmaṅhā . . . vaṅhəuš xratūm manaṅhō
 'in homage to him, I [am (here and now) asking] for the wisdom of good thought' (Y. 28.1);

YAv. *vasō.xša0rō [ahi] haoma . . .*
ər'žuxdəm [pər'sahi] vācim
 '[you are] in command at will, O Haoma . . .
 [you are asking] a word correctly spoken' (Y. 9.25);

apam napās' tā āpō . . . šōi0rō.baxtā [vī.baxša'ti]
 'Apam Napāt [distributes] those waters distributed by settlements' (Yt. 8.34);

[nāismī] daēuuō (for *daēuuī*)
 'I (herewith, by my utterance) [blame/scorn] the old gods' (Y. 12.1);

aiiaoždiia pascaēta bauua'nti
 'after that they become ritually impure' (V. 3.14);

OPers. *sakā tayaiy xaudām tigrām [bara'tiy]*
 'the Scythians who wear a pointed hat' (DB 5.22);

martiya taya patiy martiyam 0ātiy ava mām [naiy varnavataiy]
yātā ubānām hadugām [āxšnauvaiy]
 'I [do not believe] what a man says against (another) man
 until I [hear] (= have heard) the testimony of both (of them)' (DNb 21–24) [cf.
 section 6.4.3.4];

aīta adam yānam [jadiyāmiy] auramazdām
 'this I [am requesting] of Ahuramazdā (as) a request' (DPd 20–21).

7.1.1.1 Av. present indicative with 'before' = past

OAv. *parā* 'before' and YAv. *pa"ruua-* 'former, earlier' sometimes transfer the action of the present indicative into the past:

OAv. *drūjō āiiesē [hōiš.pi0ā] tamuō [parā]*
 'in the *harness of the Lie [you (plur.) *very much fattened] (your) bodies [before = of old]' (Y. 53.6) (*hōiš.pi0ā*: intensive *hai-špī-* < *spī-* 'fatten?');

YAv. *yōi [pa"ruua] mi0rəm [družiṅti]*
 'who [betrayed] Miθra [before]' (Yt. 10.45).

7.1.1.2 Young Avestan, Old Persian present *asti* for imperfect

In both Young Avestan and Old Persian, *asti* is occasionally used instead of the (rare) imperfect:

YAv. *kauuaēm x^varənō . . . yazama^dde . . .*
yaṭ [asti] ahurahe mazdā yaθa dāmaṇ daθaṭ
 ‘we sacrifice to the Fortune of the poets . . .
 which [was] Ahura Mazda’s when he set in place the creations (Yt. 19.9–10);

OPers. *yaθā taya adam xšāyaθiya abavam*
[astiy] a^ttar aitā dahayāva . . . ayauda
 ‘when I became king
 there [were] among these lands (some that) were in turmoil’ (XPh 29–32).

7.1.1.3 Old Persian present for continuing state

In Old Persian, the present is found with adverbs denoting time in the past to express states that still obtain in the present:

hacā paruviyata^h āmātā [a^hmahay]
 ‘from the beginning [we have been] distinguished’ (DB 1.7–8).

7.1.2 Present injunctive and imperfect

7.1.2.1 Present injunctive

In Old Avestan, the present injunctive is used for general (durative or iterative) actions or states typically taking place either in the divine or the human sphere, mostly without specific time reference (‘does’ or ‘would do, used to do’):

mazdā [dadāṭ] ahurō hauruuatō amər³tātas-cā / būrōiš ā . . .
 ‘Ahura Mazda [(always) gives] out of (his) plenty of wholeness and immortality’
 (Y. 31.21);

adā tašā gōuš [pər³saṭ] ašəm
 ‘then the fashioner of the cow [will (usually) ask] (or: asked?) Order’ (Y. 29.2);

tā [dəb³naotā] mašīm hujiiātōiš
 ‘thereby [you (plur.) deceive] mortal man of good living’ (Y. 32.5).

7.1.2.2 Narrative past

The common past narrative tense is the present injunctive (augment-less imperfect) in Young Avestan and the augmented imperfect in Old Persian. It also provides the background description for actions in the present injunctive. The Old Avestan texts contain too few examples of the augmented imperfect to give a precise idea of its use, but it may simply transpose descriptive and repetitive present indicative actions and states into the past. Evidence in Old Avestan for the use of the present injunctive as past narrative tense is uncertain:

OAv. *kadā [ajān] mūθrəm ahiiā madahiiā*
 ‘when [did] the urine of his intoxication [(use to) smash] . . .?’ (Y. 48.10);

0βōi [as] ārma'tiš
 'Humility [was] with you' (Y. 31.9);

YAv. *yō [as] vər³0rajqstəmō*
 'who [was] the most obstruction-smashing' (Y. 9.15);

ā'ī [mraoī] ahurō mazdā
 'then Ahura Mazdā [said]'.

In Old Persian, the (augmented) imperfect is a narrative tense referring to actions and events in the past, both successive actions and events and actions and events anterior to other actions and events in the past:

iyam gaumāta haya maguš [a-durujiya] avaθā [a-0a'ha]
 'this (picture represents) Gaumāta the Magian; he lied (and) said thus' (DBb);

yaθā ka^mbūjiya mudrāyam [a-šiyava] pasāva . . .
 'when Cambyses [had gone] to Egypt, then . . .' (DB 1.33);

auramazdā-maiy upastām [a-bara] yātā ima xšačam [ham-a-dārayaiy]
 'A huramazdā [bore] me aid until [I (had) consolidated] this empire' (DB 1.24–26).

7.1.2.3 Negation and prohibition

The negated present injunctive expresses commands and exhortations to all persons, sometimes coordinated with the imperative:

- 1st person:

OPers. *hacā aniyānā [mā tarsam]*
 '[let me not fear] another!' (DPe 20–21);

- 2nd person:

YAv. *[mā] dim [pər³sō] yim pərəsahi*
 '[don't ask] him whom you are asking!' (H. 2.17);

OPers. *paθim tayām rāstām [mā av^harda]*
 '[do not leave] the straight path!' (DNa 58–60).

- 3rd person:

OAv. *[mā] ciš aī vē drəguuatō maθraqs-cā [gūštā]*
 'but [let no] one among you [keep listening] to the poems of the one possessed by the Lie!' (Y. 31.18);

[mā] əuuīduuā a'pī [d³bāuuaiiaī]
 '[may no] one who does not know [keep deceiving] (us) hereafter!' (Y. 31.17);

YAv. *[mā] ciš [baraī] aēuuō yaī^rristəm*
 '[let no] one [carry] alone what is dead' (V. 3.14).

7.1.3 Future in -Sjia-

The functions of the future are similar to those of the prospective subjunctive, indicating imminent future, intention, or obligation. In Old Avestan, the rare future is found in strophe-initial statements only:

aṭ [frauaxšiiā]
‘and so [I shall proclaim]’ (Y. 45.6);

YAv. *nōiṭ huškō huškāi [sraēšiiete]*
‘dry (matter) [shall] not [be mixed] into dry (matter)’ (V. 8.34);

The future is often contrasted with the past and present:

YAv. *aēšqm uxdanqm . . . yāiš yauua fra-ca vaoce*
fra-ca mruie [fra]-ca [vaxšiiete]
‘of these utterances which (instr. for nom.) **have ever been uttered,**
are being uttered, and [will be uttered]’ (Y. 19.10).

Here, Old Avestan uses the aorist subjunctive (see Y. 29.4 in section 6.3.3).

7.1.4 Aorist

7.1.4.1 Aorist indicative

The augmented aorist denotes the immediate, completed, past (Hoffmann 1967, 153–55). The few Old Avestan forms are used about divine utterances, ritual acts and oral tradition, in statements and in questions:

ahurahiiā zī aṭ vō mazdā
yasnəm-cā vahməm-cā vahištəm [a-mēhmadī]
‘for, thus, [we have thought] (= made up our minds about)
the sacrifice and hymn (as) the best (thing) for you (all)’ (Y. 35.7);

nū zīṭ cašma’nī [vii-ā-dar’səm]
‘for [I] just now [caught sight] of it in (my) eye’ (Y. 45.8);

aṭ tā maniiū pa’ruiiē
yā yēmā x’afnā [a-sruuātəm]
‘thus, those two spirits/inspirations in the beginning,
who [have been heard of] (as) “the twin sleeps” (= sleeping twin foetuses)’ (Y. 30.3).

7.1.4.2 Aorist injunctive

While the temporal reference of the Old Avestan present injunctive is relatively clear, that of the unaugmented aorist injunctive is more difficult to determine. The main problem is whether a given form refers to action already completed at the time of the ‘now’ or whether it is just starting. The aorist injunctive probably does not refer to an action that took place at a specific moment in the past, but, at most, to an action that took place at an unspecified time in the past, e.g. for the first time, ‘aorist of creation’.

It is often accompanied by present indicative or present injunctive or even periphrastic constructions with the present participle, which give the context in which the aorist injunctive took or has taken place.

7.1.4.2a *Aorist injunctive expressing anteriority*

The function of anteriority (to the main verb) is seen clearly in a few instances in subordinating clauses:

aṭ 0βā mējḥī pa^hruuīm mazdā yazūm stōi . . .

vaṅhēuš p^htarēm manajhō

hīiaṭ 0βā [hēm] cašma'nī [°grabəm]

‘thus, I (now have begun to) think-**aor.inj.** of you as being the first, O Mazda, . . . (yet) youthful, father of the good thought – since [I have (just now) grasped] -**aor.inj.** you in (my) eye’ (Y. 31.8);

hīiaṭ mēzdam zaraθuštrō magauuabiiō [cōišt] parā

garō d^hmānē ahurō mazdā jasaṭ pauruiō

‘the fee which Zarathustra [had assigned]-**aor. inj.** before to the participants in the gift exchange, in the House of Song, Ahura Mazda (always) comes-**pres. inj.** (forward for/with it as) the first (in line)’ (Y. 51.15).

7.1.4.2b *Aorist injunctive expressing future*

There is one (apparent) instance of an aorist injunctive being accompanied by a future time reference (future perfect):

yastā daēuuēng aparō . . . [tarē.mqst]

‘(he) who, on account of that, in the future [shall have despised] the old gods’ (Y. 45.11);

7.1.4.2c *Contrasting aspects of present and aorist injunctive*

In Old Avestan, present injunctive and aorist injunctive forms are often contrasted, with clear aspectual difference:

aiiā nōiṭ ərəš [višiiātā] daēuuā-cinā

hīiaṭ iš d^hbaomā . . . upā.jasaṭ

hīiaṭ vər^hnātā acištəm manō

‘especially the old gods [did] not (then) [discriminate]-**aor.inj.** correctly between these two, because deception kept coming over-**pres. inj.** them, so that they would prefer-**pres.inj.** the worst thought’ (Y. 30.6);

aṭ-cā hōi [scaṇtū] . . . yasnqs-cā

. . . dāṅhō ər^hzūš paθō /

yqm daēnaqm ahurō saošiiantō dadāt

‘and so [let them (now) start pursuing]-**aor. imp.** . . . the sacrifices (to him), (Mazdā), . . . (which are) the straight paths of the gift (awaiting) the vision-soul which Ahura (always) establishes-**pres. inj.** (as) that of the revitaliser’ (Y. 53.2).

7.1.4.2d *Mythical events*

The aorist injunctive is used to describe mythical events, as well as human actions in the past. This use of the aorist injunctive is found in cosmological contexts, often to denote the first occurrence of an act, later repeated in the sacrifice:

kas-nā xʷəṅg strēm-cā [dā] aduuānəm
 ‘who, I wonder, (first) [established]-**aor.inj.** the road of the sun and of the stars?’
 (Y. 44.3);

yaθā tū ī ahura.mazdā [məṅghā]-cā
[vaocas]-cā [dās]-cā [varəš]-cā yā vohū /
aθā tōi dadʷmahī-pres. ind.
 ‘in the same way that you, O Ahura Mazda, (for the first time?) [thought] them
 [spoke], [established], and [produced]-**aor.inj.** (those) which (are) good (thoughts,
 etc.), in that way we are (here and now) establishing-**pres. ind.** them for you’ (Y. 39.4).

7.1.4.3 Negated aorist injunctive to express prohibition

The aorist injunctive is used with the prohibitive negation *mā* to express ‘do not (start)!’, often coordinated with the imperative:

OAv. *huxšaθrā xšəntəm*
[mā] n ē dušʷ.xšaθrā [xšəntā]
 ‘let (now)] those of good command command-**aor. imp.** (us)!
 [let not] those of bad command [(now) command]-**aor. inj.** us!’ (Y. 48.5).

7.1.4.4 Young Avestan and Old Persian aorist

The Young Avestan and Old Persian examples of aorist indicative and aorist injunctive may be imitations of Old Avestan:

YAv. *ā] [maqsta] yimō*
 ‘then Yima [thought]-**aor.inj.**’ (V. 2.31);

OPers. *baga vazarka auramazdā haya imām būmim [a-dā]*
haya avam asmānam [a-dā]
 ‘Ahuramazdā is a great god, (he) who [set in place]-**aor.ind.** this earth,
 who [set in place] yonder heaven’ (DE 1–11);

imā dahayāva tayā adam [a-daršiy]-aor.inj.
 ‘these lands that I [got hold of]’ (DPe 5–10).

7.1.5 Perfect

The function of the perfect in the oldest Indo-European languages was to express the state obtaining from a past action or event. In Avestan, this function is found with verbs of perception and verbs of state, as well as of transitive verbs.

The perfect has a few subjunctive forms, as well as a pluperfect. The optative forms are commonly used in conditions contrary to fact (see sections 8.2.2.6, 8.2.4.2c).

7.1.5.1 Perfects without present

The two perfects *ād-* ‘say’ (OInd. *āh-*) and *vaēd-* have no corresponding presents:

OAv. *a] tōi vīspəṅg aṅgrəṅg ašāunō [ādarʷ]*
 ‘thus, [they say/claim] (that) all the evil ones (are) followers of Order’ (Y. 43.15);

YAv. *naēciš iða zaraθuštra sūš*
yaθā hīm [ādar^o] mašiiāka
 ‘there is no “vitalisation” here, O Zarathustra,
 (such) as humans [say] it (= refer to it)’ (FrD. 3, see Hoffmann 1968);

*[vaēθā]-ca taṭ *cikaēθa-ca*
āi ašāum zaraθuštra mana xraθβā-ca cisti-ca
 ‘I [know] and am aware of that,
 O Orderly Zarathustra, by my wisdom and insight’ (Yt. 1.26).

7.1.5.2 Perfect expressing resulting state

The perfect indicative is used to indicate the result of a preceding action or state and is often used in ‘ever/never’ statements.

Perfect of intransitive and medio-passive verbs:

OAv. *yōi . . . [nōiṭ] frasaiiā vaṅhēuš [cāxnar³ (<kan-)] managhō*
 ‘those who . . . [have never taken pleasure] in the questioning of (their) good
 thought?’ (Y. 44.13);

YAv. **frēna āṅhqm nasunqm*
yā pa’ti āiia z^omā [rīriθar^o (<raiθ-)]
 ‘on account of the great quantity of the corpses
 that [have (ever) died (and now lie dead) all over this earth’ (V. 5.4).

Perfect of transitive verbs:

OAv. *ā mā aēš^omō hazas-cā rəmō [hišāiā (<hai-)]*
 ‘wrath and violence, restraint [have bound] me (and now keep me bound)’ (Y. 29.1);

YAv. *yō nō [daḍa] yō [tataša] yō [tuθruie]*
 ‘(he) who [has made] us, who [has fashioned] us, who [has *compiled] us’ (Y. 1.1);

*yō [nōiṭ] pascaēta [hušx^oafa (for *hu-šūafa < huap)]*
yaṭ maniiū dāmqn daḍiṭəm
 ‘(Sraoša . . .) who [has never slept]
 (ever) since the two spirits would set in place (their) creations’ (Y. 57.17).

7.1.5.4 The past perfect/pluperfect

OAv. *taēcīṭ . . . yōi drəguwatō mazbīš [ci-kōit-ər^oš (<kait-)]*
 ‘those too . . . who [had (always) distinguished] those possessed by the Lie with great
 (brilliant gifts?)’ (Y. 32.11) [see Jasanoff 1997];

YAv. *niš taṭ pa’ti druxš nāšā’te yaḍāṭ aβiciṭ [ja-γm-aṭ (<gam-)]*
 ‘the lie will be dispelled and destroyed there to the very place whence [it had come]’
 (Yt. 19.12).

7.1.6 The verb ‘to be’ and noun clauses

Both the subject and predicate of ‘to be’ are typically in the nominative:

YAv. *mošu taṭ [ās] nōiṭ dar"γəm*
 'quickly that [was], not long' (Yt. 5.65);

Occasionally, adverbs can be used as predicates with 'to be' (see Hoffmann 1952):

OPers. *kāra pārsa . . . [kamnam āha]*
 'the Persian army . . . [was] (just) a few' (men)' (DB 2.18–19).

7.1.6.1 The copula

A 3rd singular copula is often gapped, sometimes also the 3rd plural or infinitive, occasionally other persons when the subject is an expressed pronoun:

OAv. *huuō zī drəguuā*
 'for that one (is) possessed by the Lie' (Y. 46.6);

kaṭ vā xšaθrəm
 'what (is) your command?' (Y. 34.5);

YAv. *kaiia atjḥā ratauuō*
 'which (are) its models?' (Y. 19.18);

OPers. *tayaiy paruvam xšāyadīyā āha*
 'who had been kings before' (DB 1.9–10).

7.1.6.2 The existential verb

This present indicative *asti* usually expresses existence, but is also used in contexts where it could not easily be omitted or is emphatic. In Young Avestan, it is found in the following instances: fronted, in relative-subordinate and interrogative clauses, in some expressions with superlatives, after fronted demonstrative pronoun or adverb (many of these after *zī*), and occasionally elsewhere. The use of the 3p *hənti* is similar.

7.1.6.3 Possession

The existential verb can take a genitive to express possession:

OAv. *ciš ahī [kahiiā] ahī*
 'Who are you? [Whose] are you?' (Y. 43.7);

YAv. *[yejhe] vaēm mahi*
 '[whose] we are' = 'to whom we belong' (Vr. 11.13);

OP. *manā auramazdā [COP.] auramazdāha adam [COP.]*
 'mine (is) Ahuramazdā, Ahuramazdā's (am) I' (DSk);

[dārayavahauš] puçā aniyai-ciy [āha"tā]
 '[Darius's were] other sons, too' = 'Darius had other sons, too' (XPf 28–29).

7.2 Moods

The Old Iranian moods were used much as in other ancient Indo-European languages. The principal innovation was the use of the optative to denote habitual past action in

Young Avestan and Old Persian. In both Young Avestan and Old Persian, ‘irrealis’ is expressed by the perfect optative.

7.2.1 Subjunctive present and aorist

The subjunctive is the tense used to refer to the future, and so denotes intention, prospective action and exhortation. In general, the deliberative subjunctive is used in questions and in a variety of subordinate clauses. The aorist subjunctive has the same functions as the present subjunctive, but with the difference in aspect. The negations are *nōit* and *mā* (negative instructions). On the use of the subjunctive in subordinate clauses, see section 8.2.2.1 (relative clauses) and section 8.2.3 (adverbial clauses).

7.2.1.1 The subjunctive in main clauses

7.2.1.1a 1st person

The 1st person subjunctive usually means ‘I/we want to do, may I/we do, let me/us do’, indicating intention or impending action:

present:

OAv. *yauuōi višpāi fraēštāṅhō [āṅhāmā]*
 ‘for an entire lifespan [let us be] (your) dearest friends!’ (Y. 49.8).

aorist:

taṅ nē nūciṅ [var³šānē]
 ‘that [I am about to produce] for us’ (Y. 51.1);

ašā kaṅ 0βā [dar³sānī]
 ‘[shall I see] you through Order, I wonder?’ (Y. 28.5);

YAv.

present:

azəm tē gaē0ā [frāδaiieni] azəm tē gaē0ā [var³δaiieni]
 ‘I [shall further] your herds, I [shall make] your herds [grow]’ (V. 2.5);

aorist:

aša sraēšta [dar³sāma] 0βā haxma
 ‘[shall we (now get to) see] company with you, O haoma, through most beautiful Order?’ (Y. 60.12).

OPers.

present:

šiyāta [ahaniy] jīva utā maṛta artāvā [ahaniy]
 ‘[let me] be happy (while) alive and blessed (after I am) dead!’ (XPh 47–48).

7.2.1.1b 2nd and 3rd persons

The 2nd and 3rd person subjunctive refer to the future in various modes:

present:

OAv. *hā nō [aṇha]*
 ‘she [shall be = belong] to us’ (Y. 32.2);

kadā yauuā huuō [aṇha]
 ‘when [shall he ever be] (there)?’ (Y. 29.9);

YAv. *auuaθa tē [aṇha] vaṇhō*
 ‘thus for you [shall be] the better (reward)’ (V. 18.7);

aorist:

OAv. *kōi . . . rāmam [dāntē < da’antai]*
kēng ā vaṇhōuš [jima < jam- / gam-] manahō cistiš
 ‘who [will obtain] peace?
 to whom [will come] the insight of good thought?’ (Y. 48.11);

YAv. [*bun (< bu’an)*] *gaēθā amaršāntiš*
 ‘the creatures [shall become] indestructible’ (Yt. 19.12);

mā ciš mē āṇham zaoθranam [fraṇ^vharā] yō nōit . . .
 ‘[let] **not** anyone [consume] these libations of mine who is not . . .’ (Yt. 10.122).

7.2.2 Optative present and aorist

The optative is the mood of the imaginary and thus mainly expresses wishes, prescriptions, imaginary comparisons or examples, and unreal conditions and comparisons. The optative of the aorist is used like that of the present, but with the difference in aspect. For the use of the optative perfect, see 8.2.2.4, 8.2.4.2c.

On the use of the optative in subordinate clauses, see sections 8.2.1 (*that*-clauses), 8.2.2.1 (relative clauses) and 8.2.3 (adverbial clauses).

7.2.2.1 Optative in main clauses

The optative in main clauses expresses wishes and exhortations. These are distinguished when negated: *nōit* negates wishes and *mā* exhortations. In questions, the optative expresses ability, potentiality, and permission:

present:

OAv. *aθā tū nō gaiias-cā astəntās-cā [xiiā]*
 ‘and so [may you be] our life and boniness!’ (Y. 41.3);

nōit d^{ai} bitīm duš.sastiš ahūm [mər^ušiiā]
 ‘[may] the one of bad announcing **not** [destroy] a second time (this) existence!’
 (Y. 45.1).

OAv. *kaθā mazdā rāniō.skər^{ti}tīm gəm [iš^usōit]*
 ‘how [might] he, O Mazdā, [try to obtain] (by prayer?) a joy-bringing cow?’
 (Y. 50.2);

YAv. *vasō.xšaθrō [hiiā] ašauua*
 ‘[may] the Orderly [have] command at will!’ (Y. 8.6);

mā-ca pascaēta mazdaiiasna tqm zqm [kāraiēn] (for *-aiiāēn*) *mā āpō [hər³zaiiēn]*
 ‘afterward the Mazdayasnians [shall **not** plough] that earth **nor** [release] water (upon it)’ (V. 6.2);

OPers. *auramazdām [yadaišā]*
 ‘[you should sacrifice] to Ahuramazdā!’ (XPh 50–51);

daivā mā [yadiyaiša]
 ‘the *daivas* [shall **not** be sacrificed to]!’ (XPh 38–39);

aorist:

OAv. *“ruuāzištō huuō nā . . . [pa⁴tī.jamiā]*
 ‘(you) there, who gladden (us) the most, [may you] *in return [come] to us’
 (Y. 36.2);

kaθā ašāi drujēm [diiqm] zastaiiō
 ‘how [might I deliver] the Lie into the hands of Order?’ (Y. 44.14);

YAv. *vasas-ca tū . . . [xšāēša] hauuanqm dāmanqm*
 ‘and at will [may] you [rule over] your own creations’ (Y. 8.5);

[buiama] tē šōiθrō.pānō
mā [buiama] šōiθrō.‘ricō
 ‘[let us be] for you those who stay in the settlements!
 [let us **not** be] those who leave the settlements!’ (Yt. 10.75);

OPers. *auramazdā-taiy ja⁴tā [biyā⁴]*
utā-taiy taumā mā [biyā⁴]
 ‘[may] Ahuramazdā [be] your striker (= strike you),
 and [may] there **not** [be] family to you (= may you have no family)!’ (DB 4.55–59).

7.2.2.2 Optative of past action in Young Avestan and Old Persian

The optative is used in Young Avestan (with or without augment) and Old Persian to express habitual or repeated action in the past (see Hoffmann 1976: 605–19):

YAv. *zaraθuštrō ahunəm va‘rim [frasrāuuaiiōi⁴] . . .*
āpō va⁴hiš [frāiiazaēta] . . .
daēnaqm māzdaiiasnim [fraor³naēta]
 ‘Zarathustra [would (= used to) recite] the Ahuna Vairiia . . .
 He [would sacrifice] to the good waters . . .
 He [would choose] the vision-soul of those who sacrifice to Ahura Mazda’
 (V. 19.2);

yō bāda . . . [a-uuarōi⁴] vācim (<a-barait)
 ‘who at times [would lift up] his voice’ (Yt. 10.73);

OPers. *yaθā-šām hacā-ma aθahya avaθā [a-kunavaya⁴tā]* (thematic optative)
 ‘as was said to them by me, thus they would do’ (DB 1.17–24);

kāram vasiy [avājaniyā⁴] (<ava-a-)
 ‘he [would kill] the people in large numbers’ (DB 1.50–53).

7.2.3 Imperative

The imperative is used only in positive statements to express a command or an exhortation. The person the command or exhortation is addressed to is frequently in the vocative. In Old Avestan, a distinction is made between the present imperative: ‘keep doing!’ and aorist imperative: ‘stop/start doing!’.

Negative commands or wishes are expressed with the injunctive (see section 7.1.2.1) and the optative (see section 7.2.2.1).

aorist:

OAv. *vohū [ga'dī] manayhā*
[dā'dī] ašā dā (<da'ah) dar'gāiū
 ‘[come (now)] with good thought!
 [give (now)] with Order the gift of a long lifespan!’ (Y. 28.6);

present:

OAv. *ā īt [°uuāēnā] ahurā*
 ‘[keep looking] hither at it, O Ahura!’ (Y. 46.2);

dužuar'snayhō . . . [hēntū] . . . [xraosəntəm upā] . . .
[īratū] īš duuafšō . . . mošucā [astū]
 ‘[let them be] (men) of bad virility! [let them be *howled upon]! . . .
 [let] *torment [huddle] them off, and [let it be] soon!’ (Y. 53.8);

YAv. *[apa] druxš [nase] (<nasiia)*
[apa] druxš [duwara] [apa] druxš [vīnase]
 ‘[disappear], O Lie!
 [run away], O Lie! [lose yourself], O Lie!’ (V. 8.21);

[x'arata] narō aētəm miiazdəm
 ‘[eat], O men, this *myazda*!’ (Y. 8.2);

*x'ar'0anaqm hē [*barəntəm]*
 ‘[let] (some) of the (best) foods [be brought] to him!’ (H. 2.18);

mā hē auui pādəm [auua.hišta]
mā gātūm [nīpa'diiaṅ'ha]
 ‘[do not direct] (your) foot in her direction!
 [do not lie down] on (her) bed!’ (Yt. 17.57);

mā cim gər'zānā [pār-iiāntu] haca ahmāz nmānāz
 ‘[let them not leave] this house complaining about anybody’ (Yt. 13.157);

OPers. *[para-idiy/para-itā] avam kāram [jadiy/jatā]*
 ‘go forth! crush that army!’ (2s/2p imp.) (DB 2.20–21, 3.14–15);

ava-taiy auramazdā^hucāram [kunautuv] . . .
ava-taiy auramazdā [nika'tuv]
 ‘[let] Ahuramazdā [make] that easy for you! . . .
 [let] Ahuramazdā [destroy] that for you!’ (DB 4.76, 79–80).

7.3 Middle, passive, causative

7.3.1 Middle

The middle of transitive verbs most often has passive meaning: YAv. *aza-*, act. ‘leads, takes (away)’, mid.: ‘is led, taken away’; *vaēna-* act. ‘sees’ vs. mid. ‘is seen’, etc.).

More rarely, the middle denotes that the action is being performed in the interest of the subject (e.g. *paca-*, act. ‘cook (for someone else)’, mid.: ‘cook for oneself’; *var^oḍaiia-* act.: ‘increase (trans.), enlarge’, mid.: ‘increase (intrans.), grow’.

Note the intrans. *hišta-*, act. ‘go and stand’, mid. ‘stand, be standing’.

Middle with passive function:

OAv. *aṭ tā maniiū pa^oruiiē*
yā yēmā x^oafnā [asruuātəm]
 ‘thus, those two spirits/inspirations in the beginning,
 which [have been heard of (aor. du. 3p)] (as) “the twin sleeps”’ (Y. 30.3);

YAv. *imqm tanum yā mē [vaēna^ote] huraoda*
 ‘this body, which [is seen] to me (to be) beautiful’ (Y. 11.10);

[vər^oziiatqm]-ca iḍa vohu vāstriia
 ‘and [let] good pastures [be produced]-**pres. imp. 3s** here!’ (Vr. 15.1) [cf. section 5.3.3];

OPers. *ima frašam taya [vainataiy]*
 ‘this perfection that [is seen]’ (DNb 1–2);

middle with reflexive, etc., function:

YAv. *aīβi vastra [yāḡhaiiaḡ^oha] frā zasta [snaiiaḡ^oha]*
 ‘[put on your (sing.)] clothes! [wash your (sing.)] hands!’ (V. 18.19);

zqm caxrəm [kər^onauuāne]
 ‘[I shall make] the earth [my] wheel’ (Yt. 19.43);

ā^oṭ azəm tanūm [aguze]
 ‘then [I hid]-**aor. 1s mid.** [my] body’ (Yt. 17.55);

OPers. *avaθā xšačam [agarbāyatā]*
 ‘thus [he took] the royal command [for himself]’ (DB 1.42–43).

7.3.2 Passive

The passive, including passive middle forms, passives in *-iia-*, and 3s aor. in *-i*, is almost exclusively used when the sentence has no agent or the agent of the sentence is not expressed:

YAv. *yaθa anie yazatāḡhō [yaziḡte]*
 ‘the way (the) other deities [are sacrificed to]’ (Yt. 8.11);

kahe nō iḍa nqma [āy^orriiāṭ]
 ‘whose name among us [will be welcomed] here in song?’ (Yt. 13.50);

OPers. *vayam haxāmanišiyā [Oahayāmahay]*
 ‘we [are called] Achaemenids’ (DB 1.7);

yakā hacā g'dārā [ābariya] (or: *a-bariya*)
 ‘the sisso wood [was brought (or: carried)] from Gandhara and Carmania’
 (DSf 34–35).

3s passive in *-i*:

OAv. *vahištā īštiš [srāuuī] zaraθuštrahē / spitāmahiī*
 ‘the best *ritual [has been heard of] (= become renowned) (as that) of Zarathustra /
 Spitama’ (Y. 53.1);

YAv. *yaṭ turō jāni fraṅrase*
 ‘when the Turian Frangrasiian was struck (down)’ (Yt. 19.93).

7.3.2.1 Passive with agent

Very occasionally, passives are construed with agents (see also section 7.4.2.2 on the past participle).

OAv. passive with agent in the instrumental or with preposition *hacā*:

yā zī [vāuuər³zōi . . . daēuuāiš-cā mašiiāiš-cā]
 ‘yes, those that [have been produced . . . by old gods and men]’ (Y. 29.4);

yā išudō [dadəntē] dāθranqm [hacā ašāunō] /
 ‘the *dues in gifts that [shall be given] [on the part of (= by?) the sustainer of Order]’
 (Y. 31.14).

In Old Persian, the agent is expressed by (one repeated example of each) the genitive-dative, by the preposition *hacā*, or (possibly) by the postposition *rādiy* (see Skjærvø 1985: 215):

utā-[šām] auramazdā naiy [ayadiya]
 ‘and Ahuramazdā [was not sacrificed to by them]’ (DB 5.15–16);

taya-šām [hacā-ma aθahiya] ava akunava
 ‘what [was announced] to them [from (= by) me]; that they did’ (XPh 17–18), cf.
 **taya-šām adam aθaham ava akunava*
 ‘what I said to them, that they did’ (DNa 36–37);

nai-mā kāma taya skauθiš [tunuva³tahyā rādiy miθa kariyaiš]
 ‘it is not my desire that a weak (man) [should be done wrong on account of/by a
 mighty (one)]’ (DNb 8–9).

7.3.2.2 Old Persian alternatives to passive with agent

In Old Persian, passive with agent is avoided by using impersonal 3rd person plural constructions (cf. Middle Persian, Ch. 4, section 4.3.5.3):

OPers. *utā ciça³taxmam [ag³r:bāya³ ānaya³] abiy mām*
 ‘and [they seized] Ciçantaxma (and) [brought] him to me’ (DB 2.87–88), cf.
hauv āçina basta [ānayatā] abiy mām
 ‘that Āçina [was brought] bound to me’ (DB 1.82–83).

Alternatively, the agent can be circumscribed:

taya ištiš [ajaniya] kāra haya bābiruviya hauv akunauš
 ‘that the brick [was pounded] (into shape): the Babylonian contingent, it did (it)’
 (DSf 29–30).

7.4 Participles

Participles are construed as adjectives, usually attributive or in apposition, sometimes as predicates.

7.4.1 Present, future, aorist, and perfect participles

The present and aorist participles differ only in aspect.

7.4.1.1 Present participles

OAv. *yā vā [xšaiiqs] a-dqs drītā [aiiaṅtəm] . . . vīcīrō [hqs]*
 ‘or (he) who (because) [being in command], (by) having placed (him) there (in his home)
 would keep (someone) [coming] to (him) . . . [being] discriminating . . .’ (Y. 46.5);

fraēšiiāmahī . . .
[sʰrunuuatas]-cā [a-sʰrunuuatas]-cā
[xšaiiaṅtas]-cā [a-xšaiiaṅtas]-cā
 ‘we send (our *presentations) forth,
 both toward [those who listen] and toward [those who do not listen],
 toward [those being in command] and toward [those not being in command]’
 (Y. 35.4);

YAv. *ātārəm; . . . yazamaʹde taxməm [həṅtəm] raθaēštārəm*
 ‘we sacrifice to the fire, [being] a firm charioteer’ (Y. 62.8);

OPers. *X māhayā [jiyamnam] patiy*
 ‘of the month of X at [the waning] (= on the last day)’ (DB 2.62).

The present participles can be used with ‘to be(come)’ (see also section 8.2.4.2e):

OAv. *aṭ . . . ahuuā . . . [frīnəmnā] ahurāi ā*
 ‘thus, **we two were** there [presenting ourselves as guest-friends] to the Ahura’
 (Y. 29.5);

YAv. *nōiṭ dim yauua azəm . . . bitim vācim [paiti.pərʰsəmnō] bauua*
 ‘**I shall** never **be** [asking] him another word’ (V. 18.29);

OPers. *manahā uvaipašiyahayā dəršam [xšayamna] aʹmiy*
 ‘by (my) mind **I am** strongly [controlling] myself’ (DNb 14–15).

In Young Avestan, present participles in apposition with the subject are used with *manīia-*, *sədaīia-*:

nmānəm hō manīiaēta [para.daθō]
 ‘he may think (of himself) [(as) having given] a house’ (V. 18.28);

ā dim vātō [upa.vāuuō] saḍaiie'ti
 'then a wind seems (to be) [blowing toward] him' (H. 2.7).

7.4.1.2 Other participles

Aorist:

OAv. *acištā [dantō]*
 '(even) [giving] what are the worst (things)' (Y. 32.4);

YAv. *fraša haṃ.rāzaiata ātarš . . . u'ti auuaθa [maṃhānō]*
 'he stretched himself forward, the fire, (suddenly) [thinking] as follows' (Yt. 19.47);

future:

YAv. *haomanam-ca [har³šiiamnanam] yōi har³šiiēnte*
 'and of the haoma (plants) [being about to be filtered], which are being filtered'
 (Vr. 12.1);

perfect:

OAv. *[vīduuā] [vīdušē] mraotū*
mā [ə-uuīduuā] a'pī d³bāuuaiiaṭ
 'let him [who knows] say [to the one who knows]!
 may no one [who does not know] keep deceiving (us) hereafter!' (Y. 31.17);

yehiiā "ruuā xraodaitī . . . ašahiā [naṣuuā] (<na-ns-) paθō
 'whose soul will *shudder, [having lost] the paths of Order' (Y. 51.13);

yas-tē . . . zbaiiā . . . [apānō] (<āp-) dar³gō.jiiā'tim
 '(I), who shall be invoking you . . . [having obtained] long life' (Y. 33.5);

YAv. *vispā³biiō haca ar³za³biiō [vauuanuuā] pa'ti jasa'ti*
 'he returns, [having won], from all battles' (Y. 57.11–12);

*aṃhe haxaiiō frāiiēnte . . . naēda.ciṭ *miθō [vaox³āṃhō] (<uak-luac-)*
 'his companions will go forth, [having] never [said] anything wrong' (Yt. 19.95);

šātəm daθā'ti "ruuānəm ašaonō [rīriθānahe] (<raiθ-)
 'he makes happy the soul of the Orderly one [who has died]' (P. 23);

narəm ašauuanəm [daḍrānəm] (<dar-) humatəmca . . .
 'the Orderly man, [who has held up] his well-thought (thought) and . . .' (Vr. 2.5);

perfect participle contrasting with the present and past participles:

OAv. . . . *hu-uar³štanam . . . vər³ziiamnanam- cā vāuuər³zanam-cā*
 'of (deeds) well done, both those being performed and [those having been performed]' (Y. 35.2).

7.4.2 Past participles in -ta-

The past participle in *-ta-* is commonly used in Young Avestan as apposition, in which case it refers to an action that was completed before the action or state indicated by the

main verb of the clause or by present participles. Also when used as an adjective, it sometimes retains this temporal function:

YAv. [*fra-stər³tā*] *pā'ti bar³smən*
 [*uz-dātā*] *pā'ti haomā*
raociṇtā *pā'ti ā0ra*
srāuuaiiamnā *pā'ti ahunā* *va'riiā*
 'beside the barsom [spread out],
 beside the haoma [set up],
 beside the blazing fire,
 beside the Ahuna Vairiia being recited' (Yt. 10.91, Āfr. 4.5);

auua . . . *hanjamanəm* . . . *bar³šnuuō auua* *jhe ašnō* [*γ³matəm*]
 . . . *yazamaide*
 'to yonder gathering [having come (together)] upon the height of yonder sky
 . . . we sacrifice' (G. 2.8);

OPers. *duvarayā-maiy* [*basta*] *adāriya*
 'he was held [bound] at my gate' (DB 2.75, 89–90);

hamiçiyā [*ha³-gmatā*] [*para-itā*] . . .
hamaranam akunava
 'the conspirators, [having come together] and [gone off] . . .
 fought the battle' (DB 2.37–39).

7.4.2.1 Medial-reflexive function of the participle in *-ta-*

The past participle sometimes has medial-reflexive function and can take a direct (inner) object:

YAv. [*ā-stūtas*]-*cā* [*fra-uuar³tas*]-*cā*
āstuiē humatəm manō
 '[having assigned myself by my praise] and [having made my choice]
 (having spoken the *āstuiē* and the *frauuarānē*)
 I assign myself by my praise to well-thought thought' (Y. 12.8) [*ā-stau-* and
fra-uuar-, both middle];

upa. taca . . . *ao0ra* [*paiti.šmuxta*]
 'she came running, [shod] in shoes' (Yt. 5.64).

A past participle is frequently combined with a verb from the same root, see section 10.5 *Figura etymologica*.

7.4.2.2 Past participle with copula

In Young Avestan, the past participles are occasionally used with the copula and sometimes with a personal pronoun (rarely a noun) in the genitive indicating the agent:

YAv. *kaða nō iða ašāum* [*agatō* (<*ā-gam-*)]
a'0iiajaḡhəm ahum ā
 'how [have you come], O Orderly one, to us,
 to (this) existence without danger?' (V. 19.31);

cuuaḡtəm zruuānəm maniauuu stiš [dāta as]
 ‘for how long [had] the existence in the world of thought [been established]?’
 (FrV. 2.19);

mānaiēn ahe yaθa nā satəm-ca . . . paršanaqm [ni-jatəm hiiā]
 ‘just as if a hundred and . . . ears of grain [were to be (lying) smashed down]’
 (Yt. 13.71);

with agent:

YAv. *yaḡ [mē] auuuuuḡ daēuuaiiasnanqm [ni-jatəm]*
yaθa sārəm-a varsanaqm barāmi
 ‘that [by me is smashed down = I have smashed down]
 as many sacrificers to the old gods as I carry hairs on the head’ (Yt. 5.77);

yezica [hē] ania aya šiiaoθna [fra-uuaršta]
pa’tita hē ciθa
 ‘if [for/by him are performed = he has performed] other evil deeds,
 (then) the penalty for it (is) absolved’ (V. 3.21);

agent with attributive past participle:

yōi ka’nina (for ka’ninō) [an-upaēta (<upa-i-) mašiiānaqm]
 ‘the girls [not yet approached by men]’ (Yt. 17.55).

7.4.3 The Old Persian perfect

In Old Persian, the old perfect indicative appears to have been lost and been replaced by constructions of past participle and copula with agents. This new perfect, like the old perfect, expresses the result seen in the present of a past action or event. It is often used to ‘sum up’ past events told in the imperfect (see Lazard 1976: 184–86, Skjærvø 1985).

There are forms of both transitive and intransitive verbs. As the past participle of transitive verbs has a passive meaning (‘done, killed’), the perfect construction in these instances is formally passive.

There is no *active* perfect construction corresponding to the *passive* one, so the construction corresponds to both active and passive constructions in the present and imperfect.

7.4.3.1 Without agent

When the past participle is from an intransitive verb, the perfect corresponds to an active imperfect; when it is from a transitive verb and the agent is not expressed, it corresponds to a passive imperfect:

intransitive verb:

pārsahayā martiyahayā dūrai y arštiš [parā-gmatā]
 ‘the Persian man’s spear has gone far away’ (DNa 43–45);

transitive verb:

ava ahayāyā dipiyā naiy [ni-pištam]
 ‘that [has] not [been written (is not written)] in this inscription’ (DB 4.47);

paruv frašam [fra-mātam] paruv frašam [karta]
 ‘much perfect work [had been ordered], much [has been made]’ (DSf 56–57);

aniyaš-ciy vasiy [astiy kartam]
 ‘much else too [has been done]’ (DB 4. 46–47);

pluperfect:

xšačam taya . . . [parā-bartam āha] ava adam patipadam akunavam
 ‘I redressed the command that [had been taken away]’ (DB 1.61–62).

7.4.3.2 With agent

If an agent (noun or pronoun) is expressed, it is in the genitive-dative, and the perfect corresponds to an active imperfect. The only examples are with *karta* ‘done, made’, and the construction is formally identical with possessive constructions in which *karta* means ‘work, achievement’ (there may have been a difference in word order):

taya[-mai]y karta] utā taya-mai] [piča dārayavahauš XŠhāyā karta]
 ‘what [I have done] and what my [father King Darius has done]’ = ‘what [has been done by me] and what [has been done by my father King Darius]’ (XPc 13–14);

[avaišām] avā [nai]y astiy karta]
ya0ā [manā] . . . hamahayāyā 0arda [karta]
 ‘[they have not done] as much
 as [I have done] in one and the same year’ (‘their achievement is not as much as what I achieved in one and the same year’) (DB 4. 51–52).

7.4.4 The Old Persian potential construction

There are a few verbal constructions in Old Persian involving a past participle + forms of *kar-* ‘do’ and *bau-* ‘become’ that express either completion of a past event or the feasibility of a past event. These constructions are commonly referred to as ‘potential’ constructions (see also section 8.2.2.6 Conditional relative clauses):

active:

yātā [karta] akunavam]
 ‘until [I had finished (doing)]’ (DNa 51, XPf 45–46);

passive:

ya0ā [ka]tam abava]
 ‘when it had finished being dug (DSf 25).

7.5 Infinitives

7.5.1 Infinitive with intentional verbs

Infinitives either depend upon a verb or are used independently to express the purpose of an utterance or an action. They are used with *vas-* ‘wish’, *vīsa-* ‘be ready’, *hācaiiia-* ‘induce’, verbs meaning ‘order’, etc.:

OAv. *huuō nō . . . vaštī . . . car'kər'Orā [srāuuaiieḡhē]*
 'he wishes [to make heard] for us poems of praise' (Y. 29.8);

YAv. *yaθa azəm hācaiene . . .*
zaraθuštrəm [anu-matē] daēnaiiāi
 'that I shall induce . . .
 Zarathustra [to help] the vision-soul [along with his thought]' (Yt. 5.18);

OPers. *utā-diš atāvayam [bartanaiy]*
 'and I was able [to bear] them' (DNb 46–47);

ima stānam hauv niyaštāya [ka'tanaiy] . . .
pasāva adam niyaštāyam imām dipim [ni-paištānaiy]
 'and he gave the order [to dig] this niche.
 then I gave the order [to write] this inscription' (XV 20–25).

7.5.2 *Infinite with verbs of speaking and thinking*

Verbs meaning 'think, speak, see' can take a direct object plus infinitive ('accusative with infinitive') as the transformation of a S + V clause; the infinitive of 'to be' can then also take an accusative predicate (see also Gippert 1985):

OAv. *aṭ θβā mēḡhī . . . yazīm [stōi]*
 'thus, I think of *you* [to be = as being] youthful' (Y. 31.8) [*<* 'I think: you are young'];

YAv. *ašauuanəm tē [āfiē'diāi] mraomi*
 'I tell you a sustainer of Order [(is) to be obtained] (from a sustainer of Order)'
 (Y. 71.13) [*<* 'I tell you: a sustainer of Order is obtained'].

7.5.3 *Independent use*

Infinitives are used independently to express purpose, in which case the subject/direct object of such an infinitive may be in the dative:

OAv. *taṭ mōi [vī-cidiāi] vaocā . . .*
[vīduiē] . . . [mēḡ]-cā [dā'diāi]
 'speak that to me [for (it) to be discriminated],
 [for (me) to know] (it), . . . and [for (it) to be kept in (my) mind]' (Y. 31.5);

YAv. *frā gauue-dat. [vər'ndiāi] . . . ā hīm vaēdaiiamahī*
 'in order for the cow [to be chosen], we make her known to (you)' (Vr. 4.2);

OPers. *āiš hadā kārā . . . hamaranam [cartanaiy]*
 'he came with an army [in order to do] battle' (DB 2.67).

7.6 Verbal adjectives of necessity/gerundives

See Dative of agent (section 6.4.5.5).

8 SYNTAX IV: COMPLEX SENTENCES

8.1 Coordination

The usual method of coordination is by way of conjunctions. The most common exceptions are various kinds of parataxis and repetition.

8.1.1 Parataxis

Both coordination and, in Old Persian, subordination can be expressed by parataxis:

YAv. *vīsa'ti dim fraγrāraiō – nōi't *fraγrāraie'ti*
 'he declares himself ready to wake him (and then, but) does not wake him' (N. 1);

OPers. *avaiy mā dauštā biyā – ^hufraštā-dīš parsā*
 'you should not favour them (and = but) punish them well!' (DB 4.69).

With existential verb:

utā aniyāš-ca āha – duškartam akariya
 'and there was other matter as well (and it = that) had been made badly' (XPh 41–42);

astiy a'tar aitā dahayāva . . . – ayauda
 'there were among these lands (and they = some that) were in turmoil' (XPh 29–32).

In Old Persian and 'post-Old Persian', the verbs 'to send' and 'to command, order to do' can be construed with an unmarked clause:

pasāva adam frāišayam ūvjam – hauv āçina basta ānayatā abiy mām
 'then I sent (an order) to Elam (and) that Āçina was brought bound to me' (DB 1.82–83);

niyaštāyam – hauv araxa utā martiyā . . . uzmayā-patīy akariya'tā
 'I issued (an order) (and) that Araxa and the men . . . were impaled' (DB 3.91–92).

8.1.2 Repetition

In Avestan, nominals, verbs, preverbs, prepositions and negations are often repeated instead of being coordinated by conjunctions (frequent in incantations):

adjectives:

YAv. *asista nū aįha't haca vīsa't gāuš buiīā't*
asistəm ašəm asistəm narš ašaonō aojō
asistō āhūiriš tkaēšō
 'may the cow now be **not cut off** from this town,
 (**not cut off** =) **nor** the Order, **nor** the strength of the Orderly man,
nor the *guidance of Ahura Mazdā!' (Y. 60.3);

verbs:

*nase *da'uii druxš nase daēuuō.ciðre*
nase daēuuō.frakaršte nase daēuuō.fradā'te
apa druxš nase apa druxš duuara

apa druxš vī-nase apāxədre apa-nasīehe
 Get lost, deceiving Lie! Get lost, you demon-spawn!
 Get lost, you demon-dragged-forth! Get lost, you demon-made!
 Disappear, O Lie! Run away, O Lie!
 (Run) away and lose yourself, O Lie!
 You disappear in the northern (direction).

pronouns:

OAv. *kaṭ tōi rāzar*³
kaṭ vašī kaṭ vā stūtō kaṭ vā yasnahiiā
 ‘what (is) a *straight utterance for you?
 what do you want: what of praise or sacrifice?’ (Y. 34.12);

YAv. *imaṭ vō āpō ja’düemi*
imaṭ z’mō imaṭ ruuarā imaṭ aməšā spənta
 ‘this I request from you, O waters,
 and from you, O earths, O plants, and O Life-giving Immortals’ (Y. 65.12);

preverbs:

nī tē zā’re madəm mruiē
nī aməm nī vər’Oraynəm (etc.)
nī taṭ yaθa . . . vasō.xšaθrō fracarāne . . .
nī taṭ yaθa ta’ruuāieni . . .
 ‘I call **down** your intoxication, O tawny one,
 your strength, **and** obstruction-smashing . . .
 (I call it) **down** so that I may go forth commanding at will . . .
 and so that I may overcome . . .’ (Y. 9.17–18);

prepositions:

para auuaṅhe ašnō dāṅhōiṭ
para āpō para z’mō para ruuarāiā (etc.)
 ‘before the establishment of yonder sky,
 of the water, of the earth, **and** of the plant . . .’ (Y. 19.8);

negations:

OAv. *nōiṭ nā manā / nōiṭ sēṅhā nōiṭ xratauuō /*
naēdā var’nā / nōiṭ uxδā naēdā šīiaoθ’nā /
nōiṭ daēnā / nōiṭ ruuqno . . . haciṭē
 ‘neither our thoughts, announcements, guiding thoughts,
 preferences, utterances, actions,
 vision-souls, **nor** breath-souls go together’ (Y. 45.2);

YAv. *mā buiama šōiθrō.ricō . . .*
māda yaṭ nō ugra.bāzāuš niuuānāṭ
 ‘may we **not** be those who leave the settlements . . .
 nor indeed (may it happen) that a strong-armed one should *conquer us’ (Yt. 10.75);

OPers. *naiy āha martiya naiy pārsa naiy māda*
naiy amāxam taumāyā kaš-ci-y

‘there was **no** man, **either** Persian **or** Mede
or indeed any of Our family’ (DB 1.48–49);

abiy imām dahayāum mā ājamiyā
mā hainā mā dušiyāram mā drauga
‘against this land may there **not** come
an enemy army, famine, **or** the lie!’ (DPd 18–20).

8.1.3 Conjunctions

8.1.3.1 Coordination

The coordinating conjunctions are *utā* and enclitic *-cā* ‘and, as well’, which coordinate words (and their epithets), clauses or sentences. Common coordinations include the sequences A B-*ca* and its extension A B . . . C-*ca* and *uta* A *uta* B. Combinations include *uta* . . . -*ca*, *uta* . . . *uta* . . . -*ca*, -*ca* . . . *ca* . . . *uta* ‘and also, as well’. For Old Persian, see Klein 1988.

Connecting sentences:

YAv. *vaṅhuš sraošō . . . hē-ca iða yōiθβā astu*
‘good Sraoša . . . **and** let him have taken up his position here’ (Y. 27.6);

OPers. *manā auramazdā upastām baratuv . . . utā imām dahayāum auramazdā pātuv*
‘let Ahuramazdā bear me aid, **and** let Ahuramazdā protect this land!’ (DPd 13–16).

In Old Avestan, the form A B*cā* # C may be possible at the cesura:

ā mā [aēš^omō hazas-cā # rəmō] °hišāiiā
‘[Wrath and violence, (as well as) restraint] keep me tied up’ (Y. 29.1).

The sequence A-*ca* . . . B-*ca* can take an adjective or a pronoun, or it can itself be the subject or direct object of a verb or be governed by adpositions:

OAv. *aṭcā hōi scaṇtū manarhā uxδāiš šiaoθanāišcā . . .*
[kauuacā vištāspō zaraθuštriš spitāmō f^orašaoštrascā]
‘and so **let them follow** with thought, utterances, and actions . . .
[Kauui Vištāspa, a Zarathustra-follower (and) Spitamid, and Fraša-uštra]’ (Y. 53.2);

YAv. *ima [haomas-ca miiazdqs-ca zaoθrās-ca]*
‘**these** [haomas, myazdas, libations]’ (Y. 4.1);

aṇtar^o [māṅhəm-ca huuar^o-ca]
‘**between** [the moon and the sun]’ (Yt. 6.5);

OP. *aita-maiy aruvastam upariy [manaš-cā ušī-cā]*
‘This is my ability **in** [thought and understanding]’ (DNb 31–32).

Occasionally OPers. *utā* introduces a main clause after a subordinate clause with the function of accommodating an enclitic pronoun:

OPers. *imā uvnarā tayā . . . utā-diš atāvayam bartanaiy*
‘these (are) talents which . . . **(and) I** was able to carry them’ (DNb 45–47).

8.1.3.2 Disjunction

The disjunctive conjunctions are enclitic *vā* ‘or’ and tonic *aṭ vā* ‘or else’, negated YAv. *nauua* ‘or not’, ‘nor’.

The sequence A B-*vā* (A B-*vā* C-*vā*) is most common with clauses, but relatively seldom found with single words. Other sequences include A *vā* B *vā* (C *vā*), X *vā* A B *vā*, A *vā* *aṭ vā* B (see also section 5.3.1 Questions):

OAv. *vand¹ntī vā vānghən vā vaonar² vā*
 ‘(who) are winning, shall (now) win, **or** have (always) won’ (Y. 39.2);

yā aṣaonē vahišṭō # x²aētū vā aṭ vā vər²zēniō /
d¹riiamnā vā
 ‘(he) who is best for the sustainer of Order #
either through (his) *family **or else** (as) a member of the *household
or by (his) *community’ (Y. 33.3);

YAv. *us vā hištō nī vā pa¹ḍiiamnō*
 ‘standing up **or** lying down’ (Yt. 1.17);

OPers. *yadiy imām dipim vaināhəy imai-vā patikarā*
 ‘if you see this inscription **or** these images’ (DB 4.72–73);

negated:

uzgər²ptāt pa¹ti draonāt nauua uzgər²ptāt
 ‘at the *drōn* lifted up **or not** lifted up’ (V. 5.26);

kā hē asti ciθa . . .
nauua hē asti ciθa nauua hē asti āpər²tiš
 ‘what is the penalty (etc.) for it? . . .
 there is **neither** penalty for it, **nor** is there expiation for it’ (V. 3.38–39).

8.1.4 Coordinating particles

These are sentence-initial, -medial and -final.

8.1.4.1 Sentence-introductory *aṭ*, *ā¹t*

These are the two most common particles in Old and Young Avestan and, according to the context, may be rendered as ‘and, then, but’, and similar. They can be followed by a temporal clause introduced by OAv. *hiiaṭ*, *yadā*, YAv. *yaṭ* ‘when’. Enclitic *aṭ* is typical of the *Yasna Haptaṅhāiti* (Y. 35–41). They are also used as adversative particles:

OAv. *sraotā gəuš. āiš . . .*
aṭ tā manūi pa¹ruiiē . . .
aṭ-cā hiiaṭ tā hēm maniiū jasaētəm . . .
 ‘listen with your ears . . .
Then, those two spirits in the beginning . . .
 and **then, when** those two spirits come together’ (Y. 30.2–4);

aṭ-cā ‘as well as’, *aṭ vā* ‘or else’:

fʷδrōi . . . paʰθiiaē-cā vāstriiaēʰbiiō aṭ-cā xʷaētauuē
 ‘to (her) father **and** husband, to the foragers,
as well as the *family’ (Y. 53.4);

yasnəm . . . xšmāuuatō
aṭ vā ašā staomiā vacā
 ‘the sacrifice to one like you (pl.),
or else, by (my) Order, words of praise’ (Y. 33.8);

enclitic *aṭ*:

apō aṭ yazamaʰdē
 ‘**and so** we sacrifice (to) the waters’ (Y. 38.3);

sraēštqm aṭ tōi kəhr pəm kəlrpəm āuuāēdaiiamahī
 ‘**thus**, as your form, the most beautiful of forms, we are making known . . .’ (Y. 36.6).

YAv. *āʰt*:

āʰt yaṭ miθrəm . . . frādādqm . . . āʰt dim dadqm
 ‘**then**, when I brought forth Miθra . . . **then** I established him’ (Yt. 10.1);

adversative *aṭ*, *āʰt*:

In Old Avestan, all examples of *aṭ* introducing a second ‘half-line’ and several introducing ‘verse lines’ (exc. *aṭ vā*, *aṭ-cā*) appear to have adversative function, occasionally also when enclitic. YAv. *āʰt* often has adversative function, notably in conditional clauses introduced by *āʰt yezi* (Vid.), *yezi āʰt* (Nir.):

OAv. *acištō drəguuatqm # aṭ ašāunē vahištəm manō*
 ‘the worst (existence will be that) of those possessed by the Lie #
but for the sustainer of Order (there will be) best thought’ (Y. 30.4);

YAv. *vispe zī anīie maδāṅhō aēšma hacinṭe . . .*
āʰt hō yō haomahe maδō aša hacaʰte
 ‘for all other intoxications are followed by Wrath (i.e. darkness and night),
but the intoxication of the *haoma* is followed by Order (i.e. light and day)’ (Y. 10.8);

mā ciš baraṭ aēuuō yaṭ ʰristəm
āʰt yezi šē barāṭ
 ‘let no one carry alone what is dead;
 but *if* he carries alone something that is dead . . .’ (V. 3.14);

yō gāθā srāuuaiēʰti . . .
yezi hauuaēʰbiia ušibiia aʰβi.sʰrunuuaʰti ratufriš
yezi āʰt nōiṭ hauuaēʰbiia ušibiia . . .
 ‘he who recites the Gāθās . . .
 if he can hear (them) with his own ears, (then) he satisfies the (ritual) models,
 but if he cannot hear (them) with his own ears . . .’ (N. 8).

8.1.4.2 Emphatic particles

These include Av. *ī* (combines with various other particles for emphasis); Av. *-cī*, OPers. *-ciy* ‘even, -self, just, exactly’ (OAv., YAv. *aṭ-cī*, *aṭ-ci*, OPers. *aciy* ‘even then, so’); Av. *-cinā* (< *-cana*) ‘especially, even’; *tū* (used with imperative and optative); and *vōi* ‘*definitely’:

OAv. . . . *yazama'dē varjūš-cā ī varj'hiš-cā ī*
 ‘. . . we are sacrificing to **both** the good male ones **and** the good female ones’
 (Y. 39.3);

YAv. *a0ā a0ā cōī* (< *ca-i*) *zara0uštrō* . . . *viiāmruuīta*
 ‘and in that **very** manner, Zarathustra would renounce . . .’ (Y. 12.6);

vīspəm ā ahmā ī yaδōī (*yaδa-i*) *upa.jasōī* ‘ristqm tanūm
 ‘**just** until **when** he might come upon the dead body’ (V. 6.27);

OAv. *azēm-cī*
 ‘I for my part’ (Y. 29.10);

taṭ nē nū-cī *var'šānē*
 ‘that best (action/command) I am **just** now about to produce for us’ (Y. 51.1);

OPers. *ya0ā paruvam-ci*
 ‘**just** like before’ (DB 1.63);

OAv. *aṭ-cī* *aē'biio vahištā yōi* . . .
 ‘**but also** as the best things to those who . . .’ (Y. 31.1);

YAv. *yaṭ-ci* *tanūm apaiie'ti aṭ-ci* *dim nōi* *rāšaiēte*
 ‘**even when** it reaches the body, **even then** it does not harm him’ (Yt. 10.2);

utā vištāspa utā aršāma ubā ajīvatam
āciy . . . *dārayavaum* . . . *xšāya0iyam akunauš*
 ‘both Vištāspa and Aršāma were (still) alive.
 but, ***even so**, he made Darius king’ (XPf 19–24);

OAv. *aiiā nōi* *ar'š višiiātā daēuuā-cinā*
 ‘**especially** the old gods did not discriminate rightly between these two’ (Y. 30.6);

YAv. *mā aḍa ka0a-cina* *ṭpa'ti.jime*
 ‘may ***I not** come up against any such **at all!**’ (Vr. 22.2);

OAv. *huxša0rastū nē nā vā nārī vā xšaētā*
 ‘may a man or a woman of good command command us!’ (Y. 41.2);

ātarš vōi mazdā ahurahiiā ahī
 ‘you are **definitely** “the fire of Mazdā Ahura”’ (Y. 36.3).

8.1.4.3 *zī* ‘for’

The enclitic particle *zī* is an explanatory particle used in main, subordinate, and parenthetical clauses with the general meaning of ‘for, namely’. Av. *yezī*, which in Young Avestan is in complementary distribution with *yeidi zī*, may be from **yaṭ zī* (see section 8.2.4 Conditional clauses):

OAv. *aṭ tōi aṅhən . . . tōi 𐬀 dātā hamaēstārō*
 ‘and they shall be . . . **for** they (have been) made/established as opponents’ (Y. 48.12);

nū 𐬀𐬀 (< zī īṭ) . . . viiādarəsəm
 ‘**for I just** now caught sight of’ (Y. 45.8).

YAv. *nōiṭ 𐬀 ahmi pa’ti na’ri duua maniiu rēna auuastāṅhaṭ*
 ‘**for** not shall over this man the two spirits take up a firm stance (of competition)’
 (V. 7.51–52).

See also section 5.3.3 Parenthetical clauses.

8.1.4.4 Responsive *bā*, *bōiṭ*

The particle is found once in the *Yasna Haptaṅhāiti*, but more frequently in Young Avestan, where it is often used in answers to questions or as emphasis:

OAv. *huxšaθrō.təmāi bā aṭ xšaθrəm . . . dad’mahi-cā*
 ‘**yes**, for (you) having the greatest command,
 we then establish the command’ (Y. 35.5);

YAv. *ciθra vō buiār’s masānā . . .*
ciθrəm bōiṭ (< bā īṭ) yūžəm-ciṭ x’ar’nō . . .
 ‘may your *greatnesses be brilliant! . . .
yes, and may you, too, give brilliant munificence . . .!’ (Ny.3.11).

8.1.4.5 OAv. *θβaṭ*, YAv. *bāda* ‘*sometimes, *again and again’

OAv. *kē yā mā uxšieitī nərəfsaitī θβaṭ*
 ‘who (is he) through whom the moon is (now) **first** waxing **then** waning (is again and again waxing and waning?)’ (Y. 44.3);

YAv. *həṅti bāda māuuaiiaciṭ caθβārō aršāna*
 ‘even *I* have ***again and again/at times(?)** four males’ (V. 18.31).

8.1.4.6 Interjections

The interjections are *uštā* ‘hail! good luck!’, *nəmō* ‘hail!’, OAv. *auuōi*, YAv. *āuuōiia*, *bāuuōiia*, and *sādrəm* ‘woe!’, which take the dative, and YAv. *āi* ‘O!’, which takes the vocative.

8.2 Subordination

Subordinate conjunctions are usually derived from the relative pronominal stems (OAv. *hiia-*, *ya-*; YAv. *ya-*; OP. *ya-*, *taya*).

8.2.1 Substantival that-clauses

These function typically as the subject or direct object or appositions to the subject or direct object. The mood is according to the meaning:

OAv. *aṭ zī tōi vaxšiiā . . . [hiiaṭ akōiiā drəguuā'tē]*
 'for thus I shall tell you . . . [**that** for the one possessed by the Lie (it will end) in an evil (turn)]-DO' (Y. 51.8);

taṭ aṭ var'madī . . . [hiiaṭ ī mā'nimā'di-cā]
 'thus, we (now) prefer **that**: [**that** we may think them and . . .]-DO' (Y. 35.3);

YAv. *taṭ alumāi jasaṭ āiiaptəm [yaṭ hē puθrō us.zaiiata]*
 'that prize came to him [**that** a son was born to him]-S' (Y. 9.4);

OPers. *kārahayā naiy azdā abava [taya bardiya avajata]*
 'it did not become known to the army [**that** Smerdis had been killed]-S'
 (DB 1.31–32);

ava adam akunavam . . . [yaθā aniya aniyam naiy ja'tiy]
 'that I accomplished: [**that** they no longer kill one another]-DO' (DSe 34–36);

utā [taya būmiš akaniya fravata] . . . kāra haya bābiruviya hauv akunauš
 'and [(the fact) **that** the earth was dug down]-DO, the Babylonian people did (it)'
 (DSf 28–30) [or: *nominativus pendens*];

nai-mā kāma
[taya skauθiš tunuva'tahyā rādiy miθa^h kariyaiš]
 'it is not my desire
 [**that** a weak (man) should be done wrong by a mighty (one)]-S' (DNb 8–9) [*kariyaiš*:
 optative of wish].

8.2.2 Relative clauses

Relative clauses are the most frequent subordinate clauses in Old Avestan and probably in Young Avestan and Old Persian. They are adjectival and are often coordinated with adjectives.

A relative clause can precede, follow, or include its antecedent, which can be overt (explicit) or gapped (omitted) or just implied by the verb. It can be next to the antecedent or separated from it. The antecedent can be a noun or a (personal, demonstrative, interrogative, indefinite) pronoun. The antecedent and the relative pronoun can be in the same case or in different cases, the antecedent usually being in the case demanded by the syntax of the main clause and the relative pronoun in the case demanded by the syntax of the relative clause, but case 'attraction' (assimilation) also occurs.

Like adjectives, relative clauses can be 'substantivised' and take the place of a noun and so by themselves function as subject, direct object, etc. ('he who'-clauses). An antecedent or resumptive 'dummy' pronoun may be present, which, by its case, shows the syntactic function of the relative clause: 'I saw him who killed the dragon', '(the one) who killed the dragon, **him** I saw'. In Old Persian, the 'he who' clauses were standardised and produced a new relative pronoun *haya-*, *taya-* (cf. Adiego Lajara 2000).

As in other early Indo-European languages, the antecedent of the relative may be located inside the relative clause (included) and may be in the same case as the relative pronoun. Such relative clauses may originate from interrogative or indefinite clauses: 'which man killed the dragon?', 'whichever man killed the dragon' > 'which man killed the dragon went away' (see also Kellens and Pirart 1990, Chapter 2 on relative clauses).

See also section 6.1.1 The Young Avestan and Old Persian connecting relative and the relative particle.

8.2.2.1 The use of moods and tenses in relative clauses

All tenses and moods are used in relative clauses, but the functions of the moods can differ slightly from those in main clauses:

subjunctive present:

OAv. [yā vā mazdā ahurā pā'rijasāi] vohū manaijāhā
 '(I) [who **shall circumambulate** you (pl.)], O Mazdā Ahura, with (my) good thought'
 (Y. 28.2);

YAv. aṭ-cā tōi vaēm xiiāmā [yōi īm f'rašēm kər'naon] ahūm
 'thus, also: may *we* be those [who **shall make** it perfect], (this) existence' (Y. 30.9);

subjunctive aorist:

OAv. hātqm huuō aojištō
 [yahmāi zauuēng jīmā *kərəduš ā]
 'he is strongest of (all) who are [for whom **I shall come** (<ā jam-) to (his) calls (even if they are) weak]' (Y. 29.3);

optative present:

OAv. [yā hūm ahmāi vāstrauu' tīm stōi usiāt]
 '(someone) [who **might wish** her and (her) pasture to be for him]' (Y. 50.2);

optative aorist:

taṭ ahiia [yā taṭ upā.jamiāmā] tauua-cā haxəmə
 'that of this (one) [by which **we may come** to that], (namely) your company and . . .'
 (Y. 40.2).

8.2.2.2 Old Avestan relative clauses

The Old Avestan relative clauses are particularly difficult to parse, partly because of the poetic word order and partly because of frequent ellipsis of antecedent, verb, etc.

8.2.2.2a Relative clause precedes the main clause

Antecedent implicit in verb with overt resumptive pronoun:

[yā . . . ašəm nipā'ijhē manas-cā vohū yauuaētātē] | tuuōm . . . frō mā sīšā
 '[(you-sing.) **who** are guarding Order and good thought for the duration of a lifespan,] teach **you** me!' (Y. 28.11);

[yā frauuaxšiiā] yezī tā aōā h'ə'iiā
 '(the words) [**which** I want to pronounce] – for if **they** (are) true in this manner'
 (Y. 44.6);

[yā vā . . . pā'rijasāi . . .] mā'biio dāuuōi ahmuā . . . āiaptā
 'for (you) to give **to me**, [**who** shall circumambulate you], the spoils of both existences' (Y. 28.2).

Antecedent implicit in verb without overt resumptive pronoun:

[yōi mōi ahmāi sərəošəm dān (daʾan-3p) caiias-cā]-S
upā.jimən-3p hauruuātā amərətātā
 ‘[whosoever shall give readiness to listen to this one of mine],
 they shall come to wholeness (and) immortality’ (Y. 45.5);

[yehiā mōi . . . vahištəm yesnē paʾitī]-DO *vāēdā mazdā ahurō*
 ‘(he) [in return for whose sacrifice the best (accrues) to me], Mazdā Ahura knows
 (him)’ (Y. 51.22).

Antecedent included in relative clause with overt resumptive pronoun:

[yē zaotā ašā ərəzuš] huuō . . .
 ‘[the libator who (is) straight by (his) Order], he . . .’ (Y. 33.6);

[yās-cā tōi gʾnā . . . vaʾriiā] tā yazamaʾdē
 ‘[and the women who (are) your well-deserved ones], those we are sacrificing (to)’
 (Y. 38.1).

Antecedent included in relative clause without overt resumptive:

[yē maθrā vācəm . . . barʾitī]-S . . . zaraθuštō
 ‘the poet who is lifting (his) voice], . . . (he is I/he) Zarathustra’ (Y. 50.6);

[hiiaṭ mīždəm zaraθuštō magauuabiiō cōišṭ parā]-INSTR
. . . ahurō mazdā jasaṭ paurauiō
 ‘[the fee which Zarathustra assigned before to the masters of the gift-exchange],
 (with that) Ahura Mazdā comes (forward as) the first (in line)’ (Y. 51.15).

8.2.2.2b Relative clause follows the main clause

Overt antecedent preceding the main clause:

kē [yā mā uxšieitī nərəfsaitī 0βaṭ]
 ‘who (is he) [through whom the moon is (now) first waxing then waning]?’ (Y. 44.3);

aēm mōi idā vistō [yē nē aēuuō sāsnā gūšatā]
 ‘this one (is the one) found by me [who alone has listened to our ordinances]’
 (Y. 29.8);

sraotū mazdā ahurō [yehiā vahmē vohū frašī manaḡhā]
 ‘let Ahura Mazdā listen, [in whose hymn I have (now) asked my good thought]’
 (Y. 45.6).

Antecedent overt and included:

xšaθrəm-cā . . . [yā xšaθrā grēhmō hišʾsaṭ]
 ‘and (their) command . . . [a command by which the *glutton *seeks a seat]’
 (Y. 32.12–13).

Without overt antecedent:

hiiaṭ . . . uštā [yē ašəm dādrē]-DAT
 ‘that . . . wished-for (things for him) [who upholds Order]’ (Y. 51.8);

hiiaṭ-cā gāuš ja'diūi mraoī [yē dūraošəm saocaiiaṭ auuō]-AG

‘and when the cow is mistreated to (the point of) being smashed (by **him**) [**who** “purifies” (by burning) the *haoma*]’ (Y. 32.14).

Antecedent implicit in the verbs:

anāiš vā nōiṭ . . . yānāiš zar^anaēmā-1p . . . [yōi . . . yōiθ^amā-1p]

‘may **we** not anger you with those requests (to you), [**we who** have taken our places]’ (Y. 28.9);

dāidi-2s mōi [yē gəm tašō-2s]

‘**give** me (now), [**you who** fashioned the cow . . .!]’ (Y. 51.7);

var^aštəm-3s [hiiaṭ vasnā frašō.təməm (COP 3s)]

‘let (**that**) be produced [**which** is most perfect in exchange value]!’ (Y. 50.11).

Antecedent included in relative clause:

aṭ yūš daēuuā vspāṅhō akāṭ manāṅhō stā ciθrəm [yas-cā vā maš yazd^atē]

‘but you, O old gods, are all (of you) the *seed (issued) from an evil thought, and (so is) [**the great one who** is sacrificing to you]’ (Y. 32.3);

təm kauuā vīštāspō . . . nqsaṭ . . . [yəm cistīm . . . maṅtā]

‘Kauui Vištāspa *reached **that** (insight), [**the insight which** he thought]’ (Y. 51.16).

Relative clause coordinated with noun or adjective:

OAv. *[vāstriiāṭ vā] d^atē [yē vā nōiṭ aṅhaṭ vāstriiō]*

‘to come [to the forager] or (to him) [**who** shall not be a forager]’ (Y. 31.9);

ahuuā [astuuatas-cā hiiaṭ-cā manāṅhō]

‘of both worlds, both the one with bones and (the one) **which** (is) of thought]’ (Y. 28.2).

8.2.2.2c Case attraction of relative pronoun

There are a few instances in Old Avestan in which the relative pronoun appears to have been assimilated to the case of its antecedent, which may or not be present. We may distinguish between two cases:

- I. The verb comes after the relative clause and may be considered as part of both the main and the relative clause:

vahištəm θβā . . . [yēm ašā vahištā hazaošəm ahurəm] yāsā

‘**you the best-DO** (one), [**whom-DO** of the same pleasure as the best Order, (you) the Ahura-DO] **I ask for** . . .’ (Y. 28.8).

2. A verb may be supplied in the relative clause, e.g. ‘whom (I know as)’, ‘whom (X said to be):

*yaiiā spaniā *u^atī mrauuāṭ [yēm aṅgrəm]*

‘of which two the life-giving one shall tell (**him**) [**whom** (you know/said to be) the Evil one]’ (Y. 45.2),

cf.

[yəng . . . vōistā . . . dāōəng]

‘(those) [whom you know to follow the established rules]’ (Y. 28.10);

təm aduiānəm . . . [yəm mōi mraoš vaɪhəuš manaɪhō]

‘(teach us) that road [which you . . . tell me (is that) of good thought]’ (Y. 34.13).

8.2.2.3 Young Avestan relative clauses

The Young Avestan relative clauses show several of the general features of the Old Avestan ones, although, on the whole, they are less complex. They may precede or follow the main clause, and the antecedent and resumptive may be overt or not.

8.2.2.3a Relative clause precedes main clause

Antecedent included in relative clause:

[yqm ašauua vaɪ him ašaiqm vaēda] tqm druud əuiiduuā

‘[the Orderly fashion which the Orderly one knows to be good], (of) that the Lieful one is ignorant’ (Vr. 22.2).

8.2.2.3b Relative clause follows the main clause

Overt antecedent outside relative clause:

ciɪ auuaɪ vacō ās ahura mazda [yaɪ mē frāuuaoocō]

‘which was yonder word, Ahura Mazda, [which you said forth to me]?’ (Y. 19.1).

Antecedent not overt:

təm aoi ɪbaēšā iša pa’tiiāntu

‘0iiejā iša [yō ī daɪa]

‘0iiejā yaɪtu [yō dī daɪa]

‘by (this) invigorant let the hostilities go back against him [who has established them]

(and), by (this) invigorant the dangers (against him) [who has established them]! let the dangers go (to him) [who has established them]!’ (Y. 65.8).

Antecedent included in relative clause:

rəɪd’ti haomahe madō [yō . . . haoməm vaɪd’te mašiiō]

‘the intoxication of the haoma energises [the man who honors the haoma]’ (Y. 10.8).

8.2.2.3c asti ya- ‘here is/some of which’

Young Avestan relative clauses with *asti* (*asti yō*, etc.) express ‘some . . . others’ (cf. Middle Persian *ast ī*, Ch. 4, sections 5.2.2.8 and 5.3.4):

hazaɪrəm *0anuuaninqm asti yō gauuasnahe snāuuia jūia . . .

hazaɪrəm *išunqm . . . asti yā aɪhaēna spar’ya

‘a thousand *bows, well-made, some of which have strings made of sinews of gazelles . . .

a thousand arrows . . . some of which have stems made of bone’ (Yt. 10.128–29).

8.2.2.4 Old Persian relative clauses

In Old Persian, the relative clause most often follows its antecedent and is often embedded in the main clause:

daraniyam hacā spardā utā hacā bāxtriyā abariya [taya idā akariya]
 ‘the gold [that was wrought here] was carried from Sardis and from Bactria’ (DSf 35–37).

Resumptive not overt:

martiyā [tayai-šaiy fratamā anušiyā āhatā] avāja
 ‘the men [who were his foremost followers] (those) he killed’ (DB 3.74–75).

Resumptive overt:

dārayavauum [haya manā pita] avam xšāyaθiyam akunauš
 ‘Darius, [who (was) my father], him he made king’ (XPf 22–24).

Antecedent *nominativus pendens*:

imā dahyāva [tayā adam aqarbāyam] . . . adam-šām patiyaxšayaiy
 ‘these countries [which I seized . . .] I ruled over them’ (DNa 16–19).

See section 6.1.1 The Young Avestan and Old Persian connecting relative and the relative particle.

8.2.2.5 Indefinite relative clauses

Indefinite relative clauses are introduced by indefinite pronouns or a combination of relative and indefinite pronouns. The relationships between pronouns and antecedents are the same as for relative clauses:

OAv. *maniūm zaraθuštrō vər³ntē . . . [yastē ciš-cā spēništō]*
 ‘Zarathustra chooses the inspiration [whichever is your most life-giving]’ = ‘. . . whichever inspiration . . .’ (Y. 43.16);

[yōi mōi ahmāi s³raošəm dcañ caiias-cā] upā.jimən . . .
 ‘[whosoever shall give readiness to listen to this one of mine], they shall (now) come to . . .’ (Y. 45.5);

aṭ huuō . . . [kas-āṭ ašā huzēntuš]
 ‘thus, he there is . . . [whoever by (his) Order (is) of good lineage]’ (Y. 49.5);

uštā ahmāi [yahmāi uštā kahmāicī . . . mazdā dāiūāṭ ahurō]
 ‘wished-for (things are) in the wish for him, [to whomever Mazdā Ahura shall give (them)]’ (Y. 43.1).

8.2.2.6 Conditional relative clauses

The only example (OPers.) has the perfect optative potential:

OP. *naiy āha martiya . . . [kaš-ci-y*
haya avam gaumātam . . . xšačam dītam caxriyā-3s-perf. opt.]

‘there was **no man at all** . . .

[**who could have taken** the royal command from that Gaumāta]’ (DB 1.48–49, 53).

See also section 8.2.4.2c Conditions contrary to fact.

8.2.3 Adverbial clauses

Adverbial clauses typically function like adverbs, modifying a verb or a clause. Most of them are introduced by adverbial forms of the relative *ya-*: OAv. *hīiaṭ*, YAv. *yaṭ*, OPers. *taya* and *yaθā* ‘as, how’ (most common); local Av. *yaθrā* ‘where’, OPers. *yadā-taya*; temporal OAv. *yadā* ‘when’, YAv. *yaḍa*, and others.

8.2.3.1 Clauses of time

The conjunctions are OAv. *hīiaṭ*, *parā hīiaṭ* ‘before’, YAv. *yaṭ*, *para . . . yaṭ*, *ā . . . yaṭ* ‘till’, *v̄spəm ā aḥmāṭ yaṭ* ‘until’, *yaḍōiṭ*, OPers. *yadiy* and *yaθā* ‘when’; OAv. *yadā* ‘when’, *hakaram* ‘once’, and others.

Tenses and moods are mostly as in main clauses:

8.2.3.1a Old Avestan

aṭ-cā [hīiaṭ tā hōm maniiū jasaētəm-pres. inj.] pa^aruuīm dazdē

‘thus, also: [whenever the two spirits **come together**], (then) one *receives for the first time’ (Y. 30.4);

[hīiaṭ us . . . naptiiaēšū nafšucā . . . °jōn-aor.inj.] . . . (< gan-t < gam-t)

‘[when he **has come up** among the *great-grandchildren and grandchildren] . . .’ (Y. 46.12);

OAv. *[yadā aēšqm kaēnā jama’tī] . . .*

‘when the retribution **has come** for these (sins)’ (Y. 30.8).

Note the verb-less time clauses; whether the gapped verb is the copula cannot be verified:

yā mazdāθā [hīiaṭ-cīṭ vīdušē]

‘(the words) to which attention should be paid, [even **when** (it is) by a knowing (man)]’ (Y. 30.1);

yaθā tā aḥhən [hōḥkər^atā hīiaṭ]

‘how those shall be [when (they are) at the *accounting (< hōḥkər^ati-)]’ (Y. 31.14).

8.2.3.1b Young Avestan

[yaṭ-ciṭ ahi-pres. ind. rašnuuō ašāum upa karšuuar^a yaṭ ar^azahi] . . .

‘[O Orderly Rašnu, just when **you are** over the continent (of) Arəzahi] . . .’ (Yt. 12.9);

*išauuas-ciṭ . . . ašəmnō.vīḍō *bauuainti*

[yaθa graṇtō . . . miθnā’ti miθrō]

‘even the arrows will not find their targets
[when Miθra **dwells** angered]’ (Yt. 10.39);

yauuata gaiia juuūuaa-pres. ind. ‘as long as we two live’ (Yt. 15.40).

[*yaṭ astuuat. ər'tō fraxštāte-pres.subj. haca apaṭ kaṣaoiiaṭ*] . . .
 ‘when *Astuuat. ər'ta* shall stand forth from the Kaṣaoiia Sea . . .’ (Yt. 19.92);

nōiṭ vaṣḥō ahmāṭ šūiaoθnəm vər'zii'e'ti
 [*yaṭ dim puθrahe . . . fr'ā naēzəm nasūm pacāṭ-pres.subj.*]
 ‘he performs a deed no better than
 [when one **cooks** the corpse of one’s son on a spit]’ (V. 16.17);

tqm yazata pāuruuō . . .
 [*yaṭ dim usca uzduuqnaiiaṭ-inj. . . . Oraētaonō*]
 ‘to her Pāuruuō sacrificed
 [when *Θraētaona* **tossed** him up in the air]’ (Yt. 5.61).

Preterital optative:

[*yaṭ tum d'nim auuaēnōiš*] . . . *ā'ṭ tum niš. hiḏōiš*
 ‘[whenever **you would see** . . .] then you would sit down’ (H. 2.13);

. . . [*yauuata xšaiiōiṭ huuqθβō yimō*]
 ‘. . . [for as long as Yima **ruled**]’ (Y. 9.5).

8.2.3.1c Old Persian

[*hakaram-maiy ušīyā gāθavā hištataiy-pres. ind.*] . . .
adakaiy fratara maniyaiy afuvāyā
 [*yadiy vaināmiy-pres.ind. hamičiyam*] *yaθā* [*yadiy naiy vaināmiy-pres.ind.*]
 ‘[once **it stands** (firmly) in place in my mind] . . .
 then I feel more beyond fear
 [when **I see** something conspiratorial than [when **I do not see** (it)]’ (DNb 34–40);

utā-[taiy yāvā taumā ahatiy-pres. subj.]
auramazdā-taiy dauštā biyā
 ‘and, [for as long as **there is** strength to you = you have strength],
 may Ahuramazdā favour you’ (DB 4.71–79);

auramazdā [*yaθā avaina-impf. imām būmim*] . . .
 ‘Ahuramazdā, [when **he saw** this earth] . . .’ (DNa 31–32);

[*yaθā taya adam xšāyaθiya abavam-impf.*] *astiy* . . . [see section 7.1.1.3]
 ‘[when **I became** king], there was . . .’ (XPh 29–32);

viyaxanahayā māhə[yā] XIV raucabiš θakatā āha [*yadiy ulapatatā-impf.*]
 ‘it was on the 14th of the month of Viyaxana [when **he rose up**]’ (DB 1.38);

[*yātā adam bābirauv āham-impf.*] . . .
 ‘[while **I was** in Babylon] . . .’ (DB 2.6);

[*yaθā ka^mbūjiya bardiya avāja-impf.* (< *ava-a-ja^m*)] . . .
 ‘[when Cambyses **had killed** Smerdis] . . .’ (DB 1.31–32).

8.2.3.2 Clauses of place

The usual conjunction is Av. *yaθrā*; YAv. also *yaṭ* . . . *paiti* ‘where’; OPers. *yadā-taya* and *yadāyā* (for **yadāyadā* acc. to Schmitt 1994) ‘where’. Indefinite clauses are introduced by YAv. *yaδōiṭ* ‘wherever’ (often indistinguishable from ‘whenever’):

OAv. *hūiaṭ a'bī.g'mən* [*yaθrā cinuuatō pər'tuš*]
‘when they have come to [where the Ford of the Accountant (is)]’ (Y. 46.11);

YAv. *nasiie'ti . . . ahmaṭ haca nmānāṭ āhitiš . . .*
[*yaθra bāda upastao'ti haomahe . . . baēšazəm*]
‘pollution disappears from this house
[where he ever and again praises . . . the healing of the haoma]’ (Y. 10.7);

[*yaδōiṭ dim bā'diš'təm auua. zanqñ -subj. sūnō*] (*sūnō*-acc. for *spānō*-nom.)
‘[wherever most often dogs will notice him]’ (V. 6.45);

OPers. *utā a'tar aitā dahayāva āha*
[*yadā-taya . . . daivā ayadiya-impf.*] . . .
[*yadāyā paruvam daivā ayadiya-impf.*] *avadā* . . .
‘and among these lands, there was (one)
[where old gods had been sacrificed to] . . .
[wherever the old gods were sacrificed to], there . . .’ (XPh 35–41).

8.2.3.3 Clauses of purpose

The conjunctions are OAv. *hūiaṭ*, YAv. *yaṭ* and *yaθa*, OPers. negated *mā-taya* ‘lest’. The mood is the subjunctive:

OAv. *vaxšiiā vā* . . .
[*hūiaṭ dāθōṅg vīcaiaθā-2p-subj. adāθqs-cā*]
‘I shall tell you,
[so that you may sort out those according to the rules and those not]’
(Y. 46.15);

YAv. [*yaθa azəm ta'ruuaiteni- subj. vīspe yātauuō pa'rikās-ca*]
[*yaṭ maṣm naēciš ta'ruuaitā-subj.*]
‘[so that I shall overcome all sorcerers and witches],
[so that no one shall overcome me]’ (Yt. 1.6);

OPers. *avahaya-rādiy kāram avājaniyā*
[*mā-taya-mām xšnāsātiy-subj.*] (*xšnāsā'tiy?*)
‘therefore he would kill the people/army:
[so that it (they?) will not recognize-subj. me]’ (DB 1.51–53).

8.2.3.4 Clauses of result

The conjunctions are OAv. *hūiaṭ*, YAv. *yaṭ*, OPers. *taya*. The mood is the indicative:

OAv. *hūiaṭ iš d'baomā pər'smanēṅg upā.jasaṭ*
[*hūiaṭ vər'nātā-pres. ind. acištəm manō*]
‘because deception would come over them as they were discussing,
[so that they would prefer the worst thought]’ (Y. 30.6);

YAv. *tā hē ta^hruuaitatəm tbaēšā . . .*
[yaṭ nōiṭ āpō takāiš staiiat-pres. inj.]
 ‘those two overcame his hostilities,
 [so that he did not stem the waters from (their) courses]’ (Yt. 13.78);

OPers. *drauga-diš *hamiçiyā akunauš*
[taya imaiykāram ad^hrujiyaš-impf.]
 ‘the Lie made them conspiratorial
 [so that they lied to the people/army]’ (DB 4.34–35).

8.2.3.5 Clauses of cause

The conjunctions are OAv. *hiiaṭ*, YAv. *yaṭ* and *yaθa*, OPers. *taya* and *yaθā*. The mood is according to sense:

YAv. *[yaṭ asti-pres. ind. baβriš sraēšta]*
 ‘[because the female beaver is the most beautiful there is]’ (Yt. 5.129);

auuaθa saošiiçs [yaθa vīspəm ahūm astuuaitatəm sāuuaiiāṭ-pres.subj.]
 ‘thus (he is called) ‘Revitaliser,’ [because he will revitalise the entire bony existence]’
 (Yt. 13.129);

OPers. *avahaya-rādi-maiy auramazdā upastām abara . . .*
[yaθā naiy a^hrīka āham-impf.]
 ‘for this reason Ahuramazdā . . .
 because I was not on the Evil One’s side’ (DB 4.62–63).

In Old Avestan, a causal clause may be correlated with an instrumental of cause/reason:

OAv. *vaēdā taṭ yā ahmī mazdā anaēšō*
mā kamnafšuuā [hiiaṭ-cā kamnānā ahmī]
 ‘I know that on account of which (= the reason why) I am weak, O Mazdā:
 (it is) because of my lack of cattle [and because I have few men]’ (Y. 46.2).

8.2.3.6 Clauses of manner, degree and comparison

The common conjunction is *yaθā* ‘as, how, than’. Degree is expressed by Av. *yauuaṭ* ‘to the extent that, as long as, until’, *mānaiien ahe yaθa* ‘as if’, etc.; OAv. *yāṭ* and *yā.tē* ‘to the degree that, as far as’. The moods are as in main clauses.

Manner:

OAv. *kā īštiš . . . [yaθā vā hahmī-pres. ind.]*
 ‘what (is your) wish . . . [or (for how I sleep)]?’ (Y. 34.5);

[yaθā tū ī . . . mēnghā-cā-aor.inj.]
 ‘[in the same way that you . . . and have thought them]’ (Y. 39.4);

tā θβā pərəsā . . . [yaθā tā aṅhən-pres. subj.]
 ‘I ask you . . . [how those shall be]’ (Y. 31.14);

YAv. *[yaθa vašī-pres. ind.] . . . ida aṅhō ašauua*
 ‘[as you desire] . . . here you shall be Orderly’ (Y. 71.16);

uzjamiqñ . . . [yaθa hīš fradaθaṭ-pres. inj.]
 ‘may they come up . . . as he brought them forth’ (Y. 55.3);

OPers. [*yaθā-šām hacā-ma aθahaya-impf.*] *avaθā akunavayaʹtā*
 ‘[as was said to them by me], thus they would do (preterital opt.)’ (DB 1.123–24).

Indefinite manner:

YAv. [*yaθa kaθa-ca hē zaošō*]
 ‘[howsoever his pleasurē (may be)]’ (Yt. 4.7).

Comparison:

YAv. *yā yāskərʹstarā ahmāṭ* [*yaθa vaca framrauūāʹre-pres.ind.I*]
 ‘who are more effective at the audition [than they can be said forth with words (to be)]’ (Yt. 13.64).

‘As if’-clauses take the optative:

[*mānaiiēñ ahe yaθa hazarəm narəm ōim narəm ādərʹzaiiōiṭ-pres.opt.*]
 ‘[as if 1000 men were to fetter a single man]’ (Yt. 8.55).

Degree:

auuaṭ [*yāṭ huuarʹ auuācī -aor. pass.*] ‘as (high) [as the sun has been said (to be)]’
 (Y. 36.6).

Indefinite degree clauses take the subjunctive:

OAv. [*yauuaṭ ašā tauuā-subj.-cā isāi-subj.-cā*]
 ‘[to the extent I by the Order (of my ritual) can and am able]’ (Y. 50.11);

[*yauuaṭ cuiuaṭ-ca hē zastaēʹbiia haṅəʹruuāiqñ-subj.*] *aētauuaṭ . . .*
 ‘[however much of it they can grasp by the hands], this much . . .’ (V. 6.29).

8.2.4 Conditional clauses

The conjunctions are OAv. *hiiaṭ*, *yezī*, YAv. *yaṭ*, *yezi* (*yeidi* before *zī*), OPers. *yadiy*. Real conditions are expressed by various combinations of indicative, subjunctive, optative and imperative. The order of protasis (if-clause) and apodosis (then-clause) is irrelevant. See also Panaino 1989–90.

8.2.4.1 Conditional clauses in Old Avestan

The Old Avestan *yezī* clauses seem to be ‘truth-tests’: ‘if something is/was/will be true, then . . .’. The apodosis either contains a statement about the current state of the world or a request for knowledge, rewards or similar in the imperative:

yā frauuaxšiiā [*yezī tā aθā haʹθiiā (COP)*]
ašəm šiiāoʹnāiš dʹbqzaʹtī-pres. ind. ārmaʹtiš
 ‘(the words) which I want to pronounce, [if they (are) indeed true in this manner], (then, clearly), by (her) actions, Ārmaiti is *thickening Order’ (Y. 44.6);

[*yezi a0ā stā-pres. ind. ha'0īm*] *mazdā* . . .
aṭ taṭ mōi daxštəm dātā-imp.
 '[if you are (all) truly in that way], O Mazdā . . .
 then make that a *(permanent?) mark for me!' (Y. 34.6);

[*yezi adāiš ašā drujəm vānghanā-tī-aor.-subj.*] . . .
aṭ tōi sauuāiš vahməm vaxšaṭ-aor.-subj. ahurā
 '[if he shall indeed overcome the Lie with (his) . . ., by his Order] . . .
 then he will magnify (his) hymn with life-giving strengths for you, O Ahura'
 (Y. 48.1).

8.2.4.2 Conditional clauses in Young Avestan

8.2.4.2a *yezi-clauses: simple conditions*

These take the present indicative or subjunctive and the perfect indicative in 'if ever' conditions:

[*yezi vaši-pres. ind. . . . auuā tbaēšā ta'ruuaiiō*] . . .
a0a imā nāmānīš . . . fra.mrauua-pres. imp. (thematic form)
 '[if you wish to overcome yon hostilities] . . .
 then say forth these names!' (Yt. 1.10–11);

*yā ahurahe *xšaiiantō dašinəm upa yu'diie'nti-pres. ind.*
 [*yezi aēm bauua'ti-pres. ind. ašauua-xšnus*]
 '. . . who fight at the right side of the ruling lord
 [if (= whenever) he satisfies the Orderly ones]' (Yt. 13.63);

hazaṅrəm azəm tē zao0ranəm . . . barāni-pres. subj.
 [*yezi jum (<juuam) frapaieni-pres. subj. aoi zəm*]
 'I shall bring you a thousand libations . . .
 [if I reach alive the earth]' (Yt. 5.63);

[*yezi 0βā diduuaēša-perf. ind.*] . . .
ā tē aḡhe fra-ca stuiie-pres. ind.
 'if I have (ever) antagonised you,
 for this I am assigning myself by my praise (*ā-stuiie*?) to you and utter your praise
 (*fra-stuiie*)' (Y. 1.21).

Alternate conditions are expressed by *yezi . . . nauuāt* 'if . . . or if not':

[*yezi taṭ yaiiata-perf. ind. nauuāt taṭ yaiiata-perf. ind.*]
 '[if he has taken up his position or not taken up his position] . . .' (Āfr. 3.3).

8.2.4.2b *yezi-clauses: hypothetical conditions*

These take the optative in the protasis:

[*yezi-ca hō nā pa'ti.hiñcōiṭ-pres. opt.*] *yō* . . .
ku0a aētaṭ druxš pər'nā'te-pres. subj. (aētaṭ druxš for aētəm drujim)
 '[and if that man were to pour] who is . . .
 how shall he overcome this lie-demon?' (V. 9.47).

8.2.4.2c *Conditions contrary to fact*

Conditions contrary to fact (irrealis) take the perfect or present optative in the protasis and the perfect optative in the apodosis:

[*ye'di zī azəm nōiṭ da'diiqm-perf. opt. . . . aom stārəm*] . . .
hāu pa'rika . . . aṅhuuqm auua.hisiḍiāṭ-perf. opt.
 '[for if I had not set in place yonder star] . . .
 (then) . . . that witch would have cut off the *thread of life' (Yt. 8.52–54);

[*ye'di zī mā mašiiāka . . . yazaiiaṅta-pres. opt.*] . . .
*frā *nuruiiō ašauuaoiio . . . šušuiqm-perf. opt.*
 '[for if men were to sacrifice to me],
 (then) I would have gone forth for the Orderly men' (Yt. 8.11).

See also section 8.2.2.6 Conditional relative clauses.

8.2.4.2d *yaṭ-clauses*

In Young Avestan, conditional *yaṭ*-clauses are sometimes difficult to distinguish from temporal clauses:

[*yaṭ anusō vi'fiē'ti-pres. ind.*] . . . *kā hē asti ciṭa*
 '[if/when someone is bugged unwillingly], what is the penalty for that?' (V. 8.26);

[*yaṭ ahmi nnāne . . . nā'rika daxštauuā'ti aṅhaṭ-pres. subj.*]
 '[if/when in this house a woman becomes menstruant]' (V. 5.59);

abda-ca iḍa . . . sadaiiāṭ-pres. subj.
yaṭ iḍa pasēuš anumaiiehe paḍəm vaēnā'te-pres. subj.
 'and wondrous it will seem
 [if the footprint of a sheep should be seen here]' (V. 2.24).

8.2.4.2e *Periphrasis with present participle*

Presumably because not all verbs had perfect optative forms, a periphrasis was used consisting of a present participle feminine accusative [!] + perfect optative of 'to be':

[*yezi-ca aēte nasāuuō . . . nasuš narəm āstāraiiēntim aṅhāṭ-perf opt.*]
išar^o.štā'tiia mē vīspō aṅhuš astuuā . . . (no verb)
 'and if these corpses . . . (if this) corpse would have rendered a man guilty,
 then, pronto, my entire bony existence (would be . . .)' (V. 5.4).

8.2.4.3 *Conditional clauses in Old Persian*

In Old Persian, only the future condition type is attested:

yadiy imām dipim vaināḥay-pres. subj. imai-vā patikarā]
 '[if you see this inscription or these images], (then) . . .]' (DB 4.72–73).

See also section 8.2.2.6 Conditional relative clauses.

9 LEXIS

9.1 Word formation

Words can be derived in a variety of manners: by prefixes and suffixes and by composition, of which the following is a selection.

9.1.1 Derivation

9.1.1.1 Suffixes

Among the more common, even productive, Old Iranian suffixes are the following:

-a-: Among its many uses, this suffix makes adjectives from nouns (including derived nouns and compounds), often accompanied by full or lengthened grade of the first syllable (*vriiddhi*); when the first syllable has a long vowel, the adjective will be identical with the noun from which it is derived: Av. *maniiuuua-* ‘belonging in the other world’ < *maniiu-*; *haozq0βa-* ‘being of good lineage’ < *huzaqntu-* ‘of good lineage’; *āpa-* ‘waterlogged’ < *āp-* ‘water’; *upa^{ri}-z^{ma}-* ‘who live upon the earth’ < *zam-*; *hazaqrō.zima-* ‘a thousand-year period’ < *ziam-*; – OPers. *^huvāipašiya-* ‘own’ < *^huvaipašiya-* ‘self’; *mārgava-* ‘person from Margiana’ < *margu-* ‘Margiana’; *pārsa* ‘Persian’ < *pārsa* ‘Persia’.

-i-, -ja-: The suffix *-i-* and its thematicised form *-ja-* are used to form adjectives denoting various kinds of appurtenance, notably geographical; they are often accompanied by lengthened grade in the first syllable (*vriiddhi*): YAv. *āhū^{ri}-* ‘pertaining to/ belonging to Ahura (Mazdā); *hāuuani-* ‘(time of day) pertaining to the haoma pressing (*hauuana-*)’; *vār^oOrayni-* ‘victorious’ < *vār^oOrayna-* ‘victory, god of victory’; – OPers. *bāgayādi-* month name < *baga-* ‘god’ + *yāda-*, cf. *yada-* ‘to sacrifice’; – YAv. *ra0βiia-* ‘according to the *ratu*’; *tūr^{ri}ia-* ‘Turian’ < *tura-* ‘id.’; *hupa0man^{ia}-* ‘the fact of having good flights (**pa0man-*)’; – OPers. *māniya-* ‘(servant?) belonging to one’s house (*māna-*)’; *0anuvaniya-* ‘connected with/using a bow (< *0anuarln-*); *a0uriya-* ‘Assyrian’ < *a0urā-* ‘Assyria’, *maciya-* ‘Makranian’ < *maka-* ‘Makran’. – A special use of this suffix is to denote someone in charge of: YAv. *pasuš.hasti-*, sb. in charge of, owner of the sheep-pen (*pasuš.hasta-*). – The suffix *-i-* is also used to make patronymics: Av. *zara0uštri-* ‘son/ follower of Zarathustra (*zara0uštra-*)’; also *-āna-*, *-āni-*: YAv. *jāmāspana-* ‘son of Jāmāspa’; *ā0βiūāni-* ‘son of *Ā0βiūā*’.

-ka-, -aka-: This is the most productive suffix in Iranian, but is still relatively rare in Avestan and Old Persian: YAv. *jaⁿikā-* < *jaⁿi-* ‘woman’; *mašiiāka-* > *mašiiā-* ‘man, mortal’; *pasuka-* ‘sheep and goats’ < *pasu-*; *apərⁿnāiūka-* = *apərⁿnāiū-* ‘minor child’ (< *apərⁿna-āiū-* ‘who has not filled his age’); – OPers. *vazarka-* < *vazarln-* ‘greatness’; *arštika-* (or *ārštika-*) ‘spear-man’ < **čršti-* ‘spear’; *kaṛnuvaka-* ‘workman, artisan’ < *kaṛnu-* ‘do, make’.

-na-, -ana-: These suffixes produce various kinds of nouns and adjectives from verb roots: Av. *yasna-* ‘sacrifice’ < *yaz-*; *x^vafna-* ‘sleep’ < *x^vap-*; *haṅjamana-* ‘assembly, gathering’ < *ham-gam-* ‘come together’; *pacina-* ‘cooked meal’ < *pac-* ‘cook’; – OPers. *stāna-*, *stāna-* ‘place (for . . .)’ < *stā-*; *āvahana-* ‘settlement, town’ < *ā-vah-*; *draujana-* ‘liar, deceiver’ < *drau-*; *hamarana-* ‘battle’ < *ham-ar-* ‘form ordered (battle) lines’.

-uant- (-mant- after u): This is a very common suffix meaning ‘full of, containing’: OAv. *aojōṅghuuant-* ‘strong’ (< *aojah-uant-*); – YAv. *xratu-mant-* ‘wise’.

-tāt-: This is a productive suffix making abstract nouns from adjectives. It is also used to ‘quote’ or refer to words: Av. *ha^uruua-tāt-* ‘wholeness’ < *ha^uruua-*; *upara-tāt-* ‘superiority’; OAv. *kəuuī-tāt-* ‘being a *kauui*, the word *kauui*’ (cf. abstract suffix MPers.

-īh, Ch. 4, section 6.1.2.1); YAv. *yauuaē-tāt-* ‘that which is *yauuaē* “forever”’; *riθiiqs-tāt-* ‘fact of being *riθiiant-* “dying” (mortal)’. In Old Persian, this suffix may have been replaced by *-tā-* (OPers. *arš-tā-* vs. YAv. *arš-tāt-* ‘rectitude’).

9.1.1.2 Prefixation

Some prefixes are also preverbs, others are used with nominal forms only.

9.1.1.2a Preverbs used as prefixes

Preverbs used as prefixes typically have slightly different meanings from those of the preverbs, e.g.:

apa ‘in the back, backward’ (preverb: ‘back, backward’): *apa-kauua-* ‘with hump in the back, humpback’;

fra ‘in front’ (preverb: ‘forward, forth’): *fra-bāzu-* ‘(the length of) the arm held forward’, *fra-xšnu-* ‘with knees sticking out, knobby-kneed’;

vī ‘to the side(s), away’ (preverb: ‘to the sides, far and wide’): *vī-bāzu-* ‘(the length of) the arms held to the sides’, *vī-daēuua-* ‘keeping the demons away, discarding the demons’, *vī-xrūmañt-* ‘(a blow) that causes blood to flow out’.

9.1.1.2b Nominal prefixes

Exclusively nominal prefixes include:

a- (*an-* before vowel) ‘not, non-, lacking’, used to negate nouns and adjectives (including those made from verbs) and to make adjectives meaning ‘lacking sth.’: OAv. *a-drujiiant-* ‘non-deceiving’, YAv. *an-ašauuan-* ‘not Orderly’, *ə-uuišta-*, *ə-uuišti-* ‘unfound; lack of finding’ (*a-* + *vista-lvisti-*), *a-ḡhaošəmma-* ‘not drying out’, *a-karana-* ‘limitless’;

hu- ‘good’ and *duš-*, *duž-* ‘bad, evil’: OAv. *hu-šiti-* ‘good dwelling’, *dušiti-* ‘bad dwelling’ (<*duš-š^o*), *x^v-īti-* ‘good/easy going’ (<*hu-iti-*), *duž-iti-* ‘bad/difficult going’; *hu-mata-* ‘well thought’; OPers. *^huv-asa-* ‘having good horses’ (<*hu-asa-*); *duš-ⁱyāra-* ‘bad season, famine’.

9.1.2 Compounding

Compounds, as well as their individual elements, can be all kinds of words (nouns, adjectives, adverbs, other compounds): YAv. *daḡhu-pa’ti-* ‘landlord’, *və^r0ra-jan-* ‘obstruction-smasher’; – OPers. *daiva-dāna-* ‘idol-repository’; *tigra-xauda-* ‘pointed-hat = wearing pointed hat’, *hamarana-kara-* ‘battle-doer = battle-fighter’.

Possessive adjectival compounds (*bahuvrīhi*) are especially common: *bar^osmō.zasta-* ‘barsom-hand = with barsom in the hand’, *hu-xša0ra-* ‘having good command’.

If one member of the compound is a word that usually contains two parts, only one part can normally be used in the compound: *ahura-δāta-* ‘Ahura-placed = set in place by Ahura Mazdā’, *mazda-iiasna-* ‘Mazdā-sacrifice = someone who sacrifices to Ahura Mazdā; compounds consisting of more than two members are probably not of an inherited type: YAv. [*draoγō.vāxš*].*draojišta-* ‘who lies [with lying words] the most’.

9.1.2.1 Form of first and second members of compounds

The final vowel of the first member usually becomes *ō*, whether it as an *a-*, *ā-*, or *an-*stem or an invariable in *-a*: *daēuuō.dāta-* ‘demon-established’ (<*daēuua-*), *zruuō.dāta-* ‘time-

established (< *zruuan-*); *haptō.karšuuā'rī-* '(the earth) having seven continents (< *hapta* 'seven').

Some adjectives use an *i*-stem as the first member of a compound, e.g. *xšuiβi.išu-* '(shooting) vibrating/fast arrows' (< *xšuiβra-* 'vibrating'), *bər³zi.gāōra-* 'with loud singing' (< *bər³zant-* 'high, loud').

Nouns as second members of compounds are sometimes in the zero grade, e.g. *frādaḡ.fšu-* 'cattle-furthering' (< *pasu-* 'sheep and goats'), *ər³duua-fšnī-* 'with high bosom' (< *fštāna-* 'breat'), *spitāma-* 'with fattened (increased) strength' < **spīta-Hma-* < *ama-*).

Sometimes, the first member of a compound is in a case form:

- nom.: YAv. *afš.tacin-* '(mountains) where water flows' (< *āfš* 'water' nom. sing.), *kər³fš.x³ar-* 'body-eating, carrion (bird)' (< *kṛp-š³uar-* < *h³ar-?*);
- acc.: OAv. *ahūm.biš-* 'world-healer', *vər³ōrēm.jan-* 'obstruction-smasher'; YAv. *ašəm.stut-* 'praising Order';
- gen.: *z³mas.ciōra-* 'earth-seed, having the seed of the earth';
- dat.: *yauuāē-sū-* 'forever benefiting';
- loc.: YAv. *mā³diūdi.šad-* 'sitting in the middle'.

9.1.2.2 Adverbial *āmreḡita* compounds

The so-called *āmreḡita* compounds are adverbial phrases consisting of repeated words: OAv. *narēm narēm* 'man after man'; YAv. *nmāne nmāne* 'in house after house'.

9.1.2.3 'Dual' or 'open *dvandvas*'

These denote a pair consisting of two different objects, which may be singular, e.g. 'firewood and incense', 'Frašaoštra and Jāmāspa', or (generic) plural, e.g. 'cattle and men'. In these expressions both nouns are in the dual, so that the expression looks like '2 firewoods, 2 incenses', etc.: OAv. *gāuuā azī* 'a bull and a (fertile) cow'; YAv. *āpa uruuā're* 'waters and plants', *saḡhauuāci ar³nauuāci* 'the two (sisters) Saḡhauuāci and Arnauuāci'.

If one of the items consists of two words, one word is usually dropped: *mīōra ahura* 'Miōra and Ahura (Mazdā)', *ahuna *a³rīiamana* 'the (prayers) Ahuna (*Vairīia*) and *Airiīaman* (*Išūō*)'.

10 STYLISTIC FEATURES

Both in the Old Avestan complex metres and in the Young Avestan octosyllabic metre, syntactic units are frequently confined within a metrical unit, but clauses commonly continue through several units.

10.1 Old Avestan poetic word order

The Gathic metre is characterised by strophes divided into 3–5 metrical units ('verse-lines' . . . / . . .) split into two smaller units ('half-lines') by a caesura (. . . # . . .).

The most common departure from 'normal' word order involves the splitting up of small syntactic units, for instance, the separation of determinants from their nouns:

noun + adj.:

[*gāuš*] *vər³zēnē* [*aziiā*]

'[cow] in *household [fertile]' = 'in the *household of the fertile cow' (Y. 34.14);

[ahurō] vaēdā / [mazdā]
 ‘[Ahura] knows [Mazdā]’ = ‘Ahura Mazda knows’ (Y. 31.2);

adj./pron. + noun:

[0βahmī] mazdā [vīci0ōi]
 ‘[your] Mazda [in discrimination]’ = ‘in *your* discrimination, O Mazda’ (Y. 32.8);

noun + gen.:

[ratūm] ahurō vaēdā / mazdā [aiiā qsaiiā]
 ‘[model] Ahura knows Mazda [of these two premiums]’ = ‘Ahura Mazda knows the model of these two premiums’ (Y. 31.2).

Often a small syntactic unit, such as noun + adjective/genitive, verb + direct object, or even preverb + verb, is split between two half-lines or lines, especially in *Gā0ās* 2–3, which have first ‘half-lines’ of only four syllables. The split can take various forms: the terms can remain adjacent to one another, occupy parallel positions, be mirrored, or other; and they can be split between half-lines or verse-lines (enjambement).

Split between half-lines:

adā tašā gōuš [pər³saʔ # ašəm]
 ‘then the fashioner of the cow [asks # Order]’ (Y. 29.2);

hiiaʔ dā0ōng vī # caiia0ā adā0qs-cā
 ‘so that you can [dis # criminate] (between those) who are according to the rules and those who are not’ (Y. 46.15).

Enjambement:

maibiiō dāuuōi [aluuā # . . . / āiiaptā]
 ‘for (you) to give to me [of the two existences # . . . / the spoils]’ (Y. 28.2).

‘Inverse enjambement’ (new clause – or smaller syntactic unit – begins with a single word at the end of a verse-line):

xšiiō / darəgēm āiiū təmaḡhō
 ‘lamentations, / a long lifespan of darkness’ (Y. 31.20);

vahištā / gauuōi vər³ziiātqm
 ‘let the best (things/pastures) be produced for the cow!’ (Y. 48.5).

10.2 Parallelism and chiasmus

Sequences of similar groups of words can be arranged in parallel (A_1, B_1, A_2, B_2) or chiasmically (A_1, B_1, B_2, A_2).

10.2.1 Parallelism

Parallelism occurs within or between metrical units ($A_1 \dots B_1 // A_2 \dots B_2$, etc.) and is found in all three languages:

OAv. [kaṭ] vō {xšaθrəm} [kā] {ištiš} #
 ‘what (is) your command? what (is your) wish?’ (Y. 34.5);

yāsā . . . / [maniīduš] . . . {pa^hruuīm} # [spəntahiī] . . . {vīspəng}
 ‘I ask [of inspiration] . . . {the first} # [life-giving] . . . {all}’ = ‘I ask (you) all for the first (existence) of the life-giving inspiration’ (Y. 28.1);

hiiaṭ [hōm] {vohū} # . . . [fraštā] {manaḡhā}
 ‘when [con-] {with good} # [sulted] {thought}’ ‘when he has consulted with (his) good thought’ (Y. 47.3);

YAv. [xrūmīm] {gāuš yā caḡraḡhāxš} / [varā^hθim paṅtəm] {aza^hte}
 ‘[bloody] {cow the grazing} / [of captivity road] {is driven}’ = ‘the grazing cow is driven along the bloody road (the road which makes her bleed) of captivity’ (Yt. 10.38);

OPers. utā [avam vahayazdātam] {agčrbāya^h} utā [martiyā] . . . {agčrbāya^h}
 ‘and [Vahayazdāta] {they seized} and [the men] {they seized}’ = ‘and they seized that Vahayazdāta, and they seized the men’ (DB 3.47-49).

10.2.2 Chiasmus

In this case, the elements of parallel phrases are arranged in opposite order, ‘crossed’, as it were.

Verb + predicate/direct object, predicate/direct object + verb:

OAv. [ahmī] . . . {anaēšō / . . . kamnānā} [ahmī]
 ‘[I am] {weak . . . / few-menn’ed} [I am]’ = ‘I am weak and have few men’ (Y. 46.2).

Adjective/genitive + noun vs. noun + adjective/genitive:

OAv. kū [spəntā] {ārmā^htiš} / kuθrā {manō} [vahištəm] ‘where (is) [life-giving] {Humility}? / where (is) {thought} [the best]?’ (Y. 51.4).

10.2.3 Combined parallelism and chiasmus

OAv. {amər^htātā ha^hruuātā} # [spəništā <maniū>] . . . /
 {təuišī utaiū^htī} # [<manaḡhā> vohū]
 ‘(give me) {immortality and wholeness} [by (your) most life-giving <inspiration>],
 {*tension and *texture} [by (your?) <thought> the good]’ (Y. 51.7).

10.3 Ellipsis in Old Avestan poetry

The omission of key words, mostly verbs, is one of the most intriguing, frustrating, and tantalising aspects of the Old Avestan poetry. The omitted words can sometimes be supplied from the context itself, but very often they must be supplied from the larger context of the poet-sacrificer’s world, as manifested elsewhere in his poetry. If the omitted word is to be supplied from a still more general context, we cannot do it.

Repeated verb:

aṭ kəhrpəm utaii'itiš dadāṭ

ārma'tiš qnmā

'thus, *texture **gives** (him) form,

Humility (the earth) (gives him) breathing' (Y. 30.7);

yas-tā maṇtā pa'ruiiō . . . /

huuō xraθβā . . . ašəm

'(he) who **thought** those (thoughts) the first . . . /

he, by (his) wisdom, (thought [= produced by his thought]) Order' (Y. 31.7), cf.

yō maṇtā ašəm

'(he) who thought Order' (Y. 31.19).

Repeated noun and verb:

kaṭ tōi ašā zbaiēntē auuaghō /

. . . kaṭ tōi vohū managhā

'What **help** do you have for him **when he invokes** (you) with Order?

What (help) do you have (for him when he invokes you?) with good thought?

(Y. 49.12).

Close context:

nū im vīspā # ciθrē zī mazdāṅhō.dūm /

nōiṭ d^{ai}bitīm # duš.sastiš ahūm mər^{ai}šiiāṭ

'now, all, pay attention to *it* (*sēnghəm* 'the (good) announcement?'), for it (is)

brilliant!

let not a second time the one of bad **announcements** destroy the existence!' (Y. 45.1),

cf.

nī hīm mər^{ai}ždiiāi θbahiiā maθrāiš sēnghahiiā

'in order for it (the Lie) to be wiped out by the (poetic) thoughts of your

announcement' (Y. 44.14).

10.4 Formulaic diction in Old Persian

Old Persian prose is characterised by several of the features described above for Avestan.

Word order variation:

avaθā-šaiy aθaham paraidiy [avam kāram jadiy] {haya manā naiy gaubataiy}

'thus I said to him: "go forth! [crush that army] {which does not call itself mine}!"'

(DB 3.14–15),

vs.

avaθā-šaiy aθaham paraidiy [kāra] {haya hamiçiya manā naiy gaubataiy} [avam

jadiy]

'thus I said to him: 'go forth! {that conspiratorial [army] which does not call itself mine}, [crush it]!" (DB 2.30–31).

Same formula, different syntax:

ima taya adam akunavam [vašnā auramazdāha] {hamahayāyā θarda}

pasāva yaθā xšāyaθiya abavam

'this (is) what I did [by the greatness of Ahuramazdā] {in one and the same year} after I became king' (DB 4.3–5),

vs.

ima taya adam akunavam [vašnā auramazdāha] {hamahayāyā 0arda} akunavam
 ‘this which I did – [by the greatness of Ahuramazdā] – I did (it) {in one and the same year}’ (DB 4.40–41).

10.5 Figura etymologica

This is one of the most common stylistic features especially of Young Avestan and Old Persian and is found in a large variety of forms.

Noun + noun:

YAv. *baēuuar³ dōiθranqm vīdōiθre*
 ‘ten thousand **eyes** for the one whose **eyes** see far and wide’ (Yt. 10.82);

cīm haxa haše bara’ti
 ‘what does the **friend** carry to the **friend**?’ (Y. 62.8).

With one noun as first or last member of a compound:

daḡhēuš daḡhu-pa’tiš
 ‘the **landlord** of the **land**’ (Yt. 10.83);

ahmāi darəyqm dar³γō.jīlīm ‘(give) to him **long longevity**’ (Y. 68.11);
šōiθrō.baxtā vī.baxšā’ti
 ‘he **distributes** (the waters) **distributed** by settlements’ (Yt. 8.34).

Noun + adjective:

YAv. *ama ahmi amauuastəmō*
 ‘in **force** I am the most **forceful**’ (Yt. 14.3).

Noun + verb:

YAv. *yānəm vō yāsāmi*
 ‘I am **requesting from** you a **request**’ (Y. 65.11);

fšōhš . . . pa’ti auua.pašāḡ
 ‘one shall ***nail** (him) with ***nails**’ (V. 4.51).

With suppletive forms:

YAv. *cīm aētaiia paīti. vaca paītiāmraoḡ*
 ‘*what did he **answer** by this **answer**?’ (Y. 21.4).

Verb + past participle:

YAv. *haθra. taršta θrāḡhaiiete*
 ‘he **frightens** them (so that they are) **frightened** then and there’ (Yt. 10.101);

yō hišta’tē maniiu.stātō
 ‘(the sky) which **stands stood** in the world of thought’ (Yt. 13.2);

OPers. *haya vināθayatiy anu-dim vinastahayā avaθā parsāmiy*
 ‘he who **does harm**, according to **the harm done**, I punish him’ (DNb 25–26).

Note especially expressions for ‘doing/treating well/badly’, etc.:

YAv. *ya0a hu-bər'tō harāte*

‘when he is **carried well-carried** = well treated’ (Yt. 10.112);

OPers. *avam^hu-bartam abaram . . .*

avam^hu-frastam aparsam

‘him I **carried well-carried** (= treated well) . . .

him I **asked well-asked** (= punished well)’ (DBI.20–22);

taya duš-kartam akariya

‘that which was **done badly-done** = had been done badly’ (XPh 42–43).

11 SAMPLE TEXTS

11.1 Old Avestan

Yasna 28 is the first poem of the first *Gā0ā* (the *Ahunauuđtī Gā0ā*) and follows the *Ahuna vairiia*, which is the introductory strophe to the entire collection. The metre of the *Ahunauuđtī Gā0ā* is relatively free compared to those of the other *Gā0ās*: strophes contain three ‘verse-lines’ divided into two ‘half-lines’, the first of which, with few exceptions, has seven syllables, while the second varies, mostly between 7 to 9 syllables, but some more or less.

Y. 28.6

<i>voḥū ga'dī mananḥā dā'dī aṣā dā (da 0) dar'gāiū</i>	7 + 9
<i>ər'šuuāiš tū uxđāiš mazdā zara0uštrāi aojō ḡhuuat raf'nō</i>	7 + 9
<i>ahma'biīā-cā ahurā yā d^mbišuuatō duuaēšā ta^mruuāiāma</i>	7 + 10

Come with (your)/on account of (my) good thought! Give through (your)/on account of (my) Order the gift which bestows a long life span!

By (your) capacious utterances you, O Mazdā, (gave) support with strength to Zarathustra.

(So) give to us, too, O Ahura, (support) by which we shall overcome the hostilities of the one hostile (to us = the lord of darkness)!

Y. 28.6a

voḥū < *uahu-ū* (< *uahu-H*) ‘good’: *u*-stem adj., instr. sing. neut.

ga'dī < *ga-di* (< *gṃ-di*) ‘come!’: root aor. imper. 2s act.; pres. stem. *jasā-* < *jaća-* for *gaća-* < *gṃ-ća-* < *gam-ljam-* ‘to go, come’.

mananḥā < *man-ah-ā* ‘thought’: *ah*-stem neut., instr. sing.; < *man-* ‘to think’, cf. *manju-* (Y. 28.11c).

dā'dī < *dā-di* ‘give!’: root aor. imper. 2s act.; pres. stem. *dadā-* < *dā-* (< *daH-*) ‘to give’.

aṣā < *art-ā* ‘Order’: *a*-stem neut. instr. sing.; < *r-* ‘to fit (together)’: the Av. form *aṣā-* appears to be from < *árta-*, rather than *rtá-* (= OInd.); the meaning in the Avesta is ‘cosmic/ritual order’, never ‘truth’.

dā < *da 0* < *dā-ah* (< *daHah*) ‘gift’: *ah*-stem neut., acc. sing.; < *dā-* ‘to give’.

dar'gāiū < *dargāju* (< *dṛHga + Hju*) ‘which bestows a long life span’: *u*-stem adj., acc. neut. sing.; adjectival compound < *darga-* ‘long’ + *ju-* < *u*-stem *āju-* *l̥ju-* *l̥ju-* (*Hāju-* *l̥ju-* *Hju-* ‘time, life span, age’.

Y. 28.6b

ər^ošuuāiš < *ṛšy-āiš*: ‘capacious’: *a*-stem adj., instr. plur. neut.

tū < *tū* ‘you’: pers. pron. 2s, nom. enclitic; tonic *tuuam* (see Y. 28.11b).

uxδāiš < *uxθ-āiš* ‘utterance’: *a*-stem neut., instr. plur.; < *vak-/ vač-* ‘speak’ (cf. Y. 28.11b).

mazdā (< *mazd^haH*): ‘who places (all) in the mind, the all-knowing one’: masc. *ā*-stem (*aH*-stem), voc. sing.; adjectival compound < *mas* (< *mṛs*) ‘mind’ + *dā-* (< *d^haH-*) ‘to place’.

zaraθuštrāi < *ǰaraθuštr-āi* ‘to Zarathustra’: *a*-stem proper name, dat. sing.; adjectival compound < *zarat*-(?) ‘old(?)’ + *uštra-* ‘camel’; cf. *frašaoštra-* (see Y. 28.8b).

aojōṭhhuuaj < *auǰah-ujat* (< *auǰah-ujnt*) ‘strong’: *ujant*-stem adj., acc. sing. neut.; < *auǰah-* ‘(bodily) strength’.

raf^onō < *rafn-ah* ‘support’: *ah*-stem neut., acc. sing.; < *rap-* ‘to support’.

Y. 28.6c

ahmaⁱbiiā^o < *ahma-bja* ‘to us’: pers. pron. 1p dat.-abl.; obl. stem *ah-ma-* < *ṇh-*, cf. *nā* < *nah* ‘us’ (gen.-dat. enclitic).

-ča ‘and, too’: enclitic.

ahur-ā ‘lord’: *a*-stem, voc. sing.

yā < *ǰ-ā*: rel. pron., instr. sing. neut.

d^{ai}bišuuatō < *duǰiš-ujat-ah* (< *-ujnt-ah*) ‘hostile’: *ujant*-stem adj., gen. sing. masc.; < root noun *duǰiš-* ‘hostility’.

duuaēšā < *duǰaiš-āh* ‘hostility’: *ah*-stem neut., acc. plur.; < *duǰiš-* ‘be hostile’; note *figura etymologica*.

ta^oruuāiāma < *taru-aiā-ma* (*trH_ua^o*) ‘we shall overcome’: *aiā*-stem pres. subj. 1p; < *tar-* (< *tarH-*) ‘to pass over, cross’, cf. OInd. *tūrva-* < *trH_ua-*.

Y. 28.8

vahištəm *θβā vahištā* *yām ašā vahištā hazaošəm* 7 + 9

ahurəm yāsā vāunuš narōi f^orašaoštrāi (-*ao-* = -*a-u-*) *maⁱbiiācā* 7 + 9

yaeⁱbiias-cā *ū rāṇhaṇhōi vīspāi yaoē vaṇhēuš manaṇhō* 7 + 9

You the best, (I ask) for the best (things), (the one) whom (I know) has the same taste as best Order,

(you), the Ahura, I ask, (once) having won, for the hero Fraša-uštra and for me, and (for those) to whom you shall give it for the entire life span(?) of (someone of) good thought (or: time span of good thought?).

Y. 28.8a

vahištəm < *u_{ah}-išt-am* < *u_{ah}-u-*: superl. acc. sing. masc.

θβā < *θu-ā*: pers. pron. 2s, acc. enclitic.

vahištā < *u_{ah}+išt-ā* ‘best (things)’: acc. plur. neut.

yām < *ǰ-am*: rel. pron. acc. sing. masc.

vahištā < *u_{ah}+išt-ā*: instr. sing. neut.

hazaošəm < *ha-ǰauš-am* ‘having the same taste (as)’ + instr.: *a*-stem adj., acc. sing. masc.; adjectival compound < *ha-* (< *h_uṛ-*) ‘same’ + *ǰauša-* ‘taste’.

Y. 28.8b

ahurəm < *ahur-am*: acc. sing. masc.
yāsā < *iāc-ā* ‘ask for, request’: *a*-stem pres. ind. 1s; < *iā-sca* < *iā*-.
vāunuš < *uā-un-u-š* ‘winner’: *u*-stem adj., nom. sing. masc.; adjectival derivative of the
perf. stem *uā-un-* < *uan-* ‘to win’.
narōi < *nar-ai* < *nar-* ‘man’: *r*-stem masc., dat. sing.
f^orašaoštrāi < *fraša-uštr-āi*: *a*-stem proper name, dat. sing. masc.; adjectival compound
< *fraša-* ‘fat’ + *uštra-* ‘camel’.
ma^hbiā < *ma-bja*: pers. pron. 1s dat.; cf. OInd. *mahya-m*.

Y. 28.8c

yaē^hbūias^o < *i-ai-bjas^o*: rel. pron., dat.-abl. plur. masc.
ū < *i-t*: pers. pron. 3s acc. neut., referring to ‘best things’ neut. plur.
rāyhaṅhōi < *rāh-a-hai* ‘you shall give’: *s*-aor. subj. 2s mid.; < *rā-* ‘to give’.
vīspāi < *uicū-āi* ‘all’: *a*-stem pronominal adj., dat. sing. neut.
yaoē < *iay-ai* < *āju-* (see Y. 28.6a) ‘time, life span’: *u*-stem neut., dat. sing.
vaṅhāuš < *uāh-au-š* < *uāhu-* (see Y. 28.6a): gen.-abl. sing.
manahō < *man-ah-ah* (see Y. 28.6a): gen.-abl. sing.

The three times repeated *vahišta-* ‘best’ leads up to *vahu- manah-* ‘good thought’ (the divine inspiration, the new sunlit sky, . . .), as the desired result of the ritual; it alliterates with *vaunu-*, implying victory over the forces of darkness. The mention of *fraša-uštra-* alludes to Ahura Mazda’s reward: making the new existence *fraša-* ‘filled with the juices of fertility and life’.

Y. 28.9

<i>anāiš vā nōiṭ ahurā mazdā ašəm-cā yānāiš zar^anaēmā</i>	7 + 10
<i>manas-cā hiiaṭ vahištəm yōi vō yōi^omā das^omē stutqm</i>	7 + 9
<i>yūžəm zəuuištiiāṅhō išo xšaθrəm-cā sauuanḡqm</i>	7 + 9

May we not anger you (pl.), O Ahura Mazda, and Order with those requests (to you),
and (your) best thought, (we) who have taken up (our) positions at the *fulfillment
of (our) obligations* (to you in the form) of praises!
You are the fastest offerings and the command over the life-giving strengths.

Y. 28.9a

anāiš < *an-āiš* < *ana-* ‘those (addressed to you)’: dem. pron., instr. plur. neut.
vā < *uāh*: pers. pron. 2p acc. enclitic, cf. *vō* (Y. 28.9b).
nōiṭ < *na-it* ‘not’: negation.
ašəm < *art-am* (see Y. 28.6a): acc. sing.
yānāiš < *iān-āiš* ‘request’: *a*-stem neut., instr. plur.; < *iā-* (see Y. 28.8b).
zar^anaēmā < *jar-na-ī-ma* (< *jar-na-*) ‘may we (not) anger’: *a*-stem pres. opt. 1p;
thematized from *j^har-nā-lj^har-n-* < *j^har-na-H-lj^har-n-H-* (< *j^harH-*), cf. OInd. *hr-ṇ(ī)-*
mid.

Y. 28.9b

manas^o < *man-ah-* (see Y. 28.6a): acc. sing.
hiiaŋ < *h-i-at*: rel. pron., nom. sing. neut.
vahištəm < *uahišt-am* (see Y. 28.8a): nom. sing. neut.
yōi < *i-ai*: rel. pron. nom. plur. masc.
vā < *uah*: pers. pron. 2p gen.-dat. enclitic, cf. *vā* (Y. 28.9a).
yōiθ³mā < *i-a-iθ-ma* ‘we have taken up positions’: perf. ind. 1p act.; < *i-at-*.
das³mē < *dačmai* < *dač-ma-* ‘at the fulfillment of obligations’(?): *a*-stem masc., loc. sing.; < *dač-* ‘to fulfill one’s obligations (to give gifts for gifts)’, cf. OInd. *dāś-*.
stutqm < *stut-a’am* < *stu-t-* ‘praise’: root noun masc., gen. plur.; < *stau-lstu-* ‘to praise’.

Y. 28.9c

yūžəṃ < *iūž-am*: pers. pron. 2p nom.
zəuuīštūāṅhō < *jaṃ-ištū-āhah* < *jaṃ-ištū-ia* ‘most enduring’: *a*-stem superl., nom. plur. masc.; cf. *zāuuar³* (< *zāu-r*) ‘strength (of endurance)’ (of horses, feet); cf. OInd. *yav-iṣṭh-ya-* ‘youngest’.
išō < *iš-ah* ‘(ritual) offering’: root noun masc. nom. plur.
xšaθrəm < *xšaθr-am* ‘command’: *a*-stem neut., nom. sing.; < *xšā-* (pres. *xšaija-*) ‘to be in command, rule’.
sauuaṅhəm < *cauH-ah-a’am* ‘life-giving strength’: *ah*-stem neut., gen. plur.; ablauting *sau-ls-pā-* < *cauH-lcauH* ‘swell (with life-giving juices)’; cf. *spəṅta* (Y.8.3a below).

Y. 28.11

<i>yā aiš ašəm nipāṅhē manas-cā vohū yauuaētā’tē</i>	7 + 9
<i>tuuṣm mazdā ahurā frō mā sīšā θbahmāŋ vaocaṅhē</i>	7 + 9
<i>manūiṣ hacā θβā ē³ṅhā yāiš ā aṅhuš pa³ruiiō bauuaŋ</i>	7 + 9

(You) who with/by these protect your Order and good thought for eternity, you, O Ahura Mazda, teach me to speak according to your inspiration by your mouth (the words/announcements) by which the first existence will be here (every time)!

Y. 28.11a

yā < *i-ah*: rel. pron., nom. sing. masc.
aiš < *a-* ‘with them/these’: pers./dem. pron., instr. plur. masc.
nipāṅhē < *ni-pā-hai* ‘you guard’: root pres. ind. 2s mid.; < *pā-* ‘to protect’.
vohū < *uah-u*: acc. sing. neut.
yauuaētā’tē < *iauai-tāt-ai* ‘for that which is for a (full) time/life span, forever’: *tāt*-stem fem., dat. sing.; secondary derivative < *iauai*, dat. of *āiu-liau-* ‘time/life span’ (see Y. 28.6a).

Y. 28.11b

tuuṣm < *tu-u-am*: pers. pron. 2s nom., cf. *tū* (see Y. 28.6b).
frō sīšā < *fra-či-(ć)ša* ‘*teach’: reduplicated *a*-stem pres. imper. 2s; < *cah-lčš-* (< *caHs-lčHš-*), cf. OInd. *śās-lśikṣa-*.

mā: pers. pron. 1s acc. enclitic.

0βahmāŋt < *0uā-hm-āt* < *0βa-* ‘your’: poss. pron. 2s abl. sing. masc.

vaocajhē < *uā-uč-a-hai* ‘to speak’: inf. of reduplicated *a*-stem aor. *uā-uč-a-* < *uak/uač-* (cf. Y. 28.6b).

Y. 28.11c

maniiēuš < *manj-ai-š* ‘spirit, inspiration’: *u*-stem masc. abl. sing.; < *man-* ‘to think’, cf. *manah-* (Y. 28.6a); belongs to the other world, ‘that of thought’.

hacā < *haca* ‘from, according to’: adpos. + abl.

0βā < *0u-ā* < *0uā-* ‘your’: poss. pron. 2s, instr. sing. neut.

š^uajhā < *āh-ā* ‘mouth’: root noun neut., instr. sing. (see section 2.3.1.2a).

yāiš < *j-āiš* < *ja-*: rel. pron., instr. plur. masc./neut.; + gapped noun.

ā . . . bauuat < *ā . . . baua-t* < *ā-bau-lbū-* ‘to come about, come into existence(?)’: *a*-stem pres. subj. 3s; < *b^hau-lb^hū-* ‘become’.

aṅhuš < *ahu-š* ‘(new) life, (new) world’: *u*-stem masc., nom. sing.

pa^hruiiō < *parū-ija-h* (< *pr^hHu-ja-*) ‘first, primordial’: *a*-stem adj., nom. sing. masc., cf. *parō* ‘before’ < *pr^hH-ah*, OInd. *puras*.

Y. 53.8

<i>anāiš ā dužuuar^ošnaṅhō</i>	<i>dafšniā hēntū</i>	7 + 5
<i>zaxiā-cā vīspāṅhō</i>	<i>xraosəntqm upā</i>	7 + 5
<i>luxša0rāiš jān^orqm xrūn^orqm-cā</i>	<i>rāmqm-cā āiš dadātū</i>	8 + 7 + 5
<i>īratū īš duafšō huūō</i>	<i>dər^ozā mər^o0iiaoš mazištō</i>	7 + 7 + 5
	<i>mošu-cā astū</i>	

On account of those (actions/words of theirs) let them be there (at the judgement as men) of bad virility! (Let them) be duped

and laughable, all (of them)! Let them be booed!

By those of good command (= good rulers) let them be smashed and bled! And (but) let him give peace with these to the settled towns!

Let that greatest torment drive them off with the chain of death! And let it be soon!

Y. 53.8a

anāiš < *an-āiš* < *ana-*: dem. pron., instr. plur. masc. (see section 6.5.3.3).

ā . . . hēntū < *ā . . . h-antu* < *ā-ah-* ‘to be present’: root present imper. 3p act.; < *ah-* ‘to be’

dužuuar^ošnaṅhō < *duž-uršn-āh-ah* < *o^ouršna-* or *o^ouršnah-* ‘bad men, unmanned’: *a-lah-* stem noun/adj., nom. plur. masc. (-*āhah*).

dafšniā < *dafšn-ij-ā* ‘(easily) deceivable(?)’: *a*-stem adj., nom. plur. masc.; verbal adj. of necessity < **dafš-na-* < *dab-* ‘to deceive’ (cf. *yesniā-*, see section 4.6.2); cf. *diβža-* ‘seek to deceive’ (see section 4.1.1.1 no. 8).

Y. 53.8b

zaxiā < *jāh-ij-ā* ‘laughable’: *a*-stem adj., nom. plur. masc.; verbal adj. of necessity < **jāh-* ‘laugh’ (not in Avestan), cf. OInd. *has-*.

vīspāṅhō < *uicū-āhah* < *uicūa-*: nom. plur. masc.

xraosəntqm upā < *xrausa-ntām upa* < *upa-xrausa-* ‘cry at, boo’: *a*-stem pres. imper. 3p mid.; < **xraut*’-*sca-* < *xraud-* ‘to cry’(?).

Y. 53.8c

huxša0rāiš < *hu-xša0r-āiš* ‘who has good command’: *a*-stem adj., instr. plur. masc.; possessive adjectival compound (*bahuvrīhi*) < *hu-* ‘good’ + *xša0ra-* ‘(royal) command’ (see section 9.1.2).

jōr’raqm < *jan-rām* < *jan-* ‘to smash, smite’: root stem pres. imper. 3p mid.

xrūr’raqm < *xrun-rām* < *xrunā-/xrun-* (< *xrūr-na-H-/xrūr-n-H-*) ‘to bleed’ (trans.) < **xrūr-* (*xruH-*): *nā*-stem pres. imper. 3p mid.; cf. *xrūra-* ‘bloody’.

rāmam < *rām-ām* ‘peace’: *ā*-stem fem., acc. sing.

dadātū < *da-dā-tu* < *dā-* ‘to give’: reduplicated athem. pres. imper. 3s act.

šüētibiio < *šjat-ī-bjah* < *šjant-/šjat-* (< *šjnt-*) < *čšai-/čši-* ‘to dwell, inhabit’: pres. part. fem. dat.-abl. plur. with medio-passive meaning ‘inhabited’.

vīz’biiō < *vij-bjah* < *vič-* ‘house’: root noun fem., dat.-abl. plur., dat. function.

Y. 53.8d

iratū < *ira-tu* < *ir-a-* ‘to set in motion, send’: *a*-stem pres. imper. 3s act.; reduplicated *i-er-lī-r-a* (< *Hi-Har-/Hi-Hr-a-*) < *Hr-* ‘to move’, cf. OInd. *iyar-*.

iš < *i-nš* ‘them’: pers. pron. 3p acc.

duafšō < *duafša-h* ‘*torment’: *a*-stem masc., nom. sing. *huuō* < *hau* ‘that’: dem. pron. nom. sing. masc. (see Section 6.5.3.3).

dər’zā < *d^(l)ij^h-ā* ‘chain’: root noun instr. sing.

mər’0iiaoš < *mṛ0i-au-š* < *mṛ0iu-* < *mṛ-t-iu-* ‘death’ < *mṛ-* ‘die’: *u*-stem masc. gen. sing.

mazištō < *maj-išta-h* < *maj^h-* < *majH-* ‘big, great’: superl. nom. sing. masc.

mošuo < *mašu* ‘soon’ < *mac-š-*: adverb, cf. OInd. *makṣú*, Latin *mox*.

astū < *as-tu* < *ah-*: root present 3s imper. act.

11.2 Young Avestan

Y.8.2

x^oarata narō aētəm miiazdəm
yō dim haṇhāne ašā-ca frōr^oti-ca

Eat, men, this *myazd*,
(he among you) who has gained it for himself by (his) Order and by (his) sending (it) forth(?)!

x^oarata < *huara-ta* < *huar-a-* ‘to eat’: *a*-stem pres. imper. 2p.

narō < *nar-ah* < *nar-* ‘man’: voc. plur.

aētəm < *ait-am* < *aiša-laita-* ‘this’: dem. pron., acc. sing. masc.

miiazdəm < *miiazd-am*, a kind of food offering: *a*-stem masc., acc. sing., cf. OInd. *miyédha-*.

dīm < *di-m* ‘it’: pers. pron. 3rd pers., acc. sing. masc. (see section 3.4.1).

haṇhāne < *ha-hān-ai* < *han-* ‘to gain’: perf. 3s mid.

frōr^oti^o < *fra-rt-ī* ‘sending forth(?)’: *i*-stem fem., instr. sing.; action noun in *-ti-* < *fra-r-* ‘to set in forward motion’; see section 2.3.1.2.

Y.8.3

aməša spənta daēne māzdaiiasne
vaṅhauuas-ca vaij' hīš-ca zaoθrās-ca
yō aēšuuu mazdaiiasnaēšuuu mazdaiiasnō aojanō
ašahe rāθma jīštaiiamnō yāθβa ašahe gaēθā mər'γ'nte
auui tū dim disiiata yā apas-ca "ruuarās-ca zaoθrās-ca

O life-giving immortals! O vision-soul of those who sacrifice to (Ahura) Mazdā!
 O good (gods) and good' (goddesses)! O libations!
 (He) who, among these who sacrifice to (Ahura) Mazdā, (while) representing
 himself as someone who sacrifices to (Ahura) Mazdā,
 (as someone) *seeking victory(?) by a *composition of Order
 (yet) by sorcery destroys the living beings of Order,
 do point him out, O waters, plants, and libations!

Y.8.3a

aməša < *a-mṛt-ā* 'non-dead, immortal': *a*-stem adj., voc. plur. masc.; adjectival compound < *a*-, privative prefix + *mṛta-*, past part. of *mṛ-* 'die'.

spənta < *čūHantā* < *čūHan-ta-* 'life-giving': *a*-stem adj., voc. plur. masc.; < *č(u)uHan* (OAv. *spən-*) '*swelling' + suffix *-ta-* 'being provided with' < *čūaH-* (*spā-*) 'swell (with the juices of life)' (see Y. 28.9c, above). The *aməša spənta* 'life-giving immortals' are six beings brought forth by Ahura Mazdā during his cosmogonic sacrifice. When Ahura Mazdā is counted they become the Seven Life-giving Immortals.

daēne < *dajan-ai* 'vision-soul': *ā*-stem fem., voc. sing.; < *dāy-lā-* (< *d^haiH-l^hiH-*) 'to see', specialised in the meaning of 'seeing' in the world of thought, in the beyond; the vision-soul is also the totality of a person's good or bad thoughts, words, and deeds and is represented as a woman, pretty or ugly, accordingly.

māzdaiiasne < *māzdaiiasn-ai* 'pertaining to a *mazda-iasna-*': *i*-stem adj., voc. sing. fem.; "vriddhi" derivative < *mazda-iasna-* 'someone who sacrifices to (Ahura) Mazdā', adjectival compound < (*ahura-*) *mazdā-* + *iasna-* 'sacrifice', cf. *yaza-* < *iaj'-a-* 'to sacrifice (to), offer up in sacrifice (to)'

Y.8.3b

vaṅhauuas° < *uḥ-ay-as*° < *uḥ-u*: voc. plur. masc.

vaij' hīš < *uḥ-u-ī-š* < *uḥ-u-*: derived fem. *ī*-stem adj., voc. plur.

zaoθrās < *jauθr-ā-s*° 'libation': *ā*-stem fem. pl. of neut. sing. *zaoθra-*, voc. plur.; see section 3.1.1.

Y.8.3c

aēšuuu < *ai-šu* + *ā* < *a-*: dem. pron., loc. plur. masc.

māzdaiiasnaēšuuu < *mazda-iasn-aišu* + *ā* < *mazda-iasna-*: *a*-stem masc., loc. plur.

māzdaiiasnō < *mazda-iasn-ah*: nom. sing.

aojanō < *auj'-āna-h* < *aug-lauj'-* 'to present oneself as, say': pres. part. mid., nom. sing. masc.

Y.8.3d

aṣahe < *art-ahja* < *art-a-*: gen. sing.
rāoma < *rāomā* < *rāo-ma-* ‘*composition’: *a*-stem masc., instr. sing. (uncertain derivation and meaning).
jīštaiamnō < *jīšt-aja-mna-h* < *jīšt-aja-* ‘*seek life/victory(?)’: *aja*-stem pres. part. mid., nom. sing. masc.; < *jī-* ‘live’ or *jī-* ‘win’(?).
yāθβa < *iāθu-ā* < *iāθu-a-* ‘sorcery’: *a*-stem masc./neut.(?), instr. sing.; derivative of *iātu-* ‘sorcerer’.
gaēθā < *gaiθ-āh* ‘living beings’: *ā*-stem fem., acc. plur.; derivative of *gai-l jī-* (*gaiH-l jīH-*) ‘live’.
mər^oγ^oṇ-te < *mṛng-tai* (*mṛnx-tai?*) < *mṛ-n-k-* < *mṛk-* ‘to destroy’: *n*-infixated athem. pres. indic. 3s mid.; the form appears to be metathesised, but the exact phonetic processes involved are unclear; the OAv. form would have been **mər^oṇg^(o)-dē*.

Y.8.3e

auui . . . *disiūata* < *abi* . . . *diēja-ta* < *abi* . . . *diē-ja-* ‘to point at’ < *daiē-l diē-* ‘point’: *ja*-stem pres. imper. act.
tū < *tu*: emphatic enclitic particle; see section 8.1.4.
yā < *iā-h*: rel. pron., nom. plur. fem.
apas^o < *ap-ah* < *āp-* ‘water’: root-noun fem., voc. plur.
“ruuarās^o < *ruuar-āh^o* < *ruuar-ā-* ‘plant’: *ā*-stem fem., voc. plur.

Y.8.5

vasas-ca tū ahura mazda uštā-ca xšaēša hauuanəm dāmanəm
vasō āpō vasō “ruuarā vasō vīspa vohū ašaciθra
xšaiiamnəm ašauuanəm dāiūata axšaiiamnəm druuantəm

May you, O Ahura Mazdā, rule at will and according to wish over your own creations!
(May you rule) at will, O waters, at will, O plants, at will, O all good things whose seed is from Order!
Place the Orderly one in command, the one possessed by the Lie out of command!

Y.8.5a

vasas^o < *uācas^o* < *uāc-ah* ‘wish’ < *uāc-luē-* ‘to wish’: *ah*-stem neut., acc. sing. used as adverb; see section 3.3.
tū < *tū*: pers. pron. 2s voc. enclitic.
ahura: voc. sing.
mazda < *mazdā*: voc. sing.
uštā^o < *ušt-ā* < *ušt-i-* ‘wish’ < *uāc-luē-* ‘to wish’: *i*-stem fem., loc. sing., action noun in *-ti-*.
xšaēša < *xšai-ša* < *xšā-ī-ša*: aor. 2s opt., < *xšaija-* to rule (over: + gen.); < *xšā-*.
hauuanəm < *hay-ānā’am* < *hay-a-* ‘own’: poss. refl. pron., gen. plur. neut.
dāmanəm < *dāman-a’am* < *dām-an-* ‘creation’: *n*-stem neut., gen. plur.; < *d^hā-* ‘to place’.

Y.8.5b

āpō < *āp-ah*: voc. plur.

**ruuarā* < *ruuar-āh*: voc. plur.

vīspa < *uičy-ā*: acc. plur. neut.

vohū < *uāh-ū*: acc. plur. neut.

ašaciōra < *arta-čiōr-ā* ‘whose seed is from(?) Order’: *a*-stem adj., acc. plur. neut.; adjectival compound < *arta-* + *čiōra-* ‘seed’.

Y.8.5c

xšaiiamnəm < *xšaja-mn-am* < *xš-aja-* ‘to rule’ < *xšā-*: *aja*-stem pres. part. mid., acc. sing. masc.

ašauuanəm < *artāuan-am* < *artā-uan-lart-āun-* ‘Orderly, who sustains Order’: *uan*-stem adj. acc. sing. masc.; < *arta* + *uan-*.

dāiūata < *dāja-ta* < *dā-ja-* < *d^hā-* ‘place’: *ja*-stem pres. imper. 2p.

axšaiiamnəm < *a-xšaja-mn-am* ‘not in command, not ruling’: *a*-stem adj., acc. sing. masc.; adjectival compound < *a-* + *xšaja-mna-*.

druuantəm < *druguant-am* < *drug-uant-ldrug-uat-* (*drug-unt-*) ‘possessed by the Lie’: *uant*-stem adj., acc. sing. masc.; < *drug-ldrug-* ‘the (cosmic) Lie’ (female principle of deception), together with the Evil Spirit Ahura Mazda’s principal opponent. Originally perhaps Chaos. She deceives gods and men as to the true nature of the ordered cosmos.

Y.8.8

rauuas-ca x^vāōrəm-ca āfrīnāmi vīspaiiā ašanonō stōiš

qzas-ca dužāōrəm-ca āfrīnāmi vīspaiiā druuatō stōiš

I invite the open space and easy breathing of the entire (temporal) existence of the Orderly one.

I invite the constriction and laboured breathing of the entire (temporal) existence of the Orderly one.

Y.8.8a

rauuas^o < *raū-ah* ‘open space’: *ah*-stem neut., acc. sing.

x^vāōrəm < *hu-āōr-am* < *hu-āōr-a-* (< *H^hH-ōra-*) ‘good breathing (space), easy breathing, comfort’: *a*-stem neut., acc. sing.; nominal compound < *hu-* ‘good’ + *āōra-* < *anH-* ‘to breathe’.

āfrīnāmi < *ā-frīnā-mi* < *ā-frī-nā-lā-frī-n-* (< *fri-na-H-lfri-n-H-*) ‘to invite as (guest) friend(?)’: *nā*-stem athem. pres. ind. 1s act. (performative).

vīspaiiā < *uičy-āi-āh* < *uičy-a-*: fem. sing.

ašanonō < *artā-un-ah* < *artā-uan-lartā-un-*: gen. sing. masc.

stōiš < *stai-š* < *s-ti-* ‘(duration of) being, (temporal) existence’: *i*-stem fem., gen. sing.; < *ah-* ‘to be’.

Y.8.8b

qzas^o < *anj-ah* ‘constriction’: *ah*-stem neut., acc. sing.; cf. *angst*.

dužāθrəm < *duž-āθr-am* < *duž-āθra-* ‘bad (constricted) breathing, discomfort’: *a*-stem neut., acc. sing.; nominal compound, antonym of *hu-āθra-*.

druuatō < *druguat-ah* < *drug-uant-/drug-uat-*: gen. sing. masc.

Most of the *Young Avesta* is written in a basically octosyllabic metre, of which the following text from the hymn to Miθra is a sample. In Yt.10.12, the first line is hypermetric.

Yt.10.12

miθrəm *vo^hru.gaoiiao^ttīm yazamaⁱde*
arš.vacaŋhəm viiāxanəm
hazaŋra.gaošəm hutāštəm
baēuuar^h.cašmanəm bər^hzantəm
pər^hθu.vaēdaiianəm sūrəm
ax^hafnəm jayā^hruuāŋhəm

We sacrifice to Miθra with wide grazing grounds, of truthful speech, *eloquent, with a thousand ears, well-fashioned, with ten thousand eyes, tall, with broad outlook, strong, sleepless, waking.

miθrəm < *miθr-am* ‘Miθra’: *a*-stem masc. acc. sing.

vo^hru.gaoiiao^ttīm < *uaru-gau-iaut-im* ‘having wide grazing grounds’: *i*-stem adj., masc. acc. sing.; possessive adjectival compound < *uaru-* ‘broad’ and *gau-iaut-i-* of uncertain exact meaning; cf. OInd. *gav-yūti-*.

yazamaⁱde < *yaza-*: pres. ind. 1s mid. (see Y.8.3a).

arš.vacaŋhəm < *arš-ūacah-am* ‘of truthful speech’: *ah*-stem adj., masc. acc. sing.; possessive adjectival compound < *arš* ‘straight, truthful(ly)’ < *Hrj-š* (OAv. *ər^hš*).

viiāxanəm < *uiiāxan-am* (4 sylls.) word of uncertain meaning: *a*-stem adj., masc. acc. sing.

hazaŋra.gaošəm < *hajāhra-gauš-am* ‘having a thousand ears’: *a*-stem adj., masc. acc. sing., possessive adjectival compound < *hajāhra-* ‘1000’ and *gauša-* ‘ear’.

hutāštəm < **hu-tašt-am* ‘well-fashioned’: *a*-stem adj., masc. acc. sing.; the form with long vowel analogical with the present *tāš-ti* ‘he fashions’.

baēuuar^h.cašmanəm < *baiuar-čašm-an-am* ‘having ten thousand eyes’: *man*-stem adj., masc. acc. sing.; possessive adjectival compound < *bai-uar-* ‘10,000’ and *čaš-man-* ‘eye’ (neut.).

bər^hzantəm < *brjant-am* ‘tall’: *ant*-stem adj., masc. acc. sing.

pər^hθu.vaēdaiianəm < *pṛθu-uidājan-am* ‘with broad outlook’: *a*-stem adj., masc. acc. sing.; possessive adjectival compound < *pṛθu-* (< *pṛtH-u-*) ‘broad’ and *uid-ājan-*, derivative of *ui-dāy-* ‘see far and wide’.

sūrəm < *cūr-am* (< *cuH-ra-*) ‘rich in life-giving strength’: *a*-stem adj., masc. acc. sing. (see Y. 28.9c on *sauuajhəm*).

ax^hafnəm, *a-huafn-am* ‘sleepless’: *a*-stem adj., masc. acc. sing.; from *a-* + *huaf-na-* ‘sleep’.

jayā^hruuāŋhəm < *ja-gar-ūāh-am* ‘waking’: pf. part. act. of *garH-/grā-* ‘wake’; the diphthong *-ā^h* perhaps substituted for *-ao^h* (de Vaan 2003, §17.4.2) but perhaps with vowel metathesis, cf. OInd. *jā-gar-*.

Yt. 10.13

- a *yō paō'riiō manīiauuō yazatō*
 b *tarō harçm *āsənao'ti*
 c *pa^aruua.naēmāꞤ aməṣahe*
 d *hū yaꞤ a^aruuaꞤ.aspahe*
 e *yō paō'riiō zaraniiō.pīsō*
 f *srīrā *bar³šnauuō gər³βnā'ti:*
 g *aḏāꞤ vīspəm ādiḏā'ti*
 h *a^ariiō.šaiianəm səuiištō*

Who as the first in the other world to be sacrificed to
 rises over Harā,
 in front of the immortal
 sun with fleet horses,
 who, as the first, seizes
 the gold-adorned, beautiful heights.
 From there, he surveys the entire
 settlement of the Aryans, he most rich in life-giving strength.

a

paō'riiō < *parū-īia-h* (see Y. 28.11c): *a*-stem adj. masc. nom. sing.
manīiauuō < *manī-au-ah* 'of the world of thought: *a*-stem adj., masc. nom.sing.; <-
manīu- (see Y. 28.11c).
yazatō < *īaj¹-at-ah* 'being worthy of sacrifices' <*īaja-* (cf. Y.8.3a): *a*-stem, masc. nom.
 sing.

b

tarō < *tar-ah* 'over, beyond' < *tṛH-* 'cross over': prep. + acc.
harçm < *har-ām* 'Harā' a mountain: *ā*-stem fem. acc. sing.
 **āsənao'ti* (mss. *āsnaoiti*) < *ā-sa-nau-* 'rise' < *ā-sṇ-nau-*: *nau*-stem pres. ind. act. 3s.

c

pa^aruua.naēmāꞤ < *parūa-naim-āt* 'from in front of': *a*-stem masc. abl. sing.; compound
 < *pa^aruua-* 'before' and *naēma-* 'half, side'.

d

hū < *huṃaṅ-h* < *hū-uar/n-* 'sun': *uar/n*-stem neut. gen. sing.
yaꞤ: relative particle (see section (6.1.1.1a)).
a^aruuaꞤ.aspahe < *arūat-acū-ahīa arūat-acūa-* 'having fleet horses: *a*-stem adj. masc. acc.
 sing.; possessive adjectival compound < *arūant-* 'fleet' and *acūa-* 'horse'.

e

zaraniiō.pīsō < *j^haranīa-piç-ah* 'gold-adorned': cons.-stem adj., fem. nom. plur.;
 adjectival compound < *j^haranīa-* 'gold' and root noun *piç-* 'paint, adorn', cf. OPers.
ni-piθ- 'write'.

f

srīrā < *crīr-āh* ‘beautiful’ < *crāiH-/cīriH-*: *a*-stem adj., fem. nom.plur.

**bar³šnauuō* (ms. F1 °auua) < *barj^hn-au-ah* ‘height’: *u*-stem fem. nom. plur.; this nom. phrase is used for expected acc.; for the gender and reading, cf. Yt. 18.6 *vīspā bar³šnauuō g^hrinqm* ‘all the heights of the mountains’.

g^hβnā^hti < *grb-nā-ti* < *g^hrb^h-nā-lg^hrb^h-n-* ‘take, seize’: *nā*-stem pres. ind. act. 3s.; also *g^hruuāiia-* < *grbāya-*, cf. OPers. *garbāya-* (DNa 15–22b).

g

adā^hi < *ad^h-āt* ‘from there’;

ādīdā^hti < *ā-di-dā-ti* ‘surveys’ < *dāy- /dī-* (see Y.8.3a): red. pres. ind. act. 3s.

h

driiō.šaiianəm < *ariā-cšaiian-am* ‘settlement of the Aryans’: *a*-stem, masc. acc. sing.; compound of *ariā-* ‘Aryan, Iranian’ and *cšai-ana-*, derivative of *cšai-* ‘dwell, be settled’ (see Y.53.8c).

sauuištō < *cauH-išt-ah* ‘most endowed with life-giving strength’ (see Y. 28.9c on *sauuarhqm*): *a*-stem superl., masc. nom. sing.

11.3 Old Persian

DNa 15–22

- a *Ōātiy dārayavauš xšāyaθiya*
- b *vašnā auramazdāha imā dahayāva tayā adam aqarbāyam apataram hacā pārsā*
- c *adam-šām patiyaxšayaiy manā bājim abaraha*
- d *taya-šām hacā-ma aḏahaya ava akunava*
- e *dātam taya manā ava-dī[š] adāraiya*

King Darius announces:

By the greatness of Ahuramazdā, these (are) the lands that I seized further away from Persia.

I ruled over them. They brought me tribute.

Whatever was announced to them from/by me, that they did.

My law held them.’

a

Ōātiy < *caṅha-ti* < *caṅha-* ‘to announce’: *a*-stem pres. ind. (performative) 3s act.

dārayavauš < *dāraja-vahu-š*: *u*-stem proper name masc., nom. sing.; nominal compound < *dāraja-*, pres. stem. of *dar-* ‘to (up)hold’ and *uahu-* ‘good (things)’.

xšāyaθiya < *xšājaθ-ja-h* ‘king’: *a*-stem masc., nom. sing.; *ja*-stem ‘vridhhi’ derivative of **xšaj-aθa-* ‘rule’ < pres. *xšaja-* ‘to rule’.

b

vašnā < *uajnā* < **uajar/n-* ‘greatness’: *r/n*-stem neut., instr. sing.; cf. *uajr-ka-* (*vazarka-*) ‘great’.

auramazdāha < *ahura-mazdāh* + *ah* < *ahura-mazdā-*: gen. sing.; the gen. sing. ending *-ah* has been added to the gen. *mazdāh* < *mazdaH-ah*.

imā < *im-āh* ‘these’: dem. pron., acc. plur. fem.

dahçiyāva < *dahç-āu-ah* < *dahç-au-* ‘land’: *au*-stem fem., nom. plur.

tayā < *ta* + *ç-āh*: rel. pron., acc. plur. fem.

adam < *açam* ‘I’: pers. pron. 1s, nom.; Av. *azəm*.

agçrbāyam < *a-grçb-āçia-m* < *grçb-āçia-* < *g^hrab^h*- ‘to seize’: *āçia*-stem pres., imperf. 1s act.; cf. Yt. 10.13 f.

apataram < *apa-tara-m* < *apa-* ‘away’: adverbial acc. sing. neut., comparative.

hacā < *haca* ‘from’ (+ instr.-abl.).

pārsā < *pārsāt* < *pārsa-* ‘Persia’: instr.-abl. sing., abl. function.

c

-šām ‘them’: pers. pron. 3rd pers. enclitic, gen.-dat. plur., gen. function.

patiyaxšayaiy < *pati-a-xšaçi-ai* < *pati-xšaçia-* mid. ‘to rule over’: *açia*-stem pres., imperf. 1s mid.

manā < *mana* ‘me’: pers. pron. 1s gen.-dat., gen. function.

bājim < *bāçi-m* ‘tribute’: *i*-stem (masc./fem.?), acc. sing.

abaraha < *a-bara-n* (error?) < *bara-* ‘to carry’: *a*-stem pres., imperf. 3p; in principle, the form could be read as *ā-bara-* ‘to bring’.

d

taya < *taçi-at* ‘that which, what’: rel. pron., nom. sing. neut.; see section 3.4.4. *-ma* < *-ma-t* ‘me’: pers. pron. 1s (instr.-abl. enclitic, abl. function).

a0ahçya < *a-çahçia-t* < *çah-çia-* < *çhçia-* or < *çahç-çia-* ‘to be announced’: *çia*-stem passive of *çahça-*, imperf. 3s act.

ava < *au-at* ‘that’: dem. pron., acc. sing. neut.

akunava < *a-ku-nau-an* < *kç-nau-lkç-nu-* ‘to do’: *nau*-pres. stem, imperf. 3p act.; commonly regarded as an ‘allegro’ form with *kun-* < *kçn-*, cf. Av. *kær^o-nau-*, but perhaps analogical (rhyming) with forms such as **sru-nau-* (Av. *s^orunao-*).

e

dātam < *dāt-am* ‘law’: *a*-stem neut., nom. sing.; substantivised past part. *dāta-* ‘placed, set down’ from *d^hā-*

-dīš < *^oi-nš* ‘them’: pers. pron. 3rd pers., acc. plur. masc.; for **i-nš*.

adāraçia < *a-dāraçia-t* < *dāraçia-* ‘to hold (firmly in place)’: *açia*-stem pres., imperf. 3s act.

DNa 31–38

a *auramazdā ya0ā avaina imām būmim *yauda^otīm pasāva-dim manā frābara*

b *mām xšāya0iyam akunauš adam xšāya0iya a^hmiy*

c *vašnā auramazdāhā adam-šim gā0avā niyašādayam*

d **taya-šām adam a0aham ava akunava *ya0ā mām kāma āha*

When Ahuramazdā saw this earth being in turmoil, then he gave it to me.

He made me king. I am king.

By the greatness of Ahuramazdā, I set it down in its place.

Whatever I announced to them, that they did as was my wish.

a

auramazdā < *ahura-mazdā-h*: nom. sing.; Av. *ahurō mazdā* < *ahura-h mazdā-h*.

yaθā < *ja-θa* ‘when’: conjunction.

avaina < *a + uaina-t* < *uaina-* ‘to see’: *a*-stem pres., imperf. 3s act.

imām < *im-ā-m* ‘this’: dem. pron., acc. sing. fem.

būnim < *būm-im* ‘earth’: *i*-stem fem., acc. sing.; cf. Av. *būmī-*.

yauđa“*tūm* < *iaujā-nt-ī-m* < *iaujā-* ‘to be in turmoil, be chaotic’: pres. part., derived *ī*-stem fem., acc. sing.; cf. Av. *yaoza-*.

pasāva < *pasća + aya-t* ‘afterward, then’: adverbial compound < *pasća* ‘after’ + *aya-t* ‘that’.

-dim < *đm* ‘it’: pers. pron. 3rd pers., acc. sing. fem. enclitic; it is not known whether the form had preserved the long *ī*, cf. OAv. *hīm*.

manā < *mana* ‘me’: pers. pron. 1s, gen.-dat., dat. function.

frābara < *fra-a-bara-t* < *fra-bara-* ‘to proffer, give’: *a*-stem pres., imperf. 3s act.

b

mām < *mām* ‘me’: pers. pron. 1s, acc.

xšāyaθiyam < *xšājaθi-am*: acc. sing.

akunauš for *a-kunau* < *a-kunau-t* < *kunau-* ‘to do, make’: *nau*-pres. stem, imperf. 3s act.

a^hmiy < *ah-mi* < *ah-* ‘to be’: pres. ind. 1s act.

c

-šim < *-š-īm*: pers. pron. 3rd pers., acc. sing. fem. enclitic; cf. OAv. *hīm*.

gāθavā < *gāθau* + *ā* < *gātu-* (*gāθu-*) ‘place, platform (throne)’: *u*-stem loc. sing. + postpos. *-ā*.

niyašādayam < *ni-a-šādaja-m* < *ni-šādaja-*, caus. of *ni-šad-* < *had-* ‘to sit’: *aja*-stem (caus.) pres., imperf. 1s; the form has been analogically remade from **ni-a-hādaja-*.

d

aθaham < *a-θaḡha-m*: imperf. 1s act.

kāma < *kāma-h* ‘wish’: *a*-stem masc., nom. sing.; this noun takes a personal dir. obj.: *mām kāma^h* ‘I wish’ see section 6.4.3.4.

āha < *āha-t* < *ah-* ‘was’: imperf. 3s act.; the form is a thematised replacement for **ās-(t)*, which would probably have been reduced to **ā* in OPers.; the long initial *āh-* is < *a-Hh-*.

DNa 38–47

- a *yadi-patīy maniyāhay taya ciyakaram [āha] *avā dahayāva tayā dārayavauš xšāyaθiya adāraya*
- b *patīkarā dīdiy tayaiy gāθum bara^htiy*
- c *adā xšnāsāḡay ada-taiy azdā bavātiy pārsahayā martiyahayā dūrai y arštiš parāgmatā*
- d *ada-taiy azdā bavātiy pārsa martiya dūray hacā pārsā parataram patiyajatā*

If, again, you think – ‘How were those lands which King Darius held?’ – then look at the pictures (of the men) who carry the throne.

Then you will know, then it will become clear to you that the Persian man’s spear has gone far away.

Then it will become clear to you that the Persian man *defended himself far beyond Persia.

a

yadi- < *yadi* ‘if’: conj.

-patiy < *-pati* ‘in addition’(?): enclitic particle.

maniyāhay < *man-iā-hai* < *man-* ‘to think’: *īa*-stem. pres. subj. mid.

taya < *ta* + *īa-t* ‘that’: conj., here introducing dir. speech.

ciyakaram < *čīā-kar-am* ‘how?’ (or: ‘how many?’): adverbial compound < *čīā-* ‘how much(?)’ + *kara-* ‘work’, i.e. ‘of what kind of work is he capable?’.

āha < *āh-an* < *ah-* ‘to be’: imperf. 3p act.

b

patikarā < *pati-kar-ā* ‘pictures’: *a*-stem masc. acc. plur.; nominal compound < *pati* + *kara-* ‘sth. made to reflect’.

dīdiy < *dī-di* ‘look at!’, pres. *vaina-*: imper. 2s act.; perhaps originally an aor. stem.

tayaiy < *ta+īai* ‘who’: rel. pron. nom. plur. masc.

gā0um < *gāt-u-m* < *gātu-*: acc. sing.; see section 2.3.2.5.

bara'tiy < *bara-nti* ‘to carry’: *a*-stem pres. ind. 3p act.

c

adā < *ada* ‘then’: temporal adv.

xšnāsāhay < *xšnāčā-hi* < *xšnā-sca-* ‘to know’: pres. subj. 2s act.; *xšnā-sca-* < *jnā-*, cf. *dānā-* ‘know’ < *jā-nā-* (for *jñ-na-H-*).

-taiy < *-tai* ‘you’: pers. pron. 2s, gen.-dat., dat. function.

azdā < *azdā* ‘known, clear’ (< *ad'dhā*, OInd. *addhā*).

bavā'tiy < *bau-ā-ti* < *bau-lbū-* ‘to become’: *a*-stem pres. subj. 3s act.

d

pārsahayā < *pārsa-hja* ‘Persian’: *a*-stem adj., gen.-dat. sing. masc., gen. function.

martiyahayā < *martija-hja* ‘man’: *a*-stem masc., gen.-dat. sing., gen. function.

dūrai < *dūrai* < *dūra-* ‘far’: adverbial loc. sing.; see section 3.3.

arštiš < *ršti-š* ‘spear’: *i*-stem fem., nom. sing.

parāgmatā < *parā-gmatā* < *parā-gam-* ‘to go away’(?): intrans. perf. nom. sing. fem. (see section 7.4.3.1); cf. YAv. *γ³mata-*.

e

pārsa < *pārsa-h:* nom. sing. masc.

martiya < *martija-h:* nom. sing.

parataram < *para-tar-am* < *para-* ‘beyond’: adverbial acc. sing. neut., comparative.

patiyajatā < *pati-a-ja-ta* < *pati-jan-ljn-* (< *jñ-*) ‘to strike back’, mid. ‘defend oneself’(?): imperf. 3s mid.

DNa 48–55

- a *Oātiy dārayavauš xšāyaθiya*
- b *aita taya k̄artam ava visam vašnā auramazdāhā akunavam . . .*
- c *mām auramazdā pātuv hacā gastā utā-maiy viθam utā imām dah̄yāum*
- d *aita adam auramazdām jadiyahāmiy aita-maiy auramazdā dadātuv*

King Darius announces:

All this that has been done I did by the greatness of Ahuramazdā.

Let Ahuramazdā protect me from foulness, as well as my house and this land!

This I ask Ahuramazdā for. Let Ahuramazdā give me this!

b

aita < *aita-t* ‘this’: dem. pron., nom.-acc. sing. neut.

k̄artam < *k̄rt-am* < *kar-* ‘to do’: intrans.-pass. perf., nom.-acc. sing. neut.

visam < *uičū-am* ‘all’: nom.-acc. sing. neut.

akunavam < *a-kunaū-amr.* imperf. 1s act.

c

pātuv < *pā-tu* < *pā-* ‘to protect’: imper. 3s act.; there are no pres. forms attested of this verb, but cf. *pati-paya-^huvā* (DB 4.38) < *pati-paiā-huā* pres. imper. 2s mid.

gastā < *gast-āt* < *gasta-* ‘foulness’: *a*-stem neut.(?) instr.-abl. sing., abl. function; past part. *gas-ta-* ‘foul-smelling’ < *gnt^s-ta-*, replacing **gazda-* < **gnd^r-d^ha-* < *gand^h-* ‘to smell foul’.

utā- < *uta* ‘and’: conj.

-mai < *-mai* ‘me’: pers. pron. 1s, gen.-dat., gen. function.

viθam < *uičū-am* ‘house’: root noun fem.(?), acc. sing.

dah̄yāum < *dahy-āu-m* < *dahj-au-*: acc. sing.

d

auramazdām < *ahura-mazda-am.* acc. sing.; < *mazdaH-am.*

jadiyahāmiy < *jad-īā-mi* < *jad-ī-a-* ‘to ask’ (+ acc. of pers. + acc. of thing): *īa*-stem pres. (performative) ind. 1s act.

dadātuv < *da-dā-tu* < *dā-* ‘to give’: reduplicated pres. imper. 3s act.

ABBREVIATIONS OF TEXTS

Avestan texts

Āfr.	Āfrīnagān
FrD.	text edited in Hoffmann 1968.
H.	Hādōxt nask
Her.	Hērbedistān
N.	Nīrangistān
V.	Videvdad
Vr.	Vispered
Y.	Yasna
Yt.	Yasht

Old Persian texts

- DB Darius at Behistun (Bisutun), large inscription
 DBa . . . Darius at Behistun, smaller inscriptions
 DE Darius at Elvand
 DNa, b Darius at Naqsh-e Rostam
 DPa . . . Darius at Persepolis
 DSa . . . Darius at Susa
 DZa . . . Darius at Suez
 XPa . . . Xerxes at Persepolis
 XV Xerxes at Lake Van

BIBLIOGRAPHY AND REFERENCES**Linguistic descriptions**

The most recent descriptions of phonology, morphology (and some syntax) are Abaev and Bogoliubov (eds) 1979 (I. M. Oranskij: overview; S. N. Sokolov: Avestan and OPers; Abaev: Scythian-Sarmation); Schmitt, 1989 (Schmitt: OPers.; Kellens: Avestan); Hoffmann and Forssman 1996; Skjærvø 2006a (pre-history), 2007b (morphology, morpho-phonology). De Vaan 2003 covers a large section of the phonology. The only complete syntax is still the one in Reichelt 1909. Kellens and Pirart 1990 contains select topics of OAv. syntax. Wackernagel and Debrunner 1896–1964 is still indispensable, notably vol. II, 2 on nominal suffixes.

Bibliographies

A comprehensive bibliography of Avestan up to 1967 is found in Schlerath 1968, additional bibliographies are found in the histories of studies listed below. Hoffmann and Forssman 1996 lists bibliographical items for each chapter. De Vaan 2003 also contains an extensive bibliography. For OPers., see Kent 1953 and Brandenstein and Mayrhofer 1964, as well as Schmitt 1991 and 2000.

- Abaev V. I., M. N. Bogoliubov, V. S. Rastorgueva (1979) (eds) *Osnovy iranskogo iazykoznanija: drevneiranskije iazyki*. Moskva: Nauka (= *Osnovy I*).
- Adiego Lajara, I.–X. (2000) ‘Sobre el relativo del antiguo persa’, in: B. Forssman and R. Plath (eds) *Indoarisch, Iranisch und die Indogermanistik. Arbeitstagung der Indogermanischen Gesellschaft vom 2. bis 5. Oktober 1997 in Erlangen*. Wiesbaden: L. Reichert, pp. 1–13.
- Allegri, M., and A. Panaino (1995) ‘On the *š*-Ending in OPers. *akunauš* and Similar Forms. With a Contribution by Ilya Gershevitch’, in: B. G. Fragner *et al.* (eds) *Proceedings of the Second European Conference of Iranian Studies*. Roma: Istituto Italiano per il Medio e Estremo Oriente, pp. 1–33.
- Bartholomae, Ch. (1895) ‘Vorgeschichte der Iranischen Sprachen’ and ‘Awestasprache und Altpersisch’, in: Geiger and Kuhn (eds), pp. 1–151, 152–248.
- (1904) *Altiranisches Wörterbuch*. Strassburg: K. J. Trübner; repr. Berlin: W. de Gruyter, 1961.
- Beekes, R. S. P. (1981) ‘The Neutre Plural and the Vocalization of the Laryngeals in Avestan’, *IJ* 23, pp. 275–87.
- (1988) *A Grammar of Gatha-Avestan*. Leiden: Brill.
- Brandenstein, W., and M. Mayrhofer (1964) *Handbuch des Altpersischen*. Wiesbaden: O. Harrassowitz.

- Cantera, A. (2001) 'Die Behandlung der Lautfolge (C)rHC- im Iranischen', *MSS* 61, pp. 7–27.
- Darmesteter, J. (1892–1893) *Le Zend-Avesta*, 3 vols. Paris: E. Leroux; repr. Paris: A. Maisonneuve, 1960.
- Duchesne-Guillemin, J. (1936) *Les composés de l'Avesta*. Liège: Faculté de Philosophie et de Lettres; Paris: E. Droz.
- (1962) 'L'étude de l'iranien ancien au vingtième siècle', *Kratylos* 7, pp. 1–44.
- Fortson, B. J. (1996) 'The Double Reflex of *ai in Avestan', *IJJ* 39, pp. 41–54.
- Geiger, W., and E. Kuhn (1895–1901) (eds) *Grundriss der Iranischen Philologie*, ed. 2 vols. Strassburg: K. J. Trübner; repr. Berlin and New York, 1974.
- Gershevitch, I. (1967) *The Avestan Hymn to Mithra*. Cambridge: Cambridge University Press.
- Geldner, Karl Friedrich (1886–96) *Avesta the Sacred Books of the Parsis*. 3 vols., Stuttgart: W. Kohlhammer.
- Gippert, J. (1985) 'Verbum dicendi + Infinitiv im Indoiranischen', *MSS* 44, pp. 29–57.
- Hale, M. (1988) 'OPers. Word Order', *IJJ* 31, pp. 27–40.
- (1993) 'Tmesis and Movement in Avestan', *IJJ* 36, pp. 29–43.
- Hintze, A. (1994) *Zamyād-Yašt. Edition, Übersetzung, Kommentar*. Wiesbaden: L. Reichelt.
- (2005) "Indo-Iranian gar 'to raise aloft'. In: G. Schweiger (ed.), *INDOGERMANICA – Festschrift Gert Klingenschmitt: Indische, iranische und indogermanische Studien dem verehrten Jubilar dargebracht zu seinem hundertsechzigsten Geburtstag*. Taimering (Riekofen): Schweiger VWT, 2005 (Studien zur Iranistik und Indogermanistik; Bd. 3), pp. 237–260.
- (2007) *A Zoroastrian Liturgy: The Worship in Seven Chapters (Yasna 35–41)*. Wiesbaden: O. Harrassowitz.
- Hinz, W. (1973) *Neue Wege im Altpersischen*. Wiesbaden: O. Harrassowitz.
- Hoffmann, K. (1952) 'Zum prädikativen Adverb', *MSS* 1, pp. 42–53 (= *Aufsätze zur Indoiranistik* II, pp. 339–49).
- (1958) 'Altiranisch', in: *Handbuch der Orientalistik* I, IV, 1. Leiden and Cologne: Brill, pp. 1–19.
- (1965) 'Zu den altiranischen Bruchzahlen', *KZ* 79, pp. 247–5 (= *Aufsätze zur Indoiranistik* I, pp. 182–90, and addendum p. 338).
- (1967) *Der Injunktiv im Veda. Eine synchronische Funktionsuntersuchung*. Heidelberg: C. Winter.
- (1968) 'The Avesta Fragment FrD. 3', *IJJ* 10, pp. 282–88 (= *Aufsätze zur Indoiranistik* I, pp. 221–27).
- (1970) 'Zur awestischen Textkritik. Der Akk. Pl. mask. der a-Stämme', in: M. Boyce and I. Gershevitch (eds) *W. B. Henning Memorial Volume*. London: Lund Humphries, pp. 187–200 (= *Aufsätze zur Indoiranistik* I, pp. 274–87).
- (1975, 1976, 1992) *Aufsätze zur Indoiranistik*, 3 vols. Wiesbaden: L. Reichelt.
- (1976) 'Präteritaler Optativ im Altiranischen', in: *Aufsätze zur Indoiranistik* II, pp. 605–19.
- (1988) Avestan Language. i. The Avestan Script, ii. The Phonology of Avestan, iii. The Grammar of Avestan', in: *EnIr* III, pp. 47–62.
- Hoffmann, K., and B. Forssman (1996) *Avestische Laut- und Flexionslehre*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- Hoffmann, K., and J. Narten (1989) *Der Sasanidische Archetypus*. Wiesbaden: L. Reichelt.
- Humbach, H. (1991) *The Gāthās of Zarathustra and the Other OAv. Texts*, 2 vols. Heidelberg: C. Winter (with the collaboration of J. Elfenbein and P. O. Skjærvø).

- Insler, S. (1975) *The Gāthās of Zarathustra* (= Acta Iranica 8). Tehran and Liège: Brill.
- Jackson, A. V. W. (1892) *An Avesta Grammar in Comparison with Sanskrit*. Stuttgart: W. Kohlhammer; repr. Darmstadt: Wissenschaftliche Buchgesellschaft, 1968; repr. New York: AMS Press, 1975.
- Jamison, S. W. (1983) *Function and Form of the -āya-Formations of the Rig Veda and Atharva Veda*. Göttingen: Vandenhoeck & Ruprecht.
- (1992) 'Vedic "sá figé". An Inherited Sentence Connective?', *Historische Sprachwissenschaft* 105: 213–39.
- (1997) 'Vedic anyá- 'another, the other': Syntactic Disambiguation', in: A. Lubotsky (ed.) *Sound Law and Analogy. Papers in Honor of Robert S. P. Beekes*. Amsterdam and Atlanta: Rodopi, pp. 111–18.
- Jasanoff, J. H. (1997) 'Gathic Avestan cikōitərəš', in: A. Lubotsky (ed.) *Sound Law and Analogy. Papers in Honor of Robert S. P. Beekes*. Amsterdam and Atlanta: Rodopi, pp. 119–30.
- Kellens, J. (1971) 'L'avestique de 1962 à 1972', *Kratylos* 16, pp. 1–30.
- (1973) 'L'avestique de 1962 à 1972. Addenda et corrigenda', *Kratylos* 18, pp. 1–5.
- (1974) *Les noms-racines de l'Avesta*. Wiesbaden: L. Reichert.
- (1984) *Le verbe avestique*. Wiesbaden: L. Reichert.
- (1986) 'Comment faut-il éditer les formes vieil-avestiques de paoiriia-?', in Schmitt, R. and Skjærvø, P. O. (eds) *Studia Grammatica Iranica: Festschrift für Helmut Humbach*, München: R. Kitzinger, pp. 217–226.
- (1991) 'L'avestique de 1972 à 1990', *Kratylos* 36, pp. 1–31.
- (1994) 'Retour à l'infinitif avestique', *MSS* 55, pp. 45–59.
- (1995) *Liste du verbe avestique avec un appendice sur l'orthographe des racines avestiques par Eric Pirart*. Wiesbaden: L. Reichert.
- (1998) 'Considérations sur l'histoire de l'Avesta', *JA* 286, pp. 451–519.
- Kellens, J., and E. Pirart (1988, 1990, 1991). *Les textes vieil-avestiques*, 3 vols. Wiesbaden: L. Reichert.
- Kent, R. G. (1953) *Old Persian Grammar, Texts, Lexicon*, 2nd rev. ed. New Haven: American Oriental Society.
- Klein, J. (1988) 'Coordinate Conjunction in Old Persian', *JAOS* 108, pp. 387–417.
- Lazard, G. (1976) 'Notes de vieux-perse', *BSL* 71, pp. 175–92.
- (1984) 'La métrique de l'Avesta récent', in: *Orientalia Jacques Duchesne-Guillemain Emerito Oblata*, Acta Iranica 23, pp. 283–300.
- (1990) 'Composition et métrique dans les yashts de l'Avesta', in: G. Gnoli and A. Panaino (eds) *Proceedings of the First European Conference of Iranian Studies held in Turin, September 7th–11th, 1987 by the Societas Iranologica Europaea*. Part 1: *Old and Middle Iranian Studies*, Serie Orientale Roma. Vol. 67.1. Roma: Istituto Italiano per il Medio ed Estremo Oriente, pp. 217–28.
- Lühr, R. (1994) 'Zur Interdependenz der Methoden 'Funktionsbestimmung' und 'Rekonstruktion' – das Infinitivmorphem indoiran. *-d^hiāi', *MSS* 44, pp. 69–97 [pub. 1995].
- Mayrhofer, M. (1980) 'Zu iranischen Reflexen des vrkī Typus', in: J. Bingen et al. (eds) *Recherches de linguistique. Hommages à Maurice Leroy*. Bruxelles: Éditions de l'Université de Bruxelles, pp. 130–52 (= *Ausgewählte Kleine Schriften II*, pp. 353–75).
- (1981) 'Laryngalreflexe im Indo-Iranischen', *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 34, pp. 427–38 (= *Ausgewählte Kleine Schriften II*, pp. 292–303).
- (1996) *Ausgewählte Kleine Schriften II*. Wiesbaden: L. Reichert.
- Morgenstierne, G. (1942) 'Orthography and Sound System of the Avesta', *NTS* 12, pp. 30–82; repr. in: *Irano-Dardica*, pp. 31–83.

- Oettinger, N. (1985) 'Beobachtungen anhand des avestischen Textes Yašt 5, 90–93', *MSS* 45 (*Festgabe für Karl Hoffmann* II), pp. 183–95.
- (1992) 'Zu den Verben auf vedisch *-anyá-* und hethitisch *amje-*', *MSS* 53, pp. 133–54.
- Panaino, A. (1989–90) 'Il periodo ipotetico dell'irrealtà in avestico recenziore', *Atti del Sodalizio Glottologico Milanese* 31, pp. 124–48.
- (2007) 'Chronologia Avestica: tra cronologia linguistica e storia religiosa (Filologia e storia del testo avestico, 1.)', in: A. Panaino and V. Sadovski, *Disputationes iranologicae Vindobonenses, I*, Milan, pp. 7–33.
- Pirart, E. (1988) 'L'instrumental des neutres vieil-avestiques en *-an-*', *StIr* 17, pp. 141–46.
- (1993) 'Avestique *janiiaoš*', *IJJ* 36, pp. 337–40.
- Ravnæs, E. (1981) 'The Development of *ə*/Interconsonantal Laryngeal in Iranian', *IJJ* 23, pp. 247–73.
- Reichert, H. (1909) *Awestisches Elementarbuch*. Heidelberg: C. Winter; repr. Darmstadt: Wissenschaftliche Buchgesellschaft, 1967; Heidelberg: C. Winter, 1978.
- Schindler, J. (1987) 'Zur avestischen Kompositionslehre: *aš-* "groß"', in: G. Cardona and N. H. Zide (eds) *Festschrift for Henry Hoeningwald . . .* Tübingen: G. Narr, pp. 337–48.
- (1982) 'Zum Nom. sg. m. der *nt*-Partizipien im Jungavestischen', in: E. Neu (ed.) *Investigationes Philologicae et Comparativae. Gedenkschrift für Heinz Kronasser*. Wiesbaden: O. Harrassowitz, pp. 186–209.
- Schlerath, B. (1968) *Avesta-Wörterbuch. Vorarbeiten*, 2 vols. Wiesbaden: O. Harrassowitz.
- Schmitt, R. (1980) 'Altpersischforschung in den Siebzigerjahren', *Kratylos* 25, pp. 1–66.
- (1989) (ed.) *Compendium Linguarum Iranicarum*. Wiesbaden: L. Reichert.
- (1991) *The Bisitun Inscriptions of Darius the Great*, *Corpus Inscriptionum Iranicarum* I, I, Texts I. London: Corpus Inscriptionum Iranicarum.
- (1994) 'Altpersisch /*yadā*/ "wo"', in: P. Cipriano, P. di Giovine, and M. Mancini (eds) *Miscellanea di studi linguistici in onore di Walter Belardi*. Roma: Il Calamo, pp. 99–106.
- (1995) 'Direkte und indirekte Rede im Altpersischen', in: M. Ofitsch and Ch. Zinko (eds) *Studia Onomastica et Indogermanica. Festschrift für Fritz Lochner von Hüttenbach zum 65. Geburtstag*. Graz: Leykam, pp. 239–47.
- (1999) *Beiträge zu altpersischen Inschriften*. Wiesbaden: L. Reichert.
- (2000) *The OPers. Inscriptions of Naqsh-e Rostam and Persepolis*, *Corpus Inscriptionum Iranicarum* I, I, Texts II. London: Corpus Inscriptionum Iranicarum.
- (2007) *Pseudo-Altpersische Inschriften. Inschriftenfälschungen und moderne Nachbildungen in altpersischer Keilschrift*, Sb. d. ÖAW, phil.-hist. Kl., 762, Veröff. zur Iranistik 39.
- Skjærvø, P. O. (1985) 'Remarks on the Old Persian Verbal System', *MSS* 45, pp. 211–27 (= *Festgabe für K. Hoffmann*).
- (1994) 'Hymnic Composition in the Avesta', *Die Sprache* 36, pp. 199–243 [publ. 1996].
- (1995) 'The Avesta as Source for the Early History of the Iranians', in: G. Erdosy (ed.) *The Indo-Aryans of Ancient South Asia*. Berlin and New York: W. de Gruyter, pp. 155–76.
- (1996) 'Iranian Alphabets Derived from Aramaic', in: P. T. Daniels and W. Bright (eds) *The World's Writing Systems*. New York and Oxford: Oxford University Press, pp. 515–35.
- (1997a) 'A Perfect Participle, *vaox"ānhō*', *JAOS* 117, pp. 145–47.

- (1997b) ‘Avestica II Yokes and Spades and Remnants of the “Tripartite Ideology”’, *MSS* 57, pp. 115–28 [pub. 1998].
- (1998) Review of Kellens (1995) in *Kratylos* 43, pp. 190–91.
- (1999a) ‘Avestan Quotations in Old Persian?’ in S. Shaked and A. Netzer (eds) *Irano-Judaica* IV, Jerusalem, pp. 1–64.
- (1999b) ‘Methodological Questions in Old Persian and Parthian Epigraphy’, *BAI* 12 [pub. 2002], pp. 157–67 (review article of Schmitt, *The Old Persian Inscriptions of Naqsh-e Rostam and Persepolis*, 2000).
- (2003–2004) ‘The Antiquity of Old Avestan,’ *Nāme-ye Irān-e Bāstān. The International Journal of Ancient Iranian Studies* 3/2, pp. 15–41.
- (2005a) ‘Avestica III. Four Notes on Avestan Morphology,’ in: D. Weber (ed.) *Languages of Iran: Past and Present. Iranian Studies in Memoriam David Neil MacKenzie*, Iranica 8, Wiesbaden: O. Harrassowitz, pp. 197–206.
- (2005b) ‘Poetic and Cosmic Weaving in Ancient Iran. Reflections on Avestan *vahna* and Yasna 34.2,’ in: D. Haug and E. Welo (eds) *Haptaçahaptāitiš. Festschrift for Fridrik Thordarson*, Oslo: Novus, pp. 267–79.
- (2005–2006) ‘The Importance of Orality for the Study of Old Iranian Literature and Myth,’ *Nāme-ye Irān-e Bāstān. The International Journal of Ancient Iranian Studies* 5/1&2 (pub. 2007), pp. 1–23.
- (2006a) ‘Iran iv. Iranian Languages and Scripts’, in: *Encyclopædia Iranica* XIII/3, pp. 344–77.
- (2006b) ‘The Videvdad: Its Ritual-Mythical Significance’, in: V. S. Curtis and S. Stewart (eds) *The Age of the Parthians, The Idea of Iran* 2, London and New York, pp. 105–41.
- (2007a) ‘Avestica V. The Thematic Optative 3rd Plural in *-aiiaēn* and the Instrumental Plural of *n*-stems and Some Other Consonant Stems,’ in: Alan J. Nussbaum (ed.) *Verba Docenti. Studies in Historical and Indo-European Linguistics Presented to Jay H. Jasanoff by Students, Colleagues, and Friends*, Ann Arbor, pp. 321–28.
- (2007b) ‘Avestan and Old Persian Morphology,’ in: Alan S. Kaye (ed.) *Morphologies of Asia and Africa*, Winona Lake, Ind.: Eisenbrauns, pp. 853–940.
- Testen, D. (1997) ‘Old Persian and Avestan Phonology’, in: A. S. Kaye (ed.), *Phonologies of Asia and Africa*. Winona Lake: Eisenbrauns, pp. 569–600.
- Tichy, E. (1986) ‘Vedisch *rtāvan-* und avestisch *ašauuan-*’, *Die Sprache* 32, pp. 91–105.
- Tremblay, X. (1998) ‘Sur *parsui* du Farhang-i-ōim, *ratu*, *pərətu-*, *pitu-* et quelques autres thèmes avestiques en *-u*. Essais de grammaire comparée des langues iraniennes III’, *StIr* 27, pp. 187–204.
- (1999) ‘Ist jungawestisch *nāismi*, *nāist* Präsens oder sigmatischer Aorist?’, in: H. Eichner et al. (eds) *Compositiones Indogermanicae in Memoriam Jochem Schindler*. Prague: Enigma Corporation, pp. 537–43.
- (2006) ‘Le pseudo-gâthique. Notes de lecture avestiques II’, in: A. Panaino and A. Piras (eds) *Proceedings of the 5th Conference of the Societas Iranologica Europæa held in Ravenna, 6–11 October 2003*, vol. I. *Ancient & Middle Iranian Studies*, Rome: Istituto italiano per il Medio ed Estremo Oriente, pp. 233–81.
- Tucker, E., ‘Denominative Verbs in Avestan: Derivatives from Thematic Stems’, in J. H. W. Penney (ed.) *Indo-European Perspectives: Studies in Honour of Anna Morpurgo Davies*, Oxford/New York: Oxford University Press, 2004, pp. 548–61.
- de Vaan, M. (2001) ‘Avestan *vaēsmānda*’, *MSS* 61: 185–92.
- (2003) *The Avestan vowels*. Amsterdam and New York: Rodopi.
- (2004) ‘“Narten Roots” from the Avestan Point of View’, in: A. Hyllested et al. (eds) *Per Aspera ad Asteriscos: Studia Indogermanica in Honorem Jens Elmegård Rasmussen Sexagenarii Idibus Martiis Anno MMIV*, Innsbruck (Innsbrucker Beiträge zur Sprachwissenschaft 112), pp. 591–99.

- Wackernagel, J., and A. Debrunner. (1896–) *Altindische Grammatik*. Göttingen: Vandenhoeck und Ruprecht.
- Waters, M. W. (1999) ‘The Earliest Persians in Southwestern Iran: the Textual Evidence’, *Iranian Studies* 32: 99–107.
- Watkins, C. (2000) ‘*sá figē* in Indo-Iranian and Anatolian’, in: A. Hintze and E. Tichy (eds) *Anusantatyai. Festschrift für Johanna Narten zum 70. Geburtstag*. Dettelbach: J. H. Röhl, pp. 263–81.
- Windfuhr, G. (1971) ‘Diacritic and Distinctive Features in Avestan’, *JAOS* 91: 104–24.
- (1972) ‘Some Avestan Rules and Their Signs’, *JAOS* 92: 52–59.

Further readings

- Asmussen, J. P. (1992) ‘Codices Hafnienses’, in: *EnIr* V, pp. 886–93.
- Bailey, H. W. (1943) *Zoroastrian Problems in the Ninth-Century Books*. Oxford: Clarendon Press (2nd ed. with new introd., 1971).
- Henning, W. B. (1942) ‘The Disintegration of the Avestic Studies’, *TPhS* 1942: 40–56 [pub. 1944].
- Gippert, J. (1986) ‘Zur Metrik der Gathas’, *Die Sprache* 32: 257–75.
- Kellens, J. (1988) ‘Avesta’, in: *EnIr* III, pp. 35–44.
- Klein, J. (1985) *Toward a Discourse Grammar of the Rigveda*, 2 vols. Heidelberg: C. Winter.
- Narten, J. (1995) *Kleine Schriften*, M. Albino and M. Fritz (eds) Wiesbaden: L. Reichert.
- Pirart, E. (1992) ‘Les fragments vieil-avestiques du Y 58’, *AION* 52: 225–47.
- Schmidt, H. -P. (1974) ‘Associative Techniques and Symmetrical Structure in the Composition of Yasna 47’, in: R. N. Frye (ed.) *Neue Methodologie in der Iranistik* (dedicated to Wolfgang Lentz); Wiesbaden: O. Harrassowitz, pp. 306–30.
- Schwartz, M. (1986) ‘Coded Sound Patterns, Acrostics, and Anagrams in Zoroaster’s Oral Poetry’, in: R. Schmitt and P. O. Skjærvø (eds) *Studia Grammatica Iranica. Festschrift für Helmut Humbach*. Munich: R. Kitzinger, pp. 327–92.

CHAPTER FOUR

MIDDLE WEST IRANIAN

Prods Oktor Skjærvø

1 INTRODUCTION

1.1 Overview

Middle Persian (MPers.) and Parthian (Parth.) constitute the western branch of the known Middle Iranian languages.* The term ‘Pahlavi’ (Pahl.) refers to the Middle Persian of the Zoroastrian texts, but is also sometimes used for Middle Persian in general. The indigenous words for the languages may have been Pārsīg and Pahlawānag ‘Parthian’ (found in a Manichean text). This usage was sometimes adopted by Western scholars, as well (Herzfeld, Nyberg: Pārsīk, Pahlavīk). Still earlier, Pahlavi and Chaldeo-Pahlavi were used.

The two languages are closely similar in structure, though Parthian shares some features with its eastern neighbors, notably with Bactrian, rather than with Middle Persian (see Sims-Williams 2007). The fact that Manichean Middle Persian and Parthian were two separate languages with different linguistic affinities was first shown by Tedesco (1921).

Parthian was spoken in Parthia, east of the Caspian Sea, and became an official language under the Parthian (Arsacid) rulers of Iran (ca. 247 BCE–224 CE). It is known mainly from a large corpus of short, formulaic, Parthian inscriptions mostly on potsherds from the Parthian capital of Nisa dating from the first century BCE (see Skjærvø 1995a, 1999); a few royal Parthian inscriptions from the last couple of centuries of Parthian rule; Parthian versions of the inscriptions of the third-century Sasanian kings Ardashir I, Shapur I, and Narseh; and from the Manichean (Man.) texts found at Turfan in north-eastern Chinese Turkestan (Xinjiang) in the early nineteenth century. On the language of the mostly Aramaic inscriptions from north-western Iran and the Awrōmān document see Skjærvø (1995a).

Middle Persian is descended from Old Persian (OPers.) and is the ancestor of New Persian (NPers.). Its proper homeland was the area of Pārs in south-western Iran (Gk. Persis), and it was the official language of the local Perside kings (ca. 200 BCE–224 CE) and their successors, the Sasanians (224–651 CE). After the fall of the Sasanians, it continued to be used by the Zoroastrians long after the spoken language had become New Persian in the centuries following the Arab conquest (ca. 650). Middle Persian is known mainly from inscriptions, documents on parchment and papyrus, a translation of the *Psalms* of David (the ‘Pahlavi *Psalter*’ [Ps.]) found at Turfan, the Zoroastrian scriptures, and the Manichean texts from Turfan. A page from a Pahlavi text containing verb forms was also found there (Barr 1936).

*For reasons of space, Parthian could not be discussed as fully as Middle Persian. All descriptions and examples therefore apply to Middle Persian only, unless explicitly marked as Parthian.

The earliest inscriptions are those on the coins of the Persian rulers and on a silver bowl from the second half of the first century BCE (Skjærvø 1997b). There are royal and private inscriptions by high officials and travelers, most of them from the third century, a few from the fourth–fifth centuries; the earliest ones were bilingual (trilingual) Middle Persian, Parthian, (Greek). There are inscriptions on objects (silver bowls, seals, etc.) from the entire Sasanian period (see Brunner 1974; Skjærvø 1993). The Zoroastrian and Christian funerary inscriptions are from the late Sasanian and early Muslim period (see Gignoux 1978: 14); some are from as far away as India (that on the Thomas Cross from the seventh century, see Gignoux 1995) and Xi'an, China (from 874, see Humbach and Wāng Shìpíng 1988; Sundermann and Thilo, 1966). The papyri are probably from the sixth–seventh centuries (see Weber 1992, with bibliography pp. ix–x). A large corpus of mostly legal documents on parchment from the seventh century has recently surfaced and is currently being deciphered and studied (see, e.g. Gignoux 1991, 2001, 2002–3 with refs.; Weber 2007 with refs.).

The earliest Manichean texts may have been composed in the mid-third century, but the manuscripts are much later, and texts were still composed in Turfan in the eighth century (e.g. the *Mahnāmag* is dated in 762, see Boyce 1975: 52); the *Psalter* manuscript probably dates from the seventh century, although the text is probably older (Skjærvø 1983: 178–79).

The Zoroastrian ‘Pahlavi books’ were written down in the ninth–tenth centuries, but contain texts that were transmitted orally for centuries. The corpus can be divided roughly into three categories (see Tavadia 1956; Boyce 1968): 1. translations of Avestan texts, 2. texts with religious contents, sometimes incorporating translations from the extant or lost Avesta; and 3. secular texts, including word lists.

1.2 Writing systems

The non-Manichean Middle Persian texts are written in scripts derived from Achaemenid Aramaic (‘Imperial Aramaic’; Henning 1958; Skjærvø 1995a, 1996); the Manichean texts are written in a Syriac script more closely related to the Palmyrene script than to Estrangelo, as traditionally thought (Durkin-Meisterernst 2000a). In both, the consonants <ʿ, w, y> are also used to express vowels (*matres lectionis*, see section 2.3.1). See Tables 4.1–3.

The Persian variant of the Aramaic script evolved in Pārs in the late Parthian period, as can be seen from coins. There is a clear change at the end of the second century CE, and the standard form appears under Ardashir I’s older brother Shapur (Skjærvø 1991b, 1997b). An advanced form of the script is seen in the *Psalter* manuscript, while, in the earliest funerary inscriptions, we have the cursive forerunner of the Pahlavi script used in the later funerary inscriptions and in the Pahlavi books (‘Book-Pahlavi’). The final stage of the script is found in the documents on papyri and parchment (Weber 2007 with refs.).

In Tables 4.1–2, the order of the letter forms is right-to-left. Letter forms preceded by a line are final (or post-vocalic in the Pahlavi script). Manichean letters with a left-side flourish or lengthening lose these when connected with a following letter.

In the Pahlavi column, letters in parenthesis are alternative forms: those of <k> and <l> are used in a few arameograms.

TABLE 4.1: ALPHABETS. PARTHIAN AND MIDDLE PERSIAN SCRIPTS DERIVED FROM IMPERIAL ARAMAIC

<i>Semitic alphabet</i>	<i>Hebrew</i>	<i>Translit. values</i>	<i>Parthian inscrips.</i>	<i>MPersian inscrips.</i>	<i>Psalter</i>	<i>Pahlavi</i>
' (alef)	א	a A	𐭠	𐭡	𐭢	𐭣
b	ב	b B	𐭤	𐭥	𐭦	𐭧
g	ג	g G	𐭨	𐭩	𐭪	𐭫
d	ד	d D	𐭬	𐭭	𐭮	𐭯
h	ה	- E	𐭰	𐭱	𐭲	𐭳
w	ו	w W	𐭴	2 [=O, r]	𐭵	𐭶
z	ז	z Z	𐭸	𐭹	𐭺	𐭻
h	ח	h H	𐭼	𐭽	𐭾	𐭿
ʔ	ט	- θ	𐭿	𐭿	-	-
y	י	y Y	𐭽	𐭿	(𐭿) 𐭿	𐭿
k	כ	k K	𐭽	𐭿	𐭿	(𐭿) 𐭿
l	ל	l L	𐭽	𐭿	𐭿	(𐭿) 𐭿
m	מ	m M	𐭽	𐭿	𐭿	𐭿
n	נ	n N	𐭽	𐭿	𐭿	𐭿
s	ס	s S	𐭽	𐭿	𐭿	𐭿
' (ayin)	ע	O	> [=r]	2 [=r, w]	𐭿	𐭿
p	פ	p P	𐭽	𐭿	𐭿	𐭿
ʃ	צ	c C	𐭽	𐭿	𐭿	𐭿
q	ק	Q	𐭽	𐭿	𐭿	𐭿
r	ר	r R	> [=O]	2 [=O, w]	𐭿	𐭿
š	ש	š Š	𐭽	𐭿	𐭿	(𐭿) 𐭿
t	ת	t T	𐭽	𐭿	𐭿	𐭿

1.2.1 Merger of letter shapes

In the Parthian and Middle Persian scripts derived from Imperial Aramaic, some of the original 22 letters began merging early on, and the Book-Pahlavi script contains only 12 distinct letters, but even these are often ambiguous. In the Parthian inscriptions, <' (ayin) = r>, and , <k>, and <r = ' > are sometimes indistinguishable (e.g. in the Paikuli inscription, <OBDt> *kerd* is often written so as to look like <krd-t>). In the Middle Persian inscriptions and the *Psalter*, <w = r = O>, and, in Book-Pahlavi, <n = w = r = O>; in the inscriptions and the *Psalter*, <' (alef)>, <l>, <g>, <d>, and <y> were still distinct, but, in Book Pahlavi, <' = h>, <g = d = y>, and in the *Psalter* and Book Pahlavi, <s = yy>.

TABLE 4.2: ALPHABETS. THE PAHLAVI SCRIPT

Pahlavi	Values	Pahlavi	Values	Pahlavi	Values
𐭠	<'> = <h>, <y>	𐭡	<y> = <g, d>	𐭢	<O> = <w, n, r>
𐭣		𐭤	<d>	𐭥	<p>
𐭦	<g> = <d, y>	𐭧	<K, γ>	𐭨	<c>
𐭩	<d> = <g, y>	𐭪	<l>	𐭫	<-c> = <-p>
𐭬	<-d>	𐭭	l	𐭮	<Q> = <m>
𐭯	d, g, y, j	𐭰	<l>	𐭱	<t> = <w, n, O>
𐭲	<-E>, <mn>	𐭳	<m> = <Q>	𐭴	<š>
𐭵	<w> = <n, O, r>	𐭶	<r> = <w, O, r>	𐭷	<š> = <y'>
𐭸	<z> = <d>	𐭹	<s> = <yy>	𐭺	<t> (= <yn>)
𐭻	<h> = <'>				

TABLE 4.3: ALPHABETS. THE MANICHEAN SCRIPT

Pahlavi	Values	Pahlavi	Values	Pahlavi	Values
𐭠	' (alef)	𐭡	h	𐭢	' ('ayin)
𐭣	b	𐭤	t	𐭥	p
𐭦	β	𐭧	y	𐭨	f
𐭩	g	𐭪	k	𐭫	č
𐭬	γ	𐭭	x	𐭮	q
𐭯	d	𐭰	l	𐭱	r
𐭲	-h	𐭳	m	𐭴	š
𐭵	w	𐭶	n	𐭷	t
𐭸	z	𐭹	s	𐭺	j (ž)

In Pahlavi, which is a cursive script, the letters also have variant forms when combined and, in certain words, they take on special forms or the shapes of other letters, e.g. <z> can represent <y = d>; <y = d> can represent ; and, so, <z> can represent : <YZLWN> for <YBLWN> 'carry', etc. This is indicated in transliteration by an underbar: <YBLWN>. In Pahlavi manuscripts written in India, <š> is indistinguishable from <y>. The letter <t> is frequently used for <yy, zd>: <y 'n'> = <yzd h0> *yazdān* 'the gods'; note also <m 'yyt'> (or m 'yyyt') for <mzdysn' (= mhyst') > 'Mazdayasnian' with <t> for <zd> and <yt> for <sn>.

Note that the Aramaic letters <E> (only final), <O ('ayn)>, <Q>, and <Θ> (Aram. <t>) were used only in arameograms (see below). In Pahlavi, they merged with other letters: <E = mw>; <O = w>; <Q = m> (only in <QDM> = *abar* 'on'); <Θ = t> (e.g. Aram. <tb>, MPers. inscr. <ΘB>, Pahl. <TB> = *td* = *nēw* 'good').

In the Manichean script, all the letters are distinct and the spelling is near-to phonetic (with a few exceptions: Man. MPers. <p> = *p* and *f*; <Vc> = MPers. *Vz*, <Vc, Vj> = Parth. *Vž*).

2 PHONOLOGY

2.1 Inventory and distribution

By cross-referencing Middle Persian with Old Persian and New Persian, the vowel and consonant phonemes that can be posited for Middle Persian (see MacKenzie 1967) are set out in Tables 4.4–5. The Parthian vowel system was probably of the same type, but the consonant system was more complex.

TABLE 4.4: PHONOLOGY. VOWELS

	Front		Back	
High	<i>ī</i>	<i>i</i>	<i>u</i>	<i>ū</i>
Mid	<i>ē</i>	(<i>e</i>)	(<i>o</i>)	<i>ō</i>
Low		<i>a ā</i>		

TABLE 4.5: PHONOLOGY. CONSONANTS

	Labial	Dental	Palatal	Velar	Glottal
<i>Stops, Affricates</i>	<i>p b</i>	<i>t d</i>	<i>č ĵ</i>	<i>k g</i>	—
<i>Fricatives</i>	<i>f v</i>	<i>s z</i>	<i>š (ž)</i>	<i>x (ɣ)</i>	—
[<i>Parthian</i>	<i>β</i>	<i>ð</i>	<i>ž</i>	<i>ɣ</i>]	—
<i>Rounded</i>	—	<i>xw</i>	—	—	—
[<i>Parthian</i>	—	<i>x^o</i>]	—	—	—
<i>Nasals</i>	<i>m</i>	<i>n</i>	—	(<i>ŋ</i>)	—
<i>Continuants</i>	<i>w</i>	<i>r, l</i>	<i>y</i>	—	<i>h</i>

2.1.1 Vowels

Since the script does not distinguish between *ē* and *ī*, *ō* and *ū*, these values can only be determined by etymological considerations and their distribution in early New Persian and modern dialects of Persian (in standard New Persian they have merged into *i* and *u*), and a few instances remain uncertain.

The existence of the phonemes (*e*) and (*o*) is suggested by alternate spellings with or without *mater lectionis* and is also posited for etymological reasons in words such as *kerd* <krt'> 'did' (OPers. *karta* with syllabic *r*, Man. MPers. <kyrd>, NPers. *kard*, dial. *kerd*) and *ox* <hw'> 'mind' (Av. *oŋ'hā*, Man. MPers. <wx>; see MacKenzie 1967: 23–25). The question remains, however, whether we have any *e* and *o* contrasting with *a* ~ *i* and *a* ~ *u*, respectively, as there do not seem to be any minimal pairs/triples of the type *kird* ~ *kerd* ~ *kard* or *widurd* ~ *widord* ~ *widerd* <wtwt', wlt'> 'crossed over'.

2.1.2 Consonants

The consonant phonemes Middle Persian (γ) and (\check{z}) are found mainly in learned (Avestan) words, but perhaps also in words of non-Persian origin, and so probably represent marginal phonemes $/\gamma/$ and $/\check{z}/$ (see MacKenzie 1967: 20–23)

(η) is an allophone of n before velars (not expressed in the script). The opposition $n \sim m$ is neutralised before labials, as evidenced by the alternate spellings <np, mp, nb, mb>.

The voiceless stops and affricates are found in initial position and after voiceless sibilant, but only sporadically after vowels (other than after prefixes such as $a-$ ‘non’). In intervocalic position, they are the result of gemination (geminate, apparently, do not contrast with non-geminates, and so were not phonemic): *appār-* ‘rob’ (< **apa-pāra-*), *pat-tāy-* > *pattāy-* ‘endure (something)’, *waččag* ‘child’; *kerbag-kar* > *kerbakkar* ‘well-doer’. In final position, they are found in the suffixes *-ak* and the rare *-ik* (contrasting with *-ag* and *-ig*); as a result of early syncope in *ēk* (*yak*) ‘one’ (< **ēw-k*), *ēč* ‘any’ (< **ēw-č*), and *pāk* ‘pure’ (< **pāw-k*, cf. Parth. *pawāg*); in a few words such as *kayk* ‘flea’, *sabuk* ‘light, easy’; and after h in Man. <bšyhk>. For the comparative and superlative suffixes *-tar* and *-tom*, Man. MPers. has *-dar* and *-dom* after voiced sounds, but New Persian has forms with *-t-* throughout, and it is likely that this was already so in Pahlavi. Examples of other geminates: *passand-* ‘please’ (< **pati-sand-*), *škenn-* ‘break’ (< **škandaya-*), and *andar-rōn* ‘in-side’.

(\check{j}) is essentially limited to the initial position and the position after n : *panj* ‘five’, Man. MPers. <pnz>. Before voiced consonants, it is indistinguishable from (\check{z}): *dujden* ~ *duzden* ‘unbeliever’.

The phoneme (l) is common in non-initial position, but also occurs sporadically in initial position (judging by the New Persian equivalents of these words and the occasional explicit Pahlavi spelling with <l> with diacritic). In Manichean Middle Persian, it is rarer, being often replaced by r : Pahl., NPers. *sāl* ‘year’, *larz-* ‘tremble’, Man. MPers. *sār*, *varz-*, but Pahl., NPers., and Man. MPers. *dil* ‘heart’.

Parthian probably also had the phonemes δ and γ (mainly from older intervocalic d , g) perhaps also β (but old intervocalic b had become w) contrasting with d , g , b , cf. minimal pairs such as *pād* ‘foot’ ~ *pād* ‘protected’ and contrasting structures such as *bay-ān* ‘gods’, but *frēšt-ag-ān* ‘prophets’, as well as words such as *aβr* ‘cloud’.

Note also that Manichean Parthian usually writes <wx> where Middle Persian has <xw>, which may indicate an indivisible phoneme x° : <wxybyh> “*xēbēh*, MPers. *xvēbaš*’.

In inscripational Parthian, initial \check{c} - is written <š> before i (<šyhr> = Man. Parth. *čihr*), which perhaps reflects a local sound change. Another difference between the two dialects is the form of the abstract ending inscr. Parth. *-if* = Man. Parth. *-ift* (both probably from **-itf*).

2.1.3 Syllable structure and consonant groups

Middle Persian initial, medial, and final groups usually contain fricatives, sibilants, and r . The initial groups ‘sibilant + stop’ early on acquired prosthetic vowels, globally in Manichean Middle Persian and Parthian, but sporadically as early as the third century, judging from inscr. Parth. (*a*)*spād-bed* ‘general’, spelled <šp’dpty> and <spdpty> (cf. Man. Parth. *ispād* <šp’d> ‘army’).

Parthian also had the non-initial groups *df* and *šf* (not found in Middle Persian):

nidfār- ‘hurry’ (Pahl. <nswb|>-> *nixwār*-); *sadf-ān* ‘beings’ (<Skt. *sattva*), *kišfar* ‘continent’ (MPers. *kišwar*), *padišfar* ‘honour’ (MPers. *padixšar*).

A small problem is posed by the phonological analysis of initial <C-y-V>, which is best analysed as *Ci-* plus a non-phonemic glide: *ni(y)āy-* ‘to sing’, *zi(y)ān* ‘harm’ (there is no phonemic contrast between [CyV-], [CiyV-], and [CiyV-]). Similarly, the final -y in words such as <hy> ‘other’, <mhy> (beside <myh>) should perhaps be analysed as -ī (rather than -y); cf. the Pahlavi spellings <AHRNyḏ> and <ZKYḏ>, with <-yḏ> = -īy; here, this final will be spelled -īy, with -y reflecting the spelling with (-d), *anīy* for MacKenzie’s *any*.

2.2 Non-segmental aspects

The theory that Middle Persian may have had a system of ‘light’ and ‘heavy’ stems (like Sogdian) proposed by Back (1978: 39–61) was disputed by MacKenzie (1982: 294–95). A similar theory was also proposed by Huysse (2003), which still needs to be assessed.

Words such as *anīy* and *mahīy* may have been stressed on the first syllable, since the final syllable was later lost (*ānīy*, later *an*; *máhīy* later *meh*).

2.3 Relationship between orthography and phonology in Pahlavi

Two features complicate the phonological interpretation of Pahlavi orthography: 1. conservatism, resulting in an orthography based on that of the first two centuries of our era (cf. English, French); 2. the use of arameograms.

2.3.1 Vowels

All the Parthian and Middle Persian alphabets contain consonants only, with three of the consonants used as *matres lectionis*: <’,”> = *ā*, <w> = *ō*, *u*, *ū*; <y> = *ē*, *i*, *ī*. In the Manichean script, <y-> is used in initial, and sometimes in internal, position, to distinguish *i-*, *-ī* from *ē-* spelled <y->.

2.3.2 Consonants

In non-Manichean Middle Persian, few letters regularly spell the corresponding consonants: <m, n, z> = *m*, *n*, *z*. In initial position, all the letters usually express the corresponding consonants: <p k’> *pāk* ‘clean’, <c l’> *čār* ‘able’, <g l’> *gar* ‘mountain’, etc. In the Middle Persian inscriptions, <l’> and <r’> are used according to the whim of the scribe to spell *r*. In Pahlavi, <r’> is relatively rare and limited to certain types of words (especially arameograms); the <l’> with a diacritic stroke can be used when it spells *l* (rather than *r*).

After vowels and voiced consonants, the unvoiced stops express the corresponding voiced stops: <ḡ’> *āb* ‘water’, <w ḡ’> *wād* ‘wind’, *kerd* <krt’> ‘did’, etc. and <c> is used to spell *z* (but <nc> *nj*): <tyc’> *tēz* ‘sharp’, <lnc’> *ranj* ‘toil’ (Man. <rnz>). In intervocalic and final position after vowels, the voiced stops <b (wb)> and <d (yd), g> spell *w* and *y*, respectively: <lwb k’> *ravāg* ‘current’, <dlwb’> *drō(w)* ‘deceit’, *mayān* <myḏh> ‘middle’, <mynwd> *mēnōy* ‘the world of thought (the other world)’.

2.3.3 Historical and pseudo-historical orthography

Pahlavi orthography reflects a phonology closer to Old Persian than to Manichean Middle Persian, and consonants that had been lost or changed were still written as they were originally. Thus, what the Manicheans wrote as <by, r'ÿ, xrd, šhr, s'ř'ŕ>, i.e. *bay, rāy, xrad, šahr, sālār*, one would write <bg, l'd, hlt', štr', srd 'l> (OPers. *baga-*, *rādiy*, *xratu-*, Av. *xšaθra-*).

Analogical (also called 'pseudo-historical') spellings are also common. For instance, since old *b* and *d* had become *w* and *y* between vowels, and <d> were used for any *w* and *y*, e.g.: *ruvān* <lwb ñ> 'soul' (Av. *uruuānəm*), *ziyān* <zyd ñ> 'harm' (Av. *ziiāni-*) = Man. <rw ñ>, <zyy ñ>; since old *s* and *θ* had both become *h* between vowels, *h* could be written as <h> or <s>: *gāh* <g 'h, g's> 'throne' (Av., OPers. *gāθu-*), Man. <g'h>; *puhl*, inscr. <pwhty, pwlsy> 'bridge' (Av. *pərəθu-*) = Man. <pwht>.

In many Pahlavi manuscripts, a letter identical with <w>, but with no phonetic meaning, is written at the end of the word (the 'final stroke'). The form developed from a final <-y>, also with no phonetic (but perhaps historical) significance, used in inscriptional Middle Persian and the *Psalter*, but their distribution is different; in Pahlavi, it is only used after consonants with which it is not combined (<b, c, k, t, w/n/r>). It is commonly transliterated as <->.

In addition, *ē* is often spelled <yw>: <nywk'> *nēk* 'good', etc., and final *-ī* (*-īy*) <-yđ>: <gytyđ> *getī(y)* 'the world of the living'. In learned words, *ō* and *ū* are often spelled <wk>: <yt'hwkwlywk> *Yatā-ahū-wēryō*, an Avestan prayer.

Letter forms differ from scribe to scribe and develop throughout the centuries.

2.3.4 Arameograms

Parthian and Pahlavi (including the inscriptions and the *Psalter*) used a large number of, more or less faithfully rendered, Aramaic words, also called 'ideograms' or 'heterograms', but here: 'aramaeograms' (Skjærvø 1995a; Nyberg 1988). Iranian endings were added to the arameogram as 'phonetic complements'. To distinguish the arameograms from Middle Persian words spelled 'phonetically' arameograms are now transliterated in upper case, e.g.: <OLE> = *ōy* 'he, she, it', <GBRA> = *mard* 'man', with phonetic complements: <OLE-š'ñ'> = *avē-šān* 'they', <GBRA-ñ'> = *mard-ān* 'the men('s)', <OBYDWN-d> = *kun-ēnd* and Parth. <OBDW-nt> = *kar-ēnd* 'they do'. Parthian still has a few Aramaic plurals, where Middle Persian uses phonetic complements, e.g. MPers. <GBRA-ñ'> = *mard-ān*, Parth. <GBRYN> 'the men('s)'.
For the phonetic complements of verbal arameograms, see Skjærvø 1989, 1995a.

2.3.5 Transliteration and transcription systems

Throughout the history of Pahlavi studies, many transliteration and transcription systems have been used. This is due to two facts: the correct interpretation of the Pahlavi orthography did not become known until the inscriptions and the Pahlavi *Psalter* were fully deciphered, and the actual pronunciation of Pahlavi did not become known until the discovery of the Manichean texts.

Different opinions among scholars then resulted in basically two systems of transcription: one that remained faithful to the actual spelling as seen in the inscriptions and the *Psalter* and one that was based on the assumption that the spelling was largely historical and did not reflect the actual pronunciation, as evidenced by the Manichean texts. These

two systems were implemented in Nyberg's *Manual* (1964–74) and MacKenzie's *Concise Dictionary* (1971), respectively, which set the standards for them (see also MacKenzie 1967). Here, MacKenzie's system will be adhered to throughout with only minor departures. While the differences in transliteration systems today remain small (see Table 4.2), those between the two major transcription systems produce two rather different-looking languages, e.g.: *āp* ~ *āb* 'water', *wāt* ~ *wād* 'wind', *sakān* ~ *sagān* 'of Sakas', *mātagdān* ~ *mādayān* 'specifically', *hač* ~ *az* 'from', *mi0r* ~ *mīhr* 'Mithra'.

The main innovation in MacKenzie's transliteration system was the use of only capital letters in the arameograms, that is, also <A, E, O> for <'(alef), H (H), '(ayin)>, to which Humbach added <Θ> for <Ṭ>, e.g. <AB> vs. <'B> 'father', <BYTA> vs. <BYT> 'house', <ZNE> vs. <ZNH> (<ZNHH>) 'this', <OLE> vs. <'LH> (<'LHH>) 'he', <ΘB> vs. <ṬB> 'good'.

Another common system, especially in French publications: <' , H, ' > = <A, E, O>, <ḥ> = <h>, and <Ṣ> = <C>, e.g. <l 'ḥ> vs. <l 'h> *rāh* 'road', <'RḤY'> vs. <ORHYA> *bay* 'god, majesty', <ṢBW> vs. <CBW> *xīr* 'possessions', <YṢBH-> vs. <YCBE-> *kām* 'wish'. In some personal transliteration systems, <x> is used for <h> and <H> for <E>, e.g. <XWH-> for <HWE-> *h*- 'be'.

3 MORPHOLOGY

With the loss of Old Iranian final syllables in Middle West Iranian, the original gender and case distinctions were to a large extent lost. Nevertheless, third-century inscriptional Middle Persian and Parthian still had a two-case system, which had largely been given up in the Manichean texts, but is still present in early Pahlavi, especially in the translations of the Avesta. In contrast, the verbal system is quite archaic.

There is a rich derivational morphology, including sequences of more than one derivational marker, which often leads to nominalised constructions that straddle the boundaries between word class, noun phrase syntax, and sentence syntax, a process which is complemented by the derivational verbalisation of simple and derived nominals. As word order is relatively free, it is often difficult to distinguish adverbs, adpositions, and preverbs.

3.1 Nominal morphology

Morphological categories in nominals include number, case, person, and deixis and marginally feminine gender, definiteness and animacy.

3.1.1 Nouns

3.1.1.1 Gender

Animate and inanimate is distinguished in the interrogative pronouns *kē* 'who?' versus *čē* 'what?' and the indefinite pronouns *kas* 'somebody' and *tis* 'something' (Parth. *kē*, *čē*, *keč*, *čiš*), as well as compound forms of these (see section 3.1.4.5).

Feminine gender can be marked by *-ag* in Middle Persian proper names (Jam and Jamag) and by *-ānīy* (*-ēnīy*) in imitation of Avestan (*ahlaw* and *ahlawēnīy* 'righteous'). In Manichean Middle Persian and Parthian, *-čān* denotes female members of a group: *niyōšāg-čān*, Parth. *niyōšāg-čān* 'hearer female'.

3.1.1.2 Number and case

In the early texts (inscriptions and the *Psalter*), there are two cases and two numbers, and there are sufficient examples in Pahlavi of the old system to show that it probably survived until quite late. The cases are regularly distinguished in the plural, but in the singular only in the *r*-stem family terms ('father', etc.) and in the 1s pronoun. The enclitic personal pronouns are only used as oblique. See Sims-Williams 1981; Skjærvø 1983; Cantera Glera 1999, 2004: 270–73. Table 4.6 (the forms are also Parthian).

TABLE 4.6: NOMINAL NUMBER AND CASE MARKING IN THE THIRD-CENTURY INSCRIPTIONS

	'man'	'father'
DIRs	<i>mard</i>	<i>pid</i>
OBLs	<i>mard</i>	<i>pid-ar</i>
DIRp	<i>mard</i>	<i>pid-ar</i>
OBLp	<i>mard-ān</i>	<i>pid-ar-ān</i>

In the Pahlavi texts, the forms in *-ān* are used freely as marked general plurals, and the forms with *-ar* alternate with those without.

Archaisms include the plural (obl.) endings *-īn* and *ūn* (from original *i*- and *u*-stems, but there is much analogy): inscr. *mow-ūn* <mgwny> 'Magians' and Man. MPers. *dušmen-ūn* 'enemies' (old *u*-stems), Ps. *pidar-ūn* <ptlwny> 'fathers' (old *r*-stem), *frazend-ūn* <plcndwny> 'children' (old *i*-stem), Man. Parth. *dušmen-īn*, *frazend-īn*. On the reflexes of old final vowels before enclitics, see Sims-Williams 1981.

This case system eventually collapsed into a no-case system, with cases surviving marginally in the pronouns; in Pahlavi, the originally plural oblique case marker, *-ān*, became the unmarked plural marker, next to which a new ending *-īhā* (Man. *-īhān*) expressing individual plurality is found: *kōf-ān* 'the mountains' vs. *kōf-īhā* 'the various/individual mountains', *pad sūrāg-īhā ī kōf-ān* 'in the various holes in the mountains' (Bd. 21E.7), *pad anjoman-īhā* 'throughout the congregations' (Ps. 131 canon). The plural was used mainly when plurality was emphasised, not when a class of objects was referred to, e.g. 'stones are heavy'.

3.1.1.3 Definiteness

There are no definite and indefinite articles, but the numeral *ēw* <-1, -HD> 'one' can be used to express 'one specific X, a certain X' (see also section 3.1.4.5):

aswār-[ēw] šahryār wēn-ēm

'horseman-[one] prince we see' = 'we see/saw a princely rider';

ud nūn zan-[ēw] paydāg

'and now woman-[one] manifest (COP)' = 'and now a woman appeared';

u-š tarāzūg-[ēw] pēšīy est-ēd

'and-him scales-[one] before stand-PRES/IPF-3s' = 'and a scales stood before him' (KNRm 34–37).

3.1.1.4 Address

In Manichean texts, a particle *-ā* may be used to express address:

[*rōšn-ā*] *ǰāyēdān* ‘[O light] eternal!’ (M729 II V i, line 6, *Mir. Man.* ii, p. 333).

Man. MPers. *pid-ā-n* ‘our father!’ may be a frozen vocative *pidā* plus the old enclitic 1p pronoun *-n*.

Parthian

kaδ tō šaw-ē [*xwādāy-ā*] *amāh-iž až zād-murd bōž*

‘when you go-PRES-2s [o lord!] us-too from born-dead save-IMP-2s’

= ‘when you go, o lord, save us, too, from rebirths!’ (M5779c, lines 3–5, BBB, p. 45).

Pahlavi has the particle of address *ay* (*hay*): *ay druwand* ‘O wicked (one)!’ (Bd. 30.20).

Man. Parth. *ōn* <wn> is a particle of address or means something like ‘hail!’: *ōn xwaršēd wuzarg* ‘hail, great sun!’ (*Mir. Man.* iii, text e 15–16).

3.1.2 Adjectives and comparison

The endings of the comparative and superlative are *-tar* and *-tom* (Man. MPers. *-dar* and *-dom* after vowels and voiced consonants): simple adjective: *abēzag*, *abēzag-tar*, *abēzag-tom* ‘pure, purer, purest’; participle of necessity: (*zan-išn*)-*tar* ‘more worthy of being killed’; adverbial: *frās-tar* (from *frāz*, in *frās-tar zamān* ‘in the future’), derived adverbial: (*gōwizār-ihā*)-*tar* ‘more in detail’ see section 3.1.3. Irregular formations include the following, which preserve fossilised forms showing vowel gradation (see Old Iranian Chapter 3, section 3.1.2) *gar-ān* ‘heavy (serious)’, comp. *grāy*, superl. *grāy-ist*; comp. *frāy*, *freh* ‘more’, superl. *frāy-ist*, *frāh-ist* ‘most’; *was* ‘much’, *wēš* ‘more’; *bul-and* ‘high’, superl. *bāl-ist*. Note also *kam* ‘little’ *kem* ‘less’, *kamist* ‘least’.

There is a small set of suppletive comparatives and superlatives: *wazurg* ‘great’, *mah-īyl meh*, *mah-ist* ‘greater, greatest’; *kōdag* ‘small’, *keh* ‘smaller’; *xōb* ‘good’, *wah-īylweh* ‘better’, *pahlom*, *pāšom* ‘best’.

Forms such *abdom* ‘last’ (Av. *apa-tama-*) are mere lexical items in Middle Persian.

Parthian has *-istar* and *-istom*, but *-istar* is used for both comparative and superlative: The suffix *-tar* (*-dar*) survives sporadically: *masā-dar* ‘elder’, *abar-dar* ‘higher’. Note *bay-ān bay-istom* ‘the highest god’ (Sogd. *βay-ān βay-tam-i*). Irregular: *qmbyg*, *kamb*, *kambist* ‘little, less, least, fewer, fewest’; Parth. *kalān* ‘big’, *mas* ‘bigger’, *mas-išt* ‘biggest’.

Simple comparisons are expressed by *čiyōn* ‘like’, *ēdōn čiyōn* ‘like’, *mānāg (i)* ‘like’, and the postposition *homān-āg* ‘resembling, like’ (also *pad homānāg ī* and *ēdōn čiyōn-homānāg* ‘just like’).

3.1.3 Adverbs and particles

Adverbs are formed from simple adjectives and other adjectival forms by the productive ending *-ihā*. They can take the comparative and superlative endings.

From simple adjectives: *xūb-ihā* ‘well’; – from compound adjectives: (*abē-brahm*)-*ihā* ‘(walking) (without clothes)-ADV’ = ‘(walking) without clothes’, (*arešk-kāmag*)-*ihā* ‘(envy-desire)-ADV’ = ‘out of envious desire’; *was-ihā* ‘much/many-ADV’ = ‘greatly’; (*a-fraǰūd-az-iš*)-*ihā* ‘non-diminished-from-it)-ADV’ = ‘so that nothing has been diminished from it’ (Dk. 5.2.1, 7.1.41); (*a-padīr-išn*)-*ihā* ‘(non-reception)-ADV’ = ‘without receiving’; – from the reflexive pronouns: *xwēš-ihā* ‘by oneself’; *xwad-ihā* ‘of oneself’; – from adjectival verbal forms: verbal noun in *-āg*: (*tuxš-āg*)-*ihā-tar* ‘(striv-ing)-ADV-COMP’ = ‘more diligently’, (*dān-āg*)-*ihā-tom* ‘(knowing)-ADV-SUP’ = ‘most’

knowingly'; participle in *-ag*: (*zad-ag*)-*īhā* '(strick-en)-ADV' = 'having been struck down'.

3.1.3.1 Adverbs of time, place and manner

These include simple and combined forms:

Time or duration of event: *ahīy* 'first, originally' (Parth. *has*); Man. *az nox* 'from the beginning'; *fradom* (*pad fradomīh*), *nazdist* 'first, at first'; *abdom* 'last, at last' (*pad abdomīh* 'at last, in the end'); *nūn* 'now' (Parth. *awās*); *ahanūz* 'still, yet' (see also section 4.1); *hamēw* (later *hamē*), *hamēšag*, and *hamwār* 'always', *hamēy-īhā* 'forever'; Man. *and(om)* 'for so long'. Adverbial phrases formed with *zamān* 'time', *gāh* 'time (of the day)', *bār* 'time; fois' (Parth. *yāwar*): (*pad*) *ōylān zamān* 'at that time', (*pad*) *ham zamān* 'at the same time, immediately', *andar zamān* 'immediately', *harw gāh* 'every time', *pad ēw bār* 'once'.

Location or direction of event: *ēdar* 'here' and *ānōh* 'there' (Parth. *ēd* and *ōd*), note *pad ham ānōh* 'also in that place'; *awar* 'hither, come here!' (see section 4.3.7.3a), *bē* 'out, away'; adverbial phrases formed with *rōn* 'side, direction' (Parth. *āgōž*) and *gyāg* 'place', e.g. *anda(r)rōn* 'inside', *bērōn* 'out(side)', *az-ō(r)rōn* 'from-hither, from-until now', *ulrōn* 'upward'; *harw gyāg* 'everywhere', *kadām gyāg*, *kū gyāg* 'where?', *abar gyāg* 'on the spot'; Parth. *harwāgōž* 'to all sides, wherever'.

Time or location: *pēs* (*ō pēs*) 'before, pēsīy 'before, in front'; *pas* (Parth. *paš*) 'after, afterward, then, next', *az ān pas* 'after that, afterward', *pasīy* (*azlō pasīy*) 'after, behind' (Parth. *až paš* 'behind, after'); – bare nouns: *bāmdād* 'at dawn', *was frasang zamīg šud* 'he went a ground (= distance) of many *frasangs*'.

Manner of event: *ēdōn* 'thus, in this way', *ōwōn* (Man. *āwōn*, Parth. *awāyōn*), *ēd-ōwōn*, *ēn-ōwōn*, *ān-ōwōn* 'in this/that way'; *ōh* 'thus, in the usual way'; *ēwar* 'for sure, certainly' (also *ēwar-īhā*); *saxt* 'firmly, intensely'; – adverbial phrases formed with *ēwēnag*, *hangōšīdag* 'manner, how', and *čim* 'reason': *pad ēnlān ēwēnag* 'in this/that way', *ham-ēwēnag* 'in the same manner, likewise'; *pad ēn hangōšīdag* 'in this manner'; *imlham čim rāy* 'for this/the same reason', *pad ēdlān čim* 'for this/that reason, therefore'.

Degree: *abēr* (+ adjectives, adverbs), 'very', *was* (+ adjectives, adverbs, verbs) 'very, greatly, in great numbers' and Man. MPers. *frahūd* (= Parth.).

3.1.3.2 Adverbs formed by repetition

Middle Persian typically uses repetition to signify distribution. The repeated word is usually a noun: *gyāg gyāg* 'in various places, from place to place' (also *gyāg ō gyāg*), *āwām āwām* 'in various ages, from age to age', *jār jār* 'from time to time', Parth. *yāwar yāwar*; an adjective: *nōg nōg* 'ever anew, again and again' (Parth. *nawāg nawāg*); an adverb: *andak andak* 'little by little', *jud jud* 'separately, individually' (Parth. *yud yud*); a pronoun: *kas kas* 'each one individually'; a numeral: Man. *yak yak* 'one by one', *se se* 'three each'; phrases: *ast ī ka-ud ast ī ka* 'sometimes-sometimes'; Parth. *bād bād* 'again and again'.

3.1.4 Pronouns and deixis

3.1.4.1 Personal pronouns

Personal pronouns have full forms and enclitic forms. Case distinctions survive in the 1st singular in both Middle Persian and Parthian and, perhaps, in the 2nd singular in

Parthian. In Pahlavi, cases are only sporadically distinguished, especially in Pahlavi translations of Avestan texts. Table 4.7

TABLE 4.7: PERSONAL PRONOUNS

	Middle Persian		Parthian		Enclitic
	DIR	OBL	DIR	OBL	OBL
1s	<i>an</i> <ANE>	<i>man</i> <L>	<i>az</i> <ANE>	<i>man</i> <LY>	- <i>Vm</i>
2s		<i>tō</i> <LK>	<i>*tū</i> <ANT>	<i>*tō</i> <LK>	- <i>Vt</i>
3s		<i>ōy</i> <OLE>		<i>hō</i> <LHw>	- <i>Vš</i>
1p		<i>amāh</i> <LNE>		<i>amāh</i> <LN>	- <i>Vn</i> , - <i>Vmān</i>
2p		<i>ašmāh</i> <LKWM>		<i>ašmāh</i> <LKM>	- <i>Vtān</i>
3p	<i>ōy</i>	<i>awēšān</i> <OLEš'n'>		<i>hawīn</i> <LHwyn>	- <i>Všān</i>
		<i>awīn</i> (Man.)			

The enclitic forms of the personal pronouns are the same in Middle Persian and Parthian, except the 1st plural *-n*, which is found only in the third-century inscriptions as royal plural and, perhaps, in Manichean *pid-ā-n* 'O, our father!'. They are only used as oblique, never as subject. The vowel preceding the enclitic pronouns appears to have varied according to phonetic and morphological context (Sims-Williams 1981). After consonants, the connecting vowel was mostly *-i-* (*-im*, etc.), sometimes *-ulo-* (*-um*, *-om*, etc.).

Parthian inscriptions also have the arameograms <OLY> 'to me' and <OLYN> 'to Us', the readings of which are uncertain (**aw-im*, **aw-in?*).

3.1.4.2 Possessive pronouns

In both Middle Persian and Parthian, the oblique forms of the personal pronouns are used to express possession, also in predicative position. They also have special possessive pronouns used substantivally, but they are quite rare: Pahlavi 1s *man-īg-ān*, 1p *amā-g-ān*; Parthian 1s *man-ān*, 2s *taw-ān* (see section 4.2.3). The Parthian forms appear to be derived from the personal pronouns by means of the suffix *-ān* denoting appurtenance (see section 6.1.2.6).

3.1.4.3 Reciprocity and reflexivity

Reciprocity is expressed by the combination *ēk-did* (*didān*), Man.-Parth. *ēw-bid* (*bidān*) 'one-another'.

The reflexive pronouns are *xwad* (Man. Parth. "xad) '(my-, your-, him-, etc.) self/selves'; and *xwēbaš* (Man.), *xwēš* (Man. Parth. "xēbēh) '(my, etc.) own'. Note the adverbs *xwad-ihā* 'of oneself, (of one's own accord)', *xwēš-ihā* 'by oneself, (alone)'; *xwēšihā-gāh* 'place for being on one's own' = 'toilet'. The combination *xwēš tan* 'own body' is common, and sometimes to be analysed as a simple reflexive pronoun as in New Persian:

[*xwēš tan*] *pad bandag-ih ō kas ma abespār*

'[own body] to slave-ry to somebody do-not hand-over-IMP-2s' = 'do not give your own body over to slavery!' (AAM. 7).

[xwēš tan rāy] zan xwad xwāh
 ‘[own body for] wife self seek-IMP-2s’
 = ‘seek yourself for a wife for yourself!’ (AAM. 42).

3.1.4.4 Demonstrative pronouns

The unmarked (weakly deictic) demonstrative pronouns are identical with the 3rd person personal pronouns in both Middle Persian and Parthian. In Parthian the personal pronoun is also far-deictic. Table 4.8.

TABLE 4.8: DEMONSTRATIVE PRONOUNS

	MPers.	Parth.	MPers.	Parth.
Singular	Near-deictic <i>im</i> <LZNE>	Near-deictic <i>im</i> <ZNE>	Far-deictic <i>ān</i> <ZK>	<i>hō</i> <LHw>
Plural	<i>im-ēšān</i> <LZNE-š’n> <i>ēn</i> <ZNE> <i>ēd</i> <HNA>	<i>im-īn</i> <ZNHn> <i>ēd</i> <ZK>		<i>haw-īn</i> <LHwyn>

Spatial and temporal deixis is usually expressed by MPers. *ēn* . . . *ān*, Parth. *im* . . . *hō* ‘this . . . that’. They are also used in the sense of ‘one . . . the other’, and, by extension, *ān* can be used alone in the meaning of ‘another’:

čārag ī ardaššīr [az ān dar] xwāh-ēd
 ‘means CNCT Ardaššīr [from that door] seek-IMP-2p’
 = ‘seek a means (to deal) with Ardaššīr elsewhere!’ (KN.3.20).

Note also the deictic particle *ēn-ak* ‘look here; voici’:

ēnak āy-ēd nāwāz ī nēw
 ‘look-here comes skipper CNCT good’
 = ‘look, here comes the good skipper!’ (M224 I R 6–7, *Mir. Man.* ii, p. 322).

MPers. *ēd* expresses ‘this one in question’, referring to the written or spoken context. MPers. *im* is used in Pahlavi mainly in adverbial phrases such as *im rōz* ‘today’.

In Manichean Parthian, a few old oblique pronominal forms survive in adverbial phrases: (*ō*) *yahm* ‘until’ (< *yahmāt?), *amābar* ‘hereafter’ (< *ahmāt aparam):

ud ham haw-īn-iž kē [amābar] āžay-ēnd
 ‘and also those-too REL [hereafter] be born-PRES-3p’
 = ‘and also those who will be born hereafter’ (*Mir. Man.* iii, text b 95–96).

3.1.4.5 Interrogative and indefinite pronouns and related adverbials

Interrogative pronouns function as indefinite pronouns when doubled or in combination with demonstratives, indef. *-ēw*, or *-iz* ‘also, -ever’.

Animate/inanimate: *kē* ‘who?’; *čē* ‘what?’; paired, *kē-kē* ‘(whether) someone who–or someone who’.

Individual animate: *kas* ‘somebody’ (also noun: ‘person’); – inanimate: *tis* ‘something’ (also noun: ‘thing’); *ēč* (*hēč*) ‘any (at all)’ (Parth. *keč*, *čiš*, *ēwiž*). Combined: *kas kas*

‘each one’; *ān kas* ‘that person, that someone’; *ān tis* ‘that (some)thing’; *anīylanī-z kas/tis* ‘somebody/something else (too)’. These three occur frequently in questions and negated clauses: *kas-nē* ‘nobody, no one’; *tis-iz-nē* ‘nothing’; *ēč-nē* ‘not any = nothing’.

Inclusive: *har(w)* ‘every’ (pl. *harw-īn*); *wisp* (pl. *wisp-ān*, also Parth.), *harwisp* (*harwist*) ‘every, all’ (pl. *harwisp-īn*, Man. *harwisp-ān*); *hāmōyēn* ‘all, in entirety’; *hamāg* ‘all, entire, whole’; – composite: *harw čē* ‘everything (that)’; *harw kas* ‘everybody (who)’; *harw tis* ‘everything’; *harw N-ēw* ‘each and every N’; note *harw dō*, plur. *harw dōīn*, *dō(n)ān* ‘both’ (see section 3.1.6.1).

Selective: *kadām* ‘which?’; *kadār* ‘which (of two)?’. Composite: *kadār-iz(-ēw)* ‘whichever’; *čēgām-iz(-ēw)* ‘what-ever’.

Quantitive: *čand* ‘several; how much?’; *and* ‘so much/many’; *and čand* and *and-čand* ‘as/so (much/many)–as’ (Parth, *āwend*, *čāwend*, **čāwend*, **čāwend*); *andak* ‘a little’ – composite: *andak andak* ‘little by little’; *ēn-and*, *ēd-and* ‘this much/many’; *ān-and* ‘that much/many’; *čand-ēw* ‘several, a few’.

Adverbial: *kū* ‘where?’; *kay* ‘when?’; *čiyōn*?; *čim* ‘why?’ (see also section 3.1.3.1).

3.1.5 Adpositions

Adpositions comprise prepositions, postpositions and ambipositions. Prepositions are either simple or compounded. Postpositions are always simple, and ambipositions consist of a preposition plus a postposition or a postposed adverb/nominal. Note that prepositions are lowered when governing an enclitic pronoun attached to the first word of a clause (see section 4.2.5.1). In some instances this makes it almost impossible to distinguish between postpositions and preverbs.

Only pre/postpositions: prepositions: *az* ‘from, than’, *ō* ‘to’, *pad* ‘on, with, by means of’; Man. *hambadīz* ‘corresponding to’, *padisāy* ‘on account of’; – postpositions: *hammis(t)* ‘together with’ (usually *az-hammis*); *rāy* ‘on account of, for the sake of’, *tā* ‘except’, etc.

Prepositions that are also adverbs and preverbs: *abāg* ‘(together) with’, *abar* ‘above, on, about’, *andar* ‘in’, *bē* ‘except’ *pas*, *pasīy* ‘after’, *pēš* (*pēšīy*) ‘before’, *tar*, *tarist* ‘across, through’, *tā* ‘to, until’.

Ambipositions: *az-rāy*, *pad-rāy* ‘on account of’; *az-hammis(t)* ‘together with’; *az-rōn* ‘from (the direction of)’; *az-ō(r)rōn* ‘from–in this direction’; *ō-tarist* ‘into’; *ō-rōn* ‘toward, in the direction of’; *bē-ēnyā* ‘except’; *bē-tā* ‘except’.

Parthian: *aδ (-ham)* ‘(together) with’, *frahāh* ‘for the sake of’, *nīrδ* (*niyard?*) ‘near, with’, *parwān* ‘before’, *wasnād* ‘for the sake of’, *yad* (*yad ō*) ‘to, until’, *yahm* ‘until’ (in *yahm [yahm ō yahm] yāvēdān* ‘forever’), etc.

3.1.5.1 Nouns and adjectives as adpositions

Nouns, adjectives, and adverbs used as prepositions usually take the connector *ī* or are combined with *az* or other simple prepositions. In Pahlavi, it is not always possible to decide whether the relative particle should be present or not.

Prepositional phrases with or without the connector *ī*: with adverbs: *pēš (ī)* ‘before, in front of, in the eyes of’; *mayān (ī)* ‘between, in the middle of’; *pērāmōn (ī)* ‘around’; *az pas (ī)* ‘behind, after’ (local); *az abar (ī)* ‘above’; *azēr (ī)* ‘below’; – with adjectives: *mānāg (ī)* ‘like’.

Prepositional phrases with *az*: *pēš az* ‘before’ (time); *pas az* ‘after’ (local, temporal);

abardar az ‘farther above’; *ō(r)rōn az* ‘on this side of’; *bērōn az* ‘outside’; *parrōn az* ‘beyond’; *bēdandar az* ‘outside’, etc.

3.1.6 Numerals

Pahlavi numerals are usually written with numbers, and some are only known from the corresponding ordinals. The vocalisation of 30 and 40 is uncertain. Tables 4.9–10.

TABLE 4.9: CARDINALS

1–9		11–19		10–90		100+	
MPers.	Parth.	MPers.	Parth.	Mpers.	Parth.	MPers.	Parth.
<i>ēw, ēk, yak</i>	<i>ēw</i>	<i>yāzdah</i>	<i>ēwandas</i>	<i>dah</i>	<i>das</i>	<i>sad</i>	<i>sad</i>
<i>dō</i>	<i>dō</i>	<i>dwāzdah</i>	<i>dwādes</i>	<i>wīst</i>	<i>wīst</i>	<i>dwēst</i>	<i>dwēsad</i>
<i>sē</i>	<i>hrē</i>	<i>sēzdah</i>	<i>*hrēdas</i>	<i>sī</i>	<i>*hrīst</i>	<i>*ūrēst</i>	<i>hrēsad</i>
<i>čahār</i>	<i>čafār</i>	<i>čahārdah</i>	<i>čafārđas</i>	<i>čihl</i>	<i>*čafarast</i>	<i>čahār sad</i>	
<i>panz</i>	<i>panj</i>	<i>pānzdah</i>	<i>*panjđas</i>	<i>panjāh</i>	<i>panjāst</i>	etc.	
<i>šaš</i>	<i>šwah</i>	<i>šāzdah</i>	<i>šwahđas</i>	<i>šast</i>	<i>šašt</i>		
<i>haft</i>	<i>haft</i>	<i>hařdah</i>	<i>hařtād</i>				
<i>hašt</i>	<i>hašt</i>	<i>hařdah</i>	<i>hařtād hařtād</i>				
<i>nō</i>	<i>*nah</i>	<i>nōzdah</i>	<i>nawad</i>				
MPers., Parth. <i>hazār</i> ‘thousand’							
MPers., Parth. <i>bēwar</i> ‘ten thousand’.							

TABLE 4.10: ORDINALS

	Middle Persian	Parthian
1st	<i>fradom, naxwist (naxust)</i>	<i>afradom, naxwišt, naxwēn</i>
2nd	<i>didīg, dōwom</i>	<i>bidīg</i>
3rd	<i>sidīg, sēyom</i>	<i>hridīg</i>
4th	<i>tasom, čahārom</i>	<i>čuhrom</i>
5th	<i>panjom</i> etc.	<i>panjom</i>
9th	<i>nohom</i>	<i>nahom</i>
13		<i>hrēdesomīg</i> ‘thirteenth’
etc.		

3.1.6.1 Cardinals and ordinals

Cardinals take the plural (obl.) ending *-ān* (also Parth.), except ‘2’, which, in Pahlavi, takes *-īn* or *-ān* (<2-yn’, 2-īn’>): *ō-mān harw dō-īn (dō-nīn?)* ‘to both of us’ (Bd. 1.21). Manichean Middle Persian has *dō-nān* (= Parth.) and **si-nān* from *dō* and *sē*. Parthian also has *panj-enān*.

The forms **sinān*, and *panjenān* are presumably from old *i*-stems as in other Middle Iranian languages (Khotanese *dvīnu, draiṇu, paṇjīnu*, etc.):

az [si-nān] (āy-āg)-ān (zīw-ēn-āg)-ān im-īn gyān
‘from [three-PL] (com-er)-PL (vivi-fi-er)-PL this-PL soul’
= ‘from these three who shall come and vivify the soul’ (M31 I V 1–2, *Mir. Man.* ii, p. 37 [328])

Parthian:

avāyōn ohrmezd-bay ō [pan.j-enān] rōšn-ān dār-ēd
 ‘in-that-way Ohrmezd-god DO [five-PL] light-PL holds’
 = ‘in that way, god Ohrmezd keeps the five lights’ (KPT 941–3).

Combined numerals:

čahār sad šast ud hašt sār ‘(one thousand) and four hundred sixty-eight years’ (Šāb. 289).
pad hazār-ān ud bēwar-ān wyāg-ān ‘in thousands and ten thousands (of) places’
 (M206 II V 4–5, GW § 144).

Parthian:

pan.jwīst ‘25’, *wīst ud čafār* ‘24’, *sad ud wīst* ‘20’;
hrē hazār dwēsad haštād ud [. . .] ‘three thousand two hundred eighty-[. . .]’ (KawT 8);
šwah yāwar hrēsad ud ēwandas bēwar ud čafār sad frēsta-ān ‘six times 311 × 10,000 + 400 messengers’ (M33 I R i 2–5, *Mir. Man.* iii h 2–5).

Composition forms:

Man. MPers. *tas-kerb* ‘four-shaped’; Parth. *čuhr-pād* ‘four-footed’.

3.1.6.2 Distributives

Distributive numerals are formed by repetition (see section 3.1.3.2).

3.1.6.3 Multiplicatives

Multiplicatives are formed with the suffix *-(a)gānag*: *ē(w)-gānag* ‘single(-minded)’, *dō-gānag* ‘two-fold, double, twin’ *hazār-(a)gānag* ‘a thousand-fold’.

3.1.6.4 Fractions

Fractions include *nēm* ‘one half’, formations like *se ēk-ēw* ‘one third’, and expressions with *bahr* ‘part, share’: *ēn xwāstag se bahr ēw bahr* ‘(of) this property: (of) three parts one part (= 1/3)’ (MHD. 54.3); *az 100 bahr ēw bahr mānd* ‘of 100 parts one part (= 1%) remained’ (Bd. 14.20).

3.1.6.5 Quantity

Words specifying quantity include *was* ‘many’, Man. MPers. *frahīd* ‘many’, *abārīg* ‘the other(s), the remaining’, *and* ‘so many’, *čand* ‘some, as many’. These words are construed like normal adjectives, but note also the typical, distracted, word order:

[mardōm] andar gumēz-išn [was] kē
 ‘[people] in mix-ture [many] (COP) REL . . .’ = ‘there are many people in the Mixture who . . .’ (Bd. 1.25);

hannām-ān rōšn-ān ānōh [frahīd]
 ‘limb-PL light-PL there [many] (COP)’ = ‘there, the limbs of light (are) many’
 (BT 11, 2232–3).

3.2 Verb morphology

Morphological categories include person and number, tense, mood and various stem formations.

3.2.1 Stem formation and semantic modification

The Old Iranian conjugations (present stems) are no longer (see Old Iranian Chapter 3, section 4.1.1). The only derived stems are the causative/denominative and passive. For derivatives from verbal stems, see section 3.2.1.2.

3.2.1.1 Present and past stems

Conjugation is based upon the present and past stems of the verb. The past stem always ends in a dental stop, *-d* after vowels and voiced consonants, otherwise *-t*. Verbs with past stems ending in *-ist*, *-īd* (not Parthian), or *-ād* (rare in Middle Persian, common in Parthian) are ‘regular’, and the present stem is found by removing these endings, e.g. *dwār-ldwār-ist* ‘run’; *purs-īd* ‘ask’; *frēst-īd* ‘send’; *est-ād* (*ēst-ād*) ‘stand’.

Verbs with past stems ending in only *-d* or *-t* are ‘irregular’ and have unpredictable present stems, e.g.: *šaw-lšu-d* ‘go’; *āy-lāma-d* ‘come’; *stan-lsta-d* ‘take’; *jah-ljas-t* ‘jump’; *xwāh-lxwās-t* ‘seek, ask for’; *nišn-* (Ps., Man. *nišīy-*)/*nišas-t* ‘sit (down)’; *gūr-lgrif-t* ‘take, seize’; *kun-lker-d* ‘do, make’, etc. The past stem *mad* (= *āmad*) has no corresponding present stem; similarly Parthian *gad* (= *āyad* ‘came’, present stem *ās-*).

3.2.1.2 Causative and denominative verbs

Causative and denominative verbs derived from present stems and nouns/adjectives by *-ēn-* are common. In the *Psalter*, these causatives are written with *-n-* only; NPers. has *-ān-*, but dialects only *-n-* (see, e.g. ‘Dezfūlī and Šūštārī Dialects’ in *En Ir*, vol. VII, esp. p. 353b).

Causative verbs made from transitive verbs (doubly transitive verbs) are quite rare, e.g.: *ōzan-ēn-* ‘to cause (somebody) to kill (somebody else)’, *dān-ēn-* ‘to cause to know’ (Pahl. Y. 31.17).

Causative verbs made from intransitive verbs are common, e.g.: *šaw-ēn-* ‘to cause to go, send off’, *abesīh-ēn-* ‘to cause to perish, destroy’, *est-ēn-* ‘to make stand, establish, place’, *zīy-ēn-* ‘vivify’ (spelled <*zyw-yn-*>, cf. Man. *zīw-ēn-*).

Denominative verbs are made freely from nouns and adjectives, including those derived from verbs:

- from nouns: *tarāzūg-ēn-* ‘weigh’ (from *tarāzūg* ‘scales’), Ps. *afsōs-īy-* ‘be ridiculed’ (from **afsōs-ēn-* from *afsōs* ‘ridicule’); *rōyišn-ēn-* ‘make grow’ (Dk. 9.12.5; from *rōy-išn* ‘growth’ from *rōy-* ‘grow’);
- from adjectives: *āgāh-ēn-* ‘make aware’ (< *āgāh* ‘aware’); *paydāg-ēn-* ‘make appear, manifest’ (from *paydāg* ‘manifest’); (*ham-juxt*)-*ēn-* ‘(make a) pair’ (<*ham-juxt* ‘together-joined, partner’);
- from past stems (past participles), e.g.: *zād-ēn-* ‘cause to be born’ (*zād* ‘born’).

3.2.1.3 Compound verb phrases

Compound verb phrases with prenominals are frequent: *āgāh kerdan* ‘to inform’, *nigāh kerdan* ‘to look’, *ranj burdan* ‘to suffer’, prepositional *pad ēd dāstan* ‘to think, imagine’, etc.

3.2.1.4 Preverbs

Verbs can be modified by particles that are either exclusively used with verbs or also are pre/postpositions. As postpositions often end up before the verb, they can become indistinguishable from the verb modifiers.

The following common place words modify the sense of verbs beyond their intrinsic meaning:

abar ‘up, on(to)’ expresses upward movement and movement onto or toward. In some cases *abar* may continue an Old Iranian preverb and have no specific meaning:

pad ō ātaxš [abar mad] (= Av. *upā.jasašt* Y. 30.6)
‘next to fire [upon come-PRET-3s]’ = ‘next, he came upon the fire’ (Bd. 4.10);

[abar raft h-ēnd] pad xwēškārīh
‘[upon go-PRET-3p] to duties’ = ‘they went about their duties’ (Dk. 7.1.10);

andar ‘in(side)’ expresses movement inside (and similar) and is used idiomatically with some verbs:

wāy-iz abzār-ēw ī-š padkār [andar abāy-ist]
‘Wāy-too tool-one REL-him-IO for work [be-needed-PRET]’
= ‘Wāy, too, (was) a tool that he needed for the work’ (Bd. 1.46);

abāz ‘back(ward)’ and *frāz* ‘forth, forward’ express backward and forward movement and ‘backward’ action and return (local or temporal), repetition, remainder, and retention (hold back): *abāz est-* ‘to cease’ (e.g. rain):

pad dast [frāz] šāy-ēd [griftan]
‘with hand [forth] be-possible-PRES-3s [take-INF]’
= ‘it is possible to seize (it) with the hand(s)’ (AWN. 18.4);

frōd ‘down’ and *ul* ‘up’ express downward and upward movement, but are also idiomatic with a few verbs:

ēg ān [frōd murd]
‘then that one [down die-PRET-3s]’ (Dk. 7.3.45);

[ul] ō ān [zād] mard ī rāy-ōmand (cf. Av. *us.zaiia-* ‘up-be born’, Y.9.4, etc.)
‘[up] to that [be-born-PRET-3s] man CNCT glori-ous’
= ‘a glorious man was born to him’ (Dk. 7.2.58);

bē ‘out, away’ and aspect:

The preverb *bē*, besides being used as ‘aspectual’ modifier (see section 4.3.8.3), is still used in its original meaning of ‘out, away’, but also as a less specific modifier of the sense of the verb: *bē est-* (+ *az* ‘from’) to desist from (= stop)’.

3.2.2 Nominal verb forms

From both the present and past stem nouns and adjectives can be derived, which can function as such but which also have verbal function. The adjectival forms can be classified as participles, although their occurrences are limited. On their verbal uses, see 4.3.10. On the infinitives, see 4.3.10.2; on the agent and action nouns in *-tār*, *-tār-īhl-dār*, *-dār-īh*, see section 6.1.2.3; on verbal nouns in *-išn* from past stems, see section 6.1.2.2.

3.2.2.1 Nominal forms derived from the present stem

From the present stem, three adjectival forms with active meaning are made: the productive agent noun in *-āg*, the rare *-endag* (*-andag*) (descendant of the Old Iranian active present participle), and *-ān* (descendant of the Old Iranian middle present participle); and the participle of necessity in *-išn*. On their verbal uses, see section 4.3.10.1–3.

3.2.2.1a Agent noun in *-āg*

The form in *-āg* is an agent noun mostly used adjectivally: *šaw-āg* ‘(who/which is capable of) going’, *dān-āg* ‘one who knows, a wise man’, *saz-āg* ‘fitting, appropriate’.

Derived forms are common: *raw-āg* ‘current’: abstract noun (*raw-āg*)-*īh* ‘currency’; denominal verb (*raw-āg*)-*ēn*- ‘to make current, propagate’.

3.2.2.1b Participle in *-andag*

The form in *-endag* (*-andag*) is mostly used adjectivally: *ādur sōz-endag* / *xīwar-endag* ‘burning / consuming fire’, (*mān-endag*)-*ān* ‘those similar (to)’, *zī-ndag* (< *zīy-endag*) ‘living’.

Parthian has forms in *-end-*: *harw-īn* (*adīh-end*)-*ān* ‘all those going in’ (Huy. V, strophe 4).

3.2.2.1c Active participle in *-ān*

This form has only verbal function.

3.2.2.1d Participle of necessity in *-išn*

This form can be derived from any present stem and functions like the Latin gerundive (see section 4.3.10.1). It must be carefully distinguished from the derived noun in *-išn* (see section 6.1.2.2).

3.2.2.2 The verbal adjective in *-ag* from the past stem

The verbal adjective in *-ag* is intransitive(-passive) and expresses state resulting from a past event: *paywast-ag* ‘something connected, continuous(ly)’ (< *paywand-* ‘to connect to’), *āfrīd-ag* ‘someone/something created’, *būd-ag* ‘something that has (always) been’. On its verbal use, see section 4.3.10.3d.

3.2.2.3 Infinitives

Infinitives are either identical with the past stem (‘short’ infinitives) or identical with the past stem plus *-an* (endings *-tanl-dan*, *-īdan*, *-īstan*, *-ādan*). The two forms cannot always be distinguished in the Pahlavi script since the only orthographic difference between the short and the long infinitives is the final stroke. On their uses, see section 4.3.10.2.

3.2.3 'To be'

3.2.3.1 'To be'

The present stem is *h-*, the imperfect stem *ānā-* in the inscriptions and Manichean Middle Persian (3s *anā-d*, 3p *anā-nd*); Parthian has 3s *ahāz* (Skjærvø 1991a). The other tenses and moods, including the imperative and preterite, are supplied by *baw-* *būd* 'to be, become', with 'allegro' forms such as *b-ēd* = *baw-ēd*, *bād* = *baw-ād* (*Psalter* also 1s *b-am*); in Pahlavi, these are mostly hidden by arameograms, but the occurrence of the phonetic spelling <byt'> in series with <YHWWNy't'> suggests they were quite common.

The 3rd singular indicative copula is mostly gapped, the 3rd plural often, and the subjunctive 3rd singular sometimes. The 3rd singular *ast* usually denotes existence.

Parthian does not appear to have an imperative, and the subjunctive is used instead.

See Tables 4.11–12. Table 4.13 gives the arameograms used in the inscriptions.

On the use of 'be' as copula, existential verb, and in possessive constructions, see section 4.3.1.

TABLE 4.11: MIDDLE PERSIAN 'TO BE'

	Indicative	Imperfect	Subjunctive	Optative	Imperative
1s	<i>h-am</i> (Man. <i>h-ēm</i>)	* <i>ānā-n</i>	<i>h-ān</i>		
2s	<i>h-ē</i>				<i>bāš</i>
3s	<i>ast</i> , <i>nēst</i> 'is not'	<i>ānā-d</i>	<i>h-ād</i>	<i>h-ē</i>	
1p	<i>h-ēm</i> (Man. <i>hom</i>)				
2p	<i>h-ēd</i>				<i>baw-ēd</i>
3p	<i>h-ēnd</i>	<i>ānā-nd</i>	<i>h-ānd</i>	<i>h-ēnd hē</i>	

TABLE 4.12: PARTHIAN 'TO BE'

	Indicative	Imperfect	Subjunctive	Optative	Imperative
1s	<i>h-ēm</i> , <i>ah-ēm</i>	* <i>ah-ā-n</i>	<i>ah-ā-n</i>		
2s	* <i>ay-ē</i>			<i>ah-ē-ndē</i>	<i>baw-āh</i>
3s	<i>ast</i> , <i>nē ast</i>	<i>ahā-z</i>	<i>ah-ā-d</i>	<i>ah-ē-ndē</i>	
1p	<i>h-ēm</i> (Man. <i>h-ēm-ād</i>)				
2p	<i>h-ēd</i>				<i>baw-ēd</i>
3p	<i>ah-ēnd</i>		<i>ah-ā-nd</i>	<i>ah-ē-ndē</i>	

TABLE 4.13: MIDDLE PERSIAN AND PARTHIAN 'TO BE': ARAMEOGRAMS

	MPers.	Parth.
Present	<HWE>	<HWY>
3s	<AYTY>, <LOYTY>	<AYTY>
Imperfect	<HWYTN>	<HWE>
Preterite	<YHWWN>	<YHW>

3.2.3.2 Enclitic forms

Enclitic forms are also found, e.g. 1s *-am*, 2p *-ēd*:

guft ō awēšān ka-š dād būd hēnd kū [mard-ēd]
 ‘say-PRET to them when-he-AG create-PAST PRET-3p that [man-be-2p]’
 = ‘(Ohrmazd) said to them when he had created them: You are men’ (Dk. 7.1.9).

Cf. *Psalter*: *kōdak-am* <kwtkm> ‘I am small’ (Ps. 118.141), *kerd-am* <kltm> ‘we have been made’ (Ps. 96 canon).

3.2.4 On the endings

Forms from the present stem have endings indicating the 3 persons and two numbers. In the past tenses, the auxiliaries (‘be, stand’) have the endings, e.g.:

3s pres. ind. *šaw-ēd* ‘he goes,’ subj. *šaw-ā-d*, opt. *šaw-ē*;

3p pret. ind. *šud h-ēnd*, past pret. *šud būd h-ēnd*, perf. *šud est-ēnd*, pluperf. *šud est-ād h-ēnd*.

3.2.4.1 Notes on the present endings

In non-Manichean Middle Persian, the 1st singular is usually spelled with the phonetic complement <-m>, contrasting with <-ym> in the 1st plural, which indicates a distinction between *-am* and *-ēm* (as in NPers.). In Manichean Middle Persian the 1st singular ending is *-ēm* <-ym>.

In the Pahlavi manuscripts (especially of the Pahlavi Avesta), the 1st singular subjunctive ending is frequently written <-ḥy, -ḥd>, sometimes also <-ḥyd>, which I interpret as an archaising spelling for *-ānīy* imitating Av. *-ānī* (differently Cantera 1999, 2004: 295–96).

The optative is extremely rare, but may have ended in *-ē* (see AZ. 40 in section 4.3.8.5).

The 2nd singular ending *-ē* was originally spelled <-ydy> (Ps. *pāy-ē* <NTLWN-ydy>), which in Pahlavi became <-yḏ> and, in the oldest manuscripts, is regularly spelled with the final <-ḏ> resembling a stunted <-b>. In later manuscript, the ending is also spelled <-y’, -ḥy, -’, -y>.

The 2s optative ending *-ēš* is relatively frequent in *Dēnkard* book 9. In Indian manuscripts, it is indistinguishable from *-ē*. It needs to be investigated whether it is only used in imitation of Avestan.

The 2s imperative has no ending, but the original *-a* appears to be preserved as *-ā* before enclitics in the *Psalter* and in Manichean Middle Persian (Sims-Williams 1981): *ahrām-ā-m* <hr ḥm ʾm> ‘raise me (-m) up!’. The 2s indicative is frequently used for the imperative in Middle Persian, the 2s subjunctive in Manichaen Parthian. See also section 4.3.7.3a on *avar* ‘come!’.

A small group of verbs have only the ending *-dl-t* in the 3s: Pahl., Man. MPers. *kun-d* ‘does’, Pahl. **da-d* <dt’>, Man. MPers. <dt> and <td> (for **dah-(a)d*), Pahl. *rav-d* <lpd> ‘moves about’, *bar-d* <bld> ‘carries’, *čāš-t* ‘teaches’, Ps. *hōš-t* ‘dries’, *fragan-d* <plkyndy> from *fragan-* ‘lay the foundation’. See Gershevitch 1970.

The 1p indicative has the ending *-om* in the inscriptions (<-wmy>; it may be restricted to the imperfect) and in Manichean Middle Persian (<-wm>). Pahlavi regularly has <-ym> in both phonetically written forms and arameograms.

The 3p optative has the ending *-ēnd-(h)ē*, with *-ē* spelled <HWE> = *hē* and <HNA> = *ē* (<*ēd*>).

There are two special, ‘shorthand’ forms of the endings of the 3s and 3p derived from <-yt> and <-ynd> (see the facsimiles in Barr 1936). In some Pahlavi texts, these endings

tend to be used indiscriminately, and MacKenzie therefore introduced the transcriptions as <-x₁> and <-x₂>. Other endings are sometimes added, e.g. <YBLWN-x₂-ynd> *bar-ēnd* ‘they carry’.

For the phonetic complements of verbal arameograms, see Skjærvø 1989, 1995a.

For the endings in inscriptional Parthian, see Skjærvø 1986. For the Parthian verb in general, see Durkin-Meisterernst 2000.

3.2.5 System of tenses and moods

The Parthian and Middle Persian verbal systems are more or less identical. On the Parthian verb, see Ghilain 1939; Durkin-Meisterernst 2000b, 2002. The Old Iranian moods and inflection by six persons were preserved, but the old past tenses, with the exception of the imperfect in the earliest inscriptions, were replaced by the new, composite forms we see adumbrated in Old Persian.

From the present stem are made the present indicative and subjunctive, optative, and imperative. From the past stem are made the preterit and perfect tenses, both with the moods indicative, subjunctive and optative.

The imperfect and its accompanying past imperfect survive in the inscriptions, but only the imperfect of ‘to be’ in Manichean Middle Persian and Parthian (see section 3.2.3.1). It was then lost.

The past tenses other than the imperfect are made with the past stem plus ‘to be’ (preterite and past preterite) or ‘to stand’ (perfect and pluperfect).

There is a terminological problem with the past tenses of Middle Persian and Parthian, authors using a variety of different terms. I shall use the terminology set out in Table 4.14.

TABLE 4.14: BASIC VERB SYSTEM

From the present stem:			
present:	+ present endings	<i>šaw-am</i>	‘I go’
imperfect:	+ imperfect endings	<i>*šaw-ēn</i>	‘I went’
From the past stem:			
preterit:	+ present of ‘to be’	<i>šud h-am</i>	‘I went, have gone’
past imperfect:	+ imperfect of ‘to be’	<i>*šud anā-n</i>	‘I had gone’
past preterit:	+ preterit of ‘to be’	<i>šud būd h-am</i>	‘I had gone’
perfect:	+ present of ‘to stand’	<i>šud est-am</i>	‘I have gone’
		<i>nibišt est-ēd</i>	‘it is (stands) written’
pluperfect:	+ preterit of ‘to stand’	<i>šud estād h-am</i>	‘I had gone’
		<i>nibišt estād</i>	‘it was (stood) written’

Note that, instead of ‘preterite’, the term ‘simple past’ is also used on the model of French *imparfait* and *passé simple*; instead of ‘past imperfect’ and/or ‘past preterite’, ‘pluperfect’ is also used; and instead of ‘perfect’ and ‘pluperfect’, ‘present/past perfect’ is also used. Brunner (1977), for instance, uses the term ‘perfect’ for the preterite, ‘pluperfect’ for both the past preterite and the pluperfect, and ‘present perfect’ for the perfect.

See the complete Middle Persian paradigms in Tables 4.15–17. Table 4.18 lists the Parthian present forms only, as the past tenses follow the same pattern as the Middle Persian ones.

TABLE 4.15: MIDDLE PERSIAN PRESENT: 'to do'

	Indicative	Imperfect	Subjunctive	Optative	Imperative
1s	<i>kun-am, -ēm</i>	<i>kun-ē-n</i>	<i>kun-ā-n</i>		
2s	<i>kun-ē</i>		<i>kun-ā</i>	<i>kun-ēš</i>	
3s	<i>kun-ēd (kun-d)</i>	<i>*kun-ē (-ēd)</i>	<i>kun-ā-d</i>	<i>kun-ē</i>	<i>kun</i>
1p	<i>kun-ēm</i>	<i>*kun-om</i>	<i>kun-ā-m</i>		
2p	<i>kun-ēd</i>		<i>kun-ā-d</i>		<i>kun-ēd</i>
3p	<i>kun-ēnd</i>		<i>kun-ā-nd</i>	<i>kun-ē-nd hē</i>	

TABLE 4.16: MIDDLE PERSIAN PAST: 'to go'

	Preterit			Past preterit		
	Indicative	Subjunctive	Optative	Indicative	Subjunctive	Optative
1s	<i>šud h-am</i>	<i>šud h-ā-n</i>		<i>šud būd h-am</i>	<i>šud būd h-ā-n</i>	
2s	<i>šud h-ē</i>			<i>šud būd h-ē</i>		
3s	<i>šud-ø</i>	<i>šud h-ā-d</i>	<i>šud h-ē</i>	<i>šud būd-ø</i>	<i>šud būd h-ā-d</i>	<i>šud būd h-ē</i>
1p	<i>šud h-ēm</i>			<i>šud būd h-ēm</i>		
2p	<i>šud h-ēd</i>			<i>šud būd h-ēd</i>		
3p	<i>šud (h-ēnd)</i>	<i>šud h-ā-nd</i>	<i>šud h-ē-nd hē</i>	<i>šud būd h-ē-nd</i>	<i>šud būd h-ā-nd</i>	<i>šud būd h-ē-nd hē</i>

TABLE 4.17: MIDDLE PERSIAN PERFECT: 'to go'

	Perfect Indicative	Pluperfect Indicative
1s	<i>šud est-am</i>	<i>šud estād h-am</i>
2s	<i>šud est-ē</i>	<i>šud estād h-ē</i>
3s	<i>šud est-ēd</i>	<i>šud estād-ø</i>
1p	<i>šud est-ēm</i>	<i>šud estād h-ēm</i>
2p	<i>šud est-ēd</i>	<i>šud estād h-ēd</i>
3p	<i>šud est-ēnd</i>	<i>šud estād h-ē-nd</i>

TABLE 4.18: PARTHIAN PRESENT: 'to do'

	Indicative	Subjunctive	Optative	Imperative
1s	<i>kun-ām, -am</i>	<i>kun-ā-n</i>		
2s	<i>kun-ē</i>	<i>kun-ā</i>	<i>kun-ēndē</i>	<i>purs(-ā)</i>
3s	<i>kun-ēd</i>	<i>kun-ā-(δ)</i>	<i>kun-ēndē</i>	
1p	<i>kun-ām</i>	<i>kun-ā-m</i>		
2p	<i>kun-ēd</i>	<i>kun-ā-d</i>		<i>kun-ēd</i>
3p	<i>kun-ēnd</i>	<i>kun-ā-nd</i>	<i>kun-ēndē</i>	

3.2.6 Intransitive and transitive

The Middle Persian verb system is dominated by the opposition intransitive vs. transitive. In the present and imperfect, intransitive and transitive verbs are construed in the same way, but in the preterite and perfect (see below), intransitive verbs are construed as in the

TABLE 4.19: INTRANSITIVE AND TRANSITIVE CONSTRUCTIONS

	Present	Preterit
Intransitive:	Subject + Verb-INTR	Subject + Verb
Transitive non-agential:	Subject + Verb-PASS	Subj. + Verb
Transitive agential:	Subj./Agent + Dir. obj. + Verb	Agent + Dir. obj./Subj. + Verb
Examples:	Present	Preterit
Intransitive:	<i>an šaw-am</i> 'I-SBJ go-PRES-1s' = 'I go'	<i>an šud h-am</i> 'I-SBJ go-PRET-1s' = 'I went'
Transitive non-agential:	<i>rōz baxš-ih-ēd</i> 'day divide-PASS-PRES-3s' 'the day is divided' <i>nēkīh dād bawēd</i> 'goodness give-PASS-3s' 'goodness is given'	<i>rōz baxš-ih-ist</i> 'day divide-PASS-PRET-3s' 'the day was divided' <i>nēkīh dād</i> 'goodness give-PRET-3s' 'goodness was given'
Transitive agential:	<i>an tō wēn-am</i> 'I-SBJ you-DO see-PRES-1s' 'I see you'	<i>man tō dād h-ē</i> 'I-AG you-SBJ see-PRET-2s' 'I saw you'

present, but transitive verbs with an ergative (agential/passive) construction (Skjærvø 1985). This system is commonly referred to as a 'split-ergative' system (see section 4.3.3.1). Thus, the preterite has the following meanings:

- intransitive verb of state or action: active meaning;
- transitive verb without agent: passive meaning;
- transitive verb with agent: active and passive meaning, corresponding to active and passive constructions in English ('I killed him' = 'he was killed by me').

3.2.6.1 Intransitive-transitive verbs

Middle Persian has intransitive-transitive pairs of two main types:

1. the transitive stems are characterised by *-ā-* (OIran. causatives in *-aya-*), e.g.: *ahram-* 'go up' (Man.) – *ahrām-* 'lead up', *nibay-* *nibast* 'lie down' – *nibāy-* *nibāst* 'lay down', *nišīn-* (*nišīy-*) *nišast* 'to sit (down)' – *nišān-* (Man. *nišāy-*) *nišāst* 'to seat', etc.;
2. the present takes the suffix *-s-* (OIran. inchoatives in *-sa-*); this type is only Manichean Middle Persian and is also common in Parthian, e.g.: *buxs-* 'be saved' – *bōz-* (Parth. *bōž-*) 'save', with past stem <bwxt>, which was probably **buxt* for both stems. In Pahlavi this type was replaced by stems in *-t-*: *bōxt-* (see next).

3.2.6.2 Morphological passive

There is a small set of synthetic passive stems in Middle Persian surviving from Old Iranian, among them *kir-* 'be done', inscr. 3s *kir-ēd* <klyty> (Man. *kirīh-*); Pahlavi *abesīh-* (*abesīy-*) 'be destroyed' and *škīh-* (*škīy-*) 'be broken' (act. *škenn-* 'break'; see Skjærvø 1997b: 178–79).

Productive passive stems are formed by the suffix *-īh-* (older also *-īy-*) to the present stem:

- suffix *-īy-*: *wizand-* ‘shake’: pass. Ps. *wizand-īy-ēnd* <wcnđ-yđ-yndy> ‘they are shaken’; *kun-* ‘do’: pass. imperf. 3s inscr. *a-kir-īy* <k(y)lydy> ‘it was made’;
- suffix *-īh-*, past stem *-īh-ist* (*-īh-īd*): imperf. 3s inscr. *abzāy-īh* ‘was increased’, *gugān-īh* ‘was destroyed’.

The passive of causative and denominative verbs (see section 3.2.1.2) is formed by replacing *-ēn* with *-īh* (Skjærvø 1997a: 179–80): *afsōs-ēn-* ‘ridicule’: pass. Ps. *afsōs-īy-ēnd* <ṗsws-yđ-yndy> ‘they are ridiculed’; *rawāg-ēn-* ‘propagate’: pass. *rawāg-īh-* ‘be propagated’, *zād-ēn-* ‘give birth to’: pass. *zād-īh-* ‘be given birth to’.

A small set of verbs with past stem in *-xt-* and *-ft-* have passive present stems also in *-xt-* and *-ft-*, e.g. *bōxt-ēd* ‘he is delivered’, *nē frēft-ēd* ‘he is not (to be) deceived’ (see Dhabhar 1955). These forms correspond to Manichean Middle Persian and Parthian ‘inchoatives’ in *-xs-*, *-fs-*, e.g. *buxs-ēd* ‘he is delivered’, *wīfs-ēd* ‘he is deceived’ (see section 3.2.6.1).

3.2.6.3 Periphrastic passive

In both Middle Persian and Parthian, the passive of any transitive verb can be formed with *baw-* ‘to become’:

	<i>guft baw-ēd</i>		<i>bōxt baw-ēd</i>
beside	<i>gōw-īh-ēd</i>	beside	<i>bōxt-ēd</i>
	‘it is being said’		‘he is saved’

4 SYNTAX AND USE OF FORMS

4.1 Coordination and negation

The main coordinating conjunction is *ud* (*u-*) ‘and’ and the disjunction *ayāb* ‘or’ (Parth. *āgām*). The enclitic *-iz* ‘also, as well, even’ is used for coordination and emphasis, for which see section 5.2.2.1. The common negation is *nē*; *ma* is used with (a) the imperative, (b) the particle of exhortation *ēw* (*ma ēw*) and (c) the subjunctive of exhortation. They can be distracted from the verb.

The negations combine with indefinites such as: *kaslēč—nē* ‘no-body/not any = no-one, nothing’; *ahanūz—nē* ‘still not, not yet’; *hagrīz—nē* ‘never’; *hambun-iz—nē* ‘not at all’ (see section 3.1.4.5).

4.2 Noun phrase structure

4.2.1 Modification of nouns by adjectives and nouns

The modifying noun or adjective (pronoun, numeral) can be placed before the head noun (the older and rarer construction) or added after the head noun by means of the ‘relative connector’ *ī* (Man. *īg, ī*), descendant of the Old Persian relative connector and forerunner of the New Persian *ezāfe* (CNCT).

Dependent nouns are in the *oblique case* where retained. The ‘possessive’ relations thus expressed include all those of the ancient genitive: possession, subjective and objective genitives, etc. (see also section 4.2.4):

Without connector	With connector
Singular	
<i>[weh] dēn</i>	<i>šāh [ī wazurg]</i>
'the [good] religion'	'king [CNCT great]' = 'a/the great king'
Plural-OBL	
<i>pad [yazd-ān] nām</i>	<i>pad pušt [ī yazd-ān]</i>
'in the [gods]' name'	'with support [CNCT gods]' = 'with the gods' support'

The plural of the modifier can be used with collective nouns and, regularly, with superlatives: *abar [was-ān mardōm]* 'over [many-OBLp people]' (KN.1.45); *mard-ān pah-lom* 'man-OBLp best' = 'the best of/among men', *yazd-ān mahist* 'the greatest of/among gods'.

The degree of *agreement* between noun and adjective, including pronouns and numerals, varies throughout the history of the language as well as between the different corpora. In the earliest texts, agreement is relatively strict, but in the later texts, where the case system is loosened up or abandoned, agreement tends to serve the function of clarity (see Boyce 1964b).

Several nouns and adjectives can be connected with relative particles:

kār-nāmag [ī ardašīr] [ī pābag-ān]
 'the book-of-deeds [CNCT Ardašīr] [CNCT Pābag-(son-of)]'
 = 'the book of deeds of Ardašīr son of Pābag';

marg [ī alaksandar] [ī hrōmāyīg]
 'death [CNCT Alexander] [CNCT Roman]'
 = 'the death of Alexander, the Roman'.

Frequently, adjectives so connected are separated from the noun:

pābag rāy [pus]-ēw būd [ī abāyišnīg]
 'Pābag for [son]-one was [CNCT comely]'
 = 'Pābag had a comely son' (KN.1.24).

The adjective can be replaced by a prepositional phrase:

afsā ī dibīr [ī az harrān šahrestān]
 'Afsā [CNCT scribe] [CNCT from Harrān town]'
 = 'Afsā, the scribe, from the town of Harrān' (Bīš.9–10).

Enclitic pronouns (see below) can be attached to the relative particle:

harwisp rōz [ī-t zīndagīh]
 'every day [CNCT-your life]'
 = 'every day of your life' (Ps.122.4).

In Parthian, the relative connector *čē* (rarely *kē*) is used less than in Middle Persian and to some extent for the sake of disambiguation (Boyce 1964b; Brunner 1977: 16):

"*xad ast šahrōdār [čē] harw-īn šahr-ān]*
 'himself is ruler [CNCT all-PL realm-PL]'
 = 'he himself is the ruler of all realms' (M32a V, lines 9–10, Boyce 1975: 120, text bo 2);

rōšn-īft ay-ē [čē-mān]

‘light-ness be-2s [CNCT-us]’ = ‘you are our light’ (M83 I R, line 28, Boyce 1975, text au).

4.2.2 *Demonstrative pronouns and numerals*

Demonstrative pronouns and numerals are usually preposed, but occasionally postposed, though without the relative particle:

[pahikar ēn] man

‘[this image] (is) of me . . .’ (ANR m a);

ud [čahār-dah dar] ud [mān panz] ud [gāh sē]

‘and [fourteen doors] and [houses five] and [thrones three]’ (M98 R, lines 20–22, Hutter 1992: 11).

4.2.3 *Possessive pronouns*

The oblique forms of the personal pronouns are used to express possession, also in predicative position, while the rare possessive pronouns are used substantivally:

kē mīzd ō [manīg-ān] frāz dah-ē

‘who fee to [my-OBLp] forth give-PRES-2s’

= ‘you who give the fee to mine (= my people)’ (Y. 40.1);

ōy kē [amāg-ān] tan ud jān zīy-išn u-š yaz-am

‘he who [our-OBLp] body and soul liv-ing (COP) and-him-IO sacrifice-PRES-1s’

= ‘and I sacrifice to him who is our people’s body and soul (i.e.) livelihood’ (Y. 37.3)

ud pad [taw-ān] farrox-īh nēw framen-ām-ā

‘and at [your] fortunate-ness well rejoice-PRES-1p-Exclamation’

= ‘and we do rejoice in your good fortune!’ BBB (425–7).

Parthian:

tō žīr-īft až [man-ān] abar-dar ast

‘you-IO wis-dom than [mine] high-er there-is’

= ‘you have wisdom superior to mine’ (BT 11, lines 1401–2);

harw [taw-ān] kām kirbag ispurr būd

‘all [your] wish good-deed complete became’

= ‘all your desired good deeds were completed’ (*Mir. Man.* iii, text m 20–22)

4.2.4 *Periphrastic adnominal constructions*

The construction with preposed adjective or genitive can be replaced by *ān ī*, *ōy ī*, *ēd ī* ‘the . . . one; that of, the one of, the one belonging to, this . . . of (mine, etc.)’:

[a-sar] rōšn-īh > [ān ī a-sar] rōšn-īh

‘[without-head] light’ > [that CNCT without-head] light’

= ‘the light without beginning’ (Bd. 1.7);

[*star-ān ī*] *axtar-īg ud [ān-iz ī] nē axtar-īg*
 '[star-PL CNCT] zodiac-al and [those-too CNCT] not zodiac-al'
 = 'the stars of the zodiac and also those not of the zodiac' (Bd. 2.1);

[*ōy ī druwand*] *gannāg mēnōy*
 '[he CNCT wicked] Foul Spirit'
 = 'the wicked Foul Spirit; he, the wicked Foul Spirit' (PN. 11);

pad [ēd ī tō kerb]
 'in [this CNCT you(r) shape]'
 = 'in this shape of yours' (Bd. 13.29);

pad [ān ī xwēš (hu-kunišn)-īh]
 'by [that CNCT own (good-doing)-ness]'
 = 'by my/your/his (etc.) own good deeds' (PN. 32).

4.2.5 The use and word order of adpositions

The regular use and word order of adpositions calls for no comments.

Following are two special cases: enclitic pronouns governed by adpositions and adpositions governing complex noun phrases.

4.2.5.1 Enclitic and relative pronouns governed by adpositions

Prepositions can usually take enclitic pronouns:

warrag-ēw [abāg-iš] pad asp nišast est-ād
 'ram-one [with-him] on horse be-seated-PLUPF-3s'
 = 'a ram was sitting with him on the horse' (after KN. 3.18);

u-m kām-ist [pad pēš-iš] namāz bur-dan
 'and-me wished [to before-him] homage carry-INF'
 = 'and I wished to do homage before him' (AWN. 11.4),
 cf.

ka-m kām-ist namāz bur-dan [ohrmazd pēš]
 'when-me wished homage carry-INF [Ohrmazd before]' = 'when I wished to do homage before Ohrmazd' (AWN. 101.2).

The prepositions *az* 'from', *ō* 'to', and *pad* 'on, with' regularly govern the 3rd singular enclitic pronoun, more rarely other enclitic pronouns. In these functions, the prepositions are usually written phonetically: *az-im* <hcm> 'from me', *pad-išān* <ptš'n'> 'on + them', *aw-iš* <wbš> 'to + him' (only form):

zarduxšt [az-išān] ān ī ōy dōysar ī pad men-išn bē nē bast
 'Zarathustra-AG [from-them] that CNCT him eye CNCT in thought *BĒ* NEG closed'
 = 'Zarathustra did not close his mental eye to them' (Dk. 7.3.33).

When the enclitic pronoun is not governed by the preposition, the preposition is written with the arameogram:

ō-[mān <OL-m ḥ'> harw dō' ān]
 'to-[us-OBLp all two-OBLp]'
 = 'to the two of us' (Bd. 1.21);

abar-[iš waxšwar-īh]
 'about-[his prophe-cy]' (Dk. 7.2.61).

When used as *postpositions*, these three appear as *aziš*, *awiš*, *padiš* (spelled phonetically), where *-š* is an empty suffix. In that case, they usually govern an enclitic personal pronoun or a relative pronoun (see also Boyce 1964a):

u-[š] āsmān [aziš] bē tars-īd
 'and-[him] sky [from] *BĒ* feared'
 = 'and the sky was afraid of him' (Bd. 4.10);

spazg-īh ma kun kū-[t] dusraw-īh . . . [awiš] nē rasēd
 'slandorous-ness NEG do-IMP-2s, that-[you] infam-y [upon] NEG comes'
 = 'do not slander, so that infamy may not come [upon you]' (MX. 1.8).

In relative clauses, they may govern the relative pronoun directly or a resumptive pronoun:

pēm [kē] tan-gōhr ī zarduxšt [awiš] mad estād
 'milk [REL] body-substance CNCT Zarathustra [into] come-PLUPF-3s' = 'the milk into which the body substance of Zarathustra had come' (Dk. 7.2.46);

ān mēnōy ī-[š] gannāg-īh ī dām-ān ī ohrmazd [aziš] būd
 'that spirit REL-[him] foul-ness CNCT creature-OBLp CNCT Ohrmazd [from] be(come)-PRET-3s'
 = 'that spirit, from whom stench came to Ohrmazd's creatures' (Bd. 1.49).

4.2.5.2 Adpositions governing complex noun phrases

Nouns governed by prepositions can take all kinds of modifiers, including adjectives, possessives and relative clauses. In such cases, the postposition *rāy* comes at the end of the complex phrase:

gōšurūn [pad gōspand (abāz ō gētīy dah-išn)-īh] ham-dādestān būd
 'Gōšurūn [for cattle (back to world-of-the-living establish-ing)-ness] agreed was'
 = 'Gōšurūn agreed to the re-establishment of (= to re-establish) cattle in the world of the living' (Bd. 4A.6);

ud paydāg kū dēw-ān
 'and manifest (COP) that the demon-PL
[zan-išn ī-šān az ān xvarrah rāy]
 '[strik-ing CNCT-them-IO from that Fortune on-account-of]'
[pad kanīg (petyār-gār)-īh]
 '[for young-woman-IO (adversity-mak)-ing]'
ō hān deh 3 xēn abar bar-ēnd
 'to that village 3 plague upon carry-PRES-3p'
 = 'and it is manifest: the demons, on account of the beating they took from that Fortune, in order to provide adversity for that young woman, will bring three plagues upon that village' (Dk. 7.2.6);

spandarmad abaxšāy-ēd

‘Spandarmad (have-mercy)-PRES-3s’

abar ahlaw-ān [(duš-xwār)-īh ī-šān abar tan rāy]

‘upon righteous-OBLp [(dis-comfort)-ness CNCT-them upon body on-account-of]’

ud abar druvand-ān [pādīfrāh ī-šān abar ruvān rāy]

‘and upon wicked-OBKp [punishment CNCT-them upon soul on-account-of]’

= ‘Spandarmad (the Earth) has mercy on the righteous on account of the discomfort to their bodies and on the wicked on account of the punishment (exact) upon their souls’ (WZ. 35.47).

4.3 Semantics and use of verb forms

See Table 4.15 for a survey of verbal constructions.

4.3.1 ‘To be’

The 1st and 2nd person copula are usually expressed; in the present indicative, the 3s copula is mostly, the 3p often, and the 3s subjunctive sometimes gapped.

4.3.1.1 ‘To be’ and the existential verb

Usually *ast* denotes existence or possession, but we also find it in certain contexts (not yet entirely defined) referring to some preceding discussion or explanation, meaning something like ‘that is (*id est*)’ or ‘and that is X’:

ān ast ohrmazd

‘(Wahman said:) That/he is Ohrmazd’;

az asarag kerb ahumvar frāz būd. ast dēn

‘From the headless form, the *Ahumvar* originated, that is, the *dēn*’.

Parthian: **xad ast šahrōār* ‘he himself is the ruler’.

4.3.1.2 ‘Belonging to’

The adjectival reflexive pronoun (section 3.1.4.3) is commonly used to express possession:

ohrmazd xwēš h-am

‘Ohrmazd’s his-own be-PRES-1s’

= ‘I belong to Ohrmazd’;

pad ān ī xwēš (hu-kunišn)-īh

‘by that CNCT own (good-doing)-ness’

= ‘by his own good deeds’ (PN.32).

4.3.1.3 ‘To have’

‘To have’ is expressed by the verb of existence *h-*, pret. *baw-* and an oblique form of nominals (originally indirect object):

agar-[it] hambun-iz mihr ī amāh [ast]
 ‘if-[you-IO] at-all love CNCT us [there-is-PRES-3s]’
 = ‘if you have any love for us at all’.

In some texts, the postposition *rāy* is used to mark the possessor (as in early New Persian):

[ardawān rāy] kanīzag-ēw abāyišnīg [būd]
 ‘[Ardawān IO] servant-girl-one comely [there was]’
 = ‘Ardawān had a pretty servant-girl’ (KN. 2.1).

Parthian:

*merd-ēw ahāz *ud haft puhr [būd]*
 ‘man-one be-IMPERF-3s and seven son(s) was’
 ‘there was a man, and (he) had seven sons’ (BT 11, lines 83–84).

4.3.2 Agreement between (grammatical) subject and verb

When the subject is the plural of an animate noun (humans or animals) the predicate is usually in the plural, but in the preterite the auxiliary is often omitted. When the subject is not expressed, the auxiliary cannot usually be omitted. In sequences of verbs, the auxiliary may be added to one or more of the forms:

ān-iz-im bē [stad]
 ‘[those]-too- I-AG [seize-PRET(-3p)]’
u-m abāz-ō xwēš šahr [hišt h-ēnd]
 ‘and I-AG back to own land [let-PRET-3p]’
 = ‘those, too, I seized, and I let them return to their own land(s)’ (KKZ 13)

Collective nouns in the singular can take a plural verb:

[grōh-ēw āy-ēnd]
 ‘[group-one come-3p]’
 = ‘a group will come’ (Bd. 33.24).

When the subject is an *inanimate* noun, the predicate is usually in the plural when the individuality is emphasised, also when the noun has no plural ending:

pas [hamāg kōf] frāz [rust h-ēnd] pad 18 sāl
 ‘afterward [all mountain(s)] forth [grow-PRET-3P] in 18 years’
 = ‘afterward, all the mountains grew up in 18 years’ (Bd. 9.1).

There are occasional differences between Middle Persian and Parthian, as in the conclusion of the Hajjabad inscription (see section 8.1).

4.3.3 The ergative construction

In the ‘ergative’ construction, the performer of the action (the subject in the present) becomes the agent and is in the oblique case (frequently an enclitic pronoun), and the direct object of the action becomes the ‘grammatical’ subject and is in the direct case. The verb then normally agrees with the direct object of the action (the grammatical subject):

pas [pōrušāsp] ān hōm [kust]
 ‘next [Pōrušāsp-AG] that *hōm*-SBJ [grind-PRET-3s]’
u-[š] ō ān gāw pēm [gumēxt]
 ‘and-[he-AG] (in)to that cow milk [mix-PRET-3s]’
 = ‘next Pōrušāsp ground that *hōm*, and he mixed (it) into the milk of that cow’ (after Dk. 7.2.46);

ka frabih būd h-am ā-[t] frabihtar [kerd h-am]
 ‘when fat I was then-[you-AG make-PRET-1s]’
 = ‘when I was fat, then you made me fatter’

guft ō awēšān ka-[š dād būd hēnd] kū mard-ēd
 ‘said to them when-[he-AG create-PAST-PRET-3p] that: man-be-2p’
 = ‘(he) said to them when he had created them: You are men’ (Dk. 7.1.9).

Consecutive intransitive and transitive verbs:

When a noun is the agent of a sequence of coordinated or subordinated clauses with transitive verbs, it can be resumed by an enclitic pronoun:

[pābag] (ka-[š] nāmag [dād]) andōh-gen [būd]
 ‘Pābag-SBJ (when-[he-AG] letter [see-PRET-3s]) sorrow-ful [become-PRET-3s]’
u-[š] pad passox (ī ō ardaxšīr [kerd]) [nibišt] kū
 and- [he-AG] in answer (REL to Ardaxšīr made) [write-PRET-3s] that’
 = ‘when Pābag saw the letter, he became sad, and, in the answer he made to Ardaxšīr, he wrote that . . .’ (KN. 1.41–2).

Transitive and intransitive verbs often alternate in one and the same sentence:

ud pas [ō y wirāz] ([čiyōn-išān ān saxwan āšnūd] abar ō pāy [est-ād]
 ‘and then [that Wirāz-SBJ] (as-their that word hear-PRET-3s) on to foot
 [stand-PRET-3s]’
ud dast pad kaš [kerd] ud [guft]
 ‘and hand in arm-pit [do-PRET-3s and [say-PRET-3s]’
 = ‘and then that Wirāz, when he heard [transitive] that speech of theirs, got up
 [intransitive], placed [transitive] his hands under the armpits, and said [transitive]’
 (AWN. 1.20).

Here, the agent pronoun is left out, but could have been included, e.g. *u-š dast pad kaš kerd* ‘and-he-AG . . .’.

In late texts, New Persian-type constructions are also found, in which the preterite of transitive verbs is construed actively like the present:

ud andar [šud h-ēnd] ud ō wištāsp-šā namāz [burd h-ēnd]
 ‘and inside [go-PRET-3p] and to Wištāsp-king homage [bring-PRET-3p]’
ud frawardag bē [dād h-ēnd]
 ‘and letter *BĒ* [give-PRET-3p]’
 = ‘and they went inside, did homage to King Wištāsp, and gave (him) the letter’
 (AZ. 8)

Here *burd hēnd* and *bē dād hēnd* (for *-šān . . . burd* ‘they-AG brought’ and *-šān . . . dād* ‘they-AG gave’) reflect New Persian (active) *burd-and* and *be-dād-and*.

4.3.4.1 Reflexive pronouns in ergative constructions

Reflexive pronouns normally refer to the agent of the sentence:

[ohrmazd] nēryōsang [xwēš aštag] paydāg-ēn-īd est-ēd
 ‘[Ohrmazd-AG] Nēryōsang [own messenger] manifest-CAUS-PERF-3s’
 = ‘Ohrmazd has revealed Nēryōsang as his own messenger’ (Dk. 5.4.6).

Occasionally, they refer to the (‘logical’) direct object when it is grammatical subject:

u-m abāz ō [xwēš] šahr hišt [h-ēnd]
 ‘and I-AG back to [own] land [let-PRET-3p]’
 = ‘and I let them return to their own land(s)’ (KKZ 13).

4.3.4 The use of the tenses

4.3.4.1 Present indicative

The present indicative is the ‘unmarked’ tense and is used for events taking place in the present or soon in the future and for general statements including in subordinate clauses. It is used with the particle *ē(īw)* to express exhortation (see section 4.3.8.2) and with expressions of time to indicate the duration of an event from a point in the past until the present, ‘since’, ‘for’ (similar to New Persian):

*im rōz haft *māhīgān ast [tā] ābustan [h-am]*
 ‘this day seven month there-is [while] pregnant [be-PRES-1s]’
 = ‘today I have been pregnant for seven months’ (KN. 9.19).

4.3.4.2 Imperfect and past imperfect

In the inscriptions, the imperfect is used as a narrative past tense and still contrasts to some degree with the preterite, which is often used in subordinate clauses to express priority (Skjærvø 1985, 1989):

ud uzdēs [gugān-īh]
 ‘and idol [destroy-PASS-IPF]’
ud gilist ī dēw-ān [wišōb-īh],
 ‘and dens CNCT demon-OBLp [ruin-PASS-IPF]’
ud yazd-ān gāh ud nišēm [a-kir-īy]
 ‘and god-OBLp throne and seat [make-PASS-IPF]’
 = ‘and the idols were destroyed, and the dens of the demons were ruined and were made into a throne and seat for the gods’. (KKZ 10).

In the Manichean texts, the imperfects of ‘to be’ are used parallel with regular preterites:

ud awēšān zōr-ān kē ōy zamān hān zahag zāy-ēn-ād
 ‘and those power-PL-AG REL that time that child gave-birth’
harwisp-ān wuzarg šād-īh [būd]
 ‘all-PL great happi-ness [be-PRET-3s]’
u-š harw padīš gaš [anānd]
 ‘and-it all at joyful [be-IPF-3p]’

= ‘and all those powers (archons), who had given birth to that child at that time, they all felt great happiness, and rejoiced at it’ (KPT 510–16);

Parthian:

abāw nē hō zāwar [ahāz] kē zān-ād . . .

‘but not that power [be-IPF-3s] who-AG knew . . .’

‘but it was not that power, (the one) that knew . . .’ (M2 II R i, lines 34–35, *Mir. Man.* iii, text a).

The past imperfect is a regular pluperfect:

u-m pad ān spās

‘and-me-DO for that service

ī-m pad yazd-ān ud šābuhr šāhān šāh [kerd anā-d]

‘REL-I-AG to god-OBLp and Šābuhr king-OBLp king [do-PAST-IPF-3s]’

ān-im kunē(d) šābuhr šāhān šāh . . . kām-kār ud pādixšāy

‘that-me-DO make-IPF-3s Šābuhr king-OBLp king acting-at-will and authoritative’

= ‘and for the service I had performed for the gods and to Šābuhr, king of kings, (for) that Šābuhr, king of kings, put me in complete charge’ (KKZ 1);

[sūrīd anā-d] āz . . . u-š az nas īg dēw-ān

ud az rēm ī druxš-ān [kerd anā-d] ēn nasāh . . .

‘Āz [be-angered-PAST-IPF-3s] . . . and-she-AG from carcass CNCT demon-OBLp and from filth CNCT she-demon-OBLp [make-PAST-IPF-3s] this corpse’

= ‘Āz had been angered . . . and from the impurity of the demons and from the filth of the she-demons she had made this corpse . . .’ (S 13 a6 + S 9 a, lines 4–8, Boyce 1975: 100, text aq 2).

Parthian:

maran [kaft ahā-z] ud yōbahr abnaft

‘death [fall-PAST-IPF-3s] and sickness [retreat-PRET]’

= ‘death had fallen and sickness retreated’ (AR. VII, strophe 1).

4.3.4.3 Preterite and past preterite

The preterite and the past preterite must originally have contrasted with that of the imperfect and past imperfect, but, in Pahlavi, the preterite is the only past tense (beside the historical present). The past preterite is used to indicate that an action or incipient state was completed before something else happened in the past.

Intransitive verbs:

az mēnōy [mad h-am]

‘from world-of-thought [come-PRET-1s]’

nē pad gētīy [būd h-am]

‘not in world-of-the-living [be-PRET-1s]’

= ‘I came from the world of thought; I have not (always) been in the world of the living’ (PN. 1).

With gapped auxiliary:

u-mān ō padīrag [āmad] avēšān widerdagān ruwān
 ‘and-us to-meet [come-PRET-(3p)] those-OBL departed-OBL souls’
 = ‘and the souls of those departed came to meet us’ (AWN. 10.2).

Transitive verbs without agent:

nē abāg tan dād
 ‘not with body [establish-PRET]’
čē ka tan [dād būd] xwāb az pas būd
 ‘for when body [establish-PAST-PRET] sleep after [be-PRET]’
 = ‘(sleep) was not established together with the body, for when the body had been established, sleep came into existence after (it)’ (Bd. 19.2).

Transitive verbs with agent:

u-[š] hamāg dām ud dah-išn az ān bē [kerd]
 ‘and-[he-AG] all creature and creat-ion from that *BĒ* [make-PRET-3s]’
ud ka-[š] bē [kerd būd]
 ‘and when-[he-AG] *BĒ* [do-PAST-PRET-3s]’
ā-š andar ō tan burd
 ‘then-[he-AG] inside to body bring-PRET-3s]’
 = ‘and he made all the creatures and creations from that (body); and, when he had made (them), then he brought (them) into (his own) body’ (PR. 46.3).

4.3.4.4 Present and past perfect

The original function of the present and past perfect was to express a state resulting from a previous event, e.g. *nibišt est-ēd* ‘it is written’, with emphasis on the *presence* of writing, not on the fact that it was written at some specific moment in the past. The perfect was originally intransitive-stative and non-agential:

pad nibišt abar stān [nibišt est-ēd]
 ‘in inscription on *monument [write-PREF-3s]’
 = ‘it is/stands written in the inscription on the *monument’ (ŠKZ 27);

tā ma-dan ī petyārag 6000 sāl zamān [widerd estād]
 ‘until come-INF CNCT adversary 6000 year time [pass-PLUPF-3s]’
 = ‘until the coming of the adversary a period of 6000 years had passed (and was now over)’ (Bd. 5B.15).

Parthian:

garān mast-īft kū [xuft išt-ē]
 ‘heavy drunkenness (COP) where [sleep-PERF-2s]’
 = ‘heavy (is) the drunkenness in which you are asleep’ (*Mir. Man.* iii, text g 91).

mardōm pad ān wiyābān-īh
 ‘people by that being-led-astroy’
ō [uzdēs-[parist-išn]-īh [mad estād h-ēnd],
 ‘to (idol-[worship-ing])-ness [come-PLUPF-3p]’
 = ‘by being led astray by those, people had come to (and were then) worshipping idols’ (Dk. 7.4.72).

In Pahlavi, however, the perfect began to be construed like the preterite:

ēk čand tis ī widēmās ī dahāg pad jādūg-īh andar bābēl [kerd estād]
 ‘one some thing CNCT wonderful REL [Dahāg-AG] by sorcer-y in Babylon
 [make-PERF-PAST-3s]’
 = ‘several wonderful things that Dahāg had made by sorcery in Babylon’ (Dk. 7.4.72);

ud ka man a-franāft būd h-am [kē] wad [dāšt estād h-am],
 ‘and when I un-propagated was [who-AG] bad [hold-PLUPF-1s]’
ā-t a-franāft-tar kerd h-am
 ‘then-you-AG un-propagated-COMP make-PRET-1s’
 = ‘and when I was unpropagated by those who had considered me bad(ly) (held me in low esteem), then you made me more unpropagated’ (AWN. 17.15).

4.3.5 Passive and agency

Most passive sentences were originally agent-less, although the means of an action was regularly expressed (by the instrumental in Old Iranian, see section 6.4.7). In Pahlavi, passive constructions were also influenced by the ergative construction.

4.3.5.1 Instrument

The instrument of an action is regularly expressed by a preposition, usually *pad*:

gōw-išn ud warz [kē padiš] andar mardōm
pad waxšwar-īh wābar-īh-ist h-ēnd
 ‘say-ing and wonder [REL by] among people
 in prophet-hood true-(CAUS-)PASS-PRET-3p’
 = ‘(their) sayings and wonders, by which they were revealed among men as true prophets’ (Dk. 7.1.3).

4.3.5.2 Passive agent

Agents are occasionally used in passive constructions with *baw-*:

ēd wād murnjēn-ēd čē ka-[tān] wād murnjēn-īd
 ‘this wind destroy-IMP-2p for when-[you-AG] wind destroyed (ERG)’
ā-[tān] harwisṣ dām [murnjēn-īd baw-ēd]
 then-[you-AG] entire creation [destroy-PASS-PRES-3s]’
 = ‘destroy this wind! For, when you have destroyed this wind, then you will have destroyed the entire creation’ (after Bd. 21.4).

4.3.5.3 Unspecified agent

In subjectless sentences, a passive form can be translated using ‘one, somebody’:

kū pad gyān ī xwadāy-ān [kōxš-īh-ēd]
 ‘when upon life CNCT lord-OBLp [make-effort-PASS-3s]’
 = ‘when an attempt is made = when somebody makes an attempt upon the life of lords’ (KN.9.15).

When the agent is unknown, it can also be expressed by a 3rd plural, e.g. *gōw-ēnd* ‘they say’ = ‘it is said’; *u-[šān] ōzad* ‘they killed (him)’ = ‘he was killed’.

4.3.6 Direct and indirect object marking

Nouns as direct and indirect objects were in the oblique case in early Middle Persian and can be unmarked in later texts, as well. They can also be marked by adpositions, the direct object more frequently in Manichean, the indirect object frequently in all texts.

4.3.6.1 Unmarked and adpositional direct object

Nouns as direct object are regularly unmarked in Zoroastrian Middle Persian. In late Pahlavi, we occasionally find, presumably under the influence of New Persian, the postposition *rāy* used to indicate the direct object, even in agential constructions:

u-šān zan [rāy] nē ōzad
 ‘and-he-AG that woman-[DO] not killed’
 = ‘and he did not kill that woman’ (KN.9.22).

4.3.6.2 Direct object marked by the preposition *ō*

In the *Psalter* and Manichean Middle Persian, as well as in Parthian, the direct object is often indicated by the preposition *ō*, including in agential constructions, but the use may be restricted to certain verbs, since verbs such as ‘to see, bind’, etc., apparently do not take it. It is not clear to what extent this usage may have been influenced by the Aramaic/Syriac use of the preposition *li-*:

āfur-ēm [ō ēn xwān yōjđahr]
 ‘we bless [DO this table pure]’
ud istāy-ēm [ō tō] sārār ī hu-đadag
 and we praise [DO you] leader CNCT of-good-fortune’
 = ‘we bless this pure table, and we praise you, (our) fortunate leader’ (after M729 I R, lines 2–3, *Mir. Man.* ii, p. 330);

u-š nē padīrift [ō nabīy-ān] ud [ō payāambar-ān]
 ‘and-it-AG NEG received [DO prophet-OBLp] and [DO messenger-OBLp]’
 = ‘and it did not receive the prophets and messengers’ (Ps. 136 canon).

Parthian:

az [ō tō] bōž-ān až harwīn aβnās
 ‘I-SBJ [DO you] deliver-SUBJ-1s from all-OBLp harm’
 = ‘I shall deliver you from all kinds of harm’ (AR. VII, strophe 43).

4.3.6.3 Indirect object

In early Middle Persian, nouns as indirect objects are marked simply by being in the oblique case. Later, they were regularly marked by the adpositions *rāy* ‘for the sake of, for the benefit of’ and *ō* ‘to’.

4.3.6.4 Indirect object and possessor raising

With some verbs that are normally transitive and take a direct object plus a personal indirect object or prepositional phrase in the present, in the agential construction, the grammatical subject does not correspond to the direct object, but to the person affected by an action (Mackenzie 1964, who named it the ‘indirect affectee’). These constructions may be compared to passive constructions in English where an indirect object becomes the grammatical subject: ‘he gave me a book’ > ‘I was given a book by him’:

wēmār h-ē gōšurūn az ān ī gannāg mēnōy wēmār-īh
 ‘sick you are Gōšurūn from that CNCT Foul Spirit sick-ness’
ud kēn ī dēw-ān [abar burd h-ē] (pres.: kēn ī dēw abar tō bar-ēnd)
 ‘and malice REL demon-OBLp-AG [upon bring-PRET-2s]’ (upon . . .-2s = ‘upon you’)
 = ‘you are sick, Gōšurūn, with the Foul Spirit’s sickness and the malice the demons have brought upon you’ (Bd.4.31)

ēd ī-[t] . . . pursīd
 ‘this REL-[you-AG] asked’
u-[m] xīb-īhā [awiš guft] (for: u-m-[it] xīb-īhā [awiš guft])
 and-[I-AG] well-ADV [to] told-PRET(-2s)’
ud [handarz-ēn-īd h-ē] (pres.: ō tō handarz-ēn-am)
 and [instruction-DENOM-PRET-2s]’ (to . . .-2s = ‘to you’)
 = ‘this which you asked about and which I told (you) well and (in which) I instructed you’ (MX. 196);

u-š men-išn bē [āhōg-ēn-īd h-ēnd],
 ‘and-he-AG think-ing *BĒ* [sin-DENOM-PRET-3p]’
 = ‘and he made their thought(s) sinful’ (Bd.14.15).

4.3.7 The use of the moods

4.3.7.1 Subjunctive

The subjunctive is the mode of eventuality, and expresses future, wishes, purpose, indefiniteness (‘whatever’, etc.), etc. (see Lazard 1984, Skjærvø 1986).

4.3.7.1a Main clauses

In main clauses, the subjunctive regularly expresses exhortations to all persons. It is often accompanied by subordinate clauses also in the subjunctive:

āfrīn kerd kū
 ‘vow made that’:
šābuhr anōšag ud jāwēd-šahr [baw-ād] . . .
 ‘Šābuhr immortal and eternal-rule [be-PRES-SUBJ-3s]’
ud an-iz bandag ēd kerd [kun-ān] . . .
 ‘and I-DIR-too servant this done [do-PRES-SUBJ-1s]’
ud az ēdar drust frāz ō dar ī awēšān bay-ān [ras-ān]
 ‘and from here safely forth to court CNCT their-OBLp divinitie-s-OBLp [arrive-PRES-SUBJ-1s]’

= ‘He vowed: Let Šābuhr be immortal and rule for ever! And let me, (His) servant, also succeed in doing this . . .! And let me arrive safely at His Majesty’s court’ (ŠPs. II 5–9, cf. section 4.3.10.3e).

In Manichean Middle Persian, this is negated with the particle *ma-bēz* ‘lest, however’:

ba ma-bēz [h-ād] kas kē gōw-ād

‘but NEG-EXH [be-PRES-SUBJ-3s] somebody who say-PRES-SUBJ-3s’

= ‘but let there be nobody who shall say’ (M9 I V, line 8, *Mir. Man.* ii, p. 298).

Parthian:

kadām žamān [ah-ād] kaδ hō āxrāw društ baw-āh

‘which time [be-PRES-SUBJ-3s] when that itch healthy be(come)-PRES-SUBJ-3s’

kū pad hamag tanbār društ ud abē-dard baw-āh

‘that in entire body healthy and without-pain [be(come)-PRES-SUBJ-3s]’

= ‘when will the time come when that itch shall be healed, so that he will be healthy and painless in the entire body?’ (BT 4, lines 2047–49).

4.3.7.1b Subordinate clauses

In subordinate clauses the subjunctive is used to express future, especially in relative and final clauses:

[kē] ēn nāmag [wēn-ād] ud [pahiburs-ād],

‘who this document [see-PRES-SUBJ-3s] and [read-PRES-SUBJ-3s]’

*ān pad yazd-ān ud xwadāy-ān ud xwēš ruwān rād ud rāst ēw *baw-ēd*

‘that-one for god-OBLp and lord-OBLp and own soul generous and truthful EXH be-PRES-3s’

= ‘whoever shall see and read this document, let him be generous and truthful toward the gods, the lords and his own soul!’ (KKZ 18);

[kū tā] gōhr ī zarduxšt bē ō pidar-ān [ras-ād]

‘[so that] substance CNCT Zarathustra *BĒ* to fathers [arrive-PRES-SUBJ-3s]’

= ‘so that the substance of Zarathustra should come to his parents’

(Dk. 7.2.39);

čiyōn [dān-ā] kū kerd yazd-ān ud amā weh ōwōn kun

‘as [know-PRES-SUBJ-2s] that: done gods-IO and us-IO better thus [do-IMP-2s]’

= ‘as you shall know that (something) is done in the best way for the gods and Us, act in that way!’ (KKZ 3).

Parthian:

u-t grīw až hō pāy-ēd [kū]-š mehgār nē [kar-ā]

‘and-your soul from that protects [that]-it-DO harm NEG [do-PRES-SUBJ-2s]’

= ‘and he protects your soul from that (fire), so that it does not harm it’ (BT 11, lines 775–6);

[kē] buxtag-īft až andar [kišt ah-ād]

‘[who] deliver-ance from inside [sow-PRET-SUBJ-3s]’

= ‘(those) in whom deliverance may have been sown’ (BT 19, §54);

ōn ō tō gyān rōšn pand dah-ā-m [kū] bōγ [wind-ā]
 ‘hail to you soul light advice give-IMP-2s-me-IO [so-that] deliverance [find-SUBJ-2s]’
 = ‘hail! to you, Light soul, give me advice, so that you may find deliverance!’
 (M4a R i, lines 2–4).

In late Pahlavi texts, the subjunctive is used instead of the optative as irrealis:

[agar] nē ēdōn [kerd h-ād],
 ‘[if] NEG in-this-way [do-PRET-SUBJ-3s]’
ān rōšn-ih ō gētīy nē [tāb-ēn-īd h-ād]
 ‘that light-ness on world-of-the-living NEG [shine-CAUS-PRET-SUBJ-3s]’
 = ‘if it had not been done in this way, that light would not have been made to shine upon the world’ (Bd. 7.9).

4.3.7.2 Optative

The optative is used in main clauses to express a wish. In the Manichean texts it survives in some set formulas with 3rd singular present optative of ‘be’: Man. MPers. *ōh b-ē*, Parth. *ōh haw-ēndē* ‘so be (it)!’; MPers. *āfrīd b-ē*, Parth. *āfrīd baw-ēndē* ‘blessed be . . .!’.

The 2nd singular is found in Pahlavi, perhaps in imitation of Avestan (see section 3.2.4.1):

ohrmazd ō zardušt guft kū ōy ī dān-āg [men-ēš]
 ‘Ohrmazd-AG to Zardušt said that: he CNCT know-ing [think-OPT-2s]’
 = ‘Ohrmazd said to Zardušt: You should think the Knowing one!’ (Dk. 9.24.11);

was tō awēšān zardušt [ēz-ēš] u-šān [franām-ēš]
 ‘much you them Zardušt [sacrifice-OPT-2s] and-them [further-OPT-2s]’
 = ‘You should sacrifice to them much, Zardušt, and promote them!’ (Dk. 9.35.20)
 Cf. Y. 65.10 *āpō [jaiḍii-ōiš] zaraθuštra* ‘may you implore the waters, Zarathustra’
 and *[yaz-aēša] mē zaraθuštra* ‘may you sacrifice to me, Zarathustra’ (passim).

4.3.7.2a Impossible conditions (irrealis)

In conditional as well as relative clauses the optative is used to express impossible conditions and imaginary situations (irrealis; see Lazard 1984, Skjærvø 1986; see also Texts 8.1–2):

ud ka nē pānāgīh awēšān rāy [h-ē]
 ‘and if NEG protection their on-account-of [be- OPT-3s]’
ahrimen ud wišūdag-ān hāmōyēn ruwān ī d-ruvand-ān
 ‘Ahrimen and abortions all soul CNCT wicked’
andar dōšox bē [marnj-ēnd-ē] (-ē spelled <HNA>)
 ‘in hell *BĒ* [destroy-PRET-OPT-3p]’
 = ‘and if it were not for their protection, Ahrimen and (his) abortions would be destroying all the souls of the wicked’ (PT, p. 124 §23);

agar man nē [dād h-ē] mēnōy ī būm ud deh
 ‘if I-AG NEG [establish-PRET-OPT-3s] spirit CNCT earth and land’
harwisp mardōm bē ō ērānwēz [šud h-ēnd h-ē] xwaših ī ānōy rāy
 ‘all men away to Ērānwēz [go-PRET-OPT-3p] nice-ness CNCT there on-account-of’
 = ‘if I had not established the ‘spirit’ of earth and land, all men would have gone to Ērānwēz, because it is so pleasant there’ (Bd. 31.1);

ardaxšīr asp dō az bārag-ān ī ardawān

‘Ardaxšīr-AG horse two from steeds-OBL CNCT Ardawān’

kē pad rōz-ēw 70 frasang bē [āy-ēnd h-ē] zēn kerd

‘which in day-one 70 frasang *BĒ* [come-PRES-OPT-3p] saddle did’

= ‘Ardaxšīr saddled two of Ardawān’s horses, which would come 70 frasangs in one day’ (KN.2.14);

čē ka hamāg gēhān ābādān-īh [būd h-ē]

‘for if all world-of-the-living inhabited-ness [be-PRET-OPT-3s]’

agar-išān was-iz [zad ud stō kerd h-ēnd h-ē]

‘if-they-AG much-too [strike-PRET and ruin do-PRET-OPT-3p]’

az dām-ān wizend-īh bērōn nē [šud h-ēnd h-ē]

‘from creatures NEG harm-ing outside NEG [go-PRET-OPT-3p]’

čē-šān nē [tuwān-ist h-ē] šu-dan

‘because-them NEG [be-possible-PRET-OPT-3s] go-INF’

= ‘for, if the entire world were inhabited, even if they had smashed and destroyed a lot, harm would not depart from the creations, because it would not be possible for them to go (away)’ (Bd.28.19);

4.3.7.2b *Irrealis in questions*

The present optative is used in questions to indicate unlikely possibilities:

kē bē-z-ušān sar az hāyān [ul hannār-ē]

‘who out-too-their head from *resting-place [up raise-PRES-OPT-3s]’

u-šān ēn šād-īh [nizēh-ē]

‘and-them this happi-ness [teach-PRES-OPT-3s]’

‘who would lift their head(s) out and up from their *resting-place and teach them this happiness?’ (Šāb. esch., lines 151–3).

Parthian:

kē-m [abdāž-ēndē] až harw [. . .]

‘who-me [deliver-PRES-OPT-3s] from all . . .’

= ‘who would deliver me from all [. . .]?’ (AR. IIIb, strophe 12).

4.3.7.2c *The ‘parabolic’ optative*

The optative is commonly used in relative and comparative clauses together with their main clauses to express assumptions and imaginary comparison:

mān-āg ī ōy čiyōn

‘resembl-ing CNCT he like]

kē pus ī dōst ī (2-sāl)-ag (3-sāl)-ag pad (tōšn-dārišn)-īh [bar-ē]

‘REL son CNCT dear CNCT 2-year-ling 3-year-ling by ?-hold-ing [carry-OPT-3s]’

= ‘just like someone who were to carry his dear two-year-old (or) three-year-old son by holding him?’ (Dk. 7.2.34);

avēšān wēn-išn ōwōn tēz kū ān ī nryand-tom dōšox ōwōn wēn-ēnd

‘their sight thus sharp that that CNCT murki-est hell thus see-PRES-3p’

*čiyōn mard-ēw kē čašm abēr wēn-āg andar ēwēnag [*niger-ē] (ms. <nkyInd>)*

‘like man-one whose eye very see-ing in mirror look-OPT-3s’

*tan ī xwēš [*wēn-ē] (ms. <HZYTWNd>)*

‘body CNCT own [see-OPT-3s]’

= their sight is so so sharp that they see the most murky hell like a man with a very 'seeing' eye were to look in a mirror and see his own body' (PT, p. 124 §22);

ud āwōn čē'ōn zīr merd

'and thus like wise man'

kē draxt-ēw nēw ud bār-war tōhm [wind-ē]

'who tree-one good and fruit-ful seed [find-OPT-3s]'

= 'and just like a wise man, who were to find the seed of a good and fruit-bearing tree' (after M49 II V, lines 10–13, *Mir. Man.* ii, p. 308);

āwōn mān-āg čē'ōn iškōh merd

'thus resembl-ing like poor man'

kē duxt ī nēk [zād h-ē]

REL-IO daughter CNCT good [be-born-PRET-OPT-3s]'

ud abēr hu-čīhr [h-ē]

'and . . . very good-looking [be-OPT-3s]'

= 'like a poor man, to whom a good daughter were born, and who were very beautiful' (BT 4, 2010–13).

Parthian:

kē aš maḍyān was-ān dušmen-īn wirēxt

'who from middle many-OBLp enemies-OBL fled'

ud . . . ō dašt ēw wazurg ud wiyābān [yād-ēndē] . . . ,

'and . . . to plain-one large and deserted [come-OPT-3s]'

aḍyān aš dūr maḍyān dašt kadag ēw [wēn-ēndē]

'then from far middle plain house-one [see-OPT-3s]'

= '(like a man) who (had) escaped from the middle of many enemies, and were to come to a large plain and desert . . . and then from afar were to see a house in the middle of the plain . . .' (BT 11, 791–801).

4.3.7.3 Imperative

The imperative exhibits no special features. It is negated with *ma*. Commands to 3s and 3p and 1p are usually expressed with the particle *ēw* + indicative (see section 4.3.8.2):

ēč zamān ma pāy u-š ōzan

'any time NEG [wait-IMP-2s] and-him [kill-IMP-2s]'

= 'do not wait, but kill him!' (KN. 9.21).

The 2nd singular indicative is not infrequently used as imperative:

ma man abar [zan-ē] srid

'NEG me upon [strike-PRES-2s] Srid'

= 'do not strike me, Srid!' (Dk. 7.2.64).

Parthian:

gyān-um frih-istom ō kūfranaft ay-ē abāž [izwart-ā]

'soul-my dear-est to where go-forth-PRET-2s back [turn-PRES-2s]'

= 'O my dearest soul, return to where you went forth!' (M4b R i, lines 6–8, Boyce 1975: 160).

4.3.7.3a The ‘pseudo-imperative’ *awar*, *awar-ēd*

The original adverb *awar* (*ō*) ‘hither’ is used in expressions of greeting in the function of imperative of *āy-āmad* ‘come’ and is then inflected like an imperative: *awar*, *awar-ēd* ‘come here!’ (Nyberg 1932). It is negated by *ma*.

drīst ud weh ud pad drōd [awar-ēd]
 ‘healthy and good and in health [come-IMP-2p]!’
 = ‘welcome!’ (KN. 12.4);

[ma] drīst [awar]
 ‘[NEG-EXH] healthy [come-IMP-2s]’
 = ‘you are not welcome!’ (M3 R, line 22, Boyce 1975: 45).

Parthian:

[awar] bay ō man wēn
 ‘[come-IMP-2s] god DO me [see-IMP-2s]’
 = ‘come, god! look at me!’ (M4b I V 19).

4.3.8 Aspectual and modal particles

Aspectual and modal nuances can be rendered explicit by particles (see Lazard 1984, Skjærvø 1986):

4.3.8.1 *hamē*

The adverb *hamē* (inscr., Man. *hamēw*) ‘for ever’ is used to express ongoing, progressive, state or action:

kāmag ī gannāg mēnōy [hamē warz-ēnd]
 ‘wish CNCT Foul Spirit [*HAMĒ* perform-PRES-3p]’
 = ‘they keep doing the Foul Spirit’s wish’ (Bd. 1.25);

[hamēw nimāy-ēd ud xand-ēd]
 ‘[*HAMĒ* show-PRES-3s and smile-PRES-es]’
 = ‘he keeps pointing (at him) and smiling’ (KSM 51).

4.3.8.2 *ēw*

The particle *ēw* (inscr., Ps. <ýw>, Man. *hēb*; Pahl. *ē*), negated *ma ēw*, expresses exhortation/warning to 3rd person singular and plural and 1st person plural:

šāh-ān šāh ō ēr-ān šahr [ēw wihēz-ēd]
 ‘kings-OBL king to Iranians-OBL land [EXH move-PREE-3s]’
 = ‘let the King of Kings move hither to the land of the Iranians (= Iran)!’ (after NPi 9);

burdešnōhr [ēw baw-ēm]
 ‘thankful [EXH be-PRES-1p]’
 = ‘let us be thankful!’ (Ps. 135 canon);

drōzan [ma hēb baw-ēnd]

‘deceitful NEG [EXH be-PRES-3p]’

= ‘they [should] not [be] deceitful’ (M5794 II R, lines 15–16, *Mir. Man.* ii, p. 297);

mard uzwān az saxwan ī nē abāy-ēd guf-tan bē [ē pahrēz-ēd]

‘man tongue from speech REL NEG ought-to say-INF away [EXH keep-PRES-3s]’

= ‘a man should keep his tongue from saying things he should not say!’ (PR. 10j1).

4.3.8.3 *bē*

The particle *bē* (Man. <b’>) often appears to express the singularity and completion of an event and to be to some extent complementary to *hamē*, but this is not always the case. In Middle Persian, it is homonymous with the preverb *bē* ‘out, away’, from which it is sometimes difficult to distinguish. It can be separated from the verb, notably by the negation:

ka zamān brēhēn-ēd dām-iz ī ahrimen raw-āg [bē baw-ēd]

‘when time fashions creation-too CNCT Ahrimen mov-ing [*BĒ* be(come)-PRES-3s]’

= ‘when he fashions time, Ahrimen’s creation, too, will start moving’ (Bd. 1.36);

azēr ēn zamīg hamāg gyāg āb [bē est-ēd]

‘under this earth all place water [*BĒ* stand-PRES-3s]’

= ‘under this earth, in every place, there stands water’ (Bd. 1A.10)

dā ka hān kanīg az zarīg [ba murd]

‘until when that girl of sorrow [*BĒ* die-PRET-3s]’

= ‘until that girl died of sorrow’ (M2 V ii, lines 27–29, *Mir. Man.* ii, p. 306).

wināh andar hāsr [bē ma hil-ēd]

‘sin within “mile” [PART NEG-EXH let-IMP-2p]’

= ‘do not let a sin within a mile (of you)!’ (PN. 36).

4.3.8.4 *ōh*

The particle *ōh* ‘thus’ is used as a verbal particle meaning approximately ‘in the usual, well-known way’ (Skjærvø, forthcoming):

u-š dān-ist kū hōm [ōh ras-ēd]

‘and-he-AG knew that Hōm [*ŌH* arrive-PRES-3s]’

= ‘and he knew: the Hōm will come in the usual way’ (Pahl. Y. 9.1).

4.3.8.5 *kāč*

The particle *kāč* (*ka*) ‘if only’ is used with the preterite indicative or optative to express an impossible wish:

ruwān ī kersāsp bē ō zardušt guft kū

‘soul CNCT Kersāsp *BĒ* to Zardušt said that:’

[kāč] man hērbed-ēw [būd h-am]

‘if-only I priest-one [be-PRET-1s]’

kē-m tuwān-ēw pad pušt [h-ē] . . .

‘REL-me power-one in back [be-OPT-3s]’

ud gēhān man zišt pad čāsm [būd h-ē]
 and world me ugly in eye [be-PRET-OPT-3s]
 = ‘Kersāsp’s soul said to Zardušt: If only I were a priest and I had a power in my back (?) and the world had seemed ugly to my eye(s)’ (PR. 18f3);

[kāč ka] man az mādar nē [zād h-am]
 ‘if-only I from mother NEG [be-born-PRET-1s]’
ayāb ka zād h-am pad xwēš baxt pad rahīgīh bē [murd h-ē]
 ‘or when I was born by own destiny in child-hood *BĒ* [die-PRET-Opt-1 s]’
ayāb murw-ēw [būd h-ē] ō drayāb [ōbast h-ē]
 ‘or bird-one [be-PRET-Opt-1s] to ocean [fall-PRET-Opt-1 s]’
ayāb ašmāh bay-ān ēn frašn az man nē [purs-īd h-ē]
 ‘or you-PL divinity-PL this question from me NEG [ask-PRET-OPT-3s]’
 = ‘if only I had not been born from a mother or, once born, I had not reached childhood or I had been a bird (and) and had fallen into the ocean or Your Majesty had not asked me this question’ (AZ. 40; readings follow the manuscript MK).

4.3.9 Impersonal constructions

Various modalities can be expressed by verbal phrases involving impersonal verbs in the 3rd singular or constructions with nouns and adjectives. These constructions can take a subordinate clause in the subjunctive (see section 4.3.7.1b) or an infinitive (short or long, see section 4.3.10.2). To express necessity, a derivative in *-išn* from the present stem is used. In both cases, if an agent is expressed, it is in the oblique case, but in the first case it is an indirect object (historically also direct object), in the second case, it is an agent.

4.3.9.1 Impersonal verbs

These include *abāy-ist* ‘be proper/necessary’, *andar abāy-ist* ‘need’, *kām-ist* ‘wish’, *sah-ist* ‘seem’, *saz-* ‘be proper’, *šāy-ist* ‘can, be possible’, *wurrōy-ist* ‘believe’:

tō kē h-ē kē [man] ēdōn [sah-ēd] kū
 ‘you who are REL [me] thus [seem-PRES-3s] that’
-t harwisp xwār-īh ud āsān-īh padīš
 -you all comfort and easy in’
 = ‘who are you, in whom—it seems to me—there is all happiness and ease’ (Bd. 30.15);

kē-š [wurrōy-ēd] xwar-ēd
 ‘who-him [believe-PRES-3s] eats’
ud kē-š nē [wurrōy-ēd] nē xwar-ēd
 ‘who-him NEG [believe-PRES-3s] NEG eats’
 = ‘he who believes shall eat, and he who does not believe shall not eat’ (PR. 48.60);

bē ka-tān purs-īd ēg-[im] nē [kām-ē]
 ‘but when-you-AG asked then-[me] NEG [wish-PRES-OPT-3s]’
bē ka rāst gōw-am
 ‘except when truth speak-PRES-1s’
 = ‘but since you have asked me, I would wish for nothing except that I (should) speak the truth’ (AZ. 40).

bē [abāy-ēd] [abesīh-ād] wis ī pōrušāsp
 ‘but [be-necessary-PRES-3s] [be-annihilated-PRES-SUBJ-3s] house CNCT
 Pōrušāp’
 = ‘but it is necessary (that) the house of Pōrušāsp should be annihilated’ (Dk.7.2.57).

Active forms with a personal subject are occasionally found (*šāy-ēm* ‘we can’,
wurrōy-ēnd ‘they believe’, etc.):

pad dastwar [abāy-ēnd dāš-tan]
 ‘as spiritual-guide [be-proper-PRES-3p hold-INF]’
 = ‘they should be regarded as spiritual guides’ (Dk. 7.4.47).

4.3.9.2 Impersonal constructions with adjectives

These including expressions meaning ‘possible (for)’ (*dastan*, *tuvān*), Parth. *čār*
 ‘necessary’ (for additional examples see section 4.3.10.2):

ud ahīy framāy-ēn kī
 ‘and at-first order-IPF-1s that:’
 -[*m*] *agar [dastan h-ād]*
 ‘-me-IO if [possible be-PRES-SUBJ-3s]’
ēg ādur-ēw ēdar nišāy-ān
 ‘then fire-one here found-PRES-SUBJ-1s’
 = ‘and, at first, I gave an order (saying): if it becomes possible for me, then I shall
 found a fire here’ (Abnun);

anāy-om pid-ar-ān tawān ud dastan kū
 ‘but-my father-OBLp possible and possible that
 -*tān pāddāšin ī (ēw-rōz)-ag rāy day-ānd šād-ih ī jāyēdān*
 ‘-you-IO CNCT one-day for give-PRES-SUBJ-3p happi-ness eternal’
 = ‘but for my fathers it was possible and possible (= my fathers were able and
 capable) to give you eternal happiness as reward for one day’s (good work)’
 (*Mir. Man.* ii M95/R/4–5 [GGZ]).

4.3.10 The use of verbal nouns and adjectives

4.3.10.1 The participle of necessity in *-išn*

The participle of necessity in *-išn* takes a regular agent:

u-[š] čē [kun-išn]
 ‘and-[he-AG] what [do-NEC]’
 = ‘and what should/must he do?’ (KN.9.15);

[mardōm-ān] pad ēn sē rāh saxt [est-išn] . . .
 ‘[people-AG] on these three path firmly [stand-NEC]’
mizd ī mēnōy [bē nē hil-išn]
 ‘reward CNCT other-world *BĒ NEG* [relinquish-NEC]’
 = ‘people must stand firmly on these three paths; (they) must not relinquish (their)
 reward in the other world’ (PN.28).

4.3.10.2 Infinitive constructions

The infinitive is used largely as in English: as a noun; with verbs meaning ‘to order, command’ and ‘to send, allow’; with impersonal constructions such as ‘it is necessary to’ (often short infinitive), ‘it is time to’, etc.; with adpositions, as well as in enumerations and summaries of contents (see Sample Text 8.5).

4.3.10.2a *Infinitive as subject*

The infinitive is commonly used as subject in impersonal constructions:

čē xwad-iz pad nahčīhr [franaf-tan] būd
 ‘for himself-too to hunt [go-forth-INF] was’
 = ‘for, for himself, it was (time) to go hunting’ (M3 R, lines 14–15, Boyce 1975, text n);

[harw kas] bē [abāy-ēd dān-istan] kū az kū bē mad h-am
 ‘[every person-DO it-is-necessary know-INF] that: from where *BĒ* come-PRET-1s’
 = ‘everybody should know from where he comes’ (PT, 57 §11);

ēk ka-[š] tis [kām-ēd guft-an]
 ‘one when-[him] something’ [wish-PRES-3s say-INF]
 = ‘one (Aahunwar), when he wishes to say something’ (Dk. 9.2.4);

čē-m [ašmāh kām] kun-išn
 ‘which-I-AG [you-IO wish (COP)] do-NEC’
man (āgāh-ēn)-īdan framāy-ēd
 ‘me-IO make-known-INF order-IMP-2p’
 = ‘that which it is your wish I should do (= you wish me to do), please make known to me!’ (DD. Intro. 27.2);

ēk-ēw kū bē nē mīr-ēnd
 ‘one-one (COP) that *BĒ* NEG die-PRES-3p’
ēk-ēw kū [ōza-d] nē [šāy-ēnd]
 ‘one-one (COP) that [kill-INF] NEG [can-PRES-3p]’
 = ‘one (reason) is this: they (the *amahrsbands*) do not die; another is this: they cannot be killed’;

ka-t ahanūz [āma-d zamān nē būd]
 ‘when-you-IO yet [come-INF time NEG be-PRET-3s]’
 = ‘when it was not yet the time for you to come’ (AWN. 4.1).

4.3.10.2b *Infinitive as direct object*

The infinitive as direct object is most often found with verbs meaning ‘command’, ‘permit’, ‘begin’, ‘finish’, etc., but also with other verbs:

abāz [dwār-istan nē hišt]
 ‘back [run-INF NEG permit-ted]’
 = ‘he was not allowed to run back’ (Bd. 1A.6);

u-t [nimāy-am būd-an] ī ristāxēz ud tan ī pasēn
 ‘and-you-IO [show-PRES-1s be(come)-INF] CNCT resurrection and body CNCT final’

= ‘and I shall show you the coming about of the resurrection and the Final Body’ (AWN.5.7);

nihang-ēw azēr [nibiš-tan niwist estēd]

‘section-one below [write-INF begin-PERF-3s]’

= ‘a section (on . . .) has begun to be written below’ (Dk. 7.1.42)

u- [š] nē [tuwān būd hanzaḡ-tan] nibiš-tan

‘and-him-IO NEG [able be-PRET-3s finish-INF] write-INF’

= ‘and he was not able to finish writing’ (M1 177–8).

With *ādūg* ‘able’:

pad ān zamān čē jeh xwāh-ēd gannāg mēnōy [ādūg dād]

‘at that time what Jeh requests Foul Spirit [able give-INF] (COP)’

= ‘at that time the Foul Spirit was able to give Jeh what she requested’ (Bd. 4.7).

Parthian:

ast žamān kaḡ dēnāβar "xēbēh grīw ōh burz [čār dir-dan]

‘there-is time when faithful own soul thus high [able hold-INF]’

čwāyōn samīr kōf

‘like Samīr mountain’

= ‘there will be a time when the faithful will be able to hold his soul just as high as Mount Sumeru’ (*Mir. Man.* iii b 28–31).

4.3.10.2c Infinitive with adpositions

The infinitive is commonly used with adpositions, like the *ing*-forms in English:

[az ma-dan] ī tāzīg-ān ō ēr-ān šahr

‘[from come-INF] CNCT Arabs-OBL to Iranians-OBL land’

[(rawāg-ēn)-īdan] ī ān-išān duš-dēnīh

‘[propagate-INF] CNCT that-their (bad-*dēn*)-ness’

= ‘from the coming of the Arabs to Iran (and) the propagation of their evil religion’ (Bd. 0.2);

[pad (marnj-ēn)-īdan] tag abar kerd

‘[to destroy-INF] attack upon made’

= ‘he attacked to destroy (it)’ (Bd. 1.16);

[gumān ker-dan] ī mardōm-ān . . . [rāy]

‘[doubtful make-INF] CNCT people [in-order-to]’

= ‘in order to make people doubt’ (AWN. 1.3).

4.3.10.3 Participial constructions

The nominals derived from verbs function largely like participles in other languages.

4.3.10.3a Present agent noun in -āg.

The present agent noun in -āg occasionally has verbal function but is construed like a noun:

čiyōn awēšān dwāzdah axtar ud haft abāxtar
 ‘as those twelve constellation and seven planet’
 [*(brēh-ēn)-āg ud (rāy-ēn)-āg*] *ī gēhān h-ēnd*
 ‘[determine-*ĀG* and govern-*ĀG*] CNCT world-of-the-living are’
 = ‘as those twelve constellations and seven planets are the ones that determine (the destiny) and govern the world of the living’ (MX 21);

šahriyār ī xwarāsān-īg [dār-āg] ī dēn [winnār-āg] ī ardāw-ān
 ‘prince CNCT east-ern [hold-*ĀG*] CNCT *dēn* [establish-*ĀG*] CNCT righteous-PL’
 = ‘ruler of the East, upholder of the *dēn*, establisher of the righteous; ruler of the East, who upholds the *dēn* and settles firmly the righteous’ (MIK36 R i 8–10 [IB 6371]).

Parthian:

narah-īg [wāδ-āg]
 ‘hell-ADJ [lead-*ĀG*]
 = ‘who leads to hell’ (*Mir. Man.* iii, text g 128–37).

4.3.10.3b Present verbal adjective (participle) in -andag

The present verbal adjective in *-andag* (*-endag*) has verbal function in late Pahlavi texts (only?):

[dūr kun-andag] ast man ardwhišt amahrspand
 [far mak-ing] is I Ardwhišt *amahrspand*
kē [bahr-ēw dah-andag] ast man . . .
 ‘REL [share-one giv-ing] is I . . .’
 = ‘the one who removes, that is I, the *amahrspand* Ardwhišt; he who gives (people their) shares, that is I . . .’ (Pahl. Yt. 3.13).

4.3.10.3c Active participle in -ān

The form in *-ān* is used as a free participle and after *nīwis-tan* ‘to begin’ (more common in Manichean Middle Persian):

zarduxšt [griy-ān] passox guft
 ‘Zarathustra-AG [weep-ing] answer spoke’
 = ‘Zarathustra answered, weeping’ (PR. 36.7);

ka est-ān xēz-ān ayāb nišm-ān ayāb nibay-ān ayāb bar-ān ayāb wāz-ān
 ‘when stand-ing, ris-ing or sitt-ing or ly-ing or rid-ing or driv-ing’ (*Pursišmāhā* 10 [11]);

[nīwist] mahīy [baw-ān]
 ‘[begin-PRET-3s] big-ger [becom-ing]’
 = ‘it began to become bigger’ (BT 4, lines 568–69).

4.3.10.3d Past verbal adjective in -ag

The verbal adjective in *-ag* from the past stem is sometimes (mostly in ‘late’ texts?) used as in New Persian as a ‘conjunctive participle’ (gerund), meaning ‘after having done such and such’:

ardaxšīr 4000 mard ārāst
 ‘Ardaxšīr 4000 man equipped’
[abar avēšān taxt-ag] šabīxōn kerd
 ‘[upon them run-AG] night-raid made’
 = ‘Ardaxšīr equipped 4000 men (and), having set upon (the enemy), he made a night raid’ (KN. 5.10);

ud az kurd-ān mard 1000 ōzad
 ‘and of Kurds man 1000 killed’
[abārīg xast-ag] dast-gīr kerd
 ‘remaining [wounded-AG] captive made’
 = ‘and he killed 1000 men of the Kurds; having wounded the rest, he took them captive’ (KN. 5.10–11).

4.3.10.3e Past participle

The past participle can be used in figura etymologica with the present, to express something like ‘succeed (in doing), finish doing’:

šābuhr . . . hamē [kerd kun-ād] . . .
 ‘Šābuhr . . . ever [done do-PRES-SUBJ-3s]’
ud an-iz bandag ēd [kerd kun-ān] . . .
 ‘and I-DIR-too servant this [done do-PRES-SUBJ-1s]’
 = ‘Let Šābuhr for ever succeed in doing! And let me, (his) servant, also succeed in doing this!’ (ŠPs. II 6–7, cf. section 4.3.7.1);

harw kē weh-tar dān-ēd [guft gōw-ēd]
 ‘every who better-COMP knows [said say-PRES-3s]’
 = ‘whoever knows better let him succeed in saying it!’ (PT, 159).

5 COMPLEX SENTENCES

5.1 Clause structure and word order

The normal word order is Subject/Agent Object Verb (SOV), but the verb is frequently raised, and, not infrequently, a subject, direct object, or another part of the sentence may be lowered to the position after the verb:

ud [purs-īd] ān ī ahlaw-ān ruwān
 ‘and ask-PRET that CNCT righteous-OBLp soul’
 = ‘and the soul of the righteous ones asked’ (AWN. 4.10);

sidāgar hazārag bun būd ka frēdōn kišvar baxt [salm ud tūz]
 ‘third millennium bottom was when Frēdōn(-OBL) land shared [Salm and Tūz-AG]’
 = ‘it was the beginning of the third millennium when Salm and Tūz divided Frēdōn’s land’ (Bd. 33.3);

mard ī wēn-āg abar nigāh ē dār-ēd [deh]
 ‘man CNCT see-ing over look EXH keep-3s-IND [village]!’
 = ‘let a “seeing” man watch over the village!’ (after Dk. 7.7.29);

kōf frāz rust h-ēnd [pad 18 sāl]
 ‘mountain forth grow-PRET-3p [in 18 year]’
 = ‘the mountains grew up in 18 years’ (Bd.9.1).

5.1.1 Word order in ergative constructions

The word order in the ergative construction normally parallels that of the present (see above): Agent-Grammatical Subject-Verb. But we also find inversion, i.e. Grammatical Subject-Agent-Verb:

Agent = noun:

u-t wars ud rēš [wād-ān wišuft]
 ‘and-you(r) hair and beard [wind-OBLp-AG tousle-PRET]’
 = ‘and the winds have tousled you hair and beard’ (AZ. 86);

pas az ān ī [ardaxšīr] ān kirm [ōzad būd abāz āmad]
 ‘after-from-that CNCT [Ardaxšīr-AG] that Kirm [kill-PAST-PRET-3s] back came’
 = ‘after Ardaxšīr had killed that Kirm, he came back’ (KN. 9.1).

Agent = non-enclitic pronoun:

[kē] tō [dād h-ē] [kē] man [dād h-am]
 ‘[who-AG] you-S [make-PRET-2s] [who-AG] I-S [make-PRET-1s]’
 = ‘[who] made you, [who] made me?’ (Dk. 7.3.60).

Agent = enclitic pronoun:

ud az awēšān agrā-dar kē-[mān] ahīy [dād h-ēnd]
 ‘and than those excellent-COMP who-[we-AG] first [see-PRET-3p]’
 = ‘and he was more excellent than those whom we saw at first’ (KNRm 39);

u-[š] kōf-ihā az ān gōhr bē [rōy-ēn-īd]
 ‘and-[he-AG] mountains from that substance out [grow-CAUS-PRET]’
 = ‘and he made the mountains grow from that substance’ (PR. 46.5).

5.1.1.1 Sequence of enclitics

In chains of enclitic pronouns, the agent is in various positions:

ān-ōwōn-im-[iš] wahišt . . . [nimūd]
 ‘that-manner-me-IO-[he-AG] paradise [show-PRET-3s]’
 = ‘in that manner he showed me paradise’ (KNRb 7–9);

harw kē ēn mizdagtāzī az man padīrift hād
 ‘every REL this good-news from me receive-PRET-SUBJ-3s’
ud ō ēn āfrāh . . . ī-[m]-iš nizist hunsand hād . . .
 ‘and DO this instruction REL-[me-AG]-him-IO] taught content be-SUBJ-3s’
 = ‘everyone who may have received this good news from me and may be content with this instruction, which I taught him . . .’ (BBB., lines 12–17).

5.1.1.2 Fronted verb

If the verb is fronted, the agent can be enclitic to the verb:

guft-[iš] ohrmazd ka-š zan brēhēnīd kā
 ‘said-[he-AG] Ohrmazd-AG, when-he-AG woman fashioned that:’
dād-iz-[im] hē tō kē-t jeh-ān sardag petyārag
 ‘established-too-[I-AG] be-PRES-2s you who-your-IO whore-OBL_p species
 adversary (COP)’
 = ‘Ohrmazd said, when he fashioned the woman: I established you, too, whose
 adversary is the whore species’ (Bd. 14A.1).

Other instances of agent enclitic to the verb:

u-t anīy xīr dād-[om]
 ‘and-you-IO other possessions give-PRET-[I-AG]’
 = ‘and I gave you other possessions (as well)’ (M49 II R, line 3, *Mir. Man.* ii, p. 307).

When the verb is fronted and the agent is a noun, it is usually anticipated by a pronoun proclitic (with *u-*) or enclitic to the verb:

guft-[iš zarduxšt]
 ‘said-[he-AG Zarduxšt-AG]’
 = ‘Zarduxšt said’ (Dk. 7.4.58);

u-[šān] dīd [frawahr] ī mardōm-ān anāgīh
 ‘and-[they-AG] saw [pre-souls-AG] CNCT men-OBL evil’
 = ‘and they, the pre-souls of men, saw the evil’ (after Bd. 3.22).

Note the attachment to the preposed past preterite auxiliary *būd* in the poetic Parthian sample text VI below, Text, (8.7):

[būd-iš sānād] až žafr ahrīwar ō hō ayāg abrang-īft až kū ōsaxt
 ‘[raise-PAST-he-AG-PRET] from pool dark to that place pur-ity, from where came-
 down’
 = ‘he had raised (it) from the dark pool to that place of purity, from where it had
 come down’.

5.1.1.3 Agent resumption

When a noun is the agent of a sequence of coordinated or subordinated clauses, it may or may not be resumed by an enclitic pronoun:

Coordinate, not resumed:

pas [srōš ahlaw ud ādur yazd] dast ī man frāz [grift ud guft]
 ‘then [righteous Srōš and the divine Fire took] my hand [and said]’ (AWN.5.4),
 instead of *grift u-šān guft*.

Embedded in subordinate clause:

pābag ka-[š] pad ān ēwēnag [dīd] abd [sah-ist]
 ‘[Pābag-AG/IO] when-[he-AG] in that way [see-PRET] marvellous [seem-PRET]’
 = ‘When Pābag saw how it was, it seemed marvellous (to him)’ (KN. 1.12).

Here *Pābag* is the actual agent of *dīd* represented by *-š*, but also the indirect object of the impersonal verb *sahist* ‘it seemed’.

5.1.1.4 Fronted agent and resumptive pronoun

A similar pattern is found in main clauses, where the noun is resumed by an enclitic pronoun by means of the particle *ā*-:

ēg [ōy mar] ā-[š] ō pasīy gaw abāz [rān-ēn-īd]
 ‘then [that villain-AG] *Ā*-[he-AG] to back hands [move-CAUS-PRET]’
 = ‘then that villain brought his hands to the back’ (Dk. 7.3.6).’

5.1.2 Questions

Questions without interrogative pronouns or adverbs are indistinguishable from positive statements. For indirect question, see section 5.3.3.1.

Most often questions are introduced by interrogative pronouns and adverbs (see section 3.1.4.5). These can be placed at the beginning (but frequently *after* a raised subject, etc.) or at the end of the question:

rad ī mard-ān [kē] rad ī zan-ān [kē]
 ‘model CNCT men [who?], model CNCT women [who?]’
 = ‘who is the model for men, and who is the model for woman?’ (MX. 60.1);

war ī Jam-kard [kū kard est-ēd]?
 ‘bunker CNCT Jam-made [where? made-PERF-3s]

tan ī sām [kū gyāg nibay-ēd]?
 ‘body CNCT Sām [where? place lie-3s]’

ud mānišn ī srōš [kū]?
 ‘and dwelling CNCT Srōš [where?]’
 = ‘Where was the bunker that Jam made built? In what place lies the body of Sām?
 And where is the dwelling of Srōš?’ (MX. 61.3–5);

ān dō aswār ī ō ēn kustag-rōn mad h-ēnd [kay bē widerd]
 ‘those two horseman REL to this side-ward come-PRET-3p [when? *BĒ* pass-PRET]’
 = ‘Those two horsemen who came hither, when did they pass?’;

mīhr-druz [čand ēwēnag ast]
 ‘contract-cheater [how-many kind there-is]’
 = ‘how many kinds of contract-cheaters are there?’

Parthian:

[až kū ay-ē]
 ‘[from where? be-PRES-2s]’
 = ‘where are you from?’

Disjunctive questions are formed with *ayāb* ‘or’. A negated repeated verb can be gapped:

ēn yaz-išn ō yazd-ān ras-ēd [ayāb] ō dēw-ān
 ‘this sacrific-ing to gods-OBL arrives [or] to demons-OBL’
ud ō frayād ī ruwān ī amāh ras-ēd [ayāb nē]
 ‘and to help CNCT soul CNCT us arrives [or not]’
 = ‘does this sacrifice come to the gods or the demons? and does it come to the help of
 our soul?’ (after AWN. 1.15).

5.2 Coordination and connectors

Middle Persian and Parthian have the usual types of sentences. What is most noteworthy is the frequent use of sentence-introducing particles and sentence connectors.

5.2.1 Sentence connectors

The most common of these is the conjunction *ud* (*u-*) ‘and’, which is often introduced, it seems, to serve as ‘chair’ for an enclitic pronoun. Less frequent, but with similar function, is *ā-* (rarely independent).

Adverbs such as *ēg* ‘then’, *pas* ‘afterward, then’ and *did* ‘next’ have their full meaning:

[*ka*] *šābuhr sagān šāh . . . ēdar ō sad stīm āmad*

‘[when] šābuhr sakas-OBL king here to hundred columns came’

[*u*]-*š nān andar im xān xward*

[and]-he-AG food in this house ate’

= ‘When Šābuhr, king of Sakas, came hither to the Hundred Columns (= Persepolis), he ate food in this house’ (ŠPs. I, lines 2–6);

sē sāl urwar harw čē nē abāy-ēd [ā] nē hōš-ēd

‘three year plant all which NEG be-proper-PRES-3s then NEG dries’

= ‘for three years, all the plants that ought not to, do not dry out’ (Bd. 48.3);

ān ī ō dām ras-ēd tā fraškerd

‘that CNCT to creation arrives until perfecting’

[*ā*]-*š dānist*

[then]-he-AG knew’

= ‘what comes upon the world until the perfecting (of the existence) he knew’ (Bd. I.53);

[*agar*] *dēn nē stay-ē*

‘[if] religion NEG praise-PRES-2s’

[*ēg*]-*it ēn tigr pahikāf-om*

‘[then]-you this arrow pierce-PRES-1s’

= ‘if you do not praise the religion, then I shall drive this arrow through you’ (PR. 48.50);

[*did*] *harwisp-āgāh-īh ī ohrmazd rāy*

‘[next], omni-science CNCT Ohrmazd about

= ‘next, about the omniscience of Ohrmazd’ (Bd. 1.10).

u-š [pas] mayānag ī zamīg suft

‘and-he-AG [then] middle CNCT earth pierced’

= ‘then he pierced the middle of the earth’ (Bd. 4.10);

u-š [pas] bē ō gyāg bar-išn

‘and-he-OBL [afterward] *BĒ* to place to-be-carried’

[*pas*]-*iz bē nih-ēd*

‘[afterward]-too *BĒ* places’

= ‘and he should carry (it) to a place, and then he puts it down, too’ (after ŠnŠ 2.10)

Note the frequent use of *pas* in the Abnun inscription, which recalls that of *pasāva* in Darius’s Bisotun inscription (simplified text):

[pas] *ka abar sāl 3 šābuhr ka hrōmāy āy-ēnd*
 '[then] when in year 3 Šābuhr . . . when Romans-DIR come-PRES/IPF-3p'
 [pas] *an ēdar *anā-n*
 '[then] I-DIR here be-IPF-1s'
 [pas] *kū āšnaw-ēn kū hrōmāy āy-ēnd*
 '[then] when hear-IPF-1s that: Romans-DIR come-PRES/IPF-3p'
 [pas] *an yazd-ān *pādwēh-ēn kū . . .*
 '[then] I-DIR gods-IO implore-IPF-1s that: . . .'
 [pas] *kū āšnaw-ēn kū hrōmāy šābuhr zad l-ēnd . . .*
 '[then] as hear-IPF-1s that: Romans-DIR Šābuhr-AG smash-PRET-3p [then . . .]'.
 = 'Then, in year three of Šābuhr, when the Romans were coming, then I was here.
 Then, when I heard that the Romans were coming, then I implored the gods: . . .
 Then, when I heard that Šābuhr had smashed the Romans [then I . . .]'.

In Manichean Middle Persian, *gahīy* <ghy> is used in similar functions:

ud [gahīy] hāmšahr dudīy hāmīn baw-ēd
 'and [then] universe again summer be(come)-PRES-3s'
 = 'and then, there will again be summer in the universe' [literally: 'the universe will be summer'] (M7981 II R ii, lines 28–30, Hutter 1992: 64).

Parthian has *aḍyān* 'then' and *bid* 'next, again':

[*aḍyān*] *ohrmezd-bay padwah-ād ō mād "xwēbēh*
 '[then] Ohrmezd-god-AG implored to mother own'
 = 'then god Ohrmezd implored his own mother' (*Mir. Man.* iii, text p, lines 4–5);

[*bid*] *mārī mānī wāxt*
 '[then] Lord Mani-AG said'
 = 'then Lord Mani said' (BT 11, line 903).

5.2.2 Coordination and disjunction

Particles and conjunction connecting clauses include words meaning 'and', 'too', 'both . . . and'; 'nor', 'neither . . . nor'; 'or', 'either . . . or' and similar.

5.2.2.1 Coordination

Coordination is expressed by *ud* 'and', the enclitic *-iz* 'and, too', and the adverb *ham* 'also' (literally: 'same'), often in combination; *ud . . . ud*, *ud . . . ud . . . -iz*, *-iz . . . -iz*, *ham . . . (ud) ham* 'both . . . and'.

pādixšāy-īh bē-[iz] dah-ē [ud] bē-[iz] čāš-ē
 'king-ship *BĒ*-[too] give-PRES-2s [and] *BĒ*-[too] teach-PRES-2s'
 = 'you both give kingship and teach' (Dk. 9.35.5);

*ud [ham] awēšān andar *andarwāy āb ul bar-ēnd*
 'and [both] they in atmosphere water up carry'
[ud ham] andar zamīg frōd bar-ēnd
 'and [and] in earth down carry'

kū nēm-rōz wār-ēd

‘that mid-day rains’

= ‘and those both carry the water up into the *atmosphere and also carry (it) down into the earth, so that it rains at midday’ (Bd. 21C.20).

Coordination can also be effected by the periphrastic adnominal constructions (see section 4.2.4):

pādāšn ī ahlaw-ān [ān-iz] ī druwand-ān pādifrāh

‘reward CNCT righteous-OBLp that-too CNCT wicked-OBLp punishment’ (after AWN. 11.6), beside:

pādāšn ī ahlaw-ān pādifrāh-[iz] ī druwand-ān

‘reward CNCT righteous-OBLp wicked-OBLp-[too] punishment’

= ‘the reward of the righteous and the punishment of the wicked’

5.2.2.2 Adversative conjunctions

The conjunctions *bē* ‘but’ and Man. *anāy* express contrast and are often coupled with negations:

andar deh [nē]

‘in village [NEG]’

[bē] pad kustag-ēw ī deh widard

‘[but] to side-one CNCT village passed’

= ‘he passed, not through the village, but to one side of the village’ (KN. 2.16);

ō(r)rōn [ma] awar [anāy] ānōh pattāy

‘hither [NEG-EXH] come-IMP-2s [but] there wait-IMP-2s’

= ‘do not come here, but wait there!’ (M2 I R i, lines 12–13, *Mir. Man.* ii, p. 301).

Parthian has *abāw* and *bēž* ‘but’ expressing contrast:

[abāw] nē hō zāwar ahā-z kē zān-ād

‘[but] NEG that power be-IPF-3s REL-AG knew . . .’

[bēž] hō zāwar rōšn kē . . .

‘[but] that power light REL . . .’

[abāw]-iš hām-čihrag nēst

[but]-it of-the-same-form NEG-COP’

= ‘but it was not that power, (the one) that knew . . . but (rather) that light power which . . . but (that one) is not of the same *form as it’ (M2 II R i, lines 16–20, 34–35, *Mir. Man.* iii, text a);

čē tō ō anīy keč abwēn-āh "xad ma kar

‘what you to other person blame-PRES-2s self NEG do-IMP-2s;

[bēž] pattūd čār až masā-dar-ān

[but] endure-INF ought-to from old-er-OBLp’

= ‘do not do yourself what you blame somebody else for, but (rather) you should endure (it) from those older (than you)!’ (*Mir. Man.* iii, text b, lines 9–11).

5.2.2.3 Correlative negation

Correlative negation ‘neither . . . nor’ is expressed by *nē . . . nē, nē . . . nē-iz* ‘neither . . . nor’:

ān tarāzūg ēč kustag oḡrāy nē kun-ēd
 ‘that balance any side dip NEG does’
[nē] ahlaw-ān rāy [ud nē-iz] druwand-ān
 ‘NEG righteous-OBLp for [and NEG-too] wicked-OBLp’
 = ‘that balance does not dip to any side (i.e. is not biased), neither for the righteous
 nor for the wicked’ (after MX. 1.120).

5.2.2.4 Coordination and negation by repetition

Clauses can be coordinated by repeating a preverb or a negation. These constructions may imitate Avestan ones (see Old Iranian, Chapter 3, 8.1.2):

[bē] abar ō ān deh ān ī abārōn an-(āšt-īh) [pad-ēd]
 [BĒ] up on that village that CNCT evil non-peace [fall-PRES-3s]
bē ān ī abārōn snōh-išn bē ān ī abārōn spazg-īh
 [BĒ] that CNCT evil *bad-weather [BĒ] that CNCT evil slander-er
ud [nē-iz] az ān deh ān ī abārōn anāštīh frāz [abesīh-ād]
 ‘and NEG-too from that village that CNCT evil non-peace [be-cut-off-PRES-
 SUBJ-3s]
[ud nē] ān ī abārōn snōhišn ud nē ān ī abārōn spazgīh
 ‘and NEG that CNCT evil *bad-weather and NEG that CNCT evil slander-er’
 = ‘and upon that village evil discord, evil *bad weather, and evil slander will fall;
 and neither evil discord, evil *bad weather, nor evil slander will be cut off from that
 village’ (Dk. 7.7.17).

5.2.2.5 Disjunction

Disjunction is normally expressed by *ayāb* ‘or’ (Parth. *āgām*), *ayāb . . . ayāb* ‘whether—
 or’. In questions it is preceded by *kadār* ‘which (of two)’:

ast ka māhī [ayāb] wazay [ayāb] sang wārēd
 ‘there-is when fish [or] frog [or] stone rains’
 = ‘there are times when fish, frogs, or stones rain’ (Bd. 21 E.1);

kadār pēš dād xwarrah [ayāb] tan
 ‘which (of the two) first established Fortune [or] body’
 = ‘which was established first: the (divine) Fortune or the body?’ (Bd. 14.8);

kadār [ayāb] ahlaw [ayāb] druwand wurrōy-ēd meh
 ‘whether [either] righteous [or] wicked believes much-COMP’ [see section 4.3.9.1]
kū kas-ān [ayāb] ān ī ahlaw tuxšāg-īhā-tar wurrōy-ēd [ayāb] ān druwand
 ‘i.e. people-OBLp, whether that CNCT righteous zealous-ly-COMP believes [or]
 that CNCT wicked
 = ‘does the righteous or the wicked one believe the most: i.e. among people, does the
 righteous or the wicked one believe the more zealously’ (Pahl. Y. 31.17).

5.2.2.6 Restriction

To express ‘other than, or else’ (*bē . . .*) *ēnyā* (Man. *ba . . . anāy*) is used:

nūn [bē] ka rāst gōw-ē [ēnyā]

‘now [except] if truthfully you-speak [otherwise]

ham-dādestān nē baw-ēm

‘agreed NEG be(come)-PRES-1p’

= ‘now, unless you speak the truth, we will not agree’ (KN. 12.18);

dān-išn [ba] pad ēn dar-ān ī-m az abar nibišt [anāy]

‘knowledge [except] by these doors-OBL REL-I-AG from-above wrote [otherwise]’

ō mardōhm-ān madan nē šāy-ēd

‘to people-OBLp come-INF NEG be-possible-PRES-3s’

= ‘knowledge cannot come to people other than by these doors which I have written above’ (M9 I V, lines 9–11, *Mir. Man.*, p. 298).

5.2.2.7 Parataxis

Various relationships can be expressed without connectors (parataxis), e.g. contrast and consequence:

ēdōn bē kun-ē

‘thus *BĒ* you-do’

kas abar ōy ī did rēš ud ziyān kardan nē tuwān baw-ād

‘anybody unto him CNCT other harm and damage do-INF NEG able be-SUBJ-3s’

= ‘(if) you do thus, (then) nobody will be able to do harm and damage to another’

(Dk. 7.1.21).

5.2.2.8 Particles introducing exegetical passages

In the Pahlavi translations and commentaries on Avestan texts, the particles *ay*, *hād* and *kā* all can mean approximately ‘that is’, but, more specifically, *ay* can mean ‘the meaning is’, *hād* ‘yes (and/but)’; to point out varying opinions, the expression *ast īlkē* (*ēdōn*) *gōwēd* or *ast īlkē* ‘there is someone who (says) (thus)’ are used:

xwāršēd ī a-marg ī rāyōmand ī arwand-asp

‘sun CNCT death-less CNCT wealth-y CNCT fleet-horse(d)’

[kū]-š asp nēk;

‘[i.e.]-it horse good (COP)’

[ast kē] ēdōn gōw-ēd

‘[there-is who] thus says’

[ay] asp ī nēk bē dah-ēd

‘[the meaning is] horse CNCT good gives’

= ‘[translation of the Avestan:] the sun, immortal, shining, with fleet horses;

[Pahlavi:] i.e. it has good horses; there is one who says: the meaning is “he gives good horses” ’ (Pahl. Y. 0.9);

hōm ī ahlaw ī dūrōš

‘Hōm CNCT righteous CNCT *dūrōš*’;

[hād] dūrōš-īh-iš

‘[i.e.] *dūrōš*-ness-his’

ēd kū ōš az ruwān ī mardōm-ān dūr dār-ēd;
 ‘this (COP) that death (*ōš*) from soul CNCT people-OBLp far-away (*dūr*) holds’
rōšn guft
 ‘Rōšn said’
[hād] a-hōš-īh pad hōm bav-ēd
 ‘[yes, and] death-less-ness through *hōm* is’
 = ‘[translation of the Avestan:] Hōm, the righteous, the *durōš*’; [Pahlavi:] i.e. it is called *durōš* for this reason that it holds death far away from men’s souls’; Rōšn said: yes, and (also) because deathlessness comes about through Hōm’. (Pahl. Y. 9.2).

5.3 Subordination

Subordinate clauses usually precede or follow main clauses and other subordinate clauses, but embedding is also found (see Relative clauses).

5.3.1 Conjunctions

Among subordinating conjunctions, we can distinguish between those with one specific function and those with multiple functions, e.g. *agar* ‘if’ and *ma agar* ‘lest, let not’ (with verbs of fearing), *čē* ‘because’, *ka* ‘when, if’; *kū* ‘that, so that, than, where, etc.’, *čiyōn* ‘as, when, because; that is’, *tā* ‘so that, until’. There is further ambiguity in Pahlavi, where *ka*, *kē*, and *kū* are sometimes confused under the influence of New Persian *ke*. Composite conjunctions are common.

Nominal ‘that’-clauses introduced by *kū* are the most common subordinate clauses in Middle Persian; *kā* introduces reported speech (see Section 5.3.3.1).

5.3.2 Relative clauses

The relative pronouns are *ī*, *kē* and *čē*. They can function by themselves as subject, direct object and indirect object, and be governed by pre/postpositions. Their grammatical function (other than as grammatical subject) may be expressed by (independent or enclitic) personal or demonstrative pronouns. Thus, an enclitic pronoun attached to a relative pronoun can be independent or simply resume the relative pronoun, i.e. *kē-š* ‘who his = whose’ or ‘who . . . him/his, ‘who to him = to whom’, etc.’:

ōy [kē-š] zan nē būd
ā-š zan dah-ēnd
 ‘he who-him wife NEG was
 CNCT-him wife give-PRES-3p’
 = ‘to him who had no wife they will give a wife’ (PR. 48.63).

As enclitic pronouns tend to ‘spread’ throughout the clause, doing duty for whatever they are needed for (agent, possessive pronoun, etc.), it is sometimes difficult or impossible to determine their precise attribution:

nēk-rav-išn-īh ān mēnōy
[ī]-š/[ī-š] tan ī xwēš padiš weh bē kerd
 ‘good-go-ing-ness that spirit (COP)
 which-it body CNCT own thereby better made’
 = ‘goodness is that spirit by which he made his body better’ (Bd. 1.35).

With *[ī-š]* . . . *padiš* ‘which-it . . . thereby = with which’ or ‘which-his . . . thereby = by which . . . his’.

The relative pronoun *čē* is rare and mainly used when the referent is not specific (‘whatever, everything that, that which’), corresponding to *kē* animate references (‘whoever, someone who, he who’):

ud [čē] andar ān dar
 ‘an [what] in that chapter (COP)’
 = ‘and whatever (else is) in that chapter’.

5.3.2.1 Generic statements

Relative clauses containing generic statements take the subjunctive (cf. section 4.3.7.1b) and parabolic relative clauses take the optative (see section 4.3.7.2c).

5.3.2.2 Complex relative clauses

Relative clauses can combine with other clauses:

pas ō āb mad [ī]-m guft kū azēr ī ēn zamīg winnārd est-ēd
 ‘next to water came [which = water]-I-AG said that: (water) below CNCT this earth set up-PERF-PRES-3s’
 = ‘next he came to the water, which I said is set up below this earth’ (Bd. 4.10);

u-m anī-z kerdagān ī yazd-ān was kerd
 ‘and-I-AG other-too ritual-S CNCT gods-OBL many did’
[ī]-m agar abar ēn nāmag nibišt h-ē ēg was būd hē
 ‘[which = ritual-S]-I-AG if in this inscription write-PRET-OPT-3s then (ritual-S) much be-PRET-OPT-3s’
 = ‘and I performed also many other rituals for the gods, which, if I had written (them) here, would have been too much’ (KNRm 48-49);

čiyōn mard-ēw [kē]
 ‘like man-one [who]’
tā pad gētīy būd [aziš] hūram-tar nē būd
 ‘as-long-as (man-S) in world-of-the-living was [than (man)] happi-er NEG was’
 = ‘like a man than whom, while he was in the world, there was none happier’ (AWN. 4.6);

ān frazend [ī-š] ašmāh framūd kū bē ōzan
 ‘that child [which-him-DO] you-AG commanded that: *BĒ* kill-IMP-2s’
u-mān nē ōzad
 ‘and-we-AG (child-S) NEG killed’
 = ‘that child which you ordered us to kill we did not kill’ (KN. 10.11).

5.3.3 Nominal clauses

Nominal ‘that’ -clauses introduced by *kū* are the most common subordinate clauses in Middle Persian.

5.3.3.1 Reported speech

Pahlavi, like New Persian, does not have indirect speech; instead, reported speech is phrased as direct statements or questions introduced by *kū* and is therefore identical with nominal ‘that’-clauses. The governing verbs include verbs of utterance, thought, and various kinds of perception (seeing, hearing), but also fear, etc.:

tars-īd kū [ma agar mardōm ī deh wēn-ēnd šnās-ēnd]
 ‘he feared: Let not people CNCT this village see-PRES-3p recognise-PRES-3p’
 = ‘He feared that the people of the village might see (them) and recognise (them)’
 (KN. 2.16).

Indirect questions:

gōw kū [ān 3 tis čē]
 ‘say-IMP-2s that: [that three thing (COP) what?]’
 = ‘tell (me) what those three things are’.

Parthian:

aōyān warrag wāxt kū
 ‘then Warrag-AG said that:’
[-š az aš kū zān-ām bay] kū [čē ast]
 ‘[-it I-DIR [from where? recognise-PRES-1p lord] that: [what? COP]’
 = ‘then Warrag said: wherefrom am I to recognise, lord, what it is?’ (BT 11, lines 719–21).

Disjunctive indirect questions:

ašmāh dān-ēd kū
 ‘you-PL know that:’
[ēn yaz-išn ayāb ō yazd-ān ras-ēd ayāb ō dēw-ān]
 ‘[this sacrific-ing either to gods-OBL arrives or to demons-OBL]’
 = ‘do you know whether this sacrifice comes to the gods or the demons?’ (after ŠnŠ. 9.5).

Occasionally, the conditional conjunction *agar* ‘if’ is used in indirect questions to express ‘whether’:

frōd ward šnās-am [agar] tō h-ē spandarmad
 ‘down turn-IMP-2s know-PRES-1s [whether] you are Spandarmad’
 = ‘turn around (so) I may know whether you are Spandarmad’ (Dk. 7.4.58).

5.3.4 Adverbial clauses

The main types of adverbial clauses and their conjunctions are the following:

- Time: (*ī*) *ka* ‘when’ and *ast ī ka* ‘sometimes’; *čiyōn* ‘as when’; *ēw-bār* ‘once, as soon as’; *pas az ān ka* ‘after’; *pēš* ‘before’; *tā* (Man. *dā*) ‘as long as, until’ (also with infinitive; see also *kū tā*, below), and ‘since’ (expressing duration in the past: Germ. *seit*, Fr. *depuis*), *tā ka*, *tā (pēš) ka*, *tā ān ī ka* ‘before, until’.
- Place: *kū* ‘where’ or combinations such as *ānōh kū* ‘there where’, *ān gyāg kū* ‘the place where’, etc.

- Reason and cause: *čē* ‘because, for’; *az ān čiyōn* ‘for the reason that’.
- Purpose and result: *(kū) iā* ‘so that, in order that’.
- Comparison: *čiyōn* ‘as, in the way that’; *harw čiyōn* ‘in every/whatever way that’; *čiyōn ka* ‘as if’.
- Condition: *agar* and *ka* ‘if’, with indicative.

For examples, see on the uses of the tenses (see section 4.3.4) and the moods (see section 4.3.7).

5.3.4.1 Adverbial phrases

Adverbs in *-ihā*, especially those from compound adjectives, can qualify a verb more broadly:

daw-ēnd [abē-brahm-ihā]
 ‘they run [without-dress-ADV]’
 = ‘they run about undressed’ (M28 I V ii, line 19, Skjærvø 1995b: 251).

spāh ī kirm . . . ō bērōn mad
 ‘army CNCT Kirm to-outside came’
abāg aswār-ān ī ardaxšīr kārēzār [gyān-abespār-ihā] saxt kerd
 ‘with horsemen-OBL CNCT Ardaxšīr battle [soul-surrendering-ADV] hard did’
 = ‘Kirm’s army came out and fought (such) a hard battle with Ardaxšīr’s cavalry as to (as if they were willing to) surrender their souls’ (after KN. 6.16);

ahrimen ud māzanīg-ān ud āz spāh-bed ī abārīg dēw-ān
 ‘Ahrimen and Māzanīg(-demons) and Āz army-chief CNCT remaining demons-OBL’
[(stard homānāg)-ihā]
 ‘[(stunned like)-ADV]’
[(pad xwēš hamēstār-ih wānīd-ag)-ihā]
 ‘[(by own opposition conquered)-ADV]’
bē ō mardōm-ān dīdār-īg baw-ēnd
 ‘*BĒ* to people-OBLp seen-ADJ be(come)-PRES-3p’
 = ‘Ahrimen and the Māzanīg demons and Āz, general of the other demons, as if stunned, having been conquered by their own opponents, become visible to people’ (WZ. 35.35);

was dēw-ān druz-ān pad (bīm-nimāy-išn)-ih . . . aw-iš āmad h-ēnd
 ‘many demons lie-demons for (fear-show-ing)-ness . . . to-him came’
ud hamāg [(a-padīr-išn)-ihā] ud [(zad-ag)-ihā] [stōw-ihā] āz-iš wašt h-ēnd
 ‘and all [(non-receiv-ing)-ADV and [(strick-en)-ADV] and [defeated-ADV] from-him turned’
pad ān ī ōy abēzag (frāz gōw-išn)-ih ī abestāg
 ‘by that CNCT him pure (forth-say-ing)-ness CNCT Avesta’
 = ‘many demons and lie-demons came against him (Zarathustra) to induce fear in him, and all turned back from him without (him) receiving (them), having been struck down, and overcome, by that pure one’s recitation of the Avesta’ (Dk. 5.2.6).

5.3.5 'Subordination' by means of derived nouns and adjectives

Infinitives (see 4.3.10.2a) and the derived nouns in *-īh* (see 6.1.2.4), in particular, are often used instead of subordinate clauses, especially with prepositions like *pad* 'for', etc. and the postposition *rāy* 'in order to, on account of':

sag-ān [pad hamēstār-īh] ī gurg-sardagān
 'dogs [for oppos-ition] CNCT wolf-species-OBLp'
ud [pānag-īh] ī gōspand-ān [kerdan rāy] dād estēd
 'and [protect-ion] CNCT cattle-OBLp [do-INF in-order-to] establish-PERF-PRES-3s'
 = 'dogs have been established (in the world) for the opposition of (= in order to oppose) the wolf species and for the protection of (= in order to protect) the domestic animals' (Bd. 24.38);

andar dōšox druj ī spazg-īh [(garān-wināh)-īh rāy] dwār-išn ī pas
 'in hell lie-demon CNCT slander [(heavy-sin)-ness on-account-of] runn-ing to back (COP)'
 = 'in hell every lie-demon runs in front (of the damned), but the lie-demon of slander runs behind because it is such a heavy sin' (after MX. 1.12).

Compound adjectives are often used as alternatives for subordinate clauses, for instance, relative clauses:

ud anī-z ērbed-ān ud mowmard-ān ī ōšmurd-mānsr
 'and other-too teachers and priests CNCT recited-(holy-text)'
ī (dranjēn-īd)-zand ī uskārd-dādestān
 'CNCT memorised-zand CNCT studied-law'
ī āgāh-dēn ī (nimū-dār)-čīm ī (wizī-dār)-kirbag
 'CNCT aware-dēn CNCT (show-er)-meaning CNCT (discrimin-ator)-good-deeds'
pad čand gyāg h-ēnd
 'in several place are'
 = 'and there are other teachers and priests, too, in many places, who have recited the holy texts, who have memorised the *zand* (the Pahlavi versions), who have studied the law, who are aware of the *dēn* (tradition), who can show the meaning, and who can say what are good deeds and what are not (DD. Intro. 9.3).

5.3.6 Dating

In Middle Persian, dates are expressed by the names of the days and months and the year according to the Yazdegerd era (631 CE):

frazaft pad drōd ud šādīh ud rāmišn
 'completed in good health and happiness and joy'
andar rōz šahrīwar ud az māh ābān 908 ī yazdgerd ī šāhān šāh
 'on the day of Šahrīwar and the month of Ābān (in the year) 908 of King of Kings Yazdegerd' = 908+631 = 1639';

ēn ayādgārīhā nibištāg būd estād pad māh wahman ī andar sāl 324 rōz day pad ādur
 'these stories were written in the month of Wahman in the year 324 on the day of Day before Ādur' = 955;

ēn ayādgār andar rōz xwaršēd māh šahrīwar kadīm wihēzagīh sāl ī šaš-sad nawad-ēk nibišt

‘this story was written on the day of Xwaršēd, month of Šahrīwar in the intercalary year (according to the) Qadim (calendar) of 691’ = 1321.

Parthian inherited the dating formula seen in Old Persian, Sogdian, etc.:

pad čafār saxt šahrīwar māh

‘on four passed Šahrīwar month’

= ‘on the fourth of the month of Šahrīwar’ (*Mir. Man.* iii, text c 23–24);

pad saxt čuhram māh šahrīwar pad dōšambat ud žamān ēwandas

‘on passed fourth month Šahrīwar on Monday and hour eleven’

= ‘on the fourth of the month of Šahrīwar, on Monday, at 11 o’clock’ (*Mir. Man.* iii, text d 57–60).

6 LEXIS: WORD FORMATION

The main strategies for word formation are the use of prefixes, affixes, compounds, and combinations of these.

6.1 Nominal derivation

6.1.1 Prefixation

The following are some of the most common prefixes making adjectives from nouns or verbal stems:

- *a-* (*an-* before vowel) and *abē-* denote lack of something; antonyms include adjectives in *-ōmand* (from nouns) and *-āg* (from verbal stems): *sēj* ‘danger’: *sēj-ōmand* ‘full of danger’, *a-sēj* ‘without danger’; *dān-* ‘to know’: *dān-āg* ‘knowledgeable’, *a-dān* ‘without knowledge, stupid’.
- *ham-* ~ *jud-* ‘being of the same X’, ‘being of different/opposite X’ (Parth. *hām-* ~ *yud-*): *ham-dēn* ‘having the same religion’; *jud-dēn* ‘having a different (the opposite) religion’; Parth. *hām-tōhm-ag* ‘of the same family’, *yud-āwend-ag* ‘without equal’.
- *hu-* ~ *duš-* (*du(s)-* before *s*, *duj-* before voiced stops) ‘having good X’, ‘having bad X’ [or: ‘very bad’ if already bad]: *čašm* ‘eye’ > *hu-čašm* ‘benevolent’ (from *čašm* ‘eye’), *duš-čašm* ‘envious’; *sraw* ‘reputation’ > *hu-sraw* ‘of good fame, famous’, *dusrav* ‘of bad fame, infamous’; *gand* ‘stench’ > Man. *duj-gann* ‘foul’; Parth. *hu-čīhr* ‘pretty’, *dur-čīhr* ‘ugly’.
- *pad-*: forms adjectives expressing possession: *pad-bēš* ‘with hostility, incensed’ (cf. *a-bēš*); *pad-parr-ag* ‘with wings, winged’.

6.1.2 Suffixation

The following are common suffixes used to form adjectives and nouns from verbal stems, nouns, and adjectives.

6.1.2.1 Nouns

The two most common nominalising suffixes are *-išn* (also Parthian), which makes action nouns from present (occasionally past) verbal stems, and *-īh* (Parth. *-īf*, *-īft*), which makes abstract nouns from nouns or adjectives (including compounds). Others:

-stān makes nouns of place: *hindūg* ‘Indian’: *hindū-stān* ‘India’. Other: *dād* ‘law’: *dādestān* ‘law, judgement’; *hērbēd* ‘priest’: *hērbēd-estān*, title of Avestan texts containing rules connected with priestly studies.

6.1.2.2 Action nouns in *-išn*

The action noun in *-išn*, which differs from the passive participle of necessity (see section 4.3.10.1) by syntax and word order, can be made from any present stem and a few past stems:

- from simple present stems: *men-išn*, *gōw-išn*, *kun-išn* ‘thinking, speaking, doing’;
- from derived present stems: (*paydāg-ēn*)-*išn* ‘reaveal-ing’, (*āgāh-īh*)-*išn* ‘being made aware, being informed’, *hammōxt-išn* ‘being taught’;
- from past stems: *āmad-išn* ‘the coming’ (Man.), *zād-išn* ‘being born, birth’ (Dk. 5.2.2).

It frequently forms compound verbs with *kerdan* ‘to do’: *rām-išn ud āsāy-išn kerdan* ‘enjoy-ing and relax-ing make’ = ‘having a good time and relaxing’ (PN. 8).

6.1.2.3 Agent nouns in *-tār/dār*

This agent noun is formed from the past stem: *warz-īd* ‘work, labour, perform’: *warz-īdār* ‘one who performs/does sth., a labourer’; *xwar-d* ‘eat’: *a-(xwar-dār)* ‘not eating, without eating’; *kar-d* ‘do, make’: *margīh-(ker-dār)* ‘death-maker’; *dā-d* ‘give’: (*ātaxš-dādār*)-*tar* ‘more fire-giving’, *ma-d* ‘come’: *madār-tom* ‘who comes most (often)’.

A few forms have passive meaning: *grif-t* ‘seize’: *grif-tār* ‘prisoner’, *frēf-t* ‘deceive’: *frēf-tār* ‘easily deceived’.

6.1.2.4 Abstract nouns in *-īh*

The abstract noun in *-īh* can be made from any nominal:

- nouns: *xwadāy* ‘lord’: *xwadāy-īh* ‘(over)lord-ship, rule’;
- adjectives: *frārōn* ‘morally good’: *frārōn-īh* ‘good-ness’;
- verbal nouns and adjectives: *dān-āg* ‘knowing, wise’: (*dān-āg*)-*īh* ‘wis-dom’;
- agent nouns in *-tār/dār*: *nimū-dār* ‘show-ing’: (*nimū-dār*)-*īh* ‘guid-ance’; *bōx-tār* ‘deliver-er’: (*bōx-tār*)-*īh* ‘deliver-ance’; (*tarāzūg-ēn*)-*īdār* ‘a weigh-er’: (*tarāzūg-ēn-īdār*)-*īh* ‘weigh-ing’.
- compounds (see below): (*rāst-gōw-išn*)-*īh* ‘speaking of truthful speech’, (*dēw-ēzag*)-*īh* ‘(demon-worship)-ing’; Parth. (*abē-āstār*)-*īft* ‘being (free from sins)’.

It is common with action nouns in *-išn*: (*āgāh-īh*)-*išn* ‘being made aware’: (*āgāh-īh-išn*)-*īh* ‘the fact of being made aware’, (*hammōxt-išn*)-*īh* ‘the act of being taught’, (*āmad-išn*)-*īh* ‘the act of coming’ (Man.). In compounds: *a-(bōxt-išn)* ‘having no salvation’: (*a-bōxt-išn*)-*īh* ‘the fact of/state of having no salvation’; *a-(srāy-išn)* ‘(someone) without recit-ing’: *pad (a-srāy-išn)-īh* ‘in a state of being someone who does not recite’ = ‘while not reciting’; *anāg-(rav-išn)* ‘of evil behaviour’: (*anāg-rav-išn*)-*īh* ‘having evil

behaviour'; *gišnag-(zīy-išn)* 'having short living, short-lived': (*gišnag-zīyišn*)-īh 'short-lived-ness'. The nouns in *-išn-īh* are sometimes almost indistinguishable from those in *-išn* (see Josephson 1998).

6.1.2.4a Quotational -īh

In exegetical texts, *-īh* is often used to 'quote' a word to explain it:

mīhr ī frāgōyōd ī hazār-gōš . . .
 'Mihr CNCT "with-wide-grazing-grounds" CNCT thousand-eared'
u-š [(hazār-gōš)-īh] ēd kū
 'and-his "(thousand-ear)-ness" this (COP) that:'
-š 500 mēnōy ō sar nišīn-ēnd,
 'him-IO 500 spirit on head sit-PRES-3p'
u-š kār ī [gōš-īh] kun-ēnd
 and-him-IO action CNCT "ear-dom" do-PRES-3p'
 = '[translation of Avestan:] Mihr with wide grazing grounds, with a thousand ears . . .; and his (being said) to have a thousand ears (means) this that 500 spirits sit on his head and perform the task of being ears for him' (Pahl. Y. 4.8).

6.1.2.5 Nouns and adjectives

-ag: *nām* 'name': *nām-ag* 'written document'; (*sē-šab*)-*ag* 'lasting three nights'; with no discernable function or the unsuffixed word no longer exists: *xān*, *xān-ag* 'house', *kanār-ag* 'limit, border' (cf. *kanārag-ōmand* 'limited').

-gar (-*kar*), *-gār* (-*kār*) 'doer of X, making X': *yōjdahr* 'ritually pure': *yōjdahr-gar* 'who makes ritually pure'; *bazzag* 'bad deeds': *bazzak-kar* 'doer of bad deeds, evil-doer'; *wināh* 'sin, crime' *wināhgār* (*wināhkār*) 'sinful, criminal'.

6.1.2.6 Adjectives

-ak (-*k*) makes diminutives: *and* 'so much': *and-ak* 'a little'; *nēw* 'good': *nēk* 'good'; Man. MPers. *kōdak* and *kōdag* 'small'.

-ān forms adjectives denoting appurtenance, especially filiation: *ardašīr-ān* 'son of Ardashir', *razm ī wištāsp-ān* 'the battle fought by Wištāsp'; the expanded form *-ag-ān* commonly denotes 'son of': *šābuhr-ag-ān* 'son of Šābuhr' (also the name of the book Mani presented to king Šābuhr); *pābagān* is ambiguous, 'son of Pāb/Pābag'.

-āw Parth. suffix denoting language: *frōm-āw* '(in) "Roman" (= Greek)'.
-īg (rarely *-īk*), the most common suffix for forming adjectives from nouns: (*burz-išn*)-*īg* 'praise-worthy'; *nazd* 'near(ness)': *nazd-īk* 'near'.

-ōmand (Man. MPers., Parth. *-āvend*), forms adjectives meaning (things) 'full of . . .': *bīm-ōmand* 'fearful' (cf. *abē-bīm* 'without fear'); *kanārag-ōmand* 'limited' (cf. *a-kanārag* 'limitless'); *padixšar-āvend*, Parth. *padišfar-āvend* 'honourable'.

-gen, similar to *-ōmand*, but less common: *bīm-gen* 'scary'; Parth. *zēn* 'armour, weapons': *zēn-gen* 'weaponed'.

-ēn: adjectives denoting material and similar: *zarr-ēn* 'of gold, golden'; Parth. *ādur-ēn* 'fiery'.

-gōn 'kind, species': colour adjectives: *zarr-gōn* 'golden', and adjectives made from indefinite pronouns and adverbs, e.g. *harw-gōn*, *wisp-gōn* 'of all kinds/colours'.

-*vār*: adverbs of manner from adjectives and nouns: *sazāg* ‘fitting, appropriate’: (*saz-āg*)-*vār* ‘fittingly, appropriately’; *xwaddāy* ‘lord, ruler’: *xwaddāy-vār* ‘in a manner befitting a lord’.

6.2 Compounding

Compounds are formed by combining other words, with or without suffixes. Following are some common types.

6.2.1 Nouns

noun + noun: *kār* ‘work’ + *nāmag* ‘written document’: *kār-nāmag* ‘book of deeds’;
 noun + verbal noun: *uzdēs* ‘idol’ + *parist-* ‘to worship’: *uzdēs-parist* ‘idol-worshipper’ (*uzdēs-paristišnīh* ‘idol-worship’).

6.2.2 Adjectives

adjective + noun: *rāst-gōwišn* ‘of truthful speech’; *was-anāg* ‘who has (commits) much evil’; Parth. *xas-bōy* ‘fragrant’;

noun + noun: *drō-dādīvar* ‘whose judgements are deception (false)’;

noun + verbal noun: *gyān-abespār* ‘surrendering one’s life, reckless, desperate’; *pārag-stanišn* ‘who takes bribes’; *anāg-kerdār* ‘evil-doing, evil-doer’; (*dēv-ēz*)-*ag* ‘who sacrifices to the *dēws*’; Parth. *razm-yōz* ‘battle-stirring’;

noun, adjective + past participle: *yazd-āfrīd* ‘god-created’; *ohrmazd-dād* ‘given, established (created) by Ohrmazd’; *harwisp-pēsīd* ‘adorned with every (adornment)’; Parth. (*šary-zād*)-*ag* ‘lion cub’.

numeral + noun: (*panzdah-sāl*)-*ag* ‘fifteen-year-old’; Man. MPers. *tas-kerb* ‘four-shaped’; Parth. *čuhr-pād* ‘quadruped’.

7 LEXICON AND STYLES

7.1 Dialectology and loans

Many of the ‘Medisms’ of Old Persian have been inherited, e.g. *asp* ‘horse’. The *Psalter* and Manichean Middle Persian sometimes have purer Perside dialect forms than Pahlavi, e.g. Man. MPers. *damestān* ‘winter’ ~ Pahl. *zamestān* <z/dmst ī’>.

There are some loan words:

- from Indic: *semspār* ‘astrologer’ < *sāmvatsara* (MacKenzie 1984);
- from Greek: *stēr* ‘stater’ (via Central Asian *stadēr*); Pahl. *kālbodlkālbed* < *kalapódion* (Bailey 1943: 96);
- from Aramaic: *nakkīrā* ‘repudiating’ (Henning 1946: 732 = 1977, II: [250]).

7.2 Religious terminology

A special problem is presented by determining the meaning of the religious terminology inherited from Avestan in Pahlavi or Pahlavised form, sometimes also in Perside from inherited from Old Persian Examples:

Avestan Ahura Mazdā, the name of the supreme deity, originally meant, literally, ‘omniscient lord’, but the adjective *mazdā* (*maz-da’ ah-* ‘who places [all] in [his] mind’) is not otherwise used in the Avesta, and, already in the Young Avesta and, a fortiori, in the Old Persian inscriptions, where it is unverbated as Ahuramazdā, as well as in Pahlavi Ohrmazd, where the individual parts are no longer recognisable, it was just a name. In the tradition, however, the original epithets had been preserved as *harwisp-āgāh* ‘all-aware’ and *xwadāy* ‘lord’ (Bd. 1.1, 34 etc.).

Avestan *raēuuuant xʷarənanʰaṇt*, approximately ‘wealthy and fortunate’ (blessed by the divine fortune) has the Pahlavi equivalent *rāyōmand xwarrahōmand*; here, for instance, the Zoroastrian exegetes disagreed about the meaning of *rāyōmand*, which was interpreted as *rād* ‘generous’ (S. 1.1), *tuxšāgīh* ‘diligence’ or *abzār-ōmandīh* ‘possession of “means”’ (S. 1.11), while ‘being *xwarrahōmand*’ was equated with *xwēškārīh* ‘performance of one’s duties’ (according to station in life).

Avestan *ašauuan-* clearly referred to being in accordance with the cosmic (true) order (*aša-*) as established by god (Ahura Mazdā); the Pahlavi term *ahlaw*, on the other hand, while referring to being in accordance with the Ohrmazd’s creation and the teachings of Zarathustra, probably had a more general meaning, which we cannot define (‘righteous’ is a stop-gap translation that is often inapplicable). In addition, the Old Persian form *artāvan-*, which characterised those who went to paradise, the ‘blessed’ dead, survives as *ardā* in some phrases, e.g. *ardā zardušt*, *ardā frawahr*, where it is not much more than a standing epithet.

Avestan *daēnā-* denoted a mental faculty, probably that of ‘seeing’ in the beyond; she is also personified as sacrificer’s guide in the ritual rite and as a woman who appears to the departed in the beyond and guides him/her to the Bridge of the Accountant and beyond to paradise or hell. Pahlavi *dēn* preserves this last function in the inscriptions of Kerdīr and in the Pahlavi books, but, in addition, *dēn* also refers to the entire (orally transmitted) religious tradition, as expressed in the term *gōwēd pad dēn* ‘it says in the *dēn*’, and similar phrases. Only in the late Sasanian and Islamic periods does the term probably acquire a meaning comparable to ‘religion’ in the Christian discourse of that time.

Pahlavi *mēnōy* and *gētīy* are nouns and adjectives referring to the other and this world and their inhabitants, respectively. Translations such as ‘spirit, spiritual’ and ‘material’ associate them with modern meanings they probably did not have.

7.3 Sociolinguistic aspects: styles and levels of speech

A wide variety of styles are found in the Middle Persian corpus. Some of these are genuine Middle Persian, while others are clearly influenced by the language they were translated from. Thus, the word order in the *Psalter* follows that of the Syriac original. Some of the Manichean texts may also have been translated or adapted from Aramaic and so preserve some of the originals’ style.

The simplest style is the colloquial seen in narrative texts. It is characterised by short, paratactic sentences and sparse use of anaphora and conjunctions:

pas abar suft mad

‘then (he) bored up (through it and) came (inside)’ (Bd. 7.10).

A typical feature of Middle Persian style is variation:

war ī jam-kard [kū] kard est-ēd?
tan ī sām [kū gyāg] nibay-ēd?
ud mān-išn ī srōš [kū]

‘where is the bunker that Jam made built, in what place lies the body of Sām, and where is the dwelling of Srōš?’ (MX. 61.1).

Close to this style is the didactic prose of texts such as the *Mēnōy xrad*, the *Pand-nāmag*, and the *Bundahišn*. The most complex style is the elaborate prose of the *Dēnkard* and some other texts. The complexity of the prose of the *Dēnkard* is compounded by the fact that it is, to a large extent, based on Pahlavi versions of the *Avesta* and its exegesis.

The elevated polite style is seen in the royal inscriptions (especially the Paikuli inscription) and the *Ayādgār ī Zarērān*. The king uses, is addressed, and referred to in the royal plural and with titles such as *ōy bay* (obl. *anvēšān bay-ān*) ‘His Majesty’, *ašmāh bay-(ān)* ‘Your Majesty’. The verb *framū-dan* ‘order’ is used as a dummy verb or auxiliary for the speech and action of a superior:

ahīy az abārīg xwar-išn framāy-ē xwar-dan
 ‘first from the other food order-PRES-2s eat-INF’
 = ‘first deign eat some of the other food!’ (KN. 93).

8 SAMPLE TEXTS

The following are samples of texts from inscriptional Middle Persian and Parthian, the Middle Persian Psalter and Book Pahlavi, and Manichaean Middle Persian and Parthian, each given in parallel transliteration and transcription, together with inter-linear and idiomatic translation (ERG = ergative; AG = agent; IMP = imperfect; SBJ = subjunctive; OPT = optative).

8.1–2 Middle Persian and Parthian inscriptions of Šābuhr I at Hajiabad

tgʻhy	ZNE	LY	mzdysn	bgy	šhpwhry	(Middle Persian)
<i>tiḡrāh</i>	<i>ēn</i>	<i>man</i>	<i>mazdēs</i>	<i>bay</i>	<i>šābuhr</i>	
wʻwny	ZNE	LY	mzdyzn	ALHA	šhypowhr	(Parthian)
<i>widāwan(i)</i>	<i>im</i>	<i>man</i>	<i>mazdēzn</i>	<i>baγ</i>	<i>šābuhr</i>	
arrow-shot	this	me	Mazdean	god	Šābuhr	

= ‘This arrow-shot (is that) of me, the Mazdean divine Šābuhr,’

MLKAn	MLK A	ʻyrʻn	W	ʻnyrʻn	(Middle Persian)
<i>šāh-ān</i>	<i>šāh</i>	<i>ēr-ān</i>	<i>ud</i>	<i>an-ēr-ān</i>	
MLKYN	MLK A	ʻryʻn	W	ʻnyʻryʻn	(Parthian)
<i>šāh-ān</i>	<i>šāh</i>	<i>ēr-ān</i>	<i>ud</i>	<i>an-ēr-ān</i>	
of-Kings	king	of-Iranians	and	non-Iranians	

= ‘king of kings of Iranians and non-Iranians,’

MNW	ctry	MN	yztʻn	(Middle Persian)
<i>kē</i>	<i>ēihr</i>	<i>az</i>	<i>yazd-ān</i>	
MNW	šyhr	MN	yʻztn	(Parthian)
<i>kē</i>	<i>šīhr</i>	<i>až</i>	<i>yazd-ān</i>	
who	seed	from	gods	

= ‘whose seed is from the gods,’

BRE	mzdysn	bgy	`rthštr	MLKAn	MLKA	'yr'n	(Middle Persian)
<i>pus</i>	<i>mazdēsn</i>	<i>bay</i>	<i>ardaxšahr</i>	<i>šāh-ān</i>	<i>šāh</i>	<i>ēr-ān</i>	
BRY	mzdyzn	ALHA	`rthštr	MLKYN	MLKA	'ry'n	(Parthian)
<i>puhr</i>	<i>mazdēzn</i>	<i>baγ</i>	<i>ardaxšahr</i>	<i>šāh-ān</i>	<i>šāh</i>	<i>ēr-ān</i>	
son	Mazdean	god	Ardaxšahr	of-kings	king	of-Iranians	

= 'whose seed is from the gods, son of the Mazdean divine Ardaxšahr, king of kings of Iranians.'

MNW	ctry	MN	yzt'n	(Middle Persian)
<i>kē</i>	<i>čihr</i>	<i>az</i>	<i>yazd-ān</i>	
MNW	šyhr	MN	y'ztn	(Parthian)
<i>kē</i>	<i>šihr</i>	<i>až</i>	<i>yazd-ān</i>	
who	seed	from	gods	

= 'whose seed is from the gods.'

npy	bgy	p'pk	MLKA	(Middle Persian)
<i>nab</i>	<i>bay</i>	<i>pābag</i>	<i>šāh</i>	
pwhrypwhr	bg	p'pky	MLKA	(Parthian)
<i>puhrīpuhr</i>	<i>baγ</i>	<i>pābag</i>	<i>šāh</i>	
grandson	god	Pābag	king	

= 'the grandson of the divinity Pābag, the king.'

AP	-n	AMT	ZNE	HTYA	ŠDYTN	(Middle Persian)	
<i>u</i>	<i>-n</i>	<i>ka</i>	<i>ēn</i>	<i>tigr</i>	<i>wist</i>		
and	-we-AG	when	this	arrow	shot		
W		AMT	LN	ZNE	HØYA	ŠDYt	(Parthian)
<i>ud</i>		<i>kaδ</i>	<i>amāh</i>	<i>im</i>	<i>tiγr</i>	<i>*wist</i>	
and		when	we-AG	this	arrow	shot	

= 'And when We shot this arrow.'

ADYN	-n	LOYNY	štldl'n	W	BRBYTAn	(Middle Persian)
<i>ēg</i>	<i>-in</i>	<i>pēš</i>	<i>šahr yār-ān</i>	<i>ud</i>	<i>wispuhr-ān</i>	
then-	we-AG	before	rulers	and	princes	
		QDMTE	hštrdryn		BRBYTAn	(Parthian)
		<i>parwān</i>	<i>*šahr dār-īn</i>		<i>wispuhr-ān</i>	
		before	rulers		princes	

W	wcl'k'n	W	'zt'n	ŠDYTN	(Middle Persian)
<i>ud</i>	<i>wuzurg-ān</i>	<i>ud</i>	<i>āzād-ān</i>	<i>wist</i>	
	RBAAn	W	'z'tn	ŠDYt	(Parthian)
	<i>*wazurg-ān</i>	<i>ud</i>	<i>āzād-ān</i>	<i>*wist</i>	
and	great men	and	nobles	shot	

= 'then We shot it before the rulers, princes, great men, and nobles.'

AP-	n	LGLE	PWN	ZNE	drky	HNHTWN	(Middle Persian)
<i>u</i>	<i>-n</i>	<i>pāy</i>	<i>pad</i>	<i>ēn</i>	<i>darrag</i>	<i>nihād</i>	
		NGRYN	pty	ZK	wym	HQAYMWt	(Parthian)
		<i>pāδ</i>	<i>pad</i>	<i>ēd</i>	<i>wēm</i>	<i>awestād</i>	
and-	we-AG	foot	on	this	cleft	placed	

= 'And We placed (Our) foot/feet on this cleft (Parth. 'stone').'

AP-	n	HTYA	LCDr	ZK	cyt'k	BLA	LMYTN	(Middle Persian)
<i>u</i>	<i>-n</i>	<i>tigr</i>	<i>tar</i>	<i>ān</i>	<i>čīdāg</i>	<i>bē</i>	<i>abgand</i>	
W		HØYA	LCD	LHw	šyty	LBRA	RMYt	(Parthian)
<i>ud</i>		<i>tiγr</i>	<i>tar</i>	<i>hō</i>	<i>šīd</i>	<i>bēh</i>	<i>*abgand</i>	
and-Us-AG		arrow	past	that	cairn	away	threw	

= 'and We threw (= shot) the arrow beyond that cairn.'

BLA	OLE	wy'k	AYK	HTYA	LMYTN	(Middle Persian)	
<i>bē</i>	<i>ōy</i>	<i>gyāg</i>	<i>kū</i>	<i>tigr</i>	<i>abgand</i>		
byš	TME		ANW	HŌYA	NPLt	(Parthian)	
<i>bēž</i>	<i>ōδ</i>		<i>kū</i>	<i>ti;r</i>	<i>kaft</i>		
but	that	place	where	arrow	fell		
= 'But, the place (Parth. 'there') where the arrow was thrown (Parth. 'fell'),'							
TME	wy'k	ZK-`wgwn	LA	YHWWN	AYK	(Middle Persian)	
<i>ānōy</i>	<i>gyāg</i>	<i>ān-ōwōn</i>	<i>nē</i>	<i>būd</i>	<i>kū</i>		
there	place	that-manner	not	was	that		
	LHw-znk	ATRE	LA	YHWt	AYK	(Parthian)	
	<i>hō-zanag</i>	<i>wyāg</i>	<i>nē</i>	<i>būd</i>	<i>kū</i>		
	that-manner	place	not	was	that		
= 'there the place was not such that,'							
HT	cyt'k	cyty	HWE			(Middle Persian)	
<i>agar</i>	<i>čīdag</i>	<i>čīd</i>	<i>h-ē</i>				
'k	šyty	BNYt	HWYndy			(Parthian)	
<i>ag</i>	<i>šīd</i>	<i>*dišt</i>	<i>ah-ēndē</i>				
if	cairn	erected	were-OPT				
= 'if a cairn had been raised'							
ADYN	by-lwny		pty'k	YHWWN	HWE	(Middle Persian)	
<i>ēg</i>	<i>bē-rōn</i>		<i>paydāg</i>	<i>būd</i>	<i>h-ē</i>		
	OL LBRA ŠŌRA		'ksy	YHWt	HWYndy	(Parthian)	
	<i>ō bēh ārag</i>		<i>āgas(ī)</i>	<i>būd</i>	<i>ah-ēndē</i>		
then	to out-side		visible	would-be-OPT			
= 'then it would have been visible outside.'							
AHR	LINE	prm't	MNW	cyt'k	`wldly	cyty	(Middle Persian)
<i>pas</i>	<i>amāh</i>	<i>framād</i>	<i>kē</i>	<i>čīdāg</i>	<i>ōr-andar</i>	<i>čīd</i>	
'dyn	LN	'wpdšt	MNW	šyty	PNestr	BNYt	(Parthian)
<i>ađyān</i>	<i>amāh</i>	<i>ubdišt</i>	<i>kē</i>	<i>čīd</i>	<i>ōr-istar</i>	<i>dišt</i>	
then	We-AG	ordered-ERG	who	cairn	more-hither	erected	
= 'Then, We commanded (someone) who raised the cairn *on this side (Parth. 'further in this direction').'							
	MNW	YDE	ΘB	HWE	ZK	LGLE	(Middle Persian)
	<i>kē</i>	<i>dast</i>	<i>nēw</i>	<i>h-ād</i>	<i>ān</i>	<i>pāy</i>	
'ws	MNW	YDA	ΘB	HWY-nt		NGRYN	(Parthian)
<i>awās</i>	<i>kē</i>	<i>dast</i>	<i>nēw</i>	<i>ah-ānd</i>		<i>pāδ</i>	
'(now)	who	hand(s)	good	be-SBJ-3s/p	that	foot'	
PWN	ZNE	drky	'yw	HNHTWN		(Middle Persian)	
<i>pad</i>	<i>ēn</i>	<i>darrag</i>	<i>ēw</i>	<i>nih-ēd</i>			
pty	ZNE	wym	hyp	HQAYMWd		(Parthian)	
<i>pad</i>	<i>im</i>	<i>wēm</i>	<i>hēb</i>	<i>*awest-ēδ</i>			
= '(Now) he whose hand is (Parth. 'hands are') good, let that one place his foot on his cleft (Parth. 'stone')							
W	HTYA	OL	ZK	cyt'k	'yw	ŠDYTN	(Middle Persian)
<i>ud</i>	<i>tigr</i>	<i>ō</i>	<i>ān</i>	<i>čīdāg</i>	<i>ēw</i>	<i>wih-ēd</i>	
W	HŌTYA	OL	hw	šyty	hyp	ŠDYW	(Parthian)
<i>ud</i>	<i>ti;r</i>	<i>ō</i>	<i>hō</i>	<i>šīd</i>	<i>hēb</i>	<i>*wih-ēδ</i>	
'and	arrow	to	that	cairn	EXH	shoot'	
= 'and let him shoot the arrow to that cairn!'							

AHR	MNW	HTYA	OL	ZK	cyt'k	LMYTN	(Middle Persian)
<i>pas</i>	<i>kē</i>	<i>tigr</i>	<i>ō</i>	<i>ān</i>	<i>čīdāg</i>	<i>abgan-ēd</i>	
	MNW	HΘYA	OL	hw	šyty	z'mywd	(Parthian)
	<i>kē</i>	<i>tiγr</i>	<i>ō</i>	<i>hō</i>	<i>šīd</i>	<i>žām-ēd</i>	
	'then	who	arrow	to	that	cairn	throws (Parth. 'leads, makes reach')
= 'Then, he who shoots the arrow to (Parth. 'makes the arrow reach') that cairn.'							

OLE		YDE	ΘB		(Middle Persian)
<i>ōy</i>		<i>dast</i>	<i>nēw</i>		
LHw	-p	YDA	ΘB	HWYN	(Parthian)
<i>hō</i>	<i>-b</i>	<i>dast</i>	<i>nēw</i>	<i>ah-ēnd</i>	
	'he	indeed	hand(s)	good	(are)
= 'his hand is (Parth. 'hands are') good (indeed)!'					

8.3 Pahlavi Psalter, from *Psalm 127*

NYŠEt	cygwn	st'hy	ZY	st'dytky	PWN	d'mny	ZY	BYTA.
<i>zan-it</i>	<i>čiyōn</i>	<i>stāx</i>	<i>ī</i>	<i>stāyīd-ag</i>	<i>pad</i>	<i>dāman</i>	<i>ī</i>	<i>xān.</i>
wife-your	like	vine		praised	at	bosom	of	house
= 'Your wife is like a praised vine in the bosom of the house.'								

plndt	cygwn	nm'mk[y]	ZY	zyty	pyl'mwn	-t	hw'ny.
<i>frazend-it</i>	<i>čiyōn</i>	<i>nimāmag</i>	<i>ī</i>	<i>zayt</i>	<i>pērāmōn</i>	<i>ī-t</i>	<i>xwān.</i>
child-your	like	tree	of	olive	around	of-your	table.
= '(And) your children (are) like a tree of olive oil(?) about your table.'							

LHNAš	'wgnw	'plyny	klyty	GBRA	MNW	tlsyt	MN	MROHY
<i>*ēd-iš</i>	<i>ōwōn</i>	<i>āfrīn</i>	<i>kir-ēd</i>	<i>mard</i>	<i>kē</i>	<i>tars-ēd</i>	<i>az</i>	<i>xwadāy.</i>
this-him	thus	blessing	is made	man	who	fears	from	lord
= 'In this way he is blessed, the man who fears the Lord.'								

'plynt	OBYDWN't	MROHY	MN	chywny
<i>āfrīn-it</i>	<i>kun-ād</i>	<i>xwadāy</i>	<i>az</i>	<i>čihyōn (šihyōn)</i>
blessing-you	will-make	lord	from	Zion
= 'The Lord shall bless you from Zion.'				

W	wyn'y	PWN	klpky	ZY	'wlšlmy
<i>ud</i>	<i>wēn-ā</i>	<i>pad</i>	<i>kerbag</i>	<i>ī</i>	<i>ūrišalm</i>
and	see-SBJ-2s	by	grace	of	Jerusalem
= 'And you shall see by the grace of Jerusalem.'					

KRAwspy	YWM	ZYt	zywndkyhy
<i>harwisp</i>	<i>rōz</i>	<i>ī-t</i>	<i>zīndagīh</i>
all	day	of-your	life
= 'all the days of your life.'			

W	wyn'y	plndy	OLt	plndwn
<i>ud</i>	<i>wēn-ā</i>	<i>frazend</i>	<i>ō-t</i>	<i>frazend-ūn.</i>
and	you-see-SBJ	child	to-your	children
= 'And you shall see children unto your children.'				

8.4 Pahlavi *Bundahišn* (Bd.33.0–4)

abar wizenđ ī hazārag hazārag ō ērān-šahr mad.
 on harm that millennium millennium to Ērān-šahr came
 = ‘On the harm that befell the land of the Iranians in each millennium.’

ka gannāg mēnōy andar dwār-ist pad bun,
 when Foul Spirit in rushed in in beginning
 = ‘When the Foul Spirit rushed in in the beginning.’

nazdist hazārag ī andar gumēz-išn-īh
 first millennium in mixture
 = ‘the first millennium in the mixture.’

gāw ud gayōmard būd h-ēnd.
 Bull and Gayōmard were
 = ‘there were the Bull and Gayōmard.’

ka mašy ud mašyānī ān huspāsth kerd 50 sāl
 when Mašy and Mašyānī-AG that good-worship made 50 year
 = ‘When Mašy and Mašyānī had performed that good worship/service for 50 years.’

ēg-išān zāy-išn aziš nē būd.
 then-them birth from not was
 = ‘then there was no birth from them (-išān . . . aziš).’

andar ham hazārag pad 70 sāl
 in same millennium for 70 year
 = ‘In the same millennium, for 70 years.’

hōšang ud tahmūraf harw dō dēw-ān bē ōzad.
 Hōšang and Tahmūraf each two demons killed
 = ‘Hōšang and Tahmūraf both killed demons.’

andar hazārag sar dēw-ān jam kīrren-īd.
 in millennium head/end demons-AG Jam cut
 = ‘At the end of the millennium the demons cut Jam (in half).’

didgar hazārag bun būd
 second millennium bottom/beginning was
 = ‘At the beginning of the second millennium was.’

ažidahāg duš-xwadāy-īh abar kerd. 100 sāl bē kerd.
 Ažidahāg evil-rule began 100 year did
 = ‘when Ažidahāg began (his) evil-rule. He did (so) for a hundred years.’

čyōn hazārag sar būd frēdōn grift bast.
 as millennium end became Frēdōn-AG took bound
 = ‘When the end of the millennium came, Frēdōn seized and bound (him).’

sidgar hazārag bun būd
 third millennium beginning was
 = ‘It was the beginning of the third millennium’

ka frēdōn kišwar baxt salm ud tūz.
 when Frēdōn land divided Salm and Tūz
 = ‘when Salm and Tūz divided Frēdōn’s land (between them).’

ēg-išān *ēriĵ* *ōzad* *frazend-ān* *ud* *āwādag-ān* *abesih-ēn-īd.*
 then-they-AG Ēriĵ killed children and families destroyed
 = ‘Then they killed Ēriĵ and destroyed (his) children and family.’

andar *ham* *hazārag* *mānuščihr* *zād*
 in same millennium Mānuščihr was-born
 = ‘In the same millennium, Mānuščihr was born’

ud *kēn* *ī* *ēriĵ* *xwāst.*
 and vengeance of Ēriĵ sought
 = ‘and sought (exact)ed vengeance for Ēriĵ.’

8.5 Pahlavi *Dēnkard* (Dk.9.45.4)

(DkM II, pp. 870–71; DkDH, p. 235)

abar *stāy-išn* *ī* *pōručist* *ī* *zarduxšt-ān*
 on praise of Pōručist Zarathustrid
 = ‘On the praise of Pōručist, daughter of Zarathustra’

pad *dōš-īdan* *ī* *weh-dēn* *pad* *xrad*
 for loving of Good Dēn with wisdom
 = ‘for loving the Good Dēn with wisdom’

ud *ker-dan* *ī* *ān* *ī* *dēn* *handarz* *hunsand-īhā.*
 and making of that of dēn advice happily
 = ‘and for happily performing the advice of the Dēn.’

dā-dan *ī* *tan* *pad* *zan-īh* *ō* *zarduxšt*
 giving of body in matrimony to Zarathustra
 = ‘How she gave her body in wife-hood to Zarathustra’

ud *spurrīg* *warz-īdan* *ī* *kār* *ud* *tarsagāh-īh* *ī* *ōy.*
 and fully performing of work and respect of him
 = ‘and fully performed “work” and respect for him.’

pas *az* *zarduxšt* *zan-īh-iz* *ud* *tarsagāh-īh* *ī* *ĵāmāsp* *ker-dan*
 after Zarathustra wife-hood-also and respect of Jāmāsp mak-ing
 = ‘After Zarathustra, how she also performed wife-hood and respect for Jāmāsp.’

ud *(wazurg-mīzd)-īh* *ī-š* *az* *ohrmazd*
 and (great-reward)-ing of-her from Ohrmazd
 = ‘And the fact that a great reward (came) to her from Ohrmazd’

pad *ān* *dēnīg-īh* *ud* *yazd-ān* *xwēš-īh.*
 for that dēn-following and gods’ own-being
 = ‘for following the Dēn and being the gods’ own.’

8.6 Manichaean Middle Persian

(*Mir. Man.* i, text e, M7984 I R ii, line 34 v ii, line 1)

gahīy hān wjftōb āz garān xešm pahikand,
 then that lusty Āz heavy rage was-filled
 = ‘Then that *lusty Āz was filled with heavy rage,’

u-š niwist gām xwāh-ān
 and-it-AG began step seeking
 = ‘and it began to seek (her next) step.’

u-š framād kū pad hān dō dēs
 and-she- ordered that: by that two forms
 AG
 = ‘And she ordered: ‘According to those two forms,’

ī srīgar ud nar ī narēsah-yazd ī-m dād
 female and male of Narēsah-god that-I-AG saw
 = ‘male and female, of god Narēsah, which I saw,’

hān dō dām nar ud māyag dēs-ān
 that two creatures male and female I-shall-form
 = ‘I will form those two creatures, male and female,’

kū-m paymōg ud niyām baw-ānd . . .
 that-me garment and covering be-SBJ-3p
 = ‘(in order) that they may be a garment and covering for me.’

pas hān āz az hamāg hān zahag ī dēw-ān
 then that Āz from all that brood of demons
 = ‘Then that Āz, of all that brood of the demons’

kē az asmān ō zamīg kaft h-ēnd
 who from sky to earth fell
 = ‘that fell from the heaven upon the earth,’

hān āsarēštār ī nar āsarēštār ī māyag,
 that archon male archon female
 = ‘that male Archon (and) that female Archon,’

šagr-kerb-and axšōz ud xešm-ēn bazzag ud appar anā-nd
 lion-shaped *lusty and of-wrathful sins and thievish was-IPF
 = ‘(who) were lion-shaped, *lusty, of wrathful sins, and thievish,’

awēšān paymōxt.
 those donned
 = ‘those it donned.’

u-š xwēš niyām ud paymōg kerd h-ēnd
 and-it-AG own covering and garment made-them
 = ‘And it made them its own covering and garment.’

u-šān andar axšōz-ih-ist
 and-them inside lusted
 = ‘And she lusted inside them.’

8.7 Manichaean Parthian poetic text

(*Mir. Man.* iii, text k, M104 R lines 1–17; Boyce 1975: 127, text bx)

warm ādur-ēn frasis-tan kām-ād
cloud fiery breaking it-wished
= ‘It wished for the fiery cloud to break forth’

kū hamag pad ādur hāw-āh.
that all by fire consume-SBJ-3s
= ‘(in order) that it might consume everything by (its) fire.’

šahrδār arγāw "xēbēh padmōžan izwašt
prince noble own garment changed
= ‘The noble prince changed his garment’

ud pad sezδ-īft būd āgas ō sātān.
and in power was revealed to Satan
= ‘and was revealed to Satan in his power.’

hampad zamīg ud asmān wilarz-ād,
*immediately earth and sky trembled
= ‘Immediately, the earth and heaven trembled.’

ud samēl kaft ō žafr-ān.
and Samēl fell to depths
= ‘and Sammēl fell down into the depths.’

tarkumān rāšt būd abaxšāh-išn rōšn wasnād
interpreter true was forgiveness light for the sake of
= ‘The true interpreter had pity for the sake of the light’

čē dušmen-īn xāz-ād.
which enemies swallowed
= ‘that the enemies (had) swallowed.’

būd-iš sān-ād až žafr ahrīwar
was-he-AG raised from deep pool
= ‘He had raised-(it) from the deep dark pool’

ō hō ayāg abrang-īft až kū ōsaxt.
to that place puri-ty from where came-down
= ‘to that place of purity from where it had come down.’

namāž ō tō (wazarg-īft)-(zād-ag)
Homage to you (great-ness)-child
= ‘Homage to you, O child of greatness,’

kē abdāž-ād "xēbēh ardāw-īft.
who delivered own righteous-dom
= ‘who delivered your congregation of the righteous!’

awās-iž pāy āmōž-ag marī zakkō
now-also stay teach-er Mārī Zakō
= ‘So stay now, (our) teacher Mari Zakō,’

pān-ag wazarg čē tō čar-ag nisāg-ēn
protection great of your flock shining
= ‘(as) a great protection of your shining flock!’

ABBREVIATIONS OF TEXTS

Abnun	Inscription of Abnun, ed. Skjærvø (1992).
ANRm	Inscription of Ardaxšīr at Naqš-e Rostam, ed. Herzfeld (1924: 84–85), Back (1978: 281–82)
AR.	<i>Angad rōšnān</i> , ed. Boyce (1954).
AWN.	<i>Ardā Wirāz-nāmag</i> , ed. Haug and West (1872), Gignoux (1984).
AZ.	<i>Ayādgār ī Zarērān</i> , ed. Monchi-Zadeh (1981).
BBB.	Henning (1936).
Bd.	<i>Bundahišn</i> , ed. Anklesaria (1956).
BT 4, 11, 19	Sundermann (1973, 1981, 1997).
Bīš.	Inscription of Šābuhr at Bīšāpūr, ed. in Nyberg (1964: 124–25), Back (1978: 378–83).
Dk.7	<i>Dēnkard</i> , book 7, ed. Molé (1967).
Huy.	<i>Huyadag-mān</i> , ed. Boyce (1954).
KKZ	Inscription of Kerdīr on the Ka‘ba-ye Zardošt, ed. Back (1978: 384–489)
KN.	<i>Kār-nāmag ī ardaxšīr ī Pābagān</i> , ed. Sanjana (1896).
KSM	Inscription of Kerdīr at Sar-e Mašhad, ed. MacKenzie (1989).
M	Manichean manuscript in the Turfan collection of the Berlin-Brandenburgische Akademie der Wissenschaften.
MHD.	<i>Mādayān ī hazār dādestān</i> , ed. Macuch (1993).
<i>Mir. Man.</i> i–iii	The texts edited in Andreas (1932–1934).
NPi	Inscription of Narseh at Paikuli, ed. Herzfeld (1924), Humbach, H., and P. O. Skjærvø (1983).
Pahl. Y.	Pahlavi <i>yasna</i> , ed. Dhabhar (1949).
Pahl. Yt.	Pahlavi <i>yašts</i> , ed. Dhabhar (1927).
PN.	<i>Pand-nāmag</i> , ed. Kanga (1960).
PR.	<i>Pahlavi rivāyat</i> , ed. Williams (1990).
Ps.	The Pahlavi <i>Psalter</i> , ed. Andreas and Barr (1936).
PT.	<i>Pahlavi Texts</i> , ed. J. M. Jamasp-asana (1897).
Šāb. esch.	<i>Šābuhragān</i> , the eschatological part, ed. MacKenzie (1979–80).
ŠH	Inscription of Šābuhr at Haḡjabad, ed. MacKenzie (1978).
ŠKZ	Inscription of Šābuhr on the Ka‘ba-ye Zardošt, ed. Huyse (1999).
ŠPs. II	Inscription of Šāpūr Sagān-šāh at Persepolis, ed. Back (1978: 495–97).

LINGUISTIC DESCRIPTIONS

The most recent descriptions of phonology, morphology (and some syntax) are Rastorgueva (1966 and 1981) and Sundermann (1989), with extensive bibliographies. See also Klingenschmitt's (2000) miscellany. Older works include Haug (1870), E. W. West in Geiger and E. Kuhn (1895: 249–332); Bartholomae (1911–15 and 1916–22). For the phonology, see Weber (1997) with bibliography. For the morphology of the verb, see Henning (1933) for Middle Persian and Ghilain (1939) for Parthian. For the syntax see Brunner (1977). Introductions to Pahlavi: Nyberg's (1928–31) *Hilfsbuch* (see Henning's [1935] review) and the revised (1964–74) *Manual*. Hansen (1963) contains select texts as exercises in reading Pahlavi. Henning (1958) is a survey of the place of Middle Persian and Parthian among the Middle Iranian languages. Many of the Manichean texts

published before Sundermann (1973) were published in transliteration by Boyce (1975) accompanied by a glossary (1977). Dictionaries: MacKenzie (1971: Pahlavi) and Durkin-Meisterernst (2004: Manichean Middle Persian and Parthian).

BIBLIOGRAPHY AND REFERENCES

- Andreas, F. C. (1933) 'Bruchstücke einer Pehlevi-Übersetzung der Psalmen', ed. K. Barr, in *SbPAW*, 1933, pp. 91–152.
- (1932) 'Mitteliranische Manichaica aus Chinesisch-Turkistan. I', ed. W. B. Henning, in *SbPAW*, 1932, pp. 173–222 [= Henning (1977) I, pp. 1–48].
- (1933) 'Mitteliranische Manichaica aus Chinesisch-Turkistan. II', ed. W. B. Henning, in *SbPAW*, 1933, pp. 294–363 [= Henning (1977) I, pp. 191–260].
- (1934) 'Mitteliranische Manichaica aus Chinesisch-Turkistan. III', ed. W. B. Henning, in *SbPAW*, 1934, pp. 846–912 [= Henning (1977) I, pp. 275–340].
- Anklesaria, B. T. (1956) *Zand-Akāsīh: Iranian or Greater Bandahišn*. Bombay: Rahnumae Mazdayasnan Sabha.
- Back, M. (1978) *Die sassanidischen Staatsinschriften: Studien zur Orthographie und Phonologie des Mittelpersischen der Inschriften zusammen mit einem etymologischen Index des mittelpersischen Wortgutes und einem Textcorpus der behandelten Inschriften*. Tehran and Leiden: distrib. Brill.
- Bailey, H. W. (1943) *Zoroastrian Problems in the Ninth-Century Books*. Oxford: Clarendon Press; 2nd ed. with new intro., 1971.
- Barr, K. (1936) 'Remarks on the Pahl. Ligatures 𐭪 and 𐭫', *BSO(A)S* 8, pp. 391–403.
- Bartholomae, C. (1911–15) 'Mitteliranische Studien I', *WZKM* 25, 1911, pp. 245–62; II: pp. 389–409; III: 27, 1913, pp. 19–24, IV: pp. 347–74; V: 29, 1915, pp. 1–47.
- (1916–22) *Zur Kenntniss der mitteliranischen Mundarten I*, *SHbAW, phil.-hist. Kl.*, 1916, no. 9; II, 1917, no. 11; III, 1920, no. 2; IV, 1922, no. 6; V, 1923, no. 3.
- Boyce, M. (1954) *The Manichaean Hymn-Cycles in Parthian*. London, New York: Oxford University Press.
- (1964a) 'Some MPers. and Parth. Constructions with Governed Pronouns', in *Dr. J. M. Unvala Memorial Volume*. Bombay: Kaikhusroo M. JamaspAsa for Dr. J. M. Unvala Memorial Volume Sub-Committee, pp. 48–56.
- (1964b) 'The Use of Relative Particles in Western Middle Iranian', in *Indo-Iranicā. Mélanges présentés à Georg Morgenstierne à l'occasion de son soixante-dixième anniversaire*. Wiesbaden: O. Harrassowitz, pp. 28–47.
- (1968) 'Middle Persian Literature', in *Handbuch der Orientalistik* IV, 2, 1. Leiden and Cologne: Brill, pp. 31–66.
- (1975) *A Reader in Manichaean Middle Persian and Parthian*. Tehran and Liège: Brill (Acta Iranica 9).
- (1977) *A Word-List of Manichaean Middle Persian and Parthian*. Tehran and Liège: Brill (Acta Iranica 9a).
- Brunner, C. J. (1973) 'Middle Persian Inscriptions on Sasanian Silverware', *Metropolitan Museum Journal* 9, pp. 109–21.
- (1977) *A Syntax of Western Middle Iranian*. Delmar, New York: Caravan Books.
- Cantera Glera, N. A. (1999) 'Die Stellung der Pahlavi-Übersetzung des Avesta innerhalb des Mittelpersischen', *StIr* 28, pp. 173–204.
- (2004) *Studien zur Pahlavi-Übersetzung des Avesta*. Wiesbaden: Harrassowitz.
- Dhabhar, B. N. (1927) *Zand-i Khūrtak Avistāk*. Bombay: Trustees of the Parsi Punchayet Funds and Properties.
- (1949) (ed.) *Pahlavi Yasna and Vispered*. Bombay: S. F. Desai for the Trustees of the Parsi Punchayet Funds and Properties.

- (1955) *Essays on Iranian Subjects*. Bombay: Society for the Promotion of Researches in the Zoroastrian Religion.
- Durkin-Meisterernst, D. (2000a) 'Erfand Mani die manichäische Schrift?' in R. E. Emmerick, W. Sundermann, and P. Ziehme (eds) *Studia Manichaica. IV. Internationaler Kongreß zum Manichaismus, Berlin, 14.–18. Juli 1997*. Berlin: Akademie Verlag, pp. 161–78.
- (2000b) 'Zum parthischen Verbum', in B. Forssman and R. Plath (eds) *Indoarisch, Iranisch und die Indogermanistik. Arbeitstagung der Indogermanischen Gesellschaft vom 2. bis 5. Oktober 1997 in Erlangen*. Wiesbaden: L. Reichert, pp. 75–88.
- (2002) 'The uses of *baw-* in Parthian', in Ph. Huyse (ed.) *Iran. Questions et connaissances. Actes du IV^e Congrès européen des études iraniennes organisé par la Societas Iranologica Europaea Paris, 6–10 Septembre 1999*. Vol. 1: *La période ancienne*. Paris: Association pour l'avancement des études iraniennes, pp. 47–62.
- (2004) *Directory of Manichaean Middle Persian and Parthian*, Part I of *Dictionary of Manichaean Texts III: Texts from Central Asia and Chaina*, ed., N. Sims-Williams (Corpus Fontium Manichaeorum. Subsidia, 2). Turnhout: Peeters.
- Geiger, W., and E. Kuhn (1895–1901) (eds) *Grundriss der Iranischen Philologie*, 2 vols. Strassburg: K. J. Trübner; repr. Berlin and New York: W. de Gruyter, 1974.
- Gershevitch, I. (1970) 'The Crushing of the Third Singular Present', in M. Boyce and I. Gershevitch (eds) *W. B. Henning Memorial Volume*. London: Lund Humphries, pp. 161–74.
- Ghilain, A. (1939) *Essai sur la langue parthe* (Bibliothèque du Muséon, vol. 9) Louvain: Institut orientaliste and Publications universitaires.
- Gignoux, Ph. (1978) *Glossaire des inscriptions pehlevies et parthes* (Corpus Inscriptionum Iranicarum, Supplementary Series, vol. 1) London: Corpus Inscriptionum Iranicarum.
- (1984) *Le livre d'Ardā Vīrāz: Translittération, transcription et traduction du texte pehlevi*. Paris: Editions Recherche sur les civilisations.
- (1991) 'Une nouvelle collection de documents en pehlevi cursif du début du septième siècle de notre ère', *Académie des inscriptions et belles-lettres. Comptes rendus*, pp. 683–700.
- (1995) 'The Pahlavi Inscription on Mount Thomas Cross (South India)', in Ziony Zevit et al. (eds), *Solving Riddles and Untying Knots. Biblical, Epigraphic, and Semitic Studies in Honor of Jonas C. Greenfield*, Winona Lake, Ind.: Eisenbrauns, pp. 411–21.
- Hansen, O. (1963) *Mittelpersisches Lesebuch*. Berlin: W. de Gruyter.
- Haug, M. (1870) *Essay on the Pahlavi Language*. Stuttgart: Guttenberg, K. Hofbuchdruckerei.
- Haug, M., and E. W. West (1870) *The Book of Arda Viraf*. Bombay: Govt. Central Book Depot.
- Henning, W. B. (1933) 'Das Verbum des Mittelpersischen der Turfanfragmente', *ZII* 9, pp. 158–253 (= *Selected Papers*, vol. 1, pp. 65–160).
- (1935) Review of Nyberg, 1928–31, in *GGA*, 1935, no. 1, pp. 1–19.
- (1936) *Ein manichäisches Bet- und Beichtbuch*. APAW, 1936.
- (1946) 'The Sogdian Texts of Paris', *BSOAS* 11, pp. 713–40.
- (1977) *Selected Papers*, 2 vols. Tehran and Liège: Brill.
- (1958) 'Mitteliranisch', in *Handbuch der Orientalistik* I, IV, 1. Leiden and Köln: Brill.
- Herzfeld, E. (1924) *Paikuli: Monument and Inscription of the Early History of the Sasanian Empire*. Berlin: Dietrich Reimer.
- Humbach, H., and P. O. Skjærvø (1983) *The Sassanian Inscription of Paikuli*, pt. 3.1: *Restored Text and Translation*; pt. 3.2: *Commentary*. Weisbaden: Reichert.

- Humbach, H., and Wāng Shípíng (1988) 'Die pahlavi-chinesische Bilingue von Xi'an', in *Barg-e sabz. A Green Leaf. Papers in Honour of J. P. Asmussen* (Acta Iranica 28). Leiden: distrib. Brill, pp. 73–82.
- Hutter, M. (1992) *Manis Kosmogonische Šābuhragān-Texte: Edition, Kommentar und literaturgeschichtliche Einordnung der manichäisch-mittelpersischen Handschriften M 98/99 I und M 7980–7984*. Wiesbaden: Harrassowitz.
- Huyse, Ph. (1999) *Die dreisprachige Inschrift Šābuhrs I. an der Ka'ba-i Zardušt (ŠKZ)*, 2 vols. (= Corpus Inscriptionum Iranicarum, part 3, vol. 1, Texts 1). London: School of Oriental and African Studies.
- (2003) *Le y final dans les inscriptions moyen-perses et la 'loi rythmique' proto-moyen-perses* (Studia Iranica. Cahier 29). Paris: Association pour l'avancement des études iraniennes.
- Jamasp-asana, J. M., ed. (1897) *Pahlavi Texts*, Bombay.
- Josephson, J. (1998) 'The Middle Persian Nouns in *-išn* and *-išnīh*', *OS* 47, pp. 87–104.
- Kanga, M. F. (1960) (ed.) *Ātak Handarz i Pōryōtkēšān. A Pahlavi Text*. Bombay: M. F. Kanga.
- Klingenschmitt, G. (2000) 'Mittelpersisch', in B. Forssman and R. Plath (eds) *Indoarisch, Iranisch und die Indogermanistik. Arbeitstagung der Indogermanischen Gesellschaft vom 2. bis 5. Oktober 1997 in Erlangen*. Wiesbaden: L. Reichert, pp. 191–229.
- Lazard, G. (1984) 'Les modes de la virtualité en moyen-iranien occidental', in W. Skalmowski and A. van Tongerloo (eds) *Middle Iranian Studies. Proceedings of the International Symposium Organized by the Katholieke Universiteit Leuven from the 17th to the 20th of May 1982*. Louvain: Peeters, pp. 1–13.
- MacKenzie, D. N. (1964) 'The 'Indirect Affectee' in Pahlavi', in *Dr. J. M. Unvala Memorial Volume*. Bombay: Kaikhusroo M. JamaspAsa for Dr. J. M. Unvala Memorial Volume Sub-Committee, pp. 45–48.
- (1967) 'Notes on the Transcription of Pahlavi', *BSOAS*, pp. 17–29.
- (1971) *A Concise Pahlavi Dictionary*. London and New York: Oxford University Press.
- (1982) Review of Back, 1978, in *IF* 87, pp. 280–97.
- (1984) 'Some Pahlavi Plums', in *Orientalia Jacques Duchesne-Guillemin Emerito Oblata* (Acta Iranica 23). Leiden: Brill, pp. 383–91.
- Macuch, M. (1993) *Rechtskasuistik und Gerichtspraxis zu Beginn des siebenten Jahrhunderts in Iran. Die Rechtssammlung des Farroḥmard i Wahrāmān* (Iranica, 1). Wiesbaden: Harrassowitz.
- Molé, M. (1967) *La légende de Zoroastre selon les textes pehlevi* (Travaux de l'Institut d'études iraniennes de l'Université de Paris, 3). Paris: Klincksieck.
- Monchi-Zadeh, D. (1981) *Die Geschichte Zarēr's* (Acta Universitatis Upsaliensis. Studia Indoeuropea Upsaliensia, 4). Uppsala, Stockholm: distrib. Almqvist & Wiksell International.
- Nyberg, H. S. (1932) 'Un pseudo-verbe iranien et son équivalent grec', in A. Nelson (ed.) *Symbolae Philologicae O. A. Danielsson Octogenario Dicatae*. Uppsala: A.-B. Lundequistska bokhandeln.
- (1928–31) *Hilfsbuch des Pehlevi*, 2 vols. Uppsala: Almqvist & Wiksell.
- (1964, 1974) *A Manual of Pahlavi*, 2 vols. Wiesbaden: Harrassowitz.
- (1988) *Frahang i Pahlavīk*, B. Utas and C. Toll (ed.). Wiesbaden:
- Rastorgueva, V. S. (1966) *Sredne-Persidskii iazyk*. Moskva: Nauka.
- Rastorgueva V. S., et al. (1981) (ed) *Osnovy iranskogo iazykoznaniiia. Sredneiranskie Iazyki*. Moscow: Nauka: 'Srednepersidskij jazyk' pp. 6–145, 'Parfjanskij jazyk', pp. 147–232.
- Salemann, E. (1895) 'Mittelpersisch', in *Grundriß*, vol. I, pp. 249–332.
- Sanjana, D. P. (1896) (ed.) *The Kār-nāmē ī Artakshūr ī Pāpakān*. Bombay: Printed at the Education Society's Steam Press.

- Shaked, S. (1979) *The Wisdom of The Sasanian Sages (Dēnkard VI) by Aturpāt-i Ēmētān*. Boulder, Colo.: Westview Press.
- Sims-Williams, N. (1981) 'Notes on Manichaean Middle-Persian Morphology', *StIr* 10, pp. 165–76.
- (2002) *Bactrian documents from Northern Afghanistan*, London and New York: Nour Foundation in association with Azimuth Editions and Oxford University Press (Studies in the Khalili Collection 3; Corpus Inscriptionum Iranicarum. Part II, Inscriptions of the Seleucid and Parthian periods and of Eastern Iran and Central Asia 6).
- Skjærvø, P. O. (1983) 'Case in Inscriptional Middle Persian, Inscriptional Parthian and the Pahlavi Psalter', *St Ir* 12, pp. 69–94, 151–81.
- (1985) 'Remarks on the Old Persian Verbal System', *MSS 45 (Festgabe für K. Hoffmann)*, pp. 211–27.
- (1986) 'Verbs in Parthian and Middle Persian Inscriptions', in R. Schmitt and P. O. Skjærvø (eds) *Studia Grammatica Iranica. Festschrift für Helmut Humbach*. München: R. Kitzinger, pp. 425–39.
- (1989) 'Verbal Ideograms and the Imperfect in Middle Persian and Parthian', in *Études irano-aryennes offertes à Gilbert Lazard (Studia Iranica, cahier 7)* Paris: Association pour l'avancement des études iraniennes, pp. 333–54.
- (1991a) 'MPers. *anāid*, *anānd*, in *Corolla Iranica*', in R. E. Emmerick and D. Weber (eds.) *Papers in Honour of Prof. D. N. MacKenzie on the Occasion of his 65th Birthday on April 8th, 1991*. Frankfurt, etc.: P. Lang, pp. 190–97.
- (1991b) review of M. Alram, *Nomina propria iranica in nummis*, *Iranisches Personennamenbuch IV*. (Wien: Österreichische Akademie der Wissenschaften, 1986), *IJ* 34, pp. 131–35.
- (1992) 'L'inscription d'Abnūn et l'imparfait en moyen-perse', *StIr* 21, pp. 153–60.
- (1993) 'The Earliest Datable Inscription on a Sasanian Bowl. Two Silver Bowls in the J. Paul Getty Museum', *BAI* 7, pp. 181–92 [pub. 1994].
- (1995a) 'Aramaic in Iran', *ARAM*, vol. 6 (*Palmyra and the Aramaeans*), pp. 283–318 [pub. 1997/98].
- (1995b) 'The Manichean Polemical Hymns in M 28 I', *BAI* 9, pp. 239–55 [pub. 1997].
- (1996) 'Iranian alphabets derived from Aramaic', in P. T. Daniels and W. Bright (eds) *The World's Writing Systems*. New York and Oxford: Oxford University Press, pp. 515–35.
- (1997a) 'On the Middle Persian Imperfect', in E. Pirart (ed.) *Syntaxe des Langues Indo-iraniennes anciennes. Colloque International, Sitges (Barcelona) 4–5 mai 1993 (Aula Orientalis Supplementa, vol. 6)*. Barcelona: Editorial AUSA, pp. 161–88.
- (1997b) 'The Joy of the Cup', *BAI* 11, pp. 93–104 [pub. 2000].
- (1999) 'Methodological Questions in Old Persian and Parthian Epigraphy', *BAI* 12, pp. 157–67 [pub. 2002].
- Sundermann, W. (1973) *Mittelpersische und parthische kosmogonische und parabeltexte der Manichäer (Berliner Turfantexte, 4)*. Berlin: Akademie-Verlag.
- (1981) *Mitteliranische manichäische Texte kirchengeschichtlichen Inhalts (Berliner Turfantexte 11)*. Berlin: Akademie-Verlag.
- (1989) 'Westmitteliranische Sprachen', 'Parthisch,' and 'Mittelpersisch', in *CLI*, pp. 106–13, 114–137, 138–64.
- (1997) *Der Sermon von der Seele (Berliner Turfantexte 19)*. Berlin: Akademie-Verlag.
- Sundermann, W., and Th. Thilo (1966) 'Zur mittelpersisch-chinesischen Grabinschrift aus Xi'an (Volksrepublik China)', *Mitteilungen des Instituts für Orientforschung* 11, pp. 437–50.

- Tavadia, J. C. (1956) *Die mittelpersische Sprache und Literatur der Zarathustrier*. Leipzig: O. Harrassowitz.
- Tedesco, P. (1921) 'Dialektologie der westiranischen Turfantexte', *MO*15, pp. 184–258.
- Weber, D. (1992) *Ostraca, Papyri und Pergamente* (CII III/IV–V, Texts I). London.
- (1997) 'Pahlavi Phonology', in A. S. Kaye (ed.) *Phonologies of Asia and Africa*. Winona Lake: Eisenbrauns, pp. 601–36.
- (2007) 'Einige Bemerkungen zur Paläographie des Pahlavi', in M. Macuch, M. Maggi, and W. Sundermann (eds), *Iranian Languages and Texts from Iran and Turan. Ronald E. Emmerick Memorial Volume*. Wiesbaden: Harrassowitz, pp. 434–42.
- Williams, A. V. (1990) *The Pahlavi Rivāyat Accompanying the Dādestān ī Dēnīg*, 2 vols. (Det Kongelige Danske Videnskabernes Selskab, Historisk-filosofiske Meddelelser, 60). Copenhagen: Munksgaard.

CHAPTER FIVE

SOGDIAN

Yutaka Yoshida

1 INTRODUCTION

1.1 Overview

1.1.1 History

Sogdian is a Middle Iranian language once spoken in Sogdiana. Among the Middle Iranian languages Sogdian is classified as belonging to the North–Eastern group which also includes Khotanese, Tumshuqese, Bactrian, and Choresmian.

The land of Sogdiana is located between the two great rivers of Central Asia, the Amu-Darya and the Syr-Darya (part of modern Uzbekistan and Tajikistan). Sogdiana consisted of several oasis-states which were located along the rivers Zarafshan and Kashka-Darya, the leading cities being Samarqand, Buchara, and Kish (present day Shahr-i Sabz). The eastward movement of the Sogdians led them to colonize Tashkent and an area beyond, and consequently a Chinese pilgrim Xuanzang who went to India via Central Asia in the seventh century referred to the region between Semirechie and Iron Gate, located to the south of Kish, as Suli, i.e. Sogdiana.

The history of Sogdiana is largely obscure. It constituted a satrapy of the Achaemenian Empire, and after the conquest of the empire by Alexander the Great in the fourth century BC, Sogdiana was under the control of neighbouring super powers, such as the Kushans (first to third centuries), the Sasanians (third century), the Kidarites (fourth to fifth centuries?), the Hephtalite (fifth to sixth centuries), the Western Turks (sixth to seventh centuries), and Chinese (seventh to eighth centuries). However, it was able to enjoy a degree of independence until it was conquered by the Arabs in the eighth century. During this period of relative independence the Sogdians played an active role as international traders along the Silk Road between China and the West, with the result that the Sogdian language became a kind of lingua franca in the region between Sogdiana and China, where the Sogdians founded many trade diasporas.

The dominion by the Achaemenian Empire was important for Sogdian in that the chancellery language of the Empire, Aramaic, was introduced to Sogdiana, from which later the Sogdian script developed after the Empire had collapsed and Aramaic ceased to be used by the local scribes. By the end of the sixth century, this Sogdian script came to be written vertically rather than horizontally from right to left, and lines running from left to right.



MAP 5.1 SOGDIA AND ITS NEIGHBOURS

1.1.2 Materials

The Sogdian materials handed down to us reflect the activities played by the Sogdians in history. (On the major publications until 1995, see Gharib 1995: xiii–xxvii.) Except for the very short inscriptions discovered in Kultobe, Kazakhstan (cf. Sims-Williams and Grenet 2006) and the legends found on coins of the early centuries of CE, the earliest substantial materials are the so-called ‘Ancient Letters’ discovered by Sir Aurel Stein in a watch tower located on the route between Dunhuang and Loulan. They were sent to the addressees staying in Samarkand and oasis cities lying in between by Sogdian merchants and their families resident in China, who reported the news about their activities, and were shown to have been written in the early fourth century (cf. Grenet and Sims-Williams 1987). Numerous rock inscriptions or graffiti discovered in North Pakistan along the ancient route connecting Central Asia and North West India are similar to the Ancient Letters in the ductus and seem to have been inscribed by merchants who came there either directly from Sogdiana or by way of China.

Their activities of the sixth to seventh centuries are reflected by an epitaph dated 580 CE which was unearthed from a tomb of a Sogdian couple in Xi’an and a sale contract of a female slave dated 639 CE and discovered in Turfan (see Yoshida 2005 and Yoshida *et al.* 1989). Apparently, during the seventh to eighth centuries when the Tang Dynasty was the most powerful, Sogdian trade activities reached their height, and a great many Sogdians immigrated to China. Some of them were converted to Buddhism under the influence of Chinese Buddhists. They produced many Buddhist Sogdian texts based on Chinese prototypes, of which dozens of specimens were discovered in Dunhuang and Turfan.

The role played by the Sogdians in the propagation of the Manichaeism is not entirely clear. However, a number of Manichaean Sogdian texts were discovered in Turfan, where the Uighurs, who adopted Manichaeism in 762/3 CE as their state religion, constituted the

ruling class. Apart from the translations from the Middle Persian and Parthian originals, one finds many works originally composed in Sogdian. Christian monks resident in a monastery of Bulayīq located to the north of the Turfan basin left numerous Christian Sogdian texts based on the Syriac originals. That they were also engaged in trade is inferred from a few Dunhuang texts written by Christians. Apart from the Sogdian script, which was so to speak the Sogdian national script, Manichaean and Syriac scripts were employed respectively to write Manichaean and Christian texts.

Similarly, Brahmi script was applied to write texts translated from Sanskrit (or Tocharian as well?), either of Buddhist or medical contents. However, only a small number of Brahmi fragments have so far been encountered among the treasures unearthed from Central Asia (see Sims-Williams 1996c).

While the bulk of Sogdian materials are from East Turkestan, the land of Sogdiana also yields a handful of manuscripts and short inscriptions. The most famous are some 80 secular documents discovered at Mount Mugh, which represent part of archives of Dhēwaštīč (?–722), the last independent ruler of Penjikent.

While the Sogdians were governed by Turkish peoples, such as Turks and Uighurs, they exercised strong cultural influence upon their rulers; the Sogdian language served as a literary language for them until they began to write their own languages, first in Runic, and later in Uighur script, of which the latter originated from the cursive variant of the Sogdian script. Of the two Sogdian inscriptions unearthed in Mongolia, one, the Bugut Inscription, was erected by Turks around 600 CE and the other, the Karabalgasun Inscription, is trilingual in Sogdian, Chinese and Uighur in Runic script, and commemorates the eighth Khaghan of the Uighur empire (r. 808–822). Another stele which is similar in contents to the Bugut Inscription was discovered on a stone statue unearthed in Mogolküre in Xinjiang.

Very late inscriptions going back to the tenth to eleventh centuries, which have been discovered in Kirghizia, contain several Turkish elements; they may have been written by Turkophones who employed Sogdian as a written language (see Livšic 1996). A few late documents also containing Turkish words and phrases were discovered in Dunhuang and Turfan.

The linguistic variations found among all these materials are not trivial mainly for chronological reasons. Moreover, the bulk of Christian Sogdian texts and a great part of the Manichaean materials are written respectively in Syriac and Manichaean scripts, whose orthographies are free from historical spellings found in texts written in Sogdian script. However, despite the wide area of diffusion of the Sogdian language, very few purely dialectal differences have so far been observed. On this point see the section 7.2, Dialects and variations.

1.2 Scripts, orthography, and phonetic values

The three scripts (Sogdian, Manichaean, and Syriac) derive ultimately from the Aramaic and thus have many features in common, the most salient being that they are so-called consonant scripts, and that vowel phonemes are not always sufficiently distinguished. Inevitably some uncertainties surround the problem of the Sogdian vowel phonemes, the situation not being radically improved by the recent discovery of a few texts written in Brahmi script (see Sims-Williams 1996c). In this connection an important advantage of Syriac script is to be mentioned; vocalic points are sometimes employed to distinguish vowels of different qualities: *y*: = [ē], *y*: = [ī], *w*: = [ō], *w*: = [ū], and *ç* = [Ca]. In the texts written in Sogdian script, which evolved from the Achaemenid chancery Aramaic, one

TABLE 5.1: SOGDIAN SCRIPTS

Aramaic	"Ancient Letters"	"Sogdian" (Buddhist) Sogdian Script ¹		"Manichean" Sogdian Script ²		"Syriac" (Christian) Sogdian Script	
		<>	[]	<>	[]	<>	[]
'	𐩺		' ā, a (ə, i)		' ā, a (ə, i)		' ā, a (ə, i)
b	𐩻		β β, f		b b		b β, b
			(β =) f f		(b̄ =) β β		
g	𐩼		γ γ		g g		g g
					γ γ		
d	𐩽		D ³ -		d d		d δ
h	𐩾		h ⁴ -ǎ, Ø		h ⁴ Ø		h h
w	𐩿		w w, ō, ū		w w, ō, ū		w w, ō, ū
z	𐩿		z z, ž		z z		z z
			(z, ž =) ž ⁵ ž		j ž, j		ž ž
h	𐩿 𐩾		x x, h		h h		h h(?)
t			t ⁷ -		t t (d)		t t (d)
y	𐩿		y y, č, ĭ		y y, č, ĭ		y y, č, ĭ

k	و	وگا	k	k (g)	گ	k	k	ک	k	k ⁶
l	ل	لا	δ/L ³	δ, θ, l	خا	(k̄ =) x	x	ک	(k̄ =) x	x
m	م	مگا	m	m	د	l	l	د	l	l
n	ن	نگا	n	n, ṃ	ڈ	δ	δ, θ	ڈ	m	m
s	س	سگا	s	s	ڄ	m	m	ڄ	m	m
'	ا	اگا	' ³		چ	n	n, ṃ	چ	n	n, ṃ
p	پ	پگا	p	p (b), f	ج	s	s	ج	s	s
š	ش	شگا	c	č (j)	ڇ	'	i	ڇ	'	γ
q	ق	قگا	q ⁷	-	ڄا	p	p (b)	ڄا	p	p (b)
r	ر	رگا	r	r, r', l	ڄا	f	(p̄ =) f	ڄا	f	f
š	ش	شگا	š	š	چ	c	č (j)	چ	c	č (j), ts
l	ل	لگا	l	t (d)	ڄا	q	k	ڄا	q	k (g)
					ڄا	r	r, r'	ڄا	r	r, r'
					ڄا	š	š	ڄا	š	š
					ڄا	l	t (d)	ڄا	θ	θ

finds a certain number of ideograms, i.e. Aramaic word forms pronounced with their Sogdian equivalents (see also Chapter 4 on Middle West Iranian). For example, a word for ‘ten thousand’ is often spelled as RYPW (cf. Aramaic rbw = ribbō) and was pronounced as *brēwar*. That it was pronounced as a Sogdian word is proved by a phonetic complement which appears in its cardinal Aramaic form with the Sogdian ordinal suffix (-myk): (RYPWmyk =) *βrywrmyk, i.e. *brēwar-mīk* ‘10,000th’. (In modern transliterations one distinguishes ideograms by Roman capital letters.) Below is the chart of Sogdian, Manichaean, and Syriac scripts with their transliteration and phonetic values, which will be discussed in the following section.

- 1 Of the two major varieties, formal and cursive scripts, the former as attested in the manuscript of the *Vessantara Jātaka* is reproduced.
- 2 The frequent doubling of the letter y, w, and δ, and ʧ has no phonetic significance.
- 3 Only used in ideograms.
- 4 Only used in word-final position.
- 5 Usage varies: some scribes employ the diacritic point(s) to distinguish z (= [z] or [ʒ]) from n, others to distinguish [ʒ] from [z].
- 6 Only in knθ, kθ ‘city’ and in Syriac words.
- 7 Not used.

2 PHONOLOGY

In the following, M and C indicate sources in Manichaean script (without distinction of t and ʧ) and Christian Syriac script, respectively. Otherwise, spellings cited are those in Sogdian script (if necessary indicated by S). Light stems, on which see below, are differentiated from heavy with a hyphen (-).

2.1 Inventory and distribution

2.1.1 Vowels and diphthongs

2.1.1.1 Vowel system

Given that the three scripts, Sogdian, Manichaean, and Syriac, ultimately derive from the Aramaic consonant script, there remain uncertainties in the determination of the Sogdian vowel phonemes, which are inconsistently indicated.

TABLE 5.2: VOWELS AND DIPHTHONGS

	front	central	back	Diphthongs
high	<i>ī, i</i>	<i>(i)</i>	<i>u, ū</i>	
mid	<i>ē, e</i>	<i>(ə)</i>	<i>o, ō</i>	
low	<i>a</i>	<i>ā</i>		<i>āi āu; V + ʧ; V + ʃ</i>

There is phonemic length distinction, which is not overtly marked except for the low pair, and is otherwise inferred etymologically as well as by means of the ‘rhythmic law’ (see below): *βaγ-* (β’γ-) ‘god’, *βāγ* (β’γ) ‘garden, farm’; but *fneš-* (C fnyš-) ‘be deceived’,

fñēš (C fnyš) ‘deceive’; *witar-* (wytr-) ‘go’, *wītar* (wytr) ‘went (3s imperf.)’; *but-* (pwt-) ‘Buddha’, *pūtē* (pwt’k) ‘rotten’; *roxšn-* (rwxšn-) ‘light, bright’, *rōyn* (rwyn) ‘oil’.

Naturally, long and short vowels are likely to be different in terms of quality as well. Thus *ā* is a back vowel in contrast with *a* which is front. This is clear in view of the Uighur orthography which is based on the Sogdian spelling conventions; thus in Uighur a back vowel *a* is expressed by initial (ʷ-) and medial (-ʷ-), both being the notations for *ā* in Sogdian, whereas the front counterpart *ā* is represented by those standing for *a* in Sogdian (see Sims-Williams 1981: 358).

i and *ə* are allophones of a phoneme *a* mainly appearing in the initial position where *a* never occurs: *əpsāk* (ʷps’k) ‘garland’. Compare also the fact that foreign words beginning with *a-* are usually transcribed with (ʷ-), i.e. *ā-* rather than (-), i.e. *ə-*, e.g. (ʷswr) for Skt. *asurā*.

The high central vowel *i* is restricted to the position before *sp-*, *st-*, and *sn-* and is transcribed by the letter ‘ (‘ain) in Manichaean script: *isptyāk* (M ‘spty’k) ‘completion’. When they are unaccented, short vowels *i*, *a*, and *u* seem to have been reduced to *ə*. The alternation of (-y-), (-w-), and (zero) encountered in the spellings seems to reflect this situation: e.g. *žət-* (M jt- ~ jyt-) ‘strike’ (past stem < OIr. **jata-*); *δəštē* (M δyštyy ~ S δšt’y) ‘built’ (< OIr. **dištaka-*); *ōsəytē* (ʷs(ʷ)ty ~ ʷs(ʷ)wyt) ‘pure’ (< OIr. **ava-suxtaka-*).

2.1.1.2 Rhotacised vowels

Apart from the above-mentioned simple vowels, Sogdian possesses three rhotacised vowels, *əʳ*, *iʳ*, *uʳ* which are counted as short vowels: *məʳγ-* (mrγ-) ‘bird’, *kiʳm-* (kyrm-) ‘snake’, and *puʳn-* (pwrn-) ‘full’. This element constitutes the second element of diphthongs: *mary* = *maəʳγ* (mrγ) ‘forest’, *zern* = *zeəʳn* (zyrn) ‘gold’.

2.1.1.3 Diphthongs

As the second member of diphthongs Sogdian also possesses a nasal element *ŋ*, which probably had moraic status: *kaŋθ* (knδh) ‘city’, *paŋj* (pnc) ‘five’.

The long diphthongs *āi* and *āu* are also known: *āikūn* (ʷykwn) ‘forever’, *āurāōē* (ʷwr’δ’k) ‘fellow-traveller’, while Old Iranian **ai* and **au* have become *ē* and *ō* respectively.

2.1.2 Consonants

2.1.2.1 Consonantal system

Sogdian has the following consonants; those in parentheses are allophones or marginal phonemes.

TABLE 5.3: CONSONANTS

	labial	dental	alveolar	palatal	velar/glottal
Plosive/affr.	<i>p</i> (<i>b</i>)		<i>t</i> , (<i>d</i>) (<i>ts</i>)	<i>č</i> (<i>ǰ</i>)	<i>k</i> (<i>g</i>)
Fricative	<i>f</i> <i>β</i>	<i>θ</i> <i>δ</i>	<i>s</i> <i>z</i>	<i>š</i> <i>ž</i>	<i>x</i> <i>γ</i>
Nasal	<i>m</i>		<i>n</i>		(<i>ŋ</i>)
Liquid/Glide	<i>w</i>		<i>r</i> (<i>l</i>)	<i>y</i>	(<i>h</i>)

The typologically marked system with voice opposition found only with fricatives is due to the sound change in which the voiced plosives and affricate **b, *d, *ǰ, *g* have become fricatives *β, δ, ž,* and *γ* even in initial position: *βar-* (βr-) 'bring', *δār* (δ'r) 'to hold', *žār* (C ž'r) 'poison', *γar-* (γr-) 'mountain'. On the other hand, due to the Sogdian conservatism, the voiceless plosives and affricate **p, *t, *č, *k* even after a vowel are preserved: *āp* ('p) 'water', *wāt* (w't) 'wind', *wyāk* (wy'k) 'place', *wāč* (w'c) 'send', 'release'.

The voiced counterpart *b, d, ǰ, g* are found only after nasalized vowels, i.e. *ŋr*: *zamb* (M zmb) 'coast, river bank', *βamdē* (M βndy) 'slave', *paŋǰmīk* (M pŋjmyk) 'fifth', *saŋg* (M sng) 'stone'.

2.1.2.2 Marginal *l, h, ts,* and *ŋ*:

l and *h* are marginal phonemes only found in loan words and foreign forms: *šlōk* (šl'wk ~ šr'wk) < Skt. *śloka* 'verse'; *rahaṃd* (M rhnd) < Skt. *arhant* 'arhant'. Only sporadically do *l* and *h* appear in original Sogdian words as a result of dissimilation: *wilarz* (C wlrz ~ S wyr'rz) 'tremble' (< OIr. **wi-rarz-*, cf. unreduplicated form attested in Khotanese *rrīyz-* 'tremble'), *hunax* ~ *xunax* reflected in (M hwnx ~ xwnx) 'that'.

The marginal status of *ts* may be inferred from the fact that the Christian Sogdian form (mc') 'hither' corresponds to (mrts'r) in Sogdian script; thus Syriac *tzaddi* (c) can represent *č* as well as *ts*, so that *ts* is treated as a single segment. Compare also (M pnct) and (C pncc) both representing *panč(a)ts* 'fifteen'.

ŋ is sometimes represented by the spelling (-nn-) in Manichaean script: *sarθang* (srđnng beside srđng), i.e. *saθ'ŋg* 'leader'.

2.1.3 Syllable structure and clusters

Due to the Sogdian orthography with consonant scripts it is not always easy to know the syllabic structures. Etymological consideration and spellings lead us to assume that Sogdian allows consonant clusters consisting of a considerable number of consonants: *žuxšk-* (C žwxšq-) 'disciple', *sfrīn* (M sfryn) 'create', *xēpθī* (C xypθ-t-y) 'one's own (pl. obl.)'.

The pronunciation of clusters is occasionally eased by prothetic vowel or by metathesis: *psāk* (M ps'k) ~ *apsāk* (M 'ps'k) 'garland', *žuxšk-* (C žwxšq-) ~ *žxušk-* (žxwšk-, đrxwšk-) 'disciple', etc.

Combinations of voiced fricative and voiceless plosive or affricate are characteristic of Sogdian, *paðk-* (pðk-) 'law', *əβta* ('βt) 'seven'. In view of the similar clusters in Ossetic, one may assume that the second member was pronounced as a weak devoiced plosive, i.e. *əβda*.

2.2 Non-segmental features and rhythmic law

(In the following stress is indicated by an underlined vowel.) The position of the stress accent is governed by the so-called 'rhythmic law' (see Sims-Williams 1984) as follows:

1. A syllable containing either a long vowel or diphthong is counted as heavy while a syllable consisting only of a short vowel is light.
2. Stress falls on a stem if it contains a heavy syllable, i.e. heavy stem, whereas it falls on the suffix or ending in the case of light stems that consist only of light syllable(s). Unstressed endings and suffixes suffer reduction or loss, so that the patterns of

conjugation and declension of heavy stems are totally different from those of light stems; thus:

heavy stem: *mēθ* (C myθ) ‘day (dir. sg.)’, *mēθ-ī* (myθy) (obl. sg.);

light stem: *ram-ī* (rmy) ‘people (nom. sg.)’, *ram-u* (rmw) (acc. sg.), *ram-ya* (rmy) (loc. sg.), etc.

Enclitic and proclitic forms, even if they consist of light syllables, do not bear stress and lose their endings: *βay-a* (βγ) vs. *-βay* (-βγ) ‘o lord!’ (encl. voc.); *astī* (‘sty) vs. *ast* (‘st) ‘is’ (encl.);

Light stems ending with *-r*, *-w*, and *-y* sometimes behave as heavy when they are followed by endings beginning with a consonant: *βar-t* (βrt) ~ *βar-tī* (βrty) ‘bears’ (3 sg. pres.).

2.3 Alternations and combinatory phenomena

A considerable amount of spelling variations are observed throughout the texts; some are simply due to the difference between historical spellings and those more or less reflecting the actual pronunciation: (ə)xšēθ (‘xš’y”δ vs. M xšyδ) ‘king’ (<Olr. *xšāyaθiya).

Most conspicuous are the differences observed between Christian Sogdian forms in Syriac script and those in Sogdian script. Prothetic vowels are virtually non-existent in the former while they abound in the latter: *zβāk* (C zb’q) ‘tongue’ vs. *əzβāk* (zβ’k). Another peculiar difference between the two groups is the existence of *ṃ* in the latter and its loss in the former before continuants: *kaṃθ* (knδh) ‘town’ vs. *kaθ* (C qθ), *aṃyām* (S, M ṃy’m) ‘end’ vs. *ayām* (C’y’m), but compare *βaṃtē* (S βnt’k, M βndy, C bnty) ‘slave’. In late ‘demotic’ texts one sometimes finds the tendency to voice *t* and *p* after a voiced sound: *āb* (C’b, cf. M ’p) ‘water’, *mardəxmē* (C mrdxmy, cf. M mrtxmy) ‘man’, *saydyā* (C sydy, cf. M syty) ‘on the . . . th day’.

Similarly, post-vocalic *r = ʃ* is often lost in the Christian texts: S *zeʃn* (zʃrn) ‘gold’ vs. C *zen* (zyn); S, M *sāʃ* (s’r) vs. C *sā* (s’).

The loss of *ṃ* and *ʃ* does not change the rhythmic status of the stems. Accordingly, Christian Sogdian *kaθ* and *zen* behave as heavy stems.

Metathesis of *u* or *w*, both progressive and regressive, is so common that one almost always finds alternative forms: *δwyt-* (δwyt-) ~ *δwt-* (δwt-) ‘daughter’, *swyδīk* (swyδyk) ~ *swyδīk* (swyδyk) ‘Sogdian’, *γwāθuk* (γ’δwk) ~ *γwāθuk* (γw’δk) ‘throne’, compare OIr. **gāθu-*. This metathesis affects even loanwords: *samutr-* ~ *sumtar-* (smwtr- ~ swmtr- < Skt. *samudra*) ‘ocean’.

The cluster *čt* often becomes *št*: *sāčt* (s’ct) ~ *sāšt* (s’št) ‘it is necessary’, *βaṃpūrēšt* (< *βaṃpūrēč* with the plural ending *-t*) ‘divine virgins’.

3 MORPHOLOGY

From here on, the transcription is somewhat simplified; thus *n* or *m* for *ṃ*, *r* for *ʃ*. Also, voiced plosive allophones after *ṃ* are transcribed as *p*, *t*, *k* or *b*, *d*, *g* in conformity with spellings: e.g. *βantē* (S βnt’k) or *βandē* (M βndy) for phonemic *βaṃtē*. Occasionally transliterated forms are given, somewhat inconsistently, for the sake of clarity. Moreover, the inflectional tables are idealized in that attested endings are added to sample stems regardless of whether the forms in question are actually attested or not.

3.1 Nominal morphology

The inflections of Sogdian nouns and adjectives are identical, and the following

description also applies to adjectives, except that the latter lacks the numerative, on which see below. On the Sogdian inflection see Sims-Williams 1982 and 1990.

The Old Iranian distinction of three genders (masculine, feminine, and neuter) has been preserved, although the survival of the neuter is marginal as many old neuter nouns have shifted to masculine or feminine. Similarly, the distinction of three numbers has been preserved; however, the old dual forms have come to be used in the position immediately following a numeral, where they occur not only after the number two but also higher numbers, and thus developed in the special form called ‘numerative’ (NUM), and discovered by Sims-Williams (1979).

The case distinctions, largely reflecting Old Iranian, are: nominative (NOM, N), accusative (ACC, A), genitive-dative (G–D), locative (LOC), instrumental-ablative (I–A), and vocative (VOC), with light stems, which are reduced to the opposition of direct (DIR) vs. oblique (OBL) with heavy stems.

3.1.1 Stem classes and declensions

Sogdian nouns are classified into several stems. Apart from the distinction between light and heavy (hereafter abbreviated as LS and HS respectively), a few light stems ending with *-u* (*-w*) inflect differently from ordinal light stems.

A considerable number of stems go back to forms extended by the suffixes masc. **-aka* and fem. **-ākā*, and are referred to conventionally as the *aka*-stem and the *ākā*-stem. In Sogdian script, these two stems are often written with historical spellings, *(-ʿ)k* and *(-ʿ)kh*, respectively. The feminine counterpart of the masculine stems ending in *-ē* (< **-aka*) usually corresponds with the adjectival feminine stems ending in *-č̄*: masc. *spt-ē* (C spty) ‘complete’, fem. *spč̄-* < *spt-č̄-* (C spc-).

There are also a few indeclinable nouns ending in *-ī* (*-y*): *martī* (mrty) ‘man’.

3.1.1.1 Light stems

The following are paradigms of light stem nouns masc. *ram-* ‘people’, neut. *βayn-* ‘temple, altar’, and fem. *wan-* ‘tree’.

The regular plural forms are marked by the ending *-t-*. These plural stems are treated as feminine singular; that is, the plural form of light stems like *ram-* have the same inflection as feminine light stems like *wan-*.

TABLE 5.4: DECLENSION OF LIGHT STEMS

	Singular			Plural	
	‘people’ masc.	‘temple’ neut.	‘tree’ fem.	‘people’ masc.	‘tree’ fem.
	<i>ram-</i>	<i>βayn-</i>	<i>wan-</i>	<i>ram-t-</i> (pl.)	<i>wan-t-</i> (pl.)
NOM	<i>ram-i</i>	<i>βayn-u</i>	<i>wan-a</i>	<i>ram-t-a</i>	<i>wan-t-a</i>
ACC.	<i>ram-u</i>	<i>βayn-u</i>	<i>wan-a</i>	<i>ram-t-a</i>	<i>wan-t-a</i>
G–D	<i>ram-e</i>	<i>βayn-e</i>	<i>wan-ya</i>	<i>ram-t-ya</i>	<i>wan-t-ya</i>
LOC	<i>ram-ya</i>	<i>βayn-ya</i>	<i>wan-ya</i>	<i>ram-t-ya</i>	<i>wan-t-ya</i>
I–A	<i>ram-a</i>	<i>βayn-a</i>	<i>wan-ya</i>	<i>ram-t-ya</i>	<i>wan-t-ya</i>
VOC	<i>ram-a</i>		<i>wan-e</i>	<i>ram-t-e</i>	<i>wan-t-e</i>
NUM, N–A	<i>ram-a</i>	<i>βayn-e</i>	<i>wan-e</i>		

For a masculine *u*-stem noun, note nom.-acc. *maγu* (mγw) ‘magus’, gen.-dat. *maγw-e* (mγwy).

3.1.1.2 Heavy stems

As indicated above, the declension of heavy stems does not differentiate masculine from feminine and distinguishes only direct and oblique cases. It is to be noted that in actual texts, especially those in Sogdian script, heavy stems often take the endings of light stems by analogy and possibly because of the clarity and salience of light stem inflection for the syntactic function.

TABLE 5.5: DECLENSION OF HEAVY STEMS

	masc. <i>mēl</i> ‘day’	pl.	fem. <i>žwān</i> ‘life’	pl.
DIR	<i>mēl</i>	<i>mēl-t</i>	<i>žwān</i>	<i>žwān-t</i>
OBL	<i>mēl-t-ī</i>	<i>mēl-t-ī</i>	<i>žwān-ī</i>	<i>žwān-t-ī</i>
VOC	<i>mēl(a)</i>	<i>mēl-t-e</i>		<i>žwān-t-e</i>

3.1.1.3 Irregular plural forms, *-ar-t*, *-īš-t*

The plural forms of *βrāt* ‘brother’ and *δuyt-* ‘daughter’ have the formant *-ar-t*: *βrāt-ar-t*, *δuyt-ar-t*. Also, some animate nouns, especially light stems, have the plural markers dir. *-īš-t*, obl. *-īš-t-ī*: *βaγ-* ‘god’, pl. *βaγ-īš-t-*.

3.1.1.4 *aka-* and *ākā-* stems

Historically, intervocal *-k-* was lost, and the resulting hiatus was later contracted (see Sims-Williams 1990: 286–291); thus, nom. masc. sg. **-aki* > *-*a’i* > *-ē*; similarly: acc. **-aku* > *-*a’u* > *-ō*. In further development, while in one Christian Sogdian manuscript C2 the original case distinctions are well preserved after the contraction, in all the other texts the ending *-ē* found in the nom., gen.-dat., and loc. sg. was generalized, replacing acc. *-ō*, and instr.-abl., voc. *-ā*. The plural ending *-t* is suffixed to the nominative form, thus *-ē-t*, and inflects as the heavy stem.

The contraction in the feminine *ākā-* stems similarly resulted in minimal distinctions. The following table shows the declension patterns of masc. *martəxmē* ‘man’ in C2 and in other texts, and of fem. *xānā* ‘house’.

TABLE 5.6: DECLENSION OF *aka-* AND *ākā-* STEMS

	masc. <i>martəxm-ē</i> ‘man’			fem. <i>xānā</i> ‘house’			
	C2	other texts	pl.	sg.	pl.		
NOM	<i>martəxmē</i>	<i>martəxmē</i>	DIR	<i>martəxmē-t</i>	DIR	<i>xānā</i>	<i>xānē-t</i>
ACC	<i>martəxmō</i>	<i>martəxmē</i>	OBL	<i>martəxmē-t-ī</i>	OBL	<i>xānē</i>	<i>xānē-t-ī</i>
G-D	<i>martəxmē</i>	<i>martəxmē</i>	VOC	<i>martəxmē-t-e</i>	VOC		<i>xānē-t-e</i>
LOC	<i>martəxmē</i>	<i>martəxmē</i>					
I-A	<i>martəxmā</i>	<i>martəxmē</i>					
VOC	<i>martəxmā</i>	<i>martəxmē</i>					
NUM, N-A	<i>martəxmā</i>	<i>martəxmēlā</i>					

3.1.1.5 Indeclinables and archaic plural *-ān*

Indeclinable nouns do not inflect in the singular, but have the binary case distinctions in the plural ending: sg. dir., obl. *martī* ‘man’, pl. dir. *martī-t*, obl. *martī-tī*. Some nouns take the old genitive plural ending *-ān* (< OIr. **-ānām*): *βaγ -ān* ‘gods’ (LS), *pīd-ān* ‘elephants’ (HS), *martaxm-ān* (< **-a-ān*, *aka*-stem). This ending is restricted to stereotyped phrases such as *βaγ-ān βaxtam* ‘godliest of gods’ = Skt. *devātideva*.

3.1.2 Adjectives

3.1.2.1 Degree

The productive suffix of the comparative is *-(i)star*: *murzək-istar* ‘shorter’, *γawānčīk-star* ‘more necessary’. Somewhat obsolete is the suffix *-tar*, which is not suffixed to derived stems or compounded forms. Thus we have *dūr-tar* ‘farther’, *namr-tar-* ‘sweeter’ but *mand-γrβāk-star* ‘more stupid’, *pačxūd-γōnē-star* ‘more despicable’, etc. Irregular forms are: *mazēx* ‘big, great’ > *masyātar*, *γarf* ‘many, much’ > *fyātar*, *šir-* ‘good’ > *šyātar*, *kaβn-* ‘little, few’ > *kaṃpī*, etc.

Old superlative forms with the suffix *-tam* are attested. They are either restricted to stereotype expressions, as in *βaγ-ān βax-tam* ‘devātideva’, or lose their superlative meaning and are treated as simple adjectives (to which secondary suffixes are added: *askā* ‘high’ ~ *askā-tam-čīk* ‘highest’).

An alternative way is preposing *ēw* ‘one’ or *āḍparm* ‘whatsoever’ to the comparative form, e.g. *ēw əxšnak-istar* ‘the most excellent’, *āḍparm fratar* ‘the best’.

3.1.2.2 Elative

Sogdian possesses special elative formation with the meaning ‘so much ~, very ~’. It is formed by means of the exclamatory particles *čā-* ‘how’ and *wā-* ‘so’ ~ *wat-* (before *s-*), together with the suffixes *-t*, *-(ə)st*, in various combinations: *čā-əβīzāxuk-st* ‘so painful’, *wā-zārī* ‘so miserable’, *wā-frāk-t* ‘so early in the morning’, *wat-spēt-t* ‘so white’.

3.1.3 Pronouns and deixis

3.1.3.1 Personal pronouns

The 1s and 2s personal pronouns distinguish direct and oblique. Independent forms of the 3rd person are provided by weak demonstratives (see below):

TABLE 5.7: PERSONAL PRONOUNS

	1s	2s	1p	2p	3s	3p
DIR	<i>(ə)zu</i>	<i>təγu</i>	<i>māx</i>	<i>(ə)šmāx</i>	= dem. pron.	
OBL	<i>mana</i>	<i>tawa</i>	<i>māx</i>	<i>(ə)šmāx</i>	= dem. pron.	
ENCL	<i>-mīl-m</i>	<i>-fīl-f</i>	<i>-man</i>	<i>-fan</i>	<i>-šu, -šī, -š</i>	<i>-šan</i>

Some texts show case distinction in enclitic forms; thus 2s *-f* for acc. and inst.-abl. and *-t(ī)* for gen. -dat. 3s *-šu* is chiefly used for acc. and *-šī* for all cases including acc.

When they are dependent on adpositional elements which are also enclitic (-č ‘from’, -δ ‘with’, -t ‘to’), forms -m, -f, and -š appear: *rti-šč* (rty-šc) ‘and from him’, etc. The 1s and 2s have also fused prepositional case forms: 1s *tāmā* and 2s *tāfā* (S t’β’kh) functioning as acc. (*t-* obsolete *ət(ā)* ‘to’), and similarly *čāmā*, *čāfā* (č- ‘from’), *δāmā*, *δāfā* (δ- ‘with’), *parāmā*, *parāfā* (*par-* ‘on, by’). (See also Adpositions, section 3.1.4).

3.1.3.2 Reflexive and reciprocal forms

A feminine noun *γrīw* serves as the reflexive pronoun referring to the subject, while it also retains its original meaning ‘body’:

wāyōnč īnč kū əγrīw δārāt
 such woman to self hold.SUBJ.3s
 ‘one should keep such a woman with oneself’.

Reciprocal sense is conveyed by *ēw* ‘one’ in combination with the inflected form of *δəftγ-* ‘second’ or by *əny-* ‘other’ combined with another *əny-*. *xadanyu* ‘each other’ emphasizes the meaning: *ēw δəftya xadanyu nē γərβant* ‘they do not know each other’.

3.1.3.3 Articles, demonstrative pronouns and adverbials

Sogdian distinguishes two kinds of demonstratives: (1) one employed as the article and 3rd person pronouns (weak demonstratives), and (2) the other extended forms functioning as proper demonstratives (strong demonstratives). The latter forms are extended from the former by means of several elements.

Sogdian distinguishes three foci of deixis, each consisting of a direct and oblique base: (1) *y-/m-* ‘this (with me)’, (2) *š-/t-* ‘that (with you)’, and (3) *x-/w-* ‘that (with him)’ (see Sims-Williams 1994). The three foci are here referred to by Ich-, Du-, and Er- deixis. Strong demonstratives are extended either with *-n-* or *-ēδ-*. There are adverbial forms based on the demonstrative bases. The following tabulates the declensions of some attested samples with their typical uses.

TABLE 5.8: ARTICLES AND PERSONAL PRONOUNS

Ich-Deixis			
	masc.	fem.	pl.
NOM	<i>yu</i> (yw)	<i>yu</i> (yw)	= f.
ACC	<i>(ə)mu</i> (‘)mw)	<i>əma</i> (mh) <i>əmu</i> (‘mw)	= f.
G-D	<i>əmen</i> (‘myn)	<i>əmī</i> (‘my)	<i>mēšan</i> (myšn)
LOC	<i>əmya</i> (‘my) <i>əmī</i> (‘my)	<i>əmya</i> <i>əmī</i>	
Du-Deixis			
	masc.	fem.	pl.
NOM	<i>(ə)šu</i> (‘)šw)	<i>ša</i> (š’)	= f.
ACC	<i>ətu</i> (‘tw)	<i>əta</i> (‘th)	= f.
(On (š’), see Yoshida 2000: 82.)			

Er-Deixis			
	masc.	fem.	pl.
NOM	(ə)xu ((')xw)	xa (xh, x')	= f.
ACC	ō ('w)	wa (w')	= f.
		ō ('w)	= f.
G-D	wənī (w(y)ny) wənī ('wyn)	wya (wy') əwī ('wy)	wēšan (wyšn)
LOC	wya (wy') əwī ('wy)	wya əwī	

TABLE 5.9: DEMONSTRATIVES

n-Extension			
Ich-Deixis			
	masc.	fem.	pl.
NOM	yunē (ywn'k) ēnē ('yn'k)	yānā (y'n'kh)	yānt (C y'nt)
ACC	munō (mwn'kw)	mānā (m'n'kh)	mānt (C m'nt)
G-D	nīmant(ī) (nym'nt(y))		
LOC	əmyamant(ī) ('my'mnt(y))		
Du-Deixis			
	masc.	fem.	pl.
NOM	šunē (šwn'k)	—	šānt (š'ntt)
ACC	tunō (twnkw)	—	—
	(On šunē, cf. Yoshida 2000: 81.)		
Er-Deixis			
	masc.	fem.	pl.
NOM	xunē/xunax (xwn'k/xwn'x)	xānā (x'n'kh)	xānt (C x'nt)
ACC	ōnō/wānō ('wn'kw/w'nw)	wānā (w'n'kh)	wānt (C w'nt)
G-D	(wə)niwant(ī) ((w)nyw'nt(y))		
LOC	wyawant (wy'wnt)		
ēδ-Extension			
	Ich-Deixis	Du-Deixis	Er-Deixis
DIR	ēδ ('yδ)	—	xēδ (xyδ)
OBL	mēδ (myδ)	tēδ (tyδ)	wāδ (wyδ)

TABLE 5.10: DEMONSTRATIVE ADVERBS

	Ich-Deixis	Du-Deixis	Er-Deixis	Interrogative
“here”, etc.	maδe (mδy)	taδe (tδy)	waδe (wδy)	
“here”, etc.	maδēδ (mδyδ)	taδēδ (tδyδ)	waδēδ (wδyδ)	
“here”, etc.	marθ (mrδ)	tarθ (trδ)	ōrθ ('wrδ)	kurθ (kwrδ)
“hither”, etc.	martsār (mrts'r)	tartsār (trts'r)	ōrtsār ('wrts'r)	kurtsār (kwrts'r)

In Manichaean and Christian texts, the gen.-dat. plurals *mēšan* and *wēšan* provided with the secondary plural ending *-t* are also employed: *mēšan-dlmēšan-d-ī*, *wēšan-dl wēšan-d-ī*.

These articles, just before they were lost in the latest stage of Sogdian, came to be proclitic and prefixed to nouns: *yi-mān-ī* (C y-m'ny; *yi* < *wya*, *əwī*, loc. masc. of Er-deixis) 'in the mind'.

The articles (weak demonstratives) themselves are occasionally proposed to the extended deictics, e.g. *xu xunax yrīw* 'that body'.

The adjective *wisp-* 'all' occasionally takes pronominal endings: G-D *wisp-ne*, I-A *wisp-na* (cf. *čan wispnā-č* 'from all' and *dan wispnā-δ* 'with all'), PL. NOM *wisp-e*, PL. G-D *wispēšan*.

The functional distinction between simple forms as articles, and the extended forms as demonstratives can be shown by their use in translations. Thus, of the 70 instances of simple forms in lines 1–88 of Pelliot Sogdien 5, the short Buddhist text translated from the Chinese *Dīrghanakha-sūtra*, only one case corresponds to the Chinese *qi* 'that'. In turn, of the 32 instances of extended forms, all but three render a demonstrative.

3.1.3.4 Interrogative, relative, and indefinite pronouns

In Sogdian every interrogative pronoun or adverb can be used as relative pronoun or adverb. The following interrogative pronouns are known in Sogdian: dir. *(ə)ke* 'who?' ((^o)ky), obl. *(ə)kya* ((^o)ky'), inst.-abl. *kanāč*, *čakanāč*, or *čakanā* (kn'č, ckn'č, ckn'); *(ə)ču* 'what?' ((^o)cw); *katār* 'which' (kt'r); *katām* 'which' (kt'm).

While *(ə)ču* is exclusively for inanimate antecedents, *(ə)ke* is the most commonly used relative pronoun, used for both animate and inanimate antecedents. Since *(ə)ke* is sometimes used also for oblique cases, it may rather be referred to as a relative particle that simply connects relative clauses to main sentences. Some Christian texts employ *kat* (qt) as a relative particle.

The following interrogative adverbs are also used as the relatives. They are: *(ə)kū* 'where' ((^o)kw), *kūrθ* 'where' (kwrθ), *kūtsā(r)* 'to where, whence' (kwrts'r, C qwc'), *kaða* 'when' (kδ'), *čāf* 'how much' (c'β), *čāfar* 'how much' (c'βr), *čānō* 'how' (c'n'kw, C c'nw).

The following forms serve as indefinite pronouns in Sogdian (see Sims-Williams 1986a): animate *ādē* ('dy), *ēdē* ('ydy) 'someone', and inanimate *ā(δ)č* ('δc, 'c), *ē(δ)č* ('yδc, 'yc) 'something'. The corresponding negative forms are *nēdē* (nydy) and *nēdč* (nyδc) with some variant forms. The negated forms always appear with the negative particle (see section 3.2.4), that is to say, doubly negated: *rī-šī nēdē nē pērt* 'and-him nobody not believes = Nobody believes him'.

3.1.4 Adpositions

In Sogdian both prepositions and postpositions are common.

3.1.4.1 Fused prepositions

The inherited prepositional elements are *par-* 'on, in', *č-* 'from', *δ-* 'with', while the obsolete preposition *(ə)t* ('t) < *(ə)tā* (reflected in the 1s, 2s pronouns *tāmā*, *tāfā* (see section 3.1.3.1) has been functionally replaced by *(ə)kū* 'to, towards', originally the older relative adverb *kū* 'where' (see Sims-Williams 1986b). *wasn* (wsn) 'for the sake of' and *witūr* (wytwr) 'until, up to' are not common but are not obsolete either.

What is peculiar to Classical Sogdian is that old prepositions never appear independently but are always fused with oblique pronominal or demonstrative elements. On this point see also the two ideograms found in the Ancient letters, 'LZK = *paru* and 'NwZK = *kū* 'where' which contain the pronominal element ZK = *-u*.

TABLE 5.11: FUSED PREPOSITIONS

	No Extension		<i>n</i> -Extension		<i>ēδ</i> -Extension	
	Er-Deixis < -w-	Ich-Deixis < -m-	Er-Deixis < -want	Ich-Deixis < -mant	Er-Deixis < -wēδ	Ich-Deixis < -mēδ
<i>par-</i>	<i>paru</i> (prw)	<i>parəm</i> (prm)	<i>pariwant</i>	<i>parimant</i>	<i>pariwēδ</i>	<i>parimēδ</i>
<i>c-</i>	<i>čon</i> (c'wn)	<i>čan</i> (cnn)	<i>čiwant</i>	<i>čimant</i>	<i>čiwēδ</i>	<i>čimēδ</i>
<i>δ-</i>	<i>don</i> (δ'wn)	<i>dan</i> (δnn)	<i>δiwant</i>	<i>δimant</i>	<i>δiwēδ</i>	<i>δimēδ</i>
<i>ku-</i>	(ə) <i>kū</i> ('kw)	—	<i>kiwant</i>	<i>kimant</i>	<i>kiwēδ</i>	<i>kimēδ</i>

The plural form of Er-deixis pronoun is also found: *pariwēšan*, etc. Those of Du-deixis are only sparsely attested with *ēδ*-extension: *čitēδ*, etc. Although fused with *δ-* 'with' *δiwant/δiwēδ* and *δimant/δimēδ* have genitive and locative functions. In Christian texts are attested *pariw* (pryw) 'on him', *čiw* (cyw) 'from him', and *δiw* (δyw) 'with him' of unclear origin.

Those which are construed with acc. case (of the light stem) are *par(u)*, *parm*, and (ə)*kū*, while *čan* (or *čon*) and *dan* (or *don*) govern the inst.-abl. case.

3.1.4.2 Postpositions and circumpositions

The most common postpositions are *sā(r)* (s'r, C s') 'toward, from', *parēw* (pr'yw) 'together with', and *pidār* 'for the sake of, because of' which themselves are often preceded by prepositions: *kūlčan . . . sār*; *čan . . . sār/pidār*, *dan . . . parēw*; *kū dēn sār* 'to the religion', *čan wənī sār* 'from him'.

3.1.5 Adverbs and interjections

3.1.5.1 Adverbs

Adverbs are indeclinables: *tīm* 'moreover', *yunēδ* 'immediately', *žγart* 'quickly'. Some adverbs take the optional ending *-ī* (< oblique ending): *rāmant(-ī)* 'always', *nūr(-ī)* 'today', *xwadkār(-ī)* 'alone'. Oblique case forms of heavy stem nouns behave as adverbs or postpositions: *midān-ī* (< *midān* 'middle') 'among'. They also appear in predicative position, *a(ṃ)sāk-ī* 'suitable' (C 's'q-y; *a(m)sāk* 'equipment').

Adjectives in the neuter accusative of light stems may function as adverbs, e.g. *šir-u* 'well', *žγ-u* 'very' (< *žγ-* 'severe'), *wisp-u* 'entirely'.

3.1.5.2 Interjections

In Sogdian the following interjections have been encountered: (a) *ō* ('w, M 'w:) 'O!', (b) *ai* (M 'yy, C 'y) 'hey!', (c) *nāy* (n 'y) 'lo, see!'.

3.1.6 Numerals

3.1.6.1 Cardinal numbers

TABLE 5.12: CARDINAL NUMBERS

1–9	11–19	10–90	100+	1000+	
<i>ēw</i> C <i>yō</i>	<i>yōnts</i>	<i>δas(a)</i>	<i>sat-</i>	<i>zār(u)</i>	‘1000’
<i>əδū</i> ~ (<i>ə</i>) <i>δw(a)</i>	<i>δwāts</i>	<i>wīst</i>	<i>δwēsate</i>	<i>pančzār</i>	‘5000’
<i>əθrē, sē</i>		<i>šēs</i>	<i>šēsat</i>	<i>βrēwar</i>	‘10000’
<i>čatfār</i>	<i>čatfārats</i>	<i>čatfars</i>			
<i>panč</i>	<i>pančalš</i>	<i>pančās</i>			
<i>uxušu</i>	<i>C xušarts</i>	<i>xušašti</i>			
<i>əβt(a)</i>	<i>C βtats</i>	<i>əβtāt</i>	<i>C βtasat</i>		
<i>əšt(a)</i>	<i>əštalš</i>	<i>C štāt</i>			
<i>naw(a)</i>	<i>nōts</i>	<i>nawāt</i>			

Units precede the decades: *əβt-wīst* ‘27’, *nawa-nwāt* ‘99’. Numbers close to the decade may be expressed by subtraction: *ēw kanpī pančās* ‘one less fifty’ = ‘49’. ‘Two’ distinguishes gender and case, thus masc. (*ə*)*δwa*, fem. (*ə*)*δwe*, gen.-dat. *δiβnu*. In 2, 7–10, forms without *-a* originate from proclitics.

Some numerals show an old genitive ending *-nu*: *uxušu-nu*, *dwāts-nu*, etc.

3.1.6.2 Distributives, multiples, and fractions

The distributive marker is *-ki* or *-kankī*: *βrēwar-kī* ‘by ten thousands’. Multiples are expressed with *yāwar* ‘time’ or *wāfar* ‘so much’: *uxušu yāwar uxušu* ‘six times six’, i.e. ‘36’. ‘Half’ is by *nēmē*: *əδū nēmē* ‘two and half’.

3.1.6.3 Ordinal numbers

Ordinal numbers: *əftam-* ~ *əftamīk* ~ *əftamčik* ‘first’, *δəβti* ~ *δəβtīk* ‘second’, (*ə*)*štīk* ~ *čaštīk* ‘third’. The ordinals higher than ‘fourth’ are derived from the cardinals with the suffix *-am(i)* or with *-mīk*: *čatfār-am* ~ *čatfār-mīk* ‘fourth’ *panč-am* ~ *panč-mīk* ‘fifth’, *əšt-ami* ~ *əšt-mīk* ‘eighth’, *C štīfār-wīst-mīk* ‘twenty-fourth’, *zār-mīk* ‘one-thousandth’, etc.

3.2 Verbal morphology

3.2.1 Stem formation

3.2.1.1 Present and past stems

Verb forms are based on two basic stems, present and past (in the following indicated by present/past). Historically they derive from the Old Iranian present stems and perfect participles in **-ta*. Synchronically the two stems may differ remarkably: *kun-* ~ *wan-* / (*ə*)*kt-* ‘do, make’ (< **kṛn(a)u-*, *kṛta-*). Other stems are suppletive, e.g. *wāβhwayt-* ‘say’, *āβarlāyat* ‘bring’, *x-l(w)māt* ‘be’. The productive past formant is *-āt* (occasionally shortened to *-at*) and some verbs have both forms, i.e. two past stems: *wayt-* and *wācāt*

corresponding to the present stem *wāc* ‘allow, let, send’ and *fašt-* (< OIr. **frašta-*) and *psāt* of *ps-* ‘ask’.

3.2.1.2 Imperfective stems

Certain verbs have a distinct imperfect stem. These are characterized by the preservation, or analogical extension, of OIr. augment. Diachronically, these are those verbs where the augment was preserved between original directional prefix and present stem, and later fused with the prefixal vowel to *-ā-*, *-ī-*: *patγōš* ‘hear’ > *pat-ī-γōš* (< OIr. **pai-a-gauša-*), *framāy* ‘order’ > *fr-ā-māy* (< OIr. **fra-a-māya-*). However, verbs with **ham-* > *an/m-* have the prefix *m-*: *anaxaz* ‘rise’ > *m-anxaz*, which was later extended to verbs with initial **ā-*: *āβar* ‘bring’ > *m-āβar*. Those verbs beginning with the Old Iranian preverb **abi-* and **us-* show imperfect stems with *-ā-* and *-ī-* respectively: *əβžay-* (< OIr. **abi-ǰawya-*) ‘increase (vi.)’ > *βāžay*, *sxwāy* (< OIr. **us-xwāhaya-*) ‘take up’ > *sīxwāy*. Other verbs, without an inherited prefix, such as *βar-* ‘bear’, show no alternation between present and imperfect stems.

The formation of the imperfect stem was later extended to stems without preverbs: *snāy* (< OIr. **snāya-*) ‘wash’ > *sīnāy*, *ufs-* (< OIr. **hufsa-*) ‘fall asleep’ > *wāfs*.

3.2.1.3 Derived voice stems

One finds numerous pairs of transitive-causative and intransitive-passive present stems which are inherited from OIr.: *xwēr* ‘feed’ vs. *xwar-* ‘eat’, *āβar* ‘bring’ vs. *āβir* ‘be brought’, *fnēs* ‘deceive’ vs. *fnēs* ‘be deceived’, *sōč/suyt-* ‘burn’ vs. *suxs-* ‘be burnt’, *kun-* ‘do, make’ vs. *kir-* ‘be done, be made’. These morphological causative and passive stems are largely obsolete with the possible exceptions of the so-called inchoative stems derived from the past stem with the suffix *-s* like *suxs-*: e.g. *γafs-* ‘labour’, *wiγrās* ‘be awoken’ (see Weber 1970: 127–228).

3.2.2 Nominal forms

3.2.2.1 Infinitives and verbal nouns

Sogdian possesses two kinds of infinitives, the past and the present, and verbal noun (see Yoshida 1979).

3.2.2.1a Past infinitive

This form consists of the past stem with the ending *-e* (-y, LS), or *-ø* (HS): *βayt-e* (*βaxš-/βayt-*) ‘give’, *γōβāt* (*γōβlγōβāt*) ‘praise’. On the syntax of the infinitive see section 5.3.5.

3.2.2.1b Present infinitives

While the formation and usage of the past infinitive remains constant in all types of Sogdian, the forms of the present infinitive vary considerably from text to text:

- (a) In Christian texts one often finds infinitives consisting of the preposition *par* and the present stem with an ending *-u* (LS) or without ending (HS), as in *par ōβar-u* ‘to give’, *par wēn* ‘to see’.

- (b) In Manichaean (and some Christian) texts the present stem with the ending *-i* (-y, LS) or without ending (HS) serve as infinitives, as in *βaxš-i* ‘to give’, *framāy* ‘to order’.
- (c) In Buddhist texts the stem enlarged by the suffix **-aka* is often attested, as in *wan-ē* (wn’y) ‘to do’ and *šmār-ē* (šm’y) ‘to think’.

3.2.2.1c Verbal noun

Mainly in Christian, Manichaean and other late texts one finds a verbal noun derived from present stems with a suffix *-āmantē* (-’mnty): *patfas-āmantē* ‘reading’. In Buddhist texts the present infinitive functions as a verbal noun.

3.2.2.2 Gerund and participles

3.2.2.2a Gerund

Gerund: Present stems with the suffix *-kya* (-ky’, LS) or *-kī* (-ky, HS), serve as a kind of gerund meaning ‘having ~ed, ~ing’: *βar-kya* ‘having brought, bringing’, *wāβ-kī* ‘having said, saying’.

3.2.2.2b Present participles

- 1 Present participle (a). The most productive participle in Sogdian is that derived with the suffix *-ēnē* (LS) or *-nē* (HS) from the present stem, as in *0βar-ēnē* ‘giving, giver’, and *pattāw-nē* ‘enduring’. Heavy stems often take light stem ending as well, as *pār-ēnē* (M p’rynyy) ‘nourishing, nourisher’. This participle is mainly employed as an attributive adjective or agent noun:

fritāt ke wānī sāt šire əktyē parwēžnē
 love REL the.G-D all good.G-D action.G-D nourish.PART
 ‘love which is the nourisher of all the good deeds’.

- 2 Present participle (b). Predicative and adverbial function is often played by the form which is derived by the *aka*-suffix, as in *nūδ-ē* (S nyδ’y) ‘sitting’, but *-ēk* or *-ēsk* (-yq, -ysq) in Christian texts. The Christian form always combines with the inflected forms of (*w*)*māt*, the preterite of *x-* ‘be’, and functions as a periphrastic expression denoting progressive past: *δārēk māt* ‘he was having’.
- 3 Forms provided with the old ending of present participle enlarged with the *aka*-suffix, i.e. *-antē* are also met with. But they seem to be fossilized and are no different from adjectives: *əžuw-antē* ‘living’, *βōδ-antē* ~ *βōδ-antč* (fem.) ‘fragrant’, *wēn-antē* ‘visible’.

3.2.2.2c Passive participles

- 1 Future passive participle. This participle is derived from the present stem by means of *-ī* or *-(i)čīk* (-y, -(y)cyk): *βīr-ī* (S βyry) ‘to be obtained’, *sumb-čīk* (M swmbcyk) ‘to be bored’.
- 2 Past (passive) participle. This participle is derived from the past stem with the *aka*-suffix, masc. *-ē*, fem. marked by *-č*: masc. *parθayt-ē*, fem. *parθaytč* (S prδ’ytk, prδ’yтч) ‘stretched’ (pres. *parθenč*). Forms derived from intransitive verbs have a perfect meaning while those from transitive verbs have a passive perfect meaning: *tayt-ē* ‘(having) entered’ (*tēs/tayt-* ‘enter’), *βīrt-ē* ‘(having been) obtained’ (*βīr/βīrt* ‘obtain’).

3.2.3 *System of tenses, aspects, and moods*

The Sogdian tense and aspect system comprises present, imperfect, preterite, and perfect. There are six moods: indicative, subjunctive, injunctive, optative, imperative, and irrealis.

The perfect tense is formed periphrastically by the past participle and the auxiliary verb.

Middle endings are almost all replaced by the active, only a few of them having survived. However, 3s imperfect and optative middle endings are formally so salient that they serve as starting-points for the analogical development of new paradigms.

Passive is also formed by the past participle and the auxiliary verb. (On the traditional description of the Sogdian verbal system see Gharib 1965.)

3.2.3.1 *Forms based on the present and imperfect stems*3.2.3.1a *Major conjugations*

The best attested personal endings of the above-mentioned moods are shown in the following tables. The two verbs selected to serve for the sample paradigms are *wan-* (wn-) ‘do, make’, imperfect stem *wan-* (LS), and *patγōš-* (ptγwš-) ‘hear’, imperfect stem *pat-ī-γōš-* (HS). As is the case with the nouns, heavy stem verbs frequently take light stem endings, and in fact, the heavy stem endings *-u* of 1s and *-i* of 2s imperfect are those of the light stems.

TABLE 5.13: LIGHT STEM CONJUGATION, *wan-* ‘do’

Present	Imperfect	Injunctive	Optative	Subjunctive	Imperative
1s <i>wan-ām</i>	<i>wan-u</i>	<i>wan-u</i>	<i>wan-el-ē-m</i>	<i>wan-ān</i>	
2s <i>wan-e</i>	<i>wan-i</i>	<i>wan-i</i>	<i>wan-el-ya</i>	<i>wan-a</i>	<i>wan-a</i>
3s <i>wan-ti</i>	<i>wan-a</i>	<i>wan-a</i>	<i>wan-e</i>	<i>wan-ā-t</i>	
1p <i>wan-ēm</i>	<i>wan-ēm</i>	—	<i>wan-ēm</i>	—	
2p <i>wan-θal-ta</i>	<i>wan-θal-ta</i>	<i>wan-θal-ta</i>	<i>wan-ēθ</i>		<i>wan-θal-ta</i>
3p <i>wan-and</i>	<i>wan-and</i>	<i>wan-and</i>	<i>wan-ēnd</i>		

TABLE 5.14: HEAVY STEM CONJUGATION, *patγōš-* ‘HEAR’

Present	Imperfect	Injunctive	Optative	Subjunctive	Imperative
1s <i>patγōš-ām</i>	<i>patīγōš-u</i>	<i>patγōš-u</i>	<i>patγōš-el-ēm</i>	<i>patγōš-ān</i>	
2s <i>patγōš-e</i>	<i>patīγōš-i</i>	<i>patγōš</i>	<i>patγōš-el-ya</i>	<i>patγōš-a</i>	<i>patγōš</i>
3s <i>patγōš-t</i>	<i>patīγōš</i>	<i>patγōš</i>	<i>patγwš-e</i>	<i>patγōš-āt</i>	
1p <i>patγōš-ēm</i>	<i>pat-ī-γōš-ēm</i>	—	<i>patγōš-ēm</i>		
2p <i>patγōš-θ(a)l-ta</i>	<i>pat-ī-γōš-θ(a)l-ta</i>	<i>patγōš-θ(a)l-ta</i>	<i>patγōš-ēθ</i>		<i>patγōš-θ(a)l-ta</i>
3p <i>patγōš-and</i>	<i>pat-ī-γōš-and</i>	<i>patγōš-and</i>	<i>patγōš-ēnd</i>		

The endings *-ām*, *-ān*, and *-āt* of the heavy stem are occasionally shortened to *-am*, etc. In the texts written in Sogdian script, one occasionally finds the ending (-’ty) (subj. 3s), which is a historical spelling for *-āt* (cf. OIr. **-āti* ~ *-āt*). The plural endings of the subjunctive and injunctive are not distinguishable; or rather, the subjunctive function is expressed by the forms of the injunctive.

The difference between 2 pl. endings $-θ(a)$ and $-ta$ are largely chronological; while earlier texts written in Sogdian script tend to prefer $-θ(a)$ ($-δ(ʿ)$), most Christian texts in Syriac script show almost always $-ta$ ($-tʿ$), invariable with both light and heavy stems and indifferent to moods. Some texts do distinguish between $-θ(a)$ and $-ta$ in that the former is used in modal forms other than the indicatives (see Gershevitch 1954: § 737).

In addition to the forms listed above, the imperfect stem with optative endings, called the optative imperfect, serves as a durative or iterative past, e.g. $wāpat-e$ 'it was falling' (< $ōpat$ 'fall').

MIDDLE ENDINGS

Survivals of the middle endings are: 1p $-ēman$ ($-ʿymn$); 2s optative $-ēš$ ($-yš$), which is the base of 2p opt. $-ēšθ(a)$ ($-ʿyšδ(ʿ)$); 3s pres. $-te$ ($-ty$), 3s opt. $-ēt$ ($-yt$); 3s impf. $-t(a)$ ($-t(ʿ)$). Only the 3s present form retains its original middle-passive force: $wēn-te$ 'is seen' vs. $wēn-t$ 'he sees'.

3.2.3.1b New minor and secondary conjugations

These are innovative and include the following:

1 IMPERFECT MIDDLE

This derived from the 3s impf. middle ending (e.g. $wan-ta$ 'he did', $ās-t$ 'he took'):

1s $-t-ut$, 2s $-t-i$, 3s $-ta$, 3p $-t-ant$, e.g. $ās-t-u$ 'I took', etc.

2 OPTATIVE MIDDLE

This is modeled on the 3s $-ēt$ to which new modal endings are attached; attested are:

1s $-ēt-u$, 2s $-ēt-a$, 3s $-ēt-e$, 1p $-ēt-ēman$, 3p $-ēt-ēnt$, e.g. $βīr-ēt-u$ 'I may obtain'. This category has hitherto been called 'precativē'.

3 IRREALIS

This is marked by the suffix $-ōt$ (LS) or $-t$ (HS) to which the new modal endings are attached:

1s $-(ō)t-u$, 2s $-(ō)t-a$, 3s $-(ō)t-e$, 2p $-(ō)t-ēšta$, 3p $-(ō)t-ēnt$, e.g. 2p $wərn-ōt-ēšta$ (C wrnwtyšt) 'you might believe'.

4 ĀZ-IMPERFECT

This new imperfect is a kind of durative past. It is formed by adding the imperfect endings to the 3s past form $āz$ 'was':

1s $-āz-u$, 3s $-āz$, 1p $-āz-ēm$, 3p $-āz-ant$ ~ $-āz-ēnt$, e.g. $δār-āz-ant$ 'they were holding'. These forms are particularly numerous in Christian texts and only three verbs are encountered in other texts: $əskw-āz$ 'remained', $šaw-āz$ 'was going', and $āz$ 'was'.

5 ĒK-(W)MĀT IMPERFECT

As mentioned above, this imperfect is only found in Christian texts. It consists of inflected forms of $(w)māt$, the preterite $x-$ of 'be', and functions as a periphrastic expression denoting progressive past, e.g. $dār-ēk māt$ 'he was holding'.

3.2.3.1c *Progressive or durative particle: skun, archaic -əštan*

Progressive force is provided by the particle *-skun* (-skwn) following the inflected form of verb: *šaw-ām-skun* 'I am going'. This particle is added to the present, imperfect, optative imperfect, and the *āz*-imperfect, but is incompatible with the subjunctive, injunctive, or imperative. Its form varies from text to text: *-əskun* > *skun* > *skən* > *sk*, and *-kən* > *ku* > *k* after the ending *-t* or *-ant*: *wāβ-ant-k* 'they are saying'.

The archaic alternative form *-əštan* (ʔštn) is attested in a few Buddhist texts (see Benveniste 1966).

3.2.3.1d *Future and past prospective particle: kām*

Future sense is added by the particle *-kām* (k'm), which is suffixed to present, and less frequently to the subjunctive, optative, injunctive, imperative, and the *āz*-imperfect, e.g. *βaxš-ām-kām* 'I shall give'. The form of this particle also varies from text to text: (k'm, k'n, C -q', -g'(n) after nasal consonants).

nē parēθān-kām . . . āpatē nē kunām-kām
not sell.SUBJ.1s-FUT . . . in.bondage not make.PRES.1s-FUT
'I shall not sell (her) nor put (her) in bondage'.

Of particular interest is the combination with *āz*-imperfect found in the manuscript C2, where the construction has past prospective meaning: e.g. *zīwart-āz-kā* (zwrt'zq') 'was about to return'.

3.2.3.2 Copula and existential verb

Sogdian presents a great variety of the stems of the verb meaning 'be'. Some are inherited forms and the others, mostly 3s and 3p, are formed analogically on the basis of the former. The copula also serves as the ending in the intransitive preterite.

Indicative forms (mostly inherited from OIr.):

1s *im* (ʔym), 2s *iš* (ʔyš), 3s *(ə)sti* ~ *əst* ((ʔ)sty) and *āz* '3s was' (ʔz; *āz*-imperfect); 1p *ēm* (ʔym) or *ēman* (ʔymn), 2p *-sθ(a)* (-sδ', as preterite ending) and *sta* (imperative), 3p *ant* (ʔnt).

Some examples of analogical formation:

3s *x-ači* (xcy), *i-či* ((ʔ)ycy), 2p *an-sθ(a)* (ʔnsδ(ʔ)), *išta* (ʔyšt' based on 2s ʔyš), 3p *x-ant* (xnt), *əst-ant* (ʔstʔnt).

Optative and subjunctive forms are also secondary:

opt. 3s *əst-e* (ʔstʔy), *x-e* (xʔy), *y-e* (yʔy), *e* (ʔy), 3p *əst-ēnt* (ʔstʔynt), *x-ēnt* (xʔynt), *ēnt* (ʔynt); subj. 1s *x-ān* (xʔn), 3s *əst-āt* (ʔstʔt), *x-āt* (xʔt), *y-āt* (yʔt), *āt* (ʔt).

In principle, the forms based on *əst-* denote existence while those based on *x-* and *y-* (reminiscent of Er-deixis vs. Ich-deixis) are mainly employed as a copula or auxiliary verb (see Weber 2000).

Notice that the article *(ə)xu* sometimes functions as a copula in texts written in Sogdian script:

xa āžun duyta əxu
'the child daughter COPU = the child is a daughter'.

The preterite forms (see section 3.2.3.3a) are based on the stem *(w)māt-* 'was, were' and *(ə)kt-* 'became'.

3.2.3.3 Forms based on the past stem

The Sogdian has the preterite and potentialis (see section 3.2.3.3c) which are based on the past stem followed by auxiliary verbs. The preterite of intransitive verbs consists of the past stem and the verb ‘be’ (see above on the copula) while the transitive verb takes *δār-* ‘have’ as an auxiliary.

3.2.3.3a Preterite

The following are the preterite conjugations of *β(ᵛ)-ləkt-* (LS) ‘become’, *ēs/āyat* (HS) ‘come’, *kun-* ~ *wan-ləkt-* (LS) ‘do, make’, *wēn/wēt* (HS) ‘see’.

TABLE 5.15: PRETERITE CONJUGATIONS

	vi./pass. LS ‘became’	vi. HS ‘came’	vt. LS ‘did, made’	vt. HS ‘saw’
1s	<i>əkt-im</i>	<i>āyat-im</i>	<i>əkt-u δār-ām</i>	<i>wēt δār-ām</i>
2s	<i>əkt-iš</i>	<i>āyat-iš</i>	<i>əkt-u δār-e</i>	<i>wēt δār-e</i>
3s	<i>əkt-i</i>	<i>āyat</i>	<i>əkt-u δār-t</i>	<i>wēt δār-t</i>
1p	<i>əkt-ēm</i>	<i>āyat-ēm</i>	<i>əkt-u δār-ēm</i>	<i>wēt δār-ēm</i>
2p	<i>əkt-asθ</i>	<i>āyat-asθ</i>	<i>əkt-u δār-θ</i>	<i>wēt δār-θ</i>
3p	<i>əkt-ant</i>	<i>āyat-ant</i>	<i>əkt-u δār-ant</i>	<i>wēt δār-ant</i>

The 3s of heavy stems does not take the auxiliary verb, while light stems show the ending *-i* (*-y*). Occasionally one finds 3s feminine forms in *-a*, and neuter forms in *-u*: *nīžt-a* (C *nyžt*) ‘she went out’, *xwart-u* (*xwrtw*) ‘it was eaten’. Note that the transitive preterite of the light stems differs from the heavy in that the former takes the ending *-u* (*-w*) before the auxiliary.

The 2p ending is sometimes *-išta*: *kt-išta* (C *qyšt*) ‘you became’ and *wēδār-išta* (C *wyđryšt*) ‘you saw’.

In late texts one sometimes finds the present stem instead of the past before the auxiliary *δār*: *pēz-δārt* (C *pyzd’rt*) < *pēz/pišt-* ‘hit’ (cf. *pyšt-w-δ’rt*). In some late texts intransitive verbs are construed with *δār-*: *βāw-δārt* (C *b’wd’rt*) ‘he approached’.

Another late feature is the fusion of the past stem and the auxiliary *δār-*: *kōāram* (*kδ’rm*, C *qθ’rm*) < *əktu-δāram* ‘I did’, *framāδārt* (C *frm’d’rt*) < *framāt-δārt* (M *frm’t δ’rt*) ‘he ordered’.

MODAL FORMS

These are obtained by conjugating the auxiliary verb: subj. *əkt-āi* (‘krt-‘i) ‘he may have become’, *ōžyast-δār-an* (M ‘wjjyst-δ’rn) ‘I may have settled’.

ELLIPSIS

When preterite verbs are coordinated, it is sometimes only the last verb which has an auxiliary verb, all the others consisting only of the past stems:

āraxsāt *əti* ... *βərtu* *əti* ... *patškwāt-δār-t*
‘he resorted to and brought ... and entreated ...’.

3.2.3.3b *Ergative construction*

The earlier ergative construction is found in the Ancient Letters and some other texts written in Sogdian script. Characteristically, the agent is expressed by the gen.-dat. of the noun or enclitic pronoun, while the auxiliary verb agrees with the object/patient, and thus formally corresponds to the intransitive preterite. Compare the following phrases in standard preterite and in formulaic archaic ergative, which occur at different places in the same text, both meaning '(my) father banished me':

standard (active) preterite:

rti-mī ǝxu ǝptri . . . γarmyān ǝktu-δārt
and-me the.NOM father.NOM . . . punished make-PRET.3s

ergative:

ǝwǝn ǝptre . . . γarmyān ǝkt-im
the.G-D father.G-D . . . punished become.PRET.1s

3.2.3.3c *Potentialis system*

The potentialis system expresses both possibility and anteriority of the action expressed by the verb (see Beveniste 1954). Intransitive (or passive) verbs take the auxiliary $\beta(w)$ -*lǝkt*- 'become' while the transitive *kun-* ~ *wan-lǝkt*- 'do, make'. Light stems end with *-u* (-w) or *-a* (-') whereas heavy stems take no ending (on the origin of the potentialis see Sims-Williams 2007):

(1) Possibility

nistu β-ām (nystw β'm) *tayta β-ant-sk* (C tʏt' bntsq)
'I can sit'; 'they are able to go into';

ǝktu wan-ān ('krtw wn'n) *parāyat β-ām* (prʔʏt β'm)
'I should be able to do'; 'I can arrive'.

(2) Anteriority

Another function of these forms is to denote anteriority in the temporal clause introduced by the conjunction *čānō* 'when', and in that sense it may be called the pluperfect:

rti čānō ǝxu sudāšn waytu-wan-a rti ǝxu barāman žγart manxaz
and when the S. said-POT and the brahmin quickly rose-IPF

'when/after Sudhāshn had said (it) the brahmin got up quickly'.

3.2.3.4 Forms based on the past participle

3.2.3.4a *Periphrastic perfect*

The forms of the periphrastic perfect consists of the past participle (i.e. the past stem plus *aka*-suffix) and the existential verbs. With transitive verbs the auxiliary is sometimes *δār*, similar to the transitive preterite, but the construction is much less frequent than those with the existential verbs which have passive meaning, and one may reasonably argue that the Sogdian perfect is passive by nature. Participles agree with the subject (or with the object of transitive verbs) in number and occasionally in gender, so that morphologically

speaking the construction behaves like a nominal than a verbal formation. Modal forms are obtained by inflecting auxiliaries. Some examples are:

pšātēt dār-ēm (C pš'tyt d'rym, < pš'y) 'we have cast (them) away'.
γūtč x-ān (γwtch x'n, < γw-) 'I (fem.) should have been conducting sin'.
āžitēt wmāt-and (M ʔjytytt wm'tnd, < ʔjy) 'they had been born'.

3.2.3.4b Periphrastic passive

1 Passive construction

A periphrastic passive is formed with the auxiliary verb *β(w)-ləkt-* 'become' and the past participle which often agrees with the subject in number and gender. The construction is also nominal rather than verbal. Some examples are:

βirt-ē (masc. sg.) *βw-e* (βyrt'y βwy) 'it should be obtained'.
əktč-a (fem. sg.) *β-āt* ('krtch β't) '(if) it should be done'.
sfrīt-ēt (pl.) *wβ-and* (M sfrytyt wβ'nd) 'they will be created'.

2 Agents

In the periphrastic passive, an agent is frequently expressed by (a) a noun phrase preceded by the preposition *čan* and *par* (non-human agent). However, (b) nouns or pronouns in the genitive-dative or oblique case are also used:

(a) *čan*:

xa əžγma . . . čan māxyān sār əktč-a β-āt
 the deceit . . . from M. from done.F become.SUBJ.3s
 'the deceit would be done by Makhyan'.

(b) Oblique case:

βžik rēž-ī par0ūt-ē . . . əkt-im
 evil desire-OBL burnt . . . become-RET.1s
 'I was burnt by evil desire'.

3.2.3.5 Voice and causativity

3.2.3.5a Passive and middle

Apart from the periphrastic passive discussed above, passive or middle voice is expressed morphologically, though not productive. The old force of middle voice is perceivable in the 3s present with the ending *-te*, which has a passive sense when used with transitive verbs, e.g. *wēn-te* 'is seen' vs. *wēn-t* 'he sees'.

Sogdian has inherited passive forms. Inchoative forms similarly function sometimes as passives (for both see section 3.2.1.3 above):

(a) Old passive:

rti xa zāy zpart kira
 and the ground pure was.made (< OIr. *krya-, cf. *kun-ləkt-* 'do, make')
 'the ground was made clean'.

(b) Inchoative:

rti čimēd piḍār δū waxs-ti
and from.this because.of *dhū* is.said (cf. *wāβ/wayt-* ‘say’).
‘for this reason it is called *dhū*’.

3.2.3.5b Causativity

No special means for expressing causativity is found in Sogdian. Some inherited present stems have causative meaning (see section 3.2.1.3):

rti ō barāman xwart ətī čašant xwēr
and the.ACC Brahmin food and drink feed.IMPF.3s
‘he made the Brahmin eat (= fed him with) food and drink’.

3.2.4 Negation

Sentences are negated by placing either *nē* (ny) or *nā* (n’) before finite verbs; the both are masked by an ideogram L’. The preterite differs from the periphrastic perfect in that *nē* precedes the main verb in the former and the auxiliary in the latter: *nē āyātēm* ‘We did not come’ vs. *xrāmtē nē wmat* ‘he had not proceeded’. While *nē* negates proposition, *nā* is a prohibitive particle: *nē wāβ-ām kām* ‘I shall not speak’, *nē niyāt-ḍār-t* ‘he did not take’, *nā ps-a* ‘don’t ask’. Sometimes negation is reinforced by the indefinite pronoun or negative indefinite pronoun, e.g. *ēdč maryārt nē sumbt* ‘he does not bore any pearls’. *nē* and *əsti* combine to give *nēst* (nyst) ‘is not, there is not’.

The negation of the imperfect differs from this pattern (see Sims-Williams 1996b). It is formed not with the imperfect, but with either the present indicative or the injunctive, occasionally preceded by the enclitic element *-β(i)* (*-β(y)*):

positive:	negative:
<i>rti ḍūrī zāy šaw-a</i> (impf.)	<i>rti-βi nē ḍūrī zāy šaw-t</i> (pres.)
and far land went	and-ENCL not far land goes
‘he went far’.	‘he did not go far’.

3.2.5 Synopsis of the system of tense, aspect, and mood

The following table presents a tentative synopsis of the Sogdian verb system, summarizing the discussion above. The system distinguishes non-past and past, imperfective and perfective as well as indicative and non-indicative forms, showing an over all balance. Non-indicative forms are perfective in their default meaning. The imperfective forms of the non-past and the past forms are formed by adding *skun*, which is optional in the cases of the *āz*-imperfect and optative imperfect.

The *āz*-imperfect is largely confined to some Christian texts, except for 3s forms; it corresponds to the *ēk-(w)māt*-imperfect of other Christian texts.

The perfective use of the present tense also has future meaning, which can be reinforced by the particle *kām*. Note that the periphrastic perfect and passive are still nominal constructions and are not included in this synopsis.

TABLE 5.16: SOGDIAN SYSTEM OF TENSES AND MOODS

		Perfective		Imperfective
<ind.>	non-past	present present	future meaning + <i>kām</i> ; subj. (+ <i>kām</i>)	pres. + <i>skun</i>
	past	imperfect preterite	<i>āz</i> -impf. + <i>kām</i> (past prospective)	imperfect + <i>skun</i> opt. impf. (+ <i>skun</i>) <i>āz</i> -impf. (+ <i>skun</i>) <i>ēk</i> -(<i>w</i>) <i>māt</i> impf.
<non-ind.>	non-past		Past	
	subjunctive (+ <i>kām</i>)		preterite subj.	
	optative (+ <i>kām</i>)		preterite opt.	
	injunctive (+ <i>kām</i>)		pret. inj.	
	irrealis imperative (+ <i>kām</i>)			

4 SYNTAX AND USE OF FORMS

4.1 Use of cases

4.1.1 Light stems

The use of the six cases of light stems is comparable to those of the Old Iranian counterparts and Sogdian follows the nominative-accusative pattern. Notice that the inst.-abl. case forms never occur independently, being always preceded by the prepositions *čan*, *čon* (cnn, c'wn) 'from' or *dan*, *don* (δnn, δ'wn) 'with'. However, the original situation is obscured by several changes that Sogdian underwent; even in the most archaic texts the tendency can be observed to confuse genders and case endings and to generalize the masculine nominative form for all the other functions. Some examples follow:

(a) Genitive-Dative:

xu zātē əwən əptr-e andēk dāre
the son the.G-D father-G-D custom have.OPT.3s
'the son should possess his father's characteristics'.

(b) Nominative and Instrumental-ablative:

xu xar-i dan xar-a parēw ranft
the donkey-NOM with donkey-I-A together fights
'a donkey fights with (another) donkey'.

(c) Locative:

rti mana əwī xwārant əpkāš-ya tēs
and my the.LOC right side-LOC entered
'it entered my right side'.

4.1.2 Heavy stems

Employment of the direct and oblique cases is complicated and the most consistent feature of the oblique case is the purely negative one. It never occurs in any syntactic context which would require the nominative of a light stem (where heavy stems properly have the direct case). This is partly due to the multiple origin of the ending *-ī*, which goes back to **-ya* (i.e. the endings of loc. sg. m., gen.-dat., inst.-abl., and loc. f. sg.) (see Sims-Williams, 1982: 72–73).

The situation is ameliorated to some extent by the frequent employment of the articles, which may precede both light and heavy stems indiscriminately, and the fact that the preposition never collocates with the article is in consonance with this case-marking function of the article.

Another possible way of compensating for the poverty of Sogdian case-marking is the use of *xēpθ* ‘one’s own’ as a marker of the genitive:

xunē patkarē xēpθ wiδβāy
that statue its.own explanation
‘the explanation of that statue’.

4.1.3 Direct object and animacy hierarchy

- 1 The forms assumed by the direct object of transitive verbs may briefly be surveyed. Usually it is represented by the accusative or direct case:

ərti ō xšēšpat βay-u wadēd nīšēdand
and the.ACC Splenditenes.DIR god-ACC there seated.3p
‘there they seated the god Splenditenes’.

Sometimes, through syncretism discussed above a nominative form is substituted for the accusative:

xurn-i āse-skun
blood-NOM she was taking
‘she was taking blood’.

- 2 Partitive sense is expressed by the preposition *čan* ‘from’:

nē čan yātē xwart-δāre
not from meat you ate
‘you did not eat meat’.

- 3 As in other languages a verb meaning ‘fear’ takes the ablative object:

əžyu pačkwērt-skən čan xūβ
very much fears-DUR from lord
‘he fears the lord very much’.

- 4 The two highest in the ‘agency hierarchy’, ‘me’ and ‘you’ (sg.), show special forms for the direct object *tāmā* and *tāfā*:

əzu tāfā wēnām-kām
I you.2s see. 1s shall
‘I shall see you’.

5 Direct object marking

Regarding the animacy hierarchy, the latest stage of Sogdian, the language of the Christian Sogdian ms. C5, employs the oblique case for marking the direct objects which are both human and definite:

fšam-dārt xēp0 zātē-ī
 send.PRET.3s his.OWN son-OBL
 'he sent his son'.

Compare the ordinal direct case for the accusative function: *kū zātē sā* 'to the son'. One even finds this use applied to the oblique form of the 1s pronoun: *manā-y-ī* (C mn`yy) 'me (direct object)'. On the unique oblique case form which has developed in the language of C5 see section 7.2.1.2.

4.2 Use of verb forms

4.2.1 Possession

Possession is expressed by the gen.-dat. (or obl.) of the noun or (suffixed or independent) pronoun and the word for existence: (a), (b). Though not common, *ḍār* 'have, hold' is also known to serve as the verb: (c).

- (a) *rti-βay tawa īnč əsti əti rinčak zākt*
 and-lord! you.OBL woman is and little children
 'and, O lord! you have a wife and little children'.
- (b) *rti-mī nēst wadu nē zātē nē ḍuyta*
 and-me is.not wife not son not daughter
 'I have no wife, son, or daughter'.
- (c) *rti padka xači əti əxu zātē əwən əptre andēk ḍāre*
 and law is that.CONJ the son the.G-D father.G-D manner have.OPT.3s
 'it is customary that a son should possess his father's manner'.

4.2.2 Use of tense and aspect

The basic tense distinction in Sogdian is between the past and non-past. The former is represented by the preterite, imperfect, optative imperfect, *āz-* imperfect, and *ēk (w)māt* imperfect, the latter by the present tense. The present and past perfect forms describe the state reached through an action. The future is just a subcategory of the non-past tense and expressed by the present tense form occasionally followed by the particle *kām*. The future meaning is also conveyed by the subjunctive forms.

In Sogdian there is also the opposition between perfective and imperfective (or durative) aspect. In the non-past tense, the former is represented by the plain present form while the durative is marked by the particle *skun*. In the past tense, the imperfect and preterite represent the perfective aspect, while *āz-* imperfect, *ēk (w)māt* imperfect, imperfect followed by *skun* and optative imperfect, occasionally accompanied by *skun*, represent the imperfective (or durative) aspect.

It should be mentioned that the term 'imperfect' is used because the forms go back to the Old Iranian imperfect, and does not refer to the actual function. In the last stage of Sogdian, the form with *skun* became increasingly common, and virtually replaced the simple present form, similar to the development of Early New Persian *īamē* to

the imperfective marker *mī-*, which became obligatory in the indicative (see Telegdi 1938: 222–227). The bleaching of durative meaning is paralleled by the gradual reduction of form from (*ə*)*skun* to just *sk* or *k*.

4.2.2.1 Examples

(A) Plain form

1 Present:

ər̥ti sāt wispu ark γərβ-ām
and whole all work I.know
'I know every kind of work'.

2 Imperfect:

r̥ti xu nawē nyiγōdan pat[ī]menč
and he new garment he.wore
'he wore a new garment'.

3 Preterite:

əpyār xwaβnu wēt-/[dār-]ām
last.night dream see-RET.1s
'last night I saw a dream'.

4 Optative as imperfect:

čan əštana əxšīβti rōš[-e]
from breast milk flowed
'milk was flowing from the breast'.

5 *āz-* imperfect:

par ēw žamnu əxu bagaβām əskw[-āz] əwī rājgarī mazēxak kanθī
in one time the Bhagavan stayed the.LOC Rājgarha great city
'one time the Bhagavan was staying in the great city of Rājgarha'.

6 *ēk (w)māt* imperfect:

γānt ēšt žātēt b[-ēk māt]-ant
these thing.PL said.PL being were
'these things were being said'.

(B) Forms with *skun*

1 Present:

r̥ti əču wānō zārī rāy-e[-əskun]
and why thus miserably you.cry-DUR
'why are you crying so sorrowfully?'

2 Imperfect:

ō buti t[ī]kōš-ant[-əskun]
the.ACC Buddha they.saw-DUR
'they were looking at the Buddha'.

3 Optative as imperfect:

par patēd šaw[-ē]-nt[-askun]
 for begging they.went-DUR
 ‘they were going for begging’.

4 āz-imperfect:

šir andōxčnāk eti šmārēkēn šaw[-āz-skun]
 very sorrowful and thoughtful was.going-DUR
 ‘he was going in great anxiety and deep in thought’.

4.2.2.2 Imperfect and preterite

The difference between the imperfect and preterite is somewhat comparable with that of the German Präteritum and Perfektum (see Paul 1997). Thus, in one narrative text, the *Vessantara Jātaka*, the imperfect is used in the narrative part whereas the preterite is preferred in direct speech. The story begins with the following sentences:

mā0 patīškway wiyaša eti βaya xwatāw pāruti γrān əktim
 thus say.IMPF.3s rejoice! COMP lord! king! for.COMP pregnant become.PRET.1s

paraw mā0 eti əpyār xwaβnu wēt-δārām
 because thus COMP last.night dream see-PRET.1s
 ‘(the queen) said (imperfect): “rejoice, O lord king!, for I have become (preterite) pregnant, because last night I saw (preterite) a dream”’.

In later texts the distinction became less and less clear, and one sometimes finds texts where imperfect and preterite forms alternate without any semantic distinction. For example, Christian text C5 prefers the preterite, and even in the narrative part one often finds the preterite forms:

ət γōž-δārt piδār et nīpēs . . .
 and request.PRET.3s tablet and write.IMPF.3s . . .

ət wiδās-δārant wispu ādē
 and be.surprised.PRET.3p all someone
 ‘and he requested (preterite) the writing table and wrote (imperfect) . . .
 and all the people got surprised (preterite)’.

The merger of the two tenses may also be inferred from those preterite forms, though not common, which show the augment: *z-ī-wastant* ‘they returned’ (< *zward/zwast-*) and *w-ā-māt* for *wmāt* ‘was’.

To note, originally, the preterite tense functioned as a ‘perfect’, but when it lost its original function the perfect sense came to be supplied by the newly formed periphrastic perfect tense.

4.2.3 Use of modal forms

4.2.3.1 Functions

- 1 The imperative is used for command and prohibition, the latter being preceded by *nā*.
- 2 The optative mood has several functions:
 - (a) In main clauses it denotes exhortation, comparable with the imperative, likewise with prohibitive *nā*.
 - (b) It also expresses hypothetical sense, extended to parabolic use.
 - (c) In subordinate clauses it mainly conveys hypothetical sense and is typically used in purpose, (indefinite) relative, and conditional clauses.
- 3 The subjunctive has the following functions:
 - (a) It is used in both main clauses and in temporal clauses to express future.
 - (b) It is used in purpose, (indefinite) relative, and conditional clauses for hypothetical action.
 - (c) It also denotes injunction to the third person, and when negated *nā* is employed.
- 4 The usage of the injunctive mood has been studied by Sims-Williams whose findings are summarized as follows (Sims-Williams 1996b: 183):
 - (a) It expresses the simple past in negative sentences.
 - (b) It is used in conditional and purpose clauses.
 - (c) It expresses politeness.
- 5 The irrealis mood denotes irreal or counterfactual situations in both main and subordinate clauses. Note that the forms identified by Gershevitch as 'irrealis ii' appear to be forms of *wmāt* followed by modal forms of the auxiliary: *wmāt-e* 'might have been' (wm't'y, opt. 3s) and *wmāt-u* 'I would have been' (wm't'w, inj. 1s).

4.2.3.2 Merger of modal categories: non-indicative

It is noteworthy that the subjunctive, optative, and injunctive are all used in subordinate clauses with verbs which also take the infinitive construction. This situation may suggest that the three categories were becoming confused and were merging into a single modal category, which may be referred to as non-indicative. For this tendency, note also the new, secondary modal endings, 1s *-u*, 2s *-a*, 3s *-e*, which are the endings of the injunctive, subjunctive, and optative, respectively.

4.2.3.3 Modal particles

4.2.3.3a Hypothetical particle *xāt*

xāt (*x't*), 3s subj. form of *x-* 'be', is sometimes added to inflected forms to give or reinforce hypothetical meaning: subj. *mir-ān-xāt* '(if) I should die', preterite *k0ār-ant-xāt* 'if they have done'.

4.2.3.3b Hypothetical particle: *-(a)n*

In the texts written in Sogdian script one finds an enclitic element *-(a)n* which reinforces the hypothetical meaning of the sentence. The verbs are in the present indicative, subjunctive, optative or injunctive, that is, those denoting non-past events:

rti-n pīšt kaδ uttekin wānō mān β-āt
and-N but if U. thus mind be.SUBJ.3s

əti əma čatta waδu nē dār-āt-kām pāruti-š-an wāč-āt-kām . . .
CONJ the C wife not have.SUB.3s-FUT but-her-N divorce.SUB.3s-FUT
'however, if Uttekin should think that he would not keep Chatta as his wife but would (rather) divorce her . . . '.

4.2.3.4 Examples

Here follow examples of modal forms (see also Complex sentences, section 5.3):

1 Imperative:

əti xēpθ brāt par žəyma nā dōzām
and own brother on lie PROH condemn.IMPR.2s
'do not condemn your brother with lies'.

2 Optative:

rti-šu par šir paδiβārčyā βarya
and-it on well enquiry bring.OPT.2s
'You shall make extensive investigation on it'.

rti kaδ əxu wēšparkar xe rti-šī Orē rīt əskwe
and if he Weshparkar be.OPT.3s and-him three face exist.OPT.3sg
'if he should be Weshparkar, he should have three faces'.

3 Subjunctive:

rti əxu čan mana tans mirant-kām rti mana mazēx əktānī βāt
and he frommy affliction die.FUT.3p and me great sin be.SUBJ.3s
'they will die because of affliction to me and I will have (committed) great sin'.

rti ādē . . . xwatāmē nā əskwāt
and someone . . . selfish PROH remain.SUBJ.3s
'One should not be selfish'.

4 Injunctive:

mana wānō wāβ čan tawa rēž βēk-ti ēč nē raxnu əkte
me thus said from your desire outside-COMP something not venture.INJ.1s to.do
'he said thus to me: "I would not venture to perform anything except your wish".'

5 Irrealis:

rti-šu tayu kaδāč wā-xūpt nē frēštu-kunōte
and-him you not.at.all so-good not send-POT.IRR.2s
'you would never been able to send him such a good (letter)'.

6 Hypothetical particle *-(a)n*:

γarf wātδār βōt-kām ke-t-(a)n . . . šawāt rti-n . . . γrān arn γavāt
 many beings be.FUT REL-COMP-*n* . . . go.SUBJ and-*n* . . . heavy sin commit.SUBJ
 ‘there will be many beings who would go . . . and would commit great sins’.

4.2.4 Modal constructions

4.2.4.1 Necessity construction

Necessity is expressed by the auxiliary verb *sāč* ‘be fitting’ followed by the infinitive. The construction is impersonal and the agent, if expressed, is encoded by the enclitic pronouns or nouns in G-D or Obl.:

rti mana sāčt ite čan γara kū γaru
 and I.Obl. is.fitting to.go from mountain to mountain
 ‘I must go from mountain to mountain.’

4.2.4.2 Desiderative construction

Desire is expressed by the verb *kām* similarly followed by the infinitive. In Christian texts, *par* infinitives are used in a way similar to the past infinitives:

ke kāmēnt par-patxwāy wənī
 who(REL) desire.OPT.IMP.3p *par*-kill.INF him
 ‘who were wishing to kill him’.

4.3 Concord and agreement

4.3.1 Light and heavy stems

Naturally light stems follow the same agreement patterns as observed in Old Iranian. In contrast with the light stems, heavy and contracted stems are very relaxed with regards to the rules of agreement. To a large extent, the suffixes *-ī* (obl.) and *-t* (plur.) are treated as optional, being often omitted where clarity is unimpaired. When light and heavy stems are combined to constitute a phrase, each member behaves according to its own syntactic function, although the preference of nominative or direct case for every function is often encountered:

ō tawa roxšnu čašmu
 the.ACC you.OBL light.ACC eye.ACC
 ‘your light eyes’.

par šukča naftča zāy
 on dry.F.ACC wet.F.ACC ground.F.DIR
 ‘on the dry (and) wet ground’.

čan tawa zārčanūk βaya
 from you.OBL merciful.DIR god.I-A
 ‘from you, merciful God’.

əwī sāt tər xaki tamya
 the.LOC all.DIR harsh.NOM hell.LOC
 ‘in the whole harsh hell’.

4.3.2 Group inflexion and nouns after cardinal numbers

4.3.2.1 Group inflexion

One also finds the phenomenon of ‘group inflexion’ where only the last in a series of (usually asyndetically coordinated) words is inflected: *čan anxar paxarē-t-ī βēk* ‘with the exception of fixed stars (instead of expected pl. obl. form *anxar-t-ī*) and planets’. The rule of group inflexion is not compulsory either. Thus, in combination with heavy stem adjectives and nouns one finds all the following examples in C2 (see Sims-Williams 1982: 69). The tendency is that determining adjectives are uninflected while descriptive adjectives agree in number:

<i>βayānīk a(ṇ)βarz</i>	‘divine visitation’	(dir. + dir.)
<i>γarf γampan-ī</i>	‘much labour’	(dir. + obl.)
<i>sāt δēw-t</i>	‘all demons’	(dir. + dir. pl.)
<i>xēpθ ančaman-t-ī</i>	‘one’s church’	(dir. + obl. pl.)
<i>a(ṇ)βasč-ī pačukān-ī</i>	‘constant prayer’	(obl. + obl.)
<i>rāzyān-t wāxš-t</i>	‘symbolic words’	(dir. pl. + dir. pl.)
<i>mazyātar-t ēš-t-ī</i>	‘greater things’	(dir. pl. + obl. pl.)
<i>pēnamčīk-t-ī ēš-t-ī</i>	‘former things’	(obl. pl. + obl. pl.)

4.3.2.2 Nouns after cardinal numbers

Both singular and plural forms follow cardinal numbers, while light stem nouns sometimes take numerative forms in that position:

čafār δβar-a (num.) ‘four doors’ vs. *δwāts δβar-ta* (pl.) ‘twelve doors’;
əβt paxarē (sg.) ‘seven planets’ vs. *əβt paxarēt* (pl.) ‘id’.

4.3.3 Agreement of the subject and the predicate

The number of the predicates usually agrees with that of the subject. However, 3s, which is the unmarked number and person, occasionally appears instead of the expected plural form: *əžyantt xartēt əsti* ‘the envoys have gone (sg. *əsti* instead of *əstant* or *xant*)’.

The relaxed nature of agreement in Sogdian verbs may also be exemplified by the 3s fem. forms of the intransitive preterite of light stem verbs. One sometimes finds special feminine forms ending with *-a* (‘). But ordinary forms without gender distinction are commonly encountered:

xu nafša pyamt-a əti əkt-a anyatē əpu xwēč
 the Nafsha was.healed.F and became.F whole without pain
 ‘Nafsha was healed and became whole and without pain’.

rti īnč βēksār nižt-i
 and woman outside went.out
 ‘the woman went out’.

5 SYNTAX OF SIMPLE AND COMPLEX SENTENCES

The syntax is the least studied field in the study of Sogdian grammar. As yet there exists no general account of Sogdian syntax. On some topics see Heston 1976.

5.1 Order of constituents

5.1.1 Noun phrase structure

In Sogdian the head noun of noun phrases is preceded by the article, demonstrative, possessor noun, and adjective:

- | | | | | |
|-----|------------------------------|----------------------|---|----------------------------------|
| (a) | <i>ō tawa roxšnu čašmu</i> | ‘the your light eye’ | = | ‘your light eye’ |
| (b) | <i>yunē mana xēpō žyāvar</i> | ‘this my own heart’ | = | ‘this heart of mine’ |
| (c) | <i>par xēpōāwandī framān</i> | ‘on owner’s order’ | = | ‘according to the owner’s order’ |
| (d) | <i>xu xunax yrīw</i> | ‘the that body’ | = | ‘that body’. |

Similarly, cardinal numbers and words for unit precede nouns with this order:

nōts kapčakk nā-art yantam
‘19 kapchak of non-ground wheat’.

Nouns designating rank, title, or category usually follow the proper nouns: *smēr yari* ‘Mount Sumeru’, *uparatt šamani* ‘Monk Uparatt’. However, counter examples do exist: *šamani yansyan* ‘Monk Yansyan’, *wispišē sudāšn* ‘Prince Sudhāshn’.

A prepositional phrase modifying a noun follows it:

ō zernēnē wartan dan anyatč parštāk
‘the golden chariot with all the equipment’.

When an appositional phrase is governed by a preposition, the latter is repeated before each member (see Sims-Williams 1973):

čan širβaxč ətī ətapsarē čan farnxunt zātēt
lit. ‘from Shirvakhch and Ǝstapsarē from Farnkhunt’s (two) sons’

5.1.2 Clause structure and word order

5.1.2.1 Complementizer

A syntactic feature peculiar to Sogdian (and for that matter to Bactrian as well) is that each clause (both main and subordinate) contains an enclitic complementizer (COMP): *(ə)t(i)*, *-uti* (= ‘PZY, ZY). This stands in the second position from the beginning, to which other enclitic elements of the sentence are added (see Sims-Williams 1985b, Yakubovitch 2005). Complementation includes the most frequent *(ə)rti*, which marks the beginning of a clause and etymologically consists of the adverb *r-* (see Khotanese *rro* ‘also’) and *-ti*.

The same is found in relative clauses, where relative pronouns or adverbs are followed by the complementizer. Similarly, direct quotations are also treated as independent clauses, where the second position is occupied by the complementizer (see Weber 1971):

rti-šī *xā xwatēn mā0 patīškway mā0 ətī* *βaya mēnu*
and.COMP-him the queen thus said thus COMP lord! I.thought

čan xwēr-βayī ətī ətβt čintāman ratni nīži
from sun-god COMP seven cintāmaṇi jewel it.went.out
‘the queen said to him: “O Lord! I thought thus: <from the sun god went out the seven cintāmaṇi jewels>”’

As shown by *čan xwēr-βayī ətī* above, the element standing before the complementizer is not always a single word but one syntactic unit, whereby the vocative form is not counted.

However, this feature was declining. Thus, in such a late text as Christian manuscript C5 the complementizer is almost lost, the only vestiges being the *-t* element found in the compound conjunctions *kat* (C qt) ‘that, who (rel.)’ and *pāt* (C p’t) ‘for (conj.)’, and interrogative particle *čut(i)* ‘est-ce-que’.

5.1.2.2 Enclitics

Even as a Middle Iranian language. Sogdian still observes the Wackernagel’s law, according to which enclitics occupy the second position in the sentence. As stated above, in classical Sogdian the second position is usually occupied by the complementizer *(ə)t(i)* and other enclitics are added to it:

pār-ti-šī *xu wīnā žani framātāre*
but/for-COMP-him the lute to.play you.ordered
‘but you ordered him to play the lute’.

When in a very few cases one pronominal suffix is attached to another, the first person precedes the second and the third: *ərti-m-f-(a)č* (‘HRZYmβc) ‘to me from you’, *rti-m-šu* (‘HRZYmšw) ‘and him for me’.

Apart from the enclitic pronoun, several other enclitic elements occur in Sogdian. They are prepositional elements (*-c*, *-δ*, *-t*), *-βay* ‘O lord!’, and *-n* (hypothetical particle), *-β(i)* (perfective particle employed in the negated imperfect sentence), etc. These enclitics follow pronouns:

mā0 ətī-š-n *uβyu šamani Yansyan xwati ətī . . . par kāmē xawāt*
thus COMP-her-PART both monk Yansyan himself and . . . by desire hit.SUBJ.3s
‘so that either monk Yansyan himself or . . . may hit her as he likes’.

5.1.2.3 Word order

5.1.2.3a SOV

Sogdian is basically an OV language where heads follow the dependent elements. Thus, the basic structure of the Sogdian sentences consists of SOV. However, Sogdian is far from being a consistent OV language and attests a number of counter-examples. The unmarked order is (a) SOV, but (b) the verb initial order is also often encountered. Extended constituents, including subjects and objects, tend to follow shorter constituents: (c), (d). Sometimes, even a part of a longer constituent is extraposed toward the end of a sentence: (e). Standards of comparison precede the adjective: (f). Adverbial elements tend to stand before verbs or objects of verbs: (g) On the orders of the adposition and the auxiliary verbs see sections 3.1.4 for adpositions and 3.2.3.3–4 for auxiliaries.

- (a) *erti xu maryārtī xēpθāwand . . . ō satu dīnār zern tōž*
and the pearl's owner . . . the 100 denar gold paid
'and the owner of the pearls paid the hundred gold denars'.
- (b) *rti frāyāz əxu wispišē Sudāšn wispu ādē βayte*
and began the prince Sudāšn all things to.give
'Prince Sudāšn began to give all the things'.
- (c) *rti-šī ēs čan patīčsār ēw brāman čanstī šāw əksi əti . . .*
and-him came from in.front one brahmin very black lean and . . .
'one very black, lean, . . . brahmin came towards him'.
- (d) *rti əxu wēn γarf bikšē nāf əke-ti par patēd šawēnt-əskun*
and he saw many mendicant people who-COMP for begging were.going-DUR
'he saw many mendicant people who were going for begging (alms)'.
- (e) *ō nāft martsār tūx əfšām əti arkkarēt*
the people.PL hither quickly send.IMPV and workers
'send people and workers hither quickly'.
- (f) *čan šakara nāmartar*
from sugar sweeter
'sweeter than sugar'.
- (g) *rti yonēd βōk βīrant-kām*
and immediately salvation they.find-FUT
'and they will get salvation immediately'.

5.1.2.3b Counter-examples

Since Sogdian is not a consistent OV language, counter-examples are not uncommon. But some of them are due to the prototypes. The bulk of the Sogdian texts are translations from other languages, and the word order often reflects those of the original language or texts. This is particularly evident in translations from the Syriac New Testament. See the following example. Notice, however, that the deep-rooted order of possessor + possessed (*βar-ēt-ī žamnu* 'fruits' time') is never inverted:

Syriac

kd dyn mt' [zbn' d- p'r'] šdr l-[bdw-hy] lwt plḥ'
when but reached [time of-fruit] sent to-[servants.his] towards farmers

Sogdian

pīšt čānō parāyat [βar-ēt-ī žamnu] fšamdārt [xēpθ βantētī] kū βāγ-dārēt
but when reached [fruits' time] sent [his servants] to farm-keepers
'however, when the time of fruits (i.e. harvest) came he sent his servants to the farmers'.

Naturally, emphasized and topicalized elements tend to be placed towards the beginning of sentences:

rti xa uxušu βrēwar īnč mart mart čan əštana əxšifti rōše
and the six 10,000 woman each each from breast milk flow.OPT.IMPV.3s
'as for 60000 women, milk was flowing from each woman's breast'.

5.2 Questions

Sogdian does not seem to have a special device for forming interrogative sentences. Interrogative pronouns and adverbs follow the ordinary word order: (a), (b). However, since

the interrogatives are the topics of the statements in their very nature they tend to be placed at the beginning of sentences: (c). Yes/No questions are formed either by an ordinary order (possibly with different intonation): (d), or by putting *čut(i)* at the beginning of a sentence, of which -t(y) originates from the complementizer. Disjunctive sense is sometimes emphasized by the phrase *katār (əti) nē* ‘or not?’ placed after the sentence: (e). Rhetorical questions make use of the particle *p(u)nūkar*: (f).

- (a) *tayu pērnāstar ču əktya k0āre*
 you before what deed do.PRET.2s.
 ‘What deed did you do before?’
- (b) *xa əspiyi-ti kutsār wāčām*
 the horses-COMP where I.send
 ‘Where shall I send the horses?’
- (c) *ərt-šu čakanāč piḏār marγārt nē framāye suḑte*
 and-him whence because pearl not order.2s bore.INF
 ‘Why didn’t you order him to bore pearls?’
- (d) *nē tkōš*
 not see.INJ.2s
 ‘Don’t you see?’
- (e) *čuti xa zākt tadēḏ āyatant katār-əti nē*
 whether the children there came or-COMP not
 ‘Have the children arrived to you or not?’
- (f) *əčuti pnūkar tawa wānō nē patyōšti*
 whether pnwkr by.you thus not heard.PRET
 ‘Have you never heard thus? (Yes, you have surely heard thus.)’

5.3 Complex sentences

In Sogdian there are three kinds of subordinate clauses: (a) relative, (b) adverbial, and (c) nominal.

5.3.1 Conjunctions

Coordinate conjunctions are (*ə*)*r**t**i* ((*ʾ*)*r**t**y*, ʾḤRZY) which marks the beginning of the clauses, *ət(i)* (*ʾ*t(y) = (PZY, ZY) ‘and’, and *katār (əti)* (ktʾr, ʾWZY) ‘or’. Apart from these, several adverbs placed towards the sentence initial are employed to enhance cohesion. Some of the most frequent are: *mas* ‘also, again’, *pīšt* ‘but’, *nūkar*, *patsār* ‘then’, *tūm* ‘furthermore’. Of Old Iranian origin are connecting particles *βād . . . βād* ‘sometimes . . . sometimes’, *uβyu . . . uβyu* (wβyw, C by) ‘both . . . and . . .’.

Subordinate conjunctions are:

- preceding the main clauses: *ču* ‘if’, *čānō* ‘when, while, since’, *kaḏ(a)* ‘if’, *mant* ‘when, while’;
- following the main clauses (*əti* is the complementizer described above): *ət(i)* (*ʾ*t(y) = ʾPZY, ZY) ‘that’, *čānō at(i)* ‘as (= like), than’, *pār(u)ti* (pʾr(w)ty, later *pāt*) ‘for, (not . . .) but’, *paraw at(i)* (pʾw ʾt(y)) ‘because’, and *yiwār-əti* (yiwʾr (ʾty)) ‘however’. *kaḏ* and *ət(i)* on the one hand, and *māθ* (mʾθ = KZNH) ‘thus’ and *ət(i)* on the other, are combined to give the new subordinate conjunctions *kat* (kt, C qt) ‘that’ and *māt* (mʾt) ‘that’, respectively, which are quite common in Manichaean and Christian

texts. These conjunctions occasionally combine with adverbs to form compound conjunctions as *čiwēd piḏār pāruṭi* ‘because’, lit. ‘because of that for’, *paraw māθ ətī* ‘because’, *witūr ətī* (wytwr ’ty) ‘until’, etc.

5.3.2 Relative clauses

5.3.2.1 Correlatives and gaps

Relative clauses are introduced by relative pronouns or adverbs, to which is attached the complementizer *ət(i)* in Classical Sogdian. They always follow their head, which is usually marked by a cataphoric demonstrative. Although the following remark is not without exception, a general tendency is observable: the correlative in the relative clause is the verbal ending if nominative (a), while the direct object is gapped (b). With other cases, the relative may be in the oblique itself (*ə*)*kya* (c), but generally it assumes the relative particle (*ə*)*ke*, while the correlative is a pronoun or adverbial demonstrative filling the gap in the relative clause (d), (e) (see Provasi 1997). Comparative clauses can be classified as relative clauses (f):

- (a) *ōnō martī wiru kunāt ke-ti-šī xwati rēžāt*
that man husband make.SUBJ.3s REL-COMP-herherself please.SUBJ.3s
‘she shall make that man her husband who might be pleasing to her’.
- (b) *yunē čakraβart čintāmani dārani ke-ti əzu pβērūt-ḏārām*
this Cakravarti Cintāmaṇi spell REL-COMP I explain-PRET.1s
‘this Chakravart Chintamani spell which I explained’.
- (c) *əwən šē ratne . . . əkya-ət-mī sāče xu pāš pāt*
the three jewel . . . REL-COMP-me be.fitting.OPT.3s the honour observe.INF
‘three jewels . . . to which it may be fitting for me to pay honour’.
- (d) *xānd āfrītēt ōtākt ke ətī-šan sāk ətī patšmār nēst*
those blessed places REL COMP-their number and number is.not
‘those blessed places whose number and counting do not exist’.
- (e) *xānd āfrītēt ōtākt . . . kū-ətī wadēd mēnand xa roxšnda βayīšt*
those blessed places . . . where(REL)-COMP there stay.3p the light gods
‘those blessed places where the light gods are staying’.
- (f) *frutar māx ō əβiža βarēm čānō-ti əwən əptre xā punyān zarnē*
more we the evil endure than-COMP the father’s the merit lost
βāt
become.SUBJ.3s
‘we would rather endure sufferings than our father’s merit should be lost’.

5.3.2.2 Other relatives

(*ə*)*ke* is originally a nominative singular form used with an animate noun, but it tends to be generalized to all the contexts and behaves like a relative particle. On this point see the examples cited in the previous section.

In principle any interrogative pronoun or adverb can be used as a relative. Here follow several examples:

- (a) *wispu ark ču ətī-mī tayu framāye*
all work which (REL) COMP-me you order
‘all the work which you order me’.

- (b) *wāyōnē sāmār βīrt-đārant*
 such samādhi obtain-PRET.3p
čakanāč ətī ɔnyu βrēwar sāmār βīrt-βāt
 from.which(REL) COMP other 10,000 samādhi obtain-PASS.POT.3s
 ‘they obtained such meditation from which other ten thousand meditations can be obtained’.
- (c) *tīm āđču ɔsti katār-ətī-əmī βaxše-kām*
 still something is which(REL)-COMP-me give.OPT.2s-FUT
 ‘still there is something which you could give me’.

5.3.2.3 Indefinite relative clauses

Indefinite relative clauses have no head and precede the main clause, which is often introduced by the conjunction *rti*: (a). Sometimes indefiniteness is reinforced by *xatu* and/or *āđparm*: (b). Another way of reinforcing indefinite force is to duplicate the relatives, first by the original (in the present case *kū* ‘where’) and again by *ke*, which in this function is a relative particle rather than a relative pronoun: (c). *ču žamnya*, lit. ‘whatever time’ functions as a kind of compound conjunction meaning ‘whenever’: (d).

- (a) *rt-mas katām āžunī āžit rti xu rāmant jātismar yrīw βīrt*
 and-again which(REL) life is.born and he always jātismara body gets
 ‘again in whichever life he may be born, he will always get jātismara-body (i.e. he will remember his previous lives)’.
- (b) *əču xatu āđparm wanti rti-štī wisp sāt . . .*
 what(REL)ever ever does and-him every all . . .
 ‘whatsoever he may do, everything is . . . for him’.
- (c) *kū sār ke šawēte xu fuxār . . .*
 where(REL) toward REL go.OPT.MID.3s the blessed.one
 ‘wherever the blessed one may go . . .’.
- (d) *rti ču žamnya čan βayistān ɔxazāt rti paru butyāk*
 and what time.LOC from heaven descend.SUBJ.3s and on buddhahood
rāθ ɔštayt
 road stands
 ‘whenever he may descend from the heaven he will stand on the road of Buddhahood’.

5.3.3 Adverbial clauses

Clauses that explicitly or implicitly state logical and temporal conditions precede the main clause, and are typically introduced by conjunctions such as *kađ(a)* ‘if’, *čānō* ‘when, since’, *mant* ‘when, while’, *kū parm* ‘as long as’, etc. In classical Sogdian both subordinate and main clauses are introduced by a conjunction *rti*.

- rt-šu kađ nē waytu-đāre rt-šu yōnēđ nā wāč*
 and-him if not send-PRET.2s and-him immediately not send.IMPV.2s
 ‘if you have not sent him, do not send him immediately’.

- rt-šu kū parm nē tōžān rt-šu . . . paru wartu đārām-kām*
 and-it as.long.as not pay.SUBJ.1s and-it . . . on interest have.1s-FUT
 ‘until I pay it (= as long as I do not pay it), I shall owe (lit. ‘have’) it with interest’.

Other adverbial clauses expressing purpose, cause, etc. are accompanied by appropriate conjunctions and follow matrix sentences. Here follow examples of (a) cause and (b) purpose:

- (a) *čan šē satu əfsānx zāy əti āyatim paraw māθ-əti wānō patīyōš*
 from 3 100 mile land COMP I.came because-COMP thus I.heard
 ‘I have come from the place 300 miles far (from here), because I heard thus . . .’.
- (b) *šawām-əštān əti-š(a)č 0βār βīrān*
 I.go-DUR CONJ(=COMP)-him-from gift obtain.SUBJ.1s
 ‘I am going (there) so that I might get a gift from him’.

5.3.4 Nominal clauses and direct quotation

5.3.4.1 Nominal clauses

The most common nominal clause is the complement of a transitive verb which is introduced by *ət(i)*, *kat* (< *kaδ-ət(i)*), or *māt* (< *māθ ət(i)*): (a). An appositional clause is also introduced by the same conjunction: (b).

- (a) *xa pōdt wānō wēnant əti xa brāmānt 0 . . . xwatāw parēpant*
 the elephants thus saw that the brahmins the . . . king took-away
 ‘the elephants saw that the brahmins took the king away’.
- (b) *rti wānō padka nēst əti əwī yantāk rāθī anpatē*
 and thus judgement is.not that the bad way fall.OPT.3s
 ‘there is no such judgement that he should fall into a bad existence’.

5.3.4.2 Direct quotations

Direct quotations are treated differently from clausal complements in classical texts but later they came to take the same construction. Thus in one Buddhist Sogdian text *Vessantara Jātaka* direct quotations are always introduced by the complementizer placed in the second place (see section 5.1.2.1), while in later texts they are often introduced by *kat*:

- ərt-xu xēpθāwand wānō wāβ kat-βay munō martī. . .*
 and-the owner thus said that-lord! this man . . .
 ‘the owner said: “O Lord! (I hired) this man . . .”’

Notice that in Sogdian it is a rule that the verbs introducing direct speech or nominal complements should be preceded by a pleonastic adverb *wānō* or *māθ* ‘thus’.

5.3.5 Use of infinitives and verbal nouns

5.3.5.1 Past infinitives

Functionally speaking, infinitives and verbal nouns are comparable to subordinate sentences and their usage is described in this section.

The past infinitives are dependent on verbs such as meaning ‘order’, ‘begin’, ‘wish’, ‘be fitting’, etc. and always follow them. Below is the past infinitive dependent on an impersonal verb *sāšt/sāčt* ‘it is fitting’: (a). A subordinate clause dependent on *sāšt* is also given: (b)

- (a) *rti əxu nē βaxše ōnō əču-ti-šī nē sāčt βayte*
 and he not give.OPT.3s which (REL)-COMP-him not is.fitting give.INF
 ‘he may not give that which is not fitting for him to give away’.
- (b) *wānō sāšt kat par wēšant čūpar yān xōžēm čan βaya*
 thus is.fitting that on them upon favour request.OPT.1p from god
 ‘it is fitting that we should ask favour from the God upon them’.

One special usage of the past infinitive is to appear in the construction *čan . . . kārī* (cnn . . . k’ry) meaning ‘after having . . .’, where the light stem takes the ending *-a* or *-u* while the heavy is without ending (on the origin of the construction, see Sims-Williams, 2007):

<i>čan murtu kārī</i>	<i>mana čan nirβān witart kārī</i>
from dying after	my from nirvāṇa passing after
‘after having died’;	‘after my passing into the nirvāṇa’.

5.3.5.2 Present infinitives and verbal nouns

As stated above the form and usage of the present infinitives in Sogdian vary considerably from text to text. Their unstable nature owes much to the fact that they were developing from the nominal form based on the present stem, i.e. productive verbal nouns, to the forms more similar to the past infinitives. *par* infinitives in Christian texts are used in a way similar to the past infinitives:

<i>ke</i>	<i>kāmēnt</i>	<i>par-patxwāy wənī</i>
who(REL)	desire.OPT.IMPF.3p	<i>par</i> -kill him
‘who were wishing to kill him’.		

The present infinitives of the Buddhist texts show more characteristics of the verbal nouns than *par* infinitives of the Christian texts, where the ordinary verbal nouns are derived by means of the highly productive suffix *-āmantē*.

Compare the Buddhist Sogdian and Christian functionally equivalent:

<i>zanē</i>	<i>žamnya</i>	<i>wya zanāmantē mēθ</i>
giving.birth time.LOC		her giving.birth day
‘at the time of giving birth’;		‘the day of her giving birth’.

Forms derived with *-āmantē* are very few in Buddhist texts whereas they are much more numerous in Manichaean texts.

6 LEXIS

6.1 Sogdian lexicon

The Sogdian lexicon consists of three groups of words: (a) those inherited from Old Iranian; (b) loanwords; (c) foreign elements temporarily appearing in texts. Some of the native Iranian elements show distinctively East Iranian features. Thus, *kamθ* ‘town’, *kap-* ‘fish’ are typically East Iranian. *farn* ‘glory, fortune’ has recently been shown to be originating from the language of Scythians at the Old Iranian stage (see Lubotsky 2002).

Foreign elements are mainly Sanskrit forms in Buddhist texts and Syriac words in Christian texts (see Sims-Williams 1983b, 1988). They are quite numerous and their

number seems to be dependent on the scholarship of each translator. On the other loanwords originating from the language contact, see section 7.3.

6.2 Derivation

6.2.1 Suffixes

Derivation in Sogdian, which is an OV language, is mostly by suffix. The most productive suffixes are the following. They are so productive that loanwords come to be provided with them:

(a) Adjectives from (mainly) nouns:

- 1 *-ēnē*, f. *-ēnč* (-'yn'k; *aka*-stem) 'made out of ~': *zernēnē* < *zern* 'gold', *δārukēnč* < *δāruk* 'wood'.
- 2 *-čīk*, *-īčk* (-cyk, -yck): *γarčīk* 'mountain-like' < *γar*- 'mountain'; *cād arčīk* 'low' < *čādar* 'below'; *βēkpārčīk* ~ *βēkpārīčk* 'external' < *βēkpār* 'outside'. Compare the two adjectives derived from *zāwar* 'power': *zāwarčīk* 'helpful' vs. *zāvarkēn* 'powerful'.
- 3 *-mēnč* (-mync): *īnčmēnč* 'female' < *īnč* 'woman', *žūk yāmēnč* 'safe, sound' < *žūkyā* 'health' < *žūk* 'healthy'.
- 4 *-īk* (-yk): *δarmīk* 'of dharma' < *δarm* (< Skt. *dharma* 'Buddhist law').

(b) Abstract nouns:

- 5 *-yāk* (-y'k, from LS), *-yā* (-y'kh, M -y', from HS): *roxšnyāk* 'splendour' < *roxšn*- 'light', *βayyāk yā* 'divinity' < *βayyāk* 'id.'. < *βay*- 'god'.
- 6 *-āwē* (-'w'k, *aka*-stem): *friyāwē* 'love' < *fri*- 'dear', *patpatēnāwē* 'isolation' < *patpatēn* 'opposed'.
- 7 *-ōnī* (-wny): *mastōnī* 'drunkenness' < *mast* 'drunken', *tāyōnī* 'theft' < *tāy* 'thief'.

(c) Others:

- 8 *-ānč* (S -'nch): female counterpart of nouns meaning male human: *upāsānč* < *upāsē* 'lay-brother' (< Skt. *upāsaka*), *nəyōšākānč* < *nəyōšāk* 'Manichaeon auditor' (< Part. ngwš'g). For *upāsānč* see Tocharian B *upāsakāñca* 'id.'.
- 9 *-āu* (-'w) 'in the ~ language': *čmāu* 'in Chinese' < *čm* 'China', *syudyāu* 'in Sogdian' < *suyδ*- 'Sogd'.

6.2.2 Prefixes

Prefixes are much less common than suffixes. Productive are the following:

- 1 *āu*- ('w-) 'co-': *āupatyāp* 'sharer' < *patyāp* 'share', *āuxānēt* 'co-inhabitants' < *xānā* 'house'.
- 2 *mant*- (mnt-) privative: *mantγərβāk* 'foolish' < *γərβāk* 'wise'.
- 3 *nā*- (n-) privative: *nā-paδkčīk* 'unlawful', *nā-martəxmē* 'non-human'. The past participle is negated with this prefix: *nā-wēt* 'invisible' < *wēn/wēt* 'see'.
- 4 (ə) *pu*- (()pw-) privative: *pu-patšmār* 'innumerable' < *patšmār* 'number'.
- 5 *frī*- (S pry-) 'philo-': *frī-rwān* 'who loves one's own soul' < *rwān* 'soul', *frī-āβraxsē* 'lascivious' < *āβraxsē* 'lust'.

6.3 Composition

Several patterns of both verbal and nominal compounds are known (see Gershevitch 1945).

(a) Containing present stems:

- 1 noun + present stem + *aka*-suffix forming agent nouns: *framān-patyōšē* ‘servant, obedient’ (*framān* ‘order’ + *patyōš* ‘hear’), *yipāk-βarē* ‘angry’ (*yipāk* ‘anger’ + *βar*-‘bear’). Notice that older type of the similar compounds show as its second member a slightly different form than the present stem: *əptxwār* ‘carnivorous’ < **pitu* + *xwāra*, compare *xwar*- ‘eat’, *širxōzē* ‘friend’ < *šir*- ‘good’ + *xwēž* ‘to wish’. The following compounds contain Old Iranian acc. sg. forms as the first members: *βžangārē* ‘sinful’, *šrankārē* ‘pious’.

(b) Containing past stems:

- 2 past stem + noun forming bahuvrīhis: *suft-γōš* ‘with pierced ears’ < *sumblsuft*-‘pierce’, *əkt-əspās* ‘obedient’, lit. ‘made-service’, *βart-γamband* ‘enduring the strain (*γamban*) (pl.)’ < *βar-lβart* ‘bear’.
- 3 noun/adverb + past participle: *wāt-nyātē* ‘wind-sick, taken by wind (*wāt*)’ < *nyās/nyāt* ‘take’, *čon-žmē-miđānī-āγatē* ‘coming from among (*miđān*) the anger (*žmē*)’ < *ēslāγat* ‘come’.

(c) Containing nouns and adjectives:

- 4 noun, adjective, etc. + noun + *aka*-suffix forming bahuvrīhis: *šir-nāmē* (M *šyr-n`my* ‘good’ + ‘name’) ‘famous’ (cf. *širnām* ‘fame’), *δwāts-rītē* ‘having twelve faces (*rīt*)’, *ān-kutrē* ‘originating from the An family (*kutr*-)’.

6.4 Collocation

6.4.1 Hendiadys

Two near synonyms deriving from the same root are combined to make a kind of dvandava compound, occasionally with a conjunction *ət(i)* ‘and’. They are called etymological hendiadys and are quite common in Sogdian. Of the two members, the shorter form precedes. When they consist of the same number of syllables, the one beginning with vowel precedes: *wazti frawazti* ‘he flies about’ < *waz*- ‘fly’, *frī āfrī* (pryh ”pryh) ‘very dear’, *ankrantt ət frakrantt* ‘he cuts and chops’ < *krant* ‘cut’, etc. Synonymous hendiadys, usually asyndetic, are no less common: *yōk fsāk* ‘teaching (and) training’, *nom paδka* ‘law and judgement’, *ramē γē0* ‘cattle flock’. However, Sogdian lacks such hendiadyses as those abundant in Uighur which consist of an Indian loanword and its Uighur synonym: *čakir tilgän* ‘wheel’ < Skt. *cakra*- (see Gabain 1974: 159).

6.4.2 Phrasal verbs or idioms

Idiomatic combinations or phrasal verbs are not uncommon in Sogdian. Examples are: *0βār 0βar*-, lit. ‘gift give’, = ‘give’, *rā0 šaw*-, lit. ‘road go’ = ‘travel’, *namāč βar*-, lit. ‘honour bring’ = ‘pay homage’, *wadu parēp*, lit. ‘wife take’ = ‘get married’, *zāwar 0βar*-, lit. ‘power give’ = ‘help’. While all these are transparent and are easily analysable, the origin of

kūmārī parēs (kw̄m'ry prys) 'understand perfectly' seems to have been already opaque to the Sogdians. It is a combination of *kū mārī parēs*, lit. 'arrive at memory (*mār* 'spell')'.

7 SOCIOLINGUISTIC ASPECTS

7.1 Polite phraseology and signals

7.1.1 Honorific expressions

7.1.1.1 Verbs

Sogdian possesses several ways of expressing respectful and humble feelings on the part of a speaker. The honorific expression for 'say' is *framāy*, lit. 'to order' while the opposite is expressed by *patškway* and the unmarked lexeme is *wāβ*. Thus when the Buddha or a king says, *framāy* is employed whereas disciples or servants are *patškway*-ing:

rti nūkar əxu ānant kū βayān βaxtam buti māθ patīkway . . .
and now the A. to devātideva Buddha thus said . . .
rti əxu βayān βaxtam buti əwən ānant māθ frāmāy . . .
and the devātideva Buddha the A. thus said . . .
'then (the disciple) Ananda said to the devātideva Buddha . . .
the devātideva Buddha said to Ananda . . .'

The corresponding nouns *framān* and *patškwān* mean '(His Majesty's) word' and 'humble message' respectively.

The verb *framāy* combined with the past infinitive is also used as an honorific expression meaning 'deign to . . .'. The opposite is the verb *raxn-* 'venture to do':

framāy əti-mī ōnō wiδβāy əkte
order.IMPV COMP-me that explanation make.INF
'please explain that to me'.
əzu čānō raxnu əti-n yunē əsfrāčk əpsu
I how dare.INJ.ls COMP-PART this thing ask.INJ.ls
'How dare I ask (the Buddha) about this matter?'

Notice that in the sentence cited above, politeness is also expressed by the injunctive forms *raxnu* and *əpsu*. On the hypothetical particle *-(a)n* see section 3.2.3.1f.

The verb *xrām* 'walk' seems to have a shade of honour, i.e. 'walk with dignity':

rti-šī māθ patīškway xrām əti βaya . . .
and-him thus said walk COMP lord!
'he said to him (the prince): come, O Lord!, . . .'

7.1.1.2 Nouns and other expressions

Pleonastic use of the vocative form *βaya* 'O Lord!, sir' or its enclitic counterpart *-βay* is one of the most common ways of expressing one's feeling of honour towards the addressee:

rti-šu māθ patīškwayant zārī əti βaya sayēm
and-himthus said.3p sad COMP lord! seem.1p

pār-ti βaya māx amī uxušu βrēwar inč anyu zātē nēst
 for-COMP lord! us the 6 10,000 woman other son is.not
 ‘they said to him: we feel sad, O Lord!,
 because, O Lord!, for us 60,000 women there is no other son’.

rṭi-βay əzu mađe əkū čāčnīk xūβ ēsu rṭi-βay uβyu pōstēt patīwēdu . . .
 and-lord! I here to of.Chach king came and-lord! both letters submitted . . .
 ‘O Lord! I came here to the king of Chach. O Lord! I handed both the letters . . .’.

farn (S prn) ‘fortune, majesty’ is also sometimes used as an honorific word, compare *šmāxfarn* ‘Your Majesty’, *tavafarn* ‘Your (sg.) Majesty’.

As in New Persian 2p forms instead of 2s may be employed for expressing honour:

Dāwīd sūkβār farn-sār γarf žām namāč . . . anβarz βarta . . .
 D. monk majesty-toward many humble homage . . . care bring.IMPV.2p . . .
əzu šmāx βantē im
 I your.PL servant am
 ‘to David the monk. much humble homage to (Your) Majesty . . . (please) pay (2p)
 heed to . . . I am your (2p) servant’.

7.1.1.3 Expressions showing the speaker’s humility

In the passage cited in the last section one finds other means of expressing one’s feeling of humility: *žām* ‘humble’, lit. ‘fine’ and *βantē* ‘(your) servant’. *βantē* ‘male servant’ or *dāy* ‘female servant’ is a humble expression for ‘I’ and is often accompanied by *kaštar* ‘less’ or other synonyms:

pīšt čan xēp0 kaštar satu βrēwarmīk βantē
 sent from your less 100 10,000th servant
 ‘sent from your most trifle one millionth servant’.

Notice that among the elements discussed above *patškway* (> *patškwān*), *raxn-*, and injunctive forms are also counted as showing humility on the part of a speaker or addresser.

7.2 Dialects and variations

7.2.1 Archaic and late features

7.2.1.1 Ancient letters: oldest stage

The oldest and latest strata of the Sogdian language differ considerably from each other. The oldest stage is represented by the so-called ‘Ancient Letters’. One finds several features peculiar to the material. These include:

- 1 Some nouns that show special inflections totally unknown in later Sogdian:
 inst.-abl. *duγOrya* (δωγδρυh) < *duyta* ‘daughter’,
 gen.-dat. *māOrī* (mʹδρυh) < *māt* (mʹth) ‘mother’.
- 2 The abundance of the ergative construction which is noticeable because the construction is later displaced by the transitive preterite with the auxiliary *dār*:

əst Orē sarδ čānō-ti-m-(a)n-č *pačβant əβyart*
 is 3 year since-COMP-me-them-from answer obtained
 'It is three years since I have received an answer from them.'

As regards the phonology, several light stem forms are without vocalic endings: e.g. *əβiž* ('βyz) 'evil', *kərt* (krt) 'made', etc., but it is hard to assess this fact in the history of Sogdian phonology.

7.2.1.2 Christian Sogdian text C5: latest stage

The latest stage of the Sogdian language is represented by Christian Sogdian manuscript C5. At this stage:

- 1 No article is employed.
- 2 The conjunction *rti* never occurs and its function of connecting sentences is displaced by *ət*.
- 3 The conjunction *kat* 'that' plays the function of relative particle as well.
- 4 The nominal inflection shows strong tendency to use nominative form in all the case functions. Thus in C5 only *x-* and *y-*forms of the extended demonstratives are attested: *xēδ*, *xānt*, *ēδ*, *yānt*.

On the other hand:

- 5 The new oblique case begins to be formed by adding *-ī* to nom. sg. and to similar generalized forms. Compare the inflexions of *ram-* (masc.) 'people' and *wan-* (fem.) 'tree': sg. dir. *ram-i*, obl. *ram-i-(y)ī*; pl. dir. *ram-ta*, obl. *ram-ta-(y)ī*; sg. dir. *wan-a*, obl. *wana-(y)ī*.

Had the inflection been extended to all nouns, the language would have become a fully agglutinative type like Yagnobi, the so-called 'Modern Sogdian' (see Sims-Williams 1982: 69–70).

7.2.2 Sogdian dialects?

Linguistic differences observed within the written materials are relatively trivial in view of the wide area where Sogdian texts were discovered. Most differences can be explained as chronological or orthographic. For example the durative particle *skun* appears in the forms ('skwn, skwn, C sqn, sk'n, kn (C qn), sk, kw, and k). ('skwn) and (skwn) which are mainly attested in Buddhist and Manichaean texts are older than (sqn, kn, kw), and (k) found in Christian and other late texts and one can assume that gradual reduction of unaccented particles is reflected in the spellings: *əskun* > *skun* > (*s*)*kən* > (*s*)*kə* > (*s*)*k*. A similar reduction of the future particle *kām* is observed.

Henning once argued that the linguistic difference between Manichaean and Christian texts would most probably be due to the sociolinguistic factors (see Henning 1958: 105–108). According to him, in Turfan where the bulk of Manichaean and Christian Sogdian texts were discovered, the Manichaean church was supported by the aristocratic believers whereas ordinary people were the members of Christian church, and the differences between the two varieties are due to the social stratification of the both groups of speakers. Although Henning was reluctant to admit the existence of Sogdian dialects reflecting geographical distribution, one may be reminded that in Christian texts the two forms representing the durative past, i.e. *āz*-imperfect and *ēk(w)māt* imperfect, show

complementary distribution, and that the fact may most reasonably be explained by supposing two dialects which developed different forms for a single category (see Yoshida 1980). However, the difference is trivial and may not deserve to be called dialectal.

The Sogdian language documented by the materials known to us most probably represents the standard variety spoken in Samarqand. A Chinese pilgrim Xuanzang (?-664) who went to India via Sogdiana reported that the manners of Samarqandians were imitated by other Sogdians (see Beal 1884: 33). In fact the Bukharan dialect as cited by Islamic writers is slightly different (see Sims-Williams 1989b: 165–166).

7.2.3 Sogdian and Yaghnobi

The wide range of linguistic difference which once existed in Sogdiana may be inferred by comparing Sogdian with Yaghnobi.

- 1 The absence of the rhythmic law in Yaghnobi may be explained by its loss in the course of development.
- 2 However,
 - (a) the formation of the imperfect stem by adding the augment *a-* to any present stem: Yaghn. *piraxs-* > *a-piraxs-* ‘left’ vs. Sogd. *parəxs-* > *pārəxs*, and
 - (b) the 3rd pl. ending *-or*:
Yaghn. *a-wen-or* ‘they saw’ vs. Sogd. *wēn-ant*, cannot be due to the linguistic change.

It is worth noting that the *-r* ending of 3p is shared by the neighbouring Choresmian and Khotanese languages.

7.3 Language contact and loanwords

The land of Sogdiana was surrounded by other Iranian-speaking areas such as Bactria (to the south), Chorasmia (to the west), and Khotan (to the east). The three languages show a strong affinity to Sogdian and they share several areal features (see Sims-Williams, 1983a and idem 1989a: 169–170). However, no loanwords borrowed from Choresmian and Khotanese has so far been noticed, whereas a few Bactrian elements are known in Sogdian: *saxr* (*sxr*) ‘wheel’ and *raxā* (*rx*’kh) ‘cart’.

7.3.1 Borrowings

7.3.1.1 Contact with Bactrian or Kushan influence

Bactrian, the national language of the great Kushan empire, apparently had prestige over Sogdian, because the two forms just cited had cognates in Sogdian, i.e. *čaxr* (*cxr*) and *raθ-* (*rδθ-*), and may have been borrowed as culture words (see Sims-Williams 1996a: 50–51). Compare also *paθfar-* (*pδθr-*) ‘honour, rank’ borrowed from Bact. *πιδοφορο* [*pid̥far*] ‘honour’ (see Sims-Williams 2004: 541). Similarly *sārtpāw* (*s’rtp’w*) ‘caravan-leader’ is suspected to be a loan from Bactrian and this form consisting of Indian *sārt* (cf. Skt. *sārtha-* ‘caravan’) and Iranian *-pāw* induces one to assume that some loanwords of Indian origin were borrowed into Sogdian via Bactrian. One possible example is *pani* (*pn̥y*) ‘small copper coin’ ultimately going back to Skt. *paṇa*, which was borrowed into Bactrian as *πανα*.

Other Indian elements in Sogdian may also be due to the Kushan influence. Thus Indian loanwords attested in the Ancient Letters belong to the field of transaction, such as *mīdi* (mwdy < Skt. *mūlya*) 'price', *sārt* (< Skt. *sārtha*) 'caravan', *prastak* (< Skt. *prastha* 'pint (measure of capacity)'), etc., and the Sogdians may have learnt them from the Indian merchants travelling through the territory of the Kushan empire. Moreover, several Greek elements may also have come from Bactrian: *diðam* (dyð(y)m < διαδημα 'diadem', *nom* (nwm < νομος 'law', *draxm-* (dʀxm < δραχμη drachma', *kapiθ* (M qpyð < καπηλειον 'shop', *miðamβan* (< μεδιμνος 'measure of capacity'), etc. Of course they could equally be a legacy of Hellenism in Iranian languages in general.

7.3.1.2 Loanwords from Western Middle Iranian

Strong cultural influence came from Parthia and Sassanian Iran, and Western Iranian loanwords are numerous. For example all the words for days of a week are of Middle Persian origin: *mīr* 'Sunday', *māx* 'Monday', *wanxān* 'Tuesday', etc. (see Henning 1937: 85–86). A Middle Persian word *rōč* 'day' is usually added to the Sogdian name of a day of a month, e.g. *xurmazt rōč* 'day of Xurmazt, i.e. the first day of a month'. *maγdaβ-* 'minister' is from Parthian and is once accompanied by another Parthian word *wazark* 'great': wz'rkt mγδβt 'great ministers'.

Vocabulary concerning Christianity is also from Western Iranian: *tarsāk* 'Christian', *sūkβār* 'monk', *masiðar* 'presbyter', etc. (see Benveniste 1964).

However, the alleged contact with Old Sogdian and Old Persian has been refuted as illusory (see Sims-Williams 1989a: 171).

Some Indian elements came via Manichaean Parthian, e.g. *čaxšāpat* (< Parth. cxšʔbyd < Skt. *śikṣāpada*) 'precept', *bāšīk* 'hymn' (< Parth. bʔšh < Skt. *bhāṣā*).

7.3.1.3 Indian loanwords

However, not all the Indian elements came via Parthian. Thus, such words as *makara* 'monkey' (mkr < Skt. *markaṭa*), *nāk* 'dragon' (nʔk < Skt. *nāga*), *ratn-* 'jewel' (rtn- < Skt. *ratna*), *šaman-* 'Buddhist monk' (< *śamana* < Skt. *śramaṇa*), etc. which are fully naturalized in Sogdian may represent Indian cultural influence in general. Some phonological features found in those words are proved to be of North Western Prakrit (generally known as Gāndhārī) origin (see Sims-Williams 1983a).

Naturally, Indian elements abound in Buddhist texts. They are of three groups: (a) loanwords naturalized in Sogdian: e.g. *samutr-* (< Skt. *samudra*) 'ocean', (b) Buddhist terminologies well established in Sogdian: e.g. *bodisatβ* 'Bodhisattva' (< Skt. *bodhisattva*), and (c) nonce borrowings employed by certain translators: e.g. *prātimokš* (< Skt. *prātimokṣa* 'moral code').

7.3.1.4 Borrowing from other languages

As the traders of Silk Road, a number of Sogdians were resident in China and one might expect many borrowings from Chinese. However, they are not at all numerous, and a very few which actually found their ways into Sogdian are attested also in other Central Asian languages, such as Uighur and Tocharian (see Yoshida 1994: 379). For example a word for 'pint' *šang* which comes from Chinese 升 (Middle Chinese **šiəng*), also found in Uighur *šing*, Tocharian *šank*, and Khotanese *šamga*, *šimga*. The situation suggests that they were culture words widely in use among the peoples of Central Asia, and that

they were not necessarily borrowed directly from Chinese into Sogdian. Similarly, the languages of Turkish peoples, with whom Sogdian had close contact, lent very few words to Sogdian. One of the possible examples, other than personal names and titles, in late texts is *arxiš* ‘caravan’ (< Uighur *arqış*) which seems to have displaced a Indian counterpart *sārt* by that time.

Tocharian elements are also very few. One example *sintāp*, which occurs in a medical text, seems to be borrowed from Tocharian B *sintāp* ‘a kind of rock-salt’, ultimately derived from Indian, compare Skt. *saindhava* ‘rock salt’.

Recently, *čabiš* (cpyš), which alternates with an inherited word *sarθang* (srθng) ‘general (of army)’, has been argued to have originated from the language of the Hephthalites. The Hephthalites ruled Sogdian for some hundred years beginning in the latter half of the fifth century and their language could have influenced Sogdian, though no other Hephthalite words have been known.

7.3.2 Sogdian elements in other neighbouring languages

In stark contrast with the borrowings, Sogdian lent many words to Uighur, an Old Turkish dialect, e.g. *ažun* (< *āžun*) ‘existence’, *kānt* (< *kaṃθ*) ‘town’, *nizwani* (< *nizβānē*) ‘sin’, *tamu* (< *tam-*) ‘hell’, etc. Sogdian’s prestige over Uighur was such that the latter adopted Sogdian script for writing Uighur. A so far unnoticed Sogdian element in other Turkic languages than Uighur is *axšam* ‘evening’ which goes back to *axšām* (‘xš’mh) ‘evening meal’.

The influential Sogdian culture and language in the early Islamic period may account for relatively numerous Sogdian loanwords in New Persian (see Henning 1939): e.g. *čayz* ‘frog’ < *čayz* (cyz), *čuyd* ‘owl’ < *čyut* (cywt), etc.

The reason why virtually no Sogdian loanword is found in Chinese may be due not only to the latter’s prestige but also to its natural reluctance to borrowing foreign words in general. Only two forms which found their way into dictionaries are *chi bo* 叱撥 (Middle Chinese **tš’iet puat*) denoting a special kind of horse which is derived from *čərθpād* (cyrθp’δ) ‘quadruped’ and *mi* 蜜 (Middle Chinese **miet*) from *mūr* ‘Sunday (ultimately from Middle Persian *mihr*)’.

7.4 Language death

The latest known Sogdian text is dated to 1025 CE (see Livšic 1996: 271 and Yoshida 2004: 21). Possibly within a few centuries after that the Sogdian language was no longer spoken or written and became a dead language.

7.4.1 Death of Sogdian: Turkicization and Persianization

7.4.1.1 Turkicization

The changing relationship between Sogdian and Uighur is reflected in several late texts where many Turkish elements (words, personal names, expressions calqued on Uighur) are found. A text discovered in Dunhuang and dated to the tenth century is bilingual in Sogdian and Uighur; in it the scribe alternates between writing in Sogdian and in Uighur, and it is difficult to find reasons for the use of one or the other (cf. Sims-Williams and Hamilton 1990: 24–25):

Uighur

temči-dä bir qizil qars alyu bar
 Temči-by one red wool to.be.taken is
 ‘by Temči one red qars is to be taken’.

Sogdian

alp irkinī nəβant əšē karmīr rayzi āsī xu
 Alp Irkin.OBL by three red wool to.be.taken is
 ‘by Alp Irkin three red rayzs are to be taken’.

Sogdian appears to have been in the process of being absorbed into Turkish first in the eastern part of the Sogdian speaking area. Maḥmūd al-Kāšgarī of the eleventh century reports that in his time people from Soγd (area between Bukhara and Samarqand) were resident in Balāsāγūn (a city in modern Semirechie, Kirghiz), and that their dress and manner was that of Turks (see Dankoff and Kelly 1982: 352 and Yoshida 2009). The description of those Sogdians may point to their profound Turkicization in the eleventh century.

7.4.1.2 Persianization

The examples of the languages that al-Muqaddasī in the tenth century reports as those of Samarqand (*be-goftag-om* ‘I have said’) and Bukhara (*dānest-i* ‘you.2s knew’) are nothing but dialects of New Persian (see al-Muqaddasī 1994: 335), and show that Sogdian was no longer in use there by that time. This conclusion is also reflected in a note of Iṣṭakhri of the tenth century that in his time the inscription on the Kish gate of Samarqand was incomprehensible to the local people. It is likely that the inscription, which Iṣṭakhri took for Himyaritic, was in Sogdian (see Barthold 1958: 87). On the other hand, al-Muqaddasī also mentions a vernacular spoken in Suγd (the area lying between Bukhara and Samarqand) which sounded similar to a dialect found in the suburbs of Bukhara.

This suggests that during his lifetime Sogdian was still spoken in small villages, while New Persian was spoken in the urban areas. Today, the sole survivor of the Sogdian dialects, once spoken all over Sogdiana, is Yaghnobi, which continues to be spoken in the most remote valley of the Yaghnob (North Tajikistan).

8 SAMPLE TEXT

The following sample texts reproduce the main part of the Manichaean Sogdian ‘Story of the Pearl-Borer’, pp. 466–469 in Henning, W. B. (1945). The text is in two scripts, one in Manichaean script and the other in Sogdian. (Round brackets) indicate damaged letters; [square brackets] indicate suggested restorations of missing letters (see Henning, *ibidem*). The text is arranged here by clause divisions. For each clause there are five lines:

- 1 transliteration of Manichaean script;
- 2 transliteration of Sogdian script;
- 3 transcription;
- 4 literal interlinear translation;
- 5 idiomatic translation, mostly following Henning.

1

'rt_{xw} mr'z mr_{ty} kw x_{tw} s'r w'nw p_{ty}škw_y
 [rt_y ZK] mr'z kw 'x_{tw} s'r (w'n')[kw p_{ty}š]kw'y
 (ə)rti xu marāz martī kū əxtu sār wānō patīškway
 and the hired.labour man to judge towards thus said.IMPF
 'The hired labourer addressed the judge thus:

2

ktβ_y ywny fšy'ws c'nw t'm' w'crn kwsy_y wyn
 ktβ_y ywn'k βšy-'ws c'n'kw t'm'k ZKwy w'crn'y kwsy wyn
 kat-βay yunē (ə)fšəyāus čānō tāmā əwī.LOC wāčarn(ī) kōsī.OBL wēn.IMPF
 that-lord! this gentleman when me the bazaar side saw
 "My lord, when this gentleman saw me at the side of the bazaar,

3

'tm_y w'nw ps' kδw_{ty}h cw 'rq γrβ_y
 rt_{my} w'n'kw ps' 'YKZY cw 'rkh γrβ(?)_y
 ərtlət-mī wānō psa kaδ-uti ču ark γərβe
 and-me thus asked that what work you.know.2s
 he asked me: "Hey, what work can you do?"

4

'rtš_y 'zw w'nw p_{ty}škw_y kt β_y wyspw 'rk
 rtš_y 'zw w'n'kw p_{ty}škw'y [k](t) β_y wysph 'rkh
 (ə)rt-šī əzu wānō patīškway kat βay wyspula.ACC ark
 and-him I thus said that lord! all work
 'I replied: "Sir, whatever work

5

cw'tm_y t_yw frm'y_y 'rt_y s't wyspw 'rk γrβ'm
 cw-[ZY](m)_y t_yw pr'm'y_y r_{ty} s't wys(p)[h] γrβ'm
 ču-ət-mī tayu framāye ərti sāt wispula ark γərβām
 that-COMP-me you order and all every work I.know
 you may order me (to do), I can do it all" '.

6

't_ym_y c'nw kw x'n' s'r šykr
 rt_{my} c'n'kw kw x'n'kh s'r šykr
 ərt-mī čānō kū xānā sār šīkar
 and-me when to house towards lead.IMPF
 'When he had taken me to his house,

7

'rt_{my} wyn' fr'm'y j_{ty}
 r_{ty} my wyn'kh pr'm'y z-yt'y
 ərti-mī wīnā framāy žəte
 and-me lute ordered play.INF
 he ordered me to play on the lute'.

8

'rty	βy'ryy	prm	prw	xypδ'wndy _h	frm'n	wyn'	jyt(w) δ'rm
rty	βy'r'k	prm	prw	xypδ'w'nty	prm'nh	wyn'k	z-ytw δ'r'm
<i>ərti</i>	<i>βyārē</i>	<i>parm</i>	<i>paru</i>	<i>xēpθāwantī.OBL</i>	<i>framān</i>	<i>wīnā</i>	<i>žətu-δāram</i>
and	evening	until	on	owner's	order	lute	play-PRET.1s

'Until nightfall I played on the lute at the owner's bidding'.

9

'rtšw	x _{tw}	w'nw	pδk'	nym'y
rtšw	'x _{tw}	w'n[kw]	pδkh	nym'y
<i>ərt-šū</i>	<i>(ə)x_{tu}</i>	<i>wānō</i>	<i>paδka</i>	<i>nīmāy</i>
and-the	judge	thus	verdict	judge.IMPF

'The judge pronounced this verdict.'

10

kt	t _{yw}	w _{ny}	m _{rty}	mr'z	ptxrytδ'ryy
'YKZY	t _{yw}	ZKn	[m](r)ty	mr'z	ptxr'ytδ'r'y
<i>kat</i>	<i>ta_{yw}</i>	<i>wānī</i>	<i>martī</i>	<i>marāz</i>	<i>patxrūt-δāre</i>
that	you	the	man	hired.labour	hire-PRET.2s

"You contracted that man's (skilled) labour"
(which includes musical performance).

11

'rtšw	cqn'c	pyδ'r	mrγ'rt	nyy	frm'yy	swβty
rtβ	[ckn](c)w	pyδ'r	mrγ'rt	L'	pr'(m)[y	swβty
<i>ərt-šul-β</i>	<i>čakanāč</i>	<i>piδār</i>	<i>maryārt</i>	<i>nē</i>	<i>framāyef/frāmāye</i>	<i>suβte</i>
and-him.PART	from.what	sake	pearl	not	order.PRES.2s	bore.INF
					ordered.IMPF.2s	

"So why did you not order him to bore the pearls,

12

p'rtyšy	xww	wyn'	jnyy	frm'tδ'ryy
[p'rZYšy	ZK	wyn'kh	z-n'y	prm't δ] r'y
<i>pār-ti-šī</i>	<i>xu</i>	<i>wīnā</i>	<i>žani</i>	<i>framāt-δāre</i>
but-COMP-him	the	lute	play.INF	order-PRET.2s

"but bid him play on the lute instead?"

13

'rty	w _{ny}	m _{rty}	xww	prxyy	'spty _h	δβtyy	γwt-k'm
rty	(ZKn)	[m _{rty}	ZK]	prxy	'spt'k	δβrt'k	γwt-k'm
<i>ərti</i>	<i>wānī</i>	<i>martī</i>	<i>xu</i>	<i>parxē</i>	<i>əsp_{tē}</i>	<i>θbartē</i>	<i>γawt-kām</i>
and	the	man	the	wage	completely	given	needs-FUT

"That man's wages will have to be paid in full."

ACKNOWLEDGEMENT

In writing this chapter I have been greatly benefitted from the works of N. Sims-Williams, in particular his overview of 1989a for the section on phonology and morphology. Gershevitch 1954 has also been consulted from time to time. It is simply impossible to refer to their discoveries and contributions in all places.

BIBLIOGRAPHY AND REFERENCES

- Barthold, W. (1958) *Turkestan down to the Mongol Invasion*, 2nd ed., London: Luzac & Co.
- Beal, S. (1884) *Si-Yu-Ki. Buddhist Records of the Western World*, Vol. 1, London: Trübner.
- Benveniste, E. (1954) 'Expression de «pouvoir» en iranien', *BSL* 50: 56–67.
- (1963 [1964]) 'La vocabulaire chrétien dans les langues d'Asie Centrale', in *Atti del convegno internaz sul tema: L'Oriente cristiano nella storia della civiltà*, Rome: Accademia Nazionale dei Lincei, pp. 85–91.
- (1966) 'Le verbe *stā-* comme auxiliaire en iranien', *AO* 30: 45–49.
- Dankoff, R. and Kelly, J. (1982) (eds) *Maḥmūd al-Kāshgarī: Compendium of the Turkic Dialect*, Part 1, Cambridge: Harvard University Printing House.
- Gabain, A. von (1974) *Altürkische Grammatik*, 3rd ed., Wiesbaden: Otto Harrassowitz.
- Gershevitch, I. (1945) 'Sogdian compounds', *TPhS*: 137–149.
- (1954) *A Grammar of Manichean Sogdian*, Oxford: B. Blackwell.
- Gharib, B. (1965) *Analysis of the Verbal System in the Sogdian Language*, PhD dissertation, University of Pennsylvania.
- (1995) *Sogdian Dictionary*, Tehran: Farhangān.
- Grenet, F. and Sims-Williams, N. (1987) 'The historical context of the Sogdian Ancient Letters', in *Transition Period in Iranian History, Actes du symposium de Fribourg-en-Brisgau (22–24 Mai 1985)*, Leuven: Peeters, pp. 101–122.
- Henning, W. B. (1936 [1937]) *Ein manichäisches Bet- und Beichtbuch*, *APAW* no. 10.
- (1939) 'Sogdian loan-words in New Persian', *BSOS* 10: 93–106.
- (1945) 'Sogdian tales', *BSOAS* 11: 465–87.
- Heston, W. L. (1976) *Selected Problems in Fifth to Tenth Century Iranian Syntax*, PhD dissertation, University of Pennsylvania.
- Livšic, V. A. (1996) 'Glave 7: Sogdijskie teksty, dokumenty i epigrafika', in Ploskij, V. M. et al. (eds) *Istočnikovedenie Kyrgyzstana (c drevnosti do XIX v.)*, Biškek: Ilim, pp. 215–273.
- Lubotsky, A. (2002) 'Scythian elements in Old Iranian', in Sims-Williams, N. (ed.) *Indo-Iranian Languages and Peoples*, Oxford: Oxford University Press, pp. 189–202.
- al-Muqqadaṣī (1994) *Aḥsan al-taqāsīm fī ma 'rifat al-aqālīm* (The best divisions for knowledge of the regions), Collins, B. (trans.), Reading: Centre for Muslim Contribution to Civilization Garnet Publishing.
- Paul, L. (1997) 'Präteritum und Perfekt im Soghdischen', *IF* 102: 199–205.
- Provasi, E. (1997) 'Note sulle costruzioni relative in Sogdiano', in Ambrosini, R. et al. (eds) *Scribthair a ainm n-ogaim. Scritti in memoria di Enrico Campanile*, Opedelatto, Pisa: Pacini, pp. 815–839.
- Sims-Williams, N. (1973) 'Notes on Bactrian syntax', *IF* 78: 95–99.
- (1979) 'On the plural and dual in Sogdian', *BSOAS* 42: 337–346.
- (1981) 'The Sogdian sound-system and the origins of the Uyghur script', *JA* 269: 347–360.

- (1982) 'The double system of nominal inflexion in Sogdian', *TPhS* 1982: 67–76.
- (1983a) 'Chotano-Sogdica', *BSOAS* 46: 40–51.
- (1983b) 'Indian elements in Parthian and Sogdian', in Röhrborn, K. and Veenker, W. (eds) *Sparachen des Buddhismus in Zentralasien*, Wiesbaden: In Kommission bei O. Harrassowitz, pp. 132–141.
- (1984) 'The Sogdian "rhythmic law"', in Skalmowski, W. and van Tongerloo, A. (eds) *Middle Iranian Studies. Proceedings of the International Symposium Organized by the Katholieke Universiteit Leuven from the 17th to the 20th of May 1982*, Leuven: Peeters, pp. 203–215.
- (1985a) *The Christian Sogdian Manuscript C2*. (Schriften zur Geschichte und Kultur des alten Orients. Berliner Turfantexte XII), Berlin: Akademie-Verlag.
- (1985b) 'A note on Bactrian phonology', *BSOAS* 48: 111–116.
- (1986a) 'Sogdian **δprm* and its cognates', in Schmitt, R. and Skjaervø, P. O. (eds) *Studia Grammatica Iranica. Festschrift für H. Humbach*, München: R. Kitzinger, pp. 407–424.
- (1986b [1987]) 'Sogdian *kw* and Slavonic *ku*', *Peredneaziatskij Sbornik* 4: 116–21.
- (1988) 'Syro-Sogdica III: Syriac elements in Sogdian', in Sundermann, W. et al. (eds) *A Green Leaf. Papers in Honour of Professor Jes P. Asmussen* (Acta Iranica 28. Hommage et Opera Minora, XII), Leiden: Brill, pp. 145–156.
- (1989a) 'Sogdian', in *CLI*, pp. 173–192.
- (1989b) 'Eastern Middle Iranian', in *CLI*, pp. 165–172.
- (1990) 'Chotano-Sogdica II: Aspects of the development of nominal morphology in Khotanese and Sogdian', in Gnoli, G. and Panaino, A. (eds) *Proceedings of the First European Conference of Iranian Studies*, Part 1, Roma: Istituto Italiano per il Medio ed Estremo Oriente, pp. 275–296.
- (1994) 'The triple system of deixis in Sogdian', *TPhS* 92(1): 41–53.
- (1996a) 'The Sogdian merchants in China and India', in Cadonna, A. and Lanciotti, L. (eds), *Cina e Iran*, Firenze: Leo S. Olschki, pp. 45–67.
- (1996b) 'On the historic present and injunctive in Sogdian and Choresmian', *MSS* 56: 173–189.
- (1996c) 'The Sogdian manuscripts in Brāhmī script as evidence for Sogdian phonology', in Emmerick, R. E., Sundermann, W., Warnke, I. and Zieme, P. (eds) *Turfan, Khotan und Dunhuang. Vorträge der Tagung 'Annemarie v. Gabain und die Turfanforschung', veranstaltet von der Berlin-Brandenburgischen Akademie der Wissenschaften in Berlin (9.–12. 12. 1994)*, Berlin: Akademie Verlag, pp. 307–315.
- (2000) *Bactrian Documents I*, London: Nour Foundation in association with Azimuth Editions and Oxford University Press.
- (2004) 'The Parthian abstract suffix *-yft*', in Penny, J. H. W. (ed.) *Indo-European Perspectives. Studies in Honour of Anna Morpurgo Davies*, Oxford: Oxford University Press, pp. 539–547.
- (2007) 'The Sogdian potentials', in Macuch, M. et al. (eds) *Iranian Languages and Texts from Iran and Turan. Ronald E. Emmerick Memorial Volume*, Wiesbaden: Harrassowitz, pp. 377–386.
- Sims-Williams, N. and Grenet, F. (2006) 'The Sogdian inscription of Kultobe', *Shygys (Almaty)* 2006.1 [2007]: 95–111.
- Sims-Williams, N. and Hamilton, J. (1990) *Documents turco-sogdiens du IX^e–X^e siècle de Touen-houang*, London: Published on behalf of Corpus Inscriptionum Iranicarum by SOAS.
- Telegdi, S. (1938) 'Notes sur la grammaire du sogdien chrétien', *JA* 230: 205–233.
- Weber, D. (1970) *Die Stellung der sog. Inchoativa im Mitteliranischen*, PhD dissertation, Göttingen.

- (1971) 'Also sprach Mandrī (Bemerkungen zur direkten Rede im Sogdischen)', *IF* 76: 77–83.
- (2000) 'Zur grammatisch-semantischen Bestimmung einiger Lemmata des sogdischen Lexikons', in Emmerick, R. E. *et al.* (eds), *Studia Manichaica IV. Internationaler Kongreß zum Manichäismus, Berlin, 14.–18. Juli 1997*, Berlin: Akademie Verlag, pp. 639–647.
- Yakubovitch, I. (2005) 'The syntactic evolution of Aramaic ZY in Sogdian', *StIr* 34(2): 199–230.
- Yoshida, Y. (1979) 'On the Sogdian infinitives', *Journal of Asian and African Studies* 18: 181–195.
- (1980) 'On the dialectology of Christian Sogdian [in Japanese]', *Oriente* 23(1): 83–93.
- (1994) 'Chinese in Sogdian script [in Japanese]', *The Toho Gakuho* 66: 271–380.
- (2000) 'First fruits of the Ryukoku-Berlin joint project on the Turfan Iranian manuscripts', *Acta Asiatica* 78: 71–85.
- (2004) 'Dating of linguistic changes encountered in the texts unearthed from the Silk Road [in Japanese]', *Ex Oriente* 11: 3–34.
- (2005) 'The Sogdian version of the new Xi'an inscription', in Trombert, E. and de la Vaissière, E. (eds) *Les Sogdiens en Chine*, Paris: École française d'Extrême-Orient, pp. 57–72.
- (2006) 'Honorific and polite expressions in Sogdian', in Shirai, S. and Shogaito, M. (eds) *Philological Studies on Old Central Asian Manuscripts*, Kyoto: Department of Linguistics, Kyoto University, pp. 81–94.
- (2009) 'Turco-Sogdian features', in Sundermann, W., Hintze, A. and de Blois, F. (eds), *Exegisti monumenta. Festschrift in honour of N. Sims-Williams*, Wiesbaden, Harrassowitz, pp. 571–585.
- Yoshida, Y. *et al.* (1989) 'A Sogdian sales contract of a female slave from the period of the Gaochang Kingdom under the Rule of Qu Clan [in Japanese]', *Studies on the Inner Asian Languages IV*: 1–50.

CHAPTER SIX

KHWAREZMIAN

Desmond Durkin-Meisterernst

1 INTRODUCTION

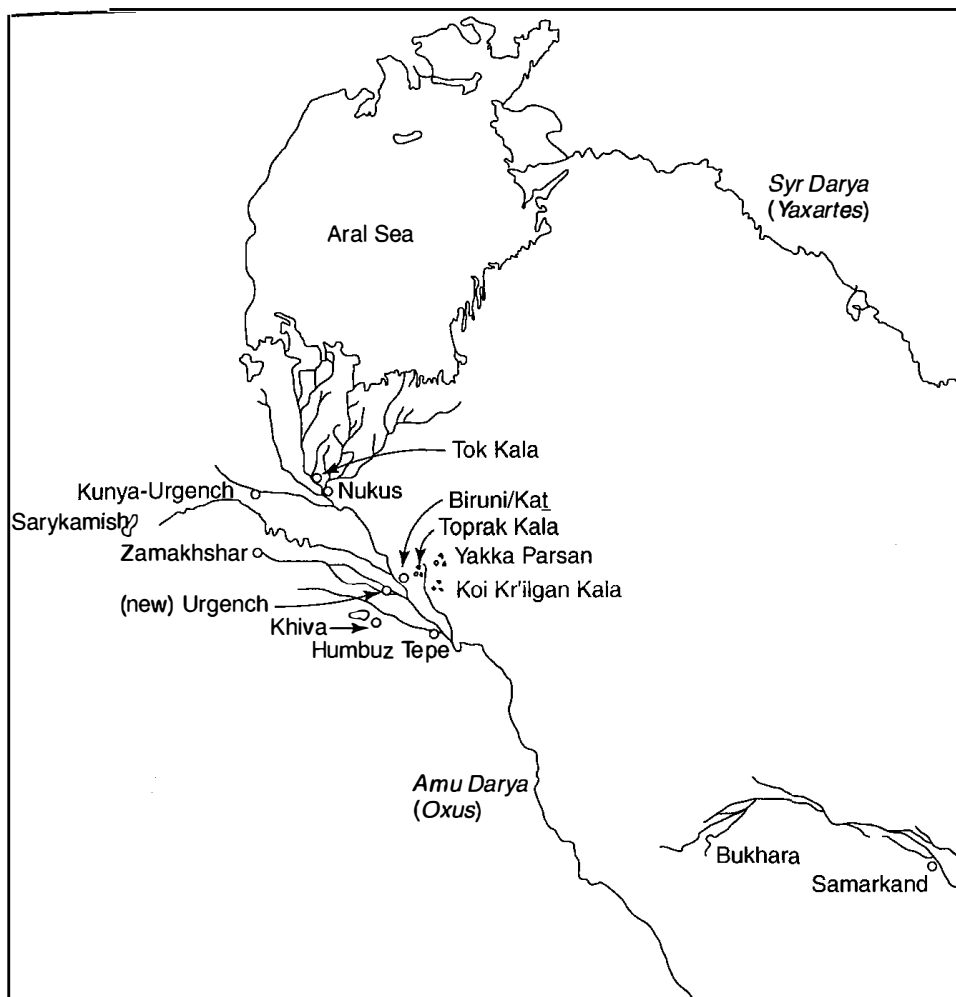
1.1 Overview

Khwarezmian is an Eastern Middle Iranian language formerly spoken around the lower reaches of the river Amu Darya (Oxus), including its delta into the Aral Sea in the north – the ancient Chorasmia. The Khwarezmian language area borders on the areas of Sogdian to the south-east and Parthian in the south, though in each case semi-steppe or desert intervenes. Khwarezmian shares features with Sogdian (note in particular the augment in the imperfect, see section 3.2.4.1a) but also with other East Iranian languages (verbal endings containing *r* occur only in Khwar., Avestan, Saka and Yaghnobi, Henning 1958: 117). The best attested form of the language is patently influenced by Arabic and New Persian. Not unlike most Middle Iranian languages, Khwarezmian came to light and to be studied only early in the twentieth century.

Khwarezmian is comparatively scantily attested, in two forms:

- 1 sources in a regional development of the Aramaic script: coin legends (Vainberg 1977), inscriptions on wooden boards (from Toprak Kala, see Livshits 1984), on ossuaries (from Tok Kala, see Tolstov and Livshits 1964) and on silver vessels found in the Urals, some leather documents (from Toprak Kala, see Livshits 1984) and on an ostrakon (from Humbuz Tepe, see Livshits and Mambetulaev 1986).
- 2 Islamic sources in Arabo-Persian script:
 - (a) quotations in books of case-law, namely the *Yatīmat ad-dahr* (YD) (645 AH 1246–7 CE, publ. Zeki Velidi 1927), the *Qunyat al-Munya* and the derivative *Risāla* (together QR) of ca. 755 AH 1354 CE (MacKenzie 1990: 5, ‘the Khwarezmian material in R. can be regarded as practically the most reliable that we have’);
 - (b) the Khwarezmian glosses in a copy of al-Zamaxšarī’s thematic Arabic dictionary *Muqaddimat al-Adab* (Zeki Velidi Togan 1951, Benzing 1968, MacKenzie 1970–72; the text is quoted as Muq., or simply by page and line references to the manuscript).

The sources date from the pre-Sasanian period (some of those in Khwarezmian script) to the Islamic period when, sometime after the fourteenth century CE, the language died out. The paucity of the pre-Islamic sources makes it difficult to form a picture of the language in these texts, though the older Khwarezmian script was apparently still in use in the eleventh century (Henning 1958: 58). Our main sources for Khwarezmian are those of the Islamic period (including older forms recorded by the polymath al-Bērūnī (362–



MAP 6.1 ANCIENT KHWAREZMIA

440 AH 973–1048 CE), a native speaker: (Henning 1958: 114–5). In the case of differences between the language of the pre-Islamic and the Islamic sources we can speak of Middle and Late Khwarezmian, but it should be remembered that Khwarezmian is attested neither in the Old nor the modern Iranian period. On the relative chronology of the language of the various sources, see Tolstov and Livshits 1964: 237–240.

There is no record of how long Khwarezmian survived our latest sources. The invasions of the Mongols (1220 CE) and Timur (1388 CE) devastated the area, the prosperity of which depended heavily on the maintenance of extensive and sensitive irrigation systems. Khwarezmian was replaced by Turkic languages (the use of Turkish is mentioned in QR 123).

Azami–Windfuhr (1972: 15–17, 36) suggested that Sangesari, while a member of the dialect groups circling the town of Semnan east of Tehran, includes a distinct subset of phonological, grammatical, and lexical features which it shares with an Eastern Iranian continuum that included Khwarezmian. (These include: $*\theta r > \dot{s}$ in $\dot{s}\dot{\alpha}$ ‘three’ and other items, Khwar. (\dot{s} y); the object marker $-d\dot{\alpha} < dar$, Khwar. $\dot{d}\dot{a}r\dot{\alpha}$; the potential construction

with auxiliary *bə-ker-*, Khwar. (k-); lexical items such as *šāš-* ‘dilapidate, scatter’, Khwar. (ššy-) ‘strew’; *āšax* ‘knucklebone, dice’, Khwar. (škk) ‘hoof’ (little shoe); *šəkēyīn* ‘woman’ (besides *žən*), Khwar. (šc(y)k) ‘female’; cf. also G. Windfuhr 1975.)

The limited extent of the material means that we have an incomplete picture of the language. For the following description the sources of the Islamic period are used. Khwarezmian appears to be of mixed origin, since it often exhibits more than one development of Old Iranian consonants (Henning 1958: 109–110); compare the different forms, with (š), (hr), and (cy) for **θr*, from the family ‘3, 13, 30’ (see section 3.1.6, Numerals).

1.2 Writing systems

As mentioned above the bulk of the Khwarezmian material is written in the Arabo-Persian script. The proper reading of numerous passages is complicated by the fact that the crucial diacritic dots, that serve to distinguish letters in the Arabo-Persian script are often neglected. In the case of isolated words, the correct reading is difficult to establish. In turn, there are also cases of mispointing, e.g. (n) for (y) in (mryndnd) for *m-arγend-eda* (m-rγnd-yd) ‘called’; underpointing (b) for (p) in (‘sb) for *asp* (‘sp) ‘horse’ and rarely overpointing, e.g. (p) for in (p’dr) for *ba-āder*; (b’dr) ‘without fire’.

New letters are: (1) (β) based on (f) with three dots (as in Early New Persian); (2) the pair (c,j) [ts, dz] based on *h* with 3 dots above (as in Pashto), e.g. (‘wrcy-) < OIr. **ā-wartaya-* and (βncy-, *βnjy-) ‘to tie’ < OIr. **bandaya-* (MacKenzie 1970, Glossary I: 543).

As mentioned below, among special uses of diacritics is that of the *shadda*, for gemination and stress.

In the earlier script Aramaic ideograms are used, e.g. ŠNT for *sarδ-* (srδ) ‘year’.

1.3 A note on transcription

The incomplete nature of the sources, the defective orthography used therein, the functional (and facultative) nature of the vowel marking in the Arabo-Persian script and uncertainty about whether unaccented syllables were reduced or lost means that we do not have enough information to transcribe Khwarezmian consistently and reliably. Providing even a tentative transcription risks falsifying the facts of the language. Therefore the basis for the following description must be the transliteration. Very often the transcription can do no more than indicate what could be expected (i.e. as an exercise in etymology or in formal analysis), which may not at all be what actually existed. Sometimes a transcription is avoided entirely. When it is offered it is always tentative, and is occasionally marked with an asterisk as being speculative, particularly to indicate problems of legibility or pointing and, in tables, to indicate reconstruction on the basis of documented forms. Old Iranian forms and likely earlier forms are marked here as ‘OIr.’ and ‘earlier’, respectively, with asterisk.

2 PHONOLOGY

2.1 Inventory and distribution

2.1.1 Vowels and diphthongs

2.1.1.1 Vowels

The vowels of Khwarezmian are defectively marked in both scripts. In the Arabo-Persian script, they may be indicated according to the traditional conventions: (1) both long and short vowels by the consonantal letters ʾ, y, w; (2) short vowels by the three superscript diacritics C^a, Cⁱ, C^u (supplemented by C^o for absence of vowel). Only before enclitics are short vowels more regularly written in order to distinguish inflectional endings:

ī . . . *xīwb-mi zādek*
 (ʾy . . . x^uwb-m z^ad^ok)
 ‘my beautiful child (masc.)’ QR 122;

yā xīwb-a-mi δuyδ-a
 (y^a xwb^a-mⁱ δ^uγ^od^a)
 ‘my beautiful daughter’ (fem.) QR 120.

Occurrences of Cⁱ or C^u in place of an expected *a* seem to indicate conditioned variants that may be indicated in the transcription by *e*, *o* respectively. However, this compromise is not extensively relied upon here.

Any attempt at establishing the phonetically distinct vowels based on such a narrow range of representational signs must therefore remain hypothetical, mostly based on parallels in related languages and sometimes on etymological considerations. With these caveats, it is possible to posit eight phonemic vowels:

TABLE 6.1: VOWELS

<i>ī</i>	<i>ē</i>	<i>ā</i>	<i>ō</i>	<i>ū</i>
	<i>i</i>	<i>a</i>		<i>u</i>

- a* *aβd* (ʾβd) ‘seven’; *aδwi* (ʾδw) ‘two’; *δast-* (δst) ‘hand’;
- ā* *ās-* (ʾs-) ‘to take’; *nāmek-* (n^am^ok) ‘name’;
- i* *iy-* (ʾy-) ‘go’; *pica* (pc) ‘father’;
- ī* *δīcek-* (δyck) ‘seen’; *nīθ-* (nyθ) ‘become’;
- ē* *ēw-* (ʾyw) ‘one’; *mēθ-* (myθ) ‘day’;
- u* *δum-* (δwm) ‘tail’; *uc-* (ʾc, ʾwc) ‘place’;
- ū* *βūm-* (βwm) ‘earth’; *βrūc-* (βrwc) ‘brows’;
- ō* *βōδ-* (βwδ) ‘scent’; *γōx-* (γwx) ‘ear’.

For simplicity the non-etymological prothetic vowel in e.g. (ʾps) ‘sheep’ may be represented by *a*, **aps-* though a reduced vowel such as shewa or even a front vowel *i* or *e*, since it seems that initial (ʾ) can express any short vowel, might be more accurate.

2.1.1.2 Semivowels and diphthongs

w: *wār-* (wʾr) ‘rain’; *yāwar-* (yʾwr) ‘time’; *γrīw-* (γryw) ‘self’;
y: *yīma(?)* (yīm) ‘I am’; *ayāc-a* (ʾyʿc) ‘come!’; *casūy-* (cswy) ‘*tasū* (weight)’.

2.1.1.3 Nasalization

It is likely that there was strong conditioned nasalization. This is indicated by the orthographical variation ($n \sim y$); e.g. ($\beta ndk \sim \beta ydk$), probably $< \beta \ddot{e}dek$ - ‘servant’, ($pnd'k \sim pyd'k$) $< p\ddot{e}d\ddot{a}k$ - ‘way’, ($rnj \sim ryj$) $< r\ddot{e}j$ - ‘toil’ (differently Henning 1958: 117), and the sporadic loss of final $-n$, that is indicated by variants such as (δyn) and (δy), probably $< \delta \ddot{e}$ ‘woman’.

2.1.2 Consonants

The inherited set of phonemic consonants appears to be as follows:

TABLE 6.2: CONSONANTS

	Labial	Dental	Palatal	Velar	Glottal
Plosive	$p \ b$	$t \ d$		$k \ g$	
Affricate		$c \ j$	$\check{c} \ \check{j}$		
Fricative	$f \ \beta$	$\theta \ \delta$		$x \ \gamma$	h
Sibilant		$s \ z$	$\check{s} \ \check{z}$		
Semivowel	w		y		
Nasal	m	n	(\ddot{n})	(η)	
Trill		r			
Lateral		l			

In the Arabo-Persian script, the letters (\check{s} , \check{c} , \check{t} , \check{z} , h , $'$), as well as postvelar (q) which are retained in Arabic loanwords, most likely do not represent a pharyngealized loan component in the inventory.

n may have had palatal and velar conditioned variants: $w\ddot{a}ng$ - ($w'nk$) ‘call’, and possibly palatals, e.g. $a\ddot{n}c$ - ($'nc$ -) ‘fear’.

h is often lost between vowels, e.g. $h\ddot{a}\beta r-n\ddot{e}-di < h\ddot{a}\beta er-na-hi-di$ ($h'\beta r-n-i-y-d'i$) ‘I gave it to you’; $\gamma\ddot{o}-\beta\ddot{a}rak$ - ($\gamma w-\beta'rk$) ‘earring’, from $\gamma\ddot{o}x$ - (γwx) ‘ear’ (Humbach 1989: 195).

Examples (not minimal pairs) are:

par - (pr) ‘on’, $bazak$ -(?) (bzk) ‘bad’;
 $f(i)c\ddot{u}r$ - ($fcwr$) ‘uncle’, $\beta(a)rak$ - (βrk) ‘fruit’;
 tan - (tn) ‘body’, $dr\ddot{o}\delta$ - ($\delta r w \delta$) ‘praise’;
 θayd - (θyd) ‘hard’, δast - (δst) ‘hand’;
 $sar\delta$ - ($sr\delta$) ‘year’, $z\ddot{i}w$ - (zyw -) ‘to live’;
 $cayd$ - ($c yd$) ‘entrance’, $\beta enjay$ - (βncy - $< \beta njy$ -) ‘to tie’;
 $\check{c}iy r$ - ($\check{c} y r$) ‘sharp’, $\check{j}uft$ - ($\check{j} ft$) ‘spouse’;
 $(\check{s}yt)$ ‘lip’, $\check{z}ad$ ($\check{z} d$) ‘resin’;
 $karc$ - (krc) ‘knife’, $garak$ - (grk usually written $krk!$) ‘flock, herd’;
 $x\ddot{a}\delta$ - ($x\delta$ -) ‘to sting’, $\gamma a\check{s}$ - ($\gamma \check{s}$) ‘tooth’, $hi\beta r$ - ($h\beta r$ -) ‘to give’;
 $ra\check{s}t$ - ($r\check{s}t$) ‘true’, $l\ddot{e}y$ - ($ly\gamma$) ‘tripping’;
 $magas$ - (mks) ‘fly’, $n\ddot{a}n$ - ($n'n$) ‘this’.

2.1.3 Syllable structure

Geminates constitute an integral part of the inventory, and are usually marked by *shadda* (indicated here by an over-bar), e.g. *ḍummen-* (ḍm̄n) ‘enemy’. But *shadda* seems also to be a device to mark stress, see section 2.2.1.

While many historical consonant groups were assimilated or reduced (e.g. *cam̄m-a* (cm) ‘eye’ < OIr. **čašman-*, *apan-a* (ʿpn) ‘co-wife’ < OIr. **hapaθnī-*), others have survived, albeit modified, such as initial *δβ*, e.g. *ḍβer-* (ḍβʿr) ‘door’ (<OIr. **dwarθi-*) and *sp* and *čk*, e.g. *spēdek-* (spydʿk) ‘white’, *čkēš-* (čkyš) ‘lie’ (<Avestan **tkaiša-*), and postvocalic *βd*, *γd*, e.g. in *aβδ* (ʿβd) ‘seven’ and *ḍuyd-* (ḍγd) ‘daughter’. More stable consonant groups are seen in *βrād-* (βrʿd) ‘brother’, *γrīw-* (γryw) ‘self’, *xubisk-* (xbsk) ‘own’, *wasn-* (wsn) ‘because of’, *ašt* (ʿšt) ‘eight’, *mizy* (mzy) ‘brain’, and *namaθk-* (nmθk) ‘salt’.

Prosthesis has occurred at different stages of the language, leading to such forms as *aβ-* (ʿβ-) < *b(a)w-* ‘be’, and *ak-* (ʿk-) < *k(u)n-* ‘do, make’.

2.2 Non-segmental features

2.2.1 Stress and unstable vowels

2.1.1.1 Unstable vowels

The incidence of consonant groups, and thus the retention or loss of the short vowels *i*, *u*, *a*, seems to depend on their position in the word. MacKenzie (1990: 94) sees ‘a general reduction of unstressed short vowels, producing consonant groups in internal but not word-initial position’, contrasting *sukund-* (sʰkʰnd) ‘oath’ QR 113 with **dʰ-skund-* (dʰsʰknd) ‘then an oath’ QR 251 < *da sukund*.

2.1.1.2 Position and marking of stress

Stress on the preceding vowel appears to be often, but irregularly, marked by the *shadda* though the system is not entirely clear (MacKenzie 1989: 270ff.). Most examples point to pre-final stress. (Stress is indicated here by an underlined vowel). When a suffix is added the stress shifts to the right, with apparent non-phonemic lengthening of the vowel: *ap̄an-a* (ʿpn) ‘co-wife’ > **pen-ā̄-mi* (pnʿm) fem. ‘my co-wife’ (QR 105; Henning 1956: 423).

2.1.1.3 Unstressed lengthening

However, a similar but perhaps unstressed lengthening is to be seen when the definite article *ī* is written together with a word with initial short vowel, without discernable difference; e.g. *ī + a* is written (y.) in **y-ā̄bar-ēw-* (y.ʿbryw) ‘the other’ < *ī abar-ēw-* QR 373, similar to *ī + ā* in *y-ā̄yat-* (y.ʿyt) ‘the beginning’ Muq. 61.2 < *ī āyat*. The latter word is clearly distinct in writing from *yā̄ yat-a* (y.ʿyt) ‘the running!’ 349.8. (In the foregoing a dot has been introduced in the analysis to separate items written together in the original.)

2.1.1.4 Pre-pausal position

Of particular interest is the indication of the vowel (y), presumably stressed (even anaptyctic), before the last consonant of a word in pause, or at the end of a sentence. (This pausal (y) is hereafter written superscript (ʸ). Compare, e.g. **zā̄dek* (zʸdk) ‘son’, in

pause *zādek* (z'dʸk), even *urg* ('wrk) 'wolf', pause *ureg* ('wr'k), and the verb *hāβer-da* (h'βrd) 'he gave', pause *hāβr-eda* (h'βr'd) (after Henning 1955: 44; 1958: 81). The phenomenon seems also to occur in the earlier script (Tolstov and Livshits 1964: 243–4).

2.3 Morphophonemic change

In addition to the variation discussed above, Khwarezmian, like other Iranian languages, has inherited a complex morphophonology, which is discussed together with morphology.

3 MORPHOLOGY

Khwarezmian is inflectional. With the noun, extensive use is made of prepositional and postpositional phrases, together with the weakly differentiated flexional endings. In the verb use is made both of inflectional endings and of suffixes added to the endings to distinguish further categories.

3.1 Nominal Morphology

3.1.1 Categories

3.1.1.1 Gender, number, and case

There are two genders, masculine and feminine, and two numbers, with traces of the dual in the number 'two' and in paired parts of the body, and a numerative. The basic distinctions between the two genders and the three numbers are most evident in the demonstrative pronouns presented below.

Gender, number, and case are marked by inflectional endings. But there is no inflectional distinction of gender in the plural, and no distinction between the nominative and accusative in either the singular or plural (but see section 3.1.4.1 postpositions). The plural has a basic and an extended form.

Even though the writing system is deficient, the inflectional system is well known and predictable. In addition, there are further clues:

- 1 frequent, but irregular diacritic vocalization by " (-), ' (-y), before enclitics, mainly personal suffixes;
- 2 palatalization, prominently of final *-k* (< OIr. **-aka-*) in adjectives and in perfect participles in (-dk, -tk), which reflects the OIr. feminine ending *-ī* and plural ending *-ayah* > *-ē* > *i*. Thus masc. (-k), fem. (-c), and masc. (-dʸk), fem. (-c): OIr. **barzaka-* > masc. *βažek-* (βžk), fem., pl. (βžc) 'long'; OIr. **mṛtaka-* > masc. **mādek* (mdʸk, fem. 'mc) 'dead'; masc. (**pr'dk*), fem. *parāc-a* (pr'c) 'divorced'.

Similarly, *-t*, *-d*, *-nd* are occasionally palatalized: fem. *yā δast-* 'the hand' > *f-ā δasc-ya* (f.' δsc-y) 'with the hand' 490.3, pl. *ī δasc-ina* (y' δsc-n) 'the hands' 387.3; *kāyad-* 'paper' > *c-ā kāyac-ya* (c. ' k'γc-y) 'from the paper' 9.4; (y' . . . 'kwnd) 'the finger' 374.3 > pl. (y' 'kwnc-n) 402.4.

In addition, a few feminines are derived by the formant *-ān-* ('n): *āsk-* ('sk) 'deer' > *āsk-ān* ('sk-'n) 'doe'; *urg-* ('wrk) 'wolf' > *urg-ān* ('wrk-'n) 'she-wolf'; *gavazn-* (γwzn) 'stag' > *gavazn-ān* (γwzn-'n) 'hind'. Similarly *hun-ān-* (h(w)n-'n) 'slave-girl', presumably 'Hunnish woman' (MacKenzie 1990: 112).

3.1.1.2 Animate terms

Animate terms, even though generally not overtly marked, are known to be feminine when referring to the female, or to the species as a collective (cf. MacKenzie 1971, Supplement: 30b), e.g. *aps* ('ps) 'sheep' is contextually masculine or feminine, collective; besides this a fem. *ameh-a* ('mh) 'ewe' exists. Gender is also lexically marked, e.g. (nrk) 'male' in (nrk h'kš) 'he-goat' Muq. 51.7–8.

3.1.1.3 Dual and numerative

A dual form, identical with the feminine singular, also occurs with '3' and '4', and so can be regarded effectively as a numerative (Sims-Williams 1979: 339–40). Due to its use for paired body parts, some of the masculine terms such as (*ī*) *δast* 'hand' and (*ī*) *pāδ* 'foot' have developed alternates that inflect as feminines, thus *yā δast-a*, *yā pād-a*.

3.1.1.4 Definiteness and indefiniteness

3.1.1.4a Definite article

Khwarezmian has a definite article, which patently originates in OIr. relative pronouns: masc. *ī* ('y), fem. *yā* (y'), dual *yā* (y'), plural masc., fem. *ī* ('y). These are written together with prepositions, e.g. (f.y, f.) probably indicating contractions *f-ī, f-ā < fa ī, fa yā*; and combine with a following vowel, both the fem. and masc. becoming (y-), e.g. *yā (a)sm-a* (y'sm^a) < 'the sky' QR 297, masc. *ī āyat* (y.ʔyt) < 'the beginning' Muq. 61.2.

The article can accompany names, e.g. *ī zayd* ('y zyd) 'Zayd' 516.1; *c-ī allāh-ān* (cy 'llh'n) 'Allah's' 103.5. Note also the article with dependent tonic personal and definite demonstrative pronouns after prepositions (MacKenzie, Glossary II, 1971: 88): *par-ī haβī* (pry hβy) 'to you' QR 85, *peš-ī nān-a* (pšy n'n) 'to that one' 238.1. (cf. *par-ā βūm-a* (pr' βwm) 'to the ground' 215.7).

3.1.1.4b Indefiniteness

Indefiniteness may be marked by *ēw-* ('yw) 'one'. Non-specific indefiniteness in both singular and plural is indicated by the absence of the articles and possessives.

3.1.1.5 Nouns and inflectional classes

The cases (nominative-accusative, genitive, possessive, ablative, locative) are partly distinguished by inflection alone and partly in combination with pre- or postpositions. Thus, the dative and the definite direct object are marked by the genitive or possessive + *δāra*. The essentials of the following paradigms were established by Henning (1955: 46 n.1).

TABLE 6.3: NOMINAL INFLECTION

	Masc.	Fem. Fem.	-k Stems	Pl.	-k Stems	extended
NOM-ACC	–	-a	-ka	-i	-ci	-ina
GEN	–	-ya	-ca	-i	-ci	-ina
POSS	-ān	-ya	-ca	-ān	-k-ān	-in-ān
ABL	-a	-ya	-ca	-i	-ci	-ina
LOC	-a	-a	-ka	-i	-ci	-ina

Generally the consonants and the long vowels are written. Short vowels are occasionally written before suffixes or postpositions (e.g. *βūm-ya δārā* (βwm-y' -δ'r) but final short vowels are unwritten except, sporadically, in the vocalized material.

3.1.2 Adjectives and comparison

Adjectives as a word class are generally unmarked, e.g. *xūb* (xwb) 'good', but there are typical derivational formants, such as (-ynk, -'n(°)k, -myn(°)k), and the prefixes (θ-) with and privative (mnd-) (see section 6 Lexis).

The comparative is formed with the suffix *-dar* (-dr), e.g. *bazag-dar* (bzk-dr) 'worse'; with the preposition *ci-* (c-) 'than', e.g.

nē(n) bazag-dar ci nān-a (ny bzk-dr c.n'n)
'this one (is) worse than that one' 280.6.

A superlative suffix is not attested.

3.1.3 Pronouns and deixis

3.1.3.1 Personal pronouns

Personal pronouns are as listed by MacKenzie (1993: 136). In the following table, a hyphen indicates pre- or suffixation. The sign '+' in the left column indicates that the form of the pronoun precedes this, e.g. (m'r) occurs with the postposition (δ'r): *mār δāra* (m'r δ'r). The pronoun of the 3s is not varied for gender.

The object suffixes are directly attached to the finite verb forms; e.g. *δāmb-eda-ma* (δ'nbyd'm) 'he hit me', etc. (δnby-) 'hit'; (cf. Henning 1958: 117 n.4; Samadi 1986: 66).

3.1.3.2 Demonstrative pronouns

These distinguish between near and far deixis, and a topical, deictically neutral pronoun 'such a one, the one' (Muq. 65.4–8, 66.1–7). All inflect for gender, number, and case. The feminine of the latter is notably distinct from the masculine in both singular and plural. A rarer near-deictic is *nyš*.

3.1.3.3 Interrogative and relative pronouns

The variants with (°) contain the interrogative particle also used in questions, but the distribution of interrogatives with (°) and relatives without seems to be disturbed.

3.1.3.4 Reflexive and reciprocal pronominals

Reflexive: **x(u)dāk-* (xd'k) 'self, same; he/she'; *γrīw-* (γryw) 'soul, self, he/she'; adjective: *xubisk-* (xbsk) 'own'; reciprocal: **any* ('ny) 'other', (nywny) 'each other'.

3.1.4 Adpositions

3.1.4.1 Pre- and postpositions

According to the cases employed with them, the adpositions fall into several sets. The postpositions occur prominently with clitic pronouns, for which see Table 6.4 Pronouns.

TABLE 6.4: PERSONAL PRONOUNS

	1s	2s	3s	Indef.	1p	2p	3p
NOM	<i>nāz</i>	<i>autak</i>		<i>ci</i>	<i>maβī</i>	<i>haβī</i>	<i>*ēβī</i>
suff.		<i>tak</i>		<i>ci</i>			
G-D, POSS	<i>mi</i>	<i>di</i>	<i>hi</i>		<i>mn(a)</i>	<i>*fn(a)</i>	<i>hin</i>
suff.	<i>-m(i)</i>	<i>-d(i)</i>	<i>-h(i)</i>		<i>-men</i>	<i>-fen</i>	<i>-hin</i>
<i>fā, par, ci</i> +	<i>-mika</i>	<i>-fβika</i>			<i>ī maβy-ān</i>	<i>ī haβy-ān</i>	<i>ī ēβy-ān</i>
<i>paš</i> +	<i>mici</i>	<i>*-fβi-ci</i>			<i>ī maβy-ān</i>		
+ <i>ḏāra</i>	<i>ma-ra-</i>	<i>tewa-ra-</i>			<i>maβy-ān-</i>	<i>*haβy-ān-</i>	<i>ēβy-ān</i>
ABL	<i>ma-c(i)</i>	<i>tewa-ci</i>			<i>mna-ci</i>	<i>*fna-ci</i>	
ACC	<i>ma</i>	<i>falβa</i>	<i>hi</i>	<i>ci</i>			<i>-na</i>
suff.	<i>-m(a)</i>	<i>-fβ(a)</i>	<i>-h(i)</i>	<i>-c</i>	<i>-men</i>	<i>-fen</i>	<i>-n(a), -(h)in(a)</i>
+ <i>bir, beš</i>	<i>ma-</i>	<i>fβa-</i>	<i>(h)i-</i>	<i>ci-</i>	<i>mn-a-</i>	<i>fn-a-</i>	<i>hina-</i>
+ <i>ci, ṭa</i>							
Transliteration							
	1s	2s	3s	Indef.	1p	2p	3p
NOM	(nʻ)	(ʻ(w)tk)		(cy)	(mβy)	(hβy)	(*ʻyβy)
suff.		(-tk)		(-c(y))			
G-D, POSS	(my)	(dy)	(hy)		(mn, mnʻ)	(*fn, *fnʻ)	(hyn)
suff.	(-m(y))	(-d(y))	(-h(y))		(-mn)	(*-fn)	(-hyn)
<i>fā, par, ci</i> +	(-mʻk)	(-β/fʻk)			(-y mβyʻn)	(-y hβy)	(-yβy)
<i>paš</i> +	(-mʻc)	(*-β/fʻc)			(-y mβyʻn)		
+ <i>ḏāra</i>	(mʻr)	(twʻr)			(mβyʻn)	(*hβyʻn)	(ʻy-βyʻn)
ABL	(mʻc(y))	(twʻc)			(mnʻc)	(*fnʻc)	
ACC	(mʻ)	(β/fʻ)	(hy)	(cy)			(nʻ)
suff.	(-mʻ)	(-β/fʻ)	(-h(y))	(-c)	(-mn)	(-fn)	(-nʻ), (-hyn, -ynʻ)
+ <i>bir, beš</i>	(-mʻ-)	(-β/fʻ-)	((h)y-)	(cy-)	(-mnʻ-)	(-fnʻ-)	(hynʻ-)
+ <i>ci, ṭa</i>							

TABLE 6.5: DEMONSTRATIVE PRONOUNS

	Near		Far		Topical	
	masc.	fem.	masc.	fem.	masc.	fem.
Singular	<i>nē(n)</i>	<i>nēn-a</i>	<i>nāwer</i>	<i>nēd-a</i>	<i>nā(n)</i>	<i>nān-a</i>
Dual	<i>nāw-a</i>	<i>nāw-a</i>	<i>nāwer-a</i>	<i>nēd-a</i>	<i>nān-a</i>	<i>nān-a</i>
Plural	<i>nāw-i</i>	<i>nāw-i</i>	<i>nāwer-i</i>	<i>nēd-i</i>	<i>nān-i</i>	<i>nān-i</i>
Singular	(ny(n))	(nyn)	(nʻwʻr)	(nyd)	(nʻ(n))	(nʻn)
Dual	(nʻw)	(nʻw)	(nʻwyr)	(nyd)	(nʻn)	(nʻn)
Plural	(nʻw)	(nʻw)	(nʻwyr)	(nyd)	(nʻn(y))	(nʻn)

TABLE 6.6: INTERROGATIVE AND RELATIVE PRONOUNS

Interrogative	Relative				
<i>a-ki</i>	((ʻ)ky)	‘who?’	<i>ki</i>	(k)	‘who?’
<i>a-ci</i>	((ʻ)c(y))	‘what?’	<i>ci</i>	(c)	‘that, which’
<i>a-*kdām-</i>	((ʻ)kdʻm)	‘which?’	<i>*akdām</i>	((ʻ)kdʻm)	‘each, any’

Prepositional *θa* (θ) 'with' tentatively included here is actually attested only in compounds (see section 6.1, Word formation). The postposition *-δāra*, which expresses both dative and definite direct object, takes the genitive in the singular, but the possessive *-ān* in the plural.

TABLE 6.7: ADPOSITIONS AND CASES

Prepositions			Postpositions				
Locative			Ablative				
<i>par</i>	(pr)	'on, for'			<i>-ber-</i> (-b'r)	'on'	
<i>piš</i>	(pš)	'beside'			<i>-biš-</i> (-bš)	'before, near, with'	
<i>wasn-</i>	(wsn)	'because of'			<i>-da</i>	(-d) 'forth'	
<i>pa</i>	(p)	'at, for, by'	<i>pac</i>	(pc)	'behind, after'		
<i>fa</i>	(f)	'in, at, to, into'	<i>fa</i>	(f)	'with, for, by'	<i>-wa</i>	(-w) 'therein'
			[<i>θa</i>	(θ)	'with']	<i>-θa</i>	(-θ) 'with'
			<i>-ci</i>	(c)	'from, than'	<i>-ci</i>	(-c) 'from'
Possessive genitive			Genitive/possessive				
<i>paš-</i>	(pš)	'after'			<i>-δāra</i>	(δ'r ^a)	

3.1.4.2 Circumpositions

There are numerous combinations with prepositions and nominal quasi-postpositions, such as:

- ci – sār* (c. – s'r) 'for' (*sār* 'towards');
fa – anbec (f. – 'nbc) '(together) with, with (regard to)' (*anbec* 'together');
fa – parama (f. – prm) 'until' (*parama*).

Examples: [*c(i)*] *ī zarnī [sār]* (c.y zrny s^ar^o) 'for the gold' QR 389; [*f(a)*] *ī *xdāk-a [anbec]* (f.y xd'k 'nbc) 'with him' QR 368; [*f(a)*] *ādīnak-a [parama]* (f.'dyнк prm) 'until Friday' QR 174. Syntactically, such adverbial noun phrases are in the ablative or locative.

3.1.4.3 Nominal inflection and adpositions

The following examples demonstrate the nominal inflection, including some of the uses of prepositions and pronominal suffixes. Square brackets mark case endings.

3.1.4.3a Masculine singular

- NOM: *ī asp[-θ]* (y 'sp) 'the horse' 107.2;
 ACC: *ī xsin[-θ]* (y xs'n) 'the thing' 82.2;
 GEN: *c-ā mađān-ya-hi kām[-θ]* (c.' mđ'n-y.h k'm) 'from (the inside of) his mouth' 103.6;
 POSS: *ī asp[-ān]* (y 'sp-n) 'of the horse' 38.5;
 ABL: *c-ī kām[-a]-hi* (c.y k'm-.h) 'from his mouth' 481.5; *f-ī haqq[-a]-hi* (f.y ḥq-.h) 'for his right' 212.5;
 LOC: *f-ī asp[-a]* (f.y 'sp) 'on the horse' 343.7; *f-ī kām[-a]-hi* (f.y k'm-.h) 'into his mouth' 351.5;
 DAT: *ī bādēs-hi [δāra]* (y b'δys.h δ'r) 'to his command' 498.8;
 Def. Obj. *ī aps [δāra]* (y 'ps δ'r) 'the sheep' 305.8.

3.1.4.3b *Masculine dual/numerative (=feminine), with numbers 2–4*

- NOM: *yā-(a)δw[-a] xsin-a* (y'δw xs'n) 'the two things' 483.7; *yā caḡār(a) pād-a spēdec[-a]* (y' c'f'r p'δ spy^{dy}c) 'the four legs white' 38.1;
 LOC: *f-ā šē pād[-a]* (f.' šy p'δ) 'on three legs' 220.1.

3.1.4.3c *Masculine plural*

- NOM: *ī *ušš[-i]* (y' š) 'the camels' 130.6; *ī sanc[-i]-hi* (y' snc-y.h) 'its stones' (*sang (snk)); 36.3; *ī pād[-ina]* (y' p'δ-n) 'the legs' 387.3; – *pl. tantum: ī jumleg[-i]* (y' jwmlk) 'the totality, all' 63.4;
 ACC: *ī *ušš[-i]* (y' š) 'the camels' 302.4; *āh[-i]* ('h-') 'sighs' QR 312; *akt[-ina]* ('ktn) 'acts, deeds ('kt)' 251.7;
 POSS: *paš-ī ps[-in-ān]* (pš.y ps-n-³n) 'after the sheep' 103.8 (MacKenzie, Glossary I, 1970: 547);
 ABL: *c-ī xsn[-i]-mi* (c.y xsn-y.m) 'from my things' QR 253; *f-ī βandec[-i]-hi* (f.y βndc-y.h) 'with his servants (βndk)' 330.2;
 LOC: *f-ī pac.xās[-i]-hi* (f.y pcx's-y.h) 'in his clothes' 192.3; *f-akt[-ina]-hi* (f.ktn-³.h) 'for his deeds' 148.7;
 DAT: *ī wād[-i δāra]* (y' w'd-¹ δ'r) 'to the winds' QR 167;
 Def. Obj. *ī jumleg[-ān δāra]* (y' jwmlk-³n.δ'r) 'the totality, all' 62.5.

3.1.4.3d *Feminine singular*

- NOM: *yā aps[-a]* (y' 'ps) 'the ewe' 326.1; *yā camm[-a]-hi* (y' cm-².h) 'his eye' 110.2;
 ACC: *yā asm[-a]* (y' 'sm) 'the sky' 359.1; *yā badw[-a] -hi* (y' bdw-².h) 'his property' 298.3;
 POSS: *yā *ušš[-ya]* (y' š-y) 'of the camel', *yā γōc[-a]* (y' γwc) (palatised) 'of the cow (γōk (γwk))', *yā (a)ps[-ya]* (y' ps-y) 'of the ewe' 51.2;
 ABL: *f-ā camm[-ya]* (f.' cm-y) 'with the eye' 88.7; *c-ā βūm[-ya]* (c.' βwm-y) 'from the earth' 96.4; *c-ā pād-ya]-hi* (c.' p'δ-y³.h) 'from his foot' 223.4;
 LOC: *par-ā *pišk[-a]-hi* (pr.' pšk-³.h) 'on its back' 48.1;
 DAT: *y(ā) ārc[-a δāra]* (y'rc δ'r) 'for work ('rk)' (palatised) 399.5;
 Def. Obj. *yā βūm[-ya δāra]* (y' βwm-y³-δ'r) 'the earth' 164.1.

3.1.4.3e *Feminine plural*

- NOM: *ī camn[-a]-hi* (y' cmn-².h) 'his eyes' 98.1, which retains an old dual (OIr. *čašman-ā*); – *pl. tantum: ī hūn[-i]* (y' hwn-y) 'blood' 117.3;
 ACC: *ī šuwec[-i]* (y' šw³c) 'the nails (šwk)' 112.8; *ī γōx[-ina]* (y' γwx-n) 'the ears' 203.3; – *pl. tantum ī cūb[-i]* (y' cw b) 'the water' 98.3;
 GEN: *f-ā maδān-ya paθak[-ina]* (f.' mδ'n pθk-n) 'in the midst of the houses' 219.6;
 POSS: *ī *smn[-ān] ud ī βūm[-in-ān]* (y' smn-³n 'wd' y' βwm-n-³n) 'of the heavens and the earths' 354.8 (MacKenzie, Glossary III, 1971: 316);
 ABL: *c-ī paθak[-ina]* (c.y pθk-n) 'from the houses' 238.2;
 LOC: *f-ī badw[-ina]* (f.y bdw-n) 'in the properties' 484.8.
 Def. Obj. *ī cūb[-ān δāra]* (y' cw b-³n δ'r) 'the water' 309.1 (MacKenzie, Glossary II: 1971, 85);

Several isolated examples appear to diverge from the paradigms which cannot be detailed here.

3.1.5 Adverbs and adverbial phrases

Some adverbs are simply lexical items without specific formant, e.g. (brsm) ‘very’, (mθ) ‘here’; So are adjectives used adverbially. However, there are denominal adverbs derived by (-wy^a); e.g. (z²d^ok-³w^oy^a) ‘as a child’ QR 116 < (z²d^ok); (mr²yk-³wy) ‘hypocritically’ 430.8 < (mr²yk) ‘hypocrisy’; (nmd-³wy) ‘in honour’ 481.8 < (nmd) (MacKenzie 1990: 98).

3.1.6 Numerals

3.1.6.1 Cardinal numbers

Cardinal numbers (listed in Muq. 31.5–34.1, including very high combined numbers).

TABLE 6.8: CARDINAL NUMBERS

1–9	11–19	10–90	100–900	Other
(¹ yw) ēw	(¹ ywnd ² s)	(δ ² s)	(s ² d) <i>sed</i>	21 (¹ yw-dws ² c)
((¹)δ ² w)	(¹ δw ² s)	(¹ ws ² c)	(¹ δwysd)	1000 (hz ² r) <i>hazār</i>
(šy)	(*hrδ ² s)	(š ² ys)	(šyzd)	
(cf ² r)	(cwrδ ² s)	(cf ² r ² s)	(cf ² rzd)	
(pnc)	(pnš)	(pnc ² s)	(pnšd)	
(¹ x)	(¹ xrδ ² s)	(¹ x ² c) <i>uxxic</i>	(¹ xzd) <i>uxxuzda</i>	
(¹ βd)	(¹ βdδ ² s)	(¹ βd ² c)	(¹ βdzd)	
(¹ št)	(¹ štδ ² s)	(¹ št ² c)	(¹ štzd)	
(š ² δ) *šāδ	(nw ² δ ² s)	(nwyc)	(nwzd)	

3.1.6.2 Ordinal numbers

Ordinals, except for ‘first’, are marked by (-²m): (ftm-yc²k) ‘1st’, (δβc-²m) ‘2nd’, *šēyem* (šym, š²ym) ‘3rd’, (cf²r-²m) ‘4th, quarter’, (pnc-²m) ‘5th, a fifth’, (¹xt-²m) [sic] ‘6th, a sixth’, (¹βd-²m) ‘7th’, (¹št-²m) ‘8th’, (š²δ-²m) ‘9th’, (δs-²m) ‘10th’.

3.1.6.3 Fractions

Fractions are usually identical with the ordinals as listed (e.g. fifth/a fifth), but note (¹rc-y²dk) ‘a third (part)’.

3.1.6.4 Distributives

Distributives are plural forms, always with (-c), e.g. (¹δw-c) ‘(by) twos’ (1–10 listed Muq. 34.8–35.1).

3.1.6.5 Multiplicatives

Multiplicatives are formed with (-δβ²γ) ‘-fold’, e.g. ((¹)δw-, šy-, cf²r-δβ²γ) ‘two-, three-, fourfold’ 47.4, 61.8. Note also (šy²wr) (almost always (š²γ²wr) ‘three times (y²wr)’ and (š²γk) ‘triple (divorce)’ (or perhaps ‘triad’).

3.2 Verb morphology

Verbs are inflected for tense, mood, and aspect. In addition there are periphrastic constructions, and modifying particles. Like other Iranian languages, Khwarezmian shows complex morpho-phonemic alternation, in the main reflecting Old Iranian distinctions of stem formation. The present stem (which includes the imperfect) is inflected as a verb, the past stem is a nominal formation.

3.2.1 Stem formation

3.2.1.1 Present and imperfect stems

There are two inflected stems, present and imperfect, the latter distinguished (as in Sogdian) by some form of the ‘augment’ (see section 3.2.5.1a Imperfect). In the limited sources available most verbs are only attested in the imperfect.

The present stems form three classes. The difference is most obvious in the 3s forms of the imperfect.

- (1) stems in consonants, 3s *-eda* (-^vd) (< OIr. **-a-ta*):

(^k-) ‘do’ > *m-ak-eda* (m-k-^vd) ‘did’;
 (βx-) ‘give’ > *bāx-eda* (β^x-^vd) ‘gave’.

This *-d* is devoiced by contraction after many stems in *-s*, *-š*, and even *-rd*, *-nd* (*-dd* > *-t*), e.g. (nγs-) ‘arrive’, (m-nγs-d/t-, m-nγs-t). The frequently attested, but unique (ws) ‘he said’ is probably an allegro form of **was-ta*.

- (2) stems in ^oy-, 3s *-(a)y-eda* (-y^vd) (< OIr. **-(a)ya-ta*):

(*čy-) ‘be(come) thirsty’ > *m-ačy-eda* (m-čy-^vd) ‘became thirsty’;
 (w^{ry}-) ‘rain (tr.)’ > *wār-ay-eda* (w^{ry}-^vd) ‘it rained’.

- (3) irregular stems in *-ā*, *-ah*, and the root *i-* ‘go’ with preverb, 3s *-ay-ta* (-yt):

(šm^h-) ‘command’ > *šāmay-ta* (š^{my}-t) ‘he commanded’ (< OIr. **fra-mā-*);
 (cy-) ‘enter’ > *cāy-ta* (c^y-t) ‘he entered’ (< OIr. **ati-ay-*).

These verbs also have an exceptional 3rd pl. in *-la*, e.g. *cāy-la* (c^{yl}) ‘they entered’ (cf. regular *-āra*). The same opposition *-t* : *-l* is also found in the present of the substantive verb *y-* ‘be’, which presents several unique forms (see Table 6.13).

3.2.1.2 Intransitive stem classes and derived causatives

The following stem pairs exemplify the complex pattern of inherited and productive stem alternation, including the pervasive opposition between intransitive and passive vs. transitive and causative (cf. MacKenzie, Glossary IV, 1972: 532–533). The marked intransitives reflect Old Iranian ‘inchoatives’ in *-s-*, and passives in *-ya-*. The marked transitives derive from old causatives in *-aya-*, with lengthening of the short root vowel, and, in turn, derive secondary causatives from the ‘inchoatives’. This process is productive and applies even to loans. Denominative verbs may include loans, such as (bsmr-) ‘slaughter’ < Ar. *bismillāh-*; (tfsyr-y-) ‘comment’ < Ar. *tafsīr* ‘commentary’; (čkyš-) ‘tell lie’ < Av. **tkaeša-* ‘false belief’.

TABLE 6.9: STEM DERIVATION

Intransitive & Passive		Transitive & Causative	
(xwr-)	'eat'	(x'r-y-)	'make eat, feed';
(nč-)	'be afraid'	(nč-m-y-)	'frighten';
(rh-)	'get free'	(rh-n-y-)	'set free';
(pcmyz-)	'flow'	(pcmyz-y-)	'make rain';
"Inchoatives"			
(pcm'x-s-)	'get dressed'	(pcm'c-)	'put on';
(xl's-)	'be grasped'	(xl-)	'grasp';
((n)γs-)	'reach'	((n)γ's-y-)	'make reach';
(yx-s-)	'be arranged'	(yγz-y-)	'arrange';
Passives			
(zr-y-)	'be tormented'	(z'r-y-)	'torment';
(βsy-)	'be bound'	(βnc-y-)	'bind';
(nbzy-)	'lie down'	(n'b'z-y-)	'lay down';
(hncy-)	'rest'	(hnc-w-y-)	'make rest';
Denominative stems			
(knbw-)	'become less'	(knb'w-y-)	'make less'.

3.2.1.3 Compound verbs

Various verbs in combination with a noun may form compound verbs. With *ak-* 'make' the noun is often a loanword (Henning 1971: 23b (g)), though not exclusively so, e.g. Ar. *inkār* ('nk'r) 'denial', *hadīθ* (hdyθ) 'telling, report', *naḍr* (nḍr) 'vow', *nikāḥ* (nkḥ) 'marriage' + *ak-* 'reject, speak, vow, marry', but also Khwar. *βend-* (βynd) 'husband' + *ak-* 'marry, take as husband'. With ('fy) 'dear', (nnk) 'shame' (Pers. *naṅg*), (šmnk) 'hope', (wf'y) 'faith' (Ar. *wafā'*) + *ḍāray-* 'hold' we have 'like (hold dear), be ashamed, hope, keep faith' (MacKenzie 1971, Glossary IV: 536).

With masc. abstract nouns in *-āw(e)k-* (-'w'k) a distinction appears to be made between a specific with singular noun (in *-k-*), and a more general meaning with the plural form (in *-ci*): (pcpkk-'wk hy-θ m-k-d) 'he opposed him' (lit. 'made opposition with him') 430.3, but ('pcpkk-'wc m-k-r) 'they contended (lit. 'made oppositions') 491.8.

3.2.1.4 Directional adpositional particles

Verbs may combine with directional-locational adpositions, which may appear before or after the verb as enclitics, usually in enclitic chains (see section 4.2.2.3). They include the following:

-ci (-cy)	'there-from, out':	<i>ās-</i> -ci ('s- -c)	'take from, out';
-da (-d')	'out, away, forth':	<i>ak-da</i> ('k-d')	'perform; say',
		<i>da was-a</i> (d' ws)	'he spoke out';
-wa (-w')	+ <i>fa</i> NP 'there-into':	<i>wa ak-</i> (w' 'k-)	'put into',
		<i>wa cāy-ta f-i wāc</i> (w' c'yt (fy w'c))	'he entered into (the matter)' 357.4;
-waθa (-wθ)	'there-inside':	<i>māneda-wa-θa</i> (m'nyd(')wθ)	'he lived therein'.

3.2.2 Nominal forms

The nominal forms of the verb can be divided into those based on the present stem and those based on the past stem.

3.2.2.1 Present stem

3.2.2.1a Present participles

The productive derivation is by (a) *-nek* (-n^yk), e.g. ((^o)βzp-n^yk) ‘shining’, (‘y-nk) ‘going’; or (b) *-en(e)k* (-yn^yk), e.g. (nps-ynk) ‘writing, writer’, (y’sw-yn^yk) ‘going, goer (away)’ < ((y)sw-) ‘go’. A few archaic present-participle-like forms ending in (-nd(k)) occur, such as (‘wšy-nd) ‘hungry’ (cf. (‘wšy-) ‘be hungry’), (zyw-nd^yk) ‘living’. Both these endings have fem. and pl. forms in (-nc).

3.2.2.1b Verbal nouns

Verbal nouns, or ‘present infinitives’ in *-āk* (-^yk) are feminine; e.g. *āβōzay-āk* (‘βwzy-^yk) ‘smelling’ (see section 6.1 Word Formation).

3.2.2.2 Past stem

3.2.2.2a Normal past participles and ‘past infinitives’

These participles and ‘past infinitives’ end in a dental, *-d* or *-t* (those in *-st* < **-tt* or < **-št*), and usually originate in an OIr. zero grade stem. The infinitives function in the potential construction (see section 4.3.2.3a), the participles mostly in compounds, e.g. (βyrd-^yδyk) ‘successful’ (lit. ‘having acquired desire’) < (βyr-) ‘acquire’.

3.2.2.2b Perfect participles

Perfect participles are formally derived from the past participle, taking final *-ek* (-k): (-dk, -tk). They have regular pl. forms in *-ci*: *-deci* (-d^yc), whereas assimilation in the feminine singular yields just (-c). They function in perfect tenses (see section 3.2.4.2), and as independent forms. Examples: *akt-ek* (‘tkk) ‘done’ < *ak-* ‘do’; *edek* (‘ydk) ‘gone’ < *iy-* ‘go’. Some participles have a nasal formant, (-nk), pl. (-nc) (MacKenzie, Glossary I, 1970: 552).

Several other derived forms are found, such as participial adjectives in *-c-ik* (-c^yk) and feminine abstract nouns in *-c-ik* (-cyk) < **-ūkā-* (MacKenzie 1972, Glossary V: 1972: 69).

3.2.2.3 Morphophonology of tense stems

In overview, the various morphophonological patterns of the two pairs of stems may be exemplified as follows (for the marking of the imperfect, see discussion section 3.2.4.1a):

TABLE 6.10: MORPHOPHONOLOGY OF TENSE STEMS

Present Stem		Participles			
Present	Imperfect	Past part.	Perf. part.		
	<i>m-</i>				
<i>ak-</i>	<i>m-ak-</i>	<i>akt-</i>	<i>akt-ek</i>	'do'	< OIr. * <i>kīn-l*kī-ta-</i>
(k-)	(m-k-)	(kt-)	(()k ^h tk)	'take'	< OIr. * <i>ā-ya-sa-l*ā-ya-ta-</i>
(s-)	(m-s-)	(yt-)	(*ytk-)	'go'	
(y-)	(m-y-)	(yd)	(ydk)	'beat, pound'	
(kwnd-)	(m-kwnd-)	(*kwst)	(kstk)		
	Lengthening				
(βyr-)	(β'r-)	(βyrd-)	(*βyrdk)	'obtain'	
(pcrm-)	(p'crm-)	(*pcrm'd)	(pcrm'dk)	'rest'	
Suppletive Stems					
(s-)	(m-s-)	(yt-)	(ydk)	'come'	< OIr. * <i>ā-i-sa-l*ā-ga-ta-</i>
(wyn-)	(wyn-)	(*δyc)	(δyck)	'see'	< OIr. * <i>waina-l*dī-ta-</i>

3.2.3 Person marking

The personal endings show tense and mood distinctions. Note that the 2p endings appear to be patterned on those of the 2s, that the subjunctive forms share *-ā-*, and that the impf.-inj. forms are marked by *-a*. (As indicated, in the writing system the use of vowel signs depends on whether the verb is in pause or not, so that before a suffix the final vowel of an ending may also be written *plene*.)

The basic inflected forms, to which modal and aspectual suffixes are added, are the following (see also Table 6.13 'be, become'):

TABLE 6.11: PERSON MARKING

	1s	2s	3s	1p	2p	3p
PRS	<i>-ām-i</i>	<i>-i</i>	<i>-ica</i>	<i>-āmn-i</i>	<i>-βi</i>	<i>-ār-i</i>
SBJ	<i>-ā, -ān</i>	<i>-īx</i>	<i>-āc</i>	<i>-āmin-i</i>	<i>-β-īx</i>	<i>-ār-i</i>
IPF/INJ	<i>-in-a</i>	<i>-ix-a</i>	<i>-ed-a</i>	<i>-āmn-a</i>	<i>-β-a</i>	<i>-ār-a</i>
IMP		<i>-a</i>			<i>-β-a</i>	
OPT						<i>-īr</i>
Transliteration						
	1s	2s	3s	1p	2p	3p
PRS	(- ^h m ^(h))	(- ^(h))	(- ^h c, -c ^(h))	(- ^h mn(y-))	(-f(y)-)	(- ^h r(y-))
SBJ	(- ^h (n))	(-yx ^(h))	(- ^h c ^(h))	(- ^h m ^h n ^h)	(-βyx)	(- ^h r ^(h))
IPF/INJ	(- ^h n, n ^(h) -)	(- ^h x ^h , - ^(a))	(- ^h d ^(h) -)	(- ^h mn ^h)	(-f)	(- ^h r)
IMP		(- ^(a))			(-β ^(h))	
OPT						(-yr)

The 1p, 2p endings are attested only before the future marker *kām* (k'm); the OPT 3p *-īr* (-yr) occurs only once.

3.2.4 System of tense, mood, and aspect

The system of tenses and moods comprises both those forms with flexional endings and those marked by the addition of a suffix to an inflected form, as well as periphrastic

perfect forms. The former include the tenses present, imperfect and (rare) injunctive (in form, an ‘unaugmented imperfect’) and the moods indicative, subjunctive, imperative and (rare) optative. Suffixed formations include the future tense, the permansive aspect and the conditional mood.

3.2.4.1 Present and imperfect forms

The present stem is the base for the present indicative, the present subjunctive and imperative, and the imperfect. The subjunctive is distinguished by a separate set of endings, as are the imperfect and the injunctive.

3.2.5.1a Imperfect

The imperfect is marked also by an ‘augment’, the form of which depends on the syllabic structure of the stem. It is generally expressed as follows: either

- (a) by substituting $-ā-$ ($-^2-$) for the first internal short vowel of a polysyllabic present stem (particularly but not only in a prefix), e.g.
 prefix, $hi-βr-$ ($h-βr-$) > $h-ā-βr-eda$ ($h-^2-βr-^y d$) ‘he/she gave’,
 $(b-stw-)$ > $(b-^2-stwd)$ ‘he/she disavowed’;
 polysyllabic, $(bsmyr)$ > $(b-^2-smyr-d)$ ‘he/she slaughtered’,
 which is also an example of both a denominative stems and an integrated loan (< Ar. *bismillāh*); or
- (b) by prefixing $(m-)$ to stems with an initial vowel, whether primary or secondary, e.g.
 $ās$ ($^s-$) > $m-ās-eda$ ($m-^2-s-^y d$) ‘he/she took’,
 $ak-$ ($^k-$) > $m-ak-eda$ ($m-k-d$) ‘he/she did’.

There are, however, many exceptions. Verbs which do not fit either of these categories, e.g. monosyllabic stems, or those with internal long vowels or diphthongs, mostly have identical present and imperfect stems. (For details, see MacKenzie 1975.)

3.2.4.1b Injunctive

The injunctive is essentially the immediate base for the imperfect, and lacks the ‘augment’, i.e. $m-$ ($m-$) or lengthening:

$ne-mi$ [$aβ-eda$] . . . ($ny-m$ ^β-d)
 not-me it becomes-INJ, i.e. ‘is does not suit me’ 314.8;

in subordinate clauses:

šxt $m-aβ-eda$. . . $ka-hi$ [$ak-eda$] $ī$ $nānām wāc$ (šxt $m-β-d$ $k-^2-h$ ^k-yd $^y n$ n $^m w$ c)
 ‘close he was-IPF that he-it do-INJ the certain thing’ 516.4, i.e. ‘he was close to doing . . .’.

3.2.4.2 Compound tenses, perfect forms

Perfect forms, both transitive and intransitive, are based on the perfect participle in ($-tk/$ $-dk$) and the verb $δāray-$ ($δ^ry-$) ‘have’ as auxiliary. The three forms that are attested happen to represent the present perfect indicative, the perfect subjunctive, and the remote perfect, or pluperfect (Samadi 1986: 296), indicated by the personal endings of $δāray-$, and include both transitive and intransitive verbs.

3.2.4.2a Present perfect

Transitive:

a-ci wašīyat . . . [aktek dāray-āmi-na] (ʿc wšyt . . . ʿktk δʿry-ʿmy-n)
 ‘what testament . . . I have made’ QR 411;

3.2.4.2b Perfect subjunctive

Transitive:

ka-hi nikāh [aktek dāray-ān-i] (kʰ nkʰ ʿktk δʿry-ʿy) (= δʿry-ʿny?)
 ‘if-(to)her marriage I should have made’ QR 205;

3.2.4.2c Remote perfect

Intransitive:

ud-hi-beš(a) f-ī išārat [ēdek dāray-eda-mi] (ʿwd hy.bš f.y šʿrt ʿydk δʿry-d. ʿmy)
 ‘and-him-to at a sign she had gone, *say’ QR 257, with problematic (-ʿmy).

3.2.4.3 Basic inflectional system

The basic inflectional system may be illustrated by forms of the verbs *ak-* (‘k-’) ‘do’, *aβ-* (‘β-’) ‘be(come)’, and the existential verb *y-* (‘y-’) ‘be’ which is the most irregular, in particular its present. (For the vocalizations and stress patterns, cf. MacKenzie 1987: 570–575; note also that much more frequent for ‘become’, and quite regular, is the verb *parwuz-* (prwz-). Even for these best documented verbs some forms are missing in the texts. Thus, the example for the remote perfect is from the verb ‘go’, 3s *ēdek dāray-eda* ‘ydk δʿryd-’) ‘had gone’ and *hōb-īr* (hwbyr) ‘may they fall’ for the optative.

TABLE 6.12: BASIC INFLECTIONAL SYSTEM, ‘do, make’

(1) ‘do, make’						
	1s	2s	3s	1p	2p	3p
PRS	<i>ak-āmi</i>	<i>ik-i</i>	<i>ak-ica</i>	<i>ak-āmmi</i>	<i>ak-(a)βi</i>	<i>*ak-āri</i>
SBJ	<i>ak-ā(n)</i>	<i>ak-īx</i>	<i>*ak-āc</i>	<i>*ak-āmini</i>	<i>*ak-βīx</i>	<i>*ak-āri</i>
IPF	<i>mi-k-ina</i>	<i>m-ik-ixa</i>	<i>m-ak-eda</i>	<i>m-ak-āmma</i>	<i>*m-ak-βa</i>	<i>m-ak-āra</i>
INJ		<i>ak-a</i>	<i>ak-eda</i>			
IMP		<i>ik-a</i>			<i>ak-βa</i>	
OPT						<i>hōb-īr</i>
PRF	<i>aktek dāray-āmi-</i>					
PFSBJ	<i>aktek dāray-ā(n)i</i>					
REM			<i>ēdek dāray-eda</i>			
Transliteration						
	1s	2s	3s	1p	2p	3p
PRS	(*kʰm)	(k, -k̄)	(*kʰc, -kʰc)	(*kʰmny)	(*kʰfy)	(*kʰr)
SBJ	(ʰk̄, -kʰn)	(kyx)	(*kʰc)	(*kʰmyn)	(*kʰβyx)	(*kʰr)
IPF	(mkʰn, mk̄nʰ)	(mkʰx)	(mk̄ʰd, mkt)	(mkʰm̄nʰ)	(*mkʰf)	(mk̄r)
INJ		(k)	(kʰd)			
IMP		(k̄ʰ, -k)			(ʰ)kf)	
OPT						(hwbyr)
PRF	(ʰtk δʰryʰmy-n)					
PFSBJ	(ʰtk δʰryʰyy) (= (δʰryʰny) ?, QR 205)					
REM			(ʰydk δʰrydʰ)			

TABLE 6.13: BASIC INFLECTIONAL SYSTEM, 'become, be'

(2) 'become, be'						
	1s	2s	3s	1p	2p	3p
PRS	<i>*aβ-āmi</i>	<i>aβ-i</i>	<i>aβ-ica</i>	<i>*aβ-ām(i)ni</i>	<i>*aβ-βi</i>	<i>aβ-āri</i>
SBJ		<i>β-īx,</i> <i>ha-β-īx</i>	<i>aβ-āc,</i> <i>ha-β-āc</i>	<i>*aβ-āmini</i>	<i>*aβ-βīx</i>	<i>β-āri</i>
IPF			<i>m-aβ-eda</i>			<i>m-aβ-āra</i>
INJ			<i>aβ-eda</i>			
IMP		<i>aβ-a</i>				
Transliteration						
	1s	2s	3s	1p	2p	3p
PRS	(*β'm)	(β, ny.β)	(*β'c)	(*β'mny)	(*βfy)	(βr)
SBJ		(βyx, h'.βyx)	(β'c, h'.β'c)	(*β'myn)	(*βyx)	(βr)
IPF			(mb'd)			(mβr)
INJ			(β'd)			
IMP		(*β)				

TABLE 6.14: BASIC INFLECTIONAL SYSTEM, EXISTENTIAL VERB

(3) 'be'						
	1s	2s	3s	1p	2p	3p
PRS	<i>yī-ma</i>	<i>yāh-i</i>	<i>ye-ti?, yā-ti?</i>	<i>yī-m(i)ni?</i>	<i>*ya-βi?</i>	<i>yī-li</i>
SBJ	<i>yāh-ā(n)</i>	<i>*yāh-īx</i>	<i>yāh-āc</i>	<i>yāh-āmini</i>	<i>*yā-βīx</i>	
IPF	<i>m-ey-na</i>	<i>m-ey-a</i>	<i>m-ey-ta</i>	<i>m-ey-m(i)na</i>		
IMP		<i>yāh-a</i>				
Transliteration						
	1s	2s	3s	1p	2p	3p
PRS	(yīm)	(y'h')	(y'tī, y't', y'ty)	(yīm)	(yf)	(y'l)
SBJ	(y'h')	(*y'hyx)	(y'h'c)	(y'h'myn)	(y'fyx)	
IPF	(mȳn)	(my)	(myī)	(mymn-)		
IMP		(y'h)				

3.2.4.4 Phasal and modal enclitic particles

3.2.4.4a *Permansive particle -ina, -ī*

The permansive is a modification of an inflected verb to signify duration and iterative action, added after the personal endings. It combines with both non-periphrastic and periphrastic forms, including the present indicative, present subjunctive, and the future. Its general form is *-ī(n)-*, except for the 3s indicative:

PRS	<i>-ina</i> (-y(ī)) after 1s, 2s, 3p; but <i>-e(n), -enw(a)</i> (ʔ-, ʔn, ʔnw) after 3s;
SBJ	<i>-ī(n)</i> (-y(n));
FT	<i>-ī-kām-ī</i> (-y)-k'm-y) (double marking).

3.2.4.4b *Future particle -kām*

The future is formed with the particle *kām* (-k'm) (as in Sogdian) added after the personal endings. Both indicative and subjunctive forms occur.

(1) Indicative:

1s *kacay-āmi-kām* (kcy'm.k'm) 'I shall throw';
 2s *aβ-i-kām* ('βk'm) 'you-2s will become'.

(2) Subjunctive:

2s *ne-k-īx[-kām]* (ny.kyx.k'm) '(if) you-2p will not do'.

3.2.4.4c Conditional particle *-manc(a)*

A conditional is formed by adding *-manc(a)* (-mnc-) (with further obscure suffixes: *-āc*, *-āh(i)*, *-(h)i*) (-'c, -'h, -y) to the imperfect (cf. MacKenzie 1971, Supplement: 41b-42a):

kāš-ka-na yāwār-ina[-manca]
 (k's-k'-n' y'w'r-yn'-mnc)

would-that-them I knew-IPF-COND, i.e. 'would that I had known' 174.7; Persian transl. *kāški dānist-am-ē*;

ka yā nānām-ya dāra m-aryand-ina[-manca-āc]
 (k' y' n'n'm-y^a δ'r^a m-'r'nd-n'.mnc-'c
 if the such-and-such-ACCf I summoned-IPF-COND

ud ma-biš m-ē-da [manca-āc]
 ('wd m'.bš m'y-d'.mnc-'c)
 and me-with she went-IPF-COND, i.e.

'if I should have summoned such-and-such a woman, and she should have gone with me . . .', QR 256, with 1s impf. *m-aryand-ina* of **aryand-* 'summon', and *m-ē-da* (m'y-d) < *m-āy-eda*, 3s of *āy-* ('y-) 'go'.

3.2.4.5 Overall system of tenses, moods, and aspect

In tentative initial synopsis, the system of tenses, including forms marked by suffixes, may be shown as follows ('+' indicates documented present and future forms which add the permansive particle):

TABLE 6.15: SYSTEM OF TENSES AND MOODS

PRS	perm.	FUT perm.	PAST	conditional	PRF	perm.	PAST
IND	+	<i>-kām</i> +	IPF	<i>-manc-</i>	IND	+	RemotePF
SBJ	+	<i>-kām</i>	INJ		SBJ		
OPT							
IMP							

3.2.5 Negation

3.2.5.1 Basic negation and prohibition

Normal negation is expressed by the particle *ne-* (ny):

ne až-ica fa nān wāh-a inkār-hi ne-k-āmi ī wašīyat-hi
 (ny 'žc f-n'n w'h) ('nk'r hy ny.k-'m 'y wšyt.h')
 'not it is worth that price' 432.6; 'contest-his not I shall do the testament' QR 405.

A prohibition is marked by the particle *ma-* (m³):

ma-γīr-a (m³-γγr) 'do not return!' YD 8;
ma-kirb-a (m^{a,3}°k¹r°b.^a) 'don't talk idly!' QR 265.

The negation of the verb 'be' in the 3rd sg. is expressed by **inci* (°nc) 'is not': *pandāk-wa* (i)nci (pnd³k w^{a3}°nc) 'there is no way thereto' QR 56.

3.2.5.2 Particle *fa* + *ne* and time reference

The particle *fa* (f³), among other functions (see section 4.3.2.3b Potential construction), appears with the negative present indicative and injunctive, expressing past and future in terms of time reference and semantic context (Sims-Williams 1996: 173). In overview, the following combinations occur:

	<i>ne</i>	<i>fa + ne</i>	<i>fa + ne</i>
Present:	pres. ref.	past ref.	fut. ref., in oaths, potential construction ('can');
Injunctive:	pres. ref.	past ref.	

(1) Present reference, without *fa*:

(a) present:

inkār-hi ne-k-āmi ī wašīyat (°nk³r hy ny.k³-m³ ī wšyt.h¹) . . .
 'I shall not contest his (-*hi*) testament' QR 405;

(b) injunctive:

ne-mi aβ-eda, ī pacxās nē(n) (ny-m³°β-d³°y pcx³s ny)
 'it does not suit me, this garment' 314.8 (Henning 1971: 8b);
 cf. impf. *m-aβ-eda* (m-β-d).

(2) Past reference, *fa* + *ne*:

(a) present:

ne-fa-ci pacxr-āmi xumer (ny.f³.c pcxr³m xwm³r)
 'I did not sleep a wink', 398, 1 lit. 'taste any sleep'.

(b) potential:

yā γrīw fa-ne-δard-k-ica (y³ γryw f³-n-δrd k-c)
 'he could not restrain himself (γrīw)' 486.7 (see section 4.3.2.3, Modal constructions).

(3) Future reference, *fa* + *ne*:

(a) present:

āsērδ fa kašteḱ ne-k-āmi (°srδ f³ kštk nyk³m)
 this year sowing I shall not do, i.e. 'I shall not sow this year' QR 229.

4 SYNTAX AND USE OF FORMS

4.1 Noun phrase structure and word order

4.1.1 Noun and adjective

Adjectives agree with their noun in gender, number, and case. When indefinite, they generally follow the noun they qualify, but may precede it. The latter position is mostly found when the definite article is present. When, less often, an attribute follows a defined noun the article is repeated, together with any governing preposition:

4.1.1.1 Indefinite

Optional sequence:

sing. ('sp 'zr) 'an old ('zr) horse' 166.4, or ('zr 'sp) 37.4;
 plur. (r'č wđney) 'old (wdnk) veins' QR 338.

4.1.1.2 Definite

Fixed sequence: article + adjective + noun,

(y' 'zr γwk) 'the old cow' 94.8; (y' ftmyck srδ) 'the first year' QR 233.

4.1.1.3 Definite, postponed adjective

(Prep.) article + noun, (prep.) article + adjective:

<i>[ī] cub-i [ī] reng-drayāc-i</i>	<i>f[-ī] zβāk-a f[-ī] turkāng-a</i>
('y cw b 'y rng-dry'cy)	(f-y zβ'k f-y trk'nk)
'discoloured water (<i>cub</i> 'water', <i>pl. tantum</i>)' 196.6;	'in the Turkish tongue' QR 123.

4.1.2 Nouns and dependent nouns

With following possessives the article logically keeps the gender of the head noun. However, when the possessive is a feminine, it attracts its own article.

m. + m.: *[ī] ḥaqīqat [ī] wāc-ān* ('y ḥqyqt' 'y w'c'n) 'the truth of the matter' 122.6;
 m. + m.: *[ī] xusr [ī] marcy-ān* ('y xsr 'y mrcy'n) 'father-in-law of the man' 2.1; but
 m. + m.: *[c-ī] bādēs-a [dī] allāh-ān* (c.y b'δys c.y 'llh'n) 'from Allah's command' 103.5;
 m. + f.: *[ī] xusr [yā] δē(n)-ya* ('y xsr y' δy) 'the woman's father-in-law' 2.1;
 m. + f.: *[ī] bār [yā] uδr-ya* ('y b'r y' wδry) 'the foetus (lit. load of the belly (fem.))'
 106.7; feminine attraction:
 f. + m.: *[yā] pac-a [yā] wāc-ān* (y' pc y' w'c'n) 'the end of the matter' 61.3.

4.1.3 Nouns and demonstrative pronouns

Demonstrative are inflected, and as adjectives agree with their noun. In that function, always with a definite article, only *n'(n)* 'that' precedes the noun. All the others follow:

nān- N:f-ā [*nān-a*] *āyd-a* (f. n'n 'γd) 'at that time' QR 240,
yā [*nān-ya*] *hunān-ya dāra* (y' n'ny hwn'ny δ'r) 'for that slave-girl' QR 115;
 N *nēn-i sukund* [*nē(n)*] (y sknd ny) 'this oath' QR 363,
yā dēn-a [*nēn-a*] (y' δy nyn) 'this woman' 65.5;
 Np *nāw-i:īkt-ina* [*nāw-i*] (y ktn n'w^h) 'these deeds' QR 281;
 N *nāwer-i:ī marc* [*nāwer*] (y mrc n'wr) 'that man' 65.6;
 N *nēd-yā dē(n)-a* [*nēd-a*] (y' δy nyd) 'that woman' 65.7.

4.1.4 Nouns and pronominal suffixes

The attachment of the pronominal suffixes follows the rule of left-most attachment, after the respective case ending, but before another enclitic.

(1) Article + noun + pronoun:

yā dūyd-a[-mi] (y' δγ^od-^{at}-mⁱ) 'my daughter' QR 12;
ī bāδēs[-hi] *dāra* (y b'δys.-h δ'r) 'to his command' 498.8.

(2) Article + adjective + pronoun + noun:

yā xūb-a[-mi] *dūyd-a, ī . . . xūb[-mi]* *zādek*
 (y' xwb-^{at}-mⁱ δ^uγ^od^a, 'y . . . x^uwb-m z^dyk)
 'my beautiful daughter, my beautiful child' QR 120, QR 122.

(3) N₁ + pronoun + N₂:

c-ā *medān-ya[-hi]* *kām* (c. mδⁿ-y-.h k'm)
 'from-the inside of-his mouth' 103.7;
f-ēw-a *c-aδw-a[-hi]* *pāδ-a* (f.yw c. 'δw^h p'δ)
 'with one of his two feet' 192.3, 429.3 (Henning 1971: 15b).

4.2 Clause structure and word order

4.2.1 Order of constituents

4.2.1.1 Basic order

The basic word order is subject (- object) – predicate. Such is the case, e.g. in nominal clauses, where the copula is as a rule omitted:

nē(n) *bazag-dar ci* *nān-a* (ny bzk-dr c.n'n)
 'this one (is) worse than that one' 280.6.

4.2.1.2 Questions

All questions are introduced by the interrogative particle *a-*, which may affect word order:

parāc-a yāh-i (pr'c y'h) 'you are divorced' YD 15;
a-yāh-i parāc-a ('-y'h' pr'c) 'Q-are you divorced?' QR 109.
a-βa-ci yi-ti(?) yā wuδ-a šē-yāwar parāc-a ('.f'.c yī y' w^uδ šy'wr pr'c?)
 Q-is (your) wife (*wuδ-a*) thrice divorced from you?' QR 197; *βa-ci* 'thee from'.

4.2.1.3 Initial negation

The negation is often clause-initial, especially in combination with the particle *fa-* (f̄); accordingly, enclitics are attached to it:

fačāw-i(na?), *fa-ne-βēr-ica* (fč̄w-y f̄-n.-βyr-yc)
‘quarries, he found not’ 358.4;

ne-mi aβ-eda, *ī picxās nē(n)* (ny.-m β-d’y p̄cx’s ny)
‘it does not suit (become) me, this garment’ 314.8, injunctive;

ne-ci inci fa nān-a, *am(u)x-āk-a* (ny.-c’nc f.-n’n’mx-’k)
‘there is not any (-ci) in that/this one, movement’ 402.2.

4.2.2 Affixes and ‘principle of anticipation’

In fact, there is considerable free variation in the order of constituents. In particular, the verb is often in topical initial position, or may precede the subject and other parts of speech. In that case there is a complex set of rules that govern multiple reference by cataphoric and anaphoric chains, which may themselves be topicalized by attachment to clause initial particles and phrases.

As in Middle Persian, Parthian and Sogdian, pronominal suffixes are attached to any first syntactic unit of a clause, including particles and conjunctions. It is one of the distinctive features of Khwarezmian that such marking is obligatory, even when followed by explicit noun phrases. This is the ‘principle of anticipation’, defined by Henning (1955: 48). Attachment includes the adverbial suffixes, such as *-wā-* (-w(̄)-), *-waθa* (-wθ) ‘therein’ and *-da-* (-d(̄)-) ‘beyond, off’.

In the following ‘=’ indicates attachment boundary.

4.2.2.1 Direct object + indirect object

hāβr-ina[=hi]-di, *yā duyd-a-mi* (h’βr-n.y.-dⁱ, y’ δγ°d-^a.-mⁱ)
‘I have given her to you (*hi-di*), my daughter’ QR 12;

4.2.2.2 Adverbial suffixes

(1) *ci*, *ci NP* (c-):

wāzay-ta[=ci], *[c-]āγudc-a*, *ī xerāx* (w’zy-t’-c, c-’γdc, ’y xr’x)
‘it came out of it, from the scabbard, the sword’ 457.4; (on (’γdc), see MacKenzie, Glossary IV, 1972: 524).

(2) *-wa*, *fa-NP* (-w, f-):

m-ās-ta=na[-wa], *[f-]ī parδāβ-a* (m’st’n’w, fy prδ’β)
‘he took them-therein, into deception’ 417, 6–7 (MacKenzie, Glossary III, 1971: 325).

4.2.2.3 Pre- and postverbal position

For pre- and post-verbal positions of the enclitic chains, compare:

rasūl=hi-0a m-afṛēs-eda bāwi(n) cek-a (rswl hy-θ m-frys-d b'wyck) 348.5,
*bāwi(n) cek-a=hi-0a rasūl *m-afṛēs-eda* (b'wyck h-yθ rswl *m-frys-d) 204.5;
 both 'he sent a messenger to him (*hi-0a* 'him-to') secretly'; and

m-afṛēs-ina=hi-0a, ī xsin (m-frys-n'.hy.-θ 'y xs'n)
 'I sent it to him, the thing' 349.3;

**m-afṛēs-ed=ē-0a* (m-frys-d.y.-θ)
 'he sent him to him' 300.7.

In the last example, =ē- is contracted from *-a-hi-hi* 'him-him-to'. Similarly, *hin-hina* contracts to a single *hina*, formally indistinguishable from *hi-na*, as shown by MacKenzie (1993: 141). In fact, up to four enclitics may form the chain, but never more than two pronouns and two adverbials.

hēδ-eda=hi-[na-da]-ber (hyδ-d'.-hy.-n'.-d'.-br)
 'he recited them (greetings) to him' 298.8 (lit. 'read-him-[them-forth]-upon').

This rare example interweaves the basic verb *hēδ-* 'read' with (a) the directional pre-verb *da* 'read forth, off'; (b) its accusative/direct object *na* 'them'; and (c) its indirect object, or rather adverbial phrase, *hi-ber* 'him to, upon'.

4.2.2.4 Clause-initial particles

The same rules apply when the chain is attached to clause-initial particles and conjunctions:

ka=βa-[na]-ci-[da m-ās-ina(!)] (k.-f'.n'.-cy.-d^a m-'s^o-n^a)
 'that I took them from you' QR 366 (lit. 'that-thee-[them]-from-[away took]').

Note *-da-* before *-ber*, but after *-ci*, in the preceding and this examples, respectively.

4.2.2.5 Agency hierarchy

When two personal suffixes are present, there is an inflexible order in their sequence, as established by MacKenzie (1993: 141), irrespective of their 'logical' order:

TABLE 6.16: ORDER OF PERSONAL SUFFIXES

	ACC,DAT 1s/p	2s/p	3s/p	DAT 2s (last)
Singular	<i>-ma-, -mi-</i>	<i>-βa-</i>	<i>-hi-, -ci-</i>	<i>-di-</i>
Plural	<i>-mena-</i>	<i>-βena-</i>	<i>-na-, -hina-</i>	
Transliteration				
Singular	(m', my)	(f')	(hy, cy)	(dy)
Plural	(mn')	(fn')	(n'/hyn')	

That is, the logic of the sequence implies a pragmatic hierarchy: (1) basically, 1 before 2 before 3; (2) this, within an overall frame where the speaker is first, and the addressee in the dative is last; (3) in addition, singular appears to come before plural.

4.3 Semantics and use of forms

4.3.1 Use of number and case

4.3.1.1 Impersonal plural

Impersonal or neutral subjects are expressed by the 3p verbal endings and pronouns (Henning 1955: 46, 1971: 8b; MacKenzie, Glossary I, 1970, 548, Glossary II, 1971: 82).

aβ[-āri] nān[-i] (ʰāβʰr nʰn)
‘be they thus’, i.e. ‘be it thus’ QR 11;

a-ci-ki[-na] *da ak-ica* (ʰ.čʰ.kynʰ dʰ ʰkč)
‘Q-of-whom-them (= it) is he talking’,

i.e. ‘who is he talking about?’ Q 293 (Henning 1971: 22b) = QR 357; *da ak-* (dʰ ʰk-) contextually ‘talk, speak’.

4.3.1.2 Dative of affectee

This feature was discussed in more detail by MacKenzie (1993: 141), who cites

dahedēna f-ā tēγ-ya (dhdyn fʰ tyγγ)
< *dah-eda-hi-na* *f-ā tēγ-ya*
‘hit-he (to-)him-them with the stick’ 234.5.

The sentence exemplifies another noteworthy semantic feature of Khwarezmian by which the recipient of the action is put in the dative as the affectee. Here, ‘hit’ can be paraphrased as ‘give blows to’, thus *-hi-na* ‘them to him’, as reflected in MacKenzie’s translation of this sentence.

4.3.2 Use of verb forms

4.3.2.1 Possession, ‘have’

(1) Existential ‘be’:

Possession is expressed by verb *aβ-* ‘be’ and the genitive-dative, usually a personal suffix:

usi(?) hi-wa m-aβ-eda f-ī wāc-a (ʰwš hy.w mβd f.y wʰc)
sense him-upon was in the matter’, i.e.
‘he had sense in the matter’ 198.7;

ka-ās maβyān-δāra xšty aβ-āc . . . (k.ʰs mβyʰn δʰr xšty ʰβʰc)
if to us a will should be, i.e.
‘if we should have a will’ QR 271.

(2) *δāray-*:

The verb *δāray-* ‘hold, keep’ is much less common in this sense:

asnādek-astar jūft δāray-ina-mi ci-βika (ʰsnʰdk ʰstr jʰft δʰry-nʰ-my c.fyk)
more pious husband I had than you
‘I had more pious husband than you’ QR 164.

4.3.2.2 Phasal constructions

4.3.2.2a *Duration, δāray-*

Duration may be expressed through the verb *δāray-* ‘hold, have’ used together with a nominal form:

āyrāci-hi δār-eda (ʿyrʿcy hy δʿryd)
 ‘he kept him awake’ 350.5.
 with *āyrācin* ‘awake’ < perfect participle *āyrādek* of the verb *āyrās-lāyrās-y-* (ʿyrʿs-l
 ʿyrʿs-y-) ‘awake (intr./trans.)’ (Henning 1971: 20b–21a; MacKenzie 1971,
 Supplement: 36b, (δʿry-)).

4.3.2.2b *Incipient action, āyāz-*

The beginning of an action is expressed by the verb *fa-. . . āyāz-* (ʿyʿz) and the ‘past infinitive’:

[m-āyāz-eda]-hi f-akt-a (m-ʿyʿz-dʿh f-ʿkt)
 ‘he began-it to do’ 343,1 (cf. Henning 1971: 19b).

4.3.2.2c *Ingressive action, ‘about to, close to’, *(-m-)ah-; (šxt β-)*

- (1) The anticipation of an action is expressed by *fa. . . *(-m-)ah-* (3s imperfect) and the ‘present infinitive’ in *-āk*:

fa δambay-āk *[m-ah-eda]* (fʿ δnby-ʿk mh-yd)
 ‘on (the point of) hitting he was’ 516.3; Persian transl. *nazdīk šud ki, xʿāst ki . . .*
 ‘got close to, was about to’ (lit. ‘wanted’) (Henning 1971: 21b).

- (2) Another means is a finite construction with the injunctive dependent on the expression (ʿšxt β-) ‘be(come) close (ʿšxt) to’:

[(ʿšxt) m-aβ-eda] ka-hi ak-eda ī nānām wāca (ʿšxt mβd kʿh ʿkyd ʿy nʿnʿm wʿc)
 ‘close he was that he-it do-INJ the certain thing’ 516.4, i.e. ‘close to doing’
 (Ar. *qarub-a ya-fʿal-u*; MacKenzie 1971, Supplement: 31b).

4.3.2.3 Modal constructions

4.3.2.3a *Optative particles ās and hā(n)*

The particles *ās* (ʿs) and *hā(n)* (hʿ(n)) have a clear modal function generally termed ‘optative’ and are found in sentences containing a subjunctive verb (MacKenzie 1971, Supplement: 30 and 40):

d[-ās]-hi 0-barakak a-β-āc (dʰ.ʿ.s.ʰi θ.brkk ʿβʿc)
 ‘may it be blessed to her (-hū)!’ QR 87;

parāc-a [hā(n)]-β-īx (pʰrʰ.ʰ.β-yx)
 ‘may you-2p be divorced!’ QR 41.

4.3.2.3b *Potential construction, -k-*

The potential is formed by the verb *-k-* ‘do’ (reduced < *ak-*, see MacKenzie 1971, Supplement: 40) as an auxiliary in combination with the perfect participle of the main verb (cf. Sogdian, Gershevitch 1954: 130ff.). Most of the few examples attested are negative and involve the particle *fā* (see section 3.2.5.2):

ka-fa-ma ne-pard[-k-i] (k-f’ m’ ny prd-k̄i)
‘for you-2s cannot restrain me (-ma)’ QR 161 (pres. *pāray-* ‘restrain’);

yā yrīw-a fa-ne-δard[-k-ica] (y’ γryw f’ n-δrd-k-c)
‘he could-PRS not hold himself’ 486.7 (pres. *δāray-* ‘hold’).

4.3.2.3c *Necessity, γuw-*

Necessity is expressed by the verb *γuw-* (*γw-*) ‘be necessary, be wanting’, combined with the dative of the beneficiary:

nān [γuw-ica tewār-a dāra] ka-hi ak-a *ī nān nānām wāc*
(n’n [γw-c] tw’r-δ’r) k’-h ʔk ’y n’n n’n’nm w’c)
‘that [behooves you] that-it you-INJ-2s do such-and-such a thing’ 516.3.

5 COMPLEX SENTENCES

5.1 *Conjunctions, particles, and clitics*

The main coordinating and correlative conjunctions are *ud* (‘wd) ‘and’; *wā* (w’) ‘or’; *wā . . . wā bā* (w’ . . . w’b’) ‘either . . . or’; the enclitic *mes* (ms) ‘too, again’; *āmādi* (‘m’d) ‘let alone’. The adverb *da* (d-) ‘then’ may connect clauses loosely. (For its use in complex sentences, and other conjunctions and phrases, see section 5.1)

5.2 *Subordination*

Subordination is marked by the relative pronouns *ki* (k) and *ci* (c), *cā* (c’), and the general conjunction *ka* (k), as well as conjunctive phrases.

Clauses expressing an objective, result or fact, including direct speech, and purpose or condition of a wish, usually follow the main clause.

Clauses expressing situations, including temporal and conditional clauses, as well as relative clauses with a cataphoric demonstrative head, usually precede the main clause, which in turn is as a rule introduced by the resumptive particle *da* (d) ‘then’.

Regarding the use of tense and mood, note:

- (1) the contextual use of the imperfect also for completed action (e.g. *pāckun-eda=mi* (p’cknd’m) ‘he (has) settled (it) to me’ QR 376),
- (2) the use of the subjunctive for explicit or implicit conditions, including relative clauses, and
- (3) the use of the injunctive for purpose, and of the future for wishes, together with ‘optative’ particles.

The conjunction *ka* is the most widely used conjunction. It marks explicative remarks and object sentences including reported speech.

5.3 Relative clauses, *ki*

In relative clauses the head is represented by the verbal ending if it is the subject, or by an anaphoric pronoun.

(1) subject:

ī šuył nē(n), ki=mi fa-hi xudāk anbec m-iy-ta, da yuydek yiti(?)
 (ʿy šyl ny, k.my f.y xdʿk ʿnbc m-y-ī dʿ yʷγdk yī)
 ‘this business that-(to)me with-his self together **was, then** fixed is’ QR 368,
 i.e. ‘that I had with him’.

(2) direct object:

ī kt-ina nāw-i ki k-i=na utak
 (ʿy ktn nʷi k. kʷyn ʿtk)
 ‘the actions **these, that** you are doing-**them**, you’ Q 217 (Henning 1971: 23a) = QR 281.

5.4 Subject and object clauses

5.4.1 Subject clauses, *cā*

Cā + subjunctive, *da* + present indicative:

cā was-āc, da-nē(n) raštāwand
 (cʷsʷc, d-ny rštʷwnd)
 ‘**what** he may say, **then-this** (is) true’ QR 400 = Q *335 (Henning 1971: 34b).

5.4.2 Object clauses, *ka*

5.4.2.1 Basic pattern

ka + present indicative:

a-cūni nā yiwār-i-y-ī, ka=hi-na-ci ya dāray-āmi?
 (ʿacwni nʷ ywʷr-y-y, ky.nʷc yʷ dʷryʷmi?)
 how that you know **that-him-them-from** I have
 ‘how do you know that – that I have them from him?’ QR 360.

5.4.2.2 Direct speech, *ka (bā)*

ka-bā, a-ciwā ma čkēš-i? (k. bʷ ʷcwʷmʷ čkyš-y?)
 ‘... **that rather**, ‘why do you give me the lie (in this matter)?’ QR 149.

5.4.2.3 Topicalized object clause, with interrogative, *cā, aci*

(1) *cā* + present indicative, *da* + imperative:

cā-k-i, nā(n)-k-a and *cā=na-k-i, da=na da-ka*
 (cʷ.kʷi nʷ.kʷ) and (cʷnʷk, dnʷdʷk)
what-you do-PRS, that **-do-IMP** **what-them-you do, then-them do**
 ‘whatever you do, that do’ QR 93, i.e. ‘do what you (want to) do’ (Henning 1971:
 22b, 24a), 3p *-na* for general reference.

(2) *a-ci* (‘c) for *cā*, *a-ci* + present perfect, *da* + imperfect:

a-ci *wašīyat ī pur-dāra aktek dāray-āmi-na, da=na tewār-dāra βānc-ina*
 (‘c wšyt ’y pr-δ’r ’ktk δ’r-y’m’y-n, d-n’ tw’rδ’r β’nc-yn)
 ‘what testament for the (my) son **I have made,** **then-that** to you **I attached**’
 Q 345 (Henning 1971: 13a) = QR 411; 3p *-na* for general reference.

5.5 Adverbial clauses

5.5.1 Temporal clauses, *ka*

Temporal clauses are introduced by temporal phrases and the conjunction *ka*, and may be followed by the resumptive *da*.

5.5.1.1 ‘at time that, when’

ufān-ya ka (‘wf’ny k) + imperfect:

ufān-ya ka=mi ḥalāl-a m-ey-ta (‘wf’ny k.my ḥl’l m-’y-t̄)
 when-(to)me lawful was
 ‘when she was lawful to me’ QR 69.

5.5.1.2 ‘as soon as’

(1) *mān-bā ka* (m’nb’k) + future:

mān-bā ka=na wēn-āmini-kām (m’n. b’k.n’ wyn-’m’ny.k’m)
 when-them we shall see
 ‘when we shall once see them . . .’ QR 286.

(2) *mān-bā ka* (m’nb’k) + future, *da* + future:

mān-bā ka=mi-hi ḥalāl-a ak-βi-kām, da=wa nyōs-ām-kām
 (m^an° .b^a k.m’y.hi ḥl’l^a ’k-fy. k’m d.w’ nyws-’m. k’m)
 as soon as-(to-)me-her lawful **you-2p will make, then** thereto **I shall listen**
 ‘as soon as (that) to-me-her you will make lawful, then-thereto I shall listen’
 QR 68.

5.5.1.3 ‘while’

Imperfect, *da* + present permansive:

y-ā xab-a m-asān-eda, d-ī nānām wāc ak-ica-na
 (y’ .xb ms’nyd, d.y n’n’m w’c ’kc’n)
 the night **he passed** **while** that certain thing **he is doing**
 ‘he passed the nights while he was doing that certain thing’ 135.5;

possibly imitating Ar. *bāta yaf’alu kaδā*, (Henning 1971: 26a), referring to the use of the present indicative for ongoing action in a past context.

5.5.1.4 'when(ever)'

- (1)
- y-āyd-a ka*
- (y'γd k) + subjunctive:

y-āyd-a ka=wa ciy-ā(n) (y'γd^a k' w cȳ')
 'when-there I (shall) enter' QR 142 = Q 98 (Henning 1971: 20a).

- (2)
- y-āyd-a ka . . . da*
- (y'γd
- ^a
- k . . .) + future,
- da*
- + future permansive:

y-āyd-a k- ās-i-kām-u, dī δēnār-a hiβr-ā-kāma
 (y'γd^a k'si-k'm^u, dy' δyn'r^a hβr'm-k'm^a)
 'whenever you will come, then the money I shall give'
 QR 143 = Q 99 (Henning 1958: 20a); < *da-dī-ī?*

- (3)
- ci-yac-ya-kāma ka*
- (cȳcy'k'm k) 'whenever' + future permansive:

ci-yac-ya-kāma ka=mi ḥalāl parwuz-āc-ina (cȳcy'k'm k'm ḥl'l prwz'cyn)
 whenever-(to)me lawful may become-3s
 Q *14 (MacKenzie 1971, Supplement: 34b), cf. QR 21 and note there p 21; here with
 permansive *-ina*.

5.5.2 Conditional clauses, *ka . . . da*

In conditional clauses the protasis is introduced by *ka* (k(?)) 'if', and the apodosis by the resumptive *da*.

5.5.2.1 Basic patterns

- (1)
- ka*
- + subjunctive,
- da*
- + present:

k-amīy-ā(n), da=tak bēzār
 (k. 'mīy-', d.tk byz'r)
 'if I (should) die, then-you (are) free' QR 404.

- (2)
- ka*
- + subjunctive,
- da*
- + imperative:

ka para pan-a=mi sukund ne-xur-īx-ī, da=ci par-mika xur-a
 (k' pr.' 'pn'mi sknd ny xr-yx-y, d^a.c'y pr.mk x^ur)
 if on co-wife-my oath you not swear, then-one on me swear
 'if you do not swear an oath (of divorce) against my co-wife, then swear one on me'
 QR 105.

- (3)
- ka*
- + future permansive,
- da*
- + future:

ka=mi-hi hiβr-ī-kām-ī ī zirnī nē(n), da=βa-ci ēw ipš-i xarj ne-k-āmi
 (k. my.h h'βr-y.k'm-ī ('y zrny ny, d.β.c 'yw'pš xrj ny.k-'m))
 'if-me-it you will give the gold this, then-you-from one farthing spending I will
 not do
 'if you will give me this gold, I do/shall not spend a farthing from you,

bā=hi aspanī xin-ām-kām
 (b'hī 'sp^anīy x'n-'m. k'm)
 but-it iron I shall buy
 but I shall buy iron (with) it' QR 157.

- (4)
- ka*
- + subjunctive,
- ud*
- ('wd) 'and' + subjunctive:

ka=hi nikāh aktek δāray-ā-yī, ud=hi-ci ak-ā(n)
 (k'h nk'h 'akt'k δ'ry-'y'-y 'wd-hy-c 'k')
 if-her marriage I had made, and-her-with I may do
 'if I had married her, and if I do it with her' QR 205; *δāray-ā-yī* palatalized < *δāray-ān-ī* with permansive *-ī*; Persian transl. *agar karde bāš-am*.

5.5.2.2 Implicit conditional, imperfect

asnādek-astar jūft δāray-ina-mi ci-βika
 ('sn'dk 'str j'ft δ'ry-n'my c.fyk)
 more pious husband I had than-you
 '(if?) I had a more pious husband than you' QR 164 = Q * 117; Persian transl. *pākdāman u pārsā-tar šōy mē-dāšt-am* (cf. MacKenzie 1971, Supplement: 36a–b; Samadi 1986: 63), *-mi* is uncertain.

5.5.2.3 Wishes, *hān*

In wishes, the specifying condition given follows the wish.

- (1)
- hān*
- + subjunctive,
- ka*
- + future:

parāc-a hā(n) aβ-āc šē-yāwar, ka=hi kābīn-a ak-ām-kām
 (p'r'c h' 'β'c šy'wr, k.h k'byn 'k-m°.k'm)
 divorced OPT may-3s be thrice if-her betrothal I shall make
 'may she be divorced thrice, if I shall betroth her' QR 151.

- (2)
- hān*
- + subjunctive,
- ka*
- + past conditional,
- manc(a)*
- (
- manc*
-):

a-hā(n)-β-īx parāc-a, ka=hi m-ik-īxa-manc-ā-hi?
 ('.h'β-yx pr'c, k'.h'y m-'k-ī'.mnc-'h'?)
 Q-OPT-may-you-2s-be divorced, if-it you should have done
 'will you be divorced if you should have done it?' QR 107.

See also section 4.3.2.3 Modal constructions.

5.5.3 Causal clauses, *ka*

These are introduced by *ka*, and normally follow the main clause:

- (1)
- ka*
- + imperfect:

pāckun-eda=mi, ka f-ā βayd-a m-anyas-ta
 (p'ckn-d'm, k. f' βyd' m-nys-t)
 he has settled-me for as gift it has come
 'he has settled (it) to me, for it has come to (be regarded as) a gift' QR 376; with imperfect for completed action.

(2) *ka-bā* ‘because’:

a=ma-ci yāh-i šē-yāwer-a parāc-a, ka-bā dah-ina=na-di
 (ʔ^a.m^a.c y^hi š^ēy^{wr}^a pr^c, k.b^ʔ dhn^ʔ.n^ʔ.d^ʔ)
 Q-me-from you-2s are thrice divorced, **for** I hit -you-2s (-*di*)
 ‘are you thrice divorced from me, because I hit you?’ QR 110.

5.5.4 Final clauses

Final clauses follow the main clause and are introduced by *ka*, and *kān-ci* (k^ʔnc(y)) (< *ka-na-ci*?), as well as *camend* (cmnd) ‘so that’; clauses or expressions implying possible action or state have the injunctive:

5.5.4.1 Basic patterns, *ka*(1) *ka* + present indicative:

kān=ci arc yād(a)k ī wašī parēx-ica, wā kaδāki? āy-ās
 (k^ʔn.c ʔrcy^ʔdk ʔy wš^ʔy pryx-yc w^ʔ kδ^ʔk^ʔ ʔy. ʔs)
so that-from third the trustee **leaves** or not? yes
 ‘so that the trustee leaves a third of it (unpaid), or not?’ ‘Yes.’ QR 413; see section 8
 Sample Text.

(2) *ka* + injunctive (or imperfect):

nān ʔw-ica tuwār-δāra, ka=hi ak-a nān nānām wāc
 (n^ʔn ʔw-c^ʔ tw^ʔr-δ^ʔr, k^ʔ-h ʔk n^ʔn n^ʔn^ʔm w^ʔc)
 ‘**that** behooves you, **that-it you do** such-and-such a thing’ 516.3.

(3) *ka* + injunctive:

y-āyādek=mi k-aβ-eda ī nānām wāc
 (y^ʔʔ^ʔδk-m k^ʔβ-yd ʔy n^ʔn^ʔm w^ʔc)
 the wish-(to-)me **that it be** the such-and-such matter
 ‘it is my wish that such-and-such a matter be’ 285.6 (Henning 1971: 19b).

5.5.4.2 ‘until, so that’, *camend*(1) *camend* (-a) ‘until, so that’ + imperfect:

ī δast f-aβrūc-a wās-eda, camend wā-šākaš-ta fa xsin-a
 (ʔy δst f-ʔβrwc w^ʔs-d cmnd w^ʔ-š^ʔkšt f-xsyn)
 the hand on eyebrow he laid, **so that he could look** at something
 ‘he laid (his) hand on the eyebrow so that he could look at something’ 497.6
 (Henning 1958: 119 n. 3: ‘er legte die Hand auf die Augenbraue, auf dass er auf etwas hinsähe’).

(2) *camend*(a) + imperfect:

δāmb-eda=hi, camend pāc-eda
 (δ^ʔnb-yd^ʔ-h, cmnd p^ʔc-yd)
 ‘he hit-him, **so that it swelled up**’ 410.3 (Henning 1958: 119 n. 3).

(3) *camend ka* ‘until finally’ + imperfect (or injunctive):

yā yrīw-a fa ne-δard-k-ica, camend ka nānān wāc was-a
 (y' rryw f' n-δrd-k-c cmnd k- n'n'n w'c ws)
 the self he cannot hold, **until that** certain thing he said
 ‘he could not hold himself, until he said that certain thing’ 486.7; negated potential
 construction, *ne-δard k-*, and impf. or inj. 3s *was-a* or < *was-ta*.

5.6 Nominalized clauses

In addition to the use of the infinitives and verbal nouns in the phasal and modal constructions discussed above, infinitival phrases also occur in various functions, either with or without *fa* ‘to, for’.

[kacay-āk] kōzay-eda
 (kcy-'k kwzy-d)
 ‘he asked for dictation’ 507.4 (Samadi 1986: 107).

[nipi.j]=hi-ci kōzay-eda
 (npc hy-c kwzy-d)
 ‘asked him to write’ 494.2 (MacKenzie 1971, Supplement: 44a).

[nipi.j] yr-eda ī yēck-ān
 (npc ryr-d 'y ryc'n)
 ‘sought the writing of the letter’ 494.6 (MacKenzie 1971, Supplement: 44a; Samadi 1986: 84);
 Persian transl. *nivištan xwāst nāma-rā*.

[fa-nipi.j]=hi-ci hābr-eda qašidek-a
 (f-npc hyc h'βr-d qšydyk)
 ‘gave him the *qašida* to write’ 344.7 (MacKenzie 1971, Supplement: 44a).

[f-āβōzay-āk-a]=hi hāβr-eda
 (f-'βwzy'k hy h'βr-yd)
 ‘he gave it for smelling’ 368.3 (Henning 1971: 11b).

6 LEXIS

6.1 Word formation

6.1.1 Derivation of nouns and adjectives

6.1.1.1 Suffixation

6.1.1.1a Verbal nouns and infinitives, all feminine

-āk-a (-'k): (βx-'k) ‘division’ 423.3 (MacKenzie, Glossary III, 1971: 325) < (βx-) ‘divide, distribute’ – (kwzy'k) ‘request’ 316.6 < (kwzy-) ‘seek, ask of’ – (*bsmyr-'k) ‘sacrifice’ 301.8 < (*bsmyr-) ‘to sacrifice’ (< *bismillāh*).

6.1.1.1b Abstract nouns, all masculine

-āw(e)k (-'w'k): (δyn'βr-'w'k) ‘religiosity, piety’ 479.8 < (δyn'βr) ‘religious’ – (jft-'wk) ‘matrimony’ QR 19 < (jft) ‘spouse’.

6.1.1.1c Adjectives

- (1) General suffixes:
-ānk (-ʾnk): (jhwδ-ʾnk) ‘Jewish’ 11.7 – (trk-ʾnk) ‘Turkish’ QR 123 – (βʾrk-ʾnʾk) ‘of a rider’ 15.4; *-ēnk* (-yнк): (βyʾδk-yнк) ‘of dawn’ 23.4 – (skrβk-yнк) ‘wooden’ 15.5 (MacKenzie, Glossary V, 1972: 59);
- (2) Adjectives of relationship (cf. Sogdian masc., fem. (-my(n)c))
-mēnk (-mynk): (xwʾr-mynk) ‘flesh-y’ 334.4 – (γδ-mynk) ‘concerning wounds’ 61.6 < ((ʾ)γδ) ‘wound’;
- (3) From verbal nouns
 (bsmyr-ʾk-mynk) ‘sacrificial’ 301.8 – masc., fem. (wδry-ʾk-mynk-c) ‘of support’ 81.4 30.6 < (wδry-) ‘be supported’.

6.1.1.1d Diminutives

-ik (-ʾk): (jʾbyc-ʾk) ‘small boy’ 85.3, 106.3 < (jʾbʾk) ‘young, child’.

6.1.1.2 Prefixation

- (1) *θa-* (θ-) ‘with’: (θ-γwndk) ‘sinful’ 249.1 < (γwndk) ‘sin’ – (θ-nmθk) ‘salty’ 324.7 < (nmθk) ‘salt’;
- (2) *ba-* (b-) ‘without’: (b-mzk) ‘tasteless’ 251.3 < (mzk) ‘taste’ – (b.zβʾk) ‘dumb’ 352.7 < (zβʾk) ‘tongue’;
- (3) *mand-* (mnd-) ‘un-’: (mnd-cγr) ‘blunt’ 124.3 < (cγr) ‘sharp’ – (mnd-mʾnynd) ‘unlike’.

These derived adjectives may function as substantives as well: *ī ba-wuδ-i* (ʾy b.wδ) ‘the bachelor’ (‘without a wife (wδ)’) 323.1.

6.1.2 Compounding

6.1.2.1 Verbal components

Present participle, *kun-enk-*: *ābādān-kun-enk* (ʾbʾdʾn-knyнк) ‘settling, making cultivated’ 105.8 – past participle, *akt-*: *akt-wēden-* (ʾkt-wyδyn) ‘large four-feathered arrow’ (lit. ‘made-piercing’) 42.7 (Henning 1971: 27b) – *mand-akt-ramān-cīk-āw(e)k* (mnd-ʾkt-rmʾn-cyk-ʾwk) ‘disobedience’ 149.4, composite of *mand-* ‘un’ + *akt-ramān* ‘done-the-order’ = ‘obedient’ + abstract suffixes *-cīk* + *-āw(e)k* – perfect participle, *akt-ek-*: *ānūt-ek-akt-ek* (*nwtk-ʾtk) ‘pupil’ (lit. ‘instructed-made’) 390.4; – *ābādān-akc-a* (ʾbʾdʾn-ʾkc) ‘settled-made, populated (fem. -c-)’ 105.8 – adjectival *-k-ānk*: *pas-aβin-k-ānk* (ps-βn-k-ʾnk) ‘sheep-shearing’ (with *karcci* (krcc, fem. (-c)) ‘shears’) 51.1.

6.1.2.2 Nominal components

aβrāy-a-zād-ek (ʾβrʾγ-zʾdk) ‘young tree’ 91.2 (lit. ‘tree-child’) – *aβar-sarδ-ek* (ʾβr-srδk) ‘grown-up, adult’ B (Henning 1971: 2a) 203.21, (‘above’ + ‘year’ + suffix) – *aδwa-sard-ek* (ʾδw-srδk) ‘two-year-old’ 45.5; *aδwa-pickar-ek* (ʾδw-pckr-yk) ‘Two-Images, Gemini’, al-Bērūnī, *Chronology of Ancient Nations* 193.4, 238.12 – *cikām-βar-ek* (ckʾm-βr-k) ‘laden with fruit’ 369.5 (‘many’ + ‘fruit’ + suffix) – *βūmya xar-cīk* (βwmyʾ xr-cyk) ‘wood-louse’, 55.2 (lit. ‘earth-donkey+diminutive’).

6.2 Sociolinguistic aspects

6.2.1 Standard and colloquial registers

The situations dealt with in the Qunya include statements made on an impulse juxtaposed with concise indications of their legal import or consequences. Much of the material there is quoted colloquial, some even intimate speech. The following unmarked conditional sentence is clearly colloquial:

ēw šāβār-a wāzay-āc, ud nāz ī abarēw šāβār
 (ʿyw šʾβr^a wʾzyʿc ud nʾz y. ʾbryw šʾβr)
 one witness he may produce, and I the other witness
 ‘let him produce one witness and I (am) the other witness’ QR 373.

6.2.2 Kinship terms

Most of these kinship terms are listed in the Muq. 1–2.

TABLE 6.17: KINSHIP TERMINOLOGY

<i>pica</i> (pc), obl. <i>picer-</i>	‘father’	<i>mād-a</i> (mʾd)	‘mother’
<i>pic-ender</i> (pcʾndʿr)	‘step-father’	<i>māderī</i> (*mʾdry)	‘step-mother’
<i>pur</i> , <i>apr-</i> ((ʾ)pr)(?)	‘son’	<i>duγd-a</i> (δγd)	‘daughter’
<i>zādek</i> (zʾdʿk)	‘child, son’	<i>nabāsenek</i> (*nbʾsnʿk)	‘grandchild’
<i>pur-ander</i> (prʾndr)	‘step-son’	<i>duγd-ander</i> (δγdʾndr)	‘step-daughter’
<i>βrād</i> (βrʾd)	‘brother’	<i>ux-a</i> (x)	‘sister’
<i>afcūr</i> ((ʾ)fcwr)	‘paternal uncle’	<i>afcūrī</i> ((ʾ)fcwry)	‘aunt’
<i>afcūr-pur-ak</i> (*ʾfcwrprk)	‘cousin’		
<i>feynd</i> (?) (fynd)	‘husband’	<i>weδ-a</i> (wʾδ)	‘wife’
		<i>ap(a)n-a</i> (ʾpn)	‘second wife’
<i>pic-βrād</i> (pcβrʾd)	‘brother-in-law’	<i>nanek-a</i> (nnʿk)	‘sister-in-law’
<i>zāmād</i> (zʾmʾd)	‘son-in-law’	<i>an(u)h-a</i> (ʾnh)	‘daughter-in-law, sister-in-law’
<i>xus(u)r</i> (xsr)	‘father-in-law’	<i>xuš(-ū)</i> (xš(w))	‘mother-in-law’

7 LANGUAGE DEVELOPMENT, LOAN COMPONENT

Words with initial *b, d, g* (though these are not always written unambiguously) instead of the expected Khwar. *β, δ, γ*, are loanwords, mostly from Early New Persian (ENP): (bwstʾn) ‘garden’ = ENP *bōstān*, (drwδ) ‘praise’ = ENP *drōδ*, (grk) ‘herd’ = NP *gallah*. The latter exhibits a regular feature of all such loans with final *-ah*, namely the replacement of this by written (-k). This may date from Middle Persian (MP), e.g. *bazak* ‘bad’: MP *bazag* ‘sinful’, but ENP *bazah* ‘sin’; (zywndʿk) ‘alive’: MP *zywndg*, but NP *zindah*. Nevertheless it is strongly represented in Arabic loans in *-ah*, e.g. (ʿy ftnʿk) ‘temptation’ 115.4 < *fiṭnaʾ*, (ʿy hdyk) ‘gift’ 356.6 < *hadīyaʾ*, (ʿy bʾdyk) ‘desert’ 480.4 < *bādīyaʾ*. Other Arabic loans are mainly unassimilated, at least in writing, e.g. (ʿyb) ‘defect, disgrace’, (šḥbt) ‘conversation’ 433.5, (tʾm) ‘food’ 240.4. Those with final *alif* often take a final (-y), e.g. (dwʿy) ‘prayer’ 235.3 < *duʿā*, (wʿy) ‘faith’ 386.7 < *wafā*, but (ʿyl) ‘oath’ 384.1 < *īlā*. Many such loans occur assimilated in combinations like (θ-šl) **θ-ašl* ‘of noble origin’ 331.3, (ʿšlyk)

‘original’ (lit. ‘with-origin’) 402.5 < ʔl ‘origin, root’, (b-’db’wk) ‘rudeness’ QR 262, cf. NP *bē-adab-ī* ‘being without manners’ < Ar. *adab* ‘good manners’.

Few Turkish loanwords have been identified: (bkny) ‘beer’ QR 177 < Turk. *bāgni*, *bāgini*; (ʔdm) ‘a felt pad beneath the saddle’ 40.3 < *ādrın*; (xč) ‘ram’ 49.5 < Turk. *qoç*.

8 SAMPLE TEXT

This sample text is the rendering of an Arabic passage (QR, MacKenzie 1990: 82–83), with interlinear version (the Arabic in transcription is placed at the end):

1. QR 413

*f-ī nān šūrat ka yā *βasanjīwān-a ī jūft δāra wašī wās-eda*
 (f.y n'n šwrt k. y' βsnw'n 'y jft δ'r wšy w'-yd)
 in the that case that the *testatrix the husband-ACC trustee appointed-IPF-3s
 ‘In the case that the *testatrix appointed (her) husband as trustee’

*f-ī nafaqān *δrūkā-nīk f- arcyād(a)k-a c-ī tarikāt-a-hi*
 (f.y nf'γ'n δrwk'n'yk f. 'rcy'dk c.y trkt-'h)
 in the wasting? illness for third from the legacy her
 ‘during (her) fatal illness for a third of her legacy,’

a=wa yiti(?) dāxil ī mahr=hi
 ('a.w yī d'x'yl 'y m^ahr.h)
 Q-therein is included the dowry-her
 ‘is her (unpaid) dowry included therein,’

k-ān=ci arcyād(a)k ī wašī parēx-ica wa kadāki? āy.ās
 (k'n.c 'rcy'dk 'y wšy pryxc w' kδ'k? 'γ. 's)
 that there from third the trustee leaves-PRS and is it so? yes
 ‘so that the trustee leaves a third of it (unpaid), or not?’ ‘Yes.’

The Arabic text is:

*fī-mā 'idā 'awšati l-mar'utu ilā zawji-hā fī maraḍi mawti-hā fi0-0ul0i min tarikati-hā
 hal yadxulu l-mahru fi l-wašiyati ḥattā <yuf>riju l-wašiyu 0ul0a-hū 'am lā? – na'm.*

ACKNOWLEDGEMENTS

I would like to acknowledge here the generous assistance I had the pleasure of receiving from D. N. MacKenzie, now sadly deceased, during the preparation of this essay. Needless to say, the mistakes are my own. In particular he would not have approved of the attempt to transcribe Khwarezmian, suggesting as it does that we know more than is in fact possible.

I would also like to express my gratitude to G. Windfuhr for many suggestions, additions, and improvements.

4 April 2005

REFERENCES AND BIBLIOGRAPHY

- Azami, Ch. A. and Windfuhr, G. (1972) *A Dictionary of Sangesari, With a Grammatical Outline*. Tehran: Franklin Book Company.
- Benzing, J. (1968) *Das chwaresmische Sprachmaterial einer Handschrift der "Muqaddimat al-Adab" von Zamaxšarī*. Wiesbaden: F. Steiner.
- (1983) *Chwaresmischer Wortindex*, Taraf, Z. (ed.), Einleitung von H. Humbach. Wiesbaden: O. Harrassowitz.
- Benzing, J. and Yüce, N. (1985) 'Chwaresmische Wörter und Sätze aus einer choresmtürkischen Handschrift der Muqaddimat al-Adab', *ZDMG* 135: 92–103.
- Bogoliubov, M. N. (1961a) 'Chastitsy v khorezmiiskom iazyke' (Particles in the Khwarezmian language), in Bogoliubov, M. N. and Iakhontov, S. E. (eds) *Iazyki narodov vostoka*, *UZLU* 12: 81–84.
- (1961b) 'O nekotorykh osobennostiakh arabo-khorezmiiskoi pis'mennosti' (About some peculiarities of the Arabo-Khwarezmian writing system), *NAA* 4: 182–187.
- (1962) 'Lichnye mestoimeniia v khorezmiiskom iazyke' (Personal pronouns in the Khwarezmian language), *UZLU* 16: 6–15.
- (1963) 'Mestoimeniia v khorezmiiskom iazyke' (Pronouns in the Khwarezmian language), *KSINA* 67: 99–103.
- (1985) 'Khorezmiiskie kalendarnye glossy v "Khronologii Biruni"' (Khwarezmian calendrical glosses in Biruni's *Chronology*), *Vla* 1985.1: 28–33.
- Durkin-Meisterenst, D. (2005) 'Khwarezmian lexicology', in Cereti, C. G. and Maggi, M. (eds) *Middle Iranian Lexicography*. Roma, pp. 147–166.
- Freiman, A. A. (1951) *Khorezmiiskii iazyk* (The Khorezmian language). Moskva and Leningrad: Izd-vo Akademii nauk SSSR.
- Frumkin, G. (1965) 'Archaeology in Soviet Central Asia. V. The deltas of the Oxus and Jaxartes: Khorezm and its borderlands', *CAR* 13: 69–86.
- Gershevitch, I. (1954) *A Grammar of Manichean Sogdian*. Oxford: B. Blackwell.
- Henning, W. B. (1936) 'Über die Sprache der Chvarezmier', *ZDMG* 90: *30*–*34* [= *Selected Papers* I, pp. 401–405].
- (1955 [1956]) 'The Khwarezmian language', in *Zeki Velidi Togan'a Armağan* [Symbolae in honorem Zeki Velidi Togan], Istanbul 1955 [appeared 1956], pp. 421–436 [= *Selected Papers* I, pp. 485–500].
- (1956) 'The structure of the Khwarezmian verb', *Asia Minor*, n.s. 5: 43–49 [= *Selected Papers* II, pp. 449–456].
- (1958) 'Mitteliranisch', in *HbO, Iranistik* I, pp. 20–130.
- (1965) 'The Choresmian documents', *Asia Minor* n.s. 10: 166–179 [= *Selected Papers* II, pp. 645–658].
- (1971/1317AH) *A Fragment of a Khwarezmian Dictionary*, MacKenzie, D. N. (ed.). London: Lund Humphries and Tehran: Publications de l'Université de Tehran.
- (1977) *Selected Papers*, Boyce, M. (ed.) 2 vols. (= *AcIr* 14–15), Tehran: Bibliothéque Pahlavi, Leiden: Brill.
- Humbach, H. (1985). 'Gōpatšāh, his country and the Khwārezmian hypothesis', *AcIr* 24: 327–334.
- (1989) 'Choresmian', in *CLI*, pp. 193–203 (review by D. N. MacKenzie, *BSOS* 54 (1991): 172–176 [= *Iranica Diversa* II, pp. 587–93]).
- Livshits, V. A. (1968) 'The Khwarezmian calendar and the eras of Ancient Chorasnia', *AAntH* 16: 433–446.
- (1984) 'Dokumenty', in Rapoport, Iu. A. and Nerazik, E. E. (eds), *Toprak-kala. Dvoretz* (Toprak-kala. The palace), *Trudy Khorezmskoi arkheologo-ètnograficheskoi èkspeditsii* 14, Moskva, pp. 251–286.

- Livshits, V. A. and Mambetullaev, M. M. (1986) 'Ostrak iz Humbuz-Tepe' (An ostrakon from Khumbuz Tepe), in *Pamiatniki istorii i literatury vostoka*, Moskva, pp. 34–45.
- MacKenzie, D. N. (1970–1972) 'The Khwarezmian glossary: 1–5, I: *BSOAS* 33 (1970): 540–458; II: 34 (1971): 74–90; III: 314–330; IV: 521–536; V: 35 (1972): 56–74 [= *Iranica Diversa* II, pp. 459–550].
- (1971/1317 AH) 'Supplement', in Henning, W. B. *A Fragment of a Khwarezmian Dictionary*. London and Tehran (see Henning 1971), pp. 28–50.
- (1975) 'Khwarezmian imperfect stems', in *Mélanges linguistiques offerts à Emile Benveniste*. Louvain: Peeters, pp. 389–395 [= *Iranica Diversa* II, pp. 551–557].
- (1987) 'Khwarezmian in the law books', Paper presented at the Second European Seminar on Central Asian Studies, London, pp. 7–10 April 1987 [= *Iranica Diversa* II, pp. 575–586].
- (1988a) Review of Samadi, M. (1986), *JRAS* 1988: 197–199.
- (1988b) 'Khwarezmian and Avestan', *East and West* 38: 81–92 [= *Iranica Diversa* II, pp. 559–570].
- (1989) 'Khwarezmian in the law books', in Fouchécour, C. -H. and Gignoux, P. (eds) *Études irano-aryennes offertes à Gilbert Lazard* Paris: Association pour l'avancement des études iranniennes [= *Iranica Diversa* II, pp. 575–586].
- (1990) (ed.) *The Khwarezmian Element in the Qunyat al-munya*. London. (reviews by Y. Yoshida, *JAOS* 112 (1992): 535–537 and P. O. Skjærvø, *BSOAS* 54 (1991): 496–505.
- (1991) Review of Schmitt, R. (1989) *Compendium Linguarum Iranicarum*, *BSOAS* 54: 172–176 [= *Iranica Diversa* II, pp. 587–593].
- (1992) 'Chorasmia. iii. The Chorasmian language,' in *Enlr* 7, pp. 517–520.
- (1993) 'Clitics in Khwarezmian', in Skalmowski, W. and van Tongerloo, A. (eds) *Medioiranica: Proceedings of the International Colloquium organized by the Katholieke Universiteit Leuven from the 21st to the 23rd of May 1990*. Leuven: Peeters, pp. 135–145 [= *Iranica Diversa* II, pp. 595–605].
- (1999) *Iranica Diversa*, Cereti, C. G. and Paul, L. (eds), 2 vols. (Serie Orientale Roma LXXXIV, 1–2). Roma: Istituto italiano per l'Africa e l'Oriente.
- Rapoport, Ju. A., Nerazik, E. E. and Levina, L. M. (2000) *V nizov'iakh Oksa i Iaktsarta. Obrazy drevnego Priaral'ia* (In the lower reaches of the Oxus and the Jaxartes. Images of the old Priaral). Moskva: Indrik.
- Samadi, M. (1986) *Das chwaresmische Verbum*. Wiesbaden: O. Harrassowitz.
- Sims-Williams, N. (1979) 'On the plural and dual in Sogdian', *BSOAS* 42: 337–346.
- (1989) 'New studies on the verbal system of Old and Middle Iranian', *BSOAS* 52: 255–264.
- (1996) 'On the historic present and injunctive in Sogdian and Choresmian', *MSS* 56: 173–189.
- Tolstov, S. P. (1948) *Drevnii Khorezm* (Ancient Khorezm). Moskva: Izdanie MGU.
- (1948) *Po sledam drevnekhorezmiiskoi tsivilizatsii* (On the trail of the Old Khwarezmian civilization). Moskva and Leningrad: Izd-vo Akademii nauk SSSR; German trans. 1953, *Auf den Spuren der altchoresmischen Kultur*, O. Mehlitz (trans.). Berlin: Verlag Kultur und Fortschritt.
- (1954) 'Le Kharezmi ancien' in Piotrovsky, B. B., Schultz, P. N., Golovkina, V. A. and Tolstov, S. P. (eds) *Ourartou Neapolis des Scythes Kharezmi* (L'Orient ancien illustré). Paris: Maisonneuve, pp. 105–166.
- Tolstov, S. P. and Livshits, V. H. (1964) 'Decipherment and interpretation of the Khwarezmian inscriptions from Tok-Kala', *AAntH* 12: 231–251.
- Tremblay, X. (2002) 'Is die Aktivendung 3Pl -āra in einigen ostiranischen Sprachen inneriranische Entwicklung oder indogermanisches Erbe?', *MSS* 62: 259–287.
- Vainberg, B. J. (1977) *Monety drevnego Khorezma*. Moskva: Nauka.

- Windfuhr, G. (1975) 'Isoglosses: A sketch on Persians and Parthians, Kurds and Medes', *AcIr* 5: 457–472.
- Yoshida, Y. (1983) Review of Benzing (1983), *Journal of Asian and African Studies* 31: 172–176.
- (1992) Review of MacKenzie (1990), *JAOS* 112: 535–537.
- Zeki Velidi, A. (1927/8) 'Hwāresmische Sätze in einem arabischen Fiqh-Werke', *Islamica* 3: 190–213.
- (1936) 'Über die Sprache und Kultur der alten Chwarezmier', *ZDMG* 90: *27*–*30*.
- (1951). *Horezmce tercümelî Muqaddimat al-adab. Khorezmian Glossary of the Muqaddimat al-adab.* = Horezm kültürü vesikalari. Documents on the Khorezmian Culture, kısım 1. Part 1. Istanbul: Sucuoğlu Matbaası.
- For further bibliography, see CLI, pp. 202–03.

KHOTANESE AND TUMSHUQESE

Ronald E. Emmerick†

1 INTRODUCTION

1.1 Overview

1.1.1 Historical background

Khotanese and Tumshuqese, properly Gyāz̄dian (Rong and Duan 1996 [2000]), are Middle Iranian languages that were once spoken in what is now known as the Xinjiang (earlier Sinkiang) Autonomous Region of China. They are treated here together because they are more closely related to each other than to any other Iranian language. Both geographically and linguistically they are classified as East Iranian languages and are thus related more closely to other East Middle Iranian languages such as Sogdian than to Iranian languages spoken in western regions.

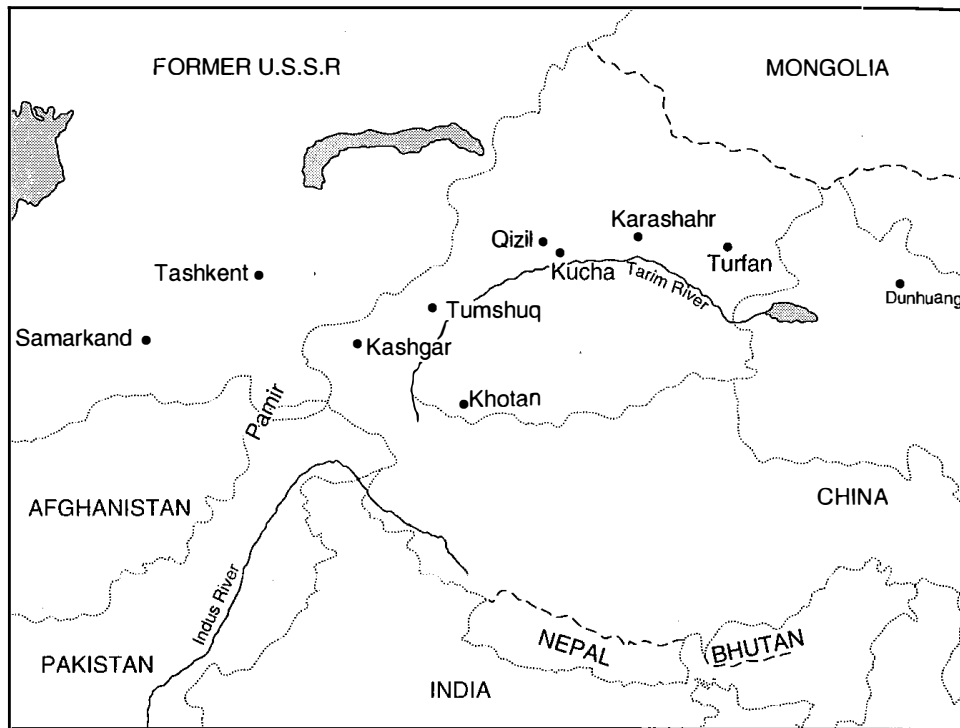
Although Khotanese and Tumshuqese are so closely related to each other that we can confidently posit a common parent language, such a language remains entirely hypothetical and is likely to be the language of a nomadic tribe that moved about Central Asia before settling in oases around the Tarim basin some time during the second half of the first millennium BCE.

After speakers of the Tumshuqese-Khotanese parent language settled on the northern side of the Taklamakan desert they appear to have had little contact with their fellow tribesmen who settled along the southern rim of the desert. The language spoken in the north developed in the course of time into what we now term Tumshuqese after the locality where the most of the extant texts were found whereas the language spoken in the south developed into what we call Khotanese since it was spoken in the ancient kingdom of Khotan.

Little is known of the history of the Tumshuqese speakers in the north and even their language is not well known since few texts have survived.

The situation is rather better in the case of Khotan, concerning which Chinese sources provide extensive information over a long period. In addition, there survive a very considerable number of documents in Khotanese. Since many of them are translations from known languages it has been possible to decipher them with confidence.

We do not know exactly when speakers of Khotanese first settled in the area, but it could hardly have been before the founding of Khotan, which is likely to have been in the third century BCE. There are traces of the presence of Khotanese speakers that date long before the earliest datable documents in Khotanese. Of particular importance is a document written in an Indian language, North-West Prakrit, that is thought to date from the third century CE. It is dated in a regnal year 'of the Great King of Khotan, King of kings' and betrays clear evidence of the Khotanese language.



MAP 7.1 KHOTAN AND TUMSHUQ

Long before the seventh century Buddhism was well established in Khotan, which by then was a leading centre of Mahayana Buddhism. Almost all the extant texts in Khotanese are coloured by Buddhism. Many of them are translations of Buddhist texts while some of them appear to be original compositions concerning Buddhist lore and doctrine.

1.1.2 Textual material

1.1.2a Khotanese

Most important for the secure decipherment of the language are the numerous works translated from Buddhist Sanskrit *sūtra* literature and a large medical text, the *Siddhasāra*, translated from Tibetan with account being taken of its Sanskrit source. There are also many secular documents dated in the regnal years of Khotanese kings. Relatively little original Khotanese literature has survived, but we have among other items a charming version of the *Rāmāyaṇa*, some lyrical verses, a substantial fragment of a metrical composition containing a love story, the description of a journey through Gilgit and Chilās to Kashmir, and some literary epistles.

The bulk of the extant documents all appear to date from the period between the seventh and the tenth century CE, but some fragments have been dated on palaeographical grounds to the fifth and sixth centuries CE (for details, see Emmerick 1992a).

1.1.2b Tumshuqese

Far less material in Tumshuqese has come to light than in Khotanese so that it remains difficult to decipher it and many details of grammar remain unknown or uncertain. The oldest text is the so-called Karmavācanā text, which is also the most important text since it is the only one for which we have bilingual evidence. It was translated from Tocharian (Schmidt 1988). In addition there are fourteen published items (see Skjærvø 1987: 77), including legal documents, fragments of Buddhist literary works, sales documents of commodities in a Buddhist monastic milieu, and several unpublished fragments in various collections.

None of the Tumshuqese documents can be precisely dated, but there are grounds for assigning the majority of them to the late seventh or the eighth century CE (for details, see Emmerick 1992a).

1.2 Writing systems

In Khotanese, besides a few numerals in Tibetan script (Maggi 1995), the extant material is written in varieties of Central Asian Brāhmī. The varieties range from highly formal to cursive. Only the formal varieties have been the object of detailed study. They were classified as: (1) Early Turkestan Brāhmī; (2) Early South Turkestan Brāhmī; (3) South Turkestan Brāhmī; and (4) Late Turkestan Brāhmī (see Sander 1986; Skjærvø 2002: lxxi–lxxii).

TABLE 7.1: THE EARLY SOUTH TURKESTAN BRĀHMĪ SCRIPT OF KHOTAN
(FROM LEUMANN 1934)

1 A	2 ¹ Ā	3 ² Ā	4 I	5 ¹ ī	6 ² ī	73 khā	74 ¹ lysā	75 rvā	76 mā	77/78 tā-ndi	
7 U	8 ū	9 R	10 E	11 O	12 AU	79 ṣṭā	80 hwā	81 jā	82 jyā	83 jā	84 jā
13 ka	14 kha	15 ga	16 gga	17 gha	18 ṅga	85 ttra	86 ² lysā	87 ñi	88 hi	89 ¹ ysi	90 li
19 ca	20 ¹ cha	21 ² cha	22 jā	23 ¹ ña	24 ² ña	91 jsi	92 ṣpri	93 ḍu	94 ṣku	95 dū	96 ru
25 ṭa	26 ṭha	27 ṭhha	28 ḍa	29 ṇḍa	30 ṇa	97 nu	98 pu	99 rru	100 ¹ rrū	101 ² rrū	102 rū
31 ta	32 tta	33 tha	34 da	35 dha	36 ṇa	103 bū	104 hū	105 ysmū	106 gu	107 śsu	108 śū
37 pa	38 pha	39 ba	40 ¹ bha	41 ² bha	42 ma	109 ¹ ru	110 ² ttu	111 ¹ pyū	112 ¹ pyū	113 grū	114 brū
43 ya	44 ra	45 la	46 va	47 100	48 200	115 ¹ stū	116 ² ttū	117 rvī	118 kye	119 klai	120 tcej
49 ṣa	50 ṣā	51 ṣa	52 ha	53 300	54 400	121 ysno	122 khyau	123 lo	124 dām	125 bim	126 ṣsim
55 1	56 2	57 3	58 10	59 20	60 30	127 rtha	128 dda	129 nna	130 kkra	131 ṣḍe	132 tva
61 4	62 5	63 6	64 40	65 50	66 60	133 yyo	134 ḍye	135 tṭye	136 ² pyu	137 gya	138 jvā
67 7	68 8	69 9	70 70	71 80	72 90	139 ryau	140 crrā	141 rddhā	142 rṣta	143 drai	144 m

The Tumshuqese documents are written in formal (Karmavācanā and Tum. VIII–IX) and cursive varieties of North Turkestan Brāhmī that are similar to those used by Tocharians. The Tumshuqese may have adopted their writing system from the Tocharians adding nine signs not used for Tocharian. The precise interpretation of a number of these signs is still to be determined. The absence of long vowels other than *ā* is noticeable. In comparison, the Khotanese alphabet and that used in the Tumshuqese Karmavācanā text, with signs found in other texts added, are as follows:

TABLE 7.2: COMPARATIVE ARRANGEMENTS OF KHOTANESE AND TUMSHUQESE ALPHABETS

Khot.	a ā ä	i ī	u ū	r	e ai ei o au
Tumsh.	a ā ä	i	u	r	e ai o au
Khot.	k k ^h	g g ^h	ñ		
Tumsh.	k	g	ñ		
Khot.	c c ^h	j j ^h	ñ		
Tumsh.	c c ^h	j	ñ		
Khot.	ṭ ṭ ^h	ḍ ḍ ^h	ṇ		
Tumsh.			ṇ		
Khot.		t t ^h	d d ^h	n	
Tumsh.	ts	t t ^h	d d ^h	n	
Khot.		p p ^h	b b ^h	m	ṃ
Tumsh.		p	b	m	ṃ
Khot.	y	r l	v	ś ś s	h
Tumsh.	y	r r l	v	ś ś s ž	h
Tumsh.	<u>k</u> γ	<u>d</u> dz	w	ž	xš

There are certain differences in the use of several features. Thus, the Khotanese writing system, while largely identical with Sanskrit, includes digraphs for sounds not represented in the Brāhmī script, e.g. *ys* for [z], and diacritics added to an akṣara for the following: (1) *ä* [ə] by two superscript dots, approximately [e] in stressed syllables (Emmerick and Pulleyblank 1993: 45–46); (2) *ei* [aə] by a superscript sign resembling a St. Andrew's cross (Emmerick 1998); (3) nasalisation by one superscript dot, transcribed by *m* with underdot or a hook below the vowel, e.g. *ṃ* or *ṁ*; (4) aspiration by two dots, transcribed by a colon (Late Khotanese); (5) breathed syllable by a hook below vowel, transcribed by a hook below the vowel or by an apostrophe at the end of the syllable.

1.3 Orthography

Khotanese texts are characterised by very considerable spelling variation that not seldom impedes interpretation. Thus, the texts vary in date and reflect different linguistic stages of the language. They belong to different genres and reflect different registers. Beside archaising formal speech we find colloquial language; besides texts carefully written by learned monks and official documents, we find personal letters, casual notes and unskilled school exercises. There appear to have been different orthographic conventions, possibly associated with different monasteries. Moreover, the texts were in many cases copied by scribes whose speech differed from the language of the text they were copying and they allowed their speech to influence their work.

The Tumshuqese Karmavācanā text is earlier than the other Tumshuqese documents probably by several centuries, and was written before the introduction of the nine special signs. In view of our limited knowledge of the language it is not safe to draw any conclusions from apparent variations in spelling.

2 PHONOLOGY

2.1 Inventory and distribution

2.1.1a Khotanese

See especially Emmerick 1981a, 1981b, 1989, 1992b, 1998, Emmerick and Maggi 1991, Emmerick and Pulleyblank 1993, Kumamoto 1995.

2.1.1a.1 Vowels

The vowel system of the older language appears to have had 11 phonemes. In the later language the distinction of quantity was replaced by one of quality with a consequent reduction to the 6 phonemes *i u e o ə a*. Also, the diphthongs *ai*, *au*, and *aə* of the older language were monophthongised.

TABLE 7.3: VOWELS—KHOTANESE

	Front	Central	Back
High	<i>ī, i</i>		<i>u, ū</i>
Mid	<i>ē, e</i>	<i>ə</i>	<i>o, ō</i>
Low	<i>a</i>	<i>ā</i>	

2.1.1a.2 Consonants

The consonant system appears to have distinguished 41 or 42 phonemes.

TABLE 7.4: CONSONANTS—KHOTANESE

	Labial	Dental	Retroflex	Palatal	Velar	Glottal
Plosives	<i>p^h</i> <i>p</i> <i>b</i>	<i>t^h</i> <i>t</i> <i>d</i>	<i>tʰ</i> <i>ʈ</i> <i>ɖ</i>		<i>k^h</i> <i>k</i> <i>g</i>	
Affricates		<i>ts^h</i> <i>ts</i> <i>dz</i>	<i>tʂ^h</i>	<i>tʃ^h</i> <i>tʃ</i> <i>dʒ</i>		
Fricatives	<i>f</i> <i>v</i>	<i>s</i> <i>z</i>	<i>ʃ</i> <i>ʒ</i>		<i>x</i> <i>ɣ</i>	<i>h</i>
Nasals	<i>m</i>	<i>n</i>	<i>ɳ</i>	<i>ɲ</i>	<i>ŋ</i>	
Trill		<i>r</i>				
Approximants	<i>w</i>	<i>rr</i>	<i>ʀ</i>	<i>y</i>		
Lateral		<i>l</i>				

Of these, *f* and *x* occur only in Late Khotanese renderings of foreign words. The exact phonetic nature of the two *r*-sounds is unknown. They are distinguished in the script by signs that are conventionally transliterated as *r* and *rr*, although *rr* is not a graphic combination of two *rs* and does not represent a long consonant. In Khotanese the retroflex phonemes were not confined to loanwords as in the case of Tumshuqese but formed an integral part of the system. Both *r*-phonemes may have had retroflex allophones.

2.1.1b *Tumshuqese*

No attempt has so far been made to reconstruct the Tumshuqese phonological system as the material available is so scanty.

2.1.1b.1 Vowels

Although the script does not reflect length distinction except in the case of *ā* and *a*, it can be assumed that it was phonemic on systemic and historical-comparative grounds. The central vowel *ä* [ə] is commonly found in unstressed syllables originally containing *a*. Other vowels may have had an allophone [ə] in unstressed syllables. Two diphthongs, *ai* and *au*, are attested. It is probable that a set of nasalised vowels also existed.

2.1.1b.2 Consonants

The comparison with Khotanese suggests a similar system. This includes the assumption of two kinds of *r*. Differences may also reflect different stages of development. Thus Tumshuqese has the voiced fricatives *β*, *v*, and *δ*, and no evidence for the glottal stop, or for *f*, and *x*. The aspirated plosives found in Khotanese could theoretically be still fricatives in Tumshuqese.

2.1.2 *Syllable and clusters*

Only in the case of Khotanese do we have any means of determining the syllabic structure of the language. Analysis of Old Khotanese metrical texts (Emmerick 1968b) reveals that non-initial syllables may begin with any single consonant phoneme including the affricates whether or not they are followed by *y*. We have no evidence regarding consonant groups containing a sibilant plus a single consonant except for *ysm* [zm], which does not make position in iambic metre so that the word *aysmū* ['azmu:] must have been divided *a-ysmū*. The word *biśśa*- 'all' < Indo-Iranian **yiśśya*- could be treated in metrical texts as having its first syllable either heavy or light. Since long consonants otherwise make position, it appears to be the case that [lbiśśa] was in the process of being simplified to [lbiśa] at the time the texts were being composed.

2.2 Non-segmental features

Since the position of the stress is not recorded in the script, it can only be inferred on the basis of historical phonology. Khotanese accentuation has been treated by M. Maggi (1992, 1993). According to him, in polysyllables the Old Khotanese stress accent fell on the first heavy syllable from the end of the word but never on the word final syllable.

A word containing only light syllables was stressed on the initial syllable. There is insufficient evidence to say much regarding the stress in Tumshuqese.

3 MORPHOLOGY

3.1 Nominal morphology

Both Khotanese and Tumshuqese preserve the morphological distinction of masculine and feminine gender, and of singular and plural number. Neuter is still distinguished in Khotanese although it has largely been replaced by masculine and feminine, while the evidence for neuter in Tumshuqese is at present insufficient (in the nom. pl., Schmidt 1988: 312).

Old Khotanese and probably also Tumshuqese had a six-case system, reduced from the eight-case system of Indo-Iranian: nominative, accusative, locative, genitive-dative, instrumental-ablative, and vocative. The two compound cases result from morphological merger but retained the syntactic functions of their components. The accusative is identical to the nominative in the plural. The vocative is *-a* in the singular of the *a*-declension and *-ā* in the singular of the *ā*-declension; elsewhere it is identical with the nominative in the singular, but with the instrumental-ablative in the plural.

3.1.1a Nominal stem classes and declensions—Khotanese

The following summary is largely confined to Old Khotanese.

Nouns and adjectives are inflected according to vocalic (primary and secondary) and consonantal declensions.

The primary vocalic declensions are:

the *a*-declension which is masculine or neuter;
and the *ā*- and *i*-declensions which are feminine.

The secondary vocalic declensions attested are:

masculine or neuter *-aa*, *-ia*, *-ua*, *-āa*, *-īa*, *ūa*;
feminine *-aā*, *-iā*, *-uā*, *-āā*; and *-ai*, *-ii*, *-au*.

They arose as a result of the loss of intervocalic consonants, and the merger of the resulting short or long vowels with the vocalic endings, mostly of the *a*- and *ā*-declensions, rarely of the *i*-declension, while the *au*-declension arose from words in an original labial.

For the consonantal declensions the following subdivisions may be made:

the masculine and feminine *r*-declensions;
the masculine and neuter *n*-declensions;
and the *nd*-declension which has natural gender.

Of these, the masculine and feminine declensions have almost exclusively consonant stem endings only in the nominative (sing. *-e*, plural in *-ā*), but are based on a thematic stem in the other cases. Isolated forms show traces of other consonant declensions, e.g. the nom. sg. *ysar-e* 'old age' < **zar-āh* (OIr. *h*-stem **zar-ah*-).

It should be noted that there is considerable variation in orthography, and that most of the forms cited in the tables are attested in other spellings, especially when reflecting

different stages. Thus, the distinction between nom. sg. *-ā* and gen.-dat. sg. *-i* in the primary *a*-declension is observed only in a few Old Khotanese texts. Endings that have spread from the *n*-declension and from the pronominal declensions (Sims-Williams 1990) are on the whole not included. (See Emmerick 1968a: 249–349 and Sims-Williams 1990.)

3.1.1a.1 Vocalic declensions:

Examples for primary and secondary vocalic declensions:

masc. *-alfem.* *-ā*: *balysa-* ‘Buddha’, *uryāna-* ‘garden’ – *kanthā-* ‘city’, *dīvatā-* ‘goddess’, *hīnā-* ‘army’;
 masc. *-aalfem.* *-āā*: *āchaa-* ‘disease’, *ysamaśśandaa-* ‘world’, *balysūñavūysaa-* ‘bodhisattva’, *ggāṭṭhaa-* ‘householder’ – *baśdaā-* ‘sin’, *brītaā-* ‘love’;
 fem. *-i*: *mulysdi-* ‘compassion’, *balysūsti-* ‘enlightenment’, *bā’yi-* ‘ray’, *hālsti-* ‘spear’, *tcārami-* ‘sphere’.

TABLE 7.5: VOCALIC DECLENSIONS—KHOTANESE

Singular	<i>-a</i> masc.	<i>-ā</i> fem.	<i>-aa</i> masc.	<i>-āā</i> fem.	<i>-i</i> fem.
NOM	<i>balys-ā</i>	<i>kanth-a</i>	<i>āch-ei</i>	<i>baśd-ā</i>	<i>mulysd-ā</i>
ACC	<i>balys-u</i>	<i>kanth-o</i>	<i>āch-au</i>	<i>baśd-o</i>	<i>mulśd-u</i>
G–D	<i>balys-i</i>	<i>kīnth-e</i>	<i>āch-ai</i>	<i>baśd-(i)ye</i>	<i>mulśd-e</i>
I–A	<i>balys-āna</i>	<i>kīnth-e jsa</i>	<i>āch-aina</i>	<i>baśd-(i)ye</i>	<i>mulśd-e jsa</i>
LOC	<i>uryāñ-a</i>	<i>kīnth-a</i>	<i>ysamaśśand-i ya</i>	<i>brīt-ya</i>	<i>balysūst-a</i>
VOC	<i>balys-a</i>	<i>dīvat-ā</i>	<i>balysūñavūys-ā</i>		
Plural	<i>-a</i> masc.	<i>-ā</i> fem.	<i>-aa</i> masc.	<i>-āā</i> fem.	<i>-i</i> fem.
N–ACC	<i>balys-a</i>	<i>kanth-e</i>	<i>āch-ā</i>	<i>baśd-e</i>	<i>hālšt-ā</i>
G–D	<i>balys-ānu</i>	<i>kanth-ānu</i>	<i>āch-ānu</i>	<i>baśd-ānu</i>	<i>hālšt-ānu</i>
I–A	<i>balys-yau</i>	<i>hīn-yau jsa</i>	<i>āch-yau</i>		<i>bā’y-yau jsa</i>
LOC	<i>uryāñ-uvo’</i>	<i>kanth-uvo’</i>	<i>ggāṭṭh-uvo’</i>		<i>tcārīm-uvo’</i>
VOC	<i>balys-yau</i>				

In *a*- and *aa*-stems, neuter differs only in the nom.-acc. pl., where they have *-e* instead of *-a*, and *-e* instead of *-ā* respectively, e.g. *kīra-* ‘act’ has *kīr-e*, and *āstaa-* ‘bone’ has *āst-e*. The same endings are found in Tumshuqese and may reflect a similar distinction (Schmidt 1988: 312).

3.1.1a.2 Declensions, *āa*- and *āā*-stems

Examples for *āa*- and *āā*-stems:

dāa- ‘fire’, *pāa-* ‘foot’; *nīāa-* ‘river’, *biśāa-* ‘tongue’, *hvāa-* ‘sweat’; *dyūa-* ‘demon’, *busūa-* ‘fuel’, *syūa-* ‘orphan’; masc. *raysā-a-*, fem. *-ā-* ‘empty’; masc. *paḍā-a-*, fem. *-ā-* ‘first’; *palāā-* ‘banner’.

Monosyllabic and polysyllabic *āa*-stems are declined differently. No such distinction is apparent in the few attested cases of the *īa*- and *ūa*-declensions. Some words have irregular declension. The forms of the *āā*-declension tend to merge with either with the *āā*- or the *au*-declension.

TABLE 7.6: DECLENSIONS OF *āa*- AND *āā*-STEMS—KHOTANESE

Singular	<i>āa</i> - stem monosyll.	<i>āa</i> -stem polysyll.	<i>āa</i> -stem. masc.	<i>āā</i> -stem. fem.
NOM	<i>dai</i>	<i>nit-ā</i>	<i>paḍ-ā; rrays-ā</i>	<i>rrays-ā</i>
ACC	<i>dau</i>	<i>nit-ā</i>	<i>paḍ-ā</i>	<i>paḍ-o; rays-au</i>
G-D	<i>dai</i>	<i>nāit-ā-yi</i>		
I-A	<i>daina</i>	<i>biš-ā-na</i>		<i>paḍ-ā-ye</i>
LOC	<i>dā-ā</i>	<i>nit-ā-ya</i>		
Plural	<i>āa</i> - stem monosyll.	<i>āa</i> - stem polysyll.	<i>āa</i> - stem. masc	<i>āā</i> - stem fem.
N-ACC	<i>p-ā</i>	<i>nāit-ā</i>	<i>rrays-ā</i>	<i>pal-e</i>
G-D	<i>p-ānu</i>	<i>biš-ānu</i>		<i>pal-yau</i>
I-A	<i>p-āyau</i>			
LOC	<i>p-vo'</i>	<i>nāit-ā-vo'</i>		

3.1.1a.3 Consonantal declensions

Examples for consonantal declensions:

nd-stems: *rre*, *rrund*- 'king', *hve'*, *hva'nd*- 'man';

n-stems: masc. *naḍe*, *naḍaun*- '(heroic) man', *šve*, *švān*- 'dog' – ntr. *ššāma*, *šāman*- 'face', *tcei'man*- 'eye', LateKh. *tsaman*-;

r-stems: masc. *pāte*, *pātar*- 'father', *brāte*, *brātar*- 'brother' – fem. *māta*, *mātar*- 'mother', *duta*, *dutar*- 'daughter'.

TABLE 7.7: CONSONANTAL DECLENSIONS—KHOTANESE

Singular	Endings	<i>nd</i> -stem	<i>n</i> -stem	<i>r</i> -stem masc.	<i>r</i> -stem fem.
NOM	-Ø	<i>rre</i>	<i>naḍe</i>	<i>pāte</i>	<i>māta</i>
ACC	- <i>u</i>	<i>rrund-u</i>		<i>pātar-u</i>	<i>mātar-u</i>
G-D	- <i>i</i>	<i>rrund-i</i>	<i>naḍaun-ā</i>	<i>pīr-ā</i>	<i>mer-ā</i>
I-A	- <i>ina</i>	<i>rrund-ina</i>	<i>naḍaun-āna</i>	<i>pyar-ina</i>	<i>mer-i jsa</i>
LOC	- <i>ia</i> , -(<i>i</i>) <i>na</i>		<i>ššamañ-a</i>		
VOC	-Ø	<i>rre</i>	<i>naḍe</i>		
Plural	Endings	<i>nd</i> - stem	<i>n</i> - stem	<i>r</i> -stem masc.	<i>r</i> -stem fem.
N-ACC	- <i>ā</i>	<i>rrund-ā</i>	<i>švān-ā</i>	<i>brātar-ā</i>	<i>dutar-ā, duñr-ā</i>
G-D	- <i>ānu</i>	<i>rrund-ānu</i>	<i>švān-ānu</i>		<i>duñr-ānu</i>
I-A	- <i>yau</i>	<i>rrund-yau</i>		<i>brātar-yau</i>	<i>dvatar-yau</i>
LOC	- <i>uvo'</i>	<i>hvañd-uvo'</i>	<i>tcamañ-ūā</i>		
VOC	- <i>yau</i>	<i>rrund-yau</i>	<i>naḍaun-yau</i>	<i>brātar-yau</i>	

3.1.2a Adjectives—Khotanese

Adjectives follow the same declensional patterns as nouns but also have some mostly optional endings adopted from the pronominal declension, and most adjectives take pronominal endings where otherwise palatalisation would be required.

The comparative suffix is *-tara-* and the superlative suffix *-tama-*, inherited from Indo-Iranian, e.g. *bata*- 'small', *battara-*, *battama-*, and the old pair *hastara-*, *hastama-* 'better,

best' to the positive *śśāra-* 'good'. Only the comparative seems to have been fully productive by the time of our texts.

3.1.1/2b Declension of nouns and adjectives—Tumshuqese

Tumshuqese seems to have developed basically the same declensional patterns as Khotanese. The most important difference is the gen.-dat. sg. ending *-ā* of the *a*-declension vs. Khot. *-i* (both from OIr. **-ahya*). Consonantal declensions similar to those in Khotanese are attested by such forms as nom. sg. *brāde* 'brother', cf. Kh. *brāte*, beside nom. pl. *brāre*, cf. Kh. *brātari*, nom. sg. *re* 'king' cf. Kh. *rre*, beside gen.-dat. sg. *ride*, cf. Kh. *rrundi*, and nom. sg. *hvaže* 'man', cf. Kh. *hve*', beside nom. pl. *hvažandi*, cf. Kh. *hva'ndi* (for the latter, see Skjærvø 1987: 84).

3.1.3 Pronouns and deixis

3.1.3.1a Personal and demonstrative pronouns—Khotanese

Some originally 1p and 2p forms of the personal pronouns are used as polite forms for the singular.

The demonstratives distinguish three degrees of deixis (see Sims-Williams 1994), all with the same set of endings:

- neutral *ś-*, *tt-*;
- near *śā-t-*, *ttu-t-u* (acc.), *ttā-t-* (reduplicated);
- remote *ś-ār-*, *tt-ār-*,

The neutral deictic demonstrative also serves as personal pronoun. Emphatic forms of the neutral pronouns are marked by the deictic particle *-ī*: *ś-ī*, *śa-ī*, *tv-ī*, *tt-ī* for *ś-ā*, *ś-a*, *tt-u*, *tt-ā*.

Occasionally double reduplications are found with the near deictic pronouns, e.g. *śā-śa*, *śā-śā*.

Pronouns are used as correlatives of the relative pronoun. Usually the interrogative-relative pronoun distinguishes animate *kye*, *ce* and inanimate-neuter *cu* in the nom.-acc. singular, but it does not distinguish number. In the following table contracted and some alternate forms are omitted.

TABLE 7.8: PRONOUNS—KHOTANESE

Singular	1	2	3 masc.	3 fem.	'who'	'what'
NOM	<i>aysu</i>	<i>thu</i>	<i>ś-ā</i>	<i>ś-a</i>	<i>kye</i> , <i>ce</i>	<i>cu</i>
ACC	<i>ma</i> , <i>muho</i>	<i>uhu</i>	<i>tt-u</i>	<i>tt-o</i>	<i>kye</i> , <i>ce</i>	<i>cu</i>
G-D	<i>mamā</i>	<i>tvī</i>	<i>tt-ye</i>		<i>kye</i> , <i>ce</i>	
I-A	<i>muho jsa</i>	<i>tvī jsa</i> , <i>uho jsa</i>	<i>tt-āna</i>	<i>tt-āñe</i>		<i>tcam-āna</i>
LOC			<i>tt-āñā</i>	<i>tt-āñā</i>		<i>tcam-āñā</i>
POSS	<i>mamānaa-</i>	<i>tvānaa-</i>	<i>tt-ye</i>			
Obl. encl.	<i>mī</i> , <i>-ṃ</i>	<i>tī</i> , <i>-e</i>	<i>-ī</i> , <i>-yā</i>			
I-A			<i>-n jsa</i> (<i>-ṃ jsa</i> , <i>-jsa</i>)*			

(Continued overleaf)

TABLE 7.8: Continued

Plural	1	2	3 masc.	3 fem.	'who'	'what'
NOM	<i>buhu, muhu</i>	<i>uhu</i>	<i>tt-ī</i>	<i>tt-e</i>	(same as singular)	
ACC	<i>maha</i>	<i>uhu</i>	<i>tt-ā</i>	<i>t-te</i>		
G-D	<i>māvu,</i> <i>mā(nu)</i>	<i>umāvu,</i> <i>umā(nu)</i>	<i>tt-ānu</i>			
I-A	<i>mih-yau</i>	<i>um-yau jsa</i>	<i>tt-yau</i>			
LOC			<i>tt-uvo'</i>			
VOC		<i>um-yau</i>				
POSS	<i>mājāa-,</i> <i>mānia-</i>	<i>umājāa-,</i> <i>umānia-</i>	<i>tt-ānu</i>			
Obl. encl.	<i>nā, -m</i>	<i>-ū</i>	<i>nā, -n, -m</i>			

*Ex.: *ttātā-n jsa = ttātā jsa (ttātā-Ø jsa)* 'they from it' (Suv 0.5), *vyātaraṇe-n jsa = vyātaraṇe jsa (vyātaraṇe-Ø jsa)* 'the predictions from it' (Suv 0.8). [Skjærvø]

3.1.3.1b Personal and demonstrative pronouns—Tumshuqese

Few pronouns are certainly attested in Tumshuqese. Their forms mostly reflect earlier stages of development from Old Iranian in comparison with Khotanese: 1s nom. *asu, azu*, encl. obl. *me*, 2s nom. *to*, gen.-dat. *tivya*, 1p *mvo*.

3.1.3.2 Reflexive and reciprocal pronouns

The reflexive possessive adjective is *h̄via-* 'one's own' for all three persons. It is attested also in Tumshuqese *havya, h̄vyä* (Emmerick 1985a); the pronoun is *uysānā-* 'self'. The reciprocal pronoun is *śśūjāta-* 'one another', lit. 'one the second'. Its equivalent in Tumshuqese may be *śodi, śode*.

3.1.3.3 Indefinite and interrogative pronouns and adverbials

Khotanese has only one indefinite pronoun *ye* 'one', which is enclitic and invariable, e.g. *ne ju ye* 'no one' where it follows the negation and the indefinite particle *ju*. It has not yet been established for Tumshuqese.

Generalised indefiniteness 'ever' may be expressed by the simple repetition of an interrogative-relative, e.g. *kāma-* 'which' in *kām-āñā kām-āñā kšīr-a* 'in whichever land'.

Insertion of the neutral pronoun is found in a few fixed phrases, e.g. *kye ś-ā kye* 'whoever' (nom. sg.), *ce tt-ā ce* (nom. pl.); *cu ś-ā cu* 'whatever' (nom. or acc. sg.). Most commonly it is indicated by particles such as *ju, buro, hanu* or *halci*, e.g. *kye hanā, ce -ju, -buro, -halci, -halci ju, -ju halci*, all 'whoever'; *cu -buro, -halci, -buro halci* all 'whatever'; *ku -buro, -halci* 'wherever'.

Alternate indefiniteness may similarly be expressed by the repetition of interrogative-relatives such as *kye - kye* 'one – some' or 'some – others':

kye dāñ-a pūt-tā kye vā ggar-na pat-īndā
'one falls into a fire, some fall from a mountain'.

The interrogative-relatives form pairs with correlatives and may themselves function as adverbs. Those of location and direction may be distinguished by deixis: near speaker, *mara(ta)* 'here'; near addressee, *ttara* 'there'; and away from speaker, *vara(ta)* 'there'. The main pairs are the following.

TABLE 7.9: ADVERBIAL DEMONSTRATIVES—KHOTANESE

<i>kāma-</i> , <i>cu</i>	‘which’	<i>ṣ-</i> ; <i>tta</i>	‘that’; ‘thus’
<i>crāma-</i>	‘of what kind’	<i>ttrāma-</i> , <i>ttandrāma-</i>	‘of that kind’
<i>canda-</i>	‘how much’	<i>ttanda-</i> , <i>ttandia-</i>	‘(just) so much’
<i>cāndāka-</i>	‘how much, little’	<i>ttāndāka-</i>	‘(so) little’
<i>cerä</i>	‘how much’	<i>tterä</i>	‘so much’
<i>cīyā</i>	‘when’	<i>ttīyā</i>	‘then’
<i>cālsto</i>	‘whither’	<i>ttolsto</i>	‘thither’
<i>ku</i>	‘where’	<i>mara</i>	‘here’
		<i>ttara</i>	‘there’
		<i>vara</i>	‘there’ (remote)

3.1.4a Adpositions—Khotanese

Khotanese has by far more postpositions than prepositions. Some adpositions occur in both positions, but there are no circumpositions. There is considerable variation with regard to case rection which is essentially determined by the semantics of the adposition and the context (see section 4.4.1 Use of cases). Most take the gen.-dat., some only or also the accusative (direction, extension, etc.), a few the instr.-abl.

3.1.4a.1 Postpositions

Postpositions found in Old Khotanese include: *āstanna* ‘etc.’, *udīssā* ‘for the sake of’, *kāḍana* ‘on account of’, *pracaina* ‘by reason of’, *iñā* ‘before’, *nimalśa* ‘behind’, *n(u)va* ‘behind’, *pīrmo* ‘at the head of’, *brum̄bāte* ‘in front of’, *vara (ta)*, *varālsto* ‘towards’, *vaṣṭa* ‘throughout’, *vaska* ‘for’, *vātā*, *vīrā* ‘in, at, on’. The last pair of interchangeable postpositions occurs in numerous idiomatic phrases, such as *hamu vātā*, *hamu vīrā* ‘always’, with nom.-acc. sg. neut. of *hama-* ‘all’. Also used as prepositions are: *baña* ‘beside’, *bendä* ‘upon’, *patāna* ‘before’.

3.1.4a.2 Prepositions

Prepositions include: *patä* ‘to(wards), in the presence of’, *odi* ‘up to, until’, *naysdä* ‘near’, *myāño* ‘in the middle of’, *dī* ‘under’, *anau* ‘without’, *vina*, *vānau* ‘without, apart from’.

3.1.4b Adpositions—Tumshuqese

In Tumshuqese there is little evidence for adpositions. The postposition *tsi* is used like Khotanese *jsa* (< OIr. **hačā*) as a marker of the instr.-abl. case. The relationship of *au* ‘up to’ (< OIr. **awa*) to Khot. *odä* is not clear (Emmerick 1985a: 20).

3.1.5 Adverbs

There are three productive ways of forming adverbs from other parts of speech: (1) by using the acc. sg. nt. of an adjective, e.g. *śśār-u* ‘well’ < *śśāra-* ‘good’, including relative adjectives, e.g. *ttand-u* ‘just a little’; (2) by means of suffixes such as *-ālsto* added to adjectives, adverbs, pronouns, adpositions, and locative nouns, e.g. *bīśś-ālsto* ‘homewards’ < *bīsā-* ‘house’ (Degener 1989a: 105–112); (3) by means of the language suffix *-au*

added to an adjective, e.g. *hvatan-au* ‘in(to) Khotanese’ < *hvatana-* ‘Khotanese’ (Degener 1989a: 172–173).

Like adjectives, adverbs may be *intensified* by preposing an adverb such as *atä* or *käde* ‘very’, *tvare* or *bihīyu* ‘extremely’ or more rarely by repetition, e.g. *mulšu mulšu* ‘at very short intervals’, *ta ta* ‘thus’.

In Tumshuqese there is little evidence for adverbs, but some adverbs found in Khotanese are attested, e.g. *mara* ‘here’ as in Khotanese, and *kari* ‘at all’, Khot. *karä*.

3.1.6a Numerals—Khotanese

3.1.6a.1 Cardinal numbers

The cardinal numbers in Old Khotanese are as follows (cited in the nom.-acc.):

TABLE 7.10: CARDINAL NUMBERS—KHOTANESE

1–9	11–19	1–9 +	10–90	100 +	
<i>śśau</i>	<i>śśīndasu</i>	<i>śśīvare-</i>	<i>dasau</i>	<i>satä</i>	‘hundred’
<i>d(u)va</i>	<i>d(u)vāsu</i>	<i>dvāvare-</i>	<i>bistä</i>	<i>ysāru</i>	‘thousand’
<i>drvai(a)</i>	<i>drvaisu</i>	<i>drvaivare-</i>	<i>därsi</i>	<i>byūrru</i>	‘ten thousand’
<i>tcahora</i>	<i>tcahulasu</i>	<i>tcahaure-</i>	<i>tcaholsä</i>	<i>kūla-</i>	‘ten million’
<i>pañjsa</i>	<i>pañjsūsu</i>	<i>puspare-</i>	<i>pañjsāsä</i>	<i>nayuta-</i>	10 ¹¹
<i>kṣāta</i> , <i>kṣei</i>	<i>kṣasu</i>	<i>kṣei</i> ‘vare-	<i>kṣaṣṭi</i>		
<i>hoda</i>	<i>haudūsu</i>	<i>hopare-</i>	<i>hodātä</i>		
<i>haṣṭa</i>	<i>haṣṭūsu</i>	<i>haṣpare-</i>	<i>haṣṭātä</i>		
<i>no, nau</i>	<i>nausu</i>	<i>novare-</i>	<i>notä</i>		

Numbers between the decads above twenty are compounds of the unit + **parah* ‘beyond’ + decade, e.g. *śśīvare-bistä* ‘21’, *puspare-bästi* ‘25’, *hopare-därsä* ‘37’. Counters of higher numbers precede, e.g. *dvī satī* ‘two hundred’, *haṣṭä ysāru* ‘eight thousand’, *tcahorehaṣṭātä ysāru* ‘eighty-four thousand’.

Numbers are inflected, except *byūrru* ‘ten thousand’. The number ‘one’ distinguishes masculine and feminine; both ‘one’ and ‘two’ have some special forms with neuter nouns.

3.1.6a.2 Ordinal numbers

Ordinal numbers, except for the first five, are regularly marked by the suffix *-ama-*, with the final *-u* of teens > *-ama-*, thus ‘first’ through ‘ninth’: *paḍauysa-*, *śāta-*, *dä(d)a-*, *tsūrāma-*, *pūha-*; *kṣei*‘ma-, *hodama-*, *haṣṭama-*, *nauma-*.

3.1.6a.3 ‘Both’ and indefinite adjectives

Khotanese *hūduva* ‘both’, etymologically originally ‘both two’, is declined like *duva* ‘two’. Indefinite adjectives include *phar-u*, *pharāka-* ‘many’, *biśśa-* ‘all’, *harbiśśa-* ‘every, all’, *hama-* ‘all’, *pana-* ‘every’, *handara-* ‘another’.

3.1.6b Numerals—Tumshuqese

Tumshuqese numbers are comparable to those in Khotanese: *šo*, *dva*, *dre*, *tsahari*, *paṛṛtsi*, *xši*, **hoda* (ordinal *hodama-*), *hašti*, *dase*; only a few decads are attested: *bista* ‘20’, *patsasu* ‘50’, and *xšišta* ‘60’. However, there is a striking difference from Khotanese with regard to the formation of the compound numerals in that the **parah* type is used for the numbers below but not above twenty (Maggi 1991); thus Tumsh. *ho-parsa-* ‘17’ vs. Khot. *haudūsu*, Tumsh. *bistyō patsyō* ‘25’ vs. Khot. *puspare-bāstā* showing inversion of digit and decad. Attested ordinals are: *tsārma-* ‘fourth’, *xšima-* ‘sixth’, *hodama-* ‘seventh’, *haštama-* ‘eighth’, *dasma-* ‘tenth’.

3.2 Verb morphology

3.2.1 Stem formation—Khotanese

3.2.1.1 Present and past stems

Verbs have two stems, present and past. They derive from the Old Iranian present stems and the past participle in **-ta*, respectively, but the original relationship is often no longer transparent: present/past stems *yan-/yāda-* ‘make, do’ (<Indo-Iranian **kṛnaŷ-*, **kṛtā-*). Some verbs have suppletive stems, e.g. *sarb-/sata-* ‘rise’, and *hīs-/āta-* ‘come’, *ah-/vatā-* ‘be’.

Many verbs distinguish intransitive and transitive stems, but have the same past participle which is then inflected according to transitivity: present *panam-/panem-* ‘rise’/‘raise’, past intr. 2s perf. *panat-ī* ‘you arose’ vs. trans. 3s perf. *panat-āndā* ‘they raised’ (same past participle *panata-*).

The two present stems relate to one another in various ways:

- (1) no formal distinction, transitivity only indicated by endings: *bar-/bar-* (act./mid.) ‘ride’/‘carry’;
- (2) length: *hvañ-/lvāñ-* ‘be spoken’/‘speak’; *hanaśś-/lhanāśś-* ‘go astray’/‘destroy’; *hamāh-/lhamīh-* ‘change’ (intr.)/‘change’ (trans.);
- (3) intransitive *-s* (‘inchoative’ present stem of Old Iranian) vs. otherwise marked transitive stem: *haspās-/haspīs-* ‘strive’/‘urge’;
- (4) palatalisation of transitive stem: *pasūjs-/pasūj-* ‘burn’ (intr.)/‘light (a lamp)’ (<**apa-saučā-* and **apa-saučāya-*);
- (5) obscured: *hatcy-/lhatcañ-* ‘be broken’/‘break’ (trans.) (<**fra-ščādaya-* and **fra-ščādaya-*);
- (6) causative marker *-āñ*: *bam-/lbam-āñ-* ‘vomit’/‘make vomit’; *ysai-/lysy-āñ-* ‘be born’/‘cause to be born’.

3.2.1.2 Directional particles

The particles *vā* and *hā* are commonly associated with verbs of motion and occasionally with other verbs. The particle *vā* expresses motion towards, *hā* motion away from the speaker:

[*vā*] *usahy-a ku-ṛṇ bis-a balys-a*
‘deign (come) [here] where my house (is), Buddha’ (Z 2.58);

ku [hā] tsut-e hām-āte
‘when he has gone [away]’.

The directional particle *ttā* is used in Middle and Late Khotanese to express motion towards the addressee, giving a system of triple deixis similar to the directional demonstratives:

to speaker, *vā*; to addressee, *ttā*; away from speaker, *hā*.

3.2.2 Nominal forms

3.2.2.1 Present participles

There are two:

- (1) Most common is that formed with the suffix masc. *-anda-*, fem. *-aṃkyā*, *-aṃcā* (historically, thematicised **-ant-*, fem. **-anta-čī-*) found with both active and middle stems: masc. *hūs-anda-* ‘sleeping’, masc. *hvar-anda-* ‘eating’; fem. *bārūñ-aṃcā-* ‘shining’. The feminine form was often used also for the masculine in the later language.
- (2) The participle in *-āna-* remains confined to the middle stems: *śś-āna-* ‘lying (down)’, *hām-āna-* ‘becoming’. The verbs *āh-* ‘to sit’ and *śś-* ‘lie’ have reinforced forms *āṇa-* ~ *āṇ-aṃda-*, *śś-āna-* ~ *ś-ān-aṃda-*.

Both types of present participle may optionally be extended by a secondary *-a-*: *hūs-anda-* ~ *hūs-and-aa-* ‘sleeping’, *āṇa-* ~ *āṇ-aa-* ‘sitting’.

In Tumshuqese, only the present participle in *-anda-* is attested: *dāy-anda-* ‘seeing’, *ras-andā* ‘knowing’.

3.2.2.2 Past participles

Past participles are based on the past stem, masc. *-ta-*, fem. *-tā-* (see section 3.2.1 Stem formation): *mār-* > *muḍa-* ‘die’; *huṣ-* > *huṣta-* ‘become dry’. There was however a tendency to add *-ta-* directly to the unmodified present stem: *murr-* > *murr-da-* ‘crush’, *paysān-* > *paysān-da-* ‘recognise’, and in some cases *-āta-* or *-āta-*: *huṣṣ-* > *huṣṣ-āta-*, *huṣṣ-āta-* ‘grow’. Occasionally there are extended forms: pres. *mār-* ‘die’ > *muḍa-* > *muḍa-a-*.

Tumshuqese shows many striking agreements with Khotanese: *āta-* ‘come’ (< **āgata-*), *vasuta-* ‘pure’ (< **awa-suxta-*), *hvata-* ‘spoken’.

3.2.2.3 Infinitives

Khotanese has two productive types of infinitive based on the present and past stems respectively: (1) present stem + *-ā* or *-i*, and (2) past stem + *-e*, whose palatalising effect is increasingly lost in the later language. Some verbs have both types: *hvāñ-* ‘speak’ > present infinitive *hvāñ-i* and > past infinitive *hvūt-e*, *hvīy-e*, but there appears to be no functional difference. Thus, both *hvāñ-i* and *hvīy-e* are used with *hamjsaṣ-* ‘be about to’ (Z 2.64 5.7). In a few instances the infinitive is substantivised and declined in the oblique cases as an *a-* stem (see section 3.2.7.3 Necessity construction).

Tumshuqese

In Tumshuqese the predominant type of infinitive is marked by *-ana-* added to the present stem (once found in Khotanese). The loc. sg. of such infinitives with the post-

positon *ā* as a whole came to be treated as a feminine noun (Emmerick 1985a: 14), e.g. instr.-abl. sg. *tsatānāy-ā tsi* ‘from killing’.

3.2.2.4 Gerundives and gerunds

The gerundive or participle of necessity is formed (1) by the suffix *-āñā-* (new formation) added to the present stem: *yan-āñā-* ‘be done’, *hvāñ-āñā-* ‘be spoken’; or (2) by *-ya-* suffixed to the lengthened verbal root. The latter is unproductive and often stands beside the new formation, e.g. *yan-āñā-* ‘be done’ beside *tcera-* (< **čārya-* from **kar-*); *hvāñ-āñā-* ‘be spoken’ beside *hvañā-* (< **hwanya-* from **hwan-*). The *ya-*gerundive may optionally be extended by secondary *-a-*, e.g. *tcera-* besides *tcera-a-* (< **čārya-ka-*), an extension found only once with *-āñā-*.

That Tumshuqese had a corresponding *āñā-*gerund is evidenced by *pāt vāñya*, probably < **pati-tav-āñā-* (in the Karmavācanā, Emmerick 1985a: 20), while the *ya-*gerund may be reflected in *ñyesyo* (Konow 1935: 27), but the context is obscure.

3.2.3 Person marking and ‘to be’

Person is marked by the forms of the copula and verbal endings. Forms based on the present stem distinguish active and middle voice; forms based on the past stem distinguish intransitive and transitive inflection, and distinguish masculine and feminine gender in some persons. Also, 3s neuter *-u* is often found in idiomatic expressions: *ttai* (< *tta + -ī*) *hāmāt-u* ‘thus it occurred to him’, *tta dist-u* ‘it appeared thus’, *tta-mā pyūšt-u* ‘thus I have heard’ (cf. Sanskrit *evaṃ mayā śrutam*).

Personal endings are distinguished in the four moods: indicative, subjunctive, optative, and imperative; a few instances of the injunctive are found in Old Khotanese. Tumshuqese has an imperfect form marked by an augment as in Old Iranian, which is not attested in Khotanese.

The verb ‘be’ is suppletive, its past stem *vāta-* deriving from Old Iranian **būta-*. The singular of the indicative present has enclitic forms.

The following is an overview of the person marking in Khotanese, combining the attested personal endings and ‘be, become’; for periphrastic verb forms see section 3.2.6.2 (note that raised *i* indicates palatalisation of the preceding syllable; “*ā~i*” = “*ā* or *i*”).

The archaic optative forms in *v-* of the copula (originally aorist optative) occur in potential and irreal conditional sentences (Skjærvø 1981; see section 3.2.6.2 for pluperfect and section 5.3.2 for conditional clauses).

In Tumshuqese only few non-indicative forms are attested, which correspond closely to their Khotanese counterparts.

TABLE 7.11: PRESENT AND PAST PERSON MARKING—KHOTANESE

Active. Present						
	1s	2s	3s	1p	2p	3p
IND	-ī mā	-i	-itā~i	-āmā	-(i)ta	-īndā~i
SBJ	-ī-ñi	-e	-ātā	-ā-ma(ne)	-ā-ta	-ā-ru~o
OPT	-ā, - ⁱ o-, ⁱ u, -ī-ne		-ī-yā, - ⁱ a, -ā		-ī-ru	-ī-ru~o
INJ	-u	-a	-(ā)ta			
IMP			-(ā)tu		-(ā)ta	-ā-ndu

(Continued overleaf)

TABLE 7.11: Continued

Middle. Present						
	1s	2s	3s	1p	2p	3p
IND	-e	-a	-(i)te	-ā-mane	-īru	-āre
SBJ	-ā(-ne)	-ā	-ā-te	-ā-mane	-ā-ta	-ā-ru~o, -ā-nde
OPT	-ī-ne		-ī-ya			-ī-ru~o
INJ			-(ā)ta			
IMP		-u-	-ūtō			
Past Indicative						
	1s	2s	3s	1p	2p	3p
INTR	-t-ā mā	-t-ī	-t-ā-θ (m) -t-a-θ (f)	-t-a mā	-t-a sta (m) -t-e sta (f)	-t-a-θ (m) -t-e-θ (f)
TRNS	-t-ai mā (m) -t-ā mā (f)	-t-ai	-t-e-θ (m) -t-ā tā (f) -t-u-θ (nt)	-t-ānd-ā mī	-t-ānd-ī sta	-t-ānd-ī

TABLE 7.12: COPULA AND 'BE, BECOME'—KHOTANESE

<i>ah-</i> 'be'. Present						
	1s	2s	3s	1p	2p	3p
IND	īmā, -mā	ī, -e	astī, -stā	mā	sta	īndā
SBJ		ā-ye	ā-ya	ā-maṃ		ā-ro, ā-mdu
OPT			ī-yā, ī-tā, ī			
IMP					ā-ndu	
<i>vāt-</i> 'be, become'. Past						
	1s	2s	3s	1p	2p	3p
IND	vāt-ā mā	vāt-ī	m. vāt-ā f. vāt-a	vāt-a mā	m. vāt-a sta f. vāt-e sta	m. vāt-a f. vāt-e
SBJ			vāt-ā-ya			
OPT	v-y-o		v-ī-ya, (v)ya			v-ī-ro

3.2.4 Conjugation classes

Four main conjugation classes can be distinguished based on the endings of the 3 pres. indicative. Of these types C and D are rare (Emmerick 1968a: 177). The table contains attested forms of *hvāñ-* 'speak', *yan-* 'make, do', *puva'd-* 'fear', *dai-* 'see', *hām-* 'become, occur', and *ysai-* 'be born'.

TABLE 7.13: CONJUGATION CLASSES—KHOTANESE

Active	A	B	C	D
	'speak'	'make, do'	'fear'	'see'
1s	<i>hvāñ-īmā</i>	<i>yan-īmā</i>	<i>puvai'-mā</i>	<i>dai-mā</i>
2s	<i>hvāñ-i</i>	<i>yañ-i</i>		<i>dai</i>
3s	<i>hvāñ-ātā</i>	<i>yīn-di</i>	<i>pvai't-tā</i>	<i>dai-yā</i>
1p	<i>hvāñ-āmā</i>	<i>yan-āmā</i>		<i>dāy-āmā</i>
2p	<i>hvāñ-ita</i>	<i>yan-da</i>	<i>puva't-ta</i>	<i>dai-ya</i>
3p	<i>hvāñ-īndi</i>	<i>yan-īndi</i>	<i>puvai'-ndi</i>	<i>dai-ndā</i>

Continued overleaf

However, there was a tendency for the transitive to predominate (see Sims-Williams 1997 on the origin of the perfect).

3.2.6.2 Periphrastic perfect and pluperfect

The periphrastic past forms are based on the past participle and forms of 'be'.

In the periphrastic indicative perfect in the 3s and 3p only forms with the negative are found:

ne hvat-e štä
'he did not speak'.

The perfect subjunctive and optative add the respective forms of the copula.

The pluperfect adds the past forms of 'be', of which the optative is extremely rare (Skjærvø 1981). The following table illustrates the pattern, citing some of the rare non-indicative forms.

TABLE 7.14: PERIPHRASTIC PERFECT AND PLUPERFECT—KHOTANESE

Perfect	Pluperfect				
IND <i>yäḍ-e sta</i>	'you (pl.) have done'	<i>yäḍ-e vät-ä</i>	'he had done'		
SBJ <i>yäḍ-e äye</i>	'you may have done'	<i>autt-ä vät-äya</i>	'it would have lasted'		
OPT <i>yäḍ-e iyä</i>	'he may, might have done'	<i>ysät-ä vya</i>	'he would/might have been born'		

Note that functionally there is no discernible difference between the non-periphrastic simple past and the periphrastic perfect. Essentially, then, the perfect has two forms:

parst-e mā 'I (have) ordered';
parst-e mā imā 'I (have) ordered'.

There is insufficient evidence to establish the perfect system in Tumshuqese.

TABLE 7.15: OVERVIEW OF TENSES AND MOODS—KHOTANESE

Present	indicative	subjunctive	optative	imperative	(injunctive)
Perfect, simple	indicative				
Perfect, periphr.	indicative	subjunctive	optative		
Pluperfect	indicative	subjunctive	optative		

3.2.7 Modal and aspectual constructions

3.2.7.1 Potential-completive construction and passive

The auxiliaries *yan-* 'make, do' and *häm-* 'become' when used with past participles have the following functions: *yan-* + past participle expresses active ability or completed action, *häm-* + past participle expresses passive ability or completed action or simple passive in the present. The past passive is expressed by 'to be'.

3.2.7.1a Ability

(a) Intransitive

ṣā vyāgr-a ne panat-a hām-āte

‘the tigress [cannot get up]’ (Suv 18.99).

(b) Transitive active

Present

ye ttuśśā-ttet-u [ne yuḍ-u yīn-dā nājsaṣṭ-u]

‘[one cannot demonstrate] (the doctrine of) emptiness’;

ku [ne] ne [yuḍ-u yan-de] samu

‘if [he] simply [cannot do] them’ (i.e., miraculous deeds) (Z 12.132), note that in the expression *riddhā yan-* ‘perform miraculous deeds’ *yan-* is middle.

(c) Transitive passive

ūc-a biśś-a [khaṣṭ-a hām-āte]

‘the water [can] all [be drunk (up)]’ (Z 2.120).

3.2.7.1b Completed action

(a) Intransitive

Present

cīyā ṣṣaman-a [āt-a hām-āre]

‘when the monks [have come]’ (Z 24.273).

(b) Transitive active

Past

cī aysur-a [purrd-u yād-ānda]

‘when [you-p had defeated] the asuras’;

cī gyast-ā balys-ā ttut-u tta tta [hvat-u yād-e]

‘when the Lord Buddha [had said] this thus’ (Suv 6.3.72).

(b) Transitive passive

Past

cīyā ṣātā tta tta [hvat-ā hāmāt-ā]

‘when this [had been said] thus’ (Suv 10.67).

3.2.7.1c Passive

Present(-future)

īmu mamā māst-ā rakṣ-a [yād-a hām-āte]

‘today by me great protection [will be made = given]’ (Suv 6.3.62), translating Sanskrit *adya mama mahat y ārakṣā kṛtā bhaviṣyati*.

Past

ce buro . . . hūñe jsa [bvata īndā]

‘as many . . . army-by [destroyed are]’ = ‘as many (lands) have been destroyed by a (foreign) army’ (6.1.24), translating Sanskrit *ye kecid . . . para-cakreṇa vā upahatā bhaviṣyanti*;

[*kāilsta-i vāt-a*] *ššārye hāirā bāt-a*

['planted-by-him was] of-good thing root' = 'by him had been planted the root of a good thing' (Suv 2.2).

3.2.7.3 Infinitive-based constructions

See Emmerick 1968a: 218–220 and Degener 1987: 284–293.

The infinitive depends on verbs predicating inception or completion, such as *hamjsaš-* 'be about to', *ākšuv-* 'begin', *byeh-* 'succeed', *parī-* 'order', or ability: *hot-* 'be able':

nāstā kye bišš-o bašda-u haut-ta vāstarna [*hvīy-ā*]

'there is no one who could-INJ [tell] the whole evil in detail' (Z 24.437);

svī ākšuv-āmā [*padīnd-e*]

'tomorrow we will begin [to make] (it)' (Z 23.43);

uysnor-a kye . . . dāt-u byeh-īndā [*pyūšt-e*]

'beings who get [to hear] the Law' (Sgh § 34.14).

3.2.7.4 Gerundives, gerunds, and necessity construction

See Emmerick 1968a: 171–176, 214–218 and Degener 1989a: 54–71, 218–266, 299–304.

The gerundive or participle of necessity, formed by the suffix *-āñ-* or by *-ya-* (see section 3.2.2 Nominal forms) expresses necessity.

Negation of both types of gerundives is by *ne*, which in the 3s may be followed by the enclitic copula, or less commonly by the privative prefix *a-*.

The neuter singular of the gerundive is used as a gerund, constructed with the gen.-dat. of the experiencing subject:

ka [*mamā*] *nary-o . . . [št-āñ-u hām-āte]*

'even if [I should have to stand = remain] in hell' (Z 2.215).

3.2.8 Diathesis

Khotanese inherited both causative and medio-passive stems (see section 3.2.1 Stem formation).

The passive can be expressed by the medio-passive endings with transitive verbs.

The middle can also have a reflexive function (Canevascini 1991). Thus, *jsan-* 'kill' is normally active as in *handar-u jsīn-dā* 'he kills another', but it is middle in *uysān-o jsan-de* 'he kills himself'.

The auxiliary verb *yan-* 'make, do' may be conjugated either in the active or in the middle, but each phrase has its particular diathesis. Thus in the phrase *kīr-e yan-* 'do (good) acts' *yan-* is always active whereas in the phrase *kādägān-e yan-* 'do evil acts' *yan-* 'do evil acts' it is always middle.

The passive can also be expressed by the potential-completive construction of *hām-* (see section 3.2.7.1c).

4 SYNTAX AND USE OF FORMS

See Canevascini 1991; Degener 1993; Emmerick 1965, 1987a; Emmerick and Skjærvø 1982: 32, 97; Herzenberg 1965: 117–137; Heston 1976; Skjærvø 1981: 461–463.

4.1 Coordination

4.1.1 Coordinating conjunctions and clitics

Coordinate conjunctions in Khotanese are: non-enclitic *u* ‘and’ and *o* ‘or’; enclitic *r(r)o* ‘and’, *vā* ‘or’, *hāḍe* ‘but’. Of these, Tumshuqese appears to have *o* ‘or’ and *ro* ‘and’. Extended forms are *o vā* ‘or on the other hand’, and alternative *o . . . o* ‘either . . . or’.

4.1.2 Enclitic particles and words

Khotanese has a common deictic particle *-ī* (see section 3.1.3a/b.1 Demonstratives), and a rare emphatic particle *-ū* (Canevascini and Emmerick in Emmerick and Skjærvø 1997: 25–28). Enclitic words include indefinite *ju*, *buro*, *hānu*, *halci*, and emphatic *gāvu*, *tāvu* after negation in Old Khotanese. In Late Khotanese the present participles of *jsā-* ‘go’, *āh-* ‘sit’, and *ṣṭ-* ‘stand’ have various functions which remain to be studied in detail.

4.2 Noun phrase structure

Khotanese and Tumshuqese are languages where dependent items precede the head. Thus the structure of the simple noun phrase as well as that of the simple genitival noun phrase are (see also 4.3.3 Enclitic pronouns): ADJECTIVE – NOUN and NP₂ – NP₁.

4.2.1 Concord and agreement

4.2.1.1 Adjectives

Adjectives agree in case, number, and gender with the nouns they qualify. For example,

āṣṣeī' ṅ-ā häten-ai śśīy-ā ham-ā hamau
‘blue, red, white (is) the same goblet’ (Z 4.33).

Here the four adjectives agree with the nom. sg. masc. of *hamau* ‘goblet’.

In later stages of the language the concord is often confused, being obscured by the reduction of the endings:

bu'ys-ye bāḍ-na
‘for a long time’ (JS 7r2, 20r2).

Here the ending *-ye* of the adjective is formally instr.-abl. sg. fem., and the ending *-na* of the noun is instr.-abl. sg. masc.

4.2.1.2 Group inflection

In the later language the final syllable *-na* of the instr.-abl. sg. masc. of an adjective is sometimes dropped as in *mirāhīnai hārna* ‘with a pearl necklace’ for **mirāhīnaina hārna*.

This phenomenon reflects the shift towards group inflection, and may have spread from the numbers, where already in Old Khotanese the members of complex numbers are either all inflected, or only the last one.

4.2.1.3 Demonstrative adjectives

Demonstrative adjectives similarly agree with their head noun:

balys-ä [ttü] hvana-u naljsond-e

‘the Buddha concluded [this] speech.’ (Z 22.334), accusative;

[ttät-äye] hvat-ye hvan-ai

‘[this] speech having been spoken’, genitive (absolute construction).

4.2.2 Dependent nouns

Dependent nouns usually precede the head noun directly. The possessive function of the gen.-dat. may be emphasised by the adjective *häv-ia-*, *häv-ia-* ‘belonging to’ which agrees in case, number, and gender with the head noun:

balys-ä (gen.) [häv-y-e] ird-i lit.

‘the Buddha’s [own] miraculous powers’ = ‘the Buddha’s miraculous powers’
(Z 14.91),

instead of **balys-ä-GEN ird-i*;

šakrr-ä [häv-ī] bārai

‘Śakra’s [own] steed’ = ‘Śakra’s steed’ (*Jātakastava* 6v3, ed. Dresden).

This usage is rarely attested in Old Khotanese but becomes more prominent later. Titles etc. precede names, e.g. *gyast-ä balys-ä* ‘the Lord Buddha’.

4.2.3 Extended noun phrase

The order of the extended noun phrase has not been fully determined but appears to be as follows:

Demonstrative – *handara* (‘other’) – Number – Adjective – Noun,

where the number may itself be preceded by indefinite adjectives such as *phar-u*, *pharāka-* ‘many’, *bišša-*, *harbišša-* ‘all’.

4.3 Clause structure and word order

See Degener 1993: 46–51.

4.3.1 Nominal sentence

Nominal sentences, which consist of a subject and a nominal predicate, omit the copula in the present and past if both positive and indicative:

rre udayan-ä sīr-ä *thu-nä päte*
 ‘King Udayana (was) joyful’ (Z 23.123); ‘you (are) our father’ (Z 23.114).

4.3.2 Order of constituents

In view of the fact that there is relatively little prose literature available that is not translated from another language, it is difficult to determine in detail the principles governing Khotanese word order.

Khotanese and Tumshuqese are SOV (subject-object-verb) languages: Khot.

balys-ä dāt-u hvat-e
 ‘the Buddha spoke = preached the Law’ (Z 13.109).

An indirect object precedes a direct object:

tt-ānu cakrravartt-ānu rrund-inu-GEN-DAT haur-u-ACC haur-ä
 ‘he would give a gift (*haur-u*) to the cakravartin kings’ (Sgh § 88.3).

4.3.3 Enclitic pronouns

The enclitic pronouns in their genitive and dative functions as possessor, experiencer, or agent are attached to the first unit of the clause. This may be the following:

(1) Adverbial:

tta-mä pyūṣṭ-u
 ‘thus-by-me was heard’ = ‘thus I have heard’.

(2) Conjunction:

<i>tta cu-te sait-tä</i> ‘well, how-to-you does it seem?’ (Sgh § 91.1);	<i>cv-ī mäst-a hastam-a kīr-e</i> ‘as-for-his great, excellent deeds’ (Z 23.47).
---	--

(3) Relative-interrogative:

<i>kye-nä rro grat-u hvāñ-äte</i> ‘and who-to-us will preach instruction?’ (Z 24.516);	<i>ku-m bis-a</i> ‘where-my-to-us (is)’ (Z 2.58).
--	---

(4) Subject itself:

thu-nä päte
 ‘you (are) our father’ (Z 23.114).

It is always attached to the negation:

nai (< ne + -ī) ne mara ram-äte
 ‘it pleases him (-ī) not at all (*ne . . . ne*) here’ (Z 22.172);

ne-n ne mara ram-äte

‘it pleases them (-n) not at all here’ (Z 22.172; 22.211);

cu nai (< ne + -ī) bei’ khās-a yan-āmā

‘why don’t we put poison in his (-ī) food?’ (Z 2.26).

The last example shows that by this rule the enclitic pronoun may be separated from its referent, lit. ‘why not-his poison in food we do’.

4.3.4 Topicalisation

It is probable that, as in most languages, focus could be given to a word by deviation from the standard word order, but further research is required to determine the details. Moreover, in texts which are mostly translations, topicalisation may follow the original. For example, in the following sentence the word order closely corresponds to Sanskrit (Sgh § 62.2, cited by Degener 1989: 49):

<i>namasāt-aimā . . . aysu</i>	<i>ttā . . .</i>	<i>balys-a</i> (Khot.)
‘worshipped . . . I	the	Buddhas’
<i>vanditās . . . me</i>	<i>te</i>	<i>tathāgatā</i> (Sanskrit)
‘worshipped . . . by me	the	Buddhas’

A common form of topicalisation is the placing of a dependent noun in initial position, marked by the conjunction *cu* ‘as’, with or without an anaphoric replacement:

cu kuṃṣa-a ṣt-e, tt-ye ysv-e

‘as far as (*cu*) sesame (*kuṃṣa*) is concerned (*ṣt-e*), its (*tt-ye*) taste (*ysv-e*)’ (Si § 3.18), from non-topicalised ‘the taste of sesame’;

cv-ī māst-a hastam-a kīr-e, biśś-ā balys-ā āṇi nijsaṣt-e

‘as for (*cv-*) his great, excellent deeds, all (of them) the Buddha displayed (while) sitting’ (Z 23.47).

4.3.5 Questions

4.3.5.1 Yes-no questions have the same word order as statements, their interrogative character being indicated by intonation:

(1) *dahak-a ysq ajīśk-ā*

‘has a boy been born or a girl?’ (Nanda 41).

(2) With reaffirmative tag, *ā ne* ‘or not’, later *o ne*:

dv-a ttaiṣy-a caiga-kṣīr-a tsvāṃd-a, ā ne

‘did the two monks go to China or not?’ (P 5538b.43 KT 3.122).

4.3.5.2 Interrogatives are as a rule in initial position:

(1) *cūḍe brem-ā*

‘why are you (sg.) weeping?’ (Z 24.235);

kve-nā rro grat-u hvāñ-äte

‘and who will preach instruction to us?’
(Z 24.516).

- (2) Optionally preceded by a discourse initiator, such as: *tta* ‘thus, so’:

tta cu-te sait-tä
‘well, how does it seem to you?’ (Sgh § 91.1).

4.4 Semantics and use of nominal forms

4.4.1 Use of cases

On the whole the syntax of the cases conforms with that of Avestan and Indo-Iranian. Particular features are the following.

4.1.1.1 Genitive-dative

- (1) As object and experiencer with verbs of fearing:

[ššūjī-ye] puvaī’-ndi
‘they fear [one another]’ (Z 11.10);

and as experiencer with such as verbs as ‘occur to, seem to someone, realise’.

- (2) With the participle of necessity:

ka [mamä] nar-yo . . . šāñ-u häm-äte
‘if it were [for me] to stay (*šāñ-u*) in hell’ = ‘even if I should have to remain in hell’
(Z 2.215);

kšāta’ pārāmat-e [hvānd-ä] . . . car-āñ-e
‘the six perfections must be practised [by a man]’ (Z 11.17).

- (3) As agent in future passive constructions:

īmu [mamä] māst-ä rakṣ-a yād-a häm-äte
‘today [by me] great protection will be given’ (Suv 6.3.62).

- (4) In past passive participle constructions:

[bišš-ānu gyast-ānu balys-ānu] āysdagād-ä
‘watched over [by all the Lord Buddhas]’ (Suv 6.1.4).

- (5) In a genitive absolute construction:

[ttät-äye hvat-ye hvan-ai]
[this spoken speech]-GEN’ = ‘this speech having been spoken’ (Sgh § 43.1).

4.4.1.2 Locative

The use of the locative is shown in the following.

- (1) For the source of motion:

kho ju ye viys-u thañj-äte [khārj-a]
‘just as one pulls a lotus out [of the mud]’ (Z 5.90).

- (2) With the adjective *bisaa-* ‘residing, located (in)’:

[*jaṃbvīy-a*] *bis-ā satv-a*
‘beings (living) [in Jambudvīpa]’ (ApS § 6);

[*tcimñ-a*] *bis-e arv-e*
‘in the eye located (medicinal plants =) medicine (that one puts) in the eye’ =
‘eye-medicine’ (Si 26.31).

- (3) In Late Khotanese, with the present participle middle *āna-* ‘sitting, dwelling’ to express motion away from the noun, instead of a simple ablative noun:

[*hvaṃny-a kṣīr-ā*] *qn-ā ṣacūvāṣṭā*
‘[in Khotanese country] dwelling to Ṣacū’ = ‘from the Khotanese country to Ṣacū’
(StH 36).

4.5 Semantics and use of verb forms

4.5.1 Possession

Possession is indicated by the genitive-dative with the existential verb, e.g. here with enclitics, *aštā-tā* ‘you have’, lit. ‘to you is’, *ništā-mā* ‘I don’t have’.

4.5.2 Tense and aspect

The basic distinction of tense is present and past. In turn, past tense distinguishes between perfect (simple and periphrastic) and pluperfect.

4.5.2.1 Future

The future is normally expressed by means of the present or the subjunctive.

4.5.2.2 Accomplishment of past action

The accomplishment of an action in the past is expressed by the potential-completive *yan-*, *hām-* construction (see section 3.2.7.1 Potential construction).

4.5.2.3 Immediacy

Immediacy is frequently expressed in Late Khotanese by adding the present participle middle *ṣṭāna-* of *ṣṭ-* ‘stand’, often together with *vaṃñā* ‘now’:

ttrāst-ai ra thu paḍā . . . u [ttrāy-i ṣṭāna] vaṃñā ṣṭāna is a participle by then, no ending ‘you formerly rescued . . . and [you (sg.) are rescuing] now’ (Avdh 12v1–2) (comparable to Buddhist Sogdian *əštən* and Yaghnobi *-išt*).

4.5.3 Use of modal forms

See Degener 1993: 51–55.

On the whole, one can distinguish between what may be termed ‘major modal’ (indicative and imperative) and ‘minor modal’ (subjunctive, optative, and Old Khot. 3s

injunctive), forms. In their usage the moods are not always clearly distinguished, the inherited system being in the process of disintegration, and there is the tendency to use only the indicative.

In general, the indicative is used for statements; the imperative for commands; the subjunctive for potential action or state, including softened commands and wishes; the optative for irreal action or state, including wishes; the injunctive for potential as well as past action or state.

Examples of the use of the moods are the following (for hypothetical statements see section 5.3.2 Conditional clauses below; for ability and necessity, see sections 3.2.7.1 Potential, and section 3.2.7.4 Gerundival constructions):

(1) Statements:

(a) Indicative:

balys-ä tt-ü hvana-u [naljsond-e]
‘the Buddha [concluded] this speech.’ (Z 22.334).

(b) Injunctive, past function:

samu ne rre pātcu [haraysd-a]
‘then no sooner [had] the king [stretched out].’ (Z 5.88).

(c) Commands, imperative:

ma ju [puva’] badr-a [ākūt-a] ggamdy-o tsāṣṭ-o
‘[fear] not, Bhadra! [Strike] the gong calmly.’ (Z 2.102).

prām-a [škīm-āṃdu]
‘[let them build] arches.’ (Z 5.96).

(2) Wishes:

(a) Subjunctive:

ka va tt-ä [pars-āro] dukh-yau jsa
‘if only [they may escape] from woes!’ (Z 5.112).

(b) Optative:

ka nā [parrīj-īyā] dukh-yau jsa
‘if (only) [he may rescue] them from woes!’ (Z 22.292).

(c) Indicative:

ma vā [pars-īmā] puṣṣo
‘nor [may I escape] wholly’ (Z 7.4).

5 COMPLEX SENTENCES

Subordinate conjunctions in Khotanese are: *ka* ‘if’, *kāmu* ‘as long as’, *ku* ‘where’, *kho* ‘how, as’, *cīyā* ‘when’, *cu* ‘when, if, because, since’. Of these, *ka* and *cu* are found in Tumshuqese.

5.1 Quotative clitic

The particle *se* is a discourse marker introducing direct speech after verbs of speaking or thinking or similar locutions, corresponding to the colon in the English translations:

tta hvat-e [se] cu nai (ne + -ī) bei khās-a yan-āmā
 ‘he spoke thus [.] “Why don’t we put poison in his food?” ’ (Z 2.26).

5.2 Relative clauses

Relative clauses are introduced by relative pronouns or relative adverbs, with or without a preceding demonstrative or demonstrative noun phrase. The case of the pronoun is determined by its function within the relative clause. Potential or generalised action or state is in the subjunctive or injunctive.

5.2.1 Indicative, factual

ništā kye yuḍ-u yīn-dā pamūt-u . . . ṣṣahān-e
 ‘there is not = no one **who** can measure-POT the virtues’ (Z 24.652).

5.2.2 Relative pronoun in genitive-dative

samu kho tye ci māta mīḍe
 ‘just as **that one whose** mother dies’ (Z 23.24).

5.2.3 Injunctive, potential

nāštā kye biśś-o baśda-u haut-ta västarna hvīy-ä
 ‘there is no one **who** could tell the whole evil in detail’ (Z 24.437).

5.2.4 Clauses with head noun

kṣīr-ä ku ne rre nāštā o ṣ-a bis-a ku hvāšt-ä mīḍe
 ‘**the land where** there is no king, or **the house where** the master dies’ (Z 23.23);

ṣ-a ṣṣav-a . . . ku nā sat-a štā hambad-a māst-ä
 ‘**the night . . . when** the full moon has not risen’ (Z 23.22).

5.3 Adverbial clauses

5.3.1 Temporal and local clauses

Temporal and local clauses are introduced by *cī*, *cīyā*, or *ku*, locational clauses by *ku*:

5.3.1.1 Temporal

cī gyast-ä balys-ä ttut-u tta tta hvat-u yād-e
when the Lord Buddha had said-POT this’ (Suv 6.3.72);

cīyā ṣṣaman-a āt-a häm-āre
 ‘when the monks have come-POT’ (Z 24.273).

5.3.1.2 Local

Local clause, subjunctive dependent on possible condition:

ka . . . nai (ne + -ī) bv-āte (ku mara āt-ā häm-āte)
 ‘if he should not realise it: **where** he has come here’ (Z 2.51).

5.3.2 Conditional clauses

Conditional clauses are introduced by *ka* or *ko*. Potential action or state is in the subjunctive, irreal in the optative:

(1) Subjunctive, potential:

ka . . . nai (ne + -ī) bv-āte (ku mara āt-ā häm-āte)
 ‘if he should not realise (where he has come here)’ (Z 2.51).

(2) Optative in protasis and apodosis, irreal:

ko satv-ā ttatvatu v-īya (nirvān-ye gāvu ni hīs-a)
 ‘if indeed a being really existed (*v-īya*), (one would never come (*hīs-a*) to Nirvāṇa)’
 (Z 6.58).

5.3.3 Consecutive and comparative clauses

Consecutive and comparative clauses are preceded by demonstrative adverbs and introduced by *kho* and a correlate relative pronoun or relative adverb:

jaṃbvīv-i trām-ā västāt-ā ku ne marata balys-ā ni vei štā
 ‘Jambudvīpa had become **such as when** no Buddha has been here’ (Z 23.22);

pan-ye tterā kāṣc-a uysnor-ā samu kho tte ci māt-a mīḍe
 ‘the anxiety of every living being (was) **as great as that of one whose** mother dies’
 (Z 23.24);

kho ju ye viys-u thaṃj-āte khārj-a
 ‘just **as** one pulls a lotus out of the mud’ (Z 5.90).

5.4 Nominalised clauses

5.4.1 Infinitival constructions

In infinitival constructions (see section 3.2.7.3), the complement may either remain with the infinitive, or be placed before the head verb

(1) Position before the infinitive:

parst-e (tt-ā dvās-u baṃṃhy-a pāhašt-e)
 ‘he ordered (to fell those twelve trees)’ (Sgh § 211.2).

(2) Position before the head verb:

uysnor-a kye . . . (dāt-u) byeh-īndä (pyūšt-e)
 ‘beings who (the Law) get (to hear)’ = ‘beings who succeed in hearing the Law’ (Sgh § 34.14).

5.4.2 Participial-adjectival clauses

In addition to the infinitival and gerundival constructions (see sections 3.2.7.2–3), and the genitive absolute (see 3.1.3a/b.1 Demonstrative adjectives), Khotanese makes frequent use of participial-adjectival clauses. They are prominently, but not only, locational:

mari ā’gy-e dīvat-e, *hvaṇm-ya kṣī’r-ä qn-ä*
 ‘here **residing** deities’ (Z 23.25); ‘**dwelling** in Khotanese country’ (StH 36);

ttāj-vā [īs-ā] devatt-a lit.
 ‘in the rivers **dwelling** deities’ (ApS V § 25).

They tend to evolve into locational enclitics:

hvaṇm-ya kṣī’r-ä qn-ä śacūvāštā
 ‘from the Khotanese country to Śacū’ (StH 36), lit. ‘**dwelling** in’.

6 LEXIS AND SOCIOLINGUISTIC ASPECTS

6.1 Word formation

6.1.1 Derivation

There is a large number of productive suffixes, but few prefixes (for productive suffixes, see Degener 1989a). The following are examples of productive affixes that derive adjectives and nouns:

6.1.1.1 Suffixation

- (1) Adjectives from nouns: (a) *-īñā-*, from nouns denoting living beings: *dahīñā* ‘belonging to a man’ < *daha* ‘man’; *kaviñā* ‘belonging to a fish’ < *kavā* ‘fish’; *-īnaa-* (fem. *-īṃgyā-*), (b) from nouns not denoting living beings (with one exception): *ggūštīnaa-* ‘flesh’ < *ggūštā* ‘flesh’; *spātāinaa-* ‘flower’ < *spātaa-* ‘flower’; (c) *-aunda-*: *puñāunda-* ‘meritorious’ < *puñā* ‘merit’; *tcei’maunda-* ‘having (seeing) eyes’ < *tcei’man-* ‘eye’; (d) *-jsa-*: *pāta’jsa-* ‘mighty’ < *pāta-* ‘might’; *rrīmajsa-* ‘dirty’ < *rrīman-* ‘dirt’. (2) Nouns from adjectives: *-ia-*: *kṛtañia-* ‘gratitude’ < *kṛtaña-* ‘grateful’; *škālśia-* ‘pride’ < *škālśa-* ‘proud’.
- (3) Nouns from nouns: (a) *-oñā-*: *arahandoñā-* ‘arhatship’ < *arahanda-* ‘arhat’; *bvākauñā-* ‘the state of being a knower’ < *bvāka-* ‘knower’; (b) *-osti-*: *dahosti-* ‘virility’ < *daha-* ‘man’; *hayūnausti-* ‘friendship’ < *hayūna-* ‘friend’.

6.1.1.2 Prefixation

- (1) Adjectives from adjectives: (a) privative *a-* ‘un-’, *ahvata-* ‘unspoken’ < *hvata-* ‘spoken’; (b) *bi-* ‘less’, *bijūndaa-* ‘lifeless’ < *jūndaa-* ‘living’; (c) *hu-* ‘well’: *hubasta-* ‘well bound’ < *basta-* ‘bound’.

- (2) Adjectives from nouns: *a-*: *aggaṃjsa-* ‘flawless’ < *ggaṃjsā-* ‘flaw’; *duṣ-*: *duṣpāta-* ‘powerless’ < *pāta-* ‘power’.
- (3) Nouns from nouns: ‘non-’: *akā’matā-* ‘non-thought’ < *kā’matā-* ‘thought’; *ggu-* ‘un-’: *gguhamāṅgatāti-* ‘unevenness’ < *hamāṅgatāti-* ‘evenness’.

6.1.2 Composition

Khotanese compounds consist almost exclusively of two members (see Degener 1987).

- (1) Substantive as prior member:
- (a) + substantive: a. *bahuvrīhi*: *dāta-āhāra-* ‘having the Law as food’ (+ *āhāra-* ‘food’); b. determinative: *kamala-rrāha-* ‘headache’ (+ *rrāha-* ‘ache’); c. dvandva: *mārā-pātara* ‘parents’ (+ *pātara-*) ‘father’.
- (b) + adjective: *cā’ya-nārmāta-* ‘produced by magic’ (+ *nārmāta-* ‘produced’).
- (c) + verbal stem. This type usually has secondary *-a-* (see section 3.1.1a Nominal stem classes): *dāta-hvāṅāa-* ‘Law preacher’ (+ *hvāṅāa-* ‘preacher’ < *hvāṅ-* ‘to preach’).
- (2) Adjective (adverb, pronoun) as prior member:
- +substantive, a. *bahuvrīhi*: *āṣei’ṅa-rraḥoṅa-* ‘having dark blue garments’ (+ *rraḥoṅa-* ‘garment’), *ṣiṣṭa-biśā* ‘tongue-tied’ (+ *biśā-* ‘tongue’); b. *karmadhāraya*: *hvatāna-kṣīra-* ‘Khotan(ese country)’ (+ *kṣīra-* ‘country’).
- + adjective: *hamṭsa-ysāta-* ‘born together’ (+ *ysāta-* ‘born’).
- (3) Numbers as prior element:
- + substantive, a. mostly *bahuvrīhi*: *drrai-pāṣkala-* ‘of three classes’ (+ *pāṣkala-* ‘class’), *śi-vāa-* ‘biped’ (+ *pāa-* ‘foot’); b. not *bahuvrīhi*, e.g. *drai-padia-* ‘threefold’ (+ *padia-* ‘kind’).
- + number, two instances only: *hūduva* ‘both’ (+ *duva* ‘two’); *śśūjāta-* ‘one another’ (< *śāta-* ‘second’).

6.1.3a Loanwords—Khotanese

See Bailey 1946, 1947, 1961: 4; Degener, 1989b, 1990; Emmerick 1985b, 1989: 228–229.

6.1.3a.1 Iranian

The oldest loanwords in Khotanese are Zoroastrian technical terms that have been adapted to a different religious environment. The best known are *urmaysde* ‘sun’ < **ahura-mazdāh* (nom. sg.), cf. Avestan *ahurō mazdā*, Old Persian *auramazdā*, and *śśandrāmatā-*, which is the Khotanese equivalent of the name of the Avestan female Holy Immortal *spəntā- ārmaiti-* but is used to designate the Buddhist goddess of fortune, Sanskrit Śrī. To this layer of borrowing may also belong the much disputed word *phārra-* ‘splendour’, the Khotanese counterpart of Old Persian *farnah-* (Emmerick in Emmerick and Skjærvø 1997: 103–104).

6.1.3a.2 Indian

The majority of the Indian loanwords in Khotanese are learned borrowings that are adopted unchanged but are inflected as Khotanese, sometimes with Khotanese suffixes.

Often it is not possible to determine whether their source is Prakrit or Sanskrit, but many of the frequent words betray by their phonological features the influence of the Northwestern or Gāndhārī Prakrit. They are extremely common in a Buddhist milieu, e.g. *ṣṣamana-*, ‘monk’, *ṣṣāvaa-* ‘hearer, disciple’, *saṃtsāra-* ‘cycle of existence’, cf. Gāndhārī *ṣamaṇa-*, *ṣavaka-*, *satsara-*.

According to the degree of assimilation there are ‘loan synonyms’, ‘loan blends’, and ‘loan homonyms’ (Degener 1989b). Thus, a word of Iranian origin may add the meaning of a similar Indian word, e.g. *sam-*, whose meaning ‘to agree’ has been ascribed to the influence of Buddhist Sanskrit and Pāli *sam-*. A loanword may be blended, e.g. Indian *pāramitā-* ‘perfection’ > *pārāmatā-*, here with the common Khotanese abstract suffix *-āmatā-*. A loanword becomes homonymous with a native word, e.g. Indian *ātama-* in the sense of ‘Buddhist tradition’ (< Sanskrit *āgama-*) became homonymous with Khotanese *āmatama-* ‘wish, desire’ (<**ākāma-*).

Borrowings from other languages are rare. (1) *Tocharian*. The only instance seems to be Old Khot. *puka-* ‘cubit’ from Toch. A *poke* ‘arm’. (2) *Chinese*. Apart from proper names and titles borrowings, few administrative terms are attested; e.g. *kiṇa* ‘16 ounces’ from Chinese *chīn (jin)*, *kṣau* ‘voucher, receipt’ from Chinese *ch’ao (chao)*. (3) *Tibetan*. Loans are rare and confined to Late Khotanese, e.g. *ṣkyaisa* ‘official gift, present’ from Tibetan *ṣkyes*. (4) *Turkish*. Except for a number of Turkish proper names and titles in Late Khotanese no loans have so far been found.

6.1.3b Loanwords—Tumshuqese

The short Buddhist Karmavācanā text shows that Tumshuqese likewise borrowed its Buddhist terminology in large measure from Gāndhārī Prakrit, e.g. *uvāsaa-* ‘layman’ (Khot. *uvāsaa-*) and *retena-* ‘jewel’ (Khot. *ratana-*). However in Tumshuqese there is much stronger influence from Tocharian than in Khotanese. The influence from Chinese is minimal, *cāṃṣyā*, from Chinese *ch’ang shih (chang shi)* designating the holder of an administrative office, a term also attested in Khotanese, Middle Persian, and Uighur (Bailey 1961: 118).

6.2 Sociolinguistic aspects

Polite Phraseology. The Khotanese verb *parī-* *parsta-* ‘order’ is used honorifically with the infinitive like Sogdian *frm’y-* and New Persian *farmāy-*: *kū ṣṣta past-ai mūd-ai* ‘Where did you deign to stay?’ In the imperative it may be rendered ‘please’: *laka mara pary-a mūṃd-ai* ‘Please stay here a little’.

The loanword *usahy-* ‘endure’ is used as an honorific instead of a verb of motion: *vā usahy-a ku-ṃ bis-a balys-a* ‘Deign (to come) to where my house (is), Buddha’ (Z 2.58).

It is customary in letters to refer to oneself as *ṅāśa-* *bīsa-* ‘humble servant’. The phrase *drūn-ā pvais-ū* ‘I ask after the health’ of persons known to the writer of a letter is sometimes abbreviated to simply *pvais-ū* ‘I ask after’.

7 DIALECTS AND LANGUAGE DEVELOPMENT

Although some fragments appear to contain rare forms, it has not so far been proved that the extant material in Khotanese and Tumshuqese reflects traces of dialect differentiation. Khotanese and Tumshuqese themselves, although closely related, must be

considered independent languages rather than dialects. They were spoken in widely separated communities that were politically independent of each other, and in view of their divergent historical development the level of mutual intelligibility must have been extremely low.

7.1a Khotanese

See Bailey 1963: vii–viii; Canevascini 1993: xvii; Emmerick 1987b; Skjærvø 1999: 294–301.

It has always been clear that Khotanese shows a striking phonological development that Bailey compares with the development of Latin to Italian and French. It has however not always been realised that the phonologically late forms are also attested very early. The texts reflect accurately the observation made by the Chinese pilgrim Hsüan-tsang (Xuan-zang) in the seventh century that the learned form of the name of Khotan was *hvatana-*, the Old Khotanese form, while at the same time the popular form was *hvaṃna-*, the Late Khotanese form. Kumamoto (1996: 90) suggests that the monastic tradition of writing in Old Khotanese did not survive the period of Tibetan sway in Khotan (ca. 790–840 CE).

Attempts have been made to distinguish between more than just Old and Late Khotanese, but unanimity prevails only with regard to the existence of two phonological systems, the details of which have gradually been established. Skjærvø has recently begun to argue in favour of three stages in the development of the language, which he terms Old, Middle, and Late Khotanese (see Skjærvø, 2002: lxx–lxxi, 2004: lxxiv–lxxv).

7.1b Tumshuqese

The scanty extant material provides no indication of the phonological development of the language. The earliest text, the Karmavācanā, has a different appearance largely due to the fact that it was written before the introduction of the nine new signs used in the later texts (cf. section 1.2, Writing systems).

7.2 Demise of the languages

We have no direct evidence concerning the demise of Tumshuqese and Khotanese, but it is probably significant that no Khotanese text can be dated to later than the tenth century, whereas the language was flourishing throughout the century. It can hardly be coincidental that Khotan fell to Moslem invasion early in the eleventh century.

8 SAMPLE KHOTANESE TEXT Z 23.22–25

<i>jaṃbvīv-i</i>	<i>trām-ä vāstāt-ä</i>	<i>ku</i>	<i>ne marata balys-ä ni vei štä</i>
Jambudvīpa such	became	where not here	Buddha not been is,
‘Jambudvīpa had become such (as)		when no Buddha has been here,’	
<i>kho ju</i>	<i>ṣ-a ṣṣav-a stūr-yau haṃtsa</i>	<i>ku</i>	<i>nä sat-a štā haṃbaḍ-a māst-ä</i>
like indeed the night	stars with	where not risen is	full moon,
‘just like the night with (its) stars		when the full moon has not risen,’	

o kšīr-ä ku ne rre nä-štä
or land where not king not-is
'or a land where there is no king,

o š-a bis-a ku hväšt-ä mīde
or the house where master dies.
or that house where the master dies.'

trām-u māñam-d-u a-näh-ä
such like protector-less
'Likewise, the whole world (was) then without a protector.'

ysamaššand-ei harbišš-ä ttīyā
world all then,

hamggargg-a harbišš-ä hayirūn-e
gatherings all pleasures
'All gatherings, pleasures,

khan-ä būsš-ä panašt-e
laughter jokes disappeared.
laughter, jokes disappeared.'

pan-ye tter-ä kāšc-a uysnor-ä
every's so-great anxiety of-living-being
'The anxiety of every living being (was)

samu kho tt-ye ci mäta mīde
just as of-him whose mother dies.
as great as that-of-one whose mother dies.'

jühän-ä stor-u uysnor-a
yearning greatly living-beings
'Beings (were) greatly yearning

balys-u väte kāšcän-a ysīr-u
Buddha for anxious fiercely
for the Buddha, fiercely anxious.'

paljsärgg-a harbišš-ä hva'nd-ä mari
tormented all men here
'Tormented (were) all men,

ä'gy-e dīvat-e yakš-a
residing deities Yakšas.
deities residing here, Yakšas.'

ACKNOWLEDGEMENTS

I would like to acknowledge here the generous assistance of Prods Oktor Skjærvø in editing this chapter by our sadly missed colleague Ronald E. Emmerick.

[The Editor]

BIBLIOGRAPHY AND REFERENCES

Abbreviated references to Khotanese texts are as recommended in Emmerick 1992a.

- Bailey, H. W. (1946) 'Gāndhārī', *BSOAS* 11.4: 764–797.
 — (1947) 'Recent work in "Tokharian"', *TPS* 1947: 139–145.
 — (1961) *Khotanese Texts* IV, Cambridge: The University Press.
 — (1963) *Khotanese Texts* V, Cambridge: The University Press.
 Burrow, T. (1937) *The Language of the Kharoṣṭhi Documents from Chinese Turkestan*, Cambridge: The University Press.
 Canevascini, G. (1991) 'Medio-reflexive verbs in Khotanese', in Emmerick, R. E. and Weber, D. (eds) *Corolla Iranica. Papers in Honour of Prof. Dr. David Neil MacKenzie on the Occasion of his 65th Birthday*, Frankfurt am Main and New York: P. Lang, pp. 23–26.
 — (1993) *The Khotanese Saṅghāṭasūtra. A Critical Edition*, Wiesbaden: Reichert.
 Degener, A. (1987) 'Khotanische Komposita', *MSS* 48: 27–69.
 — (1989a) *Khotanische Suffixe* (Alt- und Neu-Indische Studien 39), Stuttgart: Steiner Verlag Wiesbaden.
 — (1989b). 'Beispiele der Klassifikation indischen Lehngutes im Khotanischen', *SiII* 15: 41–49.

- (1990) 'Indisches Lehngut im Khotanischen', in Diem, W. and Falaturi, A. (eds) XXIV. *Deutscher Orientalistentag vom 26. bis 30. September 1988 in Köln. Ausgewählte Vorträge*, Stuttgart: Franz Steiner, pp. 381–390.
- (1993) 'Zur Syntax des Khotanischen', in Skalmowski, W. and van Tongerloo, A. (eds) *Medioiranica. Proceedings of the International Colloquium Organized by the Katholieke Universiteit Leuven from the 21st to the 23rd of May 1990*, Leuven: Peeters, pp. 45–56.
- Dresden, M. J. (1955) *The Jātakastava or "Praise of the Buddha's Former Births"*, (Transactions of the American Philosophical Society n.s. 45.2), Philadelphia: American Philosophical Society.
- Emmerick, R. E. (1965) 'Syntax of the cases in Khotanese', *BSOAS* 28.1: 24–33.
- (1968a) *Saka Grammatical Studies* (London Oriental Series 20), London, New York, etc.: Oxford University Press.
- (1968b) 'Khotanese metrics', *Asia Major*, n.s. 14.1: 1–20.
- (1979) 'The Historical Importance of the Khotanese Manuscripts', in Harmatta, J. (ed.) *Prolegomena to the Sources on the History of pre-Islamic Central Asia*, Budapest: Akadémiai Kiadó, pp. 167–717.
- (1981a) 'The vowel phonemes of Khotanese', in Brogyanyi, B. (ed.) *Festschrift for Oswald Szemerényi*, Amsterdam: J. Benjamin, pp. 239–250.
- (1981b) 'The consonant phonemes of Khotanese', *AcIr* 21: 185–209.
- (1985a) *The Tumshuqese Karmavācanā Text*, Mainz: Akademie der Wissenschaften und der Literatur; Stuttgart: F. Steiner Verlag Wiesbaden.
- (1985b) 'Tibetan loanwords in Khotanese and Khotanese loanwords in Tibetan', in Gnoli, G. and Lanciotti, L. (eds) *Orientalia Iosephi Tucci memoriae dicata* (Serie orientale Roma LVI, II 1), Roma: Istituto italiano per il Medio ed Estremo Oriente, pp. 301–317.
- (1987a) 'Auxiliaries in Khotanese', in Harris, M. and Ramat, P. (eds) *Historical Development of Auxiliaries* (Trends in Linguistics. Studies and Monographs 35), Berlin, New York: Mouton de Gruyter, pp. 271–290.
- (1987b) 'The transition from Old to Late Khotanese', in *Transition periods in Iranian history. Actes du symposium de Fribourg-en-Brisgau (22–24 Mai 1985)*, Leuven: Association pour l'avancement des études iraniennes; distr. E. Peeters, pp. 33–42.
- (1989) 'Khotanese and Tumshuqese', in *CLI*, pp. 204–229.
- (1990) 'Khotanese *ma* "not"', in Gnoli, G. and Panaino, A. (eds) *Proceedings of the first European Conference of Iranian Studies held in Turin, September 7th–11th, 1987 by the Societas Iranologica Europaea*, Part 1: *Old and Middle Iranian Studies*, Roma: Istituto Italiano per il Medio ed Estremo Oriente, pp. 95–113.
- (1992a) *A Guide to the Literature of Khotan (Studia philologica buddhica (Occasional Paper Series III))*, 2nd ed. revised and enlarged, Tokyo: The International Institute for Buddhist Studies.
- (1992b) 'The Dunhuang MS Ch 00120: Its importance for reconstructing the phonological system of Khotanese', in Cadonna, A. (ed.) *Turfan and Tunhuang: the Texts* (Orientalia venetiana IV), Florence: Leo S. Olschki Editore, pp. 145–170.
- (1998) 'Khotanese *ei*', in Sims-Williams, N. (ed.) *Proceedings of the Third European Conference of Iranian Studies held in Cambridge, 11th to 15th September 1995*. Part 1: *Old and Middle Iranian Studies*, Wiesbaden: L. Reichert, pp. 93–97.
- Emmerick, R. E. and Maggi, M. (1991) 'Thoughts on Khotanese *e* and *o*', in Emmerick, R. E. and Weber, D. (eds) *Corolla Iranica, Papers in honour of Prof. Dr. David Neil MacKenzie on the occasion of his 65th birthday on the eighth of April 1991*, Frankfurt am Main, New York: P. Lang, pp. 67–73.
- Emmerick, R. E. and Pulleyblank, E. G. (1993) *A Chinese Text in Central Asian*

- Brahmi Script. New Evidence for the Pronunciation of Late Middle Chinese and Khotanese* (Serie orientale Roma lxi), Roma: Istituto italiano per il Medio ed Estremo Oriente.
- Emmerick, R. E. and Skjærvø, P. O. (1982) *Studies in the Vocabulary of Khotanese I* (Veröffentlichungen der Iranischen Kommission 12), Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- (1987) *Studies in the Vocabulary of Khotanese II* (Veröffentlichungen der iranischen Kommission 17), Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- (1997) *Studies in the Vocabulary of Khotanese III* (Veröffentlichungen der iranischen Kommission 27), Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Gertsberg (Herzenberg), L. G. (1965) *Khotano-sakskii iazyk*, Moskva: Nauka.
- Heston, W. L. (1976) *Selected Problems in Fifth to Tenth Century Iranian Syntax*, Ph.D. dissertation, University of Pennsylvania, Philadelphia (University Microfilms International, Ann Arbor & London).
- Hitch, D. (1988) 'Penalty clauses in Tumshuqese', *StIr* 17.2: 147–152.
- Hucker, C. O. (1985) *A Dictionary of Official Titles in Imperial China*, Stanford, Calif.: Stanford University Press.
- Konow, S. (1935). *Ein neuer Saka-Dialekt* (= *SbPAW* 20), Berlin: Preussische Akademie der Wissenschaften, pp. 772–823.
- Kumamoto, H. (1995) 'Did Late Khotanese have a three vowels system?', in Fragner, B. G., Fragner, C., Gnoli, G., Haag-Higuchi, R., Maggi, M. and Orsatti, P. (eds) *Proceedings of the Second European Conference of Iranian Studies held in Bamberg, 30th September to 4th October 1991, by the Societas Iranologica Europaea*, Roma, pp. 383–90.
- Leumann, E. (1933–1936) *Das nordarische (sakische) Lehrgedicht des Buddhismus*, Text und Übersetzung, aus dem Nachlaß hrsg. von M. Leumann (= *AKM* 20), Leipzig.
- Leumann, M. (1934) *Sakische Handschriftproben*, Zürich: Pivatdruck (Augustin in Glückstadt & Hamburg).
- Lüders, H. (1913) 'Die Śakas und die "nordarische" Sprache', *SBAW* 23, pp. 406–427 (repr. in *Philologica Indica. Ausgewählte kleine Schriften von Heinrich Lüders. Festgabe zum siebzigsten Geburtstage*, Göttingen: Vandenhoeck & Ruprecht, 1940, pp. 236–255).
- Maggi, M. (1991) 'Note tumshuqesi', in *Studia linguistica amico et magistro oblata. Scritti di amici e allievi dedicati alla memoria di Enzo Evengelisti*, Milano: UNICOPLI, pp. 219–228.
- (1992) *Studi sul sistema accentuale del cotanese*. Ph.D. diss., Istituto universitario orientale di Napoli.
- (1993) 'Sull'accento cotanese', *Atti del Sodalizio Glottologico Milanese* 31: 180–187.
- (1995) 'Late Khotanese numerals from 1 to 36 in Tibetan transcription', in Fragner, B. G., Fragner, C., Gnoli, G., Haag-Higuchi, R., Maggi, M., and Orsatti, P. (eds) *Proceedings of the Second European Conference of Iranian Studies held in Bamberg, 30th September to 4th October 1991, by the Societas Iranologica Europaea*, Roma: Istituto Italiano per il Medio ed Estremo Oriente, pp. 425–443.
- Maue, D. (1996) *Altürkische Handschriften. Teil 1: Dokumente in Brāhmī und tibetischer Schrift* (Verzeichnis der orientalischen Handschriften in Deutschland 13.9), Stuttgart: F. Steiner Verlag Wiessbaden.
- (1997) 'A tentative Ssemma of the varieties of Brāhmī script along the northern Silk Road', in Akiner, S. and Sims-Williams, N. (eds) *Languages and Scripts of Central Asia*, London: School of Oriental and African Studies, University of London, pp. 1–15.

- Rong and Duan (1996 [2000]) Rong Xinjiang and Duan Qing, 'Jushideyu [On Gyāzdi]', *Zhongya xuekan* [Central Asian Studies] 5: 9–21.
- Sander, Lore (1968) *Paläographisches zu den Sanskrithandschriften der Berliner Turfansammlung (Verzeichnis der orientalischen Handschriften in Deutschland Supplementband 8)*, Wiesbaden: F. Steiner.
- (1984) 'Zu dem Projekt "Paläographie khotan-sakischer Handschriften"', in Skalmowski, W. and van Tongerloo, A. (eds) *Middle Iranian Studies, Proceedings of the International Symposium Organized by the Katholieke Universiteit Leuven from the 17th to the 20th of May 1982*, Leuven: Peeters, pp. 159–186.
- (1986) 'Brāhmī scripts on the eastern Silk roads', *StII* 11/12: 159–192.
- Schmidt, K. T. (1988) 'Ein Beitrag des Tocharischen zur Entzifferung des Tumšusqakischen', *AoF* 15: 306–314.
- Sims-Williams, N. (1990) 'Chotano-Sogdica II: Aspects of the Development of Nominal Morphology in Khotanese and Sogdian', in Gnoli, G. and Panaino, A. (eds) *Proceedings of the First European Conference of Iranian Studies held in Turin, September 7th–11th 1987 by the Societas Iranologica Europaea. Part 1: Old and Middle Iranian Studies*, Roma: Istituto Italiano per il Medio ed Estremo Oriente, pp. 275–296.
- (1994) 'The triple system of deixis in Sogdian', *TPhS* 92.1: 41–53.
- (1996) 'The Sogdian manuscripts in Brāhmī script as evidence for Sogdian phonology', in Emmerick, R. E., Sundermann, W., Warnke, I. and Zieme, P. (eds) *Turfan, Khotan und Dunhuang. Vorträge der Tagung "Annemarie v. Gabain und die Turfanforschung", veranstaltet von der Berlin-Brandenburgischen Akademie der Wissenschaften in Berlin (9.–12.12.1994)*, Berlin: Akademie Verlag, pp. 307–315.
- (1997) 'The denominal suffix *-ant-* and the formation of the Khotanese transitive perfect', in Lubotsky, A. (ed.) *Sound Law and Analogy. Papers in Honor of Robert S. P. Beekes on the Occasion of his 60th Birthday*, Amsterdam, Atlanta: Rodopi, pp. 317–325.
- Skjærvø, P. O. (1981) 'The Old Khotanese fragment H 147 NS 115 and remarks on Old Khotanese *hamdārvāto*, *patīšu*, *vya* and *ya*', *BSOAS* 44.3: 453–467.
- (1987) 'On the Tumshuqese Karmavācanā text', *JRAS* 1987: 77–90.
- (1999) 'Khotan, an early center of Buddhism in Chinese Turkestan', in McRae, J. R. and Nattier, J. (eds) *Collection of Essays 1993: Buddhism Across Boundaries – Chinese Buddhism and the Western Regions*, Sanchung, Taipei: Fo Guang Shan Foundation, pp. 266–344.
- (2002) *Khotanese Manuscripts from Chinese Turkestan in The British Library. A Complete Catalogue with Texts and Translations*, with Contributions by Ursula Sims-Williams, The British Library (Corpus Inscriptionum Iranicarum II/V, Texts VI).
- (1999 [2002]) 'Recent Khotanese Ghostwords', *Bulletin of the Asia Institute* 13, pp. 151–55.
- (2004) *This Most Excellent Shine of Gold, King of Kings of Sutras. The Khotanese Suvarṇabhāsottamasūtra*, Sources of Oriental languages and literatures 60–61, Central Asian sources 5–6, [Cambridge, Mass.]: Harvard University, Dept. of Near Eastern Languages and Civilizations.
- (2006) "Iran iv. Iranian Languages and Scripts," in *EnIr* 13.3, pp. 344–77.
- (forthc.) 'Khotanese', in *EnIr*.
- 'The end of eighth-century Khotan in its texts', *Journal of Inner Asian Art and Archaeology* 3, 2008: 119–44.

The Khotanese texts published by Bailey, transcribed by R. E. Emmerick and H. Kumamoto, and the Tumshuqese texts, transliterated by D. Maue, are available at <http://titus.uni-frankfurt.de/indexe.htm>.

CHAPTER EIGHT

PERSIAN AND TAJIK

Gernot Windfuhr and John R. Perry

1 INTRODUCTION

1.1 Overview

The focus of this chapter is Modern Standard Persian and Modern Standard Tajik. Both evolved from Early New Persian. Western Persian has typologically shifted differently from modern Tajik which has retained a considerable number of Early Eastern Persian features, on the one hand, and has also assimilated a strong typologically Turkic component, on the other hand. In spite of their divergence, both languages continue to share much of their underlying features, and are discussed jointly in this chapter.

1.1.1 Historical background

Persian has been the dominant language of Iranian lands and adjacent regions for over a millennium. From the tenth century onward it was the language of literary culture, as well the *lingua franca* in large parts of West, South, and Central Asia until the mid-nineteenth century. It began with the political domination of these areas by Persian-speaking dynasties, first the Achaemenids (c. 558–330 BCE), then the Sassanids (224–651 CE), along with their complex political-cultural and ideological Perso-Iranianate constructs, and the establishment of Persian-speaking colonies throughout the empires and beyond. The advent of Islam (since 651 CE) represents a crucial shift in the history of Iran and thus of Persian. It resulted in the emergence of a double-focused Perso-Islamic construct, in which, after Arabic in the first Islamic centuries, Persian reasserted itself as the dominant high register linguistic medium, and extended its dominance into formerly non-Persian and non-Iranian-speaking territories in the East and Central Asia.

The writing system became that of the new dominant religion, and there occurred increasing infusion of Arabic features into the lexicon, phonology and grammar (comparable to the absorption of the Norman component into English). However, throughout the evolution of the literary standards from Early New Persian to Modern Standard Persian the considerable typological changes that Persian underwent are due to both internal Persian developments, including the leveling of regional features, and to the assimilation of expanding areal cross-linguistic typological isoglosses.

1.1.2 Persian and related groups

See Chapter 2, Map 2.6.

1.1.2.1 Varieties of Persian

Overall, Persian varieties are divided into a Western group mainly in Iran and an Eastern group in Afghanistan and Central Asia, with transitional varieties.

Caucasus

The northwestern outpost of Persian is Caucasian Tat Persian spoken in an Azeri Turkic, Caucasian and Armenian environment, with three varieties: (a) Muslim and (b) Jewish (Juhuri) in Azerbaijan and Dagestan, and (c) Christian Armeno-Tat in Armenia.

Iran

(1) Persian sociolinguistic registers include:

- (a) Modern Standard Persian, the written norm in Iran (Farsi) and Afghanistan (Dari), evolved during the last few centuries;
- (b) Colloquial Persian, specifically the normalized form of Colloquial Tehrani Persian, used for most polite spoken communication, which increasingly shows reflexes in the standard language;
- (c) *Xodemuni* ‘our own’, (< *xod-emān-i*), i.e. familiar speech, the non-normalized local variant such as in Tehran.

(2) Regional and local varieties in the urban centers throughout Iran in non-Persian dialect and language areas.

(3) Khorāsān Persian varieties, representing a major distinct regional subgroup and stretching from east of Tehran to the Afghan border. Tehran to the Afghan border.

Transitional group

Varieties straddling the Iran-Afghan border: (1) closely related Kohistāni in Iran and Afghan Fārsiwāni, and (2) Sistāni on both sides of the border.

Afghanistan

See Kieffer 1983, incl. map and detailed table.

- (1) Afghan Persian, officially called Dari, mostly close to literary Persian.
- (2) Kaboli (Kāboli), increasingly become the standard Afghan vernacular.
Large regional Persian varieties include:
- (3) Herati (Herāti) near the Iranian border.
- (4) Ayamaqi (Aymāqi) near Herat (in pockets also in Iranian Khorasan).
- (5) Hazaragi (Hazāragi), stretching northeast through the Afghan center, some in Iran.
- (6) Afghan Tajik in the NE, including Afghan Badakhshani, the latter with the outpost Madaglashti near Peshawar in Pakistan.

Note that the terms “Tajik” as well as “Dari” are sometimes erroneously used collectively for all Afghan varieties.

Central Asia

Tajik, in an Uzbek Turkic environment, with four broadly defined groups (see 1.2 below).

1.1.2.2 Related varieties and dialect groups

Colonial Persian

Persian was cultivated at the courts of the Anatolian Seljuk and Ottoman rulers (from ca. 1200–1922), several of whom are known for composing Persian poetry. Probably best known among their protégés is Rumi (d. 1273), the most cherished Persian mystic poet who had come to Konya from Wakhsh near Balkh in Afghanistan. Literary Ottoman Turkish is a virtual amalgam of Turkish and Persian (with all of the latter's Arabic loan elements).

In the East, Urdu developed under heavy Persian influence. Persian first entered India with the conquest of north-west India by Ghaznavid armies in the eleventh century. Four centuries later, Persian was chosen as the court language of the Mogul rulers (1530–1857), who were major patrons of Persian literature and poets from Iran, unlike the contemporary Safavids in Iran.

It was at the courts of India and Turkey where many of the major traditional dictionaries of Persian were compiled from the fifteenth to the eighteenth centuries, including grammatical treatises. Simultaneously, there developed in India a Persian vernacular, and it was from the Indian scribes and secretaries that the English officers of the East India Company, a number of whom wrote grammars of Persian, learned their Persian, with all its local idiosyncrasies. Persian was abolished in its last official bastion—the courts of law—in 1837 by the authorities of the East India Company.

Judeo-Persian

Judeo-Persian varieties, like other Judeo-Iranian varieties, are found throughout Iranian-speaking regions, such as Jewish Tat Persian and Bukhara Tajik. In fact, Jewish merchants and travellers have been the earliest speakers who wrote in Early New Persian, and left its earliest documents. These and later documents, in a Hebrew-writing tradition parallel to that in Arabic, have been crucial for the study of the diachrony and diatopy of Persian (e.g. Lazard 1968; Paul, ed., 2003a).

Luri-type and Fars dialects

In SW Iran there are two groups which can be recognized as “Perside”, i.e. they continue numerous features that evolved from Southern Early New Persian (see 7 Diachrony), though each evolved differently:

- (1) The Luri-type dialects (Luri proper, Bakhtiāri, Boyer-Ahmadi, Mamasani-Kohgeluye).
- (2) The Fars dialects stretching from the Gulf into western and central Fars.

1.1.2.3 Number of speakers

According to various sources, the totals of native speakers of Persian and its varieties (or rather ‘ethnic’ Persians) in the three Persian-speaking countries vs. the total population (who at least use or understand Persian as a second language) are: Iran 35/70 m; Afghanistan 17/33 m (Hazāra 2.8 m; Aymāq 380,000); Tajikistan 5~6/7.2 m. Overall, the numbers are, counting speakers outside these countries, a total of 60/110 m, and of these: 35/70 m for Western Persian vs. 23/40 m for Eastern Persian. The estimates for Caucasian Tat suggest about 26,000, now reduced from a much larger number due to (partially forced) assimilation and emigration.

1.1.3 Evolution of Standard Persian

1.1.3.1 Early linguistic groups

Sources from the ninth–eleventh centuries distinguished four linguistic groups:

(1) Pārsi, the literary Middle Persian (mainly used by Zoroastrian priests).
 (2) Pārsi proper (*mutlaq*), the literary and spoken Southern Early New Persian used from Fars to Sistan, which had retained numerous Middle Persian features and vocabulary, with relatively few Arabic loans.

(3) In the North, (*Pārsi-i*) *Dari*: The term originally referred to the administrative and spoken Persian that had developed at the Sassanian court in Ctesiphon and was administratively used throughout the Empire. In Khorasan it had already replaced the local Parthian dialects even before Islamization. As a ready vehicle for the Muslim administrations, besides Arabic, it became the vehicle of the Muslim mission into Central Asia and beyond, where a Persian variety had already been used as a *lingua franca*. This process not only led to the ultimate replacement of Sogdian, Bactrian, and Khwarezmian, but also inserted into this Dari an increasingly larger Arabic loan component as well as local eastern Persian and other Iranian vocabulary.

(4) Pahlavi, Ar. *fahlavi*, lit. ‘Parthian’: The term implied the non-Persian Iranian languages, particularly in western and central Iran which was once part of the former Parthian Empire.

In addition, Early New Persian varieties must also have been spoken in Sassanian border garrisons east and west. One of these was probably the ancestor of Tat Persian in the Sassanian outpost at the Caspian gate to the Caucasus, Darband. The others, probably older, were the outpost in Central Asia out of which ultimately developed Afghan and Tajik Persian.

See also the detailed study by Perry (2009).

1.1.3.2 Arabic and Turkic

See Chapter 2, Map 2.7.

(1) Arabic and Turkic

A major contributing factor to the Arabization of Persian was the magnitude of Arab settlement in Greater Iran, and presumably the intense interaction and intermarriage between the immigrants with the local populations at both the highest and lowest social levels, and after large numbers of Arab tribes moved into the Fertile Crescent during Sassanian times, Shapur II (r. 309–379) settled some of them in Fars as well as in the hinterland of Bam and Kerman. After Islam, settlement occurred in various waves throughout, and was most extensive in eastern Iran, including Khorasan, Afghanistan, and Central Asia. At its height, the number of Arab immigrants may have totaled 250,000. While these Arabic-speaking populations were ultimately absorbed, except for isolated Central Asian Arabic pockets (eastern Iran, northern Afghanistan, Central Uzbekistan), Arabic continued as the high register literary language during the earlier centuries of New Persian, mainly as the dominant language of science and religion, and may have at least indirectly affected even syntax, particularly through extensive translation activities both from and into Arabic.

In contrast, the immigration of Turkic speakers has led to the Turkification of large regions, mainly in Azerbaijan (Azeri Turkic) and less so in Fars (Qashqā’i) and Khorasan (Turkmen). In addition, dynasties of Turkic or Turko-Mongol origin ruled in Iran from

the late tenth century to the twentieth century. Linguistically, there has been considerable Iranization of Turkic, but the intense Turko-Iranian symbiosis has also effected grammatical innovations in Persian, including the emergence of the evidential category in the verb system during the nineteenth century (for this section, see also Windfuhr 2006).

(2) Shift of cultural centers

Early New Persian Dari became the vehicle for the emerging New Persian literature. Even before the Mongol invasion in the thirteenth century, the literary center shifted westward, prominently to the city of Shiraz in Fars. This extended process led not only to the elimination of the Eastern linguistic shibboleths, but also to the demise of the Early Southern standard. With the rise of the Safavids at the beginning of the sixteenth century, the center shifted north to Turko-Persian bilingual Esfahan, and finally with the rise of the Qajars to Tehran at the beginning of the nineteenth century.

In overview, the sociolinguistic set-up from Late Sassanian times may be shown as follows: (see also section 8 Diachrony and dialectology)

	<i>Iranians</i>	<i>Persians/Non-Iranians</i>
	High	Low/Lingua franca
Late Sasanian	Middle Persian	Dari
Early Islamic	Arabic, Dari	Dari
Mongols, 13th cent.	Persian	Persian
Safavids, 16th cent.	Persian	Persian Turkic
Qajars, 19th cent.	Persian	Persian

1.2 Tajik

Tajik, or Tajik Persian (*zabon-i tojikī*, *zabon-i forsi-i tojik*) is a variety of New Persian used in Tajikistan and parts of Uzbekistan, including the cities of Bukhara and Samarkand. Since the 1920s it has been fostered as the national and literary language of the Soviet Socialist Republic (from 1991, the independent state) of Tajikistan. The Tajik and Iranian Persian speech areas are not contiguous, but lie at opposite ends of a continuum with Persian dialects of Afghanistan in between, and interrupted by areas of Turkic (Turkmen and Uzbek) speech. Spoken Tajik has been evolving independently of Persian of Iran since the sixteenth century, but the written language (which functioned as the common language of high culture, government, and diplomacy in Central Asia and India) maintained a virtually universal standard, based on Classical Persian, until the early decades of the twentieth century.

In the Soviet period, with the promotion of a more vernacular style and lexicon and the systematic introduction of Russian loanwords, language change was more rapid. The writing system was changed from Arabic to Latin in 1928, then to Cyrillic in 1939. The modern literary language (*zabon-i adabi-i hozira-i tojik*, as planned and exemplified by Soviet Tajik writers of the 1920s and 1930s, notably Sadriiddin Aini) was based loosely on the style of the old cultural center, Bukhara, and contained many Uzbek loanwords and some syntactic structures calqued on Uzbek usage. In recent decades writers from different regions of Tajikistan, and some who have traveled abroad, have introduced a more varied style, including features closer to the Persian of Iran.

Tajik dialects may be divided broadly into two groups: North-western and South-eastern, corresponding in rough topographical terms to the lowlands and highlands respectively of the Oxus basin. Several refinements of this scheme have been proposed, and much fieldwork remains to be done. The most comprehensive classification posits

four groups: (1) Northern, comprising Bukhara, Samarkand and Derbend in Uzbekistan, the Ferghana Valley (including Khujand, former Leninabad), and extending down the Varzob valley to the region of the capital, Dushanbe. (2) Central, comprising the upper Zarafshon (Zeravshan) valley. (3) Southern, stretching south and east of the capital, in Kulob and Qarotigin (Karategin) regions, including Gharm, as far as the Pamirs (Mountain Badakhshan). (4) South-eastern, in Darvoz region, on the bend of the Oxus near Rushan.

Dialects of the last two groups have strong affinities with those of the left bank of the upper Oxus in Afghanistan's province of Badakhshan. They include Tajik as the contact vernacular (called *forsi*) of Gorno-Badakhshan, extending into Afghanistan and Wakhan. In these regions the native language of a majority of the population is one of the Iranian dialects of the Pamir group (see Chapter 14a).

Northern dialects have been influenced to varying degrees by Uzbek, with which there is widespread bilingualism. A distinctive sub-variety of Northern Tajik speech, with a modest literature, is that of the Jews of Bukhara, most of whom have emigrated. Dushanbe was a small market town before its promotion to national capital in 1924 and the consequent influx of Tajiks from elsewhere in the region and of Russians and other Soviet nationalities. Though nominally included in the Northern dialect area, it is again in a state of demographic fluctuation since independence, and is best considered separately. There are approximately five million Tajik Persian speakers in Central Asia and Afghanistan, including minorities in Uzbekistan and Kirgizstan. The so-called Tajiks of south-west Xinjiang, in and around Taxkorgan (Tashqurghon) – an official nationality in China – are speakers of Pamir languages (mainly Sarikoli and Wakhi), not Persian.

The variety of Tajik described here is basically the ideal, and to an extent still transparently hybrid, Modern Literary Tajik (MLT), unless otherwise noted.

Notes

- (1) In the following Modern Standard Persian and Modern Standard Tajik are referred to as Persian and Tajik.
- (2) In examples, Persian and Tajik forms are separated by a slash: *xānd/xond* '(s)he read'.
- (3) In phonemic transcription, (a) initial glottal stop is not marked, and (b) stress is indicated by an underlined vowel.

1.3 Writing systems

1.3.1 Persian

The Persian writing system uses the Arabic alphabet, which is a consonantal system. Vowels are written as follows: the three long vowels are represented by the letter of the consonant nearest in pronunciation. Thus, the letter <y> represents both *y* and *i* [*ī*], <w> both *v* and *u* [*ū*], and <alef> both the glottal stop ' and *ā*. The three short vowels may be, but are usually not, represented by diacritics which ultimately derive from the same letters <w>, <y>, and <alef>. The main innovations in Persian are two: additional letters were created for the four Persian consonants *p*, *č*, *g*, *ž* by providing three dots to distinguish the "Arabic" letters , <j>, <k>, <z> (the dots merged into an oblique stroke in the

case of <g>). Unlike Arabic, in final position, short vowels are always represented by consonantal letters, final *o* by <w>, and both final *e* and *a* by <h>.

The Arabic orthography is retained, distinctly by the subset of letters marked A in Table 8.1a. They represent pharyngeals, a pair of non-strident interdental fricatives, and the voiceless stop *q*, which in Persian phonemically merged with the corresponding plain and strident consonants, respectively, except *z* > *z*, while *q* merged with voiced *ɣ*. Otherwise, the orthography is basically phonemic in the representation of the consonants and long vowels, but does not represent short vowels other than in final position. In rare

TABLE 8.1a: PERSO-ARABIC ALPHABET

Position in Letter Group				Name	Phonemic
Final	Medial	Initial	Separate		
ا			ا	alef	ʔ
آ			آ		ā
ب	ب	ب	ب	be	b
پ	پ	پ	پ	P pe	p
ت	ت	ت	ت	te	t
ث	ث	ث	ث	A se	s
ج	ج	ج	ج	jim	j
چ	چ	چ	چ	P če	č
ح	ح	ح	ح	A he-ye jimih	
خ	خ	خ	خ	xe	x
د			د	dāl	d
ذ			ذ	A zāl	z
ر			ر	re	r
ز			ز	ze	z
ژ			ژ	P že	ž
س	س	س	س	sin	s
ش	ش	ش	ش	inš	š
ص	ص	ص	ص	A sād	s
ض	ض	ض	ض	A zād	z
ط	ط	ط	ط	A tā	t
ظ	ظ	ظ	ظ	A zā	z
ع	ع	ع	ع	A eyn	ʔ
غ	غ	غ	غ	ɣeyn	ɣ
ف	ف	ف	ف	fe	f
ق	ق	ق	ق	A qāf	q
ک	ک	ک	ک	kāf	k
گ	گ	گ	گ	P gāf	g
ل	ل	ل	ل	lām	l
م	م	م	م	mim	m
ن	ن	ن	ن	nun	n
و			و	vāv	v/u/ow/final o
ه	ه	ه	ه	he	h/final e, a
ی	ی	ی	ی	ye	y/i/ey

A = letters occurring mostly in Arabic loanwords; P = letters found in Persian only.

The basic shapes ا د ر و do not connect to the left.

instances, an Arabic pharyngeal letter is used, such as <š> in <šad> for *sad* ‘hundred’. In Turkic loans, fronted /k/ and backed /k/ are represented by <k> and <q>, respectively.

1.3.2 Tajik

The Tajik alphabet adds six supplementary letters to the Cyrillic inventory (distinguished by means of diacritics), and despite post-Soviet spelling reforms, the Cyrillic orthography established for Tajik in 1939 remains Russian-specific in significant ways. Since examples taken from MLT will be presented in transliteration, certain peculiarities of the standard orthography must be noted here.

The yotated vowels (я, е, ё, ю: in Russian, devices to indicate that a preceding consonant is palatalized) each serve to represent the combination of the consonant /y/ and a following vowel, though /y/ is sometimes represented by й: cf. гӯяд *gūy-ad* ‘says’, юндӣ *yundī* ‘dishwater’, хонае *xonae* /xona-yel/ ‘a house’, тайёр *tauyor* ‘ready’; in the last example, the geminate /yy/ is written with a sequence of two distinct graphs. Prefixed verb stems may likewise be disguised: ояд *o-yad* or биёяд *bi-o-yad* ‘let him come’. The letter <e> additionally represents simple /e/ after a consonant: дидем *did-em* ‘we saw’; word-initially, /e/ is written with non-yotated э: элак *elak* ‘sieve’. Russian и is quasi-yotated and in Tajik always represents /yi/ after a vowel: хонаи ман *xona-i man* /xona-yi man/ ‘my house’, наистон *naiston* /nayiston/ ‘reed bed’.

The letter ӣ (with macron) is a device to distinguish accented word-final -ī (one of two morphological formatives) from unstressed final -i of the syntactic *ezāfelizofat* (EZ) clitic (see next section): дӯсти ман ‘my friend’, but дӯстӣ ‘friendship’; the macron is dropped before an EZ: дӯстии халқхо *dūsti-i xalq-ho* /dūsti-yi xalq-ho/ ‘friendship of peoples’. It is also used in monosyllables: кӣ *kī* ‘who?’ (as distinct from the unstressed subordinizer *ki*), чӣ *čī* ‘what?’, сӣ *sī* ‘thirty’, -мӣ *-mī*? (interrogative enclitic).

2 PHONOLOGY

A major factor in the development of the Persian phonology has been the impact of the Arabic loan component, which constitutes approximately 50 percent of the lexicon, with approximately 25 percent frequency of usage. Not as extensive was the impact of the Turkic loan component (and Persian-Turkic bilingualism). Also, the effects of social dynamics are difficult to determine. A particularly challenging aspect is the rather unique cultural setting in Iran, where until recently schooling involved the intensive study of classical Persian literature and prosody through all levels of education.

2.1 Inventory

The sound system of both Modern Standard Persian and Tajik are essentially the same and quite symmetrical. There are a total of 29 Persian and 30 Tajik segmental phonemes including four pairs of stops and four pairs of fricatives, two nasals, a liquid and a trill, three glides, and three pairs of vowels. The single difference is found in the diachronic development of the ENP velar fricative *ɣ* and loaned uvular *q*, which merged in Persian but not in Tajik, and that of the vowels.

TABLE 8.1b: TAJIK ALPHABETS

Cyrillic Letters	Sound	Letter name	ARABO-PERSIAN ALPHABET EQUIVA			
			Position in letter group			
			Final	Medial	Initial	Isolated
а	a				ا	ا
б	b	be	بـ	بـ	بـ	بـ
в	v	vov	وـ	وـ	وـ	وـ
г	g	gof	گـ	گـ	گـ	گـ
ғ	γ	γayn	غـ	غـ	غـ	غـ
д	d	dol	دـ	دـ	دـ	دـ
е	e					
ё	yo					
ж	ž	že	ژـ	ژـ	ژـ	ژـ
з	z	ze	زـ	زـ	زـ	زـ
з	z	zol	ذـ	ذـ	ذـ	ذـ
з	z	zod	ظـ	ظـ	ظـ	ظـ
з	z	zo	ظـ	ظـ	ظـ	ظـ
и	i, V-yi	ye				
ӣ	ī	i (-i zadanok)				
й	y		يـ	يـ	يـ	يـ
к	k	kof	کـ	کـ	کـ	کـ
к	q	qof	قـ	قـ	قـ	قـ
л	l	lom	لـ	لـ	لـ	لـ
м	m	mim	مـ	مـ	مـ	مـ
н	n	nun	نـ	نـ	نـ	نـ
о	o		وـ	وـ	وـ	وـ
п	p	pe	پـ	پـ	پـ	پـ
р	r	re	رـ	رـ	رـ	رـ
с	s	sin	سـ	سـ	سـ	سـ
с	s	se (-i se nuqta)	سـ	سـ	سـ	سـ
с	s	sod	صـ	صـ	صـ	صـ
т	t	te	تـ	تـ	تـ	تـ
т	t	to	طـ	طـ	طـ	طـ
у	u					
ӯ	ū		وـ	وـ	وـ	وـ
ф	f	fe	فـ	فـ	فـ	فـ
х	x	xe	хـ	хـ	хـ	хـ
х	h	he (-i hutti)	хـ	хـ	хـ	хـ
х	h, -a	he (-i havvaz)	хـ	хـ	хـ	хـ
ч	č	če	چـ	چـ	چـ	چـ
ҷ	j	jim	چـ	چـ	چـ	چـ
ш	š	šin	شـ	شـ	شـ	شـ
ь	ʔ	alif			ا	ا
ѣ	ʔ	ayn	عـ	عـ	عـ	عـ
э	e initial					
ю	yu					
я	ya					

TABLE 8.2: VOWEL SYSTEMS

Diachronic Development						
Modern Standard Persian	<i>e</i>	<i>i</i>	<i>o</i>	<i>u</i>	<i>ā</i>	<i>a</i>
		^		^		
Early New Persian	<i>i</i>	<i>ī ē</i>	<i>u</i>	<i>ū ō</i>	<i>ā</i>	<i>a</i>
	v		v			
Tajik Persian	<i>i</i>	<i>e</i>	<i>u</i>	<i>ū</i>	<i>o</i>	<i>a</i>
Qualitative Characteristics						
<i>Modern Standard Persian</i>			<i>Modern Standard Tajik</i>			
	Front	Central	Back	Front	Central	Back
High	<i>i</i>		<i>u</i>	High	<i>i</i>	<i>u</i>
High mid			<i>o</i>	High mid		<i>ū</i>
Mid	<i>e</i>		<i>ā</i>	Mid	<i>e</i>	<i>o</i>
Mid low		<i>a</i>		Mid low		<i>a</i>
Low				Low		
Tenseness				Tenseness		
tense/long (stable)	<i>i</i>	<i>u</i>	<i>ā</i>	stable	<i>e</i>	<i>ū</i> <i>o</i>
lax/short (unstable)	<i>e</i>	<i>o</i>	<i>a</i>	unstable	<i>i</i>	<i>u</i> <i>a</i>

2.1.1 Vowels

Both Modern Standard Persian and Tajik have reduced the eight-vowel inventory of Middle and Early New Persian to six, but in quite different ways. Whereas in the Persian of Iran the two long mid vowels *ē* and *ō* collapsed with high *ī* and *ū*, in some varieties of Tajik length was neutralized by the merger of the short and long high vowels and the rounding of long *ā*, as in the scheme in Table 8.2. The vowel systems of Persian and Tajik are thus as follows (see Windfuhr 1987: 543, and Ansarin 2008).

2.1.1a Persian

While the original opposition was one of length, the lowering of the short high vowels and the raising of the long mid vowels and merger with the old long high vowels resulted in the opposition of levels, a pair of high vowels and pair of mid vowels. Correspondingly, *ā* is lower than low mid *a*. The primary distinguishing feature, while inherently long and short, respectively, is stability, in part indicated by the lengthening of *e a o* (< *i a u*) when stressed and their shortening and partial conditioned assimilation when unstressed.

While length distinction is neutralized in Persian in most environments, it is usually found in initial stressed syllables of vocatives and imperatives (Windfuhr 1979: 136):

<i>i</i> vs. <i>e</i> :	<i>bī-adab!</i>	<i>bē-gu!</i>	‘impolite!’	‘say!’
<i>u</i> vs. <i>o</i> :	<i>hūšang!</i>	<i>hōseyn!</i>	‘Hushang!’	‘Hoseyn!’
<i>ā</i> vs. <i>a</i> :	<i>nā-kas!</i>	<i>na-gu!</i>	‘nobody!’	‘don’t say!’

Assimilation

In Persian colloquial speech the unstable vowels *e o a* are subject to assimilation in height and rounding: *kelid* > *kīlid* ‘key’, *xe yābān* > *xī yābun* ‘street’, *bolur* > *būlur* ‘crystal’, *bē-gu* > *bo-gu* ‘say!’ (*bē-* > *bo-* ~ *bū-* only in a few high-frequency stems with back vowel).

Similarly, *a* (1) tends to be raised (a) before sibilants: *madrase* > *madrese* ‘school’, *čašm*

> *čes̄m* ‘eye’ (whose *a* is retained in the venerable expression *čes̄m* ‘all right, it will be done’ abbreviated from the phrase meaning ‘by the light of my eyes’), and (b) generally in the 2s and 3s personal clitics: *ketāb-at, -as̄* > *ketāb-et, -es̄* ‘your, his book’, and (2) there tends to be length assimilation: *nahār* > *nāhār* ‘lunch’.

The two diphthongal sequences *ey* and *ow* (< *ai, au*) increasingly tend to be contracted to the long vowels [e:], and *ow* > [o:] even in higher registers: *key* > [ke:] ‘when’, *dow* > [do:] ‘run(ning), race’, *če-towr* > [četo:r] ‘how’, contrasting with short *e* and *o*: *ke* [ke] ‘general subordinating conjunction’, *do* [do] ‘two’.

2.1.1b Tajik

The Tajik central and back vowels are rounded: *ū* (orthographically <ū>) is more close than *u*, and slightly lower; *o* is more open than the Russian *o*, and without any *w*-glide.

The older phonemic contrast of length is now replaced by a contrast between stable and unstable vowels. The stable vowels, which are phonetically invariant, are *e, ū, and o*. The unstable vowels, in which the length and quality of articulation vary according to the phonetic environment, are *i, a, and u*. Thus in stressed position and unstressed closed syllables they are equivalent in length to the three stable vowels; in unstressed open syllables they may be shortened and reduced to *ə* or elided, e.g. *did* ‘(s)he saw’, *d’gar* ‘other, else’; *dud* ‘smoke’, *g^udoz* ‘melting’; *bad* ‘bad’, *b^odan* ‘body’. (In some dialects, e.g. of Bukhara, phonemic length is preserved in *ilī* and *ulū*.) The vowel *i* is lowered to *e*, and *u* to *ū*, before *h* or the glottal stop and a consonant: *istehsol* ‘production’, *mone* ‘hindrance’ (cf. *istiloh* ‘idiom, term’, *mohir* ‘skilled’); *mūhtaram* ‘respected’ (cf. *mustaqil* ‘independent’).

Note that in the following *ū* is transcribed as *ū*, and finally stressed *i* by *ī*, which reflects the Cyrillic-Tajik orthography.

2.1.2 Consonants

TABLE 8.3: CONSONANTS

	<i>Labial</i>	<i>Dental</i>	<i>Palatal</i>	<i>Velar</i>	<i>Uvular</i>	<i>Glottal</i>
Stops/Affricates						
tense/voiceless	<i>p</i>	<i>t</i>	<i>č</i> [tsh]	<i>k</i>	<i>q</i> (TAJ)	ʔ
lax/voiced	<i>b</i>	<i>d</i>	<i>j</i> [dzh]	<i>g</i>	<i>q~</i> (PERS)	
Fricatives						
tense/voiceless	<i>f</i>	<i>s</i>	<i>š</i>	<i>x</i>	<i>~γ</i> (PERS)	<i>h</i>
lax/voiced	<i>v</i> (PERS) <i>v~w</i> (TAJ)	<i>z</i>	<i>ž</i>	<i>γ</i> (TAJ)		
Nasals	<i>m</i>	<i>n</i>				
Liquids/Glide		<i>l, r</i>	<i>y</i>			

~ = conditioned variation

2.1.2.1 Stops and fricatives

The distinctive feature of the pairs of stops and fricatives in contemporary Persian is still being debated. It may be identified either as voice or as tenseness, as in this essay.

Tense *p t k* are slightly aspirated word-initially [k^hār] ‘work’.

Fronting of *k* and *g* before *i e a* is typical for Persian speakers of Azeri Turkish background, as it is in Tajik under the influence of Uzbek: Pers./Taj. [k’ar] ‘deaf’.

2.1.2.2 Velar γ and uvular q

This phoneme has triple origins: (1) the inherited lax fricative γ ; (2) the Arabic voiceless stop q and the voiced fricative γ ; and (3) the Turkic voice-neutral velar stop k before back vowels. – In Tajik, the opposition between q and γ is retained.

In Persian, q and γ have merged into a lax consonant: while systemically a lax fricative in the system, its peculiar articulation appears like a virtual compromise of its origins: (1) intervocalically it is a lax velar fricative, [γ]; (2) in initial and final position it is a lax uvular stop [q], which often has an affricate-like release before vowels (varying with the speaker and speech-situation); (3) in spoken Persian, it tends to be assimilated to a tense stop or fricative, *eqtesād* [extesād] ‘economy’, *raqs* [raxs] ‘dance’, *vaqt* [vaxt] ‘time’.

2.1.2.3 Glottal stop

The glottal stop has a triple origin: (1) the inherited automatic feature before initial vowels and in hiatus, [ʔinʔin] ‘this’, [pāʔiz/poʔiz] ‘autumn’, affixal [xānē-ʔi/xona-ʔe] ‘a house; (2) in Arabic loans, (a) the phonemic glottal stop: *soʔāl/suʔol* ‘question’ and (b) the pharyngeal stop *ayn*: *maʔnālmaʔno* ‘meaning’.

Note that in the transcription here the initial glottal is not marked.

2.1.2.4 Palatal \check{z}

This is of limited distribution, occurring in a few Persian words such as *žāležola* ‘dew’, *moželmiža* ‘eyelash’. – In Tajik, it also occurs in Russian loans; in spoken Tajik it tends to be replaced by j . (This is also a feature of Uzbek, standardized under a single grapheme; it is noteworthy that the Tajik and Uzbek sound systems are virtually identical.)

2.1.2.5 $v \sim w$

In Persian this labiodental is bilabial only in the diphthong, *au* > *ow* before consonant and pause. – In Tajik there is conditioned variation: intervocalically v is realized as bilabial w in (1) final position, (2) between rounded vowels, and (3) pre-consonantly after a : *vara \check{q}* ‘page’, *beva* ‘widow’, but *gow* ‘cow’, *suwol* ‘question’, *qawl* ‘speech’ (Taj. orthographically, <qavl>), corresponding to Pers. *gāv*, *soʔāl*, *qowl*. Afghan Persian has retained bilateral w in all environments (note that Persian and Tajik v < w).

2.1.2.6 Lateral l

Arabic rounding of l may be found in *allāh* and *yallā* ‘let’s go’.

2.1.2.7 Final position

Lax stops tend to lose release in final position, while in some dialects of Tajik b d g tend to be devoiced.

Similarly nasals and laterals tend to lack release in the environment CC-pause: [*satL*] ‘pail’, [*hokM*] ‘order, judgment’.

2.1.3 Syllable structure

Given the phonemization of the glottal stop, the syllable structure is predictable: CV, CVC, CVCC: *tāto* ‘till’, *darldar* ‘in’, *dast/dast* ‘hand’. Accordingly, the syllable boundary

is any consonant before a vowel: *dast-am* [*das-tam*] ‘my hand’, and diphthongs have the structure VC: *eyvānlayvon* [CVCCVC] ‘porch’, *qowmlqawm* CVCC ‘extended family, clan’.

2.1.4 Clusters

In clusters, tenseness tends to be neutralized: [moʃkel] ‘difficult’.

2.1.4.1 Initial clusters

The inherited initial clusters have been resolved by prothetic or epenthetic vowels, either of which could become standardized, e.g. *st-*: *stār* ‘star’ > *setārelsitōra*, *br-*: *brādar* ‘brother’ > *barādar/barodar*, but also *brū* ‘brow’ > *abrulabru*. Borrowings of an aberrant structure are assimilated by the introduction of epenthetic or prothetic vowels (to resolve initial consonant clusters), e.g. Pers. *estudiyo* ‘studio’, *teren* ‘train’, Taj. *istansa* < Russ. *stantsiya* ‘station’.

In Tajik, Russian borrowings have added alien initial clusters to the corpus, as *zveno* ‘team, unit’, *Stalinobod* (former name of Dushanbe). Such loanwords have been reproduced in their original orthographies (regarded as normative) in successive writing systems; the degree to which speakers assimilate them to native phonotactic and phonetic norms depends partly on non-linguistic factors such as their level of education and cultural predilections.

2.1.4.2 Final clusters

The Arabic impact brought a large number of final clusters that contravened Persian phonotactics, especially those involving obstruents and their clusters, notably in final position, e.g. *qotblqub* ‘pole’, *majdlmajd* ‘glory’. These were integrated into the Persian system.

In Tajik one finds a common “Eastern” feature whereby medial or final consonant clusters may be epenthized: *šukrullo* ‘thank God’, but *šukʳ* ‘thanks’, or be reduced.

2.1.4.3 Gemination

This is a distinctive characteristic of Arabic, whereas in inherited Persian items it is a marginal feature. While retained in high registers, it is eliminated in contemporary standard pronunciation, e.g. *korre* ‘foal’ vs. *kore* ‘globe’ (Ar.), *mat(t)e* ‘(mechanical) drill’, *banā* ‘building’ (Ar.) vs. *bannā* ‘architect’ (Ar.)

2.1.4.4 Glottal stop, *h* and *r*

Other than in high register or in slow articulation the glottal stop and *h* tend to result in the compensatory lengthening of the preceding short vowel: Pers. *baʳd* > [ba:d] ‘after’, or in occasional compensatory germination of continuants: Pers. *jomʳé* > [jom(:)e] ‘Friday’.

Similarly, *h* tends to be lost in postconsonantal and final position: Pers. *sobh* > [sob] ‘morning’, [sob-e zud] ‘early morning’.

In colloquial Persian, final *r* tends to be elided in final position in some high-frequency items such as *agar* > [age] ‘if’, *digar* > [dige] ‘other; moreover’.

In Tajik, particularly in southern dialects, there is a tendency to drop *h* in all positions, and to drop *r* before *d* in a few common words, especially in the past stem of the common auxiliary *kardan*, *kad-um* ‘I did’, etc.

This loss Tajik has in common with Afghan Persian, where the /a/ in this environment is additionally lowered to /a/: *ba’d* > *lba:dl* ‘after’, *šahr lša:r l* ‘town’.

2.1.5 Hiatus

2.1.5.1 Standard Persian and Tajik

No two adjacent vowels are allowed Hiatus breakers are (1) *-y-*, or (2) a glottal stop with the copula and with both indefinite and derivational *-i-*:

(1) Pers. *xāne-ye tolxona-yi to* ‘your house’, *qāli-ye bozorg/qoli-yi buzurg* ‘large carpet’; *na-y-ālna-y-o* ‘don’t come!’, less distinct when combined with *i*: *mi-ā-y-andlme-^o-o-yand* ‘they come’, *mi-ā-ⁱimlme-^o-o-em* ‘we come’, *bi-ā-lbi-^o-o* ‘come!’

(2) *xastelxasta -am, -i, -ast*, etc. ‘I, you, (s)he am, are, is, tired’, etc.; indefinite *-il-e*: *jā-iljo-e* ‘some place’; and derivative *-i*: *kojā-ilkujo-i* ‘where from’.

For intervocalic *-g-* see section 2.3 Morphophonemic alternations.

2.1.5.2 Colloquial absence of hiatus

The absence of the hiatus breakers *-y-* and glottal stop, coupled with contraction and loss of postconsonantal *h* and *-rā* > *-ro* > *-o*, is a major feature that distinguishes less formal registers of Persian (see section 2.1.5.2):

<i>ketāb-hā-yam</i> >	<i>hame-y-aš</i> >	<i>hame-y-aš-rā</i> >	<i>raft-e-i</i> >
<i>ketāb-ā-m</i>	<i>hama-š</i>	<i>hama-š-o</i>	<i>raft-i</i>
‘my books’;	‘all of it’;	‘all of it (dir. obj.)’;	‘you have left’.

2.1.5.3 Persian final *-e* as a dialect feature

Where Persian has final stressed *-e* all other varieties of Persian have *-a*, including Tajik and Dari and Kaboli of Afghanistan: Pers. *hame*, Taj. *hama*, Kab. *hama*. The examples above show that in suffixation Persian colloquial has retained the common Persian feature and not been subject to this dialectal feature of Standard Persian.

2.2 Non-segmental features

(1) Stress is word-final in nominals, including nominal verb forms: *baččelbačā* ‘child’, *did-dan/did-an* ‘to see’. Trisyllabics have secondary initial stress: *bačče-g-ilbača-g-i* ‘childhood’, *did-dan-il did-an-i* ‘worth seeing’. Transparent compounds may exhibit one or more secondary accents: *kam-sāllkam-sol* ‘juvenile’, Pers. *bonyād-gozāri* ‘founding’, trisyllabic *mājerā-ju* ‘adventure seeker, adventurous’; Taj. *kam-bayal* ‘poor (person)’ (*kam* ‘little’ + *sāllsol* ‘year’, *bayal* ‘armful’), Taj. *no-tavon-binī* ‘envy’ (*no-tavon* ‘not-able’ + *binī* ‘seeing’).

A number of common adverbials and interjections have initial or penultimate stress: *balelbale*, *ārelore* ‘yes’, *balkelbalki* ‘but (rather)’, *xeylilxele* ‘very, a lot’, *ammālammo*, *valilvale* ‘but’, *ya’nilya’ne* ‘i.e., ‘I mean’, *yāloyo* (interrog.), *zīrālzero* ‘because’, *hattā/hatto* ‘even’, but Pers. *albatte* vs. Taj. *albatta* ‘of course’, Pers. *hālā* vs. Taj. *holo*, *hole* ‘now’; further, Taj. *māylāš* ‘Good, OK’, Taj. *kānī* ‘well, then’; and conjunctions compounded

with the complementizer *kelki*: *har-kelhar-ki*. Adpositions and enclitics of all kinds are unstressed.

(2) Stress in verb forms is basically regressive, though less so in Tajik than in Persian. The subjunctive-imperative prefix *be-lbi-*, the imperfective prefix *mi-lme-*, and the negative prefix *na-*, are always stressed, the latter assuming primary stress: *mī-rav-am/lmē-rav-am* 'I won't go', *ne-mi-rav-am/na-me-rav-am* 'I won't go'. With *be-lbi-*, Pers. *be-gir-id* 'take-2p!', Tajik differs in having primary stress on the final syllable with secondary stress on the prefix: *bi-gir-éd*. Similarly, stem-initial past tense forms have stress on the stem syllable in Persian, but in Tajik may stress either stem or the (personal) ending according to context: Pers. *gofit-im/Taj. guft-em ~ guft-em* 'we said (it)'. The original stress on the personal ending in un-prefixed present forms is retained in the definite future: *xāh-am raft* 'I will go'.

2.3 Morphophonemic alternations

Synchronically there is a morphophonemic alternation between word-final *-el-a* and *-egl-ag* before the stressed suffixes *-ān/-on* (alternate plural marker), *-ānel-ona* (adjectival) and *-il-ī* (nominal and adjectival): *baččelbača* 'child' > *bačče-g-ān lbačag-on* 'children', *bačče-g-ānel/bača-g-ona* 'childish, children's', *bačče-g-ilbača-g-ī* 'childhood'. Unlike Persian, Tajik has also developed a derived participle: *kard-a* > *kard-a-g-ī* 'done' (see section 3.2.2.2).

In Tajik, this *g* is now intuited as euphonic, and is supplied even for words in *-a* of non-Persian origin (*talaba* 'student', pl. *talaba-g-on*). Note also *yak-to* 'single', *yak-to-g-ī* 'singly', while in Persian such derivation from words of Arabic origin is rare): *qā'ede* 'rule; menstruating' > *qā'ede-g-i* 'menstruation'.

Related, but irregular and rare, is suffixation of the generalizing plural suffix *-j-āt* to final vowel: *ruz-nāme* 'newspaper' > *ruz-nāme-j-āt* 'the printing media', *sabzi* 'greens, vegetable' > *sabzi-j-āt* 'vegetables' in the generic sense.

Diachronically, *-eg* originates in the Middle Persian suffix *-ag*, while the alternation *-i ~ -i-j-āt* originates in Middle Persian suffix *-īg*, which was loaned early into Arabic, with Pers. *j > g*, and, with the abstract feminine plural marker *-āt*, was later re-loaned into Persian, where it is moderately productive, where *g* thus still underlies the stem form.

In fact, most conspicuous is the complex morphophonological alternation that is inherent in the morphology of the massive Arabic loan component (see section 7.2.1).

Less evident, and frozen, is the complex Indo-European type morphophonology of Old Persian which is still evident in ablaut series like *bār* 'load' > *bar-* 'bear, carry' > *bor-d* 'carried' (< *bṛ-tá*, from the OIr. verb *bar-* 'bear, carry'), but is largely obscured (see section 3.2.1.1).

3 MORPHOLOGY

3.1 Nominal morphology

Substantives are weakly divided between nouns and adjectives; there are no characteristic stems and no case inflections. Many adjectives and primitive nouns are identical (*javānl javon* 'young; a youth', *bālābolo* 'high; top' – also an adverb, 'up'), and derived adjectives may often function as nouns (*Buxoro-ī* 'Bukharan; a Bukharan', *honar-mand/hunar-mand* 'skillful; a craftsman').

Gender is not marked in nouns or pronouns. Sex may be distinguished lexically : *morgl mury* 'fowl, hen', *xorus/xurüs* 'rooster'; more generally by selective use of a qualifier: *boz/buz* 'goat', *nar-boz/nar-buz* 'billy goat', *māde-boz/moda-buz* 'nanny goat'; *xar-e nar/xar-i nar* 'male donkey', *xar-e mādel/xar-i moda* 'female donkey'. Borrowed Arabic and Russian feminine nouns are used in context: *raqqāselraqqos-a* '(female) dancer', Taj. *studentka* '(female) student', including the feminine marking in abstract terms such as *xārej-e*, in terms like *vezārat-e omur-e xārej-e* 'ministry of foreign affairs', and female first names: m. *Tāher*, f. *Tāhere*.

Animate vs. inanimate distinction is found in subsets of interrogative-indefinite pronouns (see section 3.1.3.7a)

3.1.1 Nouns

Nouns are distinguished by having the categories of number and of definiteness ~ indefiniteness, and by syntactic status, i.e. their ability to combine in an *ezāfelizofat* (EZ) construction as both heads and modifiers by means of the particle *-el-i* and to take adpositions. Since this construction is so fundamental to the elucidation of nominal categories in Persian, a short preliminary definition is offered here (see section 4.2 Ezāfe):

(1) In a nominal EZ the modifier is a noun, pronoun or NP designating the possessor, matrix, source, agency or other relationship with the head: *dar-e xānel/dar-i xona* 'the door of the room', *xolāse-ye dāstān/xulosa-i doston* 'the gist of the tale', *raftan-e ul raftan-i vay* 'his departure'.

(2) In an adjectival EZ the modifier is an adjective or adjectival phrase qualifying the head: *havā-ye xub/havo-i nayz* 'nice weather', *čašm-ān-e por az askl/čašm-on-i pur az ask* 'eyes full of tears'.

3.1.1.1 Number

3.1.1.1a *-hāl-ho* and *-ānl-on*

Singular nouns are zero-marked, plurals by the suffixes *-hāl-ho* and *-ānl-on*. The former is universally applicable: *čiz-hāl/čiz-ho* 'things', *pesar-hāl/pisar-ho* 'boys', *asb-hāl/asb-ho* 'horses', *ma'nā-hāl/ma'no-ho* 'ideas'; *-ānl-on* is more limited in application and less frequent, denoting (1) animates: *pesar-ānl/pisar-on*, *asb-on*, *mo'allem-hālmualim-on* 'teachers'; (2) paired parts of body: *lab-ānl/lab-on* 'lips', *dast-ānl/dast-on* 'hands', etc. – a relic of the ancient dual number; (3) a few other words: *deraxt-ānl/daraxt-on* 'trees'; *setārelsitara* 'star', *setāre-g-ānl/sitora-g-on* 'stars'; it is stylistically marked as literary in register, and relatively more frequent in Tajik than in Persian.

Morphophonemic variants occur after *-āl-o* (*dānā-yānl/dono-yon* 'sages'). There is a frozen plural of *niyālniyo* 'forefather' in *niyā-k-ānl/niyo-k-on*.

3.1.1.1b Arabic loan plurals

The Arabic loan component includes a large number of plural forms:

(1) Arabicate suffixes in *-āt/-ot*, *-e-j-āt/-a-j-ot*, *-in/-in*: *xiyāl-āt/xayol-ot* 'dreams, fantasies'; *qal'e-j-āt/qal'a-j-ot* 'fortresses'; *mo'allem-in/mu'allim-in* 'teachers'.

(2) Most distinct are the Arabic "broken plurals": *olamālulamo* 'religious scholars', sg. *ālem/olim*; *axbār/laxbor* 'news', sg. *xabar/xabar* 'information, news'. For lexicalization of such plurals, see section 7.2.1 Loanwords.

3.1.1.2 Indefinite *-il-e* and *yeklyak*

Indefiniteness may be marked by the (unstressed) enclitic *-il-e*. The preposed numeral *yeklyak* ‘one’ may also function to mark indefiniteness. Both may combine, mostly in colloquial speech.

<i>mard-i āmadl</i>	<i>yek mard āmadl</i>	<i>yek mard-i āmadl</i>
<i>mard-e omad</i>	<i>yak mard omad</i>	<i>yak mard-e omad</i>
‘some man came’;	‘one/a man came’;	‘some/a man came’.

While most frequently found with singular nouns, it likewise occurs with plural nouns (see also section 5.2.1 Definiteness and specificity). Tajik also allows plural without enclitic *-i*:

<i>yek mard-hā-i āmad-andl</i>	
<i>yak mard-ho-e omad-and</i>	<i>yak piyola-ho</i>
‘some (or a certain) men came’;	‘some cups’.

Note that the admission of the plural marker distinguishes this function of *yeklyak* from its role as a numeral.

3.1.1.3 Tajik indefinite determiners

In Tajik, there are also two explicitly indefinite determiners, *yag-on* and *kadom (yak)*. These determiners may also be applied to plural NPs. (Note that this function of *kadom* is distinct from its function as interrogative-indefinite ‘which’, which determines a specific NP):

<i>yag-on rūz ba xona-i</i>	<i>mo marhamat kun-ed</i>	<i>yag-on odam-ho-e</i>
some dayto house-EZ us kindness do-p		some persons-INDEF
‘please come and visit us some day’;		‘some persons or other’;
<i>kadom yak zan-e</i>	<i>did-am</i>	
some one woman-INDEF I saw		
‘I saw some woman or other’.		

3.1.1.4 Persian referential *-e*

Colloquial Persian stressed *-e* (*-h-e* after vowel) has referential function to a person or an item spoken about by the speaker (similar to colloquial English ‘this ... here, that ... there’);

<i>pesar-e</i>	<i>čiz-i</i>	<i>na-goft</i>	<i>bačče-h-e</i>	<i>če</i>	<i>mi-xā-d</i>
boy-REF	thing-INDEF	did not say	child-REF-what	wants	
‘this/that boy didn’t say a thing’;			‘what does this/that child want?’.		

3.1.2 Adjectives

Adjectives may be used attributively, predicatively, and adverbially. In no case do they vary for number, or show agreement with any other of the properties of the head noun, unless themselves substantivized:

in gol-hā sefid astlin gul-ho safed-ast
‘these flowers are white’.

3.1.2.1 Subsets

3.1.2.1a Substantivized adjectives

Substantivized adjectives behave like nouns, but show animacy distinction in the plural, animate *-ān* vs. inanimate *-hā*: *digar-ān/degar-on* ‘the other people’ vs. *digar-hā/degar-ho* ‘the other things’.

<i>zibā-yān-e</i>	<i>hāliwudl</i>	<i>[sefid-hā-yaš]</i>	<i>zibā-stl</i>
<i>zebo-yon-i</i>	<i>holivud</i>	<i>[safed-ho-š]</i>	<i>zebo-st</i>

‘the beauties-EZ Hollywood’; ‘the white ones are beautiful’, lit.
‘the white [ones] of it’; note the sg. verb.

3.1.2.1b Morphological subsets

- (1) Several morphologically distinct classes of adjectives, often substantivized, are made up of verbal participles (see section 3.2.2 Nominal forms).
- (2) Another large notable subset is represented by Arabic loan adjectives as well as active and passive participles.

3.1.2.2 Comparative

The comparative degree is marked by the stressed suffix *-tar/-tar*: *bozorg-tar/buzurg-tar* ‘greater, bigger’. Suppletive are (1) Pers. *xub* > *beh-tar* ‘good > better’, but Taj. *xub-tar*; (2) *xeylil/xele* > *biš(-tar)/beš(-tar)* ‘much, very > more’.

The standard of comparison is introduced by the preposition *az/az* ‘from’: *az ān bozorg-tar/az on buzurg-tar* ‘bigger than that’.

In attributive position the comparative is connected with the noun in the EZ construction like the positive; in predicative position, the adjective usually precedes the copula or verbs of becoming (but may also follow the verb):

Attributive EZ	Predicative
<i>ketāb-e bozorg-tar [az ān]</i>	<i>in boluz az ān [beh-tar] astl</i>
<i>kitob-i [az on] buzurg-tar</i>	<i>in kurta az on [xub-tar]-ast</i>
book-EZ from that bigger	this blouse from that better is
‘the bigger book’;	‘this blouse/tunic is better than that’.

In colloquial Tajik the positive may also be used in a comparative function in predicative position, similar to Turkic:

in kurta az on [xub] as
this tunic from that good is
‘this tunic is better’.

3.1.2.3 Superlative

The superlative degree is marked by *-tar-in/-tar-in*.

In an attributive NP the superlative usually precedes the noun in both Persian and Tajik, but may also follow in the EZ construction in which case the superlative is

expressed as a universal comparative, using the phrase *az hamelaz hama* ‘than all’. In Tajik, the superlative may follow the head similar to the positive and comparative:

boland-tar-in kuhl kuh-e az hame boland-tarl
baland-tar-in kūh kūh-i az hama baland-tar kūh-i baland-tar-in
 ‘the highest mountain’.

In predicative position the superlative is similarly expressed as a universal comparative:

in kuh az hame boland-tar astl
in kūh az hama baland-tar ast
 ‘this mountain is higher than all = the highest’.

ELLIPTIC SUPERLATIVE

The elliptic substantivized superlative may head a partitive EZ construction, or a pronominal clitic, denoting the class of things being compared:

[boland-tar-in]-e kuh-hā-ye donyāl bozorg-tar-in-ešānl
[baland-tar-in]-i kūh-ho-i dunyo buzurg-tar-in-ašon
 highest-EZ mountains-EZ world’; biggest their
 ‘the highest mountains of the world’; ‘the biggest one of them’.

3.1.2.4 Intensive reduplicated and echoic superlative

Adjectives may form a reduplicated EZ construction to express superlative grade, ‘exceedingly’, etc. *xub-e xub* ‘totally good’, *sabz-e sabz* ‘very, thoroughly green’.

In Tajik, similar to Turkic, some qualitative adjectives (esp. of color) form an intensive by addition of a stressed pre-echoic syllable, i.e. a duplicate of the adjective’s first syllable plus a labial, voiced or unvoiced as appropriate: Taj. *sap-safed* ‘snow-white’, *sip-siyoh* ‘jet black’, *top-torik* ‘dark as dark (could be)’.

3.1.3 Pronouns and deixis

3.1.3.1 Personal pronouns and deixis

Personal pronouns are of two kinds: independent and enclitic (colloq. = colloquial):

TABLE 8.4: PERSONAL AND DEICTIC PRONOUNS

	1s	2s	3s		1p	2p	3p
INDEP.	<i>manl</i>	<i>tol</i>	<i>ul</i>	<i>veyl</i>	<i>māl</i>	<i>šomāl</i>	<i>išānl</i>
	<i>man</i>	<i>tu</i>	<i>ū</i>	<i>vay</i>	<i>mo</i>	<i>šumo</i>	<i>ešon</i> <i>vay-o</i>
Colloq.					<i>mā-hāl</i>	<i>šomā-hāl</i>	
					<i>mo-ho(n)</i> ,	<i>šumo-ho</i> ,	
					<i>mo-yon</i>	<i>šumo-yon</i>	
Far			<i>ānl</i>				<i>ān-hāl</i>
			<i>on</i>	<i>vay</i>			<i>on-ho</i>
Near			<i>inl</i>				<i>in-hāl</i>
			<i>in</i>				<i>in-ho</i>
ENCLITIC	1s	2s	3s		1p	2p	3p
	<i>-aml</i>	<i>-atl</i>	<i>-ašl</i>		<i>-emānl</i>	<i>-etānl</i>	<i>-ešānl</i>
	<i>-am</i>	<i>-at</i>	<i>-aš</i>		<i>-amon</i>	<i>-aton</i>	<i>-ašon</i>

3.1.3.1a Register and animacy

1p *mālmo* may refer to the speaker, both as the plural of majesty and as a member of a social group. In polite or formal speech, 2p *šumāšūmo* is also used to address an individual. In turn, the colloquial pluralized forms in *-hāl-ho*, Taj. also with alternate *-on*, function to refer to more than one person.

The 3rd person personal pronouns are marked for human animacy, but not gender: *ulū* ‘he, she’, *išānlešon* ‘they (human)’, in Persian also polite for an individual.

In Tajik, the 3rd person is augmented by the demonstratives *in* and *on*. *On-ho* may also refer in respectful usage to one person, and has replaced earlier *ešon* ‘they’ (human), which became an honorific and ultimately a common noun referring to religious dignitaries (*on-ho* may thus refer in respectful usage to one person).

Vej/vay. In Persian the pronoun *vej* is a highly marked polite 3s human alternate, mostly literary, and lacks a plural form, while in Tajik *vay* is both animate ‘he, she’ and inanimate ‘it’. It has the plural form *vay-o*, and may also function as an unmarked demonstrative adjective: *vay kor* ‘that matter’, *vay mard-ho* ‘those men’.

(Note also the animacy distinction in the indefinite-interrogatives *kilkī* ‘who’ vs. *čelčī* ‘what’, and *kas/kas* ‘person’ vs. *čiz/čiz* ‘thing’, and compound forms with them).

3.1.3.1b Attachment of object marker

The direct object marker *-rāl-ro* (see section 3.1.4.2) is affixed regularly in the literary language, except for the standard contraction 1s *ma-rālma-ro* (< *man-rā*). Spoken Persian and Tajik exhibit a variety of forms, notably postconsonantal *-rāl-ro* > *-ol-a*, including *man-olman-a*, *to-rolto-ra*, Taj. *vay-a*.

3.1.3.2 Possessive construction

Possession ‘mine, yours, his’, etc., is expressed by EZ constructions: Persian *māl-e*, lit. ‘possession of’, Tajik *az on-i* ‘from that of’ followed by an independent pronoun. *Az ān-i* and similar constructions are also found in local Persian dialects and earlier stages of Persian.

<i>in pul</i>	<i>māl-e</i>	<i>ki-st</i>	– <i>māl-e</i>	<i>mā-stl</i>
<i>in pul</i>	<i>az on-i</i>	<i>kī ast?</i>	– <i>az on-i</i>	<i>mo-st</i>
this money	property-EZ/ from that -EZ	who COP.3s	– property EZ/ from that-EZ	we COP.3s
	‘whose money is this?’		– it’s ours’.	

3.1.3.3 Enclitic pronouns and oblique

These interpose a euphonic *-y-* after vowels (*šahr-hā-yemān/šahr-ho-yamon* ‘our cities’), but in colloquial speech the initial short vowel of the personal clitics contracts with preceding vowels, as shown in the following example (combined with contracted *-rāl-ro*):

<i>dast-hā-yat-rā</i>	<i>be-deh-id!!</i>	>	<i>dast-ā-t-o be-d-id!!</i>
<i>dast-ho-yat-ro</i>	<i>dih-ed!!</i>	>	<i>dast-o-t-at-it!</i>
hand-p-your.s-DO	IMP.give-2p		
‘give me your hands!’.			

These enclitics function as oblique cases, including possession as in the example above, and as direct and indirect objects.

3.1.3.4 Demonstratives and quasi-articles

The demonstrative pronouns and demonstrative adjectives are identical in form; as adjectives they are preposed. In Persian there is a two-term deictic system, *inlin* ‘this’, *ānlon* ‘that’.

In Tajik there is a three-term system, the basic forms being *in* ‘this’, *on* ‘that’, as well as *vay* ‘that’; the third term is generally interchangeable with *on*, but may be more a rhetorical anaphoric word:

qin bud-agī-st vay kor
tough GERV-COP3s that work
‘must be a tough job, that’.

Inlin and *ānlon* are fairly weak demonstratives, and function virtually as definite articles.

3.1.3.4 Emphatic *ham-*

The demonstrative function is often intensified by prefixation of the emphatic particle *ham-* ‘same’, *ham-inl ham-in*, *ham-ānlham-on*:

ham-in rāh dorost astl rāh-e dorost ham-in astl
ham-in roh durust-ast roh-i durust ham-in-ast
this very road right-COP.3s road-EZ right this.very-COP.3s
‘this road is the right one’; ‘the right way is this one’.

In Tajik dialect also occurs *havay* (<*ham* + *vay*) ‘that very (one)’; other colloquial variants include *ī*, *amī*, *amu* (for *in*, *ham-in*, *ham-on*).

Compounds include: (*ham*)*in-jāl*(*ham*-)*in-jo* ‘right here’, (*ham*-)*ān vaxtl*(*ham*-)*on vaqt* ‘just then’; (*ham*-)*čon-inl*(*ham*-)*čun-in* ‘this very –, just like this, just so’, (*ham*-)*čon-ānl*(*ham*-)*čun-on* ‘that very one, etc.’; and idioms such as Pers. (*ham*-)*injur* ‘just this way, manner’, etc., Taj. (*ham*-)*in xel* ‘this sort’, (*ham*-)*on guna* ‘that kind’.

3.1.3.5 Reflexive-emphatic pronoun

3.1.3.5a Reflexive

The pronoun *xodlxud* is used possessively, ‘own’, reflexively, ‘self’, and emphatically, ‘-self’. When possessive and reflexive, it refers to the subject; when emphatic, there is no such restriction.

(1) In a possessive EZ phrase, it combines with a specifying enclitic pronoun:

medād-e xod-am šekastl
qalam-i xud-am šikast
pen-EZ own-my brake.PT.3s
‘my own pen broke’.

In the 3rd person, possessive *xodlxud* and *xod-aš*, *xod-ešānlxud-aš*, *xud-ašon* may substitute for enclitic *-aš*, *-ešānl-aš*, *-ašon* to resolve a possible conflict of scope:

barādar [-aš] -rā koštl
barodar [-aš] -ro kušt
brother his DO kill.PT.3s ‘he (A) killed his (A’s or B’s) brother’;

barādar-e [xod]-aš-rā košt
barodar-i [xud]-aš-ro kušt
 brother-EZ self-his-DO kill.PT.3s ‘he (A) killed his (A’s) brother’;

barādar-e [u]-rā košt
barodar-i [ū]-ro kušt
 Brother-EZ he-DO kill.PT.3s ‘he (A) killed his (B’s) brother’.

In this disambiguative use, *xod/xud* is not necessarily emphatic.

(2) As a reflexive pronoun it combines with the appropriate enclitic pronoun and adposition:

<i>xod-am-rā mi-šenās-aml</i>	<i>az xod-ešān mi-tars-andl</i>
<i>xud-am-ro me-šinos-am</i>	<i>az xud-ašon me-tars-and</i>
self-my-DO IPFV.know.PR-1s	from self-their IPFV.fear.PR-3p
‘I know myself’;	‘they are afraid of themselves’.

3.1.3.5b Emphatic

(1) When used emphatically, it precedes a noun or pronoun attribute in an EZ construction: *xod-e tol/xud-i to* ‘you yourself’, or with the addition of an enclitic pronoun: *xod-at/xud-at* ‘you yourself’:

Subject	Object
<i>xod-e mo’allem goftl</i>	<i>xod-e mo’allem-rā košt-andl</i>
<i>xud-i mu’allim guft</i>	<i>xud-i mu’allim-ro kušt-and</i>
self-EZ teacher say.PT.3s	self-EZ teacher-DO kill.PT-3p
‘the teacher himself said so’;	‘they killed the teacher himself’.

(2) In topical constructions, the noun or independent pronoun is placed at the head:

<i>mā xod-emān mi-dān-impl</i>	<i>mo’allem xod-aš goftl</i>
<i>mo xud-amon me-don-em</i>	<i>mu’allim xud-aš guft</i>
we self-our IPFV.know.PR-1p	teacher self-his say.PT.3s
‘we ourselves know’;	‘the teacher, he himself said so’.

3.1.3.6 Reciprocals

(1) The main reciprocal pronoun is *ye—digar/yak—digar* ‘one—another’ or *ham-digar/ham-digar* ‘each other’ (often with pronominal enclitic):

<i>yek-i pas az digar-i</i>	<i>ham-digar-ešān-rā mi-šenās-andl</i>
<i>yak-e pas az digar-e</i>	<i>ham-digar-ašon-ro me-šinos-and</i>
one-INDEF after from other-INDEF	each other-their-DO IPFV.know.PR-3p
‘one after the other/another’;	‘they know each other’; also <i>pas az yak digar</i> .

(2) The particle *ham/ham* ‘same, together’ may also function as a reciprocal pronoun synonymous with *yek digar/yak digar*, here with prepositions *bālbo* ‘with’ and *azlaz* ‘from’:

<i>[bā ham] raft-impl</i>	<i>[az ham] jodā</i>	<i>šod-andl</i>
<i>[bo ham] raft-em</i>	<i>[az ham] judo</i>	<i>šud-and</i>
together go.PT-1p	from each other separated	become.PT-3p
‘we went together’;	‘they parted’.	

3.1.3.7 Interrogatives, indefinites, and related adverbials

Animacy distinction is partially retained lexically in the basic interrogative and indefinite pronouns, and by the plural distinction animate *-ānl-on* vs. inanimate *-hāl-ho* of the indefinite pronominal determiners and of ‘other’.

3.1.3.7a Basic interrogatives and indefinite determiners

(1) Basic interrogatives

- (a) pronominal: animate *kilkī* ‘who?’, pl. *ki-hālki-ho*, inanimate *čelčē* ‘what?’, pl. *če-hālčē-ho*, *če-lča-* in compounds;
- (b) demonstrative: *kodāmlkadam* ‘which?’, pl. *kodām-hālkadam-ho*;
- (c) quantitative: *čandlčand* ‘how many?’;
- (d) adverbial: *keylkay* ‘when?’, *kojāлкуjo* ‘where’, and *kulku* ‘where is, are’ (lacking the copula);
- (e) causal: *čerālčaro* ‘why?’;
- (f) yes-no interrogative: *āyāloyo* ‘whether’ (see section 5.1.3 Questions).

(2) Basic indefinite determiners and compounds

- (a) pronominals: (aa) animate *kaslkas* ‘person, somebody’ (pl. *kas-ānlkas-on*, contextually with a connotation of persons of good or noble standing) and (ab) inanimate *čizlčiz* ‘thing, something’; (b) quantifiers: (ba) distributive *harlhar* ‘each’, (bb) *hičlheč* ‘any (at all)’ (interrogative or with negative *na-*), and (bc) collective *hamelhama* ‘all’, (bd) *čandlčand* ‘several, a few’, also *čandinlčandin*.

Examples:

ki budlkī bud ‘who was (it)?’; Pers. *če šode* ‘what happened?’, Taj. *čī gap-ast?* ‘what’s the matter?’; *kodām, har, hič, hame ruzlhar, heč, hama ruz* ‘which, every, no (not any), all day’; *čand ketāb dār-ilčand kitob dor-i* ‘how many books do you have?’, *čand(-in) bār/čand(-in) bor* ‘several times’; *key, kojā raft-ilkay, kajo raft-i* ‘when, where did you go?’; *hasan kul hasan ku* ‘where is Hasan?’; *čerā na-y-āmad-ilčaro na-omad-i* ‘why didn’t you come?’, *āyā dorost nist? loyo durust nest?* ‘isn’t that correct?’.

3.1.3.7b Pronominal quantifiers and demonstratives

The interrogatives and quantifiers function as pronouns under the following conditions:

(1) *hičlheč* functions as a pronoun by itself: *hič na-goft-amlheč na-guft-am* ‘I didn’t say anything’.

(2) *hamelhama* similarly functions as a pronoun: *hame āmad-andlhama omad-and* ‘all came’. It is also often used in EZ with a pronoun or noun, singular or plural, to express totality: *hame-yemānlhama-amon* ‘all of us’, *hame-ye donyālhama-i dunyo* ‘the whole world’, *hame-ye zan-hā-ye dehlhama-i zan-ho-i qišloq* ‘all the women of the village’ (It likewise functions as an adverb, see section 3.1.5 Adverbs).

(3) *kodāmlkadam* requires suppletion by a partitive EZ phrase or an enclitic: *kodām-e in-hālkadam-i in-ho* ‘which (one) of these/them?’; *kodām-ešān āmadlkadam-ašon omad?* ‘which (one) of them came?’ (human, pl. clitic); *kodām-aš-rā mi-deh-i?lkadam-aš-ro me-dih-i?* ‘which (of them) will you give?’ (non-human, sg. clitic); plural, *kodām-hā-yemānlkadam-ho-yamon?* ‘which (ones) of us?’.

(4) *harlhar* ‘every’ minimally requires a nominal when used pronominally: *har yeklhar yak* ‘every one’.

(5) *čandlčand* both as interrogative and non-interrogative determinative minimally requires a classifier: *čand tālčand to* ‘how many/several (items)’.

3.1.3.7c Compound indefinite forms

The basic interrogatives and basic indefinite determiners combine to form indefinites:

- (1) *hič-kodāmlheč-kadom* ‘not anyone, none’: *hič kas nistlheč kas nest* ‘there’s no one (here)’, *hič kodām-e ān-hālheč kadom-i on-ho* ‘none of them’;
- (2) *har-kas, -čizlhar-kas, -čiz* ‘each one (person, thing), whoever, whatever’; *hič-kas, hič-čiz* (colloq. *hiččī*)/*heč-kas, -čiz* ‘anyone, anything’ (with negative *na-*);
- (3) *hame-kas, -čizlhama-kas, -čiz* ‘everyone (all persons), everything (all things)’; *har-kodāmlhar-kadom* ‘every (single) one’;
- (4) *harlhar* with the pronominal interrogatives function as heads of generalized relative clauses: *har(-ān-)kelhar(-on-)kī . . .* ‘who(so)ever . . .’, *har-čelhar-čī . . .* ‘whatever . . .’; *har kojālhar kujo . . .* ‘wherever . . .’, etc.

In addition, there are compounds with *yeklyak* ‘one’: *har-yeklhar-yak* ‘everyone’, *hič-yeklheč-yak* ‘no one’, *kodām-yeklkadom-yak* ‘which one?’, *yek-čandlyak-čand* ‘a few’: *yek-čand nafarlyak-čand nafar* ‘a few people’ (with regular singular after numbers).

The indefinite clitic *-il-e*, in colloquial Persian combined with *yek*, is frequently added to the pronominals: *kodām kas-ilkas-e* ‘which one (person)’; *(yek) kas-ilkas-e, (yek) čiz-ilčiz-e* ‘something’. It is similarly added to the numeral *yeklyak* ‘one’ and several common nouns which thereby may function as indefinite pronominals: *yek-ilyak-e, šaxs-ilšaxs-e* ‘someone’ (‘person, individual’).

In terms of specificity, while the indefinite series *kas-ilkas-e, čiz-ilčiz-e*, etc., may be either non-specific (‘someone or other’) or specific (‘a certain person’), *kodāmlkadom* in its function as an interrogative determines a specific NP, with obligatory *-rāl-ro*:

<i>kodām zan-rā</i>	<i>did-id?!/</i>
<i>kadom zan-ro</i>	<i>did-ed?</i>
which woman-INDEF-DO see.PT-2p	
‘which woman did you see?’	

3.1.3.7d Other indefinites

- (1) Other pronominal indefinites include: *folān/falon* used for a person or thing not specifically named, ‘such-and-such, so-and-so’: *folān ruzl falon rūz* ‘on such-and-such day’, *āqā-ye folānloqo-yi falon* ‘Mr. so-and-so’, *ey folān! lay falon!* ‘hey, you there’, *folān-kasl falon-kas* ‘such-and-such a one, what’s his name?’, and low-respect *yāru* ‘that fellow’, lit. ‘friend’.
- (2) Other qualitative indefinites include (mostly literary) *čen-in, -ān ~ čon-in, -ānlcun-in, -on* ‘such’: *čen-ān kārčun-on kor* ‘such work, doing’, also used adverbially: *čon-ān kardlčunon kard* ‘he did thus’.
- (3) Other quantifiers include *xeylilxele* ‘much, many’, *kam* ‘little, few’, *kam-tar* ‘less, fewer’: *kam-tar kas-ilkam-tar kas-e* ‘only a few, fewer (persons)’, etc.
- (4) The determiner ‘some’ is expressed by the Arabic loan *ba’z-ilba’z-e* preposed to a plural noun: *ba’z-i Tājik-hālba’z-e Tojik-on* ‘some Tajiks’ (note animate plural *-on*).

3.1.3.7e *Compounds with čē-lča-*

Other adverbial determiners are compounds with interrogative *čēlčī* plus semantically appropriate terms, here exemplified with interrogative *čē-lčī* (literal meaning of the second component in parentheses):

- (1) time and location: *čē-vaqt/lčī vaqt* ‘when’ (‘time’); *čē-jā’ilčī jo-e* ‘where’ (‘place’);
- (2) quality and manner: (a) *čē now’lčī nav’* + N ‘what sort of’ (‘sort, species’), (b) *čē-gunelčī guna* ‘how’, + N ‘what kind of’ (‘kind, color’), (c) *čē-towr/lčī tavr* ‘how, what manner’ (‘manner, kind’), (d) Pers. *čē-jur* + N ‘how, kind of’ (‘kind’), (e) Taj. *čī xel: čē-gune hekāyat/lčī guna hikoya* ‘what sort of tale(s)?’, *čē now’ lebāsl čī nav’ libos* ‘what kind of clothing?’, etc.;
- (3) amount expressed by *-qad(a)rlqadar* ‘measure’: *čē-qadr mi-šav-ad l čī qadar me-šav-ad* ‘how much will it be?’, *čēqadr nānlčī qadar non* ‘how much bread?’.

3.1.3.7f *Digarldigar* ‘other, else’

‘Other, else’ is *digarldigar*, used as a noun and adjective both independently and in various constructions, including regular and inverted (preposed) EZ, the latter confined to fixed phrases in the modern standards:

- (1) independent use with obligatory animacy distinction *-ān* vs. *-hā*: *digar-ān* ‘the others (people)’, *digar-hā* ‘the others’;
- (2) regular construction, *mardom-e digarl mardum-i digar* ‘the other people’, *kas-i digarl kas-e digar* ‘someone else’, *yek-i digarl yak-e digar* ‘another one, someone else’;
- (3) preposed, *digar ki?ldigar ki?* ‘who else?’, *digar čī?ldigar čī?* ‘what else, what next?’.

Particular idioms include *ruz-e digarl ruz-i digar* ‘the following day’, *bār-e digarl bor-i digar*, *digar bārl digar bor* ‘once more, again; next time’.

3.1.3.7g *Tajik hybrids*(1) Turkic *kim-*

Assimilated into the Tajik literary language is a series of Uzbek-Tajik hybrids formed from Uzbek *kim* ‘who?’: *kim-kī* ‘someone (or other), anyone’, *kim-čī* ‘something, anything’, *kim-kadom* ‘some-N or other’: *vay az kim-čī no-rozi ast* ‘she’s unhappy about something,’ *kim-kadom vaqt* ‘sometime or other, whenever’, *dar kim-kujo-ho* ‘somewhere or other’. As opposed to the Persian series *kas-ilkas-e*, *jā’iljo-e*, *čiz-ilčiz-e*, etc., the *kim* series is unambiguously non-specific indefinite.

(2) Tajik indefinite determiners

In Tajik, there are also two explicitly non-specific indefinite determiners, *yag-on* and *kadom (yak)*:

[*yag-on*] *rūz ba xona-i mo marhamat kun-ed*
 some day to house-EZ us kindness IMP.do-2p
 ‘please come and visit us some day’.

These determiners may also be applied to plural NPs.

3.1.4 *Adpositions and case relations*

Case relations may be expressed by means of prepositions, postpositions, or ambipositions.

3.1.4.1 Prepositions

3.1.4.1a Primary prepositions

There are nine primary prepositions, one of which forms an EZ constructions (6), while one is partially (7) and another (9) nearly fully replaced by such construction in colloquial Persian.

(1) *belba* ‘to, in; with, by’ (dative, direction, manner, means)

This is the most frequent preposition with a wide range of meaning:

Dative: *be to mi-deh-amlba tu me-dih-am* ‘I’ll give it to you’; direction: *be šahr/ba šahr* ‘to the city’, manner: *be āsānilba osoni* ‘with ease’, lit. ‘easiness’; means: *be (zabān-e) fārsi harf be-zan!/ba (zabon-i) tojikī gap zan!* ‘speak in (the language of) Persian/Tajik!’.

(2) *dar/dar* ‘in(to)’ (locative)

In Persian colloquial speech this preposition in the sense of “inside” is frequently often replaced by *tu(-ye)*, lit. ‘the inside’, in an EZ construction; in Tajik dialects it tends to be reduced to *da*: *dar ānjāldar on jo* ‘in that place, there’, *dar tābestāndar tobiston* ‘in the summer’, *dar (~ tu-ye) sanduq gozāštdar sanduq guzošt* ‘he put into the box’; idiomatic: *dar in suratldar in surat* ‘in this case’, *do dar doldu dar du* ‘two by two’.

Omission of *be* and *dar* in Persian. In the locational and directional sense these two may be omitted: (*dar*) *tābestān xeyli garm mi-šav-ad* ‘it gets quite warm in the summer’, (*dar*) *tehrān zendegi mi-kon-ad* ‘he lives in Tehran’; *man (be) sinemā mi-rav-am ~ mi-rav-am (be) sinemā* ‘I’m going to the movies’.

(3) *az/laz* (*a* in Taj. dialects) ‘from, than; along’ (ablative, source, partitive; passage)

Ablative: *az injālaz injo* ‘from here’, *az emruzlaz imrūz* ‘from today’; than: *qeyr az tolyayr az tu* ‘other than you’, *bozorg-tar az tolbuzurg-tar az tu* ‘older than you’; source: *az tarslaz tars* ‘out of fear’; partitive: *por az talālpur az talo* ‘full of gold’; passage: *az pelle bālā raftlaz zina bolo raft* ‘he climbed the stairs’; *az rāh-e Tehrānlaz roh-i Dušanbe* ‘by way of Tehran/Dushanba’.

(4) *tālto* (up)to, until’ (goal), ‘for, as far/long as’ (extent, duration)

tā Tehrānlto Dušanba ‘till Tehrān/Dushanba’, *tā hālālto holo* ‘until now’; *tā andāze-lto andoza-e* ‘to an extent’; *tā dah ruz xabar-i na-dāšt-amlto dah ruz xabar-e na-došt-am* ‘for (a period of) 10 days I didn’t get any news’.

Note that with human objects the locational-directional prepositions require the adverbial noun *piš/peš* ‘front, before’ in the sense of “presence”: *dar, be az piš-e mādarldar, ba, az, to pēš-i mādar* ‘(being) with, (going) to, (coming) from mother’.

(5) *bālbo* ‘with’ (comitative, instrumental)

Comitative: *bā tolbo tu* ‘with you’, *bā ajalelbo ajala* ‘with, in haste’; instrumental: *bā čakošlbo čakuš* ‘with the hammer’, *bā asb raftlbo asb raft* ‘he left on horseback’.

(6) *barā-yelbaro-i* (benefactive, purposive, causative, in origin *ba rāy-e* ‘for the sake of’)

Benefactive: *in barā-ye to-stlin baro-i tu ast* ‘this is for you’, *barā-ye man tā’rif kon!! baro-i man naql kun!* ‘tell me!’, lit. ‘narrate for me’; purposive: *barā-ye čelbaro-i čī* ‘what for?’; causative: *barā-ye garmā* ‘because of the heat’, but not so used in Tajik.

Pers. colloquial also has *vāse-ye* (benefactive, purposive).

(7) *bilbe* ‘without’

While still active in Tajik, in Persian this preposition has been largely reduced to a privative preformative (see section 7.1.3 Word formation) and replaced by *be-dun-e* in an EZ construction: *bi ~ be-dun-e pedar o mādar/be pidar-u modar* ‘without father and mother’.

(8) *barlbar* 'on(to)' (locative)

This preposition occurs mostly in fixed phrasing:

banā bar inlbano bar, 'based on this, accordingly'. *bar akslbar aks* 'on the contrary'.

In Persian it has been mostly replaced by *ru-ye*, lit. '(sur)face of', in an EZ construction, *ru-ye mizlrū-yi mez* 'on(to) the table'. In Tajik it is still used in basic locative and directional phrases: *bar kūh barf bud* 'there was snow on the mountains'; *bar po istod* 'she stood up, got to her feet'.

(9) *čunlčün* 'like, similar to'

pesar-i čun tolčun tu pisar-e 'a son like you'; *čun man o tolčun man u tu* 'like me and you'. This preposition is mostly replaced by the EZ construction.

3.1.4.1b Secondary prepositions

There are many more secondary prepositions.

(1) Common compound prepositions of the structure N + Preposition. These include *hamrāh bālhamroh bo* '(together) with (person)', *joz az* 'except' (also without *az/az*), *piš az/peš az* 'before', *pas az/pas az* 'after', and prominently those based on a borrowing from Arabic such as *qabl az/qabl az* 'before', *ba'd az/ba'd az* 'after', *ebārat az/liborat az* 'consisting of'; Pers. *rāje* be/Taj. *oid ba* (or *oid-i*) 'about, concerning', *qeyr az/ya'yr az* 'other than, except'.

The preposition *az/az* may alternate with the EZ without and with a difference in spatial relationship; thus both *birun-elberun-i* and *birun azlberon az* 'outside (of)', *birun-e ~ az šahr/berun-i ~ az šahr* 'outsidel/out of town'; Pers. *ba'd az* 'after', but Taj. *ba'd az ~ ba'd-i*. However, *piš/peš* 'front, before' when locational requires an EZ construction: *dar, be, az piš-e xāneldar, ba, az peš-i xona* 'at, to, from the front of the house', but *az/az* when temporal: *piš az zohrl/peš az zuhr* 'before noon'.

(2) There are numerous noun-adverbs or adjective-adverbs forming EZ constructions with the focal noun or phrase, in combination with a primary preposition, mainly *dar, be, az* (of which the first two tend to be omitted, particularly in colloquial):

rulrū 'on' ('face, surface'), *dar, be, az ru-ye divār/rū-yi devor* 'on, onto, (from) off the wall'; others include *miyānlmiyon* 'between' ('middle, waist'), *miyān-e man o tolmiyon-i man-u tu* 'between you and me'; (*dar*) *miyān-e kuh-hāldar miyon-i kūh-ho* 'amid the mountains'; Pers. *tu, darun/Taj. darun* 'in' ('interior'), *tu-ye, darun-e xāneldarun-i xona* 'inside the house'; *tah/tağ* 'under(neath), base of, end of' ('bottom'), *tah-e deraxl/tağ-i daraxt* 'underneath the tree'; *zirl/zer* 'under(neath)', *bālā/bolo* 'above' ('upper part') and *zirl/zer* 'below' ('lower part'), *bālā-ye, zir-e mizl bolo-yi, zer-i mez* 'above, under the table'; *mānand-elmonand-i* 'like', lit. 'resembling', commonly *mesl-elmišl-i N* ('likeness of'), *mānand-e, mesl-e tolmonand-i, mišl-i tu* 'like you'.

3.1.4.2 Postposition *-rāl-ro*

3.1.4.2a Specific direct object

The principal function of *-rāl-ro* is to mark a definite and any otherwise specific noun or noun phrase as the direct object:

<i>ketāb-rā</i>	<i>xarid-aml</i>	<i>doxtar-e</i>	<i>dust-eš-rā</i>	<i>didl</i>
<i>kitob-ro</i>	<i>xarid-am</i>	<i>duxtar-i</i>	<i>dost-aš-ro</i>	<i>did</i>
book-DO	buy.PT-1s	daughter-EZ	friend-his-DO	see.PT.3s
'I bought the book'; 'he saw the daughter of his friend'.				

In Persian colloquial, *-rā* > *-ro* > postconsonantal *-o*: *u-rā gereft-and* > *u-ro gereft-and* ‘they caught him’, *xāne-rā xarid-and* > *xuna-ro xarid-and* ‘they bought the house’ (with *xāne-* > *xuna-*); *in-rā be-gir* > *in-o be-gir* ‘pick up this one’.

In Tajik, *-ro* has the colloquial reflexes *-ra*, *-a*, *-ya*, this last after a vowel: *kūza-ya ovard-am* ‘I brought the jug’.

3.1.4.2b Temporal and local extension

In addition, *-rā* may mark focused temporal and local extension:

<i>[emšab]-rā</i>	<i>injā</i>	<i>bāš-idl</i>	<i>[in hame rāh]-rā</i>	<i>raft-andl</i>
<i>[imšab-ro]</i>	<i>injo</i>	<i>boš</i>	<i>[in hama roh]-ro</i>	<i>raft-and</i>
tonight-DO	here	be.IMP-2s/p	this all	way-DO
‘stay here for the night’;			‘they walked all this way’;	

<i>ru-ye</i>	<i>yax-rā</i>	<i>āb</i>	<i>rixt-and</i>
face-EZ	ice-DO	water	pour.PT-3p
Pers. ‘they poured water on the ice’.			

3.1.4.2c Tajik circumpositional *-ro*

In addition to marking the direct object, *-ro* (or rather its colloquial reflexes) can be found in several other uses. Thus, it may form circumpositions with nouns governed by prepositions:

<i>baro-i kī</i> >	<i>baro-i man</i> >	<i>az</i>	<i>xandidan</i>	<i>murd-im</i> >
<i>[baro-i] ·kī[-ra]?</i>	<i>man[-a]</i>	<i>[a (= az)]</i>	<i>xandidan[-a]</i>	<i>murd-em</i>
sake-EZ	who-for	sake-EZ	I-for	from
‘for whom?’	– for me’;			laugh.INF-for
				die.PT-1p
				‘we died from laughing’.

3.1.4.2d Northern Tajik EZ with *-ro*

In Northern dialects of Tajik, a construction using *-ro* widely replaces the Persian type of EZ: The word order is that of the equivalent Uzbek NP *muallim[-ning] kitob-i*, lit. ‘of-the-teacher his book’:

<i>pisar-i</i>	<i>man</i> >	<i>man[-a]</i>	<i>pisar-am</i>	<i>kitob-i</i>	<i>muallim</i> >	<i>muallim [-a]</i>	<i>kitob-aš</i>
son-EZ	me	I-for	son-my	book-EZ	teacher	teacher-for	book-his
	‘my son’;				‘the teacher’s book’.		

Note that with this construction the definite direct object is not marked with *-ro*.

<i>Zaydullo-ra</i>	<i>palink-o-š</i>	<i>peš-i</i>	<i>usto</i>	<i>ovard-am</i>
Z.-for	shoe.p-his [no DO]	before-EZ	cobbler	bring.PT-1s
‘I took Zaydullo’s shoes to the cobbler’s’.				

The use of *-ro* in the preceding two constructions is reminiscent of the earlier stages of Persian (predicative) dative-possessive construction: *ma[-rā] pesar-i hast* ‘I have a son’, lit. ‘to-me a son is’.

3.1.4.3 Other Tajik postpositions

The use of postpositions is a feature that distinguishes Tajik from Persian.

(1) MLT is relatively conservative, admitting formally only *barin* ‘like, resembling’: *man [barin] odam* ‘a person like me’ (vs. Pers. *ādam-i mesl-e, mānand-e man*, with indefinite *-i*).

(2) Other postpositions are:

(a) *qatī* ‘with’ (found in Southern dialects and Afghanistan as a preposition): *tu [qatī]* ‘with you’, *[bo] qošuq [qatī]* ‘with a spoon’ (here as an ambiposition with the synonymous preposition *bo*); and (b) *da* (< *dar*) ‘in, at, to’: *ow[-da] raft* ‘she’s gone to (fetch) water’.

(3) Other postpositions used widely in the Northern dialects are direct borrowings from Uzbek, e.g. *-dan* ‘from’.

3.1.5 Adverbs

3.1.5.1 Nouns

Nouns with inherent locative and temporal meaning are typically found in adverbial function, frequently without adpositions, as mentioned above: *bāzār raft/bozor raft* ‘she went to (the) market’, *xāne nist-and/xona nest-and* ‘they’re not (at) home’.

Iterative-approximative *-hāl-ho*: Such nouns may be marked by plural *-hā* in its function to express iteration or approximation: *šab-[hā] kār mi-kon-ad/šab[-ho] kor me-kun-ad* ‘he works nights’, *in-jā[-hā]/in jo[-ho]* ‘around here somewhere, hereabouts’.

3.1.5.2 Adverbs proper

Adverbs proper include the following:

- (1) A small set of invariable, unmodified adverbs, including *hanuz/hanūz* ‘still, yet’, *hamišelhameša* ‘always’, *hargez/hargiz* ‘ever’ (with negated verb, ‘never’, in Pers. coll. *hič vaxt* ‘any time’), *faqat* ‘only’, Pers. coll. *hey* ‘continuously, repeatedly’.
- (2) A small set of those which selectively admit of limited qualification includes the words for ‘now’: *hāl/holo, hole, aknun ~ konun* (literary) *laknun*, Taj. *hozir*, modified *ham-in hālā, ham-aknun* ‘right now’, *al’ān* ‘now, this moment’. *az, tā hālā/az, to holo* ‘from now on, until now’; *tā aknun/to aknun* ‘up till now’, *az al’ān* ‘from right now’; further, *bas/bas* ‘enough’ (*bas-ā/bas-o* ‘many, much’, literary): *az bas kelaz bas ki . . .* ‘so much that . . .’

3.1.5.3 Adjective-adverbs and noun-adverbs

Essentially all adjectives may function as adverbs. In addition, a large number of nouns function as noun-adverbs. Accordingly, the majority of adverbs are identical with or derived from adjectives or nouns.

- (1) Common adjective-adverbs are: time, *zud/zud* ‘early, soon’, *dird/der* ‘long, late’ (*dir bāz/der boz* ‘long ago, for a long time’; Taj. *dar-rav* ‘immediately’; location, *dur/dur* ‘far, distant’, *nazdik/nazdik* ‘near (to), at hand’; manner, *tanhāl/tanho* ‘alone, only’, *āhestel/ohīsta* ‘slow, slowly’, *yavāš* ‘slow, slowly, restrained’, *tond/tez* ‘quick, quickly’.
Adjectival adverbs admit of the comparative: *zud-tar/zud-tar* ‘sooner, quicker’, *har če zud-tar* ‘as quickly as possible’.
- (2) Common noun-adverbs are: *bālā/bolo* ‘up, above, upstairs’, *pā’in/poyon* ‘below, down(stairs)’; *em-ruz/lim-rūz* ‘today’, *em-sāllim-sol* ‘this year’ and related time

- relational terms; *nāgāhān/nogahon* ‘suddenly’ (< *nā-gāh* ‘inopportune time’), *gāh-gāhi*/Taj. *goho* ‘sometimes’ (in origin a plural, *goh-ho*).
- (3) Abstract suffix *N-īl-ī*. Adverbs may also be formed by the periphrastic structure: (a) *be N-ilba N-ī* for manner and (b) *dar, az N-ildar N-ī* for location, the derived quality of the nominal assuming the stressed abstract suffix *-il-ī* (see section 7.1.2 Word formation): *be zud-ilba zud-ī* ‘soon, quickly’, lit. ‘with quickness’, *be sefid-i-ye barf/ba safed-ī-yi barf* ‘as white as snow’, *dar bist kilometri-ye Tehrān/dar bist kilometrī-yi Dušanbe* ‘(at a distances of) 20km from Tehran/Dushanbe’, (*be*) *hamegil(ba) hamaḡī* ‘altogether’ (*hame* ‘all’).
- (4) Reduplications of substantives, adjectives, and participles are also specifically adverbial: *gāh-gāh(i)/goh-goh(ī)* ‘from time to time, sometimes’ (abstract *-i-l-ī*), *tond tond gozašt/tez guzašt* ‘it raced by’, *raft-e-raft-elraft-a-raft-a* ‘gradually’.
- (5) Derivational are: (a) a few adjectives with suffix *-ānel-ona* (see section 7.1.2 Word formation) form a few adverbs of time: *šab-ānelšab-ona* ‘at night, by night’, *xoš-baxt-ānelxoš-baxt-ona* ‘luckily, fortunately’; *mota* ‘assef-ānelmutaassif-ona’ ‘regretfully’; (b) a few active participles in *-ānl-on*, e.g. : *davān-davān/dav-on-dav-on* ‘at a run’.
- (6) Arabic borrowings: (a) Adverbs in *-an*. These constitute the largest class of dedicated, morphologically marked adverbs: *moḡqat-an/muḡqat-an* ‘temporarily’, *ettefāq-an/littifoq-an* ‘incidentally’, *ba’d-an/ba’d-an* ‘later’, *fowr-an/fawr-an* ‘immediately’, *asl-an/asl-an* ‘at all, ever’, with neg: ‘not at all, never’. (b) Common adverbials modifying adjectives, adverbs and measure such as *nesbat-an/nisbat-an* ‘relatively’, *kāmel-an/komil-an* ‘completely’, Pers. *taqrib-an*/Taj. *taxmin-an* ‘approximately’. (c) Common Arabic adverbial constructs include *lā-bodllo-bud* ‘undoubtedly’, *be-lā-fāsele/bi-lo-fosila* ‘without interruption, immediately’, *belāxarel biloxira* ‘finally’.

3.1.6 Numerals

3.1.6.1 Cardinal numbers

TABLE 8.5: CARDINAL NUMBERS

1–9		11–19		10–90		100–900	
Pers.	Taj.	Pers.	Taj.	Pers.	Taj.	Pers.	Taj.
<i>yek</i>	<i>yak</i>	<i>yāzdah</i>	<i>yozdah</i>	<i>dah</i>	<i>dah</i>	<i>(yek)sad</i>	<i>(yak)sad</i>
<i>do</i>	<i>du</i>	<i>davāzdah</i>	<i>duvozdah</i>	<i>bist</i>	<i>bist</i>	<i>devist</i>	<i>dusad</i>
<i>se</i>	<i>se</i>	<i>sizdah</i>	<i>sezdah</i>	<i>si</i>	<i>sī</i>	<i>sisad</i>	<i>sesad</i>
<i>čahār</i>	<i>čor, *čahor</i>	<i>čahārdah</i>	<i>čordah</i>	<i>čehel</i>	<i>čil, *čihil</i>	<i>čahārsad</i>	<i>čorsad</i>
<i>panj</i>	<i>panj</i>	<i>pānzdah</i>	<i>ponzdah</i>	<i>panjāh</i>	<i>panjoh</i>	<i>pānsad</i>	<i>panjsad</i>
<i>šeš</i>	<i>šaš</i>	<i>šānzdah</i>	<i>šonzdah</i>	<i>šast</i>	<i>šast</i>	<i>šešsad</i>	<i>lšašsad</i>
<i>haft</i>	<i>haft</i>	<i>hefdah</i>	<i>hafdah</i>	<i>haftād</i>	<i>haftod</i>	<i>haftsad</i>	<i>haftsad</i>
<i>hašt</i>	<i>hašt</i>	<i>heždah</i>	<i>haždah</i>	<i>haštād</i>	<i>haštod</i>	<i>haštsad</i>	<i>haštsad</i>
<i>noh</i>	<i>nuh</i>	<i>nuzdah</i>	<i>nuzdah</i>	<i>navad</i>	<i>navad</i>	<i>nohsad</i>	<i>nuhsad</i>

* literary, archaic

Numbers are construed in order from highest to lowest, with the connective enclitic *-ol-u* ‘and’ after each range: *bist-o yek/bist-u yak* ‘21’, *si-o yek/sī-u yak* ‘31’, etc.; 1000: *hazār/hazor*, or *yek hazār/ryak hazor*; 2000: *do hazār/du hazor*, etc.; 4963: *čahār hazār-o nohsad-o*

šast-o selčor hazor-u nuh-sad-u šast-u se; ‘zero’ is *sefr/sifr* (< Arabic), in Taj. also or *noł* ‘nil’ (< Russian).

SINGULAR AND PLURAL MARKING

Numerical quantifiers precede the numerand and canonically require ‘the singular, including cardinal numbers as well as indefinite and interrogative quantifiers (*čand/čand* ‘several; how many’: *čahār, čand darviščor, čand darveš* ‘four, several dervishes’. The exception is the indefinite *ba’z-ilba’z-e* ‘some, several’ which obligatorily requires the plural: *ba’z-i deraxt-hālba’z-e daraxt-ho* ‘several trees’.

3.1.6.2 Ordinal numbers

Ordinal numbers are formed with the suffix *-oml-ūm, -yūm*; after vowels: Pers. *avval*, Taj. *yak-ūm, do-v-oml/du-yūm, se-v-oml/se-yūm, si-yoml/si-yūm* ‘first, second, third, thirtieth’; note that Persian has *avval* (< Arabic), but regular *yek-om* with decades: *bist-o-yek-om* ‘21st’.

Like regular adjectives, ordinals follow the noun qualified in EZ construction: *sāl-e se-v-oml/sinf-i se-yūm* ‘third year/grade’.

Ordinals themselves may be substantivized and constitute the head noun in an EZ phrase to express a date: *avval-e xordādlyak-ūm-i may* ‘the first of Xordād/the first of May’, and may be joined to a pronominal enclitic: *čahār-om-aš-rā na-did-amlčor-ūm-aš-ro na-did-am* ‘I didn’t see the fourth one (of them)’.

3.1.6.3 Fractions

In Persian fractions are expressed by cardinal numerator + ordinal denominator: *yek čahār-om, do se-v-om* ‘one-fourth, two-thirds’. Quite differently, Tajik uses a partitive collocation in the inverse order, ‘from [cardinal denominator] [cardinal numerator]’: *az se du*, lit. ‘from three (equal parts) two’ = ‘two-thirds’, etc., omitting the preposition if the numerator is ‘one’: *panj-yak* ‘one-fifth’, *dah-yak* ‘one-tenth; ten percent’.

3.1.6.4 Distributive adverbs

Distributive adverbs are formed by simple repetition of the number, with or without the classifier *-tāl-ta*: *čahār čahār/čor čor* ‘four by four, in fours’; *do tā do tāldu-ta du-ta (du-to du-to)* ‘two by two, in twos’; or by suffixing stressed *-i* in Persian: *yek-i yek-i* ‘one by one’, *yek-tā-i*, ‘singly’, and in Tajik by suffixing *-g-ī* to the number + classifier stem: *yak-ta-g-ī (yak-to-g-ī)* ‘singly, one by one’.

3.2 Verb morphology

Verbs are inflected for person, number, tense, aspect, and mood. Further distinctions are made by aspectual-modal particles and periphrastic constructions.

3.2.1 Stem formation

3.2.1.1 Present and past stems

Every verb has two stems: Stem I (traditionally called the “present stem”, from which are formed present tenses, and present and active participles) and Stem II (the “past stem”)

from which are formed past tenses, past and passive participles, and other nominal forms. Stem II always ends in a dental, *-d* or *-t*.

Verbs are divided into two form classes: (1) regular (weak), in which Stem II is derived from Stem I by suffixation of *-id* (as *ras-lras-* > *rasid-lrasid-* ‘arrive’), and (2) irregular (strong), in which the form of one stem is not predictable from that of the other (reflecting Old Iranian ablaut patterns, obscured and mixed since Middle Persian). There are a few pairs of verbs with both irregular and regularized Stem II forms, such as *gard-*, *gašt- ~ gardid-* ‘turn, become’. A good number of irregular verbs have regularized forms in *-id*, side by side with the irregular Stem II retained in the higher registers. Note that “regular” and “irregular” refer only to stem relationships; except for the copula, all conjugations and other verb forms are regularly derived from these two stems.

Similar to other Indo-European languages, there are small subsets which have irregular formants and those where the original morphophonemic change is still partially recognizable. The following lists some subsets with select members (omitting Tajik forms; parentheses = higher register; * = final consonant of Old Iranian verbal root + principal marker *-tá-*):

<i>-id</i>	<i>ras-</i> > <i>ras-id-</i> ‘reach, arrive’, <i>keš-</i> > <i>keš-id-</i> ‘stretch, pull’;
<i>-d ~ -t</i>	<i>xān-</i> > <i>xān-d-</i> ‘read, call, sing’, <i>koš-</i> > <i>koš-t-</i> ‘kill’;
<i>-ād</i>	<i>ist-</i> > <i>ist-ād-</i> ‘stand (up)’, <i>oft-</i> > <i>oft-ād-</i> ‘fall’;
<i>-est</i>	<i>dān-</i> > <i>dān-est-</i> ‘know’; <i>tavān-</i> > <i>tavān-est-</i> ‘can’;
<i>-ā > -ud</i>	<i>farm-ā-</i> > <i>farm-ud-</i> ‘command, order’, <i>rob-ā-</i> > <i>rob-ud-</i> ‘rob, steal’;
<i>ar > or-d</i>	<i>bar-</i> > <i>bor-d-</i> ‘bear, carry’, <i>šomār-</i> > <i>šomor-d-</i> ‘count’ (< * <i>ār-tá-</i>);
<i>b > f</i>	<i>yāb-</i> > <i>yāf-t-</i> ‘find’, <i>kub-</i> > (<i>kuf-t-</i>) > <i>kub-id-</i> ‘pound’ (< * <i>p-tá-</i> > <i>-ftá-</i>);
(<i>y</i>) > <i>s</i>	<i>šu(y)-</i> > <i>šos-t-</i> ‘wash’, <i>ru(y)-</i> > (<i>ros-t-</i>) > <i>ru(y)-id-</i> ‘grow’ (< * <i>d-tá-</i> > <i>-stá-</i>);
<i>s > š</i>	<i>nevis-</i> > <i>neveš-t-</i> ‘write’, <i>ris-</i> > (<i>reš-t-</i>) > <i>ris-id-</i> ‘spin’ (< * <i>ā-tá-</i>);
<i>z > x</i>	<i>sāz-</i> > <i>sāxt-</i> ‘build’, <i>rīz-</i> > <i>rīxt-</i> ‘drop, pour’ (< * <i>k-tá-</i> > <i>xtá-</i>).

Typically, most of the statistically frequent verbs are also most irregular, including:

<i>kon-</i> > <i>kar-d-</i> ‘do, make’;	<i>zan-</i> > <i>za-d-</i> ‘hit’;
<i>šav-</i> > <i>šo-d-</i> ‘become’;	<i>band-</i> > <i>bas-t-</i> ‘bind, tie, close’;
<i>deh-</i> > <i>dā-d-</i> ‘give’;	<i>gu-</i> > <i>gof-t-</i> ‘say’;
<i>gir-</i> > <i>geref-t-</i> ‘take’;	<i>rav-</i> > <i>raf-t-</i> ‘go, leave’.

Suppletive

<i>ā-</i> vs <i>āma-d-</i> ‘come’;	<i>bin-</i> vs <i>di-d-</i> ‘see’.
------------------------------------	------------------------------------

3.2.1.2 Denominal verbs

Verbs may be formed by suffixing regular *-id* to the noun or nominal stem: *nām-*, *nāmid-nom-*, *nom-id-* ‘name’ (< *nom* ‘name’), Pers., Taj. *fahm-*, *fahm-id-* ‘understand’ (< *fahm* ‘understanding’), Taj. *favt-lfavtid-* ‘pass away’ (< *favt* ‘death’; in origin Arabic action nouns). In both Persian and Tajik this procedure is no longer very productive.

3.2.1.3 Derived causative verbs

Intransitive verbs may be transitivized, and both intransitive and transitive verbs may be made causative by suffixation of *-ān/-on* to Stem I and *-ān-(i)d/-on-(i)d-* to Stem II: *ras-/ras-*, *rasid-/irasid-* ‘arrive’ > *ras-ān-/ras-on-*, *ras-ān-(i)d-/ras-on-(i)d-* ‘make arrive, take to’, *xor-/xūr-*, *xord-/xūrd-* ‘eat’ or ‘drink’ > *xor-ān-/xūr-on-*, *xor-ān-(i)d-/xūr-on-(i)d-* ‘make eat or drink, feed’. Irregular causatives include Pers. *nešin-*, *nešast-* ‘sit (down)’ > *nešān-*, *nešānd-* ‘set (down)’, but Taj. *šin-*, *šišt-* ‘sit’ > *šin-on-*, *šin-on-id-* ‘set, plant’.

A few causative-type derivatives are commonly used with little or no modification of the meaning of the basic verb:

Pers. *suz-/sūz-*, *suxt-/sūxt-* ‘burn (tr., intr.)’ > *suz-ān-*, *suz-ān-d-* ‘(make) burn’; Taj. *firist-*, *firistod-* > *firist-on-*, *firist-on-id-* ‘send’. Others represent a different specialization of the base verb: *keš-/kaš-*, *kešid-/kašid-* ‘pull, draw (tr.), stretch (intr.)’ vs. *keš-ān-/kaš-on-*, *keš-ān-d-/kaš-on-id-* ‘take away, carry off’.

TAJIK

While in Persian derived causativation is only partially productive, in Tajik it is fully so, including:

- (1) make-other-causatives: *dūz-/dūxt-* ‘sew’, *dūz-on-/dūz-on-id-* ‘have something sewn’;
- (2) transitive denominal and deadjectival verbs: *mukofot-on-*, *mukofot-on-id-* ‘reward’ (< *mukofot* ‘reward’), *elektr-on-*, *elektr-on-id-* ‘electrify, power’ (< *elektrifika* ‘electric’), *xušk-on-*, *xušk-on-id-* ‘dry’ (< *xušk* ‘dry’);
- (3) causativation of transitive compound verbs with *kun-*, *kard-* ‘do, make’: *remont kun-on-*, *kun-on-id-* ‘have (something) repaired’.

3.2.1.4 Preverbs

The three most frequently occurring preverbs are *bar/bar* ‘up’ (Pers. colloq. *var*), *dar/dar* ‘in; out’, *foru(d)/furū(d)* ‘down’; further, *bāz/boz* ‘re-, again’ (mostly colloq. *vālvo*), and literary also *farā* ‘forth’.

In Tajik, but not in Persian, the first three preverbs have become inseparably attached to the stem of the most common verbs of motion, and are preceded by the negation *na-* and imperfective marker *me-*:

Pers. <i>bar ā-</i> ,	<i>āmadl</i>	<i>dar āvar-</i> ,	<i>āvord</i>	<i>foru ā-</i> ,	<i>foru āmad-</i>
Taj. <i>bar o-</i> ,	<i>omad-</i>	<i>dar or-</i> ,	<i>ovard-</i>	<i>furo-</i> ,	<i>furomad-</i>
	‘go up, out’;		‘bring in, out’;		‘come down, descend’;

but

Pers. <i>bar mi-ā-yam</i>	<i>dar mi-āvar-am</i>	<i>foru mi-āy-am</i>
Taj. <i>me-[bar]-o-yam</i>	<i>me-[dar]-or-am</i>	<i>me-[furo]-yam.</i>
IPFV-out-come.PR-1s	IPFV-out-bring.PR-1s	IPFV-down.come.PR-1s
‘I go up, out’;	‘I bring in, out’;	‘I come down’.

An exception to the Tajik coalescence is *dar gir-*, *girift-* ‘catch, take (fire, etc.)’; *dar na-me-gir-ad* ‘it isn’t catching’.

For light verb constructions, see section 5.5.6.

3.2.2 Nominal forms

While forms derived from Stem I are only minimally productive, and function mainly as adjectives and nouns, those derived from Stem II are fully productive.

3.2.2.1 Forms derived from Stem I

3.2.2.1a Present participle, PR-*andel-anda*

This participle is derived by the suffix *-andel-anda*: *ā-yandelo-yanda* ‘coming; future’. The latter is both an adjective and a noun: *hafte-ye āyandelhafta-i o-yanda* ‘next week’, *dar āyandeldar o-yanda* ‘in (the) future’, whereas *nevisandel/navis-anda* ‘writer’ and most others have evolved into agent nouns.

3.2.2.1b Active participle, PR-*ānl-on*

This participle is derived by the suffix *-ānl-on*: *rav-ānlrav-on* ‘going’.

3.2.2.2 Forms derived from Stem II

3.2.2.2a Infinitive PT-*an*

The infinitive is formed by addition of *-an*: *gereft-an/girift-an* ‘to take, taking’, *šod-an/šud-an* ‘to become, becoming’ (primarily an action noun, which does not function in the verbal system).

3.2.2.2b Short infinitive PT

The “short infinitive” is identical with the past stem (Stem II) of limited application in the verbal system.

3.2.2.2c Past (perfect) participle PT-*a*

The past, or perfect, participle (active or passive) is derived by the suffix *-el-a*: *gereft-el/girift-a* ‘(having been) taken’, *šod-el/šud-a* ‘(having) become’. They may also function as adjectives or nouns: *gozaštel/gozašta* ‘past, (the) past’.

3.2.2.2d Tajik participle PT-*agī*

From the past participle is derived a second “past participle” by the addition of stressed *-g-ī* to the base: *šud-agī* ‘having become’, *girift-agī* ‘having (been) taken’, *guzašt-agī* ‘having passed’.

3.2.2.2e Future participle or gerundive PT-*anil-anī*

This participle is derived by stressed *-il-ī* from the infinitive, or verbal noun: intransitive, *raft-an-i/raft-an-ī* ‘about to go’, transitive, *did-an-ī/did-an-ī* ‘worth seeing’, *bāvar-na-kard-an-ī/bovar-na-kard-an-ī* ‘unbelievable’.

The latter participles play several specialized roles in verb formation and NP syntax (see sections 3.2.6b.4 Tajik Conjectural Mood, 5.5.4 Conject verbs, and 5.5.2 Modal constructions).

3.2.3 Person marking and 'be'

The usual present tense of the verb 'be' comprises an affirmative independent and an enclitic copula paradigm. The personal endings differ from the latter only in the 3s. The other tense and modal forms are supplied by *bāš-*, *bud-/ boš-*, *bud-*).

TABLE 8.6: 'BE' AND PERSONAL ENDINGS

	1s	2s	3s	1p	2p	3p
Independent	<i>hast-aml</i> <i>hast-am</i>	<i>hast-il</i> <i>hast-ī</i>	<i>hast, astl</i> <i>hast, ast</i>	<i>hast-impl</i> <i>hast-em</i>	<i>hast-idl</i> <i>hast-ed</i>	<i>hast-andl</i> <i>hast-and</i>
Copula	<i>-aml</i> <i>-am</i>	<i>-il</i> <i>-ī</i>	<i>ast, -stl</i> <i>ast, -st</i>	<i>-impl</i> <i>-em</i>	<i>-idl</i> <i>-ed</i>	<i>-andl</i> <i>-and</i>
<i>Personal Endings</i>						
PR	<i>-aml</i> <i>-am</i>	<i>-il</i> <i>-ī</i>	<i>-adl</i> <i>-ad</i>	<i>-impl</i> <i>-em</i>	<i>-idl</i> <i>-ed- ~ eton</i>	<i>-andl</i> <i>-and</i>
PT	<i>-aml</i> <i>-am</i>	<i>-il</i> <i>-ī</i>	<i>-Øl</i> <i>-Ø</i>	<i>-impl</i> <i>-em</i>	<i>-idl</i> <i>-ed ~ -eton</i>	<i>-andl</i> <i>-and</i>
Imperative		<i>-Øl</i> <i>-Ø</i>			<i>-idl</i> <i>-ed</i>	

The endings, but not the copula, insert *-y-* after vowels: Pers. *mi-gu-yam* 'I say', but *emrikā'i and* 'they are American', *xaste im* 'we are tired'; Tajik *me-gū-yam* 'I say', but *dono-yand* 'they are knowledgeable'. *Ast* contracts with final vowels in formal speech: *kojā astlkujo ast* > *kojā-stlkujo-st* 'where is it?'.
hast- functions both as the existential verb and as an alternate of the copula in persons other than 3s:

<i>man</i>	<i>hast-am, to</i>	<i>bo-row!</i>	<i>xub-aml/xub-am ~ xub hast-aml</i>
<i>man xub</i>	<i>hast-am, tu</i>	<i>raw!</i>	<i>xub hast-am</i>
I	well EX-1s	you.s go.IMP.2s	well-EX-1s
	'I will stay', lit. 'am here, you go!';		'I am well, fine';

but 3s *hast/hast* is only used in existential sentences:

dar in otāq do panjere hast!
dar in xona du tireza hast
in this room two window EX.3s
'there are two windows in this room'.

NEGATION

The negative form of both independent and enclitic forms of 'be' are based on *nist-/nest-*, inflected like *hast-*, *nist-am*, *nist-i*, *nist*, etc. *nest-am*, *nest-ī*, *nest*, etc.

Dialect and vernacular variants include:

	Pers.	Taj.
3s <i>ast/last</i>	> <i>-e</i>	= <i>ast, as</i>
3s <i>-adl-ad</i>	> <i>-e</i>	<i>-a</i>
2p <i>-idl-ed</i>	> <i>-in</i>	<i>-et, -e</i>
3p <i>-andl-and</i>	> <i>-an</i>	<i>-an</i>

Pers. *xub ast* > *xub-e* ‘it’s good’, *mi-ras-adlme-ras-ad* > *mi-res-elme-ras-a* ‘arrives’.

Honorific use: 2p *-idl-ed* and 3p *-andl-and* may be used to address or refer to a single person in polite or formal usage, while Tajik 2p *-eton* (cf. personal clitics) is used only for the plural (see section 3.1.3.1 Personal pronouns).

TAJIK *-ak*

In some Tajik dialects, such as Varzobi, occur forms with an apparent reflex of the nominal diminutive affix *-ak*, with affective connotations:

in-aš *raft[-ak]*, *dil-am* *sūxt*
 this-REF left-AK heart-my burn.PT.3s
 ‘she [left], and I was devastated’.

3.2.4 Aspectual-modal markers and negation

3.2.4.1 Prefix *mi-lme-*

Aspecto-temporal distinctions are expressed primarily by the presence or absence of the stressed imperfective marker *mi-lme-* added to Stem I and Stem II forms. With Stem I forms, *mi-lme-* distinguishes present/future indicative from subjunctive/optative: *mi-rav-aml me-rav-am* ‘I go, am going, will go’ vs. *be-rav-am/rav-am* ‘let me go’. For the modal use of *mi-lme-* with Stem II forms, see sections 3.2.5.9–10 Counterfactual forms.

3.2.4.2 Prefix *be-lbi-*

Stressed *be-* marks the subjunctive and imperative in Persian *be-rav-am* ‘(that) I go’, *be-rav-id* ‘go-2p!’

- (1) *be-* > *bi-* before initial stem vowel: *bi-ā-* ‘come’, *bi-andāz-* ‘throw’; the vowel may be lost: *bi-andāz-and* > *b-endāz-an(d)* ‘let them throw’.
- (2) In colloquial there is assimilation: *be-gozār* > *bo-gzār* ~ *be-zār* ‘let (me . . .)’, *be-rav* > *bo-row* ‘go-2p!’, *be-kon* > *bo-kon* ‘2s do!’

be- is mutually exclusive with the preverb *bar-*: *mi-xāh-am bar gard-am* ‘I want to return’, lit. ‘that I return’, but is optional in compound verbs, particularly in dependent constructions:

ejāze *be-deh-id* *xod-am-rā* *mo'arrefi* *(be-)kon-am*
 permission IMP.give-2p self-my-DO introduced SBJ.make.PR-1s
 ‘allow me to introduce myself’, ‘lit. ‘that I introduce’.

In Tajik, however, *bi-* is vestigial. It occurs regularly only as a morphological suppletive in Stem I forms of the two common verbs *o-lomad-* ‘come’ and *or-lovard-* ‘bring’: *me-bi-oyam* ‘I come, am coming’; *bi-or*, *bi-or-ed* ‘bring (it)’.

It may be used with Stem I to mark the subjunctive: *agar bi-gū-yad* . . . ‘if he says . . .’, but this usage is characterized as poetic or archaizing (for *agar gū-yad*). As an optional prefix of the imperative it occurs with some common verbs in polite spoken Tajik: *bi-šin(-ed)* ‘sit-2s(2p) down!’ (cf. 2p *šin-ed* ~ *šin-eton*), *bi-don(-ed)* ‘know-2s(2p)!’. Before a stem beginning with *b-* the vowel is modified to *u*: *bu-baxš-ed-2p* ‘excuse 2p’, *bu-bin-ed* ‘see-2p’.

3.2.4.3 Negation *na-/na-*

The negative prefix *na-*, Pers. *ne-* before *mi-*, precedes other prefixes and takes primary stress: *ne-mi-gu-yam/na-me-gū-yam* ‘I won’t say’, *ne-mi-āvar-aml na-me-bi-or-em* ‘we won’t bring (it)’, *na-raft/na-raft* ‘3s did not go’. It likewise precedes the periphrastic verb forms, e.g. *na-raft-e ast/na-rafta ast* ‘has not gone’.

3.2.5 Basic system of tense, aspect, and mood

Persian and Tajik, like other varieties of Persian, share a core set of basic forms of the verb, and especially in their personal inflection offer analogies with other Indo-European verbal paradigms. The basic aspect system has a triple distinction of imperfective, perfective-aorist, and resultative-stative forms.

The following comparative tables show the basic verb forms and their functions shared by Persian and Tajik, with the example of the verb for “go, leave”, present stem *rav-*, past stem *raft-*; perfect participle *raft-elraft-a*, all with 3p marking, except 2p imperative, first in list form, second in terms of their categorical vectors.

While there are nine basic finite forms, three of those forms have more than one function, of which two forms have both indicative and non-indicative functions: (1) the imperfective past *mi-raft-/me-raft-*, also functions as imperfective counterfactual; (2) the resultative-stative, *raft-e bud-/raft-a bud-*, also functions as perfective counterfactual. (3) one form, *raft-*, the simple past, often also referred to preterite, functions not only in past, but also in present and future contexts, as well as in potential conditions, for which reason the term “aorist” has been suggested by Windfuhr.

Future. All indicative forms may function in future contexts, as there is no paradigmatic future formation. (For the discussion of the Persian verb system, cf. also Windfuhr 1979: 83–126; 1982; 2006).

3.2.5.1 Present-future

mi-rav-ad/me-rav-ad ‘goes, is going, will go’

The present indicative marked by *mi-/me-* is the general present imperfective form with a wide range of functions. Depending on the context and adverbial cues, it may refer to actions that are:

- (1) habitual-iterative and generic, ‘they (always, generally) go, leave’;
- (2) progressive, ‘they are going, leaving’;
- (3) intentional, ‘they are about to go, leave’;
- (4) future, ‘they will go, leave’.

3.2.5.2 Imperfect

mi-raft/me-raft ‘was going, used to go, would go’

This imperfective form has both indicative and non-indicative functions.

As an indicative, the imperfect is the general past imperfective form with a wide range similar to the present indicative. Depending on the context and adverbial cues, it may refer to events that are:

- (1) habitual-iterative and generic, ‘they (always, generally) went, left’ ~ ‘would go, leave’ ~ ‘used to go, leave’;

TABLE 8.7: PERSIAN AND TAJIK CONJUGATIONS

	IMPERFECTIVE	PERFECTIVE	RESULTATIVE-STATIVE	
Indicative				
	1	2	3	4
	<i>Present</i>	<i>Imperfect</i>	<i>Preteritel Aorist</i>	<i>Present Perfect</i>
1s	<i>mi-rav-am</i> <i>me-rav-am</i>	<i>mi-raft-am</i> <i>me-raft-am</i>	<i>raft-am</i> <i>raft-am</i>	<i>raft-e am</i> <i>raft-a am</i>
2s	<i>mi-rav-i</i> <i>me-rav-ī</i>	<i>mi-raft-i</i> <i>me-raft-ī</i>	<i>raft-i</i> <i>raft-ī</i>	<i>raft-e i</i> <i>raft-a ī</i>
3s	<i>mi-rav-ad</i> <i>me-rav-ad</i>	<i>mi-raft</i> <i>me-raft</i>	<i>raft</i> <i>raft</i>	<i>raft-e ast</i> <i>raft-a ast</i>
1p	<i>mi-rav-im</i> <i>me-rav-em</i>	<i>mi-raft-im</i> <i>me-raft-em</i>	<i>raft-im</i> <i>raft-em</i>	<i>raft-e im</i> <i>raft-a em</i>
2p	<i>mi-rav-id</i> <i>me-rav-ed</i>	<i>mi-raft-id</i> <i>me-raft-ed</i>	<i>raft-id</i> <i>raft-ed</i>	<i>raft-e id</i> <i>raft-a ed</i>
3p	<i>mi-rav-and</i> <i>me-rav-and</i>	<i>mi-raft-and</i> <i>me-raft-and</i>	<i>raft-and</i> <i>raft-and</i>	<i>raft-e and</i> <i>raft-a and</i>
Neg.	<i>ne-mi-rav-</i> <i>na-me-rav-</i> + PR Ending	<i>ne-mi-raft-</i> <i>na-me-raft-</i> + PT Ending	<i>na-raft-</i> <i>na-raft-</i> + PT Ending	<i>na-raft-e</i> <i>na-raft-a</i> + COP
	6	2a	3a	7
	<i>Subjunctive</i>	<i>Counterfactual</i>	<i>Aorist</i>	<i>Subjunctive</i>
	<i>be-rav-</i> <i>rav-</i> + PR Ending	<i>mi-raft-</i> <i>me-raft-</i> + PT Ending	<i>raft-</i> <i>raft-</i> + PT Ending	<i>raft-e bāš-</i> <i>raft-a boš-</i> + PR Ending
Neg.	<i>na-rav-</i> <i>na-rav-</i> + PR Ending	<i>ne-mi-raft-</i> <i>na-me-raft-</i> + PT Ending	<i>na-raft-</i> <i>na-raft-</i> + PT Ending	<i>na-raft-e bāš-</i> <i>na-raft-a boš-</i> + PR Ending
	8			9
	<i>Imperative present 2p</i>			<i>Imperative perfect/past 2p</i>
	<i>be-row, be-rav-idl</i> <i>(bi-)raw, (bi-)rav-ed</i>			<i>raft-e bāš-idl</i> <i>raft-a boš-ed</i> <i>(na-raft-e bāš-id)</i> <i>(na-raft-a boš-ed)</i>
Neg.	<i>na-row, na-rav -idl</i> <i>na-raw, na-rav-ed</i>			

- (2) progressive, less commonly, ‘they were going, leaving’;
 (3) intentional ‘they were about to go, leave’;
 (4) past future ‘they would go, leave (the next day, etc.)’.

Examples for intentional and past future contexts:

to ke čiz-i [ne-mi-dād-i]
tu ki čiz-e [na-me-dod-ī],
 you.s that thing-INDEF not-IPFV-give.PT-2s,
čerā hamān dam-e dar na-goft-il
čaro hamon dam-i dar na-guft-ī
 why that-same at-EZ door not-say.PT-2s
 ‘since you [were not going to ~ would not give] me anything, why didn’t you say so
 right at the door?’

fardā piš-e mādar-aš mi-raft va be u mi-goft
fardo peš-i modar-aš me-raft-u ba ū me-guft
 tomorrow before-EZ mother-his would go and to her IPFV-say.PT.3s,
 ‘tomorrow, he would go to his mother and would tell her . . .’

For the counterfactual function of the imperfect see section 3.2.5.9.

3.2.5.3 Preterite, or “aorist”

raftraft ‘he went’, ‘there he goes’

This is the perfective form of the verb system and used mainly in past, but also in present and future contexts. Due to the renderings of the predominantly past occurrences into English and other languages, it is traditionally often referred to as (simple) past.

(1) It states that an action was performed and (by implication) completed, usually in the past, irrespective of its duration, and is the usual tense of narrative:

mā xeyli dur peyāde raft-impl
mo xeledur piyoda raft-em
 we very far on foot go.PT-1p
 ‘we walked a long way’.

(2) It may also designate an action that is presented as completed at the moment of speech:

<i>to bāš-i,</i>	<i>man raft-aml</i>	<i>hasan kū?</i>	–	<i>āh, āmad!</i>
<i>tu boš-ī,</i>	<i>man raft-am</i>	<i>hasan ku?</i>	–	<i>oh, omad!</i>
you.s be.IMP/-2s	I go.PT-1s	Hasan where.is		aha come.PT.3s
‘you stay here, I am on my way’,		‘where is Hasan – there he is!’,		
lit. ‘I went’;		lit. ‘he came’.		

(3) It may also assert the completion to an action or event in the future, or refer to action or state that will necessarily have been completed by the time of the action proposed:

šāyad mā ham raft-impl
šoyad mo ham raft-im
 perhaps we also go.PT-1p
 ‘we will most likely go, too’, lit. ‘perhaps we went, too’.

Such use is typically found as a precondition for the action in the main clause:

<i>havā ke sard šod . . . /</i>	<i>vaqti āmad, be-gul</i>
<i>havo ki xunuk šud . . .</i>	<i>vaqt-e ki omad, bi-gū</i>
weather that cold become.PT.3s	time-L that come.PT.3s say.IMP.2s
‘when/once the weather becomes cold, . . .’;	‘as soon as he comes, tell me’.

(4) As such, the aorist also may have subjunctive function:

agar na-bud, zang be-zanl
agar na-bud, zang zan
 if not be.PT.3s bell ring.IMP.2s
 ‘if he is not there, call.2s me!’,

where the assumed condition is stated as a completed action or state.

3.2.5.4 Present perfect

raft-e ast/raft-a ast ‘has gone, is gone’

(1) The present perfect tense is formed from the past participle plus the enclitic copula (see Table 8.6). In its “traditional” function as a resultative, the perfect designates an action viewed as having been completed, and its consequences still in force, at the time of speaking:

barnāme-rā be kolli ejrā nemud-e andl
plan-ro tamoman ijro namud-a-and
 plan-DO completely execution make.PART make-3p
 ‘they have fully implemented the plan’.

(2) The perfect may refer not only to the result, but also to the process. This is readily shown with change of state verbs, where English makes overt distinctions:

<i>pušid-e and ~ bud-andl</i>	<i>bačče-rā dast gereft-e budl</i>
<i>pošid-a and ~ bud-and</i>	<i>bača-ro dast girift-a bud</i>
put on.PART be.3p be.PT-3p may imply	child-DO hand take.PART be.PT.3s
‘they have/had put on (a coat),’	‘(s)he was holding the child by the hand’,
or	or
‘they are/were wearing (a coat)’;	‘(s)he had taken the child by the hand’.

The perfect has an additional function which will be illustrated under section 3.2.6b.1 Evidential mode.

3.2.5.5 Pluperfect or distant past

raft-e bud/raft-a bud ‘had gone, was gone’

The pluperfect, or distant past, is formed from the past participle and the simple past of *bud-an* ‘be’.

(1) Anterior past action

It refers to an action or state in the past already completed by a certain time in the past, often that of the start of another past action (often rendered by the simple past in English):

ketāb-i rā, ke be man dāde bud-id, xānd-aml
kitob-e-ro, ki ba man dod-a bud-ed, xond-a tamom kard-am
 the book-DO [that to me give.PART be.PT.-2p] read.PT-1s/read.PART finished
 make.PT-1s
 ‘I read/have finished reading the book that you gave me’

(for Tajik, *xond-a tamom kard-am* see section 5.5.3 Aktionsart constructions).

Sometimes a past action is not mentioned, but is understood from the context:

be šomā ādat kard-e bud-impl
ba šumo odat kard-a bud-em
 to you.p adjustment make.PART be.PT-1p
 ‘we are used to you’, lit. ‘we had habituated to you’;

this is a tag explaining the speaker’s reaction to an action just completed.

The pluperfect form, in the context of the time frame mentioned, locates the action as definitively prior to any recent discussion:

u pārsāl āmade budl
ū porsol omad-a bud
 he last year come.PART be.PT.3s
 ‘as you know, etc. he came last year’.

(2) Remoteness

This tense may also be used to underline the remoteness of an action:

devist sāl-e piš az in
dusad sol-i peš az in
 two-hundred year-EZ before from this
pedar-ān-e mā be injā kučid-e bud-andl
bobo-yon-i mo ba injo kūčid-a omad-a bud-and
 fathers-EZ us to here migrate.PART come.PART be.PT-3p
 ‘two hundred years ago our forefathers migrated here’

(for Tajik *kučid-a omad-a bud-and* see section 5.5.4 Conjunct verbs).

(3) Change-of-state verbs

The pluperfect tense of change of state verbs such as *istād-anlistod-an* ‘stand up, stand’, *nešast-anl šišt-an* or *šist-an* ‘sit down, sit’, *xābid-an/xobid-an* ‘lie down, lie’ and ‘go to sleep, sleep’, *pušid-an/pošid-an* ‘put on, wear’, designates past durative time without a sense of prior action or remoteness:

Piruz dar bālāxāne-yaš nešast- budl
Afandī dar boloxona-aš nišast-a bud
 P./A. in upper story-his sit/PART be.PT.3s
 ‘Piruz/Afandi was sitting in the upper story of his house’.

For the counterfactual function of the pluperfect see section 3.2.5.10.

3.2.5.6 Imperative

(1) present imperative: *be-rav-id/rav-ed* ‘go-2p’

The imperative is based on Stem I: 2s -Ø, 2p -*idl-ed*: *be-gir, be-gir-id! gir-, gir-ed* ‘take, take!’.

(2) perfective-resultative imperative: *raft-e bāš-id/raft-a boš-ed* ‘be gone-2p’.

The perfective-resultative imperative is as elsewhere rarely found and both semantically and contextually restricted, mostly found with change-of-state verbs: *nešast-e bāš-idl/šist-a boš-ed* ‘stay seated’, *raft-e bāš/raft-a boš* ‘be gone’.

3.2.5.7 Present-future subjunctive

be-rav-adl (bi-)rav-ad ‘he may go, (that) he go’

The present-future subjunctive (also known as the “aorist subjunctive”, following mainly Turkological nomenclature), is formed on Stem I + personal endings. Persian has the prefix *be-*, while Tajik mainly the plain stem (see Prefix *be-/bi-* section 3.2.4.2).

(1) Independent use

Independently the subjunctive is used as optative, or hortative, and as suggestion:

boland be-šav-iml
xez-em!
 tall SBJ-become.PR-1p/(SBJ).rise.PR
 'let's get up!';

<i>na-tars-iml</i>	<i>be-rav-adl</i>	<i>be-rav-am?l</i>
<i>na-tars-em</i>	<i>rav-ad</i>	<i>rav-am?</i>
not.SBJ-fear.PR-1p	SBJ.go.PR-3s	SBJ.go.PR-1s
'let's not be afraid';	'let him/her go';	'shall I go?'

bāš-adlboš-ad 'let it be, OK', Pers. colloquially *bāš-e*.

(2) Dependent use

Given its function to refer to potential action or state, the subjunctive typically follows modal verbs and expressions:

<i>mi-xāst-am</i>	<i>xāne</i>	<i>be-rav-aml</i>
<i>me-xost-am</i>	<i>xona</i>	<i>rav-am</i>
IPFV-want.PT-1s	house	SBJ-go.PR-1s
'I wanted to go home', lit. 'that I go'.		

As such it occurs in modal constructions and subordinate clauses implying potential actions or states.

Time reference is usually irrelevant, the action expressed by the subjunctive being potentially future in relation to the time of the main verb.

3.2.5.8 Past (perfect) subjunctive

raft-e bāš-adlraft-a boš-ad 'he may have gone, be gone', in past contexts 'might have gone, may be gone'

The past, or perfect, subjunctive is constructed with a past participle + the subjunctive of 'be'. It most commonly occurs in doubting or questioning, desiring or regretting a past action (accordingly rendered in English by a variety of modal formations, depending on the context):

<i>man</i>	<i>boyad</i>	<i>šaš-sola</i>	<i>šud-a</i>	<i>boš-am</i>	
I	must	6-year-old	become.PART	SBJ.be-1s	
<i>ki</i>	<i>ma-ro</i>	<i>padar-am</i>	<i>ba maktab . . .</i>	<i>burd-a</i>	<i>mond</i>
that	I-DO	father-my	to school	take.PART	do.PT.3s
'I may, must have been six years old when my father took me to school . . .'					

for *mond* see section 5.5.4 Conjunct verbs.

<i>bāvar</i>	<i>ne-mi-kard</i>	<i>ke</i>	<i>do</i>	<i>gusfand gom</i>	<i>šod-e</i>	<i>bāš-andl</i>
belief	not-IPFV.do.PT.3s	that	two	sheep	lost	become.PART (SBJ).be-3p
'he still didn't believe that two sheep were, may have, could have been lost'.						

Thus it often follows precative particles such as Pers. (*xodā*) *na-kon-ad* 'may (God) not do', Tajik *na-xod* 'God forbid', *ma-b-ād-ā kelma-b-od-o* (*ki*) 'let it not be', hopefully not', *kāškelkoški* 'would that', and the modal *bāyad/boyad* 'must' in its resumptive sense:

na-kon-ad ke šomā harf-e ma-rāl
na-xod šumo gap-i ma-ro
 God forbid that you.p word-EZ I-DO
jeddi gereft-e bāš-id?! /
jiddī fahmid-a boš-ed? /
 seriously take.understand.PART (SBJ).be.PR-2p
 ‘did you really take me seriously?’

3.2.5.9 Imperfect counterfactual

mi-raft/me-raft ‘might, would go, might, would have gone’, ‘if he went, had gone’.

In its counterfactual function the imperfect is an irrealis and tense neutral. As such it is used in both present-future and past contexts and may imply a present or past irreal action:

(1) to express an unrealized desire, after *koški* ‘would that’:

kāške ne-mi-āmad /
koški na-me-omad /
 would that not-IPFV-come.PT.3s
 ‘if only she had (not) come!’

(2) in both clauses of a counterfactual conditional sentence (see section 6.5.3.7 Conditional clauses):

agar mi-āmad, xod-etān mi-did-id /
agar me-omad-ed, xud-aton me-did-ed
 if IPFV-come.PT-2p self-your IPFV-see.PT-2p
 ‘if you came, you would see for yourself’.

In Persian it may also follow *bāyad* for which see section 5.5.2 Modal constructions.

3.2.5.10 Pluperfect counterfactual

raft-e bud/raft-a bud ‘would have gone, if he had gone’

The pluperfect may also express unfulfilled past conditions in either protasis or apodosis, or both, of a counterfactual conditional sentence (see also section 6.5.3.7 Conditional clauses).

kāške zud-tar āmad-e bud-and /
koški zud-tar omad-a bud-and
 would that early-COMP come.PART be.PT-3p
 ‘would that they had arrived sooner’;

agar zud-tar āmad-e bud-i, mi-did-il /
agar zud-tar omad-a bud-ī me-did-ī
 if early-COMP come.PART be.PT-2s IPFV-see.PT-2s
 ‘if you had come earlier, you would have seen it’.

In Tajik, but not in Persian, it may add the prefix *me-* in pluperfect conditional function: *raft-a [me-]bud-am*, etc.

The following two tables provide an overview of the basic categorical vectors and of the conjugations of the Persian and Tajik verb systems.

TABLE 8.8: BASIC PERSIAN AND TAJIK VERB SYSTEM, 3s

	Imperfective	Perfective	Resultative-stative
Present	<i>mi-rav-ad</i> <i>me-rav-ad</i>		<i>raft-e astl</i> <i>raft-a ast</i>
Past	<i>mi-raft</i> <i>me-raft</i>		<i>raft-e budl</i> <i>raft-a bud</i>
Preterite/Aorist		<i>raft</i>	
Subjunctive	<i>be-rav-ad</i> <i>rav-ad</i>		<i>raft-e bāš-adl</i> <i>raft-a boš-ad</i>
Counterfactual	<i>mi-raft</i> <i>me-raft</i>		<i>raft-e budl</i> <i>raft-a bud</i>
Imperative 2p	<i>be-rav-id</i> <i>rav-ed</i>		<i>(me-) rafta bud</i> <i>raft-a boš-ed</i>

3.2.5.11 Defective verbs

The verbs *bāš-*, *bud-lboš-*, *bud-* ‘be’ and *dār-*, *dāšt-lдор-*, *došt-* ‘have’ are by nature stative and as elsewhere are irregular and partially defective. They do not take *mi-lme-*, but may do so as part of compound verbs.

3.2.5.11a ‘be’

This verb also lacks the perfect subjunctive and past perfect, **bud-e bāš-*, **bud-e bud-*, while the present and past imperfective *mi-lme-* forms have limited use.

(1) Stem I form *bāš-lboš-* supplies the present subjunctive and imperative of the copula, and is also found in a good number of frozen idiomatic expressions, including: subjunctive in optative function: *zende bāš-idl/zinda boš-ed* ‘may you live! Bravo!’, imperative:

<i>mard bāš!</i>	<i>xaste</i>	<i>na-bāš-idl</i>
<i>mard boš!</i>	<i>monda</i>	<i>na-boš-ed</i>
man IMP.be.2s	tired	not.IMP.be-2p
‘be a man!’;	‘don’t be tired’, i.e. ‘take it easy’, a casual greeting.	

<i>Persian</i>	<i>Tajik</i>
<i>bāš-ad</i>	<i>na-boš-ad</i>
(SBJ.)be-3s	not(SBJ.)-be-3s
‘let it be, OK;	‘in that case’, lit. ‘may it not be’;

<i>na-bāš-ad</i>	<i>ke . . .</i>	<i>heč na-boš-ad</i>
not(SBJ.)-be-3s	that	any not(SBJ.)-be-3s
‘I hope not that . . .’;		‘at least, in any case’, lit. ‘may it not be at all’.

TAJIK

<i>na-boš-ad, pagoh</i>	<i>me-o-yed</i>
well then tomorrow	IPFV-come.PR-2p
‘well then, why don’t you come tomorrow?’;	

heč na-boš-ad, yak rūz mehmon-i mo šav-ed
 any not(SBJ.)-be-3s one day guest-EZ we (SBJ.)become-2p
 'at least stay with us one day'.

PRECATIVE

na-boš-ad man kasal šav-am
 not(SBJ.)-be-3s I ill (SBJ.)become.PR-1s
 'God forbid I should fall ill'.

(2) The imperfective form, *mi-bāš-/me-boš-*, may substitute for the existential *hast-* under felicity conditions, e.g. typically literary, but also to lend weight to a statement or when the verb is widely separated from its antecedent:

salim, ke hame-ye šomā u-rā mi-šenās-id,
Salim, ki hamā-i šumo ū-ro me-šinos-ed,
 S. that all-EZ you.p he-DO IPFV.know.PR-2p
ozv-e ān hey'at [mi-bāš-ad]!
a'zo-i on komsomol [me-boš-ad]
 member-EZ that committee IPFV-be.PR-3s
 'Salim, whom you all know, is a member of that committee/the Komsomol'.

(3) In Tajik, but not in Persian, *me-* in its counterfactual function may be added to Stem II: *bud-am*, etc. ~ *me-bud-am*, etc.

3.2.5.11b 'have'

The meaning 'have' of this verb is derived from its basic meaning 'keep, hold'. When used in its primary sense (which implies an imperfective-durative state), this verb does not admit the prefix *mi-/me-* with either stem, while the resultative-stative form *dāšte bāš-/došt-a boš-* functions also as the present subjunctive:

ketāb dār-adl pul na-dāšt-idl
kitob dor-ad pul na-došt-ed
 'he has books'; 'you had no money';

momken ast dāšt-e bāš-andl ān-rā dāšt-e bāšid!
mumkin ast došt-a boš-and on-ro došt-a boš-ed!
 'it is possible they have it' or 'had it'; 'have, keep it!'

It may take *mi-/me-* like any other verb when forming part of an idiomatic compound verb with a nominal or adjectival component:

u-rā dust mi-dāšt-am
vay-ro dūst me-došt-am
 he-DO friend IPFV.have.PT-1s
 'I used to like/love him';

pošt-e bām-rā tir-hā-ye mohkam negāh mi-dār-andl
bom-ro bolor-ho-i yafs nigoh me-dor-and
 roof-DO beam-p-EZ stout support IPFV.keep.PR-3p
 'the roof is held up by stout beams'.

3.2.6a *Extended verb system of Modern Standard Persian*

3.2.6a.1 Evidential or non-witnessed mode

While not found in Early New Persian, Modern Standard Persian has a sub-system of evidential forms. (Windfuhr 1979: 90 [“reported” speech, Table Verb system of contemporary literary Persian], 1982, 1985, 1987; Lazard 1985, 2000; and Jahani 2000a). It is morphologically signaled not by a single marker, but its forms are derived from the perfect participle. The evidential subset is categorically parallel to the sub-system of the past. This results in the double function of the form *raft-e ast*, as both present resultative-stative and evidential perfective past.

TABLE 8.9: PERSIAN EVIDENTIAL FORMS, 3s

	Imperfective	Perfective	Resultative-stative
Present	<i>mi-rav-ad</i>	<i>raft</i>	<i>raft-e ast</i>
Evidential	<i>mi-[raft-e] ast</i>	<i>[raft-e] ast</i>	<i>[raft-e bud-e] ast</i>
Past	<i>mi-raft</i>	<i>raft</i>	<i>raft-e bud</i>

This evidential mode is confined to the past. Its function and range are similar to those described for Turkish and in general by Johanson (2000). That is, it marks an action or state as non-manifest. This generally includes such qualifications as second-hand knowledge, conclusion, the latter often encountered in historical surveys and personal reminiscence, among others, and is rendered into Western languages by various tags, such as “they say, evidently, apparently”.

be zabān-e Māzandarāni še’r mi-[goft-e] and
 in language-EZ Māzandarāni poetry IPFV.say.PART COP-3p
 ‘(the records suggest that) they used to write poetry in the Mazandarani language’;

yād-am mi-ā-yad
 to memory-my IPFV-come.PR-3s
sāl-hā piš dar bāre-ye in xeyli motāle’e mi-[kard-e] am
 year-p before about-EZ this much research IPFV-do.PART COP.1s
 ‘I remember now that years ago I used to study this very intensively’;

mi-guy-and diruz dar yek tasādoḡ dah nafar [košt-e
 IPFV-say.PR-3p yesterday in one accident ten person kill.PART
šod-e] and
 become.PART COP.3p
 ‘they say yesterday ten people got killed in a single accident’;

ma’lum šod zud-tar [āmad-e bud-e] ast
 evident become.PT.3s soon-COMP come.PART be.PART COP.3s
 ‘it became evident from what he said, that he had come earlier’.

3.2.6a.2 Progressive

The progressive is not yet fully integrated into literary Persian. The means for this innovation is the verb *dār-/dāšt-* ‘to keep, have’ and refers to an event as ongoing

or imminent. It is preposed to the imperfective forms of the main verb, including the evidential imperfect.

TABLE 8.10: PERSIAN PROGRESSIVE, 3s

	Progressive		Imperfective	Perfective	Resultative-stative
PR	<i>dār-ad</i>	<i>mi-rav-ad</i>	mi-rav-ad	raft	raft-e ast
PT	<i>dāšt</i>	<i>mi-raft</i>	mi-raft	raft	raft-e bud
Evidential	<i>dāšt-e ast</i>	<i>mi-raft-e ast</i>	mi-raft-e ast	raft-e ast	raft-e bud-e ast

Functionally, it disambiguates the progressive and imminent-future functions of the imperfective, but has not yet emptied the latter of that function. Moreover, it is indicative only and cannot be negated. It precedes the main verb and may be separated from the latter. Significantly, both auxiliary and main verb are inflected, but may be separated:

al'ān dār-ad *howsele-am* *sar* *mi-rav-ad* (standard)
al'ān dar-e *howsela-m* *sar* *mi-r-e* (colloquial)
 now PROG.PR-3s patience-my head IPFV-go.PR-3s
 'my patience is running/about to run out now';

xorus *dāšt* *dāne* *mi-čid* *ke . . .*
 rooster PROG.PT.3s grain IPFV-pick.PT.3s when
 'the rooster was picking up grains, when . . .';

zāheran *rānande* *dāšt-e* *dowr mi-zad-e* *ke . . .*
 apparently driver PROG.PART-EV turn IPFV-hit.PT-EV.3s when
 'apparently the driver was about to make a turn when . . .'

3.2.6b Extended verb system of Modern Standard Tajik

Tajik has considerably expanded the common Persian system both in number of forms and range of functions, in addition to the inherited system of aspect, tense, and mood. There are three innovative subsystem which intersect with the basic system, and with each other.

Like Persian, Tajik has developed a sub-system of the evidential mode, based on the perfect participle, e.g. *rafta*, but expanded it to include also present tense forms and non-indicative forms.

Likewise similar to Persian, it developed a progressive sub-system which, however, is not based on an inflected converb, but on a doubly participial compound consisting of the perfect participle of the main verb plus the perfect participle of the verb 'stand', reflecting the pervasive participialization of Tajik. Here again, unlike Persian this new system is expanded to the non-indicative and, in one case, to the evidential mode.

In addition, Tajik has a less developed conjectural sub-system (see sections 5.5.3 Aktionsart constructions and 5.5.4 (Tajik conjunct verbs).

3.2.6b.1 Evidential mode

raft-a

There are four complex tense forms all structured on the form of the perfect participle *raft-a*. Similar to their function in Persian, they denote an action or occurrence known to

the speaker not by direct observation but from collateral sources – hearsay or logical inference (quotative, inferential); or a situation unexpectedly discovered, or a fact only just appreciated as relevant (mirative). The indirect nature of the experience may (or may not) be signaled by an epistemic verb such as *šunav-šunid-* ‘hear’ or *xabar gir-ğirift-* ‘be informed’, or a phrase such as *ma’lum šud* ‘it became known’.

3.2.6b.1a Perfect as evidential form

[*raft-a*] *ast* ‘(evidently, etc.) went, left’ and ‘has, is (evidently) gone’

The perfect tense, apart from its role as resultative, also functions as an evidential past (preterite, aorist) as well as evidential perfect (resultative):

mi-gū-yand ki kulliyot-i in navisanda [bar omad-a] ast
 IPFV-say.PR-3p that collected works-EZ this writer out-come.PART-EV COP.3s
 ‘they say that collected works of this writer’s have appeared, are out’;

rūz-i raft-an-aš kasal šud-a boš-ī,
 day-EZ go.INF-his ill become.PART (SBJ.)be.PR-2s

dušanbe [raft-a]-y
 Monday go.PART-EV COP.3s
 ‘if you fell ill on the day of his departure, then it was Monday that he left’
 (inference); [*raft-a*]-y < *rafta-ast*.

Unlike Persian, the perfect of ‘be’ and ‘have’, *bud-a-* and *došt-a-* refers to present time when used in evidential contexts (generally, as miratives):

odam-i xub [bud-a]-ast či zan-i xub-e
 person-EZ good be.PART-COP3s what woman-EZ good-INDEF
[došt-a]-ed!
 have.PART-EV COP.2p
 ‘he’s a good man as it turns out’; ‘what a good wife you have!’;

pul-i mayda-am na-[bud-a]-ast – šumo pul-i
 money-EZ small-to me not be.PART-EV.COP.3s you.2p money-EZ
non-ro dili-ed!
 bread-DO give.IMP-2p
 ‘I don’t have change – you pay for the bread’

(this sentence also exemplifies an idiom of ‘be’ with the sense ‘have’).

3.2.6b.1b Evidential pluperfect

[*raft-a bud-a*]-*ast* ‘(evidently) had gone’

Similar to its function in Persian, this tense corresponds to the regular pluperfect, while connoting quotation, inference or sudden realization:

loiha-i on peš az jang tayyor [šud-a bud-a]-ast
 projectEZ that before from war prepared become.PART be.PART.EV-COP.3s
 ‘his project was evidently ready before the war’, lit. ‘had been readied’.

3.2.6b.1c *Evidential durative*

me-[raft-a]-ast 'is (evidently) going, goes, will go, be going' and 'was (evidently) going, used to go, would go, be going'.

This form appears to be indifferent to tense, and may designate past or present habitual or progressive action, as well as future (or intended) action:

(1) Present:

in navozanda-g-ī xudd-i mullo-g-ī barin [bud-a]-ast-u
 this playing like-EZ mullah-ing like be.PART-EV-COP.3s-and
na-navoz-ed, az xotir me-[bar-omad-a]-ast
 not.SBJ-play-2pfrom memory IPFV.out.come.PART.EV COP.3s
 'playing a musical instrument is like reading-and-writing: if you don't play, you forget how to';

mā'lum-ast ki ū pagoh me-[raft-a]-ast
 known be.PT.3s that he tomorrow IPFV-go.PART.EV-COP.3s
 'it's known that he is going-EV tomorrow'.

(2) Past:

me-gū-yand ki peš-tar boy-ho-i qišloq
 IPFV-say.PR-3p that before-COMP richman-p-EZ village
sol-e yak-du bor ba šikor me-[bar-omad-a]-and
 year-INDEF one-two times to hunt IPFV-out.come.PART.EV COP-3p
 'they say that formerly the rich men of the village used to go hunting once or twice a year'.

3.2.6b.1d *Evidential progressive*

[raft-a istod-a bud-a]-ast 'is (evidently) going' and 'was (evidently) going'.

This is equivalent in tense function to two evidential tenses, the present progressive (*raft-a istoda-am*) and the past progressive (*raft-a istod-a bud-am*):

(1) Present:

Šodi dar xona, ba xob kard-an tayyor-ī [did-a istod-a
 Shadi at home to sleep make.INF preparing see.PART PROG.PART
bud-a] ast
 be.PART.EV COP.3s
 'Shadi is now at home; he is preparing to go to bed'.

šumo yak asar-i naw [nivišt-a istod-a bud-a] ed
 you.p one opus-EZ new write.PART PROG.PART be.PART.EV COP.2p
 '(I see) you are, have been writing a new opus'.

(2) Past:

vay kitob [xond-a istod-a bud-a] ast
 he book read.PART PROG.PART be.PART.EV COP.3s
ki man dar-ro taq taq [kard-a] am
 that I door-DO knocking make.PART COP.1s
 'he was evidently reading a book when I knocked at the door'.

3.2.6b.2 Progressive forms

The progressive forms are marked by the perfect participle of the main verb and the perfect participle *ist-*, *istod-* ‘stand’: *raft-a istoda* followed by the various tenses of ‘be’. By their semantics, the stative verbs, *dor-ldoš-* ‘have’ and *boš-/bud-* ‘be’ do not have progressive tenses. For evidential, subjunctive and conjectural forms of the progressive, see the respective subheadings.

3.2.6b.2a Present progressive

[*raft-a istod-a*]-*ast* ‘is going’

This form refers to an action now in progress, sometimes to habitual action:

man maktub [nivišt-a istod-a]-am
I letter write.PART PROG.PART-COP.1s
‘I am writing a letter’;

bača-ho bozi [kard-a istod-a] and
children play do.PART OROG.PART COP.3p
‘the children are playing’;

mo dar institut fann-ho-i gunogun-ro [omūxt-a istod-a] im
we in institute subject-EZ various-DO learn.PART.PROG.PART COP.1p
‘we are learning various subjects at the institute’.

3.2.6b.2b Past progressive

[*raft-a istod-a*] *bud* ‘was going’

This tense designates an action observed in progress in the past, often at the start of a second action:

Nuralī asb-aš-ro ba mex [bast-a istod-a] bud,
N. horse-hisDO to nail tie.PART PROG.PART COP.1s
ki Hoji-umar namo-yon šud
when H.-U. appear.PR.PART become.PT.3s
‘Nurali was tethering his horse to the stake when Hojjumar appeared’.

3.2.6b.3 Subjunctive forms

3.2.6b.3a Durative perfect subjunctive

[*me-raft-a*] *boš-* ‘may have gone’ and ‘might have gone’

The durative subjunctive is indifferent to present and past distinction, and expresses conjecture, doubt, apprehension, regret, etc. in relation not only to progressive, habitual or iterated actions in the past, but also to current or potential actions of a durative nature:

(1) Present

ba kujo [me-raft-a] boš-ad?
Ro where IPFV.go.PART (SBJ)be.PR-3s
‘where might she be going (I wonder)?’

(2) Past

boyad šodī ham az in kor-ho-i modar-i mehrubon-aš
 must Sh. also from these doing-p-EZ mother-EZ kind-his
zavq [me-girift-a] boš-ad
 pleasure IPFV.take.PART (SBJ.)be.PR-3s
 ‘Shodi must have derived pleasure from everything his dear mother did’.

3.2.6b.3b Present progressive subjunctive

[raft-a istod-a] boš-ad ‘may be doing’

This tense is rarely used in MLT, but is widespread in Northern dialects. It has similar modal connotations to the past (perfect) subjunctive and the durative subjunctive, in respect of a progressive action in past or present-future time, and occurs in concessive clauses:

on-ho az mašina xele dur kor [kard-a-istod-a] boš-and ham,
 they from machine very far work do.PART PROG.PART (SBJ.)be.PR-3p also
šavqun-i xomūš-na-šav-anda-i on na-me-mond
 noiseEZ silent-not-becoming-EZ that not-allow.PT.3s
ki ovoz-i hamsoya-i xud-ro šunav-and
 that saying-EZ neighbor-EZ self-DO (SBJ.)hear.PR-3p
 ‘although they were working quite far from the machine,
 its constant noise did not allow them to hear what the person next to them was
 saying’.

3.2.6b.4 Conjectural mood

raft-agī-st- ~ raft-agī-y-

This mood is used to express an unsubstantiated conjecture or assumption. It is a peculiarity of Northern dialects that has been assimilated into MLT. It is constructed upon the past, or perfect participle, in *-ag-ī* and forms of ‘be’. Only present tense forms are used. These are either formed with the contracted existential verb: *raftagī-st-* (*-st < hast-*), or with the copula: *raftagī-y-*. Colloquially, these forms are contracted (*-agī + personal endings beginning with a- and e- > -agī- > -age-*, and further all *> -agi-*):

<i>raftagī-st-am</i>	<i>raftagī-st-ī</i>	<i>raftagī-st</i> (zero)	<i>raftagī-st-em</i>	<i>raftagī-st-ed</i>	<i>raftagī-st-and</i>
<i>raftagī-yam</i>	<i>raftagī-yī</i>	<i>raftagī-st</i> (<ast)	<i>raftagī-yem</i>	<i>raftagē-yed</i>	<i>raftagī-yand</i>
<i>raftagem</i>	<i>raftagī</i>	<i>raftage</i> (<a(st))	<i>raftagem</i>	<i>raftag-et(on)</i>	<i>raftagen</i>
<i>raftagim</i>	<i>raftagi</i>	<i>raftagi</i>	<i>raftagim</i>	<i>raftagid,-it(on)</i>	<i>raftagin</i>

While theoretically the conjectural mood may have all tense, modal, and aspectual forms, only four forms are used in Tajik.

3.2.6b.4a Past conjectural

[raft-agī]-st ~ raft-age ‘I suppose he went’

The tense connotes various degrees of supposition or conjecture, and may be translated with the help of an adverb such as “probably” or modal idioms “may, might, must have”

<i>ū az šahr [omad-agī]-st</i>	<i>faromūš kard-aged (< [kardagī]-yed)</i>
he from city come.CNJECT-COP	forgotten make.CNJECT-COP.2p
‘he may have come from the city’;	‘you’ve probably forgotten’.

However, the tense of the verb ‘be’ invariably has present reference:

agar hamon faranjidor Pūlod na-boš-ad, mayonarav-aš [bud-agī]-st
 if that-same veiled figure Pulād may not be go-between-his has been-EV
 ‘if that veiled figure is not Pulad, it must be his go-between’.

3.2.6b.4b Present-future conjectural

[me-raft-agī]-st ‘might be going, about to go’

Constructed with the imperfective prefix *me-*, this form expresses a conjecture about a potential or a current (habitual or iterated) action:

pagoh [me-omad-agī]-st
 tomorrow IPFV.come.CNJECT-COP.3s
 ‘he’ll probably come tomorrow’;

gazeta-i muallim-on-ro ū har rīz [me-xond-agī]-st
 journal-EZ teacher-p-DO he each day IPFV.read.CNJECT-COP.3s
 ‘I’ll bet he reads the Teachers’ Gazette every day’.

3.2.6b.4c Present progressive conjectural

[raft-a istod-a]-gī-st- ‘might be going’

This progressive form corresponds to the indicative present progressive:

balki vaḡ dar yagon čariš ba xun-u xok-aš oyušt-a šud-a,
 perhaps he in some meadow in blood-and dirt its welter.PART become.PART
jon [kand-a istod-agī]-st
 soul tear.PART PROG.CNJECT-COP.3s
 ‘quite likely, in some meadow, weltering in gore and grime, he lies dying’.

Dialect reflexes of these forms are subject to various contractions, e.g. *kar-soda-y < kard-a istod-a ast* ‘might be doing’ *raf-sod-ag-em < raft-a istod-ag-ī-am* ‘might be going’.

3.2.6b.4d Pluperfect counterfactual conjectural

[raft-agī] me-bud ‘(if) he might have gone’

This form extends the range of the conjectural to parallel the pluperfect in its counterfactual function, further marked by the the prefix *me-* as a non-indicative marker (see also Imperfect section 3.2.2.2d).

agar čašm-i Yodgor-ro andeša-i oyanda-i sioh torik
 if eye-EZ Y.-DO thought-EZ future-EZ black dark
na-[kard-agī] me-bud
 not make.CNJECT IPFV.be.PT.3s
 ‘if the thought of a black future had not darkened the vision of Yadgar . . .’ ~
 ‘if Yodgor’s vision had not been clouded by the prospect of a black future . . .’

3.2.6b.5 Synopsis of Tajik verb forms

The following is an overview of the innovated forms. Note that several forms are tense neutral, that is, they may refer to both a present and past context.

TABLE 8.11: TAJIK INNOVATED VERB FORMS, 3s

<i>Progressive forms</i>		
Present progressive	<i>raft-a istod-a-ast</i>	
Past progressive	<i>raft-a istod-a bud</i>	
(see also evidential, subjunctive and conjectural forms)		
<i>Evidential mode</i>		
Evidential durative	<i>me-raft-a-ast</i>	(tense neutral)
Evidential progressive pluperfect	<i>raft-a istod-a bud-a-ast</i>	(tense neutral)
<i>Subjunctive forms</i>		
Durative perfect subjunctive	<i>me-raft-a boš-ad</i>	(tense neutral)
Present progressive subjunctive	<i>raft-a istod-a boš-ad</i> (rare)	(tense neutral)
<i>The conjectural mood</i>		
Present-future conjectural	<i>me-raft-agī-st</i>	
Past conjectural	<i>raft-agī-st</i>	
Pluperfect counterfactual conjectural	<i>raft-agī me-bud</i>	
Present progressive conjectural	<i>raft-a istod-agī-st</i>	

3.2.7 Synopsis of Persian and Tajik verb systems

Table 8.12 shows the inherited basic verbs forms (in italics) and the innovative forms (non-italic) and their categories; several forms are listed more than once according to their functions. The arrangement of the table highlights the fact that in Tajik the progressive and the evidential mode have become major parameters of the system, intersecting with the tense parameter, present and past, here represented by pairs of tense forms in each subset. Note also that the conjectural forms constitute a subset together with the imperfective conditional (imperfect form in its counterfactual function).

4 SYNTAX AND USE OF FORMS I: NOUN PHRASE STRUCTURE

The NP may consist minimally of a single nominal, which may be a noun, independent pronoun, infinitive: *xānd-an āsān ast/xond-an oson ast* ‘reading is easy’, or any item that is nominalized or used in nominal function.

4.1 Coordination

4.1.1 Single and paired conjunctions

The two basic conjunctions are: (1) coordinating *valva*, enclitic form *-o, -vol-u, -yu,-v-u*, in open and closed compound, respectively, e.g. *man va tolman va tu* ‘I and you’ vs. *man-o tolman-u tu* ‘I-and-you’, so in fixed compounds such as number, e.g. *bist-o dolbist-u du* ‘22’; (2) disjunctive *yāl/yo* ‘or’.

Paired coordinators of homogeneous NPs (and VPs) include: *valva . . . valva* ‘both . . . and’, *hamlham . . . hamlham* ‘both . . . and’, *čelčī* ‘whether . . . whether, both . . . and’; *yāl/yo . . . (va) yāl(va) yo* ‘either . . . or’; *xāhl/xoh . . . xāhl/xoh* ‘be it . . . or, whether . . . or’; *nalna . . . (va) nal(va) na* ‘neither . . . nor’:

TABLE 8.12: SYNOPSIS OF VERB SYSTEMS, 3s

Stem forms: *rav-*, *raft-*, *raft-a* 'go, leave'; = same in form; * rarely used

PROGRESSIVE RAFTA ISTODA	BASIC	EVIDENTIAL RAFTA
IMPERATIVE, 2p		
PR	<i>rav-ed</i>	
PT	<i>rafta boš-ed</i>	
SUBJUNCTIVE		
PR * <i>[rafta istoda]</i> boš-ad	<i>rav-ad</i>	me- <i>[rafta]</i> boš-ad =
PT	<i>rafta boš-ad</i>	me- <i>[rafta]</i> boš-ad =
IMPERFECTIVE		
<i>Indicative</i>		
PR <i>[rafta istoda]</i> -ast	<i>me-rav-ad</i>	me- <i>[rafta]</i> -ast =
PT <i>[rafta istoda]</i> bud	<i>me-raft</i>	me- <i>[rafta]</i> -ast =
<i>[rafta istoda]</i> buda-ast EV		<i>[rafta]</i> istoda buda-ast PROG
<i>Non-Indicative</i>		
Conjectural - <i>agī-st-</i>	Counterfactual	Conjectural - <i>agī-st-</i>
PR * <i>[rafta istoda]</i> gī-st	<i>me-raft</i> =	me- <i>[rafta]</i> gī-st
PT	<i>me-raft</i> =	<i>[rafta]</i> gī-st
PERFECTIVE		
Preterite/Aorist	<i>raft</i>	<i>[rafta]</i> -ast =
RESULTATIVE-STATIVE		
<i>Indicative</i>		
PR PF	<i>rafta ast</i> =	<i>[rafta]</i> -ast =
PT PF	<i>rafta bud</i>	<i>[rafta]</i> buda-ast
<i>Non-Indicative</i>		
	Subj., Counterf.	Conjectural - <i>agī-</i>
PR PF	<i>rafta boš-ad</i> SBJ	
PT PF	<i>rafta (me-)bud</i> C-F	<i>[rafta]</i> gī me-bud

PERSIAN

magar [na man na to] lāyeq-e in kār hast-im?
 but not I not you/s fit-EZ this work EX-I
 'but are neither I nor you are fit for this task?';

TAJIK

ba in koyaz-(h)am [na dast-i to] ras-ad, [na dast-i man]
 with this paper also not hand-EZ you.s (SBJ.)reach-3s not hand-EZ me
 'and let [neither my hand] [nor yours] touch this paper'.

The contrastive construction *na tanhāl/na tanho . . . balkelbalki ~ hamlham* 'not only . . . but also' may link NPs, VPs or clauses:

[na tanhā] man, [balke] doxtar-am [ham] u-rā didl
[na tanho] man, [balki] duxtar-am [ham] ū-ro did
 not alone I but daughter-my also he-DO see.PT.3s
 'not only I, but my daughter too saw him'.

Adjectives are similarly concatenated, or are connected by *-e/i* if they qualify the preceding N + ADJ phrase:

<i>ketāb-e bozorg o mofid</i>	<i>zan-e zibā-ye mahru</i>
<i>kitob-i buzurg u mufid</i>	<i>zan-i zebo-i mahrū(y)</i>
book-EZ big and useful	woman-EZ beautiful-EZ moon-faced
'a big and useful book';	'a beautiful moon-faced woman'.

4.1.2 Concatenation and group inflection

Nouns and noun phrases (and any other constituents) may be concatenated, typically by adding *-o/-u*, to each constituent except the last, or to the penultimate one alone, or without addition. Grammatical markers such as the direct object *-rā/-ro* are normally added only to the last of several closely coordinated group of nominals:

[u(-v)-o mā-o šomā]-rā mo'ayyan kard-and
[vay-u mo-v-u šumo]-ro ta'in kard-and
 he-and we-and you.p-DO identified do.PT-3p
 'they designated [him and us and you];

[ketāb, daftar va qalam]-rā ru-ye miz gozāšt-am
[kitob, daftar va qalam]-ro rū-i stol mond-am
 book notesbook and pen-DO face-EZ table place.PT-1s
 'I laid [the book, notebook and pencil] on the table'.

4.2 Constituents following the head: *Ezāfe*

Dependent nominals, including adjectives, nouns and nominal phrases, follow the head noun and are linked in the EZ construction by the connective clitic *-e/i*, as described under Morphology, allowing multiple embedding.

4.2.1 Basic structure

4.2.1.1 NP levels

The EZ noun phrase has two levels and two basic types, depending on the syntactic status of the complement:

(1) adjectival, noun + adjective:

ketāb-e bozorg
kitob-i buzurg
 book-EZ big
 'big book'.

(2) compound, noun (+ adjective) + noun (+ adjective):

ketāb-e pedar
kitob-i pidar
 book-EZ father
 'book of father/father's book'.

(3) phrasal: noun + adverbial phrase:

qarn-hā-ye piš az eslāmī
sada-ho-i peš az islom
 centuries-EZ before from Islam
 ‘the centuries before Islam’.

(4) clausal: noun + nominalized phrase or clause, prominently in Tajik:

pesar-e [az madrase gorixt-e]
pisar-i [az madrasa gurexta-e]
 boy-EZ from school flee.PART
 ‘boy fled from school’, i.e. ‘truant boy’.

4.2.1.2 NP clitics

There are three sets of NP clitics which attach according to level:

(1) the indefinite *-il-e*, (2) the personal enclitics, (3) the postposition *-rāl-ro*. The former two function on the level of the adjectival noun phrase, the latter functions on the level of the NP as a whole.

(1) Thus, as a rule, *-il-e* is attached to the singular and plural N, but to the adjective in an adjectival EZ:

<i>so’āl-il</i>	<i>so-āl-hā-il</i>	<i>[so’āl-e xub]-i</i>	<i>[so’āl-hā-ye xub]-i</i>
<i>suvol-e</i>	<i>suvol-ho-e</i>	<i>[suvol-i xub]-e</i>	<i>[suvol-ho-i xub]-e</i>
‘a question’,	‘some/certain questions’;	‘a good question’,	‘some, certain good questions’.

In case of several adjectives it may be attached to the noun directly, without *-il-e*:

Esfahāni-hā [mardom-ān-e besyār mehmāndust]
 Isfahani-p people-p-EZ very hospitable
va mehrabān, samimi va širin-zabān]-i and
 and kind sincere and sweet-tongued-INDEF are

vs.

Esfahāni-hā [mardom-ān]-i hast-and [besyār mehmāndust]
 Isfahani-p people-p-INDEF are very hospitable
va mehrabān, samimi va širin-zabān]
 and kind sincere and sweet-tongued
 ‘the Isfahanis are (a) very hospitable, kind, sincere and sweet-tongued people’.

(2) The pronominal clitics are similarly attached:

<i>so’āl-aml</i>	<i>so-āl-hā-yaml</i>	<i>[so’āl-e xub]-at</i>	<i>[so’āl-hā-ye xub]-at</i>
<i>suvol-am</i>	<i>suvol-ho-yam</i>	<i>[suvol-i xub]-at</i>	<i>[suvol-ho-i xub]-at</i>
‘my question’,	‘my questions’;	‘your good question’,	‘your good questions’.

These two sets are mutually exclusive.

(3) The postposition *-rāl-ro* is clitic to the NP as a whole in rightmost final position irrespective of its internal structure, be it a single N, simple NP or an extended NPP as a whole.

The basic EZ structures are thus as follows (N = noun, A = adjective, NPP = compound noun phrase; 3s -aš exemplifies personal clitics; all with nouns in the singular):

TABLE 8.13: CLITIC ATTACHMENT

N	=	<i>ketāb</i> [-aš or -i] [-rā]/			
		<i>kitob</i> [-aš or -e] [-ro]			
		book his, her a	DIROBJ		
		'his, her book' or 'a book'.			
NA	=	<i>[ketāb-e bozorg]</i> [-aš or -i] [-rā]/			
		<i>[kitob-i buzurg]</i> [-aš or -e] [-ro]			
		book-EZ large his, her a	DIROBJ		
		'his, her large book' or 'a large book'.			
NPP	=	<i>ketāb-e bozorg-e</i> [mo'Allem-e javān] [-aš or -i] [-rā]/			
		<i>kitob-i buzurg-i</i> [mu'allim-i javon] [-aš or -e] [-ro]			
		book EZ large EZ teacher EZ young	his, her or a	DIROBJ	
		'the large book of his young teacher'.			

4.2.1.3 Alternate attachment of indefinite marker

The noun alone (or with plural marker) without any determiner, modifier, or enclitic, is either non-specific or definite, and can be used either in a generic or non-generic sense. Without a determiner they are usually indefinite.

While *-i/-e* is usually added to the end of the EZ phrase, in the adjectival EZ it may optionally be attached to the head noun or NP, in which case the EZ particle is not used:

<i>kas-e</i>	<i>digar-il</i>	or	<i>kas-i</i>	<i>digarl</i>
<i>kas-i</i>	<i>digar-e</i>		<i>kas-e</i>	<i>digar</i>
person-EZ other INDEF			person-INDEF other	
'another person'			'someone else'.	

(1) *-il/-e*: As discussed under Morphology (section 3.1.1.2), both singular or plural nouns may be followed by the indefinite (specific or non-specific) clitic:

<i>ketāb-il/kitob-e</i>	<i>ketāb-hā-il/kitob-ho-e</i>
'some ~ any book' or	'some ~ any books' or
'a book';	'some ~ certain books'.

(2) *yek/yak*. Alternatively, the noun may be preceded by the quasi-indefinite article *yek/yak* 'one, a' and may further be followed by the indefinite enclitic *-il/-e*. The admission of plural and the indefinite enclitic distinguishes the quasi-article *yek/yak* from its role as numeral:

Singular:	Plural:
<i>[yek] fenjān[~i-]</i>	<i>[yek] fenjān-hā [~i-]</i>
<i>[yak] payola [~e]</i>	<i>[yak] payola-ho [~e]</i>
'a (certain) cup';	'some (certain) cups'.

Tajik also allows plural without enclitic *-i*: *[yak] piyola-ho*.

4.2.2 Types of noun phrases and complements

4.2.2.1 Overview

Fundamental for identifying the function of the *ezāfe* construction is the fact that it encompasses not only nouns, but also adjectives. Heads may be nouns, including verbal nouns, as well as adjectives and participles, prominently those belonging to the Arabic loan component and, rarely, a pronoun. Complements may similarly be nominal and adjectival, or participial, or may be sentential nominalizations. (A good number of the latter is further reduced to quasi-compound words; see section 7.1 Word formation.)

<i>kār-e</i>	<i>pesar, u, sāxtan</i>	<i>kār-e</i>	<i>moškel, anjām-šodel</i>
<i>kor-i</i>	<i>pisar, vay, soxtan</i>	<i>kor-i</i>	<i>muškil, anjom-šuda</i>
work-EZ boy,	he, building	work-EZ difficult	having been completed
‘the work of the boy, him, building’;		‘difficult, completed work’.	

In Tajik, moreover, the nominal EZ in combination with verbal nouns, especially the infinitive and other non-finite verb forms, has greatly extended its range of application (arguably under the influence of Uzbek), and may condense into a single sentence what would appear more naturally in Persian or English as a subordinate VP (see section 6.5.4 Tajik Turkic-type nominal clauses).

4.2.2.2 Nominal heads and valence relationships

The nominal EZ can be seen as essentially the nominalization of underlying case relationships, or neutralized valences:

Predicate	<i>āsmān-e ābilosmon-i obi</i> ‘blue sky’;
Event	<i>ruz-e enqelābrūz-i inqilob</i> ‘the day of revolution’ ~ ‘revolution day’;
Possessor	<i>ketāb-e Hasanlkitob-i Hasan</i> ‘the book of Hasan’ ~ Hasan’s book’;
Agent	<i>kār-e mardomlkor-i mardum</i> ‘the work of people’;
Patient	<i>qatl-e Hoseynlqatl-i Husayn</i> ‘the murder of Hoseyn’;
Purpose	<i>dāru-ye geripldāru-yi gripp</i> ‘influenza medicine’;
Goal	<i>rāh-e Tehrānlroh-i Dušanbe</i> ‘the road of/to Tehran, Dushanbe’ ~ Tehran, Dushanbe road’;
Location Time	<i>mardom-e injā, emruzlmardum-i in jo, imrūz</i> ‘people (of) here, of today’;
Origin	<i>ahl-e Tehrānlahl-i Dušanbe</i> ‘inhabitant of Tehran, Dushanbe’;
Source, Cause	<i>āb-e češmelob-i čašma</i> ‘water of well’ ~ well-water’, <i>bim-c mowj</i> ‘fear of the waves’;
Substance	<i>gombad-e talālgunbad-i- talo</i> ‘dome of gold’;
Element	<i>anbuh-e sā’el-ānlanbūh-i so’il-on</i> ‘crowd of pilgrims’;
Part	<i>do nafar-e ān-hāldu nafar-i on-ho</i> ‘two (persons) of them’.

Predication includes the following relationships, also marked by *-el-i*:

Genus-species	<i>gol-e sorx/gul-isadbarg</i> ‘rose (flower)’;
Class-gender	<i>dust-e doxtar, pesar/dūst-i duxtār, pīsar</i> ‘girl-, boyfriend’;
Class-name	<i>asb-e nar, mādelasb-i nar, moda</i> ‘male, female horse’; <i>šahr-e Tehrān/šahr-i Dushanbe</i> ‘Tehran/Dushanbe city ~ city of Tehran/Dushanbe’;
Address-last name	<i>āqā-ye, xānom-e Javādīljanob-i, xonum-i Javodī</i> ‘Mr., Mrs. Javadi’
Address-first name	<i>āqā-ye Ahmadljanob-i Ahmad</i> ‘Mr. Ahmad’, <i>xānom-e Maryam/xonum-i Maryam</i> ‘Mrs. Maryam’;
First-family name	<i>Ahmad-e Javādīl/Ahmad-i Javodī</i> ‘Ahmad Javadi’, <i>Maryam-e Javādīl/Maryam-i Javodī</i> ‘Maryam Javadi’;
Series-number	<i>dars-e panj, panj-om/dars-i panj, panj-um</i> ‘lesson five, fifth lesson’; <i>sā’at-e selso’at-i se</i> ‘three o’clock’.

A typical result of neutralization may be overt identity: *dust-e doxtar/dūst-i duxtār* may be both predicative ‘girlfriend’ and ‘a friend who is a girl’, as well as possessive ‘a friend of the girl’. Similarly, *ešq-e dust-ān/išq-i dūst-on* may be both subjective ‘love by the friends’ and objective ‘loving (the) friends’.

4.2.2.3 Possessor

4.2.2.3a Preposed possessor

The possessor can be topicalized by inversion and anaphoric pronominal clitic:

<i>esm-e tolistm-i to</i> name-EZ you.s	<i>ketāb-e Hasan/kitob-i Hasan</i> book-EZ H.
>>	
<i>[to] esm [-at]/tu ism-at</i> you.s name-your-s ‘your name’;	<i>[Hasan] ketāb [-aš]/Hasan kitob-aš</i> H. book-his ‘Hasan’s book’.

Syntactically such NPs are clause initial:

<i>be-bin-am,</i>	<i>[to] esm[-at]</i>	<i>či-st?!</i>
<i>bin-am,</i>	<i>[tu] ism[-at]</i>	<i>či-st?</i>
SBJ.see.PR-2s you.s name-your-s what-COP.3s		
‘let me see, what’s [your] name?’		

4.2.2.3b Tajik Turkic-type possessor inversion

In the Northern dialects of Tajik, a topical possessor inversion may serve to bind a complex NP such as a reduced relative clause. This construction uses the postposition *-a* (<*-ro*) which at earlier stages of Persian had also a dative function, but has become the unmarked construction, essentially a Turco-form version of the EZ (see Postposition *-rā, -ro* section 3.1.4.2):

<i>kitob-i in mard</i>	‘book of this man’ (Tajik) >
book-EZ this man	
<i>in mard[-a] kitob-aš</i>	‘this man-to book-his’.
this man[-RA] book-his	

4.2.2.4 Adjectives and ordinals

4.2.2.4a Superlatives and ordinals

These two canonically precede the head:

<i>behtar-in ketābl</i>	<i>avval-in ruz-e tābestānl</i>
<i>bihtar-in kitob</i>	<i>avval-in rūz-i tobiston</i>
‘the best book’;	‘the first day of summer’.

In Persian, superlative adjectives always precede the head, as do ordinals when in focused position.

<i>bozorg-tar-in māšin</i>	vs. <i>māšin-e bozorg-tar</i>	‘the largest car’ vs. ‘the larger car’;
<i>dovv-om-in sālgard</i>	~ <i>sālgard-e dov-om</i>	‘the second anniversary’.

In Tajik, (1) superlatives, (2) ordinal adjectives in *-in*, (3) simple ordinal adjectives in *-um*, (4) adjectives in *-ngī*, and (5) phrases with the postposition *barin* may either precede or follow the head:

<i>kalon-tar-in šahr</i>	~ <i>šahr-i kalon-tar-in</i>	‘the biggest town’;
<i>oxar-in bor</i>	~ <i>bor-i oxar-in</i>	‘the last time’;
<i>du-yūm poyezd</i>	~ <i>poyezd-i du-yūm</i>	‘the second train’;
<i>dina-ngī rūz-noma</i>	~ <i>rūz-noma-yi dina-ngī</i>	‘yesterday’s paper’;
<i>man barin šaxs</i>	~ <i>šaxs-i man barin</i>	‘a person like me’.

4.2.2.4b Preposed affective adjectives

When used affectively or evaluatively, adjectives may precede the noun, notably when exclamatory. When evaluative, Persian often has indefinite *-i*:

<i>dust-e bičāre-ye man!</i>	<i>pesar-e xub(-i) ast!</i>
<i>dūst-i bečora-i man</i>	<i>pisar-i xub buda ast</i>
friend-EZ poor-EZ I	boy-EZ good(INDEF) is/seems to be
>	>
<i>[bičāre] dust-e man!</i>	<i>[xub] pesar-i ast!</i>
<i>[bečora] dūst-i man</i>	<i>[xub] pisar bud-a ast</i>
poor friend-EZ I	good boy is/appears to be
‘my poor friend!’; ‘he’s/seems to be a good boy’, non-evidential-mirative.	

4.2.2.4c Pronouns as heads

Pronominal heads of EZ constructions are marginally productive and mostly confined to deplorable qualifications: *man-e bičāre*, *falak-zademan-i be-čora*, *falak-zada* ‘poor, fate-stricken me’, etc.

4.2.2.5 Nominalized complements

4.2.2.5a Verbal nouns as heads

Verbal nouns, in both Persian and Arabic, behave essentially like nouns, those derived from transitive verbs exhibiting ambivalence between agent and patient, i.e. subject and direct object (here using roughly equivalent Persian infinitives and Arabic verbal nouns):

Transitive	Intransitive
<i>bar gozid-an-e ~ entexāb-e Hasan</i>	<i>rasidan-e ~ vorud-e vazir</i>
<i>bar guzidan-i ~ intixob-i Hasan</i>	<i>rasidan-i ~ vurud-i vazir</i>
‘choosing ~ choice-EZ H; both ‘choice by Hasan’ and ‘choosing Hasan’.	‘the arriving ~ arrival-EZ minister’.

Archaizing diction allows the direct object to be attached (e.g. in sub-headings):

<i>košt-an-e Rostam [Sohrāb-rā]</i>	
<i>kuštan-i [Suhrob-ro]</i>	
kill.INF-EZ R. S.-DO	
‘Rostam’s killing of Sohrab’, lit. ‘the Sohrab killing of Rostam’.	

4.2.2.5b Compound participles and adjectives as heads

A few adjectives and a relatively limited set of participles forming compound verbs behave similarly to nouns and verbal nouns in neutralizing valence relationships. They consist mainly of Arabic active and passive participles. Common examples include the following (see Ahadi 2001: 105–109):

Persian adjectives	
<i>dočār beldučor ba</i> ‘affected by’:	<i>dočār-e eškāl ast</i> ‘is encountering difficulties’;
<i>šifte az/šeftra az</i> ‘enamored by’:	<i>šifte-ye češm-hā-yaš ast</i> ‘is enamored by her eyes’;
Arabic participles	
<i>motavallid dar/mutavallid dar</i> ‘born in’:	<i>āšeq bālošiq bo</i> ‘being in love with’:
<i>motavallid-e Irān ast</i>	<i>āšeq-e doxtar ast</i>
<i>mutavallid-i Tojikiston ast</i>	<i>ošiq-i duxtar ast</i>
‘was born in Iran/Tajikistan’;	‘is in love with the girl’;
<i>-rā monker/ro munkir</i> ‘denying s.th.’:	<i>montazer-elmuntazir-i</i> ‘waiting for’:
<i>monker-e xodā ast</i>	<i>montazer-e dust ast</i>
<i>munkir-i xudo ast</i>	<i>muntazir-i dūst ast</i>
‘is a denier of God’;	‘is expecting the friend’.
<i>tahvil beltahvil ba</i> ‘handing over to’:	
<i>tahvil-e u kard</i>	
<i>tahvil-i vay kard</i>	
‘he handed it over to him’.	

A distinct subtype are reduplicated adjectives:

xub-e xub/xub-i xub ‘very good’, lit. ‘good of good’, etc.

4.2.2.5c Participial phrases as complements

(See also section 6.3.7 Tajik nominalized relative clauses)

Participial relative clauses as complements tend to be idiomized compounds and adjectival, particularly in Persian:

<i>mard-e</i> [dast-šekast-e]	<i>ādam-e</i> [kešti-šekast-e]
<i>mard-i</i> [dast-šikast-a]	<i>odam-i</i> [kašti-šikast-a]
man-EZ hand-break.PART	person EZ ship-break.PART
‘the man with a broken hand’;	‘ship-wrecked person’;
but not * <i>mard-e māšīn-suxte</i> ‘the man with a burned motor’.	

Prepositions can be retained:

forsat-hā-ye [az dast raft-e]
fursat-ho-i [az dast raft-a]
 opportunities EZ from hand give.PART
 ‘missed opportunities’, lit. ‘given out of the hand’;

ham-ān pesar-e [az madrase gorixt-e]-i *bud-am ke...*
 very same boy-EZ from school flee.PART-INDEF be.PT-1s who ...
 ‘I was again that same truant boy who ...’

4.2.2.5d Tajik nominalized relative clauses

In Tajik copular relative clauses may be nominalized as a possessive EZ construction:

mard(-e) ki mū-yaš safed ast ‘a man whose hair is white’ >
mard-i [mū-yaš safed]
 the man-EZ hair his white
 ‘the white-haired man’.

4.2.2.5e Tajik constructions with participle in -agī

While in Persian participial constructions are limited, Tajik has constructions with the participle in *-agī* that are fully productive (see also section 6.3.7.1):

korgar-on-i [az Evropa boz-gašt-agī]
 worker-p-EZ from Europe return.PARTII
 ‘workers having returned from Europe’,
 cf. Persian: *dānešju-yān-e* [az farang bar-gašt-e]
 ‘students having returned from Europe’.

4.3 Constituents preceding the head noun

4.3.1 Overview

Determiners including demonstratives, indefinite and interrogative pronouns, quantifiers and numerals with or without classifiers and their interrogatives, and others precede the head noun:

DEM ~ INDEF-INTERROG – CLASS – NUM – NOUN

Examples are: *ān mardlon mard* ‘that man’; *har haftelhar hafta* ‘every week’; *če now’/čē xe?* ‘what sort?’; *do kilo guštldu kilogramm gūšt* ‘two kg of meat’, etc. (see also examples under Morphology).

4.3.2 *Classifiers*

4.3.2.1 Numerative classifiers

Numerative classifiers constitute a small set of terms that follow the cardinal number or the indefinite and numeral interrogative *čandlčand* ‘several; how many?’ and precede the head nominal without any connective marker: Numeral – Classifier – Noun.

In Persian the head noun is in the singular following the singular rule after numbers, while Tajik allows the plural as an option when the numerand denotes a conventionally or contextually defined group. A classifier is obligatory in the absence of a head nominal, but optional otherwise. The choice of these classifiers is conditioned by the semantic class of the counted nominal, with the basic distinction of human and non-human.

The general semantically unmarked classifier is unstressed *tā/to* ‘unit, item’. In Persian, there is the distinction of plurality vs. singularity: *tā* follows numbers larger than one, *do tā pesar, medād* ‘two boys, pencils’, *čand tā pesar, medād* ‘several boys, pencils’ and ‘how many boys, pencils’, as opposed to *dāne* ‘grain’ for single non-human items, *yek dāne medād* ‘a single pencil’.

In Tajik, *to* (colloquial *ta*) has no such restrictions and is widely used for all classes of numerands and all numbers: *yak to zan* ‘one [item] woman’, *sad to kurta* ‘a hundred shirts’, while *dona* ‘grain’ is the classifier for smallish inanimate objects.

Most semantically distinguished classifiers are the same in Persian and Tajik, which once constituted an extensive series; some of those still in use are the following:

(1) Human: *nafar/nafar* ‘individual’ (Ar.) and *tan* ‘body, person’, *haft nafar pirezanihaft nafar kampir* ‘seven old women’; *panj tan/panj tan* ‘five (bodies)’; Tajik plural numerand of contextually definite group:

az hamin šaš nafar-on faqat ba du nafar
 from this-same six persons only to two person
dast na-me-ras-on-em
 hand not-IPFV.reach.CAUS.PR-1p
 ‘of these six men we will leave unharmed only two’.

(2) Prominent part of whole, larger animals: (a) *sar/sar* ‘head’ or (b) *ra’s/ra’s* (Ar.), *se sar boz/se sar buz* ‘three head of goats’.

(3) Prominent part for whole, inanimate: (a) houses, *bāblbob* ‘door’ (Ar.), *dah bāb xānel dah bob xona* ‘group of ten houses’; (b) books, *jeld/jild* ‘skin, book cover’ (Ar.), *se jeld ketāblse jild kitob* ‘three volumes’.

(4) Small-sized objects, including small animals, plants, and other smallish things: (a) *dānel dona* ‘grain’ *dah dona bodring dih-ed* ‘give (me) ten cucumbers’, *čand dona anor* ‘several pomegranates’ and ‘how many pomegranates?’, or (b) *adad* ‘number’ (Ar.), *se adad mix* ‘three nails’.

(5) Shape: *taxte* ‘flat board’, etc., *yek taxte qāli* ‘one (item) ‘carpet’.

(6) Small sets: (a) Pers. *dast* ‘hand’, and *yek dast lebās* ‘one suit’, *se dast sandali* ‘set of three chairs’; (b) small groups (animal, people, objects): Pers. *daste* ‘handful’, mostly with EZ, *yek daste(-ye) gol* ‘a bunch of flowers’; Tajik has *dasta* for both (a) and (b) without and with EZ, *yak dasta gul, dasta-i jinoyatkor* ‘a criminal group’.

(7) Buildings, apartments: *dastgāh/dastgoh* ‘compound unit, set’, *se dastgāh emārāt/se dastgoh imorat* ‘three buildings, flats’.

The classifier is obligatory only if the number is mentioned without the numerand: *čand (tā) ketāblčand (to) kitob – se tā/se to* ‘how many books – Three’, *šomā čand nafar bud-id?!šumo čand nafar bud-ed? – se nafar/se nafar* ‘how many were you? – Three’.

Numbers and classifiers may also be joined with a dependent nominal or pronouns, frequently enclitic: *har do-(tā)-yaš qašang ast* *har du-(to)-yaš naʔz ast* ‘both of them’, lit. ‘of it’, ‘are nice’, *har se tā-yešān* *har se to-yašon* ‘all three of them’, *yek-dāne-yaš-rā bi-yandāz/lyak-to-yaš-ro parto* ‘toss me one of them’.

4.3.2.2 Mensurative classifiers

Mensurative classifiers identifying weight, length, and other measures designate a portion of a mass and combine freely with mass nouns. Syntactically they behave like numerative classifiers: *do fenjān čāy/du payola čoy* ‘two cups (of) tea’, *čahār-o nim kilo sib-zaminilčoru nim kilo kartoška* ‘two and a half kilo (of) potatoes’, *panjāh kilometr rāhl panjoh kilometr roh* ‘fifty kilometers (of) way, road’.

Mensurative noun phrases contrast with partitive EZ constructions: *se qatre xun/se qatra xun* ‘three drops of blood’ vs. *se qatre-ye xun/se qatra-i xun* ‘three drops of the, that blood’.

4.3.2.3 Kind and manner classifiers

Kind and manner classifiers likewise behave like numerative classifiers. These include *now/nav* ‘kind, sort; breed, species’ way, method, manner, mode’: Pers. colloq. *tovr/taʔvr* ‘sort, kind, type; manner, style’; Pers. *jur*, ‘sort, kind, variety; manner’; *raqam/raqam* ‘number’; ‘type, kind, sort’: *če now parandelči naw paranda* ‘what kind of bird?’; Pers. *če-tovr ādam-i ast?* ‘what kind of person is he?’; Pers. *in jur miz* ‘this kind of table?’; *in če raqam ketāb astlin čī raqam kitob ast* ‘what type of book is this?’.

5 SYNTAX AND USE OF FORMS II: CLAUSE STRUCTURE

5.1 Order of constituents

The canonical sequence in main, coordinate and dependent clauses is:

(SUBJECT) – (OBJECT) – VERB.

The parentheses indicate that both subject and object may be lexically omitted since the person of the subject is incorporated in the verb, and the object of transitive verbs may be contextually implied.

Thus the simplest clause may comprise a single finite verb:

<i>āmad-andl</i>	<i>dozdid-andl</i>
<i>omad-and</i>	<i>duzdid-and</i>
‘they’ve come, are here’;	‘they stole it’.

With three-valence verbs such as “give”, the definite or specific direct object generally precedes the beneficiary (indirect object), but follows it in preverbal position if indefinite. Note that in the preverbal position of the unmarked object *ketāb/kitob* exemplifies the generic function of unspecified nominals in preverbal position as part of compound verbs:

<i>(Ahmad) [ketāb-rā] be man dād</i>	<i>(Ahmad) be man [ketāb] dād</i>
<i>Ahmad [kitob-ro] ba man dod</i>	<i>Ahmad ba man [kitob] dod</i>
‘Ahmad gave me the book’;	‘Ahmad gave me a book, books.’

Word order is flexible and allows for “scrambling” (particularly in spoken registers). Typically, such re-ordering requires non-canonical marking patterns involving shifts of stress (and pauses). The more frequent and less marked of such patterns are discussed in the following.

5.1.1 *OSV and OVS*

5.1.1.1 Object raising

For emphasis on agent OSV may occur:

sag-rā man košt-aml
sag-ro man kušt-am
 dog I killed (it)
 ‘I was the one who killed the dog’.

5.1.1.2 Concomitant subject lowering

For emphasis on the patient, or the activity, OVS may occur with concomitant subject lowering:

sib-rā ne-mi-xar-id šomā?!
seb-ro na-me-xar-ed šumo?
 the apple-DO not-IPFV.buy.PR-2p you.p
 ‘aren’t you going to buy any apples?’

Note the contrastive use of the direct object marker *-rā/-ro* even with indefinite non-specific objects, approximately ‘as to the apples there (in addition to yoghurt, etc.)’.

5.1.2 *Adverbials*

Adverbs of time and sentence adverbials are often placed first in the sentence. The unmarked order of other is:

TIME – MANNER – LOCATION/DIRECTION,

and is determined by the valency of the verb, such as the means and direction of “going” in the following example:

ma’mul-an, mā har ruz bā otobus be madrase mi-rav-im
odat-an, mo har ruz bo avtobus ba madrasa me-rav-em
 usually we every day with bus to school we go
 ‘we usually go to school every day by bus’.

Otherwise the position of adverbials is relatively free, such as the focused locative preceding the preverbal beneficiary of “giving” instead of following the initial time adverbial:

diruz Ahmad ketāb-rā [dar ketāb-xāne] be man dād
dina-rūz Ahmad kitob-ro [dar kitob-xona] ba man dod
 yesterday Ahmad the book in library to me he gave
 ‘Ahmad gave me the book in the library yesterday’.

Particularly in colloquial speech direction, but not location, follows the verb, usually without the preposition *belba* ‘to’:

raft madrasel
raft madrasa
 ‘she went to the school’.

Adverbials pragmatically connecting and/or qualifying the sentence as a whole (often Arabic loan adverbials and phrases) are typically in first position:

ettefāq-an man ham did-am *zāher-an fardā mi-ā-yad*
ittifoq-an man ham did-am *az aft-āš pagoh me-o-yad*
 ‘incidentally, I saw it, too’; ‘from the looks of things, he’ll be coming tomorrow’.

5.1.3 Questions

Word order in questions is normally the same as for statements.

(1) Yes-no questions, *āyā/oyo*

In yes-no questions interrogativity is established suprasegmentally by means of intonation and pitch contour, which ends at mid-height rather than falling as in statements.

In addition to the primary contour, Common Persian is one of the languages that has developed an overt question marker, the particle: *āyā/oyo* (initial stress) ‘whether, isn’t it that’, which may be followed by the clipped tag question *yā na/yo na* ‘or not’:

āyā to hamrāh mi-ā-i yā na?!/
oyo tu hamroh me-o-yi yo na?
 Q you.s along IPFV.come.PR-2s or not
 ‘are you coming along, or not?’.

In colloquial Tajik, the Uzbek interrogative clitic *-mī* is added to the final word:

seb na-me-xar-ed-mī?
 apple not-IPFV-buy.PR-2p-Q
 ‘won’t you buy some apples?’

(2) Surprise and indignation, *magar/magar*

In questions connoting surprise or indignation (including rhetorical questions) the particle *magar* in the sense of ‘but is it not true that, but’ may precede or follow the subject:

magar u ān-hā-rā na-did-e ast!
ū magar on-ho-ro na-did-a-ast?
 but he those-DO not-see.PART COP.3s
 ‘(but) hasn’t he seen them?’

magar may also imply possibility in questions or statements (for its use as a conjunction ‘unless’ see section 6.5.3.7):

magar šomā ham mi-dān-id!
magar šumo ham me-don-ed?
 whether you.p also IPFV.know.PR-2p
 ‘maybe you know, too?’

(3) Interrogative pronouns and adverbs

Persian has *in situ* constructions. Interrogatives for parts of speech appear in the same slot as in statements when unmarked, but may appear at the beginning when focused (before or after an adverbial phrase):

Subject	Direct object
<i>[ki] emruz mi-ā-yad/</i>	<i>šomā ān-jā [če kas-i]-rā did-id?/</i>
<i>[kī] imrūz me-o-yad</i>	<i>šumo on-jo [čī kas]-ro did-ed?</i>
who today IPFV.come.PR-3s	you.p there what person-DO see.PT-2p
‘who will come today?’;	‘what person, whom did you see there?’
>> <i>emruz [ki] mi-ā-yad?</i>	>> <i>[če kas-i]-rā ān-jā did-id?</i>

Adverb

<i>ketāb-rā [kojā] gozāšt-e i?/</i>	<i>bahār emsāl [key] mi-ā-yad/</i>
<i>kitob-ro [kujo] mond-a i?</i>	<i>lbahor insol [kay] me-o-yad?</i>
book-DO where put.PART COP.2s	spring this year when IP ^F V.come.PR-3s
>> <i>kojā ketāb-rā gozāšte i?</i>	>> <i>key emsāl bahār mi-ā-yad?</i>
‘where did you leave, put the book?’;	‘when will spring come?’

5.1.4 Exclamations

(1) Vocative

The vocative has primary stress on the first syllable, and may be preceded by the particle *eyley* or *āy*:

<i>bačče, injā bi-āl</i>
<i>bača, in jo bi-o</i>
child here come-IMP.2s
‘boy, come here!’.

(2) Selective reference

Surprise and admiration are marked by *če N-i* ‘what (a) N’ and *če ADJ* ‘how ADJ’ (similar to questions):

<i>če havā-ye xub-i!</i>	<i>če xub!</i>
what weather-EZ good-INDEF	what good
‘what a fine weather!’;	‘how good!’.

(3) Tags

These include (a) *-(h)ā(n)* ‘take note, careful, beware’ and (b) *dige*, literary *digar* ‘already, simply’, lit. ‘other’:

<i>goft-am-ā!</i>	<i>goft-am-dige!</i>
‘I said so, didn’t I!’;	‘I’ve already told you!’.

5.2 Definiteness and subject and object marking

5.2.1 Marking of definiteness and specificity

The marking of a NP as definite, indefinite or specific is interdependent with case relations, which in turn are expressed chiefly by syntactic means.

5.2.1.1 Definiteness

The stem form of a noun has two uses, apart from paratactic mention in lists, titles, etc. These are:

(1) Generic, when focus is on the concept it conveys or the generic referent, rather than a specific member of the class or number of items:

<i>dar maqāze [ruznāme] hast?!</i>	<i>ketāb be-xān!!</i>
<i>dar magazīn [gazeta] hast?</i>	<i>kitob xon!</i>
in store newspaper EX.3s	book read.IMP.2s
‘are there newspapers in the store?’; ‘read a book!’, i.e. any book, or books in general;	

<i>pedar-hā-yemān [dehqān] and!</i>
<i>padar-ho-yamon [dehgon]-and</i>
father-p-our peasant COP.3p
‘our fathers are peasants’, lit. ‘peasant’.

(2) Definite, when a referent has been mentioned, or is contextually defined, for instance as a proper noun, independent or clitic pronoun, demonstrative, or equivalent interrogatives. A definite count noun denoting a plural requires a plural marker and, usually, plural verb agreement:

Singular

<i>deh az injā xeyli rāh nist!</i>	<i>pedar-bozorg āmad!</i>
<i>qišloq az injo xele roh nest</i>	<i>bobo omad</i>
village from here much way not.COP.3s	grandfather come.PT.3s
‘the village isn’t far from here’;	‘grandfather has come, arrived’.

Plural

<i>deh[-hā] xeyli rāh nist[-and]!</i>
<i>qišloq[-ho] xele roh nest[-and]</i>
village-p much way not.COP.3s
‘the villages are not far’.

5.2.1.2 Specificity

The specific NP occupies a status between the indefinite NP, the referent of which is known to neither speaker nor listener, and the definite NP, the referent of which is known to both. It constitutes information known to the speaker, but introduced for the first time to the listener; with this form, the speaker signals his ability or willingness to give further information.

A specific NP is marked with *-il-e* and/or *yek/yak*, and as a direct object it requires the marker *-rāl-ro*, while an indefinite noun normally does not.

5.2.2 Marking of subjects

5.2.2.1 Verbal agreement

(1) 3rd person. The Classical Persian norm for 3rd person verbal concord (inherited from Indo-European) is that inanimate subjects have singular agreement and anaphora. In

turn, humans and sentient beings (including higher animals, notably horses) have plural agreement: this is still a valid pattern, but the use of singular and plural is determined by the speaker's perception of the degree of individuation of the subject:

<i>dust-hā āmad-andl</i>	<i>ketāb-hā injā-st ~ (hast-)andl</i>
<i>dūst-ho omad-and</i>	<i>kitob-ho injo-st ~ (hast-)and</i>
friend-p come.PT-3p	book-p here is (EX)-3p
'the friends have arrived-3p';	'the books are here'.

Similarly, collective nouns and pronouns, as well as plural inanimates, may have a singular or plural verb, more commonly in Tajik:

human collective-plural	inanimate plural-singular
<i>mardom dar āmad-andl</i>	<i>šāx-hā-ye deraxt-hā mi-jombidl</i>
<i>mardum daromad-and</i>	<i>šox-ho-i daraxt-on me-jumbid</i>
people out.come.PT.3p	bough-p-EZ tree-p IPFV.shake.PT.3s
'the people came out',	'the boughs of the trees shook',
individually, in small groups;	all together.

(2) Combined persons. Subjects combining 1st person with either or both the other two have 1st plural agreement, subjects combining 2nd and 3rd person have 2p plural agreement:

<i>man o šomā mi-dān-impl</i>	<i>man o Ahmad bā ham raft-im</i>
<i>man u šumo me-don-em</i>	<i>man u Ahmad bo ham raft-em</i>
I and you.p IPFV.know-1s	I and A. together go.PT-1s
'I and you we know';	'I and Ahmad went together'.

(3) Honorific marking

Social convention requires respectful reference to 3rd persons with plural agreement, particularly in their presence. Similarly, the addressee is called by the plural pronoun *šomā/šumo* 'you' with the corresponding verb ending, while the speaker may use the self-deprecating plural *mālmo* 'we' (see section 3.1.3.1 Personal pronouns):

<i>ostād šaxs-e bozorg-i bud-andl</i>	<i>šomā bud-idl mā če mi-dān-impl</i>
<i>ustod šaxs-i buzurg-e bud-and</i>	<i>šumo bud-ed mo ċī me-don-em</i>
professor person-EZ great be.PT-3p	you.p be.PT-1p we what IPFV.know.PR-1p
'the professor was a great man';	'it was you'; 'what do I know?'

5.2.2.2 Copular predicates and ellipsis

As noted under Morphology, the copula form of 'be' in the present tense may be either enclitic or derived from the stems *hast-lhast-*, *bāš-lboš-*, or *nist-lnest-*, and in past tenses from the stem *bud-lbud-*.

In Persian colloquial the present tense forms are obligatory, except with certain question words and phrases: *ku?*, *kojā?* 'where (is): Hasan *ku*, *kojā?* 'where is Hasan?'. Exclamatory rhetorical phrases include deictic forms like *inak*: *inak—parčam!* behold/here (is) the banner!.

In colloquial Tajik, present tense forms may be more readily omitted, especially in questions: *Nodir kujo?* 'where's Nadir?', including basic identification statements: *man tojik* 'I'm a Tajik'.

5.2.2.3 Specified and unspecified subjects

Since the person is expressed in the verb ending, the minimal clause consists of an inflected verb, and there may be no overt subject noun phrase:

sib mi-xar-amlseb me-xar-am
apple IPV.buy.PR-1s
'I'll buy (~ some) apples'.

Unspecified human subjects are expressed by the 3p verb ending (see section 5.5.7.2 Agency and causation), while impersonal inanimate subjects are expressed by the 3s verb ending only: e.g. *bāš-adlboš-ad* 'let it be, OK', *mi-šav-adlme-šav-ad* 'it's possible, one can (do it)', and prominently in indirect verb constructions (section 5.3.2).

'Weather-verbs' such as *bār-*, *bārid-lbor-*, *borid-* 'pour from above' and similar verbs require a subject noun: *bārān*, *barf mi-bār-adlboron*, *barf me-bor-ad* 'it rains, snows' ('it pours down rain, snow').

5.2.3 Marking of direct objects

5.2.3.1 Definite direct objects

Definite objects are obligatorily marked with *-rāl-ro*:

ketāb-rā xarid-aml u-rā did-aml
kitob-ro xarid-am vay-ro did-am
book-DO buy.PT-1s he-DO see.PT-1s
'I bought the book'; 'I saw him'.

5.2.3.2 Indefinite specific and non-specific direct objects

An indefinite noun as direct object is not normally marked with *-rāl-ro*. However, while a *specific* NP is also marked with *-il-e* and/or *yeklyak*, as a direct object it requires the object marker:

(yek) zan[-i]-rā did-aml
(yak) zan[-e]-ro did-am
one woman-INDEF-DO see.PT-1s
'I saw a certain woman';

(yek) ketāb-e xub[-i]-rā xānd-aml
(yak) kitob-i xub[-e]-ro xond-am
one book-EZ good-INDEF-DO read.PT-1s
'I've read a good book'.

Such sentences as a rule require, or imply, following explanatory clauses, typically relative clauses, such as, for the latter example, 'a good book – which I might now identify or describe for you, if you show interest'.

5.2.3.3 Scale of specificity

Specificity is given overtly when a noun or noun phrase is, or contains, a name, personal or demonstrative or interrogative pronoun, and implicitly when a specific or certain item

or set is referred to by the speaker. Accordingly, *-rāl-ro* also attaches to items or sets that are marked by the indefinite *-il-e*, when specific, such as *kas-i-rā did/kas-e-ro did* ‘he saw someone, a certain person’; in comparison:

<i>gorg-rā košt-and</i>	‘they killed the wolf’;
(<i>yek</i>) <i>gorg-i-rā košt-and</i>	‘they killed a certain wolf’;
(<i>yek</i>) <i>gorg-i košt-and</i>	‘they killed some wolf’;
<i>gorg košt-and</i>	‘they killed wolves, were wolf-killing’;
<i>gorg-hā košt-and</i>	‘they killed many wolves’, <i>-hā</i> with emphatic stress.

5.3 Use of pronominal clitics

The enclitic series of the pronouns is used primarily to express attributive possession and its metaphors (see section 3.1.3.2):

ketāb-am/kitob-am ‘my book’, *varaḡ-hā-yaš/varaḡ-ho-yaš* ‘its pages’.

This use includes the Turco-form possessive construction in the northern Tajik dialects discussed above:

in mard[-a] kitob[-aš]
 this man-RA book-his’
 ‘this man’s book’.

In addition, the clitics function both as direct objects and indirect objects, including the experiencer. In Tajik they also function in a possessive construction, including nominalized clauses (see section 6.3.7).

5.3.1 Pronominal direct and indirect objects

The pronominal clitics may also express the direct or indirect object of a verb to which they attach.

(1) direct object, *did-am-etān/did-am-aton* ‘I saw you (*-etān/-aton*)’.

The pronominal clitic may replace the direct object phrase, and is attached to the verb, and in compound verbs usually to the preverbal nominal or to the verb itself (mostly colloquial, see section 5.5.6.2):

[u-rā] did-am/[vay-ro] did-am
did-am[-aš]/ did-am[-aš]
 ‘I saw him’;

Persian	Tajik
<i>[u-rā] bidār kard-am/</i>	<i>[vay-ro] bidor kard-am</i>
<i>bidār[-aš] kard-am/</i>	<i>bidor[-aš] kard-am</i>
<i>bidār kard-am[-aš]/</i>	<i>bidor kard-am [-aš]</i>
‘I woke him up’, <i>bidār kardan/bedor kardan</i> , lit. ‘make awake’.	

In spoken Tajik these forms often elide the initial vowel and add the object marker *-a* (< *ro*): *me-gir-an[-š]* ‘they’ll catch him (*-š*)’; *did-am-t[-a]* ‘I saw you (*-t*)’.

(2) indirect object, *goft-am-aš/guft-am-aš* ‘I told him (*-aš*)’.

These replacements are found with a small number of simple and compound verbs with beneficiary valence:

Persian	Tajik	Persian	Tajik
[<i>be u</i>] <i>goft-aml</i>	[<i>ba vay</i>] <i>guft-am</i>	[<i>be u</i>] <i>komak kard-aml</i>	[<i>ba vay</i>] <i>yorī dod-am</i>
<i>goft-am[-aš]</i>	<i>guft-am[-aš]</i>	<i>komak[-aš] kard-aml</i>	<i>yorī[-aš] dod-am</i>
‘I said to, told him’;		‘I gave him help, helped him’.	

5.3.2 Indirect experience

Indirect verb constructions express bodily or mental experience as happening to, or overcoming, a person without his or her doing. They are most frequent in casual styles, and appear to be an expanding class.

The verb is a compound of a noun or adjective plus ‘be’ or ‘become’ and a small set of other verbs in the impersonal 3s, while the experiencer is referred to by the personal clitics, at earlier stages of Persian also by the independent personal pronoun plus *-rā* in its dative function.

The source is usually indicated by *az* ‘from’. Some sensations and events can only be expressed that way, while others can be expressed either by an indirect construction or by verbs which imply the active involvement of the speaker, allowing for options. Among the commonly used indirect verb constructions, and their active correspondents, there are prominently those expressing bodily sensations, emotions, and mental activity. The following examples illustrate the pairing of active expressions with indirect construction, and the typical pairs of verbs which are similar to those used in the quasi-active/passive pairs of other types of compound verbs:

Active	Experiential
<i>man gorosne [am]</i>	<i>gorosne [-am] ast</i>
I hungry COP.1s	hungry-to me COP.3s
‘I am hungry’;	‘I feel hungry’;
<i>man bāvar ne-mi-kon[-am]</i>	<i>bāvar[-am] ne-mi-šav-ad</i>
I belief not-IPFV.make.PR-1s	belief-to me not-IPFV.become.PR-3s
‘I don’t believe it’;	‘I cannot believe it’;
<i>qahr kard[-im]</i>	<i>qahr[-emān] šod</i>
quarrel make.PT-1p	quarrel-to usbecome.PT.3s
‘we quarrelled’, lit. ‘made quarrel’;	‘we got to quarrelling’;
[<i>ān-rā</i>] <i>dust dār-im</i>	[<i>az ān</i>] <i>xoš[-am] mi-ā-yad</i>
that-DO friend have.PR-1p	from that well to me IPFV.come.PR-3s
‘we like him, it’, lit. ‘have as friend’;	‘I like it’, lit. ‘from it it comes well’;
	cf. earlier [<i>ma-rā</i>] <i>xoš āmad</i> .
Compound with <i>yād</i> ‘memory’, all ‘I remember’	
[<i>ān-rā</i>] <i>yād dār[-am]</i>	[<i>ān</i>] <i>yād[-am] hast</i>
that-DO memory have.PR-1s	that memory-to me EX.3s
‘I have it in memory’;	‘that is to me in memory’;
[<i>ān-rā</i>] <i>yād mi-ā-var[-am]</i>	[<i>ān</i>] <i>yād[-am] mi-ā-y-ad</i>
that-DO memory IPFV.bring.PR-1s	that memory-to me IPFV.come.PR.3s
‘I bring to memory’;	‘it comes to my memory’;

Typically colloquial is the expression *čē-t-e?* ‘what is the matter with you?’ – *čiz-i-m nist!* ‘nothing!’ (< *čē-at ast* ‘what to you is?’ and *čiz-i-am nist* ‘a thing to me is not’).

5.4 Subject and object complements

Nominal complements implying a property are non-specific (generic) and are unmarked, whether intransitive or transitive. The verb is a verb of state and persistence or change, most typically ‘be’, including copula, the existential *hast-lhast-*, neg. *nist-lnest-*, the subjunctive *bāš-lboš*, past all *būd-lbud-*. Other verbs include: intransitive *mān-*, *māndlmon-*, *mond-* ‘remain’, *šav-*, *šod-lšav-*, *šud-* ‘become’; transitive *kon-lkun-*, *kard-*, *dān-*, *dānest-l don-*, *donist-* ‘know, recognize’, *entexāb k.lintixob k.* ‘elect’, etc.

<i>qabl-an čupān bud-andl</i>	<i>u-rā dust mi-dān-aml</i>
<i>qabl-an podabon bud-and</i>	<i>ū-ro dūst me-don-am</i>
before shepherd be.PT-1p	he-DO friend IPFV.know.PR-1s
‘before they were shepherds’;	‘I consider him a friend’.

The same rules applies to adjectival complements.

However, when modified by a specifying nominal EZ plural marking is obligatory:

<i>qabl-an čupān-hā-ye mā bud-andl</i>	<i>ān-hā-rā dust-hā-ye mā mi-dān-ml</i>
<i>qabl-an podabon-ho-i mo bud-and</i>	<i>on-ho-ro dūst-ho-i mo me-don-im</i>
before shepherd-p-EZ we be.PT-3p	they-DO friend-p-EZ we IPFV.know.PR-1p
‘before they were our shepherds’;	‘we consider them our friends’.

5.5 Complex predicates

The following general characterizations can be made:

The combined morphology and categorical functions of this system suggest the following parameters:

Aspect. The primary parameter of the system is aspect, intersecting with the parameters of tense and mood. It is defined by the polarity between imperfective and perfective, framing the central perfective-aorist.

The *perfective preterite (aorist)* is the unmarked form and refers to an event simply as completed, and is central to the system. It refers most frequently to past events, and is traditionally equated with a Western preterite, or simple past. However, it may also refer to present or future contexts, and, in addition, may have subjunctive function, notably in subordinate clauses.

The imperfective is undifferentiated and refers to an event simply as not completed. Thus, the indicative present and past may refer to punctual, habitual-iterative or durative-progressive events, in present and past.

The resultative-stative refers to prior action, result, and temporary state in present and past contexts.

Mood. Basically, real action is marked by the indicative, potential action or state by the subjunctive, and unlikely and unreal action by the counterfactual. Both subjunctive and the counterfactual have binary distinctions, imperfective and perfective/resultative-stative. In form, the imperfective counterfactual is identical with the imperfect, while the subjunctive is morphologically aligned with the present. Both occur in independent clauses, as well as in subordinate clauses.

5.5.1 Possession

Possession is expressed by the verb *dār-*, *dāšt-/dor-*, *došt-* ‘have, keep’.

In Tajik, the pronominal clitics in their dative function together with ‘be’ may also express (alienable) possession of a predicate (see Pronominal direct and indirect objects section 5.3.1):

safar pul[-aš] boš-ad agar, me-ti-yam-t[-a]
 S. money-his may be if, IPFV.give.PR-1s-you.s-RA
 ‘if Safar has any money, I’ll give it you.’

5.5.2 Modal constructions, verb + verb

5.5.2.1 Definite future

xāh-am raft/xoh-am raft ‘I shall, will (definitely) go’.

This modal construction with *xāh-* ‘will’ is isolated and expresses definite future, contrasting with the unmarked imperfective present. It is formed from the unmarked present stem of the verb *xāh-/xoh-* ‘will, want, wish’ marked for person plus the short infinitive: *(na-)xāh-am raft/(na-)xoh-am raft* ‘I will (not) go’; compound verbs: *kār (na-)xāh-and kard/kor (na-)xoh-and kard* ‘they will (not) work’.

In Tajik, more so than in Persian, this construction is regarded as “literary” and generally avoided in favor of the simple present-future. See also section 8.3.2 Uncertainty constructions.

5.5.2.2 Volition

Volition is expressed by fully inflected *xāh-*, *xāst-/xoh-*, *xost-* ‘want, wish’ followed by the subjunctive.

<i>bačče če</i>	<i>mi-xāh-adl</i>	<i>mi-xāst-am</i>	<i>xāne be-rav-aml</i>
<i>bačča či</i>	<i>me-xoh-ad</i>	<i>me-xost-am</i>	<i>xona rav-am</i>
child what	IPFV-want.PR-3s	IPFV-want.PT-1s	house SBJ-go.PR-1s
‘what does this/that child want?’		‘I wanted to go home’, lit. ‘that I go’.	

5.5.2.3 Intention and expectation

(1) [*raft-an-ī*] (*hast-)*am ‘I am about to, supposed to’

Verbs of motion, and particularly *raft-an/raft-an* ‘go, leave’, are formulated in a construction formed with the future participle, or gerundive (see section 3.2.2.2e Nominal forms) and forms of ‘be’ to express intention.

In Persian this construction is marginal. Used predicatively, it is found with a semantically highly restricted set of verbs and contexts: intransitive, rarely used in tenses other than the present, and rarely negated: *man raft-an-i hast-am* ‘I have to, am about to go’.

In Tajik there are no such restrictions. It is used with both intransitive and transitive verbs, occurs with *boš-/bud-* ‘be’ and *šav-/šud-* ‘become’ (and colloquially, without a verb), may be negated, allows for any inflected form of the verb; examples of negated, evidential, and progressive forms are:

Intransitive:

man raft-an-ī-yam ~ *hast-am raft-an-ī šud-em*
 I go.GERV-COP-1s ~ EX-1s go.GERV become.PT-1p
 'I'm going, I'm off'; 'we were all set to go'.

Transitive:

heč hujum [kard-an-ī] nest-and [*guft-an-ī*] *na-[bud-a]-ast*
 any attack make.GERV not-COP-1p say.GERV not.be.PART.EV-COP.3s
 'they are not going to attack'; 'it appears she's not about to tell';

ma-ro ba dušman [dod-an-ī šud-a] istod-a and
 I-DO to enemy give.GERV become.PART PROG.PART COP-3p
 'they're getting ready to give me to the enemy'.

This productive construction is thus distinct from the use of the predicative use of the gerundive in both Persian and Tajik with forms derived from the transitive verb which may be understood as a passive: Persian *xeyli did-an-i ast* 'it is very much to be seen, worth seeing', Tajik *on jo heč did-an-ī nest* 'that place isn't worth seeing'.

(2) *rav-ona* 'about to, supposed to go'

In Tajik, the most commonly found form in the adjectival derivative is *-ona*:

man estalinobod ravona bud-am ke na-šud
 I St. go.ADJ be.PT-1s that not-become.PT.3s
 'I was about to/supposed to go to Stalinabad, but it didn't work out'.

In Persian, there is *ravāne šod-an; kard-an* 'set out, start; send off', but no corresponding construction.

5.5.2.4. Obligation and presumption

5.2.2.4a Obligation and necessity

Obligation and necessity are expressed by means of the impersonal defective verb *bāyad/boyad*, neg. *na-bāyad/na-boyad* (less frequent variants with milder force are (*mi-*)*bāyest-(i)/me-boist* with *mi-/me-* in its counterfactual function and the frozen archaic counterfactual enclitic *-i*, respectively).

(1) Impersonal generalized statements

bāyad raft/boyad raft 'one (we, etc.) must go'

This construction is formed with the short infinitive (= Stem II).

(2) Personal expressions

bāyad be-rav-ad/boyad rav-ad 'he must go'

This construction is formed with the present subjunctive:

bāyad fardā parvāz be-kon-aml
boyad fardo parvoz kun-am
 must tomorrow flight SBJ.make.PR-1s
 'I have to fly tomorrow';

mi-bāyest in deraxt-rā az bix be-kan-impl
me-boist in daraxt-ro az bex kand-a bar-or-em
 would have to this tree-DO from root pull.PART up-SBJ.bring.PR-1p
 ‘we ought to uproot this tree’.

Frequent other expressions are *lāzem ast*, *budllozim ast*, *bud* ‘it is necessary’ and *majbur hast*, *budlmajbur ast*, *bud* ‘he is obliged, forced to . . .’, etc., similarly followed by the subjunctive.

(3) Counterfactual forms

These are marked by the imperfect in its counterfactual function in the dependent verb, given that *bāyad* does not have a functional past form (similar to English “must”).

(a) imperfective form

bāyad mi-raft/boyad me-raft ‘he had to, should have gone’

bāyad zud-tar mi-raft-il
boyad zud-tar me-raft-ī
 must early-COMP IPFV.go.PT-2s
 must earlier would leave
 ‘you had to leave, should have left earlier’.

This construction also refers to an action that had to be done instead of another:

tāze qabl-aš ham [bāyad mi-raft-am], barādar-am rā madrase be-gozār-am
toza qabl-aš ham [boyad me-raft-am], uka-am-ro maktab mon-am
 ‘before that I had to go drop my brother off at school (and did not come here right way)’.

(b) resultative form (rare)

bāyad raft-e bud ‘he had to, should have gone (earlier, before then)’.

(4) Tajik infinitival construction

In Tajik, obligation in any tense is expressible by a circumlocution such as *darkor* (‘in the act, appropriate’) preceded by a nominalized clause construction:

xobond-an-i vay-ro darkor bud
 laying-EZ he-DO necessary be.PT.3s
 ‘we had to lay him down’.

5.5.2.4b Presumption

bāyad raft-e bāš-ad/boyad raft-a boš-ad

Presumption about a state or the completion of an action is also expressed by *boyad* and is followed by the perfect subjunctive or the present subjunctive of ‘be’. With change of state verbs the completion usually corresponds to a present in English:

bāyad dar xāne bāš-ad/ bāyad tā hālā rasid-e bāš-and/
boyad dar xona boš-ad boyad to holo rasid-a boš-and
 must in house (SBJ) be-3s must till now arrive.PART (SBJ.)be-3p
 ‘she must be at home’; ‘they must have arrived by now’-;

bāyad ānjā istād-e bāš-and/
boyād on jo istod-a boš-and
 must there stand.PART (SBJ.)be-3p
 ‘they must have stood up = be standing there now’.

5.5.2.5 Ability

Ability is expressed by the fully inflected verb *tavān-*, *tavānest-/tavon-*, *tavonist-* ‘be able’. Similar to *bāyad/bojad* there are two constructions:

(1) Generalized impersonal statements

mi-tavān raftme-tavon guft

This construction uses the bare present tense form plus the short infinitive, highly marked as literary:

mi-tavān goft
me-tavon guft
 IPFV-can say.INF2
 ‘one can (not) say’.

(2) Personal expressions

mi-tavān-ad be-rav-adlme-tavonist rav-ad ‘he can go’ (present form)

This construction has fully inflected modal and dependent verb forms in the subjunctive with person agreement. The cover term “ability” includes permission as well:

<i>ne-mi-tavānest-am</i>	<i>be xāb</i>	<i>be-rav-aml</i>	<i>mi-tavān-am</i>	<i>be-rav-am?!</i>
<i>na-me-tavonist-am</i>	<i>xob</i>	<i>rav-am</i>	<i>me-tavon-am</i>	<i>rav-am?</i>
not-IPFV-can.PT-1s	to sleep	SBJ.go.PR-1s	IPFV-can.PR-1s	SBJ-go.PR-1s
‘I could not go to sleep’;			‘can I go, leave?’.	

(3) Tajik participial constructions

[raft-a] me-tavon-ad (present form)

Though the Persian subjunctival construction is available in Tajik, the preferred idiom is a participial construction. Specifically, instead of the dependent verb following the modal in the subjunctive, this is a conjunct construction where the past, or perfect participle precedes the modal verb: *[raft-a] me-tavon-am* ‘I can go’ (here with the modal in the evidential mode in the first example, and in the progressive in the second):

vay [tojikī gap zad-a] na-tavonist-a-ast
 he Tajik talk hit.PART not can.PT-EV-COP.3s
 ‘he couldn’t speak Tajik, as far as I know’.

5.5.2.6 Possibility

The fully inflected verb *šav-*, *šod-* ‘become’ may express possibility as well as permission. It also provides the pendant to impersonal (*ne-*)*mi-tavān* ‘one can’. Whereas the latter implies the ability proper to living beings, reflecting the original meaning ‘be strong’ of *tavān-/tavon-*, *šodan/šudan* implies not so much physical as logistical or moral feasibility reflecting its original meaning ‘go, proceed’ (cf. German *es geht* ‘is possible, can be done’, lit. ‘it goes’).

(1) Generalized impersonal statements

mi-šav-ad raft | *me-šav-ad raft* ‘one can go’

In this construction the modal is followed by the short infinitive:

[*ne-mi-šav-ad*] *maze-ye ān-hā-rā* [*farāmuš kard*]
not-IPFV-become. PR-3s taste-EZ those-DO forgotten make.INF2
‘one cannot forget the taste of those’;

čerā na-raft-i?– *na-šod*
why not-go.PT-2s not-become.PT.3s

‘why didn’t you go? – it didn’t work out, was not possible’.

Tajik

man ba maskaw raft-an-ī bud-am, lekin na-šud

I to Moscow go.GERV be.PT-1s but not-become.PT.3s

‘I intended to, was supposed to, was ready to go to Moscow, but it didn’t come off’.

(2) Personal construction

mi-šav-ad be-rav-ad | *me-šav-ad rav-am* (present form)

Persian also has a personal construction, similar to *tavānestan*, where the 3s impersonal modal is followed by the inflected dependent verb in the subjunctive:

ne-mi-šav-ad *be-rav-am*
not IPFV-become.PR-3s SBJ-go.PR-1s
‘it is not possible for me to go ~ can’t I go?’, colloquial *ne-mi-š-e be-r-am?*

5.5.3 Aktionsart constructions, *V + V*

Aktionsart constructions are mostly confined to colloquial speech in Persian (see Windfuhr 1979: 102–105), but are less restricted in Tajik. They are expressed by a small number of verbs with dependent finite verb, infinitive, or participle. Of these, the infinitive construction is inherited from earlier stages of Persian, which is retained in Tajik, but in Persian tends to be assimilated to post-auxiliary position. Typically, most Persian constructions are double finite, while Tajik constructions are participial (see section 5.5.3), as is the case with the progressive construction discussed in section 3.2.6a/b.2:

Persian		Tajik		
(<i>dār-am</i>)	<i>mi-rav-am</i>	<i>rafta</i>	<i>istoda</i>	<i>am</i>
PROG.PR-1s	IPFV-go.PR-1s	go.PART	PROG.PART	COP.1s

both: ‘I am about to leave, leaving’.

5.5.3.1 Inception, non-intentional

In Persian, *raftan* ‘go, leave’ followed by the subjunctive expresses the notion of ‘be going to’ with the connotation of possibility or likelihood, usually in combination of *dāštan*:

mi-rav-am yavās yavāš be in natije be-res-am ke . . .
IPFV-go.PR-1s slow slow at this result SBJ-arrive.PR-1s that
‘I am slowly beginning to understand that . . .’;

mi-rav-ad farāmuš be-šav-ad
 IPFV-go-3s forgotten SBJ-become.PR-3s
 '(this old tradition) is now in the process of being forgotten'.

5.5.3.2 *Inception*

Ali šoru' kard [be neveštan]/
Ali [ba nivištan] šurū' kard
 A. begin made [to writing-INF]
 'Ali began to write'.

5.5.3.3 *Inception, completive*

Persian uses the infinitive in the literary, but TMA marking in the colloquial register; Tajik uses only the former:

Persian

(1) infinitive

asb david-an gereft
 horse run-INF take.PT.3s
 'the horse began to gallop'.

(2a) Past

gereft-and xābid-and
 take.PT-3p sleep.PT-3p

(2b) Imperative

hama-š na-gir be-xāb
 all-its not take.IMP-2s SBJ-sleep.IMP-2s
 'they took to sleeping, fell asleep'; 'don't fall asleep all the time'.

Tajik

korkar-i xud-aton-ro kard-an gir-ed!
 work-EZ own-your-p-DO do-INF take.IMP-2p
 'you get on with/begin your work!'.

5.5.3.4 *Inception, ingressive*

The incipient function of the verb *xāstan* 'want' derives from its modal connotations. This innovative construction is to be distinguished from the long-established, but isolate definite future construction with *xāh-* + past stem: *xāh-am raft/xoh-am raft* 'I will go' (cf. section 8.3.2 Uncertainty constructions). While in colloquial Persian this construction is restricted to ingressive function, in Tajik it is not.

Persian

bičāre mi-xād be-mir-e
 poor IPFV-want.PR-3s SBJ-die.PR-3s
 'the poor fellow is about to die'.

Tajik

Sobir čahār-čub girifta, xost ki bi-rav-ad
 S. frame take.PART want.PT.3s CL SBJ-go.PR-es
 'Sobir picked up the frame and made to go'.

Some of these verbs may be combined in a double Persian finite construction:

māšin *dār-e* *mi-r-e* *be-suz-e*
 car PROG.PR-3s IPFV.go.PR-3s SBJ-burn.PR-3s
 ‘the car is about to begin to burn’.

5.5.4 Tajik conjunct verb constructions

Also known as serial verbs, these are a salient feature of Tajik, especially in the literary language and the Northern dialects, where they seem to have originated as calques on Turkic usage. The category may still be evolving and expanding. In fact, the progressive *kard-a istod-a-*, and the ability construction *kard-a tavon-Itavonist-* discussed above, are grammaticalized instances of the type, and other uses will be noted.

In its most systematic manifestation, the construction joins an inflected form of one of a class of auxiliaries, which provides an Aktionsart or adverbial nuance, to the non-finite “past participle” of the semantically (and grammatically) main verb. There are about 16 such modals, the most commonly used being:

(1) *dih-/dod-* ‘give’ (other-benefactive):

nom-aš-ro *navišt-a* *me-dih-am* *ašula* *xond-a* *dod-and*
 name-his-DO write. PART IPFV-give.PR-1s song sing.PART give.PT-3p
 ‘I’ll write its name down for you’; ‘they performed songs’,
xon-/xond- ‘sing’.

(2) *gir-/girift-* ‘take’ (self-benefactive):

afsona-ro *navišt-a* *girift-am* *mavj-i* *kūtoh-ro* *došt-a* *me-girift-em*
 tale-DO write.PART take.PT-1a wave-EZ short-DO hold.PARTIPFV-take.PT-1p
 ‘I copied down the tale’; ‘we used to get the short wave’, *dor-/došt-* ‘hold’.

(3) *bin-/did-* ‘see’ (tentative):

non-ro *xūrd-a* *did*
 bread-DO eat. PART see.PT.3s
 ‘he tasted the bread’, *xūr-/xūrd-* ‘eat’.

(4) *šav-/šud-* ‘become’ (completion of an action):

kitob-ro *xond-a* *šud-am*
 book-DO read. PART become.PT-1s
 ‘I finished the book’.

(5) *firist-/firistod-* ‘send’ (sudden, violent action):

xand-a *kard-a* *firistod*
 laugh.PART do.PART send.PT-1s
 ‘he burst out laughing’.

(6) *mon-/mond-* ‘stay, remain’ (sudden action followed by continuing state; non-completion or forestalling of an action, depending on the sense of the main verb):

zan-aš kasal šud-a mond man tanho pušt-i on adam-ro did-a mond-am
 'his wife fell ill'; 'I only saw that man from the back'.

(7) *rav-raft-* 'go' (motion away from the speaker; finality; progressivity, depending on the sense of the main verb):

bača david-a raft daftar-am gum šud-a raft
 child run.PART go.PT.3s notebook-my lost become.PART go.PT-1s
 'the boy ran away'; 'my notebook got lost';

ohista-ohista sihat šud-a me-rav-ī
 little little health become.PART IPFV-go.PR-2s
 'little by little you'll get well'.

(8) *o-lomad-* 'come' is similarly used for motion toward the speaker.

5.5.5 Light verb constructions

Light verb constructions (LVC), also referred to by the ill-defined term "compound" verbs, consist of a nominal and a limited set of simple ("light") verbs. They represent a class of complex predicates that is distinct from modal, aktionsart, and conjunct verb constructions, even though lexically their sets of function verbs partially overlap. Most prominent are transitivity-voice pairs of LVCs with 'do' and 'become': *xošk kardan* :: *šodan* 'make :: become dry'. Essentially, then, LVCs are analytical types of de-nominal and de-adjectival verbs as opposed to derivation, such as *fahm-*, *fahm-id-* 'understand' (Ar. *fahm* 'understanding'), *xošk-*, *xošk-id-* 'dry (out), drain' (Pers. *xošk* 'dry'; see 3.2.1.2 above). For many centuries the emergence of LVCs, now estimated at over a thousand, has been the only source of verbal innovation, including basic verbal concepts, e.g. *dust dāštan/dūst doštan* 'love, like,' lit. 'have as friend', *kār kardan/kor kardan* 'work,' lit. 'do work', and in turn has led to the decreasing use of simple verbs, of which only some 150 have retained higher frequency.

These constructions consist of nominals that provide the main lexical meaning and about 18 light (simple) verbs that carry the full inflectional load (including auxiliaries). The nominals include adverbials, nouns, adjectives, or nominalized forms of verbs, such as verbal nouns and active and passive participles—many of which are integrated Arabic loans—both with or without adpositions. The LVCs therefore range from relatively little to considerable morphological, syntactic, and semantic complexity. Thus, Ahadi (2001: 53) found eleven syntactic subclasses of varying complexity, consisting of one- to four-place NPs, determined by the specifics of both the verb and the obligatory complements (even though some of the latter may be contextually omitted under definable conditions). Others distinguish more or fewer subclasses.

Syntactic criteria for testing the syntactic properties and subclassification include prominent place of verbal rection, permissibility of insertions, deletions, and permutations, and semantic shifts (for the insertion of the amplifier *-hā* see Windfuhr 1979: 33). While to a certain extent these hundreds of LVC can be classified, ultimately each has its own syntactic and semantic range. Moreover, they may vary from speaker to speaker. The following example shows the set of the insertion options related to *fekr* 'thought, thinking' plus *kardan* 'do, make'. This LVC has partially replaced the simple literary verb *andišidan* 'think, reflect; plan, devise; fear, be apprehensive' in terms of semantic range (cf. Meyer-Ingwersen 1974):

Nominalization, EZ

<i>fekr kardan-e u</i>		'his thoughts, thinking'.
Insertions permitted		
<i>(yek) fekr-i</i>	<i>mi-kon-am</i>	'I'll think of s.th.', lit. 'make a thought';
<i>fekr-hā-(i)</i>	<i>mi-kon-am</i>	same, action amplified by plural <i>-hā</i> ;
<i>in fekr-rā</i>	<i>mi-kon-am ke</i>	'I think this that . . .';
<i>fekr-e xub(-i)</i>	<i>mi-kon-ad</i>	'he has good ideas', lit. 'makes a good thought';
<i>fekr-e pul-eš-rā</i>	<i>mi-kon-am</i>	'I'm thinking of the costs of it', lit. 'its money-DO';
<i>be pul-aš fekr</i>	<i>mi-kon-am</i>	'I think of the costs (that might be involved', lit. 'at its money'.

Preverbal

<i>xub fekr</i>	<i>mi-kon-ad</i>	'he has very good ideas', lit. 'thinks well';
<i>hamin-rā fekr</i>	<i>mi-kon-am</i>	'I am thinking the same-DO';
<i>barā-ye to kār-i fekr</i>	<i>mi-kon-am</i>	'I am thinking of some things for you to do (a job', etc.).

The example of *fekr kardan* contrasts with the LVC *otu zadan* 'iron (clothes)', lit. 'hit with iron', where the insertion of the direct object marker *-rā* dissolves the LVC: *otu-rā zad* 'he hit the iron (in anger)', etc.

There are predictable correlative transitive vs. intransitive pairs using these and other auxiliaries (as had been noted from the earliest Western grammars of Persian onward), the semantically least specified pair being *kardan* vs. *šodan*:

<i>dorost kařdanl</i>	<i>dorost šodanl</i>
<i>durust kardan</i>	<i>durust šudan-</i>
'make correct, prepare s.th.';	'be corrected, prepared'.
<i>gerd āvordanl</i>	<i>gerd āmadanl</i>
<i>gird ovurdan</i>	<i>gird omadan</i>
'amass, convene' ('bring around');	'gather, congregate' ('come around').
<i>rāh andāxtanl</i>	<i>rāh oftādanl</i>
<i>roh andoxtan</i>	<i>roh uftodan</i>
'set in motion, start' ('throw on road')	'set out, start' ('fall on road').

From a semantic perspective, there are two kinds of light verbs:

(1) Those involving metaphorical use of a nominal, adverbial, etc., where the semantic load is spread fairly evenly between both elements and the auxiliary may vary idiomatically:

<i>sar kardan/sar kardan</i>	<i>az barlbar kardan</i>
'begin', lit. 'make start'	'learn by heart', lit. 'make by memory';
<i>sar dādan/sar dodan</i>	<i>dar bar gereftan/giriftan</i>
'let go', lit. 'give start', < <i>sar</i> 'head, start';	'embrace', lit. 'take into bosom', < <i>bar</i> 'bosom; memory', not used independently.

(2) Transparent conversions of a nominal using the same set of "dummy" auxiliaries, the semantic load remaining with the nominal:

<i>rāst konl</i>	<i>qofl kon- ~ zan-l</i>	<i>tahiye kon- ~ bin-</i>
<i>rost kun-</i>	<i>qulf kun- ~ zan-</i>	<i>tahiya kun- ~ bin-</i>
‘straighten’,	‘lock, bolt’,	‘prepare’,
lit. ‘make straight’;	lit. ‘make ~ hit lock’;	lit. ‘make ~ see to preparation’.

In literary style, the dummy *kon-*, *kard-lkun-lkard-* may be replaced by *nemā-*, *nemudl namo-*, *namud-* (originally ‘appear; show’), and *šav-lšav-*, *šod-lšud-* by *gard-lgard-*, *gašt-lgašt-* ~ *gardid-lgardid-* (originally ‘turn’):

<i>vazife-yešān-rā ejrā</i>	<i>‘nemud-e andl</i>	<i>vazife-yešān ejrā gašt ~ gardidl</i>
<i>vazifa-ašon-ro ijro</i>	<i>namud-a-and</i>	<i>vazifa-ašon ijro gašt ~ gardid</i>
duty-their-DO	deed do.PART COP-3p	duty-their deed become.PT.3s
‘they have carried out their duty’;		‘their duty was carried out’.

Semantically, LVCs frequently mark general categories such as “duration”. Thus, the two verbs *zadan* ‘hit’ and *kešidan* ‘draw, pull, stretch’ are paired as *dast zadan* vs. *dast kešidan* both ‘touch’. However, *kešidan* connotes continuous (moving) touch, and *zadan* once or quickly repeated action, including making sounds: *vāks zadan* ‘shoeshine’ (*vāks* ‘wax’); *harf zadan* ‘speak’ (*harf* ‘word’); *violon zadan* ‘play violin’; *gul zadan* ‘deceive’ (*gul* ‘deception’).

5.5.5.1 Reflexivity

While the light verbs tend to become semantically bleached, they also tend to develop a more generalized meaning. Thus, *kešidan* belongs to the subset of verbs that are both transitive and reflexive, as in *derāz kešid* ‘(s)he stretched out, lay down’. So does *xordanl xurdan* ‘eat, drink’, which as a light verb connotes “take to self”, and is one of the partners of *zadan*: *gul xordan* ‘be deceived’.

5.5.5.2 Experiencer-subject LVCs

In Persian, a distinct and open subset of LVCs consists of reflexive, or indirect, verb construction where the affected experiencer is marked by a personal enclitic in its dative function: *az ān xoš-aš āmad* ‘he liked it’, lit. ‘from that it became him well’; *qahr-emān šod* ‘we began to quarrel’, lit. ‘to us happened quarrel’; *sard-am ast* ‘I am cold’, lit. ‘to me it is cold’; coll. *če-t-e?* ‘what’s the matter with you?’ (< *če-at ast* ‘what is to you’). Today, experiencer LVCs are particularly frequent in non-formal speech, but they continue an impersonal construction well-documented throughout the diachronic stages of Persian, where it alternated with *-rā* in dative function: *ō-rā xwaš āmad*, lit. ‘it became him well’.

5.5.6 Voice and agency

5.5.6.1 Passive voice

(1) The passive voice is formed for all tenses with the past (perfect) participle of the main verb and the inflected auxiliary *šav-*, *šodlšav-*, *šud-* ‘become, get (intrans.)’:

<i>košt-e mi-šav-adlkušt-a me-šav-ad</i>	<i>košt-e šodlkušt-a šud</i>
‘(s)he is being, will be killed’	‘(s)he was killed’
or ‘is, will be getting killed’;	or ‘got killed’.

This includes the evidential forms and the Tajik conjectural forms, e.g. Tajik *ovard-a me-[šud-a]g-i-st* ‘it is probably being taken’.

(2) In compound verbs, *šav-*, *šod-lšav-*, *šud-* replaces the transitive auxiliary *kon-*, *kard-ikun-*, *kard-* ‘do, make’.

Tajik prefers the “long passive”, i.e. with passivization of *kard-an*: *kard-a šav-lšud-*, which is at best rare in Persian:

<i>ān</i>	<i>ketāb</i>	<i>tarjome</i>		<i>šodl</i>
<i>on</i>	<i>kitob</i>	<i>tarjuma</i>	<i>(kard-a)</i>	<i>šod</i>

that book translation (do.PART) become.PT.3s
‘the book was translated’.

(3) The passive construction has no overt agent, though the agent may be referred to indirectly, where felt necessary, by adverbial EZ phrases such as *az taraf-elaz taraf-i* ‘from the side of’, *be vāsete-yelba vosita-i* ‘by means of’, *az, be dast-e* ‘from, by the hand of’:

<i>in</i>	<i>nāme</i>	<i>az</i>	<i>taraf-e</i>	<i>edāre</i>	<i>nevešt-e</i>	<i>šodl</i>
<i>in</i>	<i>maktub</i>	<i>az</i>	<i>taraf-i</i>	<i>rafiq-am</i>	<i>nivišt-a</i>	<i>šud</i>

this letter from direction-EZ office/my friend write.PART become.PT.3s
‘this letter was written by the office/by my friend’.

The passive is comparatively little used in both Persian and Tajik, and is relatively more restricted semantically in Persian than in Tajik, and typically a feature of literary and bureaucratized diction.

The most prominent means of indicating an unspecified agent is the use of third person plural:

<i>ān-rā</i>	<i>tarjome</i>	<i>kard-andl</i>	<i>ān</i>	<i>tarjome</i>	<i>šodl</i>
<i>on-ro</i>	<i>tarjuma</i>	<i>kard-and</i>	<i>on</i>	<i>tarjuma</i>	<i>šud</i>

that-DO translation make.PT-3p that translation make.PART become.PT.3s
‘they translated it’; ‘it was (made) translated’.

5.5.6.2 Scale of agency and causation

The following illustrates the position of the passive construction on the scale of agency, using past tense forms (Windfuhr 1979: 105–113):

(a) Agent specified, causation:

<i>pesar</i>	<i>āb-rā</i>	<i>rixtl</i>	‘the boy poured the water out’.
<i>pisar</i>	<i>ob-ro</i>	<i>rext</i>	

boy water-DO pour.PT.3s

(b) Agent unspecified, causation:

<i>āb-rā</i>	<i>rixl-andl</i>	‘they, someone poured the water out’
<i>ob-ro</i>	<i>rext-and</i>	~ ‘the water was poured out’.

water-DO pour.PT-3p

(c) Agent unspecified, causation:

<i>āb</i>	<i>rixl-e</i>	<i>šodl</i>	‘the water was poured out’.
<i>ob</i>	<i>rext-a</i>	<i>šud</i>	

water pour.PART become.PT.3s

6.1.2 Coordinating conjunctions and clitics

The coordinating single and paired conjunctions are the same as those used to join NPs, but may have additional functions.

(1) Independent *valva* ~ enclitic *-ol-u*, etc. These denote a second event contemporaneous with or closely following the first, or a consequence of it, the latter emphasized by *valva*:

Same subject:

<i>in soxan-e talx-aš-rā goft</i>	<i>va az otāq birun</i>	<i>raftl</i>
<i>in suxan-i talx-aš-ro goft</i>	<i>va az xona bar-omad-a</i>	<i>raft</i>
this speech-EZ bitter-his-DO say.PT.3s	and from room out come.PART	go.PT.3s

‘he said these bitter words and left the room’.

Different subject:

<i>“keš” goft-am-o</i>	<i>gonješk be havā parvāz kard</i>	<i>raftl</i>
<i>“kiš” guft-am-u</i>	<i>čumčuk ba havo parvoz karda</i>	<i>raft</i>
“pshsh” say.PT-1s	and sparrow to air flight	make.PT.3s/ go.PT.3s
		PART

‘I shouted “kish”, **and** the sparrow flew up into the air’, *karda raft*, cf. conjunct verbs.

(2) The conjunctions *ammālammo*, *valilvāle*, *lākenllekin* ‘but, however’ (all borrowed from Arabic) introduce the second of two contrasted clauses:

<i>havā abri šod,</i>	<i>vali bārān na-bāridl</i>
<i>havo abr šud,</i>	<i>lekin boron na-borid</i>
weather cloudy/cloud become.PT.3s	but rain not rain.PT.3s

‘it became cloudy, but it didn’t rain’.

(3) *hamlham* and *kelki*

Ham. The clitic *ham* ‘also, too’ may have focusing-contrastive function, as does the generalized complementizer *kelki*. Note that the use of the latter is typically colloquial (see section 6.4.5 Idiomatic *ke*):

<i>to ān kār-rā</i>	<i>bo-kon,</i>	<i>man ham ~ ke</i>	<i>mi-rav-aml</i>
<i>to on kor-ro</i>	<i>kun,</i>	<i>man ham ~ ki</i>	<i>me-rav-am</i>
you.s that work-DO	SBJ.do.IMP.2s	I also though	IPFV.go.PR-1s

‘you do that work, I though ~ but I am leaving’.

In colloquial speech, the members of the preceding sets of linkers may occasionally be used in sequence, which can be rendered only approximately in English. Therefore these linkers mark three successive syntactic levels.

The *kelki*-clause here can be attached after the clitic *hamlham* (but not the coordinating conjunctions), which suggests that syntactically it is a subordinate clause.

(4) Anaphoric clitic *-aš*

The 3s clitic *-aš*, or the independent pronoun *ān/on*, refers to a preceding situation, mainly in colloquial speech:

<i>ba'd-eš ~ ba'd az un mā raft-iml</i>	<i>āxer-eš raftl</i>	<i>ku-š-eš? (Persian)</i>
<i>ba'd-aš ~ ba'd az on mo raft-em</i>	<i>oxar-aš raft</i>	
after-it after from that we go.PT-1p	final-it go.PT.3s	where-there-his?
'there(-aš)-after, then we left';	'in the end he left';	'where is he?'

(5) Anaphoric clitic *-rā*

The phrase clitic *-rā* in its general function of topicalizing reference (see section 3.1.4.2a) introduces clipped clauses:

Hasan-rā – ne-mi-dān-am če šode ast
 'as to Hasan, I don't know what happened'.

6.2 Subordinate clauses and linking

Subordinate clauses fall into three major types: (1) relative (property) clauses; (2) complement (propositional) clauses; and (3) adverbial clauses. Unmarked relative clauses are embedded in the matrix clause. Unmarked complement clauses, both subject and object clause as well as purpose clauses, are post-verbal. Unmarked adverbial clauses fall into two sets: those that precede the main clause, and those that follow it. This unmarked distinction follows the logical or temporal sequence of the events or situations in the respective main and subordinate clause.

Typologically significant is the existence of a generalized complementizer, *kelki* (partially comparable to English *that*).

This complementizer is an enclitic, and may introduce most types of subordinate clauses. As such, it precedes them when they follow the matrix clause, but is inserted after their first constituent component when they precede the matrix clause. Moreover, it combines with demonstrative pronouns and/or nouns in conjunctive constructions. Significant for the function of this complementizer are two facts: it may also attach to most semantically specified conjunctions (e.g. *čun kelčun (ki)* 'because'), and second there are pragmatic-semantic conditions where it is said to be optional. This suggests that *kelki* is not semantically empty.

Subordinate clauses may be linked as follows: (1) without morphological linker (asyndetic), (2) with the general complementizer *kelki*, and (3) with complex conjunctive constructs consisting of a nominal head + *kelki*. The head may be:

- (a) minimally a demonstrative prop, *inlin* or *ānlon*: *in ke lin ki*;
- (b) an EZ construction with the pronominal prop: *N-e in ke lN-i in ki*;
- (c) a noun + linker *-i*, with optional prop: *(ān) N-i kel(on) N-e ki*.

Depending on the role of the head in the matrix clause, these phrases may be modified by primary and secondary adpositions (see section 3.1.4.1 Prepositions), including the direct object marking postposition *-rāl-ro*: *in-rā kelin-ro ki*, *N-e in-rā kel N-i in-ro ki*, *N-i-rā kelN-e-ro ki*.

In 'this' is the default demonstrative prop in most contexts, whereas *ān* 'that' tends to be preferred in the head of relative clauses. The choice may be semantically significant: *in* pointing to the present or immediate situation, or anticipated a future event, and *ān* referring to a more remote situation or past state. – Both demonstratives may have focused forms marked by stressed *ham-* 'same', *haminlhamin*, *hamānlhamon* 'this, that very'.

Focusing, topical reordering and extraposition obligatorily requires the use of such a prop, minimally *kelki* (cf. also Aghaei 2006).

Note: In the following the interlinear annotation will be simplified.

6.3 Relative clauses

6.3.1 Restrictivity and specificity

Relative clauses (RCs) distinguish between restrictive and non-restrictive types. The former requires the linker *-il-e* before the generalized complementizer *kelki*, the latter does not. This distinction corresponds to the EZ construction marked by the linker *-el-i* vs. zero linkage in appositional NPs (Windfuhr 1979: 65):

Relative linkage	EZ linkage	
<i>N-i kelN-e ki</i>	<i>N-el-i</i>	complement
<i>N-keN-ki</i>	<i>N-</i>	complement

Like other nouns, the head noun of relative clauses may be marked by the indefinite *-il-e*, which is homophonous with the relative linker. When combined, these two markers coalesce into a single *-il-e*.

In addition, heads may be further marked by a demonstrative pronoun (*ān* being the default choice), or by indefinite *yek/yak* ‘one, a’. The following shows the basic patterns, here using Persian only (DEM = demonstrative; I = indefinite marker; L = linker).

TABLE 8.14: ORDER OF RELATIVE CLAUSES

	DEM	HEAD I	L	PROPERTY	PREDICATE
<i>Restrictive</i>					
(1a)	<i>ān</i>	<i>doxtar</i>	<i>-i</i>	<i>[ke Ali-rā dust dār-ad]</i>	<i>injā-st</i>
(1b)	<i>yek</i>	<i>doxtar -i</i>	<i>-i > -i</i>	<i>[ke Ali-rā dust dār-ad]</i>	<i>injā-st</i>
<i>Non-restrictive</i>					
(2a)	<i>ān</i>	<i>doxtar</i>		<i>[ke Ali-rā dust dār-ad]</i>	<i>injā-st</i>
(2b)	<i>yek</i>	<i>doxtar-i</i>		<i>[ke Ali-rā dust dār-ad]</i>	<i>injā-st</i>

Restrictive:

(1a) ‘(that) the girl who loves Ali is here’ (not the other one).

(1b) ‘a (certain) girl who loves Ali is here’ (not another one).

Non-Restrictive:

(2a) ‘(that) the girl, who I think, etc. loves Ali, is here’.

(2b) ‘a girl, who I think, etc. loves Ali, is here’.

6.3.2 Gapping and anaphora

There are no dedicated relative pronouns. Agreement and case roles of the head inside the relative clause are indicated by the combination of the general complementizer *kelki* and anaphoric reference. The latter is marked in the verbal ending if same subject, and by an anaphoric pronoun otherwise; non-anaphoric repetition is an emphatic option. The occurrence of the anaphoric phrases ranges from obligatory to optional. When

optional, occurrence marks emphasis (examples here from Persian only; translation is literal; parentheses = optional; double parentheses = strongly emphatic):

Basic cases:

Subj., <i>N who</i> , opt.:	<i>doxtar-i ke ((u)) Ali-rā dust dār-ad</i> 'the girl that she Ali loves';
Dir. obj., <i>N whom</i> , opt.:	<i>doxtar-i ke Ali ((u-rā)) dust dār-ad</i> 'the girl that Ali her loves';
Ind. obj., <i>N to whom</i> , obl.:	<i>doxtar-i ke Ali be u gol dād</i> 'the girl that Ali to her flowers gave';
Abl., Comp., <i>N from/than whom</i> , obl.:	<i>doxtar-i ke Ali az u kuček-tar ast</i> 'the girl that Ali from her younger is';
Poss., <i>N whose</i> , obl. (mostly clitic):	<i>doxtar-i ke Ali dust-aš ast</i> 'the girl that Ali friend-her is'.

Adverbial:

Time, <i>N when</i> , opt.:	<i>ruz-i ke ((dar ān)) in ettefāq oftād</i> 'the day that (on) this happened';
Place, <i>N where</i> , opt.:	<i>deh-i ke (dar ān) zendegi mi-kard</i> 'the village that in it she lived';
Reason, <i>N why</i> , none:	<i>ellat-i ke harf na-zad</i> 'the reason that she didn't talk';
Manner, <i>N how</i> , none:	<i>towr-i ke ((be ān)) in kār-rā anjām dād</i> 'the manner that in it this work she accomplished'.

Clausal:

CL, <i>which</i> , none:	<i>aslan harf na-zad, ke xeyli ajib ast</i> 'at all she didn't talk, which is very strange'.
--------------------------	---

6.3.3 Ordering of restrictive RCs

Restrictive relative clauses distinguish an unmarked and a marked order: (1) When unmarked, they immediately follow the head phrase, and are thus embedded within the matrix clause as in the example above: (2) When marked, they are post-verbal, and have focusing-contrastive function. The post-verbal position is also typical for relative clauses that identify a characteristic of the head, irrespective of whether the head is indefinite or definite: (*ān*) *N(-il-e)*:

(1) Unmarked:

<i>ān doxtar-i [ke Ali-rā dust dār-ad] raft</i> that girl-L that A.-DO friend have.PR-3s go.PT.3s '(that) the girl whom Ali loves left'.
--

(2) Contrastive:

<i>ān doxtar-i raft [ke Ali-rā dust dār-ad]</i> that girl-L go.PT.3s that A.-DO friend have.PR-3s go.PT.3s '(that) the girl [whom Ali loves] left', not the other one.
--

(3) Topical:

yek ketāb-i-rā *towsiye* *me-kon-am* [*ke xod-am xānd-am*]
yak kitob-e-ro *tawsiya* *me-kun-am* [*ki xud-am xond-am*]
 one book-INDEF-DO commendation IPFV.make.PR-1s that self-my read.PT-1s
 'I'll recommend a (certain) book that I have read myself'.

Copular relative clauses are necessarily post-verbal because the copula is enclitic. However other forms of 'be', such as subj. *bāš-/boš-*, past *bud-/bud-* and the existential verb *hast-/hast-* 'be there' do allow for positional options:

ān doxtar-i-st [*ke Ali-rā dust dār-ad*]
 that girl-L COP.3s who Ali-DO friend has
 'she is the girl who loves Ali'.

6.3.4 Use of moods

Inside relative clauses, the moods express the three degrees of reality in the verb system, from factual to potentially factual to counterfactual, all necessarily pragmatic responses to situations, whether with explicit reference or as general statements (then often with present subjunctive). Typically, relative clauses may be relativized adverbial clauses (see section 6.5 Adverbial clauses).

(1) Indicative

(a) Facts:

duxtarča [*ki dar gird u peš-i xud modar-i xud-ro na-yoft*]
 little girl [that in around and front-EZ self mother-EZ self-DO not.find.PT.3s
ba girya daromad
 to cry.INF in come.PT.3s
 'the little girl, who did not find her mother anywhere near, began to cry', implicitly also 'when' and 'because she did not find'.

(b) Referential facts:

in-jā kas-i-st *ke kār mi-kon-adl*
injo kas-e-st *ki kor me-kunad*
 here person-L COP.3s that work IPFV.do.PR-3s
 'here is someone who works, is working'.

(c) General statements:

Indicative

kas-i *ke kār mi-kon-ad* *xošhāl astl*
kas-e *ki kor me-kun-ad* *xursand ast*
 person-L that work IPFV.do.PR-3s happy COP.3s
 'someone who works is happy'.

(2) Subjunctive

(a) General characteristic:

in-jā kas-i nist ke kār be-kon-ad
injo kas-e nest ki kor bi-kun-ad
 here person not.COP.3s that work SBJ.do.PR-3s
 'there is no one here to work';

(b) General comparison:

mesl-e mard-i ke aslan kār-i na-karde bāš-adl
misl-i mard-e ki aslan kor-e-ro na-karda boš-ad
 like-EZ man-L that at all work-INDEF-DO not.do.PART (SBJ)be.PR-3s
 'he looks like a man who has done absolutely nothing';

(c) Purpose:

be zabān-i ke mā ham be-fahm-im harf be-zan-id
ba zabon-e ki mo ham bi-fahm-em gap zan-ed
 in language-L that we too SBJ.understand.PR-1p word speak.IMP-2p
 'speak the/a language that we can understand';

(d) Wish:

ketāb-i mi-xāh-am ke gerān na-bāš-adl
kitob-e me-xoh-am ki qimat na-boš-ad
 book-INDEF:L IPFV.want.PR-1s that expensive not (SBJ.)-be.PR-3s
 'I want a book that hopefully is not expensive';

(e) Condition and consequence:

on ki ba mo dastdarozi kun-ad, albata, bozi kun-ad
 that that to we handraising SBJ.make.PR-3s surely, play SBJ.make.PR-3s
 'he who raises his hand against us surely risks his head', Tajik.

(4) counterfactual condition

čiz-i be man foruxt ke xod-aš ne-mi-xaridl
čiz-e ba mo furūxt ki xud-aš na-me-xarid
 thing-INDEF-L to I sell.PT.3s that self-his not.IPFV.buy.PT.3s
 'he sold me something that he himself wouldn't buy'.

6.3.5 General-indefinite RCs

(1) Indefinite relatives '-ever'

Clauses introduced with the distributive pronouns *har* 'each, every' designate a restricted class of entity, defined only by the clause itself. The heads are relative constructs with indefinite pronouns and adverbial nouns, and have the structure:

(a) *har kelharki* 'whoever' vs. *har-čelhar-čī* 'whatever', *har kas kelhar kas ki* 'whoever (person)' vs. *har ān čelhar on čī* 'whatever'; *har čandlhar čand* 'however much, although, notwithstanding'; (b) *har jā kelhar jo ki* 'wherever (place)'; *har kojā kelkujo ki* 'wherever'; *har vaqt kelhar vaqt ki, vaqt-i kelvaqt-e ki* 'whenever'; *har towr-i kelhar tawr-e ki* 'however (way)'. Note that these are the only type of relative heads that distinguish

human and non-human gender (see section 3.1.3.7 Indefinite pronouns). The use of tense and mood is the same as in other relative clauses.

(2) Specific indefinite heads (optional demonstrative requiring *kelki*):

har ān ke dānā bāš-ad tavānā bāš-ad
har on ki dono boš-ad, tavono boš-ad
 each that that knowledgeable (SBJ.)be.PR-3s powerful (SBJ.)be.-3s
 ‘whoever is knowledgeable is wise; free variation of saying:
har ke dānā bov-ad tavānā bov-ad, which retains the archaic subjunctive stem *bov-* for *bāš-*.

(3) Non-specific pronominal heads

Non-specific heads do not require *kelki*:

har kojā ke be-rav-i, be-gu ke mā ki hast-im
har kujo ki rav-i, falimon ki mo ki-st-em
 wherever that SBJ.go.PR-2s IMP.say/understand that we who EX.1s
 ‘wherever you go, tell people who we are’;

har kas-i-rā mi-did, az u pul qarz mi-kard
har kes-e-ro me-did, az ū pul qarz me-kard
 each person-L-DO IPFV.see.PT.3s from he money loan IPFV.make.PT.3s
 ‘he borrowed money from whomever he met’ (Lazard 1992: 235).

(4) Non-pronominal indefinite heads:

Similar to other post-verbal object clauses, which do not require *kelki*, lexical heads that introduce object clauses do not require the complementizer. (In such cases, there tends to occur a pause when uttered):

mi-xāst-am jā-i bud-am- na ādam-i bāš-ad, na dard-esar
me-xost-am jo-e bud-am- na odam-e boš-ad, na dard-i sar
 IIPFV.want.PT-1s place-L be.PT.1s -not person-INDEF (SBJ.)be.-3s not headache
 ‘I wished I were somewhere where there were no people and no worry’;

barā-ye ān kār šab-i mi-xāh-ad
baro-i on kor šab-e darkor
 for-EZ that work night-L IPFV.want.PR-3s/needed

– *sobh na-dāšte bāš-ad*
 – *subh na-došta boš-ad*
 – morning not.have.PART (SBJ.)be.3s
yā ruz-i mi-xāh-ad – *āftāb-āš*
yo ruz-e darkor ast oftoḡ-aš
 or a day IPFV. want.PR-3s/needed COP.3s – sunshine-its

foru na-rav-ad

furū na-rav-ad

does not(SBJ.)-go.PR-3s

‘in order to do that, one needs a night that has no morning, or a day when the sun never sets’.

6.3.6 *Head incorporation*

Head incorporation is an alternative to the anaphoric strategy: the case rection of the correlating NP in the subordinate clause is moved to the head noun phrase; this process tends to trigger focused lexical repetition in the main clause:

- (1) No change of sequence
 (a) Direct object *-rāl-ro*:

- (1) indefinite-specific object:

<i>magar [zan-i</i>	<i>ke]</i>	<i>ensān u-rā dust mi-dār-ad]</i>	<i>almās-o javāher-ast?</i>
<i>magar [zan-i-rā</i>	<i>ke]</i>	<i>ensān dust mi-dārad]</i>	<i>almās-o javāher-ast)?</i>
but a woman-L DO	that	man her loves	diamond and jewels is

‘but is the woman that a person loves diamonds and jewels?’

(Ali Dashti, cited in Lazard 1992: 232)

- (2a) Prepositional phrase, Persian

<i>bāyad dānest [zabān-i</i>	<i>ke]</i>	<i>mā [bā ān]</i>	<i>emruz soxan mi-gu-yim . . .</i>
<i>bāyad dānest [bā zabān-i</i>	<i>ke]</i>	<i>mā</i>	<i>emruz soxan mi-gu-yim] . . .</i>
must know with language-L	that we	with it	today speech say

hamān zabān ast ke . . .

same language is that

‘one must realize that the language with which we communicate today . . . is the same language which . . . (Kasravi Tabrizi, cited in Lazard 1992: 232)

- (2b) Tajik

<i>[</i>	<i>ba hamon jo-e ki az ān]</i>	<i>omad-i</i>	<i>bar gard!</i>
<i>[az jo-e</i>	<i>ki]</i>	<i>omad-i], [ba hamon jo]</i>	<i>bar gard</i>
from the place	that	you came	to same place return

‘go back to where you came from!’.

6.3.7 *Tajik nominalized relative clauses*

Two types of nominalized clauses were discussed in sections 4.2.2.5d, e. Both are EZ constructions where the dependent verb is nominalized as a perfect participle. The latter is the perfect participle in *PT-e/PT-a* (*raftelrafta*) which alternates with the participle in *PT-agi* (*raftagi*) in Tajik. While in Persian the participial EZ construction is of limited length, Tajik has developed an elaborate system that goes far beyond the inherited pattern and allows for full-length participial clauses.

Typologically significant for this pattern is the fact that participialization does not involve the complete loss of tense, mood, and aspect: while tense distinction is lost (as in a number of finite forms, see Table 8.8), modal and aspect marking are retained. That is, the participial forms may be marked by the prefix *me-* for imperfectivity and non-indicative mood as well as by the participle *istoda* for progressivity. Given the two Tajik participles, the system of the participial forms are shown in Table 8.15 (*rav-/raft-* ‘go, leave’).

TABLE 8.15: TAJIK FINITE AND PARTICIPIAL FORMS

Finite forms			
<i>rafta istoda ast</i>	<i>me-rav-ad</i>	<i>rafta ast</i>	PRS
<i>rafta istoda bud</i>	<i>me-kard</i> , and <i>(bi-)kun-ad</i>	<i>rafta bud</i>	PST SBJ
Participial forms in <i>-a</i>			
<i>rafta istod-a</i>	<i>me-rafta</i>	<i>rafta</i>	
begin of progress or in progress	general process or continuing state, and modality	process concluded or temporary state	
Participial forms in <i>-agī</i>			
<i>rafta istod-agī</i>	<i>me-raft-agī</i>	<i>raft-agī</i>	
begin of progress or in progress	general process or continuing state, and modality	process concluded or temporary state	

The distinction between these two sets is partially a matter of register, the *-agī* forms being notably more frequent in colloquial speech.

As is the case with the reduced participial construction shared by both Persian and Tajik, the passive interpretation of these constructions with transitive verbs is partially suspended, and voice interpretation depends on the context, including the animacy of the subjects. Most distinct from Persian is the fact that in these constructions the agent-subject may be marked by the possessive clitics.

6.3.7.1 Level of nominalizations

Compared to finite subordination, the process from finite clause to full nominalization and Turkification involves not only the deletion of conjunction and nominalization, but, as mentioned, the replacement of the finite subject verb ending (and pronominal subject) over several stages with a possessive pronoun, and, in a final step, the inversion of head and clause. Schematically, the sequence of these stages may be shown using *-agī* forms (SUBJ = subject; VB = verb form; END = finite ending; POSS = possessive clitic):

Unmarked: *raft-agī*

(0) *ob-e [ki mo az hawz ovard-em]*
water-L that we from pool bring.PT-1p
'the water [that we brought from the pool]'

>>

(1) *ob-i [mo az hawz ovard-agī]*
(2) *ob-i [az hawz ovard-agi-amon]*
(3) *[az hawz ovard-agi-amon] ob*

Indicative or subjunctive: *me-raft-agī*,

corresponding to present or past according to context

(0) *ob-e [ki mo az hawz me-ovard-em]*
'the water that we from pool IPFV.bring.PT-1s
'the water that we are ~ were to bring from the pool'

>>

(1) *ob-i [mo az hawz me-ovard-agī] EZ*
(2) *ob-i [az hawz me-ovard-agi-amon]*
(3) *[az hawz me-ovard-agi-amon] ob*

Progressive: *rafta istod-agī*

corresponding to present or past according to context

(0) *ob-e [ki mo az hawz ovarda istoda em]*
 water-L that we from pool bring. PART PROG.PART COP-1p
 >>

(1) *ob-i [mo az hawz ovarda istod-agī]*

(2) *ob-i [az hawz ovarda istod-agi-amon]*

(3) [az hawz ovarda istod-agi-amon] **ob**

As is true for any other relative clause, the case rection of a nominalized relative clause in the matrix clause is irrelevant for its internal structure, and is indicated by a preposition preceding its head, and by *-rā* following the entire clause.

Further examples

(1) *raft-agī*

Finite *ke navisond-id* > *navisond-agi-aton* 'which you dictated', lit. 'caused to write':

mazmun-i maktub-ho-i [navisond-agi-aton]-ro ba ū naql kard-em
 content-EZ letters-EZ [having dictated-yo ur.p]-DO to he repeat make.PT-1p
 'we repeated to him the content of the letters [you had dictated]'

(2a) *me-raft-agī*, indicative

mon ki on urf-u odat-ho-i
 allow that that traditions and custom-p-EZ
[zan-ro kaniz, mard-ro yulom me-kard-agī]
 [woman-DO slavegirl, man-DO slave IPFV.making.]
az miyon bar-došta šav-ad
 from among abolish.PART (SBJ)-become-2s
 'let those traditions that make women into servant girls and men into slaves be abolished'; gerundial 'the traditions of making . . .'

(2b) *me-raft-agī*, modal function

man zan-i [sar me-dod-agī] na-dor-am
 I wife-EZ [release IPFV-giving] not.have.PR-1s
 'I don't have a wife to divorce', i.e. 'I'm not about to divorce my wife'.

(3a) *rafta istod-agī*, active interpretation

(indicated by presence of personal clitic)

qozi-ho poy-ho-i [ba taht-i šikam kašida-istod-agi-ašon]-ro ham
 the qadis foot-p-EZ [to under-EZ belly being pulling-their]-DO even
ba kor dar-ovarda, du-poya gurext-and
 to work bring.PART] two-footed flee.PT-3p
 'even the qadis brought their legs [that they had been pulling under their bellies] into action, and fled hotfooting'.

(3b) *karda istod-agī*, passive interpretation

(indicated by lack of possessive clitic)

agar etz-i [vay yod doda istod-agi]-ro
 if thing-EZ [he memory give.PART PROG]-DO
durust gufta na-tavon-ad
 correct say.PART not(SBJ.)-can-3s
 'if he cannot correctly repeat something that he is being taught . . .'

In the ultimate reduction, these participializations are compacted to nouns:

[ba qur'on amal namud-agī] kas [dar xona bud-agi]-ho
 [to Koran act doing] person [in room being/-plural
 'a person acting acc. to the Koran'; 'those who are/were in the room';

[qalam be javob na-me-girift-agī]
 [pen without permission not-IPFV.take.ADJ]
 'who is never taking the pen without permission'.

6.4 Complement clauses

Complement clauses comprise subject and object clauses as well as purpose (or final) clauses. The former two are typically in the indicative, and will be discussed in the following, while purpose clauses are typically in the subjunctive and will be discussed under Adverbial clauses (see section 6.5).

6.4.1 Subject and object clauses

Both subject and object clauses have the same basic syntactic structures and follow the same rules of permutations (see also Aghaei 2006). The order of matrix clause and complement clause defines three levels of functional markedness:

- (1) In the unmarked order, the complement clause, with or without the complementizer *kelki*, follows the matrix clause.
- (2) In the focused order, a cataphoric demonstrative, mostly *in*, with or without a lexical head, is inserted in the matrix clause, and the complement clause is introduced by the complementizer.
- (3) In the topical order, there is inversion: the complement clause precedes the matrix clause, and is introduced by the cataphoric demonstrative plus complementizer, *in kelin ki* (identical to one of the heads introducing relative clauses), with object marker *-rāl-ro* inserted if object clause.

Example patterns

((1) = object clause; (2) = subject clause; (3) = relative clause; AG = agent; parentheses: optional). Note that agents are optional since subjects are obligatorily marked in the verb ending.

TABLE 8.16: FOCUSED AND TOPICAL ORDER

(1) Basic unmarked pattern, post-verbal:

(AGENT)	PREDICATE	+	COMPLEMENT
(a) (<i>Hasan</i>) Hasan	<i>mi-dān-ad</i> knows	+	[<i>ke Ali in xāne-rā sāxt</i>] that Ali this house-DO built
(b) (empty)	<i>doruq-ast</i> lie is	+	[<i>ke Ali in xāne-rā sāxt</i>] that Ali this house-DO built

(a) '(Hasan) knows that Ali built this house';
(b) 'it is a lie that Ali built this house'.

(2) Focused pattern, post-verbal:

(AG) IDEM (HEAD)	PREDICATE	+	COMPLEMENT
(a) (<i>Hasan</i>) <i>jin-rā</i> Hasan this-DO	<i>mi-dān-ad</i> knows	+	[<i>ke Ali in xāne rā sāxt</i>]
(b) <i>jin</i> (<i>xabar</i>) this information	<i>doruq-ast</i> is lie	+	[<i>ke Ali xāne-rā sāxt</i>]

(a) 'Hasan does know (this), that Ali built this house';
(b) 'this (matter) is a lie, that Ali built this house'.

(3) Topical pattern, preposed:

PROP IDEM (HEAD)	Topical inversion COMPLEMENT (AG)	PREDICATE
(a) <i>jin</i> (<i>xabar</i>)(- <i>rā</i>)	[<i>ke Ali in xāne-rā sāxt</i>] >	(<i>Hasan</i>) <i>mi-dān-ad</i>
(b) <i>jin</i> (<i>xabar</i>)	[<i>ke Ali xāne-rā sāxt</i>] >	<i>doruq-ast</i>

(a) '(this matter) that Ali built this house, Hasan knows (about) it';
(b) 'this matter) that Ali built this house, that's a lie'.

6.4.1.1 Use of TMA

Epistemic matrix verbs of observation (cognition) of facts (see, know, understand) as well as verbs and expressions governing potential actions or states (modal verbs; want, must, can; decide, try, etc.) do not map their tense onto the dependent verb. Rather the former require the indicative (present or present perfect), and the latter the subjunctive (present subjunctive and perfect/past subjunctive):

fahmid-am ke ādam-e xub-i ast
fahmid-am ki odam-i xub-e ast
understand.PT-1s that person-EZ good-INDEF COP.3s
'I realized that he was a good person';

did ke hame rafta and
did ki hama rafta and
see.PT.3s that all go.PART COP.3p
'he saw that all had left';

tasmim geraft-am be ān-jā be-rav-am
qaror dod-am ba on-jo rav-am
decision take/give.PT-1s to there SBJ.go-PR-1s
'I decided to go there'.

6.4.2 Embedded speech

6.4.2.1 Levels of integration

Embedded speech may be direct, indirect, or integrated indirect.

(1) Direct speech

Persian and Tajik partially differ in the representation of direct speech. In Persian, the speech string follows the verb of saying, which may be introduced by the general complementizer *kelki*:

mo'allem goft ke al'an daftar-hā-yetān-rā be-gir-id
 teacher say.PT.3s that now notebook-p-your.p-DO SBJ.take.PR-2p
 'the teacher said, now take your notebooks'.

(2) Indirect speech and questions

mo'allem goft ke hālā daftar-hā-yemān-rā be-gir-im
muallim guft ki aknun daftar-ho-yamon-ro gir-em
 teacher say.PT.3s that now notebook-p-your.p-DO SBJ.take.PR-1p
 'the teacher told us to take our notebooks now'.

The interrogative marker *āyālo* 'whether' is used in with indirect questions, with or without the complementizer *kelki*, paired *ke āyā . . . yā . . . , ki . . . yā . . .*:

ma'lum nist ke āyā u hamrāh mi-ā-yad, yā nal
ma'lum nist ki oyo vay hamroh me-o-yad, yo na
 clear not.COP.3s that whether he along IPFV.come.PR-3s or not
 'it isn't clear whether he is coming along or not'.

(3) Implicit indirect speech

Direct speech may be further integrated with verbs of cognition:

šomā xod-etān mi-dān-id ke u kojā raft
šumo xud-aton me-don-ed ki ū kujō raft
 you.p self-your.p IPFV.know.PR-2p that he where go.PT.3s
 'you know where he went';

mi-dānest-am ke mo'allem-emān dir bi-ā-yad
me-donist-am ki muallim-amon der bi-o-yad
 IPFV.know.PT-1s that teacher-our late SBJ.come.PR-3s
 'I knew that our teacher would be late'; here the subjunctive implies a response to the question about the teacher's absence.

6.4.3 Adpositional and EZ complement clauses

Distinct subsets of complement clauses are those that depend on complex verbs consisting of adpositional noun phrases or EZ constructions based on adjectives, participles, and nouns, many of which are Arabic loans:

(1a) preposition (a) *azlāz* 'from' (source, cause), *tarsidan azlāz* 'be afraid of', *xošhāl budan azlāz* 'be happy about'; (1b) preposition *bālbo* 'with', *movāfeq*

budan bālmuvofiq budan bo ‘agree with’ (Arabic participle); (2) EZ construction, *montazer-e . . . budan ~ entezār-e . . . dāštan/muntazir-i . . . budan ~ intizor-I . . . doštan* ‘be expecting, waiting for’ ~ ‘have the expectation of’ (Arabic participle and noun).

In these subsets, the retention of adpositions and EZ construction marks topicalization.

(a) source *az* ‘from’:

xošhāl-am ke āmad >>

az ān *xošhāl-am ke āmad*

‘I am happy that he came, is here now’ >>

‘what I am happy about is that he came’, lit. ‘from that I am happy that . . .’;

(b) objective EZ:

montazer bud-am ke bi-ā-yad >>

muntazer-e ān bud-am ke bi-ā-yad

‘I was waiting for him to come’ >>

‘what I was waiting for was for him to come’, lit. ‘I was waiting that, that he come.

6.4.4 Tajik Turkic-type nominal clauses

In Tajik there are not only the Persian type, as above, but also at least three variants of a Turkic conjunct-type of construction, which are more frequent in literary Tajik. In addition there are finite two-clause and nominalized one-clause constructions.

(1) Two-clause constructions

(a) The matrix clause follows the speech string, and the verb of saying is incorporated parenthetically within the speech string, or more typically is enclitic to it (rather than clause-final):

[*man*] *guft* [*ba šahr na-me-rav-am*]
I say.PT.3s to town not.IPFV.go.PR-1s
‘I, she said, am not going to town’;

[*uka, bisyor kalon-girā na-kun-ed!*] *guft* *Šokir ba Ahmad*
little brother, very high-mighty not.IMP.do-2p say.PT.3s Shokir to Ahmad
‘little brother, don’t be so high-and-mighty, said Shokir to Ahmad’.

(b) with nominalized forms

(ba) Clitic conjunct participle. If the matrix verb of saying is anything other than *guft-an* ‘say’, a conjunct construction using the participle *guft-a* is followed by an inflected form of the specific speech verb:

[*xud-i ū ki-st?*] *gufta man az Rahim Qand porsid-am*
[self-EZ he who COP.3s] say.PART I from Rahim Qand ask.PT-1s
‘[who is he, actually?] I asked Rahim Qand’.

(bb) A loose conjunct construction using either *guft-a* or *gū-yon* frequently marks an idealized speech string which may be in mixed indirect and direct speech, as a sentential complement (often a purposed phrase or clause).

on-ho grup-a kard-and, [hamin rais-a na-me-xoh-im]
 they group-DO make.PT-3p this very chairman-DO not-IPFV-want.PR-1p
gufia
 say.PART
 ‘they organized [in order to protest (the appointment of) this chairman]’;

[xūjin-ro ham ba šumo [[mukofot-gūyon] dih-ad]
 saddlebag-DO also to you.p reward say.PART.PR (SBJ.)give.PR-2p
 ‘he’ll give you the saddlebag too, [as a reward]’, lit. ‘. . . “reward” saying . . .’

This is similar to a typically Turkic construction, using *dep* ‘saying’ in Uzbek.

(c) Quotative-evidential phrase.

A speaker repeating someone else’s words to a third party may end with a quotative *buda-ast*:

Pulod ako,
 Polod sir,
[modar-am xūrok tayyor kard-and, roš-ed] bud-a-ast
 mother-my meal prepared made-3p hurry-2p be.PART.EV-COP3s
 ‘Pulod, I’m to tell you [my mother has cooked a meal, and to hurry up]’

(2) Single-clause constructions

In these constructions the verb of the subject and object clause is an infinitive and the agent is converted to a possessive enclitic, and/or *xod* ‘self’:

mo me-don-em ki [kujo me-rav-em] >>
[kujo raftan-i xud]-ro me-don-em
 ‘we know where we are going’ >>
 lit., ‘we know our where-going’;

vay guft ki ba qišloq me-rav-ad >>
vay [ba qišloq raftan-aš]-ro guft
 ‘he said that he was going to the village’ >>
 lit. ‘he said our to the village-going’.

Double nominalization:

darkor nest ki bi-gūy-em [ki in duxtar ki-st] >>
[kī budan-e in duxtar]-ro guftan darkor nest
 ‘it is not necessary that we tell who this girl is’ >>
 lit. ‘saying the who-being of this girl-DO is not necessary’.

6.4.5 “Idiomatic” clitic *ke*

In the Persian colloquial register, the complementizer *ke* may have so-called “idiomatic”, i.e. asseverative function. This is because it does not seem to introduce dependent clauses, but is inserted after the first syntactic constituent of main clauses, mostly brief responses or observations. Pragmatically, it appears to have strong or weak emphatic as well as contrastive function and accordingly tends to be rendered in English by ‘though, however, but, why’ and similar, depending on the context:

in ke kār-i na-dār-e
 this KE work-INDEF not.have.PR-3s
 ‘this, though, does not require any major effort’, or, smoother:
 ‘why, that’s no problem’ (Hayyim, *Farhang-e yekjeldi-e fārsi-englishi*, p. 691).

It has been difficult to identify the syntactic role of this particle. Similar to others, Öhl and Korn (2008: 142) regard this *ke* as homonymous with, but distinct from, the complementizer *ke*, and suggest *Enunziativmarker* for its function.

However, the very designation “idiomatic”, combined with use in the colloquial speech register, should be taken as a hint that there may be no enigmatic non-subordinating function of *ke* involved at all (discussed in Windfuhr 2004). Rather, one expects elliptic constructions, which are typical for that register. Following these observations, it becomes apparent that what is elliptic here are contextually understood stereotype subject and object matrix clauses. The general tenor of such introductory clauses can be easily imagined, given the context of discourse: *ma’lum ast* ‘it’s obvious’, *be nazar-am* ‘in my opinion’, *nist* ‘isn’t it (true)’, *mesl-e in ke* ‘it seems like’ for subject clauses, and *mi-bin-id* ‘you see’, etc., for object clauses. The second position of *ke* is explained by the ellipsis: with the deletion of the introductory phrase, this enclitic necessarily shifts to second position, which may be the clause itself. The syntactic process in the example cited above and some other examples may be shown as follows:

Non- “idiomatic”	“Idiomatic”
<i>be nazar-am [ke šām xord-i]</i>	<i>šām ke xord-i?</i>
in view-my that supper eat.PT-2s	
‘it appears that you did have supper’	‘you did have supper, didn’t you?’
<i>mi-bin-id [ke jā na-dār-e]</i>	<i>jā ke na-dār-e!</i>
IPFV.see.PR-2p that place not-have.PR-3s	
‘you see (yourself) that there is no space!’	‘but there is no space!’
(in response to the suggestion to sit down),	
<i>nist [ke javāb na-dāšt?]</i>	<i>javāb ke na-dāšt</i>
not.COP.3s that answer not.have.PT.3s	(no question contour)
‘isn’t it (true) that there was no answer?’	‘there was no answer; right?’
<i>motma’en am [ke in kār-i na-dār-e] in ke kār-i na-dār-e</i>	
sure COP.1s that this work-INDEF not-have.PR-3s	
‘I’m sure that this doesn’t require much effort’	‘why, that’s no problem’.

6.5 Adverbial clauses

The basic rules for the sequence of matrix and adverbial clauses depends on the logical or temporal relation of their events or states: when the latter precede, the adverbial clauses are in first position; when they follow, their clauses are in second position. Thus, causal, temporal, conditional, and concessive clauses, as well as circumstantial clauses of place, manner, and degree, precede, while consecutive, resultative, and purposive (final) clauses, as well as causal-explanatory, interruptive-temporal, and conditional-exceptional clauses, follow. However, as is the case with relative and complement clauses, both focusing and

topical inversions are frequent, depending on the pragmatic context. In Tajik, these finite adverbial clauses compete with nominalized adverbial clauses.

There are four types of subordinators: parataxis; the general complementizer *kelki* (roughly comparable to English conjunctive 'as'); a small set of three primary conjunctions, with or without the complementizer, as well as two interrogatives; conjunctive phrases with the complementizer.

Parataxis and kelki occur prominently in colloquial speech. Being lexically empty, they may substitute for a number of other subordinators, depending on the pragmatic context:

pul ke mi-deh-i, mi-ā-y-am
 money KE IPFV-give.PR-2s, IPFV.come.PR-1s
 '(as) you give me money, I come', more specifically,
 'because ~ when ~ if ~ even though you give me money for this, I will come'.

The primary conjunctions include:

- (a) *čunlčun* 'because, since', also 'when' (only literary);
- (b) *vaqt-i/vaqt-e* 'when' (< conjunctive phrase *vaqt-i kel/vaqt-e ki*).
- (c) *tālto*, (ca) temporal and conditional duration or end point, 'till, until, as long ~ soon as' (note *tā na-lto na-* 'un-less'); (cb) comparison 'than';
- (d) *agarlagar* 'if';
- (e) *magar/magar* 'unless';
- (f) *zirālzero* 'for this reason, because';

Conjunctive phrases are numerous, and syntactically identical with the heads of relative clauses (see section 6.3) such as:

(PREP) N- <i>il-e</i>	<i>kelki,</i>	<i>(dar) vaqt-i kel/vaqt-e ki</i> 'at time that' = 'when';
(PREP) N- <i>el-i</i>	<i>in kelki,</i>	<i>(be) ellat-e in kel (ba) illat-i in ki</i> 'for the reason that' = 'because';
PREP	<i>in kelki,</i>	<i>bā in kelbo in ki</i> 'with this that' = 'although';
N *PREP	<i>in kelki,</i>	<i>ba'd az in kelba'd az in ki</i> 'after from this that' = 'after doing . . '

* The nominal may be a noun, adjective, or adverb.

6.5.1 Position and use of TMA

(1) Position. *Kelki* is an unstressed enclitic. In sentence-initial subordinate clauses it is typically inserted after the first constituent component of the clause; in sentence-final position, it is attached to the preceding matrix clause (similar to the enclitic coordinator *-ol-u* 'and'). Other clausal heads may optionally be in second position as well (see section 5.1 Order of constituents).

(2) Mood. The use of mood is predictable: Events or states require the indicative when real (factual), the subjunctive when potential, and the counterfactual when unlikely:

<p>Real <i>mi-xāh-i, mi-rav-iml</i> <i>me-xoh-ī, me-rav-em</i> you want IPFV.go.PR-1p ‘if you really want to, we will go’;</p> <p>Counterfactual <i>man mi-raft-aml</i> <i>man me-raft-am</i> I IPFV.go.PT-1s ‘I would go (if I were you)’.</p>	<p>Potential <i>raft-am, tā ~ ke az u be-pors-aml</i> <i>raft-am, to ~ ki az ū (bi-)purs-am</i> go.PT-1s till ~ that from he SBJ.ask.PR-1s ‘I went in order to ask him’.</p>
---	--

(3) Aspect is predictable as well. Notably the preterite (i.e. the perfective aorist form in the verb system) is required when an event or state is completed prior to that of the matrix clause (while English uses the present). Conversely, incomplete events or states require the subjunctive, notably those headed by *piš az in kelpēš az in ki* ‘before’, *be jā-ye in kelba jo-yi in ki* ‘instead of’, *bedun-e in kelbidun-i in ki* ‘without’:

<p><i>agar u-rā did-i, be-gu</i> <i>agar ū-ro did-ī, bi-gū</i> if he-DO see.PT-2s IMP.say-2s ‘once you see him, tell me’;</p>	<p><i>bedun-e in ke čiz-i be-gu-yad, raft</i> <i>be on ki čiz-e (bi)-gū-yad, raft</i> without this that anything SBJ.say.PR-3s go.PT.3s ‘he left without saying anything’.</p>
--	---

The basic unmarked patterns of clauses, together with verbal modality (the Persian *consecutio temporum*), may be shown as follows (cf. Windfuhr 1987: 539/2009: 457). Note that in each column, the subordinators for pre-matrix position may also occur in post-matrix position.

TABLE 8.17: PATTERNS OF CONJUNCTIONS AND MOODS

<i>Preceding Main Clause</i>		<i>Following Main Clause</i>					
Cause	<i>čunlčun</i>	‘because’	IND	Explanation	<i>zirā, zero</i>	‘(that is) because’	IND
Time	<i>vaqtilvaqte</i>	‘when’	IND	Interruption	<i>kelki</i>	‘when (suddenly)’	IND
Point/Stretch	<i>tāto</i>	‘as soon/ long as’	IND	End point	<i>tāto</i>	‘until (finally)’	IND
Condition	<i>tāto,</i> <i>tā na-l</i> <i>to na-</i>	‘until’ ‘unless’	SBJ SBJ	Purpose	<i>tāto</i>	‘in order to, so that’	SBJ
Condition	<i>agarlagar</i>	‘if’	SBJ, CND	Exception	<i>magar/</i> <i>magar</i>	‘unless, if not’	SBJ

6.5.2 Pre-matrix clauses

For these, see also the paradigm of relative clauses in Table 8.14.

6.5.2.1 Locative clauses

These are plain relative clauses:

be jā-i ke raft-i (be ān-jā) bar gard!!
ba jo-'e ki raft-i (ba on jo) bar gard!
 to place that go.PT-2s to there IMP.return.2s
 'go back to where you come from!'

6.5.2.2 Temporal clauses

These are mostly introduced by *vaqt-i (ke)/vaqt-e (ki)*. The noun may be substituted by similar time-nouns, such as *mowqe'-i kelzamon-e ki* 'the time, moment when'. Time extension may be indicated by *moddat-i kelmuddat-e ki* 'period, duration'. These may be preceded by prepositions, *dar, az Tldar, az T* 'at T, from T, since'. Punctual events are introduced by *hamin kelhamin ki* 'the very same that, just as', and simultaneous events by *dar hāl-i keldar hol-e ki* 'in the situation that, while'. Generalized events use *har vaqtlhar vaqt* 'each, every time, whenever'.

Posterior and anterior events use the adverbial phrases *ba'd ~ pas az ānlba'd ~ pas az on* 'after from that, after (he did ~ doing . . .)' and *pas ~ piš az ān kelpas ~ peš az on ki* 'before from that, before (he did ~ doing . . .)', which require the subjunctive as they refer to potential actions:

piš az ān ke ān kār-rā be-kon-i, fekr kon!
peš az on ki on kor-ro kun-ī, fikr kun!
 before from that that that work SBJ.do.PR-2s thought IMP.make.PR.2s
 'before you do that, think!';

to vazife-ye xod-at-rā tamām na-kon-i,
to vazifa-i xud-at-ro tayyor na-kun-ī,
 until dutyEZ self-your-DO accomplished not(SBJ.)-make.PR-2s
be kuče na-row!!
ba kuča baromada na-rav!
 to alley out.come.PART not.IMP.go.2s
 'until you've done your homework, don't go outside!'

6.5.2.2a Since-clauses

For the expression of the time since an event is taking or has taken place, Common Persian uses a subject clause construction of the type 'it BE(-COME) X time that':

panj sāl ast (~ mi-šav-ad) ke u-rā na-dide am
panj sol ast (~ me-šav-ad) ki ū-ro na-dide am
 five years COP.3s IPFV.become.PR-3s that he-DO not.see.PART COP.1s
 'I haven't seen him in ~ since ~ for about five years'.

6.5.2.3 Conditional clauses

These are introduced by *agarlagar* 'if', *har gāhlhar goh* 'as soon as, when, if'. In colloquial Tajik speech *agarlagar* may be placed at the end of the protasis, after the verb, similar to Uzbek.

Real conditions require the indicative, followed by the indicative present or imperative:
Present indicative

agar u-rā dust mi-dār-i, komak kon!
agar ū-ro dūst me-dor-i, yori dih!
if he-DO friend IPFV.have.PR-2s help do/give.IMP.2s
'if you really like him, help him!'.

Preterite (perfective aorist)

agar u-rā did-i, be-gul
agar ū-ro did-i, gūy
if he-DO see.PT-2s IMP-2s
'once you see him, tell me'.

Potential conditions require the subjunctive, also followed by the indicative present or imperative:

agar u-rā be-bin-id, salām-e ma-rā be-ras-ān-id!
agar ū-ro bin-ed, salom-i ma-ro ras-on-ed
if her SBJ-see.PR-2p see salām-EZ we-DO SBJ-taketo.PR-2p
'if you see her, give her my regards'.

Counterfactual conditions have the imperfective past (in its modal function) in both clauses:

agar hamrāh-e man mi-raft-id, u-rā mi-did-id!
agar hamroh-i man me-raft-ed, ū-ro me-did-ed
if along-EZ I IPFV.go.PT-2p he-DO IPFV.see.PT-2p
'if you were to come with me, you would see him' or
'if you had come . . . you would have . . .'

Past anterior time may be specified by the use of a pluperfect in its modal function in one clause:

agar ham-rāh-i man raft-e bud-id . . .)
agar ham-roh-i man raft-a bud-ed . . .)
if along-EZ I go.PART be.PT-2p
'if you had gone with me . . .'

In Tajik, logical, or quasi-conditionals (which connote 'since, as appears, it is the case that', etc., use the durative perfect subjunctive (not found in Persian) for present reference:

agar ham-digar-ašon-ro dūst me-došt-a boš-and, digar čī?
if together-their-DO friend IPFV.have.PART (SBJ).be.PR-3p further what
'if they love each other, what of it?'

They use the past subjunctive for past reference in the protasis:

in kitob-ro xond-a boš-ed agar, čaro na-ovard-ed?
this book-DO read.PART (SBJ).be.PR-2p if why not.bring.PT-2p
'if you've read this book, why didn't you bring it?'

6.5.2.4 Concessive clauses

These are introduced by the following phrases, all ‘though, although, despite the fact that’: *agar-čelagar-či*, lit. ‘if, though’; *har čand kelhar čand(-e ki)*, lit. ‘however much that’; *bā (vojud-e) in kelbo vujud-i) in ki*, lit. ‘with the existence of this that’. They usually take the present or past subjunctive, according to time reference. The matrix clause may be introduced by *valilvale*, *ammālammo*, *likenllekin* ‘but, still’:

agar če qablan šomā-rā na-did-e bāš-am,
agar či peš-tar šumo-ro na-did-a boš-am,
 although earlier you.p-DO not.see.PART (SBJ).be.PR-1s
 (vali) *esm-e šomā be man āšnā budl*
 (vale) *nom-i šumo ba man ma'lum bud*
 (yet) name-EZ you.p to me familiar/clear be.PT.3s
 ‘even though I hadn’t seen you before, still your name was familiar to me’.

Tajik post-clausal *-ham* ‘also, even’

Frequent and peculiar to Tajik is use of the postposed particle *ham* as a concessive:

havo xunuk na-boš-ad ham,
 weather cold not(SBJ).be.PR-3s also,
barf bo dona-ho-i kalon-kalon-i laklakī me-borid
 snow with piece.p-EZ big-big-EZ fluffy IPFV.rain.PT.3s
 ‘although the weather was not cold, the snow was falling in large, fluffy flakes’;

note the generalized present subjunctive instead of the perfect subjunctive.

6.5.2.5 Causal clauses

These are introduced by: *čun kelčun (ki)* ‘since, as’; *mādām kelmodom-e ki* ‘while, since, inasmuch’; *barā-ye in kelbaro-i in ki* ‘for the sake of, because’ (see also section 6.5.3.4 Purpose clauses) with tense and aspect according to sense:

mādām ke ū āmad, be u be-gu-id ke in-jā tu bi-ā-yadl
modom-e ki ū omad, ba ū gū-ed ki injo dar-o-yad
 while EZ KE he come.PT.3s to he say.IMP-2p that here in SBJ.come.PR-3s
 ‘since he’s arrived, tell him to come in’.

6.5.3 Post-matrix clauses

6.5.3.1 Causal-explanatory clauses

These are introduced by: *čun kelčun (ki)*, lit. ‘because (that)’; *zirā/zero (kelki)*, lit. ‘for this (that)’; *čerā kelčaro ki*, lit. ‘what for that’; *banā bar in kelbino bar in ki*, lit. ‘built on this that’, all ‘because, therefore’:

ketāb-etān-rā be u na-dād-am, zirā ke u-rā na-did-aml
kitob-aton-ro ba ū na-dod-am, zero ki ū-ro na-did-am
 your book-DO to he not.give.PT-1s because he-DO not.see.PT-1s
 ‘I didn’t give him your book, because I didn’t see him’.

6.5.3.2 Temporal clauses

Post-matrix temporal clauses are syntactically focused clauses:

mi-xāst-am xāb be-rav-am ke kas-i dar zadl
me-xost-am xob rav-am ki kas-e dar zad
 IPFV-want.PT sleep SBJ-go.PR-1s that someone door knock.PT.3s
 'I was about to go to bed when someone knocked at the door'

6.5.3.2a Subordinator switch

A distinct syntactic operation involves the switch of the subordinator between two clauses. This is found typically in subsets where an ongoing or existing situation, or an event that is not yet completed, is interrupted by another event. When the latter is in post-matrix position, it is introduced by a simple *kelki*:

dar ān-jā nešaste bud-am, u dar āmadl
dar on jo nišasta bud-am, ū daromad
 I there sit.PART be.PT-1s he come.in.PT.3s
 'I was sitting there, and he came in'

>>

vaqt-i dar ān-jā nešaste bud-am, u dar āmadl
vaqt-e dar on jo nišasta bud-am, ū daromad
 unmarked, 'while I was sitting there . . .'

>>

dar ān-jā nešaste bud-am, ke u dar āmadl
dar on jo nišasta bud-am, ki ū daromad
 marked: '. . ., when he suddenly came in'

6.5.3.3 Result clauses

These are introduced by *tālto*:

ān-jā nešaste bud-am, belāxare āmad
on jo nišasta bud-am, biloxara omad
 there sit.PART be.PT-1s finally come.PT.3s
 'I sat there, until he finally arrived'

6.5.3.4 Purpose clauses

These are introduced by the following, all 'so that, in order to, to (do)': *kelki*, *tā kelto* (*ki*), *barā-ye in kelbaro-i in kilbaro-i on ki*. Of these, *tālto* is mainly literary. All require the present subjunctive, irrespective of the matrix tense:

āmad-am, ke ~ tā u-rā be-bin-aml
omad-am, ki ~ to ū-ro bin-am
 I came that ~ till he-DO SBJ-see.PR-1s
 'I came to see him';

čandān kon, ke jā bāš-ad
čandon kun, ki jo boš-ad
 thus do.IMP.2s that place (SBJ).be.PR-3s
 'make it so, that there is space'

Colloquially, the purpose clause may be juxtaposed without a conjunction:

piš-e mehmān-at bo-row, az to nārāhat na-šav-ad
peš-i melmon-at rav, az to xafa na-šav-ad
 front-EZ guest-your.s IMP.go.2s from you.s offended not.(SBJ.)-become.PR-2s
 ‘go over to your guest, so he won’t be offended’. See also section 6.4.2 Embedded
 speech).

6.5.3.5 Consecutive clauses

These are introduced by *tā (in ke)to (in ki)* ‘until, as far as’, *čonān . . . kelčunon . . . ki, be qadr-i . . . kel (ham-)in qadar . . . ki* ‘so (much) that’, *towr-i ke* ‘in a manner that’, so’:

raft-o raft, tā be jangal-i rasidl
raft-u raft, to ba jangal-e rasid
 go.PT.3s-and go.PT.3s till to forest-INDEF reach.PT.3s
 ‘on he went until he reached a forest’;

u čonān tond pašme mi-čīn-ad, ke hame hayrān mi-mān-andl
ū čunon tez paxta me-čīn-ad, ki hama hayron me-mon-and
 she so fast cotton IPFV.pick.PR-3s that all astounded IPFV.remain.PR-3p
 ‘she picks cotton so fast that everyone is astounded’.

6.5.3.6 Comparative clauses

In Persian, these are introduced by *tālto*. Tajik has instead a single-sentence infinitival construction, *az INF dida* ‘seen from (his doing)’:

u bištar mi-xāb-ad, tā kār mi-kon-ad
 he more IPFV.sleep.PR-3s than work IPFV.do.PR-3s
 ‘he sleeps more than he works’;

vs. Tajik

ū az kor kardan did-a beštar xob me-rav-ad
 he from work do.INF see.PART more sleep IMP.go.PR-3s.

6.5.3.7 Conditional exception clauses

While *agarlagar* may introduce post-matrix conditional clauses, exceptional conditions are introduced by its negation, *magarlmagar* ‘unless’:

man ne-mi-rav-am magar to ham bi-ā-yi
man na-me-rav-am magar tu ham ā-yi
 I not.IPFV.go.PR-1s unless you.s also SBJ.come.PR-2s
 ‘I will not go, unless you come, too’.

6.5.4 Tajik Turkic-type nominal clauses

As shown in some of the discussions above, the final clausal constructions represent only part of Tajik’s syntactic repertory. Most of them can be recast in the Turkic mold as a single sentence containing one or more non-finite verb forms representing nominalized VPs.

The means of nominalization is the nominal EZ which has greatly extended its range of application (arguably under the influence of Uzbek), by the conversion of the verb phrase of finite subordinate clauses to non-finite verb forms, especially (1) participle II in *-agī* and (2) infinitives (gerunds), thereby condensing into a single sentence what would appear more naturally in Persian or English as a matrix-plus-subordinate clause.

6.5.4.1 Participial relative clauses in *-agī*

(1) In the basic pattern, the augmented past participle in *-agī* (which is either active or passive in voice, and may be tensed analogously with the complex tenses which it helps to form, may be used in EZ and other constructions to perform the functions of a relative clause. Note that these participles retain the modal-aspectual distinction of perfective progressive, and conjectural (see section 6.3.7.1 above):

Perfective form	Progressive form
<i>kitob-i</i> [<i>man ovard-agī</i>] book-EZ [I having brought] 'the book that I brought';	<i>in duxtar-i</i> [<i>kitob xond-a istod-ag-ī</i>] this girl-EZ [book reading-PROG] 'the girl who is reading the book';
Conjectural form	
<i>zan-i</i> [<i>sar me-dod-agī</i>] wife-EZ [to divorce-PR-CONJECT] 'a wife [to divorce]' (<i>sar dod-an</i> 'let go').	

(2) Inversions and topicalized relative constructions

The nominalization of a relative clause ranges from retention of the agent and word order, to the replacement of the agent pronoun to a pronominal clitic, to the fully Turkicized inversion:

HEAD-EZ [Clause] > [Clause] HEAD	
<i>kitob-i</i> [<i>man ovard-agī</i>]	book-EZ [I having brought] 'the book that I brought'.
<i>kitob-i</i> [<i>ovard-agi-am</i>]	book-EZ [my having brought-]
[<i>ovard-agi-am</i>] <i>kitob</i>	[my having brought] book.

6.5.4.2 Infinitival clauses

The infinitive, often in combination with conjunct verb forms, can participate in quite complex, nested NPs as sentential complements.

6.5.4.2a Object clauses

Infinitival object clauses are marked by the direct object marker *-ro* and precede the main clause:

(1) 'how do you know that doing this will not be worthwhile?'
az kujo me-don-i [(ki) az in kor foida-yaš na-bar-o-yad] >>
[az in kor foida na-bar-omad-an-aš]-ro az kujo me-don-ī?

from where do you know [that from this work benefit-its may not come out]
 [from this work benefit not-issuing-its]-DO from where do you know?

(2) ‘of course one needn’t tell anyone who this girl is’.
albatta darkor nest [ki gū-yem] [in doxtar kī-st] >>
albatta [kī bud-an-i in duxtar] -ro guftan] darkor nest

of course necessary not is [that we say [this girl who is]]
 of course [[who being-EZ this girl]-DO saying] necessary not is

(3) ‘this excessive wheat isn’t worth for us to take back’.
in gandum-i ziyodatī na-me-arz-ad [boz gašt-a girift-a bar-em] >>
in gandum-i ziyodatī [ba [boz gašt-a girift-a burdan]] na-me-arz-ad

this wheat excessive is not worth [having taken back that we carry]
 this wheat excessive to [returned CONJUNCT taking] is not worth
 (*ba . . . arzid-an* ‘be worthy of, to’).

6.5.4.2b Adverbial clauses

In adverbial infinitival clauses, the infinitival clause is headed by a nominal replacing the conjunction, here with the example of a temporal clause:

‘when the lamp was being taken out of the room, its glass broke’;
dar vaqt-i [ki lamp-ra az xona bar-ovard-and]] šiša-aš šikast >>
[dar vaqt-i [az xona bar-ovardan-i lamp-ra]] šiša-aš šikast

at the time [that the lamp-DO from the room they took out] glass-its broke
 [at the time-EZ [from the room taking out-EZ the lamp-DO]] glass-its broke

6.5.4.3 Embedded infinitival clauses

With the concomitant gapping of an anaphoric independent or enclitic pronoun, infinitival clauses may further be embedded by an EZ construction:

‘we heard the sound of that man’s footsteps (*po*) as he was coming downstairs’.
[dar vaqt-e ki on kas poin me-furomad] sado-i po-i on kas-ro šiunid-em >>
[[dar vaqt-e poin furomadan-i on kas] sado-i po-i; vaγ]-ro šiunid-em
[sado-i po-i [poin furomadan-i on kas]]-ro šiunid-em

[[at time **when** that one **was coming down**] sound-EZ foot-EZ **that one**]-DO we heard
 [at **time-EZ coming down-EZ that one**] sound-EZ foot-EZ **him**-DO we heard
 [sound-EZ foot -EZ [**coming down-EZ that one**]]-DO we heard
 ‘we heard the sound of the footsteps of that one’s coming downstairs’.

Note the EZ embedding of infinitival temporal clause and the gapping of the anaphoric pronoun.

6.5.4.4 Persian conjunct participle

Persian allows for the participialization of finite verbs in sequences of clauses by reducing finite verbs to participles in all clauses except the last (known as *hazf-e fe’l* ‘ellipsis of tense-mood-aspect and person marking’). This strategy rarely includes subordinate clauses. It is well known from classical texts, and in contemporary Persian is confined to literary use. The following is an (abbreviated) example from an Iranian journal about old

dialect poetry (*fahlaviyāt*) found in various sources (M. Rezāyati Kische Khāle, *Guy-eššenāsil/Dialectology* 2.1 (Feb. 2006): 129).

Doktor Ali Ašraf Sādeqi, ke sāl-hā-st tashih-e fahlaviyāt . . . rā
Dr. Ali Ashraf Sadeqi who years is (that) emendation-EZ fahlavi poetry-DO

vajh-e hemmat-e xod qarār dāde,
direction-EZ effort-EZ-self own resolve having given

axiran . . . in yāzdah fahlavi-rā . . . tashih karde
recently this eleven fahlavis-D correction having made

va dombāl-e har do-beyti tahqiq-i mostowfi afzude
and after-EZ each two-liner research-INDEF comprehensive having added

va be bahs-i kolli pardāxte ast
and to discussion-INDEF general has undertaken

‘Dr. Ali Ashraf Sadeqi, who for many years has made the emendation of the *fahlaviyāt* the focus of his scholarly work, has recently not only corrected these eleven *fahlavis*, but has also been engaged in extensive discussions of the problems involved in general.’

7 LEXIS AND SOCIOLINGUISTICS

7.1 Word formation

The basic procedures of derivation and compounding are the same in Tajik and Persian. However, as the result of a degree of cultural separation during the last five centuries, and a flurry of conscious language planning in Tajik over the past 75 years, there are considerable differences of detail.

7.1.1 Homonymy and conversion

There is a fair amount of tolerance for homonymy, which is disambiguated syntactically. A word such as *xāb/xob* may be a noun with two distinct meanings, ‘sleep’ and ‘dream’; it may additionally function as a predicative adjective: *xāb ast/xob ast* ‘he is asleep’. The noun *ārāmlorom* ‘quiet, calm, tranquility’ may also be the adjective (both predicative and attributive) ‘quiet, calm, tranquil’.

Conversion without morphological change occurs most readily from adjective to noun status: *mard-e javān/mard-i javon* ‘young man’, *yek javān/yak javon* ‘a youth’, *javān-ān/javon-on* ‘young people, the young’.

7.1.2 Suffixation

Suffixation is the principal means of lexical derivation; more than 40 nominal suffixes may be identified. The following list is limited to those regarded as most productive:

-ilī (-g-il-g-ī after the vowel -a-): forms abstract nouns of quality from adjectives and type nouns, of activity from agentives, etc.: *pir-ilpir-ī* ‘old age’ < *pir/pir* ‘old; old man’; *zende-g-ilzinda-g-ī* ‘life’ < *zendelzinda* ‘alive, living’. Such nouns readily evolve concrete senses, as *širin-ilširin-ī* ‘sweets, candy (sweetness)’ < *širin/širin* ‘sweet’.

-kārīkōr, -garī-gar, -čīl-čī: form agent nouns and agentive adjectives from nouns of entity or activity (occasionally, adjectives): *xeyānat-kārīxīyonat-kōr* ‘traitor, treacherous’ < *xeyānat/xīyonat* ‘betrayal’; *kār-garīkōr-gar* ‘(manual) worker’ < *kārīkōr* ‘work’; *hīle-garīhīla-gar* ‘deceitful, cunning’ < *hīle/hīla* ‘trick, deception’; *xedmat-čīlxīzmat-čī* ‘employee, clerical worker’ < *xedmat/xīzmat* ‘service, employment’; Tajik *a’lo-čī* ‘honor student’ < *a’lo* ‘superior, excellent’.

-akl-ak:

(a) added to entity nouns and NPs, forms diminutives, often endearing or derogatory in connotation: *pesar-akl/pisar-ak* ‘laddie, kid’ < *pesar/pisar* ‘boy’; Taj. *modar-i mehrubon-ak* ‘mommy dear’ < *modar-i mehrubon* (EZ phrase) ‘kind mother’; Tajik has an intensive form, *-akak*: *kam-akak* ‘a pinch, spot’ < *kam* ‘little (in quantity)’;

(b) added to verbal agentives, it forms concrete instrumentives: Taj. *ob-poš-ak* ‘watering can’ < *ob* ‘water’ and Stem I of *pošid-an* ‘spray’; *band-ak* ‘pin, clasp’ < Stem I of *bast-an* ‘fasten’.

-čel-ča: forms diminutives from nouns and adjectives (neutral or endearing): *ketāb-čel kitob-ča* ‘booklet’ < *ketāb/kītib* ‘book’; Taj. *duxtar-ča* ‘lassie’ < *duxtar* ‘girl’; Taj. *safed-ča* ‘whitish’ < *safed* ‘white’.

-el-a: a nominal suffix from several ultimate sources, still productive in some functions (metonymic or specifying): *dast-eldast-a* ‘handle’ < *dast* ‘hand’; *xandelxanda* ‘laugh(ter)’ < *xand-id-an/xandid-an* ‘laugh’; *geryelgiry-a* ‘weeping’ < *gerist-an/girist-an* ‘weep’. Numerical expressions are made adjectival or adverbial: *do-ru-yeldu-rū-ya* ‘two-faceted, two-faced, hypocritical’ < *doldu* ‘two’, *rulrū(y)* ‘face’; *panj-sāl-elpanj-sol-a* ‘of five years (adj.)’, a five-year-old, five-year plan’ < *panj/panj* ‘five’, *sol* ‘year’.

-stānlston (after a consonant, *estānl-iston*) forms locative nouns, esp. names of countries: Pers. *kudakestan* ‘kindergarden’ < *kudak* ‘child’, *Baluchestānl/Balochiston*, Baluchistan’.

The following suffixes form adjectives from nouns:

-il-ī (after a vowel, *-g-il-g-ī* or *-v-il-v-ī*): the most general and neutral relative adjective, originating in both Middle Persian *-īk* and Arabic *-īyy*, and readily substantivized: *Irān-il’eron-ī* ‘Iranian, an Iranian’; *xāne-g-ilxona-g-ī* ‘domestic, household’; *šowra-v-ilsovet-ī* ‘deliberative; Soviet’ < *šowrālšavro* ‘council’; Taj. *partiya-v-ī* ‘(of the) Party’ (*billet-i partiya-v-ī* ‘Party card’).

Taj *-angī* (after a vowel, *-ngī*): forms relative adjectives, especially in colloquial Tajik: *dina-ngī* ‘yesterday’s’ < *dina* ‘yesterday’.

-ānel-ona (after a vowel, *-g-ānel-g-ona*) ‘characteristic of ~’: forms adjectives applied to non-humans, frequently used as nouns or adverbs: *āqel-āneloqil-ona* ‘intelligent, sensible’ < *āqelloqil* ‘intelligent’ (of a person); e.g. *kār-e āqel-ānelkor-i oqil-ona* ‘sensible action’; *bačče-g-ānelbača-g-ona* ‘children’s; childhood ~; childlike, childish’ < *baččelbača* ‘child’; *māh-ānelmoh-ona* ‘monthly; salary’ (see also section 3.1.5 Adverbs).

-nākl-nok ‘possessing the quality of ~’, *xatar-nāklxatar-nok* ‘dangerous’ < *xatar/xatar* ‘danger’: a very productive suffix in Tajik, generating adjectives or nouns from intangibles: *foida-nok* ‘useful, advantageous’ < *foida* ‘benefit’; *xarakter-nok* ‘characteristic, specific’ < *xarakter* ‘character(istic)’; *sado-nok* ‘vowel’ < *sado* ‘sound, voice’.

7.1.3 Prefixation

Productive nominal prefixes (mostly identical with prepositions) which form adjectives and related nouns are:

bālbo- ‘with’: *bā-este’dādībo-iste’dod* ‘talented’ < *este’dādīste’dod* ‘talent’.

bi-lbe- 'without': *bi-adablbe-adab* 'discourteous, impolite' < *adabladab* 'manners, courtesy'.

nā-lno- 'not, un-': *nā-dān/no-don* 'ignorant, ignoramus' < Stem I of *dānest-an/donist-an* 'know'; *nā-omidlno-umed* 'hopeless, desperate' < *omidlumed* 'hope'.

ham-lham- 'together', a particle connoting sharing, reciprocity or 'cooperation': *ham-rāhlham-roh* 'companion, fellow-traveler' < *rāhlroh* 'road, journey'; *ham-āhanglham-ohang* 'harmonious, consonant' < *āhanglohang* 'melody'.

kam-lkam 'little' (quantity): *kam-zur/kam-zur* 'weak' < *zur* 'strength; Pers. *kam-harf* 'taciturn' < *harf* 'word'.

ser- 'full, sated', rare in Pers.: *sir-āblser-ob* 'saturated' < *āblob* 'water'; Taj. *ser-odam* 'crowded, populous' < *odam* 'person'; *ser-masraf* 'often used, in demand, popular' < *masraf* 'consumption, utilization'.

por-lpur- 'full': *por-mā'nālpur-mā'no* 'meaningful' < *mā'nā* 'meaning' (literary variant of *mā'ni*); *por-šokuhl pur-šukūh* 'splendid' < *šokuhl/šukūh* 'splendour'.

Prepositional EZ constructions (the following all Arabic loanwords) include:

yejyr-e-qānun-i/γayr-i-qonun-i 'illegal' < *qānun/qonun* 'law'.

zedd-elzidd-i- 'against, anti-, counter-': *zedd-e-havā'i/zidd-i-havo-i* 'anti-aircraft' < *havālhavo* 'air'.

ba'd-i- 'after': *ba'd-i jang-i* 'postwar' < *jang* 'war'.

Tajik *to-* 'up to, until': *to-maktab-i* 'pre-school' < *maktab* 'school'; *to-inqilob-i* 'pre-revolutionary' < *inqilob* 'revolution'. Note that this prefix qualifies a ready-formed adjective, as do the preceding three. (This use of *tā* is unknown in Persian, and was probably influenced by identical Russian use of *do-* 'up to, until'.)

7.1.4 Compounding

Compounding is most frequently effected by simple juxtaposition of stem nominals, or a nominal and a verbal radical, modifier before head. Such compounds are of two kinds:

Determinative, where the compound is of the same lexical category as the head (a noun): *hafte-nāmelhafta-noma* 'weekly (periodical)' < *haftelhafta* 'week' + *nāmelnoma* 'document, letter'; *piš-raftlpeš-raft* 'progress' < *piš/peš* 'forward' + *raftlraft*, Stem II of *raft-anlraft-an* 'go'; *dast-nevisldast-navis* 'manuscript' < *dastldast* 'hand' + *nevislnavis*, Stem I of *nevešt-anlnavišt-an* 'write'. Structures of this last type are more commonly agentives or instrumentives: *kafš-duzlmūza-dūz* 'cobbler' < *kafš* 'shoe', *mōza* 'boot' + Stem I of *duxt-anldoxl-an* 'sew'; *bārān-sanjlboron-sanj* 'rain gauge' < *bārānlboron* 'rain' + Stem I of *sanjid-anlsanjid-an* 'measure'.

Possessive, generally where the modifier is an adjective or adverb, and the resulting compound refers to a third entity possessing, or characterized by, the compound quality (fundamentally, an adjective): *čahār-pālčor-po* 'quadruped, animal' < *čahār/lčor* 'four' + *pālpo* 'foot, leg'; *seyāh-čašmlsiyoh-čašm* 'black-eyed' < *seyāhlsiyoh* 'black' + *čašmlčašm* 'eye'; *xoš-zabānlxuš-gap* 'well-spoken, courteous' < *xoš/xuš* 'good' + *zabān* 'tongue, language', *gap* 'speech'.

Modifier and head are occasionally reversed: *del-saxtdil-saxt* 'cruel' < *delldil* 'heart' + *saxtl/saxt* 'hard'; this is usual when the modifier is a participle: Pers. *ru-bast-e* 'veiled' < *ru* 'face' and Stem I of *bast-an* 'tie, close', Taj. *rū-toft-a* 'disobedient' < *rū* 'face' + past participle of *toft-an* 'turn (away)'.

There are a few common *copulative* compounds, where two nouns or verb stems are joined by addition of enclitic *-ol-u* 'and' to the first element: *āb-o havālob-u havo*

'weather', lit. 'water-and-air'; *raft-o-āmadraft-u-omad* 'visit(ing), traffic' < Stem II *raft-anraft-an* 'go', and Stem I of *āmad-anlomad-an* 'come'; Taj. also *raft-u-o(y)*.

Stems II of verbs. Adjectives and adverbs are similarly formed by compounding or reduplication of nouns with a connective element: *pey dar pey/pay dar pay* 'continuous(ly)' (*pey/pay* 'track, (in) pursuit, after'); Taj. *xel-ma-xel* 'various, of all sorts' (*xel* 'sort, kind').

Adjectives may be intensified by similar means: *dur-o-derāz/dur-u daroz* 'lengthy' (*dur* 'far', *daroz* 'long'), Taj. *garm-o-garm* 'extremely hot' (*garm* 'warm, hot'), but note Persian EZ construction *garm-e garm*.

By adding prefixes and suffixes to these structures, neologisms can be formed: *havā-peymā-bar* 'aircraft carrier' (*havā-peymā* 'air-plane', *bar-* 'bear, carry').

Tajik can generate neologisms of almost Russian length and complexity, a development that has been encouraged from the Soviet period on. Examples are: *mablay-jūdo-kun-ī* 'appropriation, disbursement of funds' (sum-separate-mak-ing); *avtomobil-kor-kard-a-bar-or-ī* 'automobile production' (auto-work-done-out-bring-ing).

Echoic reduplication is used in several ways, mostly in the colloquial language. Collective or generalized nouns are generated by repetition with a change of initial, to add a rhyming or echoic nonsense-word: *ketāb-metāb/kitob-mitob* 'books and papers'; Pers. *eynak-meynak* 'spectacles, glasses'; Taj. *bača-kača* 'kids and the like' < *bača* 'child'; *oš-poš* 'food and stuff' < *oš* 'food'; *mayda jūyda* 'odds and ends' (*mayda* 'small'; *jūyda, mitob*, etc. have no independent meanings).

7.2 Loanwords

Arabic vocabulary is the oldest and still the largest foreign element in the lexicon. In literary Persian it constitutes about 50 percent, in spoken about 25 percent.

7.2.1 Arabic

In Iran, official organizations, including the Farhangestān-e zabān-e Fārsi, targeted especially this stratum for replacement by native vocabulary during the 1930s and 1940s, and in Tajikistan during the Soviet language reforms of the same period; however, the presence of many everyday Arabic words for which there is no ready Persian equivalent (e.g. *ketāb/kitob* 'book', *havā/havo* 'air, weather') has mitigated the effects of purification.

Some specialized uses of Arabic and Arabicate plural forms continue to be used in Iran and have survived the general Soviet condemnation of archaisms in Tajikistan, since they have been lexicalized and fill a useful niche. Words with the suffixes *-āt/ot*, *-j-āt/j-ot*, Taj. *-v-ot* denote collectivities: *heyvānāt/hayvon-ot* 'animals, fauna', *mive-j-āt/meva-j-ot* 'fruit(s)', Pers. *sabz-ij-āt*, Taj. *sabza-v-ot* 'vegetables'. In Tajik, some plurals have developed a singular meaning: *taškil-ot* 'organization', *hašar-ot* 'insect' (with regular plurals *llarbāb-hā, taškil-ot-ho* and *hašar-ot-ho*).

Arabic "broken plurals" are often lexicalized, with collective or singular meanings: *atrāflatrof* 'environs, neighborhood' (cf. *taraf* 'side, direction'), *a'zāla'zo* 'member' (of an institution), *a'zā-ye badan/a'zo-i badan* 'parts of the body'; Pers. *arbāb* 'landowner', boss (cf. *rabb* 'Lord [God]'); Taj. *talaba* 'student' (cf. *tolib* '(religious) student').

7.2.2 *Turkic*

While Persian has integrated a considerable number of Turkic nominals (including a few of Mongol origin) such as *otāqlutoq* Pers. 'room' vs. Taj. 'house, place' and titles like *āqāl ogo* 'sir, Mr.', the Northern Tajik dialects are replete with pre-Uzbek Turkic and Uzbek vocabulary, even at the level of function words; a number of common loans have entered the literary language and are also widespread in Southern speech, e.g. *boy* 'rich', *tūy* 'wedding, circumcision celebration', *yaroq* 'weapon', *yordam* 'help', *qišloq* 'village' (Pers. *qešlāq* 'winter quarters'), and several kinship terms such as *uka* 'younger brother', *yunga* 'sister-in-law' (see section 7.4.3).

7.2.3 *Russian*

In Tajik, the large Russian component pervades the speech of the cities, reinforced by code-switching and official communication (forms, notices, interaction in banks and post offices, etc.). The written language may be highly Russianized in political, economic and bureaucratic documentation or journalism, while in other cultural contexts it will display much more Persian vocabulary. At the height of the language reform movement in 1989–90, conservatives complained that unbridled substitution of Persianisms for established Russian loans (*doniš-goh* for *universitet* 'university', *havo paymo* for *samolyot* 'airplane', etc.) was making the language incomprehensible to the man in the street.

Abbreviations and acronyms are a feature of Soviet Russian bureaucratic language that have readily been adopted (in Russian) and adapted (to Tajik), e.g. *VABK*, for *Viloyat-i avtonom-i Badaxšon-i Kūh-ī* 'the Gorno-Badakhshan Autonomous Region'.

7.3 Lexical distribution, Persian–Tajik

In the everyday Persian and Perso-Arabic vocabulary, even frequently-used words have been differentially distributed between Iran and Central Asia. Some Tajik words and expressions are either completely foreign to Iranians, or perceived as archaic or literary; the meanings of some common words have shifted considerably. In the following selection of shibboleths, words in parenthesis are Standard Persian equivalents (which sometimes coexist as variants in Tajik): *dina* (*diruz*) 'yesterday', *pagoh* (*fardā*) 'tomorrow', *be-goh* (*'asr*) 'evening; yesterday', *tira-moh* (*pā'iz*) 'autumn', *daryo* (*rud, rud-xāne*) 'river', *bahr* (*daryā*) 'sea', *paxta* (*panbe*) 'cotton', *tireza* (*panjare*) 'window', *xel* (*jur*) 'sort, kind', *tayyor* (*hāzer*) 'ready', *hozir* (*halā*) 'now', *vazn-in* (*sang-in*) 'heavy', *sang-in* (*sang-i*) '(of) stone', *kalon* (*bozorg*) 'big, great, old', *mayda* (*xōrd*) 'tiny; small change', *xurd* (*kuček*) 'small, little, young', *kampir* (*pir-e zan*) 'old woman', *nayz*, *nek* (*xub, qašang*) 'good, nice', *ganda* (*xarāb*) 'bad', *kasal* (*mariz, bimār*), 'sick', *mond-a* (*xaste*) 'tired', *pok* (*tamiz*) 'clean', *iflos* (*kasif*) 'dirty', *aftid-an* (*oftād-an*) 'fall', *šiš-t-an* (*nešast-an*) 'sit', *mond-an* (*gozāšt-an*) 'let, put', *partoft-an* (*andāxt-an*) 'throw', *yundošt-an* (*jam kard-an*) 'gather'.

This applies even to quite recent neologisms: Taj. *madaniyat*, Pers. *farhang* 'culture'; Taj. *ittifoq*, Pers. *ettehād* '(labor) union'. Many of the Tajik shibboleths are also common to Persian of Afghanistan (Kaboli, Dari).

Nominal compounds and conjunct verbs, the productivity of causatives, and adjectival formatives such as *-nok* and *-g-ī* constitute another large area of lexical difference with

Standard Persian. Compound neologisms, even if formed from native lexical stock and not calqued on Russian (for Tajik) or French (for Persian) may not coincide, either lexically or structurally: compare Taj. *bayn-al-xalq-ī*, and Pers. *beyn-al-melal-i* ‘international’, where the difference lies merely in the selection of the central (Arabic) lexeme; Taj. *yaroq-partoi*, Pers. *xal’-e selāh* ‘disarmament’, where the Uzbek-Tajik compound is a modern calque on the Arabicate compound, lit. ‘casting off arms’, still current in Standard Persian; Taj. *zioī*, Pers. *rowšan fekr* ‘(liberal) intellectual’, in which the underlying idea of ‘enlightened thinker’ is expressed in an Arabic-Persian derivative and a Persian-Arabic compound respectively.

7.4 Sociolinguistic aspects

7.4.1 Registers and *ta’ārof*

A crucial strategy of spoken and written communication is *ta’ārof* ‘formal mutual recognition’. It aims, in addition to its illocutionary function such as making a request, at the perlocutionary effect of self-presentation and respect, inferior and superior, respectively. It is one of the poles of the “communicative contexts in Iranian culture which are marked in terms of two continua: one of personal and communicative intimacy, from ‘inside’ (*bāten*) to ‘outside’ (*zāher*), and another of social hierarchy, from contexts indicating hierarchical relationships to those indicating equality” (Beeman 1988: Abstract). Linguistically, it is marked by a highly conventionalized phraseology and address terminology depending on the context, accompanied by appropriate metalinguistic behavior and gestures, including proper pacing and packaging of objectives, and is applied by speakers from the lowest to the highest speech registers.

Personal pronouns and endings are in the plural (see section 3.1.3.1a), *šomā* VB-*id* ‘you.p are . . .’ for the addressee, *išān* VB-*and* ‘they are . . .’ for a third party, while the speaker may use *mā* Vb-*im*, ‘we are . . .’, reflecting the role of the speaker as member of a social group, rather than as an individual. Lexically most notable is the use of *farmudan* ‘to order, command’ for any superior’s verb of action such as *be-farmā-id* ‘please (DO)’. Typical is the use of light verb constructions for simple verbs: superior raising includes *tašrif* ‘(your) honor’ in *tašrif āvordan* ‘bring honor’ = *āmadan* ‘come’, *tašrif bordan* ‘take honor’ = *raftan* ‘go, leave’, *tašrif dāštan* ‘have, keep honor’ = *budan* ‘be there, present’. Inferior lowering includes (*be*) *xedmat-e* ‘to the service of’ = *be, piš-e* ‘to, to the presence of’, e.g. *xedmat residan* ‘arrive at service’ = *āmadan* ‘come to’; *arz kardan* ‘make presentation’ = *goftan* ‘say, remark’, often in subjunctive, *arz kon-am* ‘let me remark’, and may use an oath-like address such as *qorbān* ‘your honor’ < *qorbān-e šomā be-rav-am* ~ *be-šav-am* ‘(I am ready) to be sacrificed for you’. Superior is addressed as *janāb(-e) āliljanob-i oli* ‘Sir’, ‘gentleman, aristocrat’, lit. ‘high side’, vs. *in jāneblin jonib* ‘I’, lit. ‘this side’, and *sarkār/sarkor* ‘Sir’ or ‘Madam’, lit. ‘leader, head of work’, vs. *bandelbanda* ‘I’, lit. ‘slave’ little used today, and others. All these are contextually selected from a paradigm of options for each, ranging from higher to lower.

Inferior typically introduces requests with one or a series of phrases, or their substitutes, rarely up to seven, and usually in this relative sequence: *janāb āqā-ye X, mā’zerat mi-xāh-am, agar ejāze be-deh-id, xāheš mi-kon-am, agar momken bāš-ad, lotf-an, bi zahmat, be-farmā-id . . .* ‘your excellency Mr. X, I beg your pardon, if you permit me, I ask of you, if it be possible, kindly, (if it is) no trouble, please . . .’

7.4.2 Kinship terms

Persian kinship terms for some basic consanguineal and affinal relationships are single terms, but others are composite descriptive terms expressed by EZ constructions.

In Tajik, alternate or additional terms (some of Uzbek origin) include: *oča* ‘mother’; *ako*, *aka* ‘elder brother’, *dodar*, *uka* ‘(younger) brother’; *apa*, *aya* ‘elder sister’; *yanga* ‘aunt (uncle’s wife), sister-in-law’; *taγoī*~*amulamak* ‘uncle’, *padar-šuy*, *xusur* ‘father-in-law (of wife)’, *modar-šuy*, *xušdoman* ‘mother-in-law (of wife)’.

TABLE 8.18: PERSIAN KINSHIP TERMS

	Male	Female		
-3	<i>jadd-e bozorg</i>	<i>jadde-ye bozorg</i>	G.-great-grandparents	
-2	<i>jadd</i> Ar.	<i>jadde</i> Ar.	Great-grandparents	
-1	<i>pedar(-e) bozorg</i>	<i>mādar(-e) bozorg</i>	Grandparents	
-1	<i>pedar</i>	<i>mādar</i>	Parents	
-1	<i>amu</i> Ar.	<i>ame</i> Ar.	Paternal uncles, aunts	
-0	<i>dā'i</i> Tk.	<i>xāle</i> Ar.	Maternal uncles, aunts	
+1	<i>pesar</i>	<i>doxtar</i>	Children	
+2	<i>nave (navāde, navāse)</i>		Grandchild	
+3	<i>natije</i>		Great-grandchild	
+4	<i>nabire</i>		G.-G.-grandchild	
+5	<i>na-dide</i>		G.-G.-G.-grandchild	
	Male	Female		
	<i>dāmād</i>	<i>arus</i> Ar.	Bridegroom	Bride
			Son-in-law	Daughter-in-law
	<i>šowhar</i>	<i>zan</i>	Husband	Wife
		<i>havu</i>		Co-wife
	<i>pedar, mādar-e šowhar</i>	<i>zan</i>	Parents-in-law	
	<i>barādar, xāhar-e šowhar</i>	<i>zan</i>	Siblings-in-law	
	<i>zan-e barādar</i>		Brother’s wife	
	<i>šowhar-e</i>	<i>xāhar</i>	Sister’s husband	
	<i>bājnāq</i> Tk.		Husbands of sisters	
		<i>jāri</i>	Wives of husbands	
	<i>zan-e amu, dāi</i>		Uncles, aunts-in-law	
	<i>šowhar-e</i>	<i>ame, xāle</i>	Uncles, aunts-in-law	

7.4.3 Modes of address

Between members of the same peer and solidary groups (age, sex, occupation) address is essentially informal, given names being widely used, often with the endearing suffix *-jānljon* (*Ahmad-jānlAhmad-jon*, *Zohre-jānlZuhra-jon*).

Family members traditionally address one another in kinship terms rather than by their given name, again often adding *-jānljon*: *bābālbobo* ‘grandfather’, Pers. also ‘daddy’, coll. ‘fellow’; *bibilbibī* ‘grandmother’ (also mistress of house, etc.); *dada* ‘father’, Pers. also ‘daddy, mamma’, Pers. *āqā-jān* ‘daddy’. Most of these may be used figuratively to address unrelated acquaintances and strangers of the appropriate age and sex.

During the Soviet period in Tajikistan, Russian and Soviet modes of address became fashionable, including *rafiq* ‘comrade’ preceding the surname and the Russian-style name and patronymic, as *Ahmad Ibrohim-ovič*. Russian-style surnames were also widely adopted (*Yusup-ov*, *Nazir-ova*). Since the 1980s many Tajiks have re-Persianized their

surnames by means of relative or patronymic suffixes such as *-ī* and *-zoda*. In the change from *Yusupov* to *Yusuf-zoda* or *Yusuf-ī* the substitution of *f* for *p* further represents re-Persianization of an Uzbek reflex of the Perso-Arabic *Yusuf*.

8 DIACHRONY AND DIALECTOLOGY

8.1 Phases

The grammatical changes and typological shifts in the evolution of New Persian over fourteen centuries has been substantial, though shifts were mostly gradual and individual rather than bundled together. They involved temporal, social, as well as geographical parameters. Overall, between Middle Persian and contemporary Persian, the decisive monument for the history of Persian was the *Shahname* by the eleventh-century poet Ferdowsi of Tus in Khorasan.

In particular, recent detailed investigations of textual occurrence and relative frequency of features and their typology have led to clearer insights into the historical and geographical dynamics of their loss and innovation. In overview, the following table, based on Paul (2000b), shows the successive phases of innovation and relative consolidation, in correlation with related dynasties:

700–900	Transition from Middle Persian
900–1100	Earliest NP, Ghaznavids; <i>Shahname</i> of Ferdowsi of Tus
1100–1300	ENP, Mongols: in twelfth century: relative dialectal homogeneity and standardization, Persian international language in Eastern Caliphate
1300–1600	“Classical Persian”, Turko-Mongols to Safavids: considerable standard reached; earlier features may occur in emulation of “Classical” style – thereafter slower shifts
1600–1900	Early Modern Persian, Qajars
1900–present	Modern Persian, Pahlavids, Islamic Republic: rapid change in written register, influenced by spoken Tehrani Persian

The complex socio-linguistic dynamics of the development during the early phases of Persian are detailed in Perry (2009).

8.2 Comparative dialectology

The common Early New Persian vowel system had six members: *i ī ē*, *u ū ō*, *a ā*. There occurred shifts in the three main varieties of Persian: In Iranian Persian, the vowel system was restructured by the turn of the seventeenth century: The tense long mid vowels were raised, while the lax short high vowels were lowered: *ē ō*, *i u* > *ī ū*, *e o*: *bē* > *bi* ‘without’, *ō* > *u* ‘he, she’; *ki* > *ke* ‘which, that’, *-u* > *-o* ‘and’. More recently, word-final *-a* was raised to *-e*, but *-a* is still retained before enclitics in the Spoken Modern Standard: *ba* > *be* ‘to’, *hame-y-aš-rā*, but coll. *hama-š-o* ‘all of it’. In Tajik Persian, the restructuring involved the loss of length distinction under Turkic pressure, whereby the inherited short and long high vowels merged, and long *ā* was rounded to *o*. The fewest changes are found in Afghan Persian, where the lax high short vowels were lowered: *i o* > *e o*, and *ā* > *â*. The development of the vowels is shown in the diagram given overleaf (see also Table 8.2).

Tajik		i	e		u	û		a	o
		↑	↑		↑	↑		↑	↑
Early NP		i	ī	ē	u	ū	ō	a	ā
		↓	↓	↓	↓	↓	↓	↓	↓
Afghan		e	ī	ē	o	ū	ō	a	ā
		↓	↓		↓	↓		↓	↓
Iran		e	.ī		o	ū		a	ā

8.3 Innovative constructions, summary

The development in the morphosyntax of these three varieties is the history of increasing differentiation, particularly in their verb systems through the development of new constructions. These include the progressive Aktionsart and the evidential mode. For Afghan Persian see also Farhadi and Perry 2009).

8.3.1 Progressive

Persian uses *dāštan* in its basic meaning ‘keep, hold’, where both auxiliary and dependent verb are finite. Tajik and Afghan use conjunct verb constructions with *istoda* ‘standing’ and *raftan* ‘go’, respectively: Pers. *dār-ad ketāb mi-xān-ad*, Taj. [*kitob-ro xond-a*] *istoda ast*, Afgh. [*ketāb-ra xānd-a*] *mē-rav-ad* ‘he is reading the book’.

8.3.2 Uncertainty constructions

Unlike Persian, Tajik and Afghan have developed modal constructions expressing uncertainty. In Tajik, the conjectural is based on participle II in *-agī* in its gerundive sense. In Afghan, the dubitative is based on the lexical marker *xāt*, a generalized form of 3s *xāh-ad* ‘it will probably be’ < *xāh-* ‘want, will’: [*zad-a*] *xāt bud-om* ‘I might hit’. These contrast with the literary definite future construction with the short infinitive, *xāh-ad raft* ‘he will go’, inherited from Early New Persian (for more detail, see Perry 2002).

8.3.3 Evidentiality

Evidentiality, while found in both Iranian and Afghan Persian, is fully developed only in Tajik, again clearly by interference from Turkic where it is expressed by the single marker *emiš*. Just as the verb forms of Turkic are mostly based on participles, so in Tajik one finds the development of participial formations with so-called converbs, where the participial main verb is followed by a varied set of verbs whose meaning is generalized to express various Aktionsarten.

8.4 Isoglosses West vs. East

8.4.1 Conjunct constructions and complementizer

Morphosyntactically, the innovative progressive and evidential constructions, and the uncertainty constructions in Tajik and Afghan Persian, differ drastically: they are

double-finite construction in Persian, and other western Persian varieties, but are nominalized conjunct verb constructions in Afghan and Tajik Persian. The earliest example of such Eastern conjunct constructions, which was integrated into common Persian, is the passive conjunct construction with *šodan* ‘go’: *ān kušta šud* ‘that one was killed’. These features evidence the membership of Tajik in a vast isoglossic area, to which belong not only Turkic, but also Hindi/Urdu (Windfuhr 2006: 279–281).

8.4.2 Persian homoglossia

In Tajikistan, the linguistic situation is bilingual, and the relationship between the literary standard and local spoken Tajik registers is locally homoglossic rather than diglossic in the sense of two distinct grammatical inventories.

Similarly, in Iran all registers share the same morphological and syntactical features and rules, though high and low differ in complexity. Phonologically, the lower register is characterized by the loss of a number of phonological distinctions which however involve only some 15 merger rules leading predictably from standard to low. The main distinction is lexical and phraseological. There are only the following shibboleths of exclusive use: high *niz* ‘also’ vs. shared *ham* ‘also’; *xiš* ‘oneself, own’ vs. shared *xod* ‘self, own’; low *vāse* ~ *vāsiye* ‘for’ vs. shared *barā-ye* ‘for (the sake of)’. Only the low aspectual particle *hey* ‘continuously’ (reminiscent of ENP *hamē*) does not have a corresponding high pendant.

Perry (2003) has shown that the major coding distinctions between the two registers of Persian amount to a difference of degree rather than kind. Those involve the selection, addition, deletion, reordering, or the re-glossing of elements familiar to both registers, but not exclusive grammatical features. They include:

- (1) Expanded and versatile use of the pronominal enclitics:
 - (a) use of personal enclitics with prepositions: *be-h-eš goft* ‘said to him’;
 - (b) use of 3s enclitic for subject, *raft-eš* ‘he went’;
 - (c) post-verbal position of the enclitic direct object: *bord-am-āš* ‘I took it’.
- (2) The double-finite progressive construction with *dāštan* ‘keep, hold’.
- (3) Violations of SOV: *raft bāzār* ‘he went to the market’, *na-did-am bābā-t-o* ‘I didn’t see your dad’.
- (4) The “idiomatic” use of *ke*.
- (5) The omission of the locative prepositions *be* ‘to’ and *dar* ‘in(to)’.
- (6) The nominal referential suffix *-e*.

However, all of these are to be found in earlier styles of Persian, notably in Classical poetry.

8.4.3 Topical references to Chapter 2

See also the sections in Chapter 2 **Dialectology and Topics** where Persian serves as the example for **the linguistic evolution**: the Iranian verbal quincunx system (section 3.4); **marking of aspect** (section 3.5); noun phrase and syntactic arguments, definiteness (section 4.1), genericity (section 4.1.2), and differential marking of the direct object (section 4.4); and clause complementation (section 4.5).

9 SAMPLE TEXTS

9.1 Persian

The following are the first six sentences of a report on an archaic custom in the region of Khomeyn, Central Lorestan, performed during the ten coldest days of winter (ca. Jan. 25–Feb. 5). A strong person is chosen to go into the Alvand mountains to assure the return of warmth (Anjavi Shirāzi, Seyyed Abolqāsem (1352) *Jashn-hā va ādāb va mo'taqedāt-e zemestāni* [Winter customs and beliefs], Tehrān). He is referred to as *Korde* 'that Kurd', which clearly reflects not only vague memory, but also its ancient connotation of the term *kord*, "mountain people".

Linguistically noteworthy in this passage, particularly in the initial sentences, is the use of the Persian evidential forms, here reflecting both the researcher's second-hand knowledge and the villagers' fading memory (L = linker in relative heads).

Korde be kuh

'(Sending) the Kurd to the mountain(s)'

Be towr-i ke mardom-e mantaqe-ye Rebāt-e Morād-e Xomeyn
the way-L that people-EZ region-EZ Rebāt-EZ Morād-EZ Xomeyn
'As the people of the region of Rebāt-e Morād in the Khomeyn district'

revāyat mi-kon-and,
narrative IPFV-make.PR-3p
'tell it,'

tā yek-sad-o panjāh sāl piš marsum bud-e
till one-hundred & fifty year(s) before customary be.PART.EV
'they had a custom until some 150 years ago'

hame sāle dar ayyām-e "Korde be kuh" yek nafar-rā
all year.ADJ in days-EZ K. one person-DO
'that during the days of K., one person'

ke besyār qavi va nirumand bud-e, entexāb mi-kard-e and.
that very strong & powerful be.PT.EV.3s choice IPFV-make.PART.EV-COP3p
'they would choose who was very strong and powerful'.

Tamām-e ahl-e mahall be u qazā va pušāk mi-dād-e and
whole-EZ people-EZ place to him food & clothing IPFV-give.PART.EV-COP3p
'All people of the village would give him food and clothing'

va u-rā bā tufang va lebās-e garm va āzuqe-ye dah-ruze
& him-DO with gun & dress-EZ warm & provision-EZ ten-days.ADJ
'and him, with gun, warm clothes and a ten-day provision,'

mojahhaz mi-kard-e va be onvān-e "Amu Kord Ali"
supplied IPFV.PART.EV & by title-EZ A.
'they would outfit, and, as "A."'

be kuh-e Alvand mi-ferestād-e and.
to mountain-EZ Alvand IPFV-send.PART.EV-COP3p
'would send him to Mt. Alvand'.

U in dah ruz-rā be kuh mi-raft
 he these ten days-DO to mountain IPFV-go.PT.3s
 'He would go to the mountain for these ten days,'

va bar mi-gašt-e ast.
 & back IPFV-turn.PART.EV COP.3s
 and return.'

Va asr-e ruz-e dah-om marāsem-i-rā
 & afternoon-EZ day-EZ ten-th ceremonies-INDEF-DO
 'And in the late afternoon of the tenth day, certain ceremonies

be xāter-e bar gaštan-e u dar jelo-ve u anjām mi-dād-e and
 for sake-EZ return.INF-EZ him in front-EZ him performance IPFV-
 give.PART.EV-COP3p
 'they would perform in front of him in appreciation of his return,'

va zemn-e rixtan-e ājil va noql va nabāt be sar-e u,
 & meantime-EZ pour.INF-EZ dried fruit & sweets & candy on head-EZ him
 'and, showering him with dried fruit, sweets, and candy,'

esteqbāl-o pazirā'i-ye garm-i az u be amal mi-āvord-e and.
 welcoming & reception-EZ warm-INDEF from him to action IPFV-
 bring.PART.EV-COP3p
 'they would give him an extensive reception and warm welcome.'

agar ān sāl kam-bārān va bad-i mi-šod-e
 if that year little-rain.ADJ & bad-INDEF IPFV-become.PART.EV
 'But if that year turned out to have little rainfall and a bad crop,'

šaxs-e digar-i-rā barā-ye in kār entexāb mi-kard-e and.
 Person-EZ other-INDEF-DO sake-EZ this work choice IPFV-make.PART.EV-
 COP3p
 'they would choose another person for this task.'

Zemn-an mardom se āš ham barā-ye salāmati-ye Amu Kord Ali mi-paz-and.
 meantime people 3 soup also sake-of health-EZ A. IPFV-cook.PR-3s
 'Also, for the sake of the health of A., people would cook three soups (one before his departure, one during his absence, and one on his return).'

Ham-čonin mo'taqed and ke agar āš na-paz-and
 same-thus convinced are-3p that if soup not.SBJ-cook.PR-3p
 'Also, they believe that if they would, lit. 'do', not cook those soups,'

va sag-hā-ye mahalle ham dar ayyām-e, Korde be kuh zuze be-keš-and
 & dog-p-EZ place also in days-EZ K. howl SBJ-continue.PR-3s
 'and also if he dogs of the village howl during the days of K'

va ow'ow kon-and va mardom kaš-hā-ye xod-rā
 & bark SBJ.make-3p & people shoe-p-EZ own-DO
 'and bark, and if (regarding their shoes) people'

pošt-o-ru na-gozār-and, yek-i az bozorg-ān-e ābādi
 back-&-front not.SBJ-place.PR-3p one-INDEF from elder-p-EZ village

mi-mir-ad.

IPFV-die.PR-3s

‘would, lit. ‘do’, not turn them around, one of the elders of the village will die’.

Pas bāyad āš be-paz-and va agar sag-hā ow’ow kard-and,
thus must soup SBJ-cook.PR-3p & if dog-p bark make.PT-3p

‘Therefore they have to cook the soups; and, once the dogs bark. lit. ‘barked’,

kafš-e xod-rā vārune kon-and va pošt-o-ru be-gozār-and
shoe-EZ own-DO turned (SBJ.)make.PR-3p & back-&-front SBJ-place.PR-3p
‘have to turn their shoes upside down and backwards.’

9.2 Tajik

The following is a paragraph from Sadriddin Ayni’s autobiographical novella *Maktabi kūhna* [The old(-fashioned) school], as published in the Ayni anthology *Aknun navbati qalamast* [Now it is the turn of the pen], Dushanbe, 1978, p. 126. Abbreviations and markings in the interlinear gloss:

A verb in upper case translates a non-finite conjunct form; + = connective in adjectival E phrase; CONJ = non-finite conjunct form; EVID/QUOT = evidential use of perfect tense, here two examples with quotative connotation; INT = intentional construction with future participle; BEN = benefactive verb in conjunct construction; PR-PROG = present progressive; SBJ = subjunctive; IMP = imperative.

The *duo* ‘prayer’ referred to in this text is a pious formula written on scraps of paper, which would be immersed in water, where the ink would dissolve, and the water would be drunk as a treatment; the *nazr-i domullo* is a charitable gift, here the fee, given to the *mullo* for his cures.

man az roh-i dur, az Obkena omad-am, guft on zan.
I from way-EZ far, from Obkena come.PT-1s said that woman.
‘I’ve come all the way from Obkena (Ābgina), the woman said.’

ma-ro hamsoya-amon šarofboy,
I-DO neighbor-our Sharofboy
‘My neighbor, Sharofboi (Sharāfbāy),’

ki zan-aš zoyid-a na-tavonist-a
that wife-his DELIVER-CONJ not-been-able-EVID/QUOT
‘his wife can’t deliver her baby’

dard-i saxt kašid-a istod-a-ast,
pain-EZ severe DRAW-CONJ PR-PROG,
‘and is in great pain,’

yak tanga dod-a
one tanga having-given-PART
‘gave me a tanga (“penny”)

baro-yi duo-yi kušoyiš firistod-an-ī šud,
 for prayer-EZ opening about-to-send-INT became.3s,
 ‘and was going to send me to fetch a childbirth prayer.’

zan-on-i hamsoya-g-on-i digar ham
 women-EZ neighbors-EZ other also
 ‘My other neighbors’ wives, too,’

ki har kadom dard-mand bud-a-and,
 that each which suffering they-have-been-EVID/QUOT,
 ‘who all have various aches and pains,’

baro-yi dard-ho-yi xud duo farmud-and
 for-EZ pains-EZ self prayer they-ordered
 ‘asked me to bring back prayers for their aches, too’

va har vaqt pul yob-em
 and any time money we-find-SUBJ
 ‘and as soon as they could afford it,’

nażr-i domullo-ro me-dih-em,
 fee-EZ master-ro we-give,
 ‘they’d send you the money’

burd-a me-dih-ī guft-and.
 TAKE-CONJ you.s-give-BEN they-said.
 ‘through me, they said.’

man ham az baro-yi savob
 I also from sake-EZ (spiritual) reward
 ‘So just to do a good deed’

ba in qadar roh po-yi piyoda xest-a omad-am
 by this much way foot-EZ on-foot RISE-CONJ I-came.
 ‘I’ve come all this way on foot’

šumo duo-ho-ro dihed,
 you.2p prayers-DO give-IMP,
 ‘So, give me the prayers.’

boz har vaqt on-ho pul dihand,
 again any time they money give-3p-SUBJ,
 ‘Then when they give me the money’

ovard-a me-dih-am . . .
 BRING-CONJ I-give-BEN . . .
 ‘I will bring it and give it to you . . .’

kadom-i in duo-ho ba kadom dard ast?
 which-EZ this prayers to which pain is?
 ‘which prayer is for which pain?’

guft-a on zan porsid.
 SAY-PART that woman asked.
 ‘the woman asked.’

BIBLIOGRAPHY AND REFERENCES

- Aghaei, B. (2006) *The Syntax of ke-Clause and Clausal Extraposition in Modern Persian*, Ph.D. dissertation, University of Texas at Austin.
- Ahadi, S. (2001) *Verbergänzungen und zusammengesetzte Verben im Persischen: eine valenztheoretische Analyse*, Wiesbaden: Reichert.
- (2002) *New Persian Language and Linguistics. A Selected Bibliography up to 2001*, Wiesbaden: Reichert.
- Alamolhoda, M. (2000) *Phonostatistics and Phonotactics of the Syllable in Modern Persian*, Helsinki: The Finnish Oriental Society.
- Ansarin, A. A. (2004) 'An acoustic analysis of modern Persian vowels', *SPECOM-2004z*; 315–318—http://www.isca-speech.org/archive/specom_04/spc4_315.pdf
- Bahār, M. T. (1958/1337) *Sabkshenāsi yā tārikh-e tatavvor-e she'r-e fārsi*, Tehrān: Amir Kabir.
- Barakaeva, G. B. (1968) *Luyat-i mukhtasar-i tojikī-anglisī* (Concise Tajik-English dictionary), Dushanbe: Universiteti Davlatii Tojikiston.
- Bashiri, I. (1994) 'Russian loanwords in Persian and Tajiki languages', in Marashi, M. (ed.) *Persian Studies in North America: Studies in Honor of Mohammad Ali Jazayeri*, Bethesda, MD: Iranbooks, pp. 109–139.
- Beeman, W. O. (1986) *Language, Status, and Power in Iran*, Bloomington: Indian University Press.
- (1988) 'Affectivity in Persian language use', *Culture, Medicine and Psychiatry* 12.1: 9–30.
- Birnbaum, S. A. (1950) 'The verb in the Bukharic language of Samarkand', *Archivum Linguisticum* 2.1: 60–73; 2.2: 158–176.
- Bosson, G. (1985) *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*, Tübingen: Narr, 1985.
- Cejpek, J. (1956) 'Die verbale Periphrase als ein wichtiges Unterscheidungsmerkmal zwischen Neupersisch and Taġikisch', *AO* 42.2: 171–182.
- Dehkhodā, A. A. (1994–1995/1372–1373) *Loqatnāme* (Dictionary), Tehrān: Mo'assese-ye Enteshārāt va Chāp-e Dāneshgāh-e Tehrān (some 26,000 pp., also digitalized version).
- Doerfer, G. (1963–1975) *Türkische und mongolische Elemente im Neupersischen. Unter besonderer Berücksichtigung älterer neupersischer Geschichtsquellen, vor allem der Mongolen- und Timuridenzeit*, 4 vols., Wiesbaden: F. Steiner.
- (1967) *Türkische Lehnwörter im Tadschikischen (AKM 37.3)*, Mainz: Harrassowitz.
- (1992) 'Central Asia xiv. Turkish-Iranian language contacts', in *EnIr* 5, pp. 231–233.
- Efimov, V. A. E. (1965) *Iazyk afganskikh khazara. Iakaulangskii dialekt* (The language of the Afghan Hazara. Yakaulang dialect), Moskva: Nauka.
- Eilers, W. (1959) *Deutsch-persisches Wörterbuch*, Wiesbaden: Harrassowitz.
- Estaji, A. and Bubenik, V. (2007) 'On the development of the tense/aspect system in Early New and New Persian', *Diachronica* 24.1: 31–55.
- Farhadi, R. A. G. (1975) *Le persan parlé en Afghanistan. Grammaire du kâboli accompagnée d'un recueil de quatrains populaires de la région de Kâbol*, Klincksieck: Paris (rev. of 1955).
- (1981/1360) *Zabān-e tājiki-ye Māvarā ol-nahr*, Kābul: Akadmi-i Ulum-i J. D. A., Markaz-i Zabānhā va Adabiyāt.
- Farhadi, R. and Perry, J. (2009) 'Kāboli', *EnIr* 15.3, pp. 276–280.
- Folli, R., Harley, H., and Karimi, S. (2005) 'Determinants of event type in Persian complex predicates', *Lingua* 115.10: 1365–1401.
- Gippert, J. (2003) 'Early New Persian as a medium of spreading Islam', in Paul, L. (ed.) *Persian Origins*, pp. 31–47.

- Glassman, E. H. (1972) *Conversational Dari*, Kabul: The Language Orientation Committee, International Afghan Mission (rev. ed.).
- Gomeshi, J. (2003) 'Plural marking, indefiniteness, and the noun phrase', *Studia Linguistica* 57.2: 47–74.
- Griunberg, A. (1963) *Iazyk severnoarzarbaidzhanskikh Tatov* (The language of the North-Azerbaijani Tats), Leningrad: Izd. Akademii Nauk SSSR.
- Griunberg, A. and Davydova, A. (1982) 'Tatskij jazyk', in *Osnovy* 3.1, Moskva, pp. 231–288.
- Haider, H. (1985) 'Chance and necessity in diachronic syntax – Word order typologies and the positions of Modern Persian relative clauses', in Fisiak, J. (ed.) *Papers from the 6th International Conference on Historical Linguistics, Poznan, 22–26 August 1983*, Amsterdam: John Benjamins, pp. 199–216.
- Hayyim, S. (1985) *The Larger English-Persian Dictionary*, Tehran: Farhang-e Mo'āser (numerous impressions).
- Henny, J. L. (1984) 'Enclitics in Pahlavi and Early Classical Persian: A theoretical analysis', in Skalmowski, W. and van Tangerloo, A. (eds) *Middle Iranian Studies. Proceedings of the International Symposium Organized by the Katholieke Universiteit Leuven from the 17th to the 20th of May 1982*, Leuven: Peeters, pp. 84–94.
- Ioannesian, Iu. A. 1999. *Geratskii dialekt iazyka dari sovremenogo Afganistana* (The Herat dialect of modern Afghan Dari), Moscow: Vostochnaia Literatura RAN.
- Jahangiri, N. (2000) *A sociolinguistic study of Persian in Tehran*, Tokyo: Inst. for the Study of Languages and Cultures of Asia and Africa.
- Jahani, C. (2000a) 'Expressions of indirectivity in spoken Modern Persian', in Johanson, L. and Utas, B. (eds) *Evidentials*, pp. 185–208.
- (2000b) 'Restrictive relative clauses in Classical and Modern New Persian and the marking of the antecedent', *Orientalia Suecana* 49: 33–56.
- (2008) 'Expressions of future in Classical and Modern New Persian', in Karimi, S. et al. (eds) *Aspects of Iranian Linguistics*, pp. 152–175.
- Johanson, L. and Utas, B. (eds) (2000) *Evidentials: Turkic, Iranian and Neighbouring Languages*. Berlin: Mouton de Gruyter (Empirical Approaches to Language Typology 24).
- Kalbāsi, Irān (2001/1380) *Sākhṭ-e eshteqāqi-ye vāzhe dar Fārsi-ye emruz* (The derivational structure of the word in Modern Persian), Tehrān: Pazhuheshgāh-e 'Olum-e Ensāni va Motale' āt-e Farhangi.
- Karimi, S., Samiiān, V. and Stilo, D. (eds) *Aspects of Iranian Linguistics*, Newcastle upon Tyne: Cambridge Scholars Publishing.
- Karimov, H. et al. (1970) *Zabon-i adabi-i hozira-i tojik: sintaksis*, Dushanbe: Nashriēt-i Irfon.
- Key, G. (2008) 'Differential object marking in a Medieval Persian text', in Karimi et al. (eds) *Aspects of Iranian Linguistics*, pp. 227–247.
- Khānlari, P. N. (1988/1366) *Tārikh-e zabān-i Fārsī*, Tehrān: Enteshārāt-e Now (fourth printing).
- Kieffer, C. M. (1983) 'Afghanistan v. Languages', *EnIr* 1.5, pp. 501–516.
- (1989) 'L'avancée du persan vers l'Est: le cas des isolats persanophones les plus orientaux en Afghanistan,' *StIr* 18.2: 221–236 (vernaculars of Laghmān, Sorkhrōd, Kandībāgh, Ūrgūn).
- (2004) 'Hazāra iv. Hazāragi dialect,' *EnIr* 12, pp. 90–93.
- Koutlaki, S. A. (2002) 'Offers and expressions of thanks as face enhancing acts: *tæ'ārof* in Persian', *Journal of Pragmatics* 34.12: 1733–1756.
- Lazard, G. (1956) 'Caractères distinctifs de la langue Tadjik', *BSL* 52: 117–186.
- (1963) *La langue des plus anciens monuments de la prose persan*, Paris: Klincksiek.

- (1968) 'La dialectologie du judéo-persan', *Studies in Bibliography and Booklore* (Cincinnati) 8: 77–98.
- (1970) 'Études quantitative de l'évolution d'un morphème: la postposition *rā* en persan', in Cohen, D. (ed.) *Mélanges Marcel Cohen*, The Hague: Mouton, pp. 381–388.
- (1982) 'Le morphème *rā* en Persan et les relations actancielles', *BSL* 77.1: 197–208.
- (1985) 'L'inférentiel ou passé distancié en persan', *StIr* 14, 1985, pp. 27–42.
- (1990a) *Dictionnaire persan-français*, Leiden, New York: E. J. Brill.
- (1990b) 'Lumières nouvelle sur la formations de la langue persane: un traduction du Coran en persan dialectal et ses affinités avec le judéo-persan', in Shaked, S. and Netzer, A. (eds) *Irano-Judaica II, Studies Relating to Jewish Contacts with Persian Culture Throughout the Ages*, Jerusalem: Ben-Zvi Institute, pp. 184–198.
- (1992) *Grammar of Contemporary Persian*, Costa Mesa, Calif: Mazda Publishers (tr. of *Grammaire du persan contemporain*, Paris: Klincksieck, 1963).
- (1993) *The Origins of Literary Persian, The Annual Norūz Lecture by Distinguished Scholars of Iranian Studies*, Bethesda, Maryland: The Foundation for Iranian Studies.
- (1995) *La formation de la langue persane*, Paris: Peeters.
- (1996) 'Darī', *EnIr* 7.
- (1999) 'Pahlavi, Pārsi, Dari', in *EnIr* 9, pp. 158–162.
- (2002) 'Le médiatif: considérations théoriques et application à l'iranien', in Johanson, L. and Utas, B. (eds), pp. 209–28.
- (2002) 'La formation du persan sous l'influence de la langue arabe au IV/X siècle', *CSI* 26: 11–19.
- (2003) 'Du pehlevi au persan: diachronie ou diatopie?', in Paul, L. (ed.) *Persian Origins*, pp. 95–102.
- Lorentz, M. (1964) 'Die partizipialen Wendungen – ein Unterscheidungsmerkmal des Tağikischen gegenüber dem Persischen', *Mitteilungen des Instituts für Orientforschung* 10: 133–139.
- Mahootian, S. (1997) *Persian. A Descriptive Grammar*, London: Routledge.
- Matras, Y. (2002) 'Kurmanji complementation: Semantic-typological aspects in an areal perspective', *STUF* 55: 49–63.
- Meunier, A. and Samvelian, P. (1997) *La postposition rā en persan. Ses liens avec la détermination et sa fonction discursive, Cahiers de Grammaire 22*, Toulouse: Université de Toulouse–Le Mirail.
- Meyer-Ingversen, J. (1974) 'Fekr. Ein Beitrag zur Transformationsgrammatik des Neupersischen', in Frye, R. (ed.) (1974) *Neue Methodologie in der Iranistik*, Wiesbaden: Harrassowitz, pp. 131–282.
- Nercissians, E. (2001) 'Bilingualism and diglossia: Patterns of language use by ethnic minorities in Tehran', *International Journal of the Sociology of Language* (Berlin) 148: 59–70.
- (1985/1364) 'Dozabānegi dar Irān' (Bilingualism in Iran), *Majalle-ye Zabān-šenāsi* 2.2: 73–80.
- Öhl, P. and Korn, A. (2008) 'Performanzbasierte und parametrische Wandel in der linken Satzperipherie des Persischen. Der Subordinationsmarker *ke* und die Interrogativpartikel *āyā*', *Die Sprache* 46: 137–202.
- Oranskij, I. M. (1975) *Die Neuiranischen Sprachen der Sowietunion*, translated by W. Winter, The Hague, Paris: Mouton & Co.
- Paul, L. (1999) 'Neuere Tendenzen in der Entwicklung des Persischen am Beispiel der Teheraner Freitagspredigten', *StIr* 28.2: 275–290.
- (2000a) 'The language of the *Shahnameh* in historical and dialectal perspective', ILEX Foundation, *Papers Presented at the Second Ferdowsi Conference*,

- Center for the Great Islamic Encyclopedia, Tehran, Iran, August 21–28, 2000*, http://www.ilexfoundation.org/pdf/ferdowsi_2/ludwigpaul300.pdf
- (2000b) ‘A linguist’s fresh view on “Classical Persian”’, *CSI* 26: 21–34.
- (ed.) (2003a) *Persian Origins – Early Judaeo-Persian and the Emergence of New Persian. Collected Papers of the Symposium, Göttingen 1999*, Wiesbaden: Harrassowitz.
- (2003b) ‘Early Judaeo-Persian in a historical perspective: The case of the prepositions *be*, *u*, *pa(d)*, and the suffix *rā*’, in Paul, L. (ed.) *Persian Origins*, pp. 177–194.
- (2008) ‘Some remarks on the Persian suffix *-rā* as a general and historical issue’, in Karimi, S. *et al.* (eds) *Aspects of Iranian Linguistics*, pp. 328–337.
- Perry, J. R. (1985) ‘Language reform in Turkey and Iran’, *International Journal of Middle East Studies* 17.3: 295–311; reprinted with corrections in Atabaki, T. and Zürcher, E. J. (eds) *Men of Order: Authoritarian Modernization in Turkey and Iran*, London, New York: I. B. Tauris, pp. 238–259.
- (1989) ‘Uzbek influence on Tajik syntax: The converb constructions’, in Clyne, P. R., Hanks, W. F., Hofbauer, C. L. (eds) *The Elements: A Parasession on Linguistic Units and Levels. Papers from the Conference on Non-Slavic Languages of the USSR*, Chicago: Chicago Linguistic Society, The University of Chicago, pp. 448–461.
- (1991) *Form and Meaning in Persian Vocabulary: The Arabic Feminine Ending*, Costa Mesa, Calif.: Mazda Publishers.
- (1996) ‘From Persian to Tajik to Persian: Culture, politics and law reshape a Central Asian language’, in Aronson, H. I. (ed.) *NSL.8. Linguistic Studies in the Non-Slavic Languages of the Commonwealth of Independent States and the Baltic Republics*, Chicago: Chicago Linguistic Society, The University of Chicago, pp. 279–305.
- (1997) ‘Script and scripture: The three alphabets of Tajik Persian, 1927–1997’, *Journal of Central Asian Studies* 2.1: 2–18.
- (2000) ‘Epistemic verb forms in Persian of Iran, Afghanistan and Tajikistan’, in Johanson, L. and Utas, B. (eds) *Evidentials*, pp. 229–257.
- (2002) ‘Grammaticalization in process: The evolution of the Afghan Persian adverb *xāt*’, in Cavoto, F. (ed.) *A Linguist’s Linguist: A collection of Papers in Honour of Alexis Manaster Ramer*, München: Lincom Europa, II, pp. 317–27.
- (2003) ‘Persian as a homoglossic language’, *CSI* 27: 11–28.
- (2005) *A Tajik Persian Reference Grammar*, Leiden-Boston: Brill.
- (2007) ‘Persian’, in Kaye, A. S. (ed.) *Morphologies of Asia and Africa*, 2 vols., Winona Lake, Ind.: Eisenbrauns, vol. 2, pp. 975–1019.
- (2009) ‘The origin and development of Literary Persian’, in Bruijn, J. T. P. (ed.) *A History of Persian Literature, Vol. I. General Introduction to Persian Literature*, London, New York: I. B. Tauris, pp. 43–70.
- Pisowicz, A. (1985) *Origins of the New and Middle Persian Phonological Systems*. Kraków: Nakładem Uniwersytetu Jagiellońskiego.
- (2003) ‘The possible influence of Kurdish dialects on colloquial Persian: the definiteness suffix *-é*’, *FO* 39: 177–180.
- Rajā’i, A. Alī (1964/1343). *Yāddāshti dar bāre-ye lahje-ye Bokhārā’i* (French title page: *Remarques sur le dialecte de Bukhara*), Mashhad: Chāpkhāne-ye Dāneshgāh-e Mashhad.
- Rastorgueva, V. S. (1954) ‘Kratkii ocherk grammatiki tadjhikskogo iazyka’, in Eadem (ed.) *Tadjhiksko–russkii slovar*, Moskva: Gos. Izd-vo inostranykh i natsional’nykh slovarei, pp. 529–570.
- (1963) *A Short Sketch of Tajik Grammar* [= *Kratkii ocherk grammatiki tadjhikskogo iazyka*], Paper, H. H. (transl. & ed.), Bloomington, Indiana and The Hague: Research Institute for Inner Asian Studies, Indiana University.

- (1964) *Opyt sravnitel'nogo izucheniia tadzhikskikh govorov*, Moskva: Nauka.
- Rastorgueva, V. S. and Kerimova, A. A. (1964) *Sistema tadzhikskogo glagola*, Moskva: Nauka.
- Rakhimi, M. Kh. and Uspenskaia, L. Kh. (1954) (eds) *Tadzhiksko-russkii slovar'*, Moskva: Gos. Izd-vo inostranykh i natsional'nykh slovarei.
- Rzehak, L. (2001) *Vom Persischen zum Tadschikischen. Sprachliches Handeln und Sprachplanung in Transoxanien zwischen Tradition, Moderne und Sowjetmacht (1900–1956)*, Wiesbaden: Reichert.
- Sadat-Tehrani, N. (2005) 'Indifference *ke*-construction in Modern conversational Persian', in *Proceedings of the 2004 annual conference of the Canadian Linguistic Association*, <http://http-server.carleton.ca/~mojunker/ACL-CLA/pdf/Sadat-Tehrani-CLA-2004.pdf>
- Sāfi, Qāsem (2003/1982) *Bahār-e adab: Tārikh-e mokhtasar-e zabān va adabiyāt-e Fārsi dar shebḡāre-ye Hend va Pākestān*, Tehrān: Mo'assase-ye Enteshārāt va Chāp-e Dāneshgāh-e Tehrān.
- Samvelian, P. (2008) 'The Ezafe as a head-marking inflectional affix: Evidence from Persian and Kurmanji Kurdish', in Karimi *et al.* (eds) *Aspects of Iranian Linguistics*, pp. 339–361.
- Soper, J. (1996) *Loan Syntax in Turkic and Iranian: The Verb System of Tajik, Uzbek and Qashqay*, Bodrogligeti, A. J. E. (rev. and ed.), Bloomington, Indiana: Eurolingua.
- Steingass, F. J. (1970) *A comprehensive Persian-English dictionary, including the Arabic words and phrases to be met with in Persian literature*, Beirut: Librairie du Liban (first impression 1892).
- Utas, B. (1991) 'New Persian as an interethnic medium', in Svanberg, I. (ed.) *Ethnicity, Minorities and Cultural Encounters*, Uppsala: Centre for Multiethnic Research, pp. 103–111.
- (2006) 'A multiethnic origin of Persian', in Johanson, L. and Bulut, C. (eds) *Turkic-Iranian Contact Areas*, Wiesbaden: Harrassowitz, pp. 241–251.
- Vivier-Muresan, A. (2006) 'Le code de politesse iranien (*ta'ārof*) ou la fiction du lien social', *L'Homme* 180, pp. 115–138.
- Windfuhr, G. L. (1979) *Persian Grammar: History and State of its Study*, The Hague, Paris and New York: Mouton.
- (1980) 'Verbal deictic matrices in Persian', paper given at the 190th Meeting of the American Oriental Society, San Francisco, April.
- (1982) 'The verbal category of inference in Persian', *Acta Iranica* 22: 263–287.
- (1986) 'A spatial model for tense, aspect and mood', *Folia Linguistica* 19.3: 415–461 (verb systems of Persian, French, Russian, Arabic, Turkish, Creole).
- (1987) 'Persian', in Comrie, B. (ed.) *The World's Major Languages*, London: Croom Helm, pp. 523–546 (rev. eds. 2009).
- (1987) 'Convergence: Iranian Talyshi and Turkic Azari', in Bashir, E., Deshpande, M. M., and Hook, P. E. (eds) *Select Papers from Sala-7, South Asian Languages Roundtable Conference, Ann Arbor, Michigan, May 1985*, Indiana, Bloomington: University Linguistics Club, pp. 385–405.
- (1997) 'Persian phonology', in Kaye, A. S. (ed.) *Phonologies of Asia and Africa*, Winona Lake, Ind.: Eisenbrauns, pp. 663–677.
- (2004) 'The Iranian languages: Typology in space and time', Historical Linguistics Colloquium, Department of Linguistics, University of Michigan, two lectures, March 2004, unpublished ms.
- (2006) 'Language change and modeling modal axes: Irano-Turkic convergence', in Johanson, L. and Bulut, C. (eds) *Turkic-Iranian Contact Areas. Historical and Linguistic Aspects*, Wiesbaden: Harrassowitz, pp. 252–283.

CHAPTER NINE

ZAZAKI

Ludwig Paul

1 INTRODUCTION

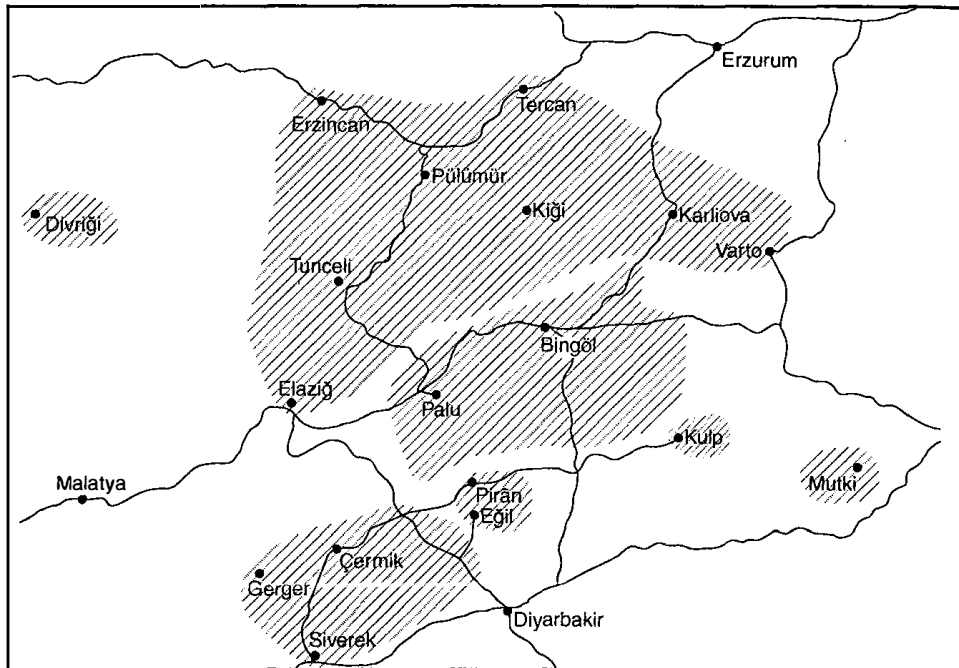
1.1 Overview

The name Zazaki (Zāzākī) is an exoterm, originally with pejorative connotation. The endoterm Dim(i)li is used, in particular by Sunni Zaza, besides other self-designations by various Zazaki groups. Geographically, Zazaki is spoken in eastern Turkey at the north-western borders of Kurdish-speaking areas, in a triangle defined by the cities of Siverek, Erzincan and Varto, as well as in an exclave of several villages near Mutki west of Bitlis, which may suggest a formerly wider distribution of Zazaki. Linguistically, Zazaki is a North-West Iranian language, more closely related to Gōrānī and the (Iranian) Āzari dialects than to Kurdish.

The history of Zazaki studies began in 1856, with P. I. Lerch's recordings of about 40 pages of text (including a German translation) in Zazaki, which at that time was still considered a Kurdish dialect. In 1932, the first grammatical description of Zazaki was attempted by Hadank (ed. of Mann 1932), which also served to establish Zazaki, among linguists, as a language in its own right. No grammar in a modern sense was published before Todd (1985). Yet another 13 years passed before two *comprehensive* grammars of Zazaki were written (Paul 1998; Selcan 1998). The present description is an abbreviated and revised version of the former.

There are no reliable statistics about the number of Zazaki speakers. In SE Anatolia they might number between 1.5 and 2 million. About the same number of Zaza may have emigrated to the urban centers of Western Anatolia, and to Western Europe, during the last 40 years. These numbers include all *ethnic* Zaza, however, many of whom (esp. the younger generation) have been assimilated to Turkish or Kurdish meantime. Another (apparently declining) part of the Zaza, although still speaking their mother tongue, have traditionally considered themselves as Kurds speaking a dialect of Kurdish. There have been ongoing emotional discussions among Zazas and Kurds for the past 20–25 years as to whether the Zaza form a separate people, or nation, and accordingly whether their form of speech may, or must, be called a separate (non-Kurdish) language. In most parts of the Zaza's home country there also live many Turks or Kurds, e.g. 50 percent of Turks in the city of Erzincan, or 50 percent of Kurds in Siverek and Varto. The most important cities with a predominant Zaza population are Çermik, Tunceli and Bingöl.

In the Republic of Turkey, Zazaki shares with (Kurmanci) Kurdish the status of a minority language, but (like Kurdish) it is not granted the rights that minority languages usually have in Europe. It has been ignored by state officials, or its existence has been denied until the 1980s. From 1989 on, publication in Zazaki (and Kurdish) has been partly allowed in Turkey, but the actual government policy has rather discouraged, or



MAP 9.1 ZAZAKI

prevented by force, regular and free publication in Zazaki. Since the 1980s, the main impetus for writing in Zazaki has come from the Zaza diaspora (exiles and guest workers) living in Western Europe (mainly Germany and Sweden).

Zazaki is divided into many quite heterogeneous dialects, most of which may be attributed to either of three (a northern, south-western, and south-eastern) dialect areas. Actually, Zazaki “consists” of dialects only, since no single dialect, or standardized mixture of dialects, can claim to be used (if in writing only) by a majority of speakers. The northern dialect area is defined not only linguistically, but also religiously, as all “northern” Zazas are Alevi Shi‘îs, while the southern Zazas are Shafi‘î Sunni.

Zazaki has been in contact with many other languages in history. The most important single language to influence Zazaki over time has been (Kurmanci) Kurdish. Other languages whose contact with Zazaki may still be seen today include Arabic, Persian, Armenian (esp. for the northern dialects), and Turkish whose influence on Zazaki has been rapidly increasing for the past decades.—In the following, NX refers to the dialect in Berz 1988.

2 PHONOLOGY

2.1 Inventory and distribution

The Çermik-Siverek dialects of Zazaki have the following *phonemic* system of 8 vowels and 30 consonants, including 2 semi-vowels.

2.1.1 Vowels and diphthongs

2.1.1.1 Vowel system

TABLE 9.1: ZAZAKI VOWELS

	Front	Back
Close	<i>ī i</i>	<i>u ū</i>
Mid	<i>ē</i>	<i>o</i>
	<i>e</i>	
Open	<i>ā</i>	

2.1.1.2 Diphthongs

In the Çermik-Siverek dialects of Zazaki there are only falling diphthongs ending in *-y* (*āy, ey, oy, ūy*) or *-w* (*āw, ew, ēw*). When a diphthong is followed by a vowel, it loses its diphthongal character, e.g. *rāy-ā mi* is syllabic *rā-yā mi* ‘my way’. Therefore, diphthongs are not considered phonemic here.

2.1.2 Consonants

TABLE 9.2: CONSONANTS

	Labial	Dental	Palatal	Velar	Uv.	Phar.	Glottal
Stops/Affricates :	<i>p b</i>	<i>t ʈ d</i>	<i>č ǰ</i>	<i>k g</i>	<i>q</i>		
Fricatives	<i>f v</i>	<i>s ʃ z</i>	<i>š ž</i>	<i>x γ</i>		<i>ħʕ</i>	<i>h</i>
Nasals	<i>m</i>	<i>n</i>					
Laterals, plain/vel.		<i>l l</i>					
Vibrants, flap/trill		<i>r ř</i>					
Semi-vowels	<i>w</i>		<i>y</i>				

2.2 Non-segmental features

As a rule, Zazaki words have their stress on the last syllable of the stem, while endings and suffixes are unstressed (stress is indicated here by underlined vowel). But note the following exceptions:

Stressed are the case endings *-er* and *-ān*; the negative prefixes *nē-* and *me-*, and the modal prefix *b(i)-*. Unstressed are the primary postpositions *-ā, -di(r), -rā, -rē, -ro*; the substantive verb; the particle *do* (future); and the clitic conjunctions *ū* ‘and’ and *žī* ‘also’. Vocative stress is on the penultimate syllable, where possible. Unpredictably, a small set of words has non-final stress, e.g. *ħepīs* ‘jail’ (anaptytic *-i-*, cf. Ar. *ħabs*), *kārdī* f ‘knife’, *mākī* ‘female’, *wīnā* ‘so’. Difference of stress may be phonemic: *nīnā* ‘she does not come’ ~ *nīnā* ‘of these’ (obl. pl. of dem.).

3 MORPHOLOGY

3.1 Nominal morphology

3.1.1 Nouns

Zazaki nouns distinguish the following inflectional categories: gender (masculine and feminine); number (singular and plural); case (direct and oblique); animacy; and definiteness. The distinction of gender exists only in the singular, including the number ‘one’, and natural gender in the vocative. The categories of animacy and definiteness operate in restricted areas.

3.1.1.1 Gender

In their simplest forms nouns are unmarked for gender, except for the following subsets:

- (1) Masculine are: (a) nouns in *-e*: *āyre* ‘mill’, *girwe* ‘work’; (b) infinitives in *-iš*.
- (2) Feminine are: (a) abstract nouns in *-ey*, e.g. *rind-ey* ‘kindness’; (b) infinitives in *-er*; (c) tree names in *-ēr*: *sā-y-ēr* ‘apple tree’.

Also, singular feminines with final consonant are often marked by a “euphonic” *-i*: *āw-i* (f) *ā xo sere ki* ‘hold the water (f) to your head’. However, gender is always overtly distinguished in the oblique cases and in the *ezāfe* construction, including nouns allowing for natural gender distinction, such as *wēr* ‘owner (m/f)’: *wēr-ē* (m) *dew-i* (OBLf) ‘the owner (m) of the village’ vs. *wēr-ā* (f) *por-ī* (OBLm) ‘the owner (f) of the hair’.

3.1.1.2 Number and case

Both singular and plural distinguish two basic cases, direct (DIR) and oblique (OBL). Any noun designating a plural number always takes the plural endings, while nouns used in the generic sense and collective nouns take the singular: *šār do pey bi-huw-o* ‘the people (s) will laugh about it’ (*pey huw-*). Feminine oblique *-er* is facultative. Vocalic endings insert hiatus *-y-*; *-ī* replaces masc. *-e*, otherwise *-ī > -y* after vowel.

TABLE 9.3: BASIC NOMINAL INFLECTION

	masc.	masc.	fem.	fem.	
DIRs	<i>-ø</i>	<i>-e</i>	<i>(-i)</i>	<i>-ā</i>	
OBLs	<i>-ī/-y</i>	<i>-ī/-y</i>	<i>(-er)(i)</i>	<i>-ā, -er-(i)</i>	
DIRp	<i>-ī/-y</i>	<i>-e-y</i>	<i>-ī/-y</i>	<i>-e-y</i>	
OBLp	<i>-ān</i>	<i>-(ey)-ān</i>	<i>-ān</i>	<i>-ān</i>	
	‘house’	‘camel’	‘girl’	‘story’	‘sister’
DIRs	masc. <i>bān</i>	masc. <i>deve</i>	fem. <i>keynek(-i)</i>	fem. <i>meselā</i>	fem. <i>wā</i>
OBLs	<i>bān-ī</i>	<i>dev-ī</i>	<i>keynek(-er)(-i)</i>	<i>meselā, mesel-er(-i)</i>	<i>wā-r(-i)</i>
DIRp	<i>bān-ī</i>	<i>deve-y</i>	<i>keynek-ī</i>	<i>mesel-e-y</i>	<i>wā-y</i>
OBLp	<i>bān-ān</i>	<i>dev-ān</i>	<i>keynek-ān</i>	<i>mesel-ān</i>	<i>wā-y-ān</i>

A small subset of nouns has a secondary oblique case with a “kinship-*r*”: *mālmā-r* ‘mother’, *pīlpē-r* ‘father’, *birālbirā-r* ‘brother’, *wālwā-r* ‘sister’. These are obligatory in those *ezāfe* constructions where the noun phrase is other than the syntactic subject or direct object (see section 4.2 Noun phrase structure).

3.1.1.3 Vocative

The vocative may distinguish gender, with a forward stress shift up to the pre-penultimate where possible: masc. -Ø, fem. -ē, (rarely) -ī or -Ø; plural -ēno.

3.1.1.4 Definiteness and unity

A noun in its simplest form may express either definite or indefinite singular or unity, or be used in a generic sense:

[lāžek] šī-n-o ĵi-rē [lāžek] b-en-o
 ‘[the boy] goes (away)’; ‘she has (‘to her is being’) [a boy]’;

[piyāz] weš perey ke-n-o
 ‘[onions] make (keno) good money’ (note ms -o).

Indefinite plurality may be expressed by the singular or plural in the direct case:

[čīnā-y-o-m new-e-DIR-SM] bī-y-ār-ē
 ‘bring [new clothes]’;

plurality,

[enĵil-ē-p wišk-ī-DIR-p] we-n-ē
 ‘they eat [dried figs]’.

In addition, indefiniteness and unity may be expressed by two clitics or the numeral ‘one’:

(1) -ē (unstressed), -ēn before vocalic clitics only: *ĵā-y[-ē]-rā* ‘from some place’, *hetē ĵā-y[-ēn]-ā* ‘to some place, (to) somewhere’. It supersedes the oblique endings: *šew-ē* ‘(for) one night’. In function, it corresponds more or less to the English indefinite article, and, depending on the context, either indefiniteness or unity may be more prominent: *bān-ē(n)* ‘some, any house’, or ‘one house’ (not two or three).

(2) The suffix -nā ‘another’ implies also an indefinite meaning. It likewise supersedes inflectional endings: oblique *merdim-[nā]-rē* ‘to another man’.

(3) *žew/žū* ‘one’ (m/f) may be thus used as well:

[žew] ĵāmērd ū [žū] ĵīn-y-ā ĵi
 ‘[a] young man and [one] wife of his (ĵī)’.

3.1.2 Adjectives

There is no strict boundary between adjectives and substantives in Zazaki. Adjectives may regularly be used as substantives, and some substantives as adjectives, while some nominals are difficult to assign to either class:

sere-y-ē [ē bēnāmūs-ī-OBL]
 ‘the head of that [dishonourable] (man)’;

merdim-ē-do [dew-iž]
 ‘a (-ē) certain man, a peasant’ (‘a certain [peasant-ish] man’);

lāžek-ē-do bol [ĵāmērd]
 ‘a very [brave] boy’ (< *ĵāmērd* ‘hero’)

(for typical adjectival derivation, see section 6 Lexis).

3.1.2.1 Gender, case and number

Like nouns, adjectives distinguish gender and show a two-case system. Unlike nouns, masculines in *-e* have corresponding feminines in *-ē*, and singular feminines never take obl. *-er*.

TABLE 9.4: ADJECTIVAL INFLECTION

	‘big’		‘white’		‘black’	
	masc.	fem.	masc.	fem.	masc.	fem.
DIRs	<i>gird</i>	<i>gird(-i)</i>	<i>si p-e</i>	<i>si p-ē</i>	<i>si yā</i>	<i>si yā</i>
OBLs	<i>gird-ī</i>	<i>gird(-i)</i>	<i>si p-ī</i>	<i>si p-ē</i>	<i>si yā-y</i>	<i>si yā</i>
DIRp	<i>gird-ī</i>		<i>si p-e-y</i>		<i>si yā-y</i>	
OBLp	<i>gird-ān</i>		<i>si p-ān</i>		<i>si yā-y-ān</i>	

When used predicatively, adjectives may optionally take the case ending of the noun to which they refer:

mā pēro weš-ī-p yē
‘we are all well-p’, but

īnsān-ī-p ben-ē āyā-s
‘the people-p become awake (*āyā*)’, i.e. ‘wake up’.

For attribution, see Section 4.2 Noun phrase structure.

3.1.2.2 Comparison

The simple form of an adjective may regularly be used in comparative or superlative functions: comparative, *nāme-y birā-y [werd]-ī* ‘the name of the [small(er)] brother’; superlative, *šeytān-o [gird] ti yē* ‘the [big(gest)] Satan is you’.

More explicit options are *dehā* ‘more’ or *en* ‘most’ (< Turkish *daha*, *en*):

zūr-ā wārezā-y [dehā gird] bī
‘the lie of the nephew was [bigger]’;

keynek-ā [en gird]-i riš-en-o ŋor
‘he sends the [biggest] girl upwards’.

The comparative relation is expressed either by the postposition *-rā*: *mi-rā gird* ‘bigger than me’ (lit. ‘big from me’); or by the adjectival suffix *-ēr(ī)*: *mi gird-ēr(ī)* ‘bigger than me’ (cf. English *my elder*). The former is used generally for simple predication, the latter for all other types of comparative sentences:

kund [mi āqil-ēr] viž-y-ā
‘the owl has turned out [smarter than me]’.

The superlative relation is expressed by postpositional *miyān-di* ‘among’:

[nē (teyr ū ŋūr)-ī-OBL miyān-di] en sifte šewšewok-i ye-n-ā
‘[among these birds], as the first comes the bat’.

The noun of reference may be resumed by a pronoun, such as the general 3rd person referential *ji*:

[*(jin ū jināwir)-ī-OBL miyān-di*], *en tersinok-ē ji arwēš-ī b-en-ē*
 '[among the animals], the most fearful of them are the hares'.

3.1.3 Pronouns and deixis

3.1.3.1 Personal pronouns

The Zazaki personal pronouns distinguish three persons in the singular and plural, and direct and oblique cases, except in the 1p, 2p. Gender is distinguished in the 3s. There is also an invariable weak (but not enclitic) oblique form. Note that Zazaki (like Northern Kurdish) lacks personal suffixes.

TABLE 9.5: PERSONAL PRONOUNS

	1s	2s	3sm	f	1p	2p	3p
DIR	<i>ez</i>	<i>ti</i>	<i>o</i>	<i>ā</i>	<i>mā</i>	<i>šimā</i>	<i>ē</i>
OBL	<i>mi(n)</i>	<i>to</i>	<i>ey</i>	<i>āy</i>	<i>mā</i>	<i>šimā</i>	<i>inān</i>
weak				<i>ji</i>			<i>ji</i>

The 1s *min* occurs before enclitic vowels only, e.g. *ti lāž-ē min ē* 'you are my son'.

3.1.3.2 Demonstrative pronouns

The demonstrative pronouns have one distant form, identical with the strong 3rd person personal pronouns, and one proximate form distinguished only by initial *n-*. There are also two corresponding demonstrative adjectives which precede the noun they qualify, and are morphologically less differentiated than the former.

TABLE 9.6: DEMONSTRATIVES

	'this'		'these'	'that'	'those'	
	masc	fem.	plural	masc.	fem.	plural
DIR	<i>no</i>	<i>nā</i>	<i>nē</i>	<i>o</i>	<i>ā</i>	<i>ē</i>
OBL	<i>ney</i>	<i>nāy</i>	<i>nīnā(n)</i>	<i>ey</i>	<i>āy</i>	<i>inān</i>
ADJ	<i>no</i>	<i>nā</i>	<i>nē</i>	<i>o</i>	<i>ā</i>	<i>ē</i>
OBL	<i>nē</i>	<i>nā</i>	<i>nē</i>	<i>ē</i>	<i>ā</i>	<i>ē</i>

3.1.3.3 Reflexive and possessive pronoun

The Zazaki reflexive marker *xo* 'self, own' is invariable, and co-referent with the syntactic subject only. In non-ergative sentences, it must be used for any "oblique" verbal or nominal complement:

(1) Present direct object

[*xo*] *kuw-en-ā 'erd-ā*

'she throws [herself] to the ground'.

(2) Adverbial

ye-n-o [ā xo]

'he comes [to himself]'.

(3) Possessive of direct object

[meselā xo] mi-rē vāž-i

‘tell me [your story]’.

Similarly in past ergative sentences, *xo* occurs only if co-referential with the subject/agent, in which case the verb always shows masculine singular agreement:

nē kese-y mā-rā tersā-y
these turtle-s us-from fear-ed-3p
‘these turtles were afraid of us’,

[xo] ešt dol-ī miyān ū [xo] dol-di nimit.
self threw-3sm lake-OBL inside and self lake-in hid-3sm
‘they threw themselves inside the lake and hid [themselves] in the lake.’

Accordingly, co-reference with the direct object/patient requires the respective personal pronoun:

pī-y-ē [to]-OBL ti-DIR-PAT polāt kerdī bī-2s?
‘had [your] father steeled you?’

In subordinate clauses co-reference with the subject of the main clause is expressed by the simple referential pronoun:

čī-y-o ki [jī-rē] lāzim o, herīn-en-ē
‘what (-ever) is necessary [to them], they buy’.

3.1.3.4 Reciprocal pronoun

A reciprocal relationship within a sentence is indicated by invariable *pē* ‘each other’ which always shows (like reflexive *xo*) singular masculine agreement:

nē wirnā [pē] vīn-en-ē šimā wirnā helāl [-ē pē] yē
‘they both see [each other]’; ‘you both are *helāl* [to one another]’;

mā [pē] kerd (ms) vīnī
‘we have lost [one another]’.

3.1.4 Adpositions

3.1.4.1 Primary adpositions

3.1.4.1a Simple preposition

The simple prepositions immediately precede the nominal form they qualify, which is normally in the oblique case; thus PREP + OBL (mostly). The following are attested:

<i>ā</i> ‘to, towards’;	<i>hetā(n)</i> ‘until, up to’;	<i>zey</i> ‘like, according to’;
<i>bē</i> ‘without, except’;	<i>nezdī</i> ‘towards, near’;	
<i>bī</i> ‘with, through’;	<i>pey</i> ‘behind’.	

3.1.4.1b Simple postpositions

The simple postpositions may be attached to nominals and secondary postpositions. They usually require the oblique, thus OBL + POST (mostly). Some of these postpositions more or less correspond to a case of traditional grammar:

-ā ‘with, through’ (instrumental);	- <i>di(r)</i> ‘in, on’ (locative);
- <i>rā</i> ‘from, out of; along; through’ (ablative);	- <i>rē</i> ‘for, to’ (dative);
- <i>ro</i> ‘(up)on’; - <i>o</i> ‘(up)on’ (only after ‘ <i>erd</i> m ‘ground’ and <i>rāy</i> f ‘way’).	

3.1.4.2 Secondary adpositions

3.1.4.2a Secondary prepositions

Secondary prepositions are nouns, adverbs, and other expressions with locational meaning, partly borrowed, that syntactically form *ezāfe* constructions, e.g.:

<i>bin-ē</i> N	‘under N’ (< <i>bin</i> ‘bottom’);
<i>čoršme-y-ē</i> N	‘around N’ (< <i>čoršme</i> ‘surrounding’);
<i>bāhd-ē</i> N	‘after N’ (< Ar. prepos. <i>ba’d</i> ‘id.’);
<i>gore-y-ē</i> N	‘according to N’ (< Turk. postpos. <i>göre</i> ‘id.’).

3.1.4.2b Secondary postpositions

Secondary postpositions are nouns, adverbs, and other expressions with locational meaning, such as *bin* ‘under’ (‘bottom’), *ser* ‘on(to)’ (‘head’), *zere* ‘in(to)’, that syntactically form *ezāfe* constructions. They express motion, but state with the postposition *di(r)* ‘in’, e.g. *bīr-ī bin* ‘(down) into the well’, but *bīr-ī bin[-di]* ‘[at] the bottom of the well’.

3.1.4.3 Circumpositions

Prepositions may be combined with simple postpositions to form circumpositions that always require the OBL. Two plain prepositions, *bi*, *zey*; and four that contain an *ezāfe*, *hendē*, *hetē*, *qāndē*, *tewdē*, may optionally be combined with the postposition *-ā*, without altering their meaning:

[<i>bi</i>] <i>ling-ān-dē xo-y[-ā]</i>	[<i>qān-dē</i>] <i>čičī-y[-ā]</i>
‘[on (‘by’)] your own feet’;	‘[for] what’.

Similarly, the prepositions *ā* and *bi*, as in *ā . . . rā*, *ā . . . ro* and *bi . . . ro*, do not alter the meaning of the postpositions: *ā xo-rā* ‘on themselves’. (*bi*) . . . *ā* and *-rā* form a large number of modal and causal adverbial phrases (*-rā* often with plural):

<i>bi hērs-ā</i> ‘furiously’, <i>bi huwatey-ā</i> ‘laughingly’, <i>rāšt-ā</i> ‘really’;
<i>kēf-ān-rā</i> ‘with joy’, <i>ters-ān-rā</i> ‘for fear’.

3.1.5 Adverbs

Any adjective where this is semantically possible may have *adverbial* sense: *dūrī* ‘far’, *tāy* ‘(a) few’, *weš* ‘well’, *germ* ‘warm’, etc. In their turn, some adverbs may be used as attributive adjectives, e.g. *winā* ‘such’ (< *winā* adv. ‘so’).

Some important *temporal* adverbs include: *emser* ‘this year’, *pār* ‘last year’, *ewro* ‘today’, *vizēr* ‘yesterday’, *mešti* ‘tomorrow’, *nūkā* ‘now’, *rew* ‘soon, fast’, *tīm* ‘always’, *finā* ‘again’, *finē* ‘once’, *finē-rā* ‘suddenly’. Some important *local* adverbs include: *jēr* ‘below’, *jor* ‘above’, *tiyā* ‘hither, here’, *wār* ‘down(ward)’, *wižā* ‘thither, there’. *tiyā* and *wižā* may be combined with simple postpositions: *tiyā-rā* ‘from here’, and also be used as masculine nouns, *tiyā* ‘the (area) here’. A particular group of local adverbs are the “absolute” prepositions, such as *pey* ‘thereby’, and circumpositions such as *pirā* ‘thereof’, *tā* ‘thereby’.

Nouns, when used for denoting units of time are in the OBL: *šefāq-ī* ‘in the morning (m)’. The suffix *-nā*, attached to certain nouns and adverbs, may form temporal adverbs that point to the future: *ser-nā* ‘next year’, or give the meaning ‘already’ in the past: *vizēr-nā* ‘already yesterday’.

3.1.6 Numerals

3.1.6.1 Cardinal numbers

The Zazaki cardinal numbers are the following:

TABLE 9.7: CARDINAL NUMBERS

1–9	11–19	10–90	100+	
<i>zew/žū</i> (m/f)	<i>žondes</i>	<i>des</i>	<i>se</i>	100
<i>di(di)</i>	<i>diwēs</i>	<i>vīst</i>	<i>di sey</i>	200
<i>hīrē</i>	<i>hīrēs</i>	<i>hīris</i>	<i>pānsey</i>	500
<i>čihār/čehār</i>	<i>čārēs</i>	<i>čewres</i>	<i>hezār</i>	1000
<i>pānž/pānĵ</i>	<i>pānĵēs</i>	<i>pānčās</i>		
<i>šeš</i>	<i>šīyēs</i>	<i>šeštī</i>		
<i>ħewt</i>	<i>ħewtēs</i>	<i>ħewtāy</i>		
<i>ħešt</i>	<i>ħeštēs</i>	<i>ħeštāy</i>		
<i>new</i>	<i>newēs</i>	<i>newāy</i>		

Other numbers are additive, using *ū* ‘and’: *vīst ū zew* ‘twenty-one’, etc.

The cardinal numbers immediately precede the noun they qualify. With most numbers, the noun takes the plural ending:

di ʿesker[-ī]-DIR-p bār ke-n-ē *ez do hīrē estor[-ān-]-OBL-p vež-ā*
 ‘two soldiers are loading’; ‘I shall get out three horses’;

nāme-y-ē nē ħewt birā-y[-ān]-OBL-p
 ‘the name of these seven brothers’.

With multiples of 10 (between 20 and 100) and multiples of 100, nouns more often remain in their singular form:

plural,	singular,
<i>ħewtāy dew[-ān]-OBL-p, -agent werd-i-f</i>	<i>vīst merdim-s kiš-y-en-o-3s</i>
‘70 camels have eaten (f, implied is a melon)’;	‘20 men get killed’.

Enumerators assume the grammatical marking, here direct object:

[hīrē čivāl-ī āltūn-ī]tir-en-ē *ti [di se-y kāyit pere-y] bi-d-ē mi*
 ‘they steal [three bags (of) gold]’; ‘(if) you give me [200 Lira (of) money]’.

3.1.6.2 Ordinal numbers

The ordinal numbers do not seem to be uniform within the Çermik-Siverek dialects of Zazaki. While according to Todd (1985: 83) they are derived by suffixing *-(i)n* to the cardinal number, for the dialect NX this is true only for the numbers ‘2’, ‘3’, and ‘10’: *did-in*, *hūr-in*, *des-in*, while the remainder are derived by *-ī*: *pānj-ī* ‘5th’, *ḥewt-ī* ‘7th’, *ḡondes-ī* ‘11th’, *pānjēs-ī* ‘15th’.

3.1.6.3 Quantifiers

Like the cardinal numbers, most quantifiers occur uninflected before the noun which they qualify:

bol ‘much, many, lot of’, *bol perey qezenč ke-n-o* ‘he earns a lot of money’;
čend ‘several, some’, also ‘how many’, *čend meng-ī tepiyā* ‘after some months’;
hergi ‘each, every’, *hergi meng-i* ‘every month’;
pēro ‘all’, *pēro od-ān-rā* ‘through all rooms’.

Alternatively, *bol* may be inflected and follow the noun; similarly, *bol*, *pēro* and *heme* ‘all’ may also follow the noun, but in apposition and with case agreement only (but not number):

bi-b-ē *wēr-ē* *čī-y-ē bol-ī-OBL-S*
 ‘(that you) become owners of many things’.

3.2 Verb morphology

Verbs are inflected for person, number, tense, aspect, and mood, and show distinction in transitivity and voice. Further distinctions are made by aspectual-modal particles.

3.2.1 Stem formation

3.2.1.1 Subjunctive, past, and present stems

The Zazaki verbal forms are based on three stems: subjunctive, present, and past.

The subjunctive and past stems generally continue inherited Iranian present stems, while the present stems are derived from the Zazaki subjunctive stems by the formant *-(e)n*.

Past stems are mostly irregular and end either in a dental, with various degrees of changes: subj. *kiš-*, past *kišt-* ‘kill’; subj. *roš-*, past *rot-* ‘sell’; or they end in a long vowel (due to loss of final *-d*), e.g. subj. *d-*, past *dā* ‘give’. “Regular” past stems end in *-ā*, e.g. subj. *res-*, past *resā-* ‘arrive’.

Present stems show likewise certain irregularities, mostly loss of final *-r* in a small subset, e.g. subj. *ker-*, pres. *ke-n-*, past *kerd-* ‘do, make’, but also subj. *vāž-*, pres. *vā-n-*, past *vāt-* ‘say’.

Two verbs are suppletive: subj. *yer-*, pres. *ye-n-*, past *āme-* ‘come’ and subj. *vīn-*, pres. *vīn-en-*, past *dī-* ‘see’.

3.2.1.2 Causative and passive stems

Most verbs with subjunctive formant *-y-* are intransitive, and verbs with subjunctive ending in *-n* are transitive and regular, e.g. intransitive subj. *rey-*, pres. *rey-en-*, past

reyā- ‘escape’; *girwey-*, pres. *girwey-en-*, past *girweyā-* ‘work’ (< *girwe* ‘work’); transitive, subj. *birn-*, pres. *birn-en-*, past *birnā-* ‘cut’.

Similarly, morphological passive forms are marked by *-(i)y* or *-ey*, e.g. *vāž-* ‘say’ > *vāž-y-*; *tir-* ‘steal’ > *tir-iy-*; *vīn-* ‘see’ > *vīn-ey-*.

Causative stems are derived by *-n*, rarely *-ān*, e.g. *geyr-* ‘go around’ > *geyr-n-* ‘show around’; *ters-* ‘be afraid’ > *ters-ān-* ‘frighten’.

Both causative and passives may be derived from the same stem, e.g. *‘ediz-y-* ‘tire’ vs. *‘ediz-n-* ‘(cause to) tire’; *šik-y-* ‘break’ vs. *šik-n-* ‘(cause to) break’.

3.2.1.3 Preverbs

Generally, the preverbs modify the meaning of the simple verb. There are two kinds, “true” preverbs and “absolute” circumpositions, such as the following:

- ā-* ‘back, open’, e.g. *ā-geyr-* ‘come back’; *ā-b-lā-ker-* ‘open’ (intr./tr.), *ā-birn-* ‘separate’;
- rā-* ‘back, open, forth’, e.g. *rā-čār-n-* ‘bring back’, *rā-b-l rā-ker-* ‘be opened, open (tr.)’, *rā-kew-* e.g. ‘lie down, go to sleep’;
- ro-* ‘down’, e.g. *ro-nen-* ‘put down’, *ro-qilot-nen-* ‘swallow’, *ro-šen-* ‘sit (down)’;
- de-* ‘into’, e.g. *de-ker-* ‘put into’, *de-kew-* ‘fall into’, *de-niš-n-* ‘set up’;
- we-* ‘out’, e.g. *we-čīn-* ‘choose’, *we-ker-* ‘kindle’ (‘bring out [fire]’).
- pā* ‘at, onto’ (< *bi-ā*), e.g. *pā nān-* ‘place (at/onto)’;
- pirā* ‘on’ (< *bi-rā*), e.g. *pirā d-* ‘put on’.

While true preverbs always precede the verb, e.g. *šew-ē, no [rā-]kew-n-o* ‘one night, he sleeps’, absolute circumpositions precede the verb if subjunctive and imperative, but otherwise follow it, e.g. *čīnā-y-o pāk [pirā] dē* ‘put on clean clothes’, but *čīnā-y dā-n-ē [pirā]* ‘they put on the clothes’.

Given that the absolute forms originate in empty circumpositions, an overt noun or pronoun takes the place of the prepositional component, e.g. *xo* ‘self’ in *čīnā xo-rā di* ‘put on (some) clothes’.

3.2.1.4 Compound verbs

Compound verbs typically consist of a nominal component and a small set of verbs that carry the inflection. Most basic are *b-* ‘be, become’ and *ker-* ‘do, make’ for intransitive and transitive function, e.g. *hewn vīn-* ‘dream’, lit. ‘see a dream’. To many English simple verbs there correspond Zazaki combinations of nominal complements with a subset of verbs. The nominal components may function as:

- (1) direct object: *zūr-ī ker-* ‘make lies’ = ‘lie’; *veyn d-* ‘give a call’ = ‘call’;
- (2) allative: *žā yer-* ‘come to the place’ = ‘be fulfilled’;
- (3) postpositional complement: *hewn-ā š-* ‘go with-sleep’ = ‘fall asleep’.

Similar to these is a small set of *ezāfe* constructions with dependent “logical object” which as a whole likewise function as syntactic direct objects, e.g.

- goštār-ey-ā N ker-* ‘make hearing of N’ = ‘listen to N’,
- or allative, e.g. *gāzī-d-ā N yer-* ‘come to the help of N’.

A third, still smaller group consists of compounds of ‘be’ with “*miš-* participles” borrowed from Turkic, e.g. *diš-miš b-* ‘make thinking’ = ‘think’.

Causatives compounds are formed with the verb *d-* ‘give’ + infinitive in *-en*, e.g.

mi sere-y-ē ĵi dā ĵi-kerd-en-i
‘I had his head cut off’.

Passives are formed with *yer-* ‘come’.

3.2.2 Nominal forms

3.2.2.1 Infinitives

Two infinitives, ending in *-(y)iš* (m) and *-(y)en(-i)* (f), are derived from the past stem, e.g. *geyrā-y-iš* ‘to go around’, *rā-kewt-iš* ‘go to sleep’; *ĵi-kerd-en(-i)* ‘to cut off’. Both infinitives are freely exchangeable in NX.

3.2.2.2 Past participle

Past participles consist of the past stem + endings: *-e* (m), *-ā* (f), *-ey* (p).

3.2.2.3 *miš*-participle

There is also a *miš*- participle borrowed from Turkic.

3.2.3 Person marking and ‘be, become’

3.2.3.1 Person marking

In initial overview, the person markers in the various tenses and moods are shown in the following table:

TABLE 9.8: PERSON MARKING

	1s	2sm	f	3sm	f	1p	2p	3p
COP	<i>-ā</i>	<i>-ē</i>	<i>-ā</i>	<i>-o</i>	<i>-ā</i>	<i>-ē</i>	<i>-ē</i>	<i>-ē</i>
PRES	<i>-ā</i>	<i>-ē</i>	<i>-ā</i>	<i>-o</i>	<i>-ā</i>	<i>-ē/-im</i>	<i>-ē</i>	<i>-ē</i>
SUBJ	<i>-ā(n)</i>	<i>-ē</i>	<i>-ā</i>	<i>-o</i>	<i>-o</i>	<i>-ē(-mi)</i>	<i>-ē</i>	<i>-ē</i>
IMPR		<i>-i</i>					<i>-ē</i>	
PRET	<i>-ā</i>	<i>-ī</i>	<i>-ā</i>	<i>-ø</i>	<i>-i</i>	<i>-ī/-ī-m(i)</i>	<i>-ī</i>	<i>-ī</i>
PERF	<i>-ā</i>	<i>-ē</i>	<i>-ā</i>	<i>-o</i>	<i>-ā</i>	<i>-ē</i>	<i>-ē</i>	<i>-ē</i>
PPRF	<i>-(i) bī-y-ā</i>	<i>-ī bī</i>	<i>-i bī-y-ā</i>	<i>-(i)-bīli</i>	<i>-i bī</i>	<i>-ī bī</i>	<i>-ī bī</i>	<i>-ī bī</i>

The 1p has both the generalized plural endings *-ē*, *-ī* and *-iml-mi*, either alone or in various combinations (*-iml-mi* are rarely found in NX).

3.2.3.2 Existential verb

The existential verb is *est-*. In NX only 3rd pers. forms are found:

	3sm	3sf	plural
PRES	<i>est-o</i>	<i>est-ā</i>	<i>est-ē</i>
PRET	<i>est-(i) bī</i>	<i>est-i bī</i>	<i>est-ī bī</i>

All other tenses of the copula and the existential verb are supplied by the verb *b-* 'be, become'.

The stative meaning 'be' is distinguished from non-stative 'become' in the subjunctive and past: subj. *b-* vs. *bīy-* 'shall be/become'; past *bī-* vs. *bīy-ā-* 'was/became':

subjunctives:

mā do wext-o ārē [bī-y-ē]
'then we will (*do*) gather' ('get together');

past:

xīznā ā-[bī-y-ā-y-ā]
'the treasury was ['became'] opened' (*ā-b-*).

3.2.4 Aspectual-modal and negative affixes and particles

3.2.4.1 Modal *bi-*

The modal prefix *b(i)-* marks two sets of moods, the subjunctive and imperative (subjunctive stem), and the past subjunctive (imperfect past in *-ē*). There are various contractions with verb-initial components: *bīr-* 'come'; *b-ānǰ-* 'draw'; *bīr-* 'eat'; note *bi-wāz-* 'want'. *b(i)-* is superseded in the present subjunctive and imperative by a preverbal particle or nominal (*ārē-d-* 'collect' *bār ker-* 'load'), and is lacking irregularly with some other (mostly monosyllabic) verbs (*ber-* 'carry', *res-* 'arrive', *šir-* 'go', *vāž-* 'say', *vīnder-* 'stand').

3.2.4.2 Negative prefixes *ne-* and *me-*

The negative prefix is *nē-* (*nī-y-* before vowel), the prohibitive prefix is *me-*; both supersede *b(i)-*. The negative forms of the copula have the prefix *nīy-*; those of the existential verb have *či-nēlī-*:

pī-y-o zey to wā či-nē-b-o
'a father like you (are) should not (*wā nē*) exist'.

3.2.4.3 Imperfective particle *-ē*

The invariable enclitic *-ē* is attached to the subjunctive past to mark the imperfect. Combined with *b(i)-* it marks the past subjunctive.

3.2.4.4 Future particle *do*

Proposed *do*, often immediately following the subject, marks the simple future in combination with the present subjunctive, and the past conditional with the past subjunctive.

3.2.4.5 Optative particle *wā*

Proposed *wā* marks the (present) optative in combination with the present subjunctive, and the past optative with the past subjunctive (*b(i)- + ē*).

3.2.5 System of tense, mood, and aspect

Zazaki has two basic tenses, present and past, and two basic aspects, imperfective and perfective. Moods include indicative, subjunctive, imperative, and optative (“conditional”), expressing various functions. Transitivity is distinctly marked in past tenses (see section 4.4.1.1 Ergative).

3.2.5.1 Forms based on present stem

3.2.5.1a Present indicative

The present indicative is formed from the present stem, as shown in the following paradigms:

TABLE 9.9: PRESENT INDICATIVE

	<i>ker-</i> ‘do’	<i>d-</i> ‘give’	<i>wer-</i> ‘eat’	<i>šir-</i> ‘go’
1s	<i>ke-n-ā</i>	<i>dā-n-ā</i>	<i>we-n-ā</i>	<i>ši-nā</i>
2sm/f	<i>ke-n-ē/ke-n-ā</i>	<i>dā-n-ē/dā-n-ā</i>	<i>we-n-ē/we-n-ā</i>	<i>ši-n-ē/ši-n-ā</i>
3sm/f	<i>ke-n-ol/ke-n-ā</i>	<i>dā-n-ol/dā-n-ā</i>	<i>we-n-ol/we-n-ā</i>	<i>ši-n-ol/ši-n-ā</i>
1p	<i>ke-n-ē (ke-n-im)</i>	<i>dā-n-ē</i>	<i>we-n-ē</i>	<i>ši-n-ē (ši-n-im)</i>
2p, 3p	<i>ke-n-ē</i>	<i>dā-n-ē</i>	<i>we-n-ē</i>	<i>ši-n-ē</i>

3.2.5.1b Continuous present

The continuous present is formed from the present indicative and the proposed copula (but 1s has more often *-o* instead of *-ā*).

TABLE 9.10: CONTINUOUS PRESENT

1s	<i>ez olā</i>	<i>ke-n-ā</i>	1p	<i>mā</i>	<i>yē ke-n-ē</i>
2sm/f	<i>ti yē</i>	<i>ke-n-ē / ti yā ke-n-ā</i>	2p	<i>šimā</i>	<i>yē ke-n-ē</i>
3sm/f	<i>o yo</i>	<i>ke-n-o / ā yā ke-n-ā</i>	3p	<i>ē</i>	<i>yē ke-n-ē</i>

3.2.5.1c Present subjunctive

The present subjunctive is formed from the subjunctive stem with the prefix *b(i)-*. It is missing before compound verbs and a small number of simple verbs. The final *-n* in the 1s occurs occasionally; the 1p variants *-(ē)-mi* and *-im-ē* occur only rarely and irregularly in NX. Used independently, the subjunctive is used as an “exhortative” or “deliberative” forms for the 1s, 1p.

TABLE 9.11: PRESENT SUBJUNCTIVE

	‘go’	‘give’	‘eat’
1s	<i>šir-ā(n)</i>	<i>bi-d-ā</i>	<i>būr-ā</i>
2sm/f	<i>šir-ē/šir-ā</i>	<i>bi-d-ē/bi-d-ā</i>	<i>būr-ē/būr-ā</i>
3s	<i>šir-o</i>	<i>bi-d-o</i>	<i>būr-o</i>
1p	<i>šir-ē ~ ši-mi</i>	<i>bi-d-ē ~ bi-dim-ē</i>	<i>būr-ē ~ būr-ēmi</i>
2p, 3p	<i>šir-ē</i>	<i>bi-d-ē</i>	<i>būr-ē</i>

3.2.5.1d *Future*

The future preposes the particle *do* to the subjunctive: *e(z) do šir-ā, ti do šir-ē* ‘I, you will go’.

3.2.5.1e *Optative*

The optative preposes the particle *wā*: *wā šār nē-vīn-o* ‘may the people not see (it)’.

3.2.5.1f *Imperative*

The imperative is formed from the subjunctive stem with the prefix *bi-*, neg. *me-*, and the personal endings 2s *-i*, 2p *-ē*. Irregular verbs whose subjunctive stem ends in *-r* have often irregular and unstable imperative forms, e.g. *bi-gīr(-i)* ‘take!’, *me-gī(r)* ‘don’t take!’.

3.2.5.2 Forms based on past stem

The following shows the inflections and the personal endings of intransitive verbs, which also mark the past patient in the ergative construction (see section 4.4.1 Use of cases).

3.2.5.2a *Preterite*

The preterite (simple past) is formed from the past stem and the appropriate personal endings. The 1p endings with *-m-* occur only spontaneously in most Çermik-Siverek dialects (similar to the present). The ending *-ē* of the 3sf is irregular.

TABLE 9.12: INTRANSITIVE PRETERITE

	‘fall’	‘go around’	‘come’	‘be’
1s	<i>kewt-ā</i>	<i>geyrā-y-ā</i>	<i>āme-y-ā</i>	<i>bī-y-ā</i>
2sm/f	<i>kewt-īkewt-ā</i>	<i>geyrā-y/geyrā-y-ā</i>	<i>āme-y/āme-y-ā</i>	<i>bī/bī-y-ā</i>
3sm/f	<i>kewt/kewt-i</i>	<i>geyrā/geyr-ē</i>	<i>āme-elām-ē</i>	<i>bī, bī/bī</i>
1p	<i>kewt-ī ~ kewt-īm</i>	<i>geyrā-y</i>	<i>āme-y ~ āme-y-mi</i>	<i>bī ~ bī-mi</i>
2p, 3p	<i>kewt-ī</i>	<i>geyrā-y</i>	<i>ām-e-y</i>	<i>bī</i>

3.2.5.2b *Imperfect*

The imperfect tense is formed from the past stem and the invariable suffix *-(y)ē* for all persons. Accordingly, the plural forms of the imperfect are formally identical with those of the perfect, and may be distinguished from the latter only by the context. The imperfect gives the sense of a repeated or habitual action, or continuous state in the past:

meng-i-rā finē žī ez [āme-y-ē], mi jī-rē tāy čī-y-ē werd-ī
 a month once also I [came] I-OBL-AG for him some things of eating
 [ārd-ē]
 [brought]

‘once a month I came . . . and I brought him some things to eat’;

hewā-rā kutik [lāwā-y-ē]

‘a dog (*kutik*) [was (just) barking] from the air’.

3.2.5.2c *Past subjunctive*

The past subjunctive, *b(i)-* + imperfect, expresses impossible conditions in the protasis of conditional clauses (see section 5.5.2 Conditional clauses).

3.2.5.2d Past conditional

The past conditional preposes *do* to the past subjunctive. It expresses impossible conditions in the apodosis.

3.2.5.2e Past optative

The past optative preposes *wā* to the past subjunctive. It expresses impossible past conditions in the apodosis, and unfulfillable wishes, e.g.

wā ez bi-merd-ē
‘(oh), if only I had died’.

3.2.5.3 Perfect forms

3.2.5.3a Present perfect

The present perfect (Perfect I) is formed from the past stem and the copula. Plural persons are not distinguished.

TABLE 9.13: INTRANSITIVE PERFECT

	‘fall’	‘flee’	‘go’
1s	<i>kewt-ā</i>	<i>remā-y-ā</i>	<i>šī-y-ā</i>
2sm/f	<i>kewt-ē / kewt-ā</i>	<i>remā-y-ē / remā-y-ā</i>	<i>šī-y-ē / šī-y-ā</i>
3sm/f	<i>kewt-o / kewt-ā</i>	<i>remā-y-o / remā-y-ā</i>	<i>šī-y-o / šī-y-ā</i>
plural	<i>kewt-ē</i>	<i>remā-y-ē</i>	<i>šī-y-ē</i>

Basically, the perfect tense is used to denote a past action that is connected, through its result, with the present moment, e.g. *eʃel-ē to [āme-y-o]* ‘your last hour [has come]’. In usage, however, the perfect is no longer clearly distinguished from the preterite. Sometimes, both are used in one and the same context, or the perfect is used where a preterite would be expected:

to cī-y-ē [dī-y-o] yā cī-y-ē āme-y-o āqil-dē to
you-OBL-AG anything seen-PERF-3ms or anything come-PRET to mind of you
‘[have you seen] anything, or [has] anything [come] to your head?’.

3.2.5.3b Non-witnessed function

Another characteristic usage of the perfect (maybe developed recently, following the Turkic “*miş*-tense”) is that as “narrative”, indicating that the speaker has not witnessed what he reports:

vā-n-ē, dēw finā [āme] (preterite in “aorist” function)
‘they say: “The Diw has come again” (in the sense of ‘he is here’)

vs.

“*dēw [āme-y-o]* (perfect in non-witnessed function)
‘(they report to Heso:)
“The Diw has come (reportedly)” ’.

3.2.5.3c *Perfect II*

While the past participle is used mostly as an adjective, three verbs whose present tense connotes directed motion of living beings (*rā-kew-* ‘go to sleep’, *ro-sen-* ‘sit down’, *vinder-* ‘stop, stand’) form a second perfect tense that expresses a temporary state, e.g. *dār-i bin-dī žew lāžek* [*rā-kewt-e y-o*] ‘under the tree, a boy [is sleeping]’; *keynek-ēn ā ro-ništ-ā* ‘a girl [is sitting]’. By contrast, the simple perfect of these verbs expresses the act of motion.

3.2.5.3d *Pluperfect*

The pluperfect seems to be quite non-uniform in the various Zazaki dialects. In NX, it is formed for most persons from the inflected preterite, together with the likewise inflected preterite of the auxiliary verb ‘be’:

pī-y-ē ez polāt [*nē-kerd-i-bī-y-ā*]
father of me I-PAT steel had not made-1s
‘my father had not steeled me’;

mi sond werd-i bī mā piyā āme-y bī
‘I had sworn’ ‘we had come together’.
(cf. Persian *sowgand xordan*);

In the 1s and 2sf, the inflected auxiliary is attached to the uninflected past *stem* of the main verb. The optional *-(i)-* of the main verb in the 1s, 3sm is euphonic.

TABLE 9.14: INTRANSITIVE PLUPERFECT

	‘had remained’			‘had arrived’	
1s	<i>mend-(i)</i>	<i>bī-y-ā</i>		<i>resā</i>	<i>bī-y-ā</i>
2sm/f	<i>mend-ī</i>	<i>bī</i>	<i>mend-i bī-y-ā</i>	<i>resā-y</i>	<i>bī</i>
3sm/f	<i>mend-(i)</i>	<i>bi, bī</i>	<i>mend-i bī</i>	<i>resā</i>	<i>bi, bī</i>
1p	<i>mend-ī</i>	<i>bī(-mi)</i>		<i>resā-y</i>	<i>bī(-mi)</i>
2, 3p	<i>mend-ī</i>	<i>bī</i>		<i>resā-y</i>	<i>bī</i>

3.2.5.4 Overview of system of tenses

In overview, the tense and mood system of Zazaki may be summarized in the following table, using 3sm, 2s forms of *geyr-* ‘go around’ and *ro-ništ-* ‘sit down’.

TABLE 9.15: SYSTEM OF TENSES AND MOODS: *geyr-* ‘go around’, *ro-ništ-* ‘sit down’

	INDICATIVE	NON-INDICATIVE	
		<i>bi-geyr-i 2s</i>	IMPR
PRES	<i>geyr-en-o</i>	<i>bi-geyr-o</i>	SUBJ
CONTIN	<i>o geyr-en-o</i>	<i>do bi-geyr-o</i>	FUT
		<i>wā bi-geyr-o</i>	OPT
PRET	<i>geyrā</i>	<i>bi-geyrā-y-ē</i>	PT SUBJ
IMPF	<i>geyrā-y-ē</i>	<i>do bi-geyrā-y-ē</i>	PT COND
		<i>wā bi-geyrā-y-e</i>	PT OPT
PERF-I	<i>geyrā-y-o</i>		
PERF-II	<i>ro-ništ-e-y-o</i> ‘he is sitting’		
PPERF	<i>geyrā bī</i>		

4 SYNTAX AND USE OF FORMS

4.1 Coordinating conjunctions and clitics

The most important coordinating conjunctions include: *ū* ‘and’; *yā(n)* ‘or’; enclitic *zī* ‘also’; *finā zī* ‘nevertheless’; *bāhd-o (zī)* ‘then, after’, *wext-o* ‘then’; *qānd-ē* *joy* ‘therefore’, *yānī* ‘that is’; *emā* ‘but’.

4.2 Noun phrase structure

Demonstratives and numbers generally precede the head noun; dependent nominals follow it.

4.2.1 Ezāfe construction and case marking

Dependent nominals are connected to their head noun by enclitic particles (EZ) that distinguish gender, number, and case. Adjectival markings differ slightly in the 3sm. Adjectives assume the grammatical marking of their head noun; dependent nouns are in the basic oblique case (OBL₁).

The marker *d-* distinguishes *ezāfe* constructions that function as secondary syntactic cases (OBL₂), that is, any cases other than the syntactic subject/agent and direct object/patient. Constructions with allative function and with primary postpositions (ALL/POST) have optional *d-*, except 3s *do*.

TABLE 9.16: EZĀFE PATTERNS

	<i>Adjectival</i>			<i>Nominal</i>		
	Head	EZ	ADJ-	Head	EZ	N-OBL
DIR _m	N- <i>o</i>	- <i>o</i>	ADJ- <i>o</i>	N- <i>o</i>	- <i>ē</i>	
OBL _{1m}	N- <i>o</i>	- <i>ē</i>	ADJ- <i>ī</i>	N- <i>o</i>	- <i>ē</i>	
OBL _{2m}	N- <i>o</i>	- <i>dē</i>	ADJ- <i>ī</i>	N- <i>o</i>	- <i>dē</i>	
ALL/POST _m	N- <i>o</i>	- <i>do</i>	ADJ- <i>ī</i>	N- <i>o</i>	-(<i>d</i>) <i>ē</i>	
DIR _p	N- <i>o</i>	- <i>ē</i>	ADJ- <i>ī</i>	N- <i>o</i>	- <i>ē</i>	
OBL _{1p}	N- <i>ān</i>	- <i>ē</i>	ADJ- <i>ān</i>	N- <i>ān</i>	- <i>ē</i>	
OBL _{2p}	N- <i>ān</i>	- <i>dē</i>	ADJ- <i>ān</i>	N- <i>ān</i>	- <i>dē</i>	
ALL/POST _p	N- <i>ān</i>	-(<i>d</i>) <i>ē</i>	ADJ- <i>ān</i>	N- <i>ān</i>	-(<i>d</i>) <i>ē</i>	
DIR _f	N- <i>o</i>	- <i>ā</i>	ADJ-(<i>i</i>)	N- <i>o</i>	- <i>ā</i>	
OBL _{1f}	N-(<i>er</i>)	- <i>ā</i>	ADJ-(<i>i</i>)	N-(<i>er</i>)	- <i>ā</i>	
OBL _{2f}	N-(<i>er</i>)	- <i>dā</i>	ADJ-(<i>i</i>)	N-(<i>er</i>)	- <i>dā</i>	
ALL/POST _f	N-(<i>er</i>)	-(<i>d</i>) <i>ā</i>	ADJ-(<i>i</i>)	N-(<i>er</i>)	-(<i>d</i>) <i>ā</i>	

Words ending in a vowel usually insert a glide *-y-* before an *ezāfe* vowel, with optional contraction of masc. *-e-y-ē* and fem. *-ā-y-ā*: *nāme-y-ē* ~ *nāmey* ~ *nāmē*; *keynā-y-ā* ~ *keynā*.

4.2.1.1 Adjectival *ezāfe* constructions

(1) Primary cases

DIR:	OBL _{1m} , direct object:
<i>birī-y-o gird</i>	[<i>embāz-ē bīn-ī</i>] <i>vīn-en-ō</i>
‘the big brother’;	‘he sees the other friend’.

(2) Secondary cases, OBL

(a) direct object: *veyn dā-n-ā [legleg-ān-dē-p bīn-ān]* ‘(she) calls [the other storks]’; (b) adposition: *hetē [’esker-ān-dē-p bīn-ān]-ā* ‘to the other soldiers’.

(c) allative: *nobet-i ye-n-ā [keynek-er-dā-fwerd-ī]* ‘the turn comes [to the little girl]’.

(3) Primary postposition, ex. without *d-*:

[jā-y-ān-ē-p pāk-ān]-di ‘at (-*di*) clean places’.

Certain adjectives that obligatorily require qualification by a noun may take this with the *ezāfe*:

piš-ē gūnī ‘full of blood’; *keynā mi [lāyiq-ē to] nī-y-ā* ‘my daughter is not [worthy of you]’.

4.2.1.2 Nominal *ezāfe* constructions

(1) Primary cases

(a) direct case: *beg-ē diyārbekir-ī-OBL* ‘the Beg of Diyarbakir’; (b) direct object: *lāžek [mirīčik-er-ā-f xo] gē-n-o* ‘the boy takes [his bird]’;

(c) past subject/agent: *o ki [pī-y-ē-m šimā] mā-rē kerd-o* ‘what [your father] has done to us’.

(2) Secondary cases

(a) indirect object: *ārweš-ī mis-n-en-o [lāž-dē-m xo]* ‘he shows [his son] the hare’; (b) adverbial cases: *xātir-dē (m) to-OBL* ‘good bye’ (‘for your memory’);

(c) allative: *ver-ē xo dā-n-ē [āyre-y-ē (m) āw-i-OBL-f]* ‘they proceed [to the water mill]’ (‘mill of water’); (d) primary adposition: *[bizēk-dā-f mā-di] ey čičī dī-y-o?* ‘what has he seen [in (-*di*) our kid]?’

(3) ‘Kinship-*r*’

(a) primary case: *[birā-y-ē xo] ber-ē wižā* ‘bring [your brother] there’; (b) secondary case: *[birā-r-dē ji]-rā* ‘from his brother’.

4.2.1.3 Group inflection

Hendiadys connected by *ū* 'and' often inflect the second element only:

- (1) Adjective:

(ṭeyr ū ṭūr)-o-m bīn 'the other birds';

- (2) Dependent noun (with both elements inflected):

xeber-i dā-n-ē [(mā-r ū pē-r)-dē-p bizēk-er-i]
'they inform [(the mother and father = parents) of the kid]'.
'they inform [(the mother and father = parents) of the kid]'.

4.2.1.4 Multiple *ezāfe* constructions

In multiple *ezāfe* constructions, where the head noun (N_1) is followed by a combination of dependent nouns and adjectives, each qualifier is preceded by its appropriate *ezāfe*. Basically, secondary *ezāfe* constructions normally should be in OBL and marked by *d-*, but may be not under certain conditions.

- (1)
- $N_1 + N_2 + ADJ_1$

An adjective qualifying the head noun (ADJ_1) is invariably in final position. It attracts the EZ of the head noun, and is linked to N_2 without *d-*, whether the whole phrase is in the primary or secondary case:

- (a) subject:

wā [o āyā (-y-ē šimā)-y-o bēnāmūs] b-īr-o
'may [that infamous Agha of you] come';

- (b) allative:

šī-n-ē [ā dew (-dā xidir āyā)-y-ā-f bīn-i]
'they go [to that other village of Xidir Agha]';

- (c) agent:

[lāž (-ē min)-ē werd-ī-OBL₁-m] vāt-i bī
'[my small brother] had said'.

- (2)
- $N_1 + N_2 + N_3$
- or
- ADJ_2

A further dependent noun (N_3) or adjective qualifying N_2 may or may not be marked by *d-* if the noun phrase as a whole is in a primary syntactic case:

- (3)
- $N_1 + N_2 + ADJ_2$

- (a) with
- d-*
- :

hetē [nē 'esker-ān-(dē pādīšāh-dē bīnī-y)]-ā rem-en-ā
'(it) flees to the soldiers (of the other king)';

- (b) without
- d-*
- :

res-en-ē [tāy 'esker-ān-(ē pādīšā-y-ē bīn-ī)]
'they reach [some soldiers (of the other king)]'.

(4) $N_1 + N_2 + N_3$

- (a) with *d-*: direct object,
 NP, *este-[y-ē gān-(dē mi)] zīn-[ē estor-ān-(dē xo)] girē dā-n-ē*
 ‘the bones of my body’; ‘they tie up the saddle [of their (*xo*) horses]’;

- (b) without *d-*:
 subject, *qeček-[ē embiri-y-ān-(ē mā)] āme-y*
 ‘the children [of the neighbor of us] came’;
 direct object, *heq-[ē rāy-(ā to)] ez dā-n-ā*
 ‘I pay your fare’ (‘fare [of road of you]’).

4.2.1.5 Indefinite head noun

Indefinite *ezāfe* construction, marked by *-ē(n)* or *-nā*, always insert *d-*:

(1) Adjectival noun phrases

- (a) predicative:
kāmōre [lāžek-ē-do bol jāmērd] o
 ‘K. is [a very brave boy]’;

- (b) direct object:
[pirēn-ē-do diřāt]-e dā-n-o xo-rā
 ‘he puts on [a torn shirt]’;

- (c) agent:
[keynek-ē-dā zey āšmī] govend-i ānt-ē
 ‘[a girl like the moon] has danced the round dance’.

(2) Genitival noun phrases

- (a) subject:
[derd-ē-dē to] est-o
 ‘there is a pain of yours’ =
 ‘you have a pain’;

- (b) with *-nā*, direct object:
[birā-r-nā-dē nīn-ān] kišenē
 ‘they kill (another brother of these)’.

4.2.1.6 Inalienable possession

With body parts or articles of dress in the sense of inalienable possession, the possessor pronoun and *xo* often precede the head noun unconnected if the phrase has a locative or allative sense:

kewt [mi dest]
 ‘(it) has fallen [(in) my hand]’ (not *kewt dest-ē mi*);

miričik-i dā-n-o [in-ān dest] *[xo ser-ī miyān-di] plān-ān benā ā-n-ā*
 ‘he gives the bird [(into) their hand]’; ‘[inside (her) own head] she makes plans’.

4.2.1.7 Elliptic *ezāfe* forms

The head noun of both the nominal and adjectival *ezāfe* may be elided, giving the group a possessive sense, or a substantival sense:

no lāžek [ē nāhmān beg-ī] nī-y-o *[o bīn] vā-n-o*
 ‘this boy is not [(the one) of Nāhmān Beg]’; ‘[the other (one)] says’.

4.3 Clause structure and word order

Every simple sentence contains a subject and a predicate. The other details of its structure depend mainly on the valence of the verb that is part of the predicate: direct object; indirect object; complements (that depend on the valence of the verb), either a simple case, or adpositional.

For the word order of the main constituents the following rules apply:

Subject + Object + Verb, specifically Direct Object + Verb.

4.3.1 Pre- and postverbal placements

Dependent complements without postposition follow the predicate; those with postpositions may take any post-subject position (but for most verbs there is one characteristic position in which they occur). Thus, indirect objects marked by *-rē* precede, those without it follow the VP. These rules apply to past ergative sentences respectively.

4.3.1.1 Preverbal placement

ez-DIR [āy-ā] zewž-iy-ā-y-ā *tī yē [mi-rā] čičī wāz-en-ē?*
 ‘I [with her] got married’; ‘you [from me] what you want?’;

‘erebū z. činā *[xo-rā] vež-en-ā*
 ‘Erebu Z. the clothes [from self] takes’.

4.3.1.2 Postverbal placement

(1) Oblique cases, allative

lāžek šī-n-o [dew-dā xo]
 ‘the boy goes [to his village]’;

(2) Postpositional phrases

mā-yā lāžek-ī ūr-ā nān-ī dān-ā [sere-dē bizēk-er-i-ro]
 ‘the mother of the boy the rolling-pin throws [on the head of the kid]’.

4.3.2 Other conditions for word order

4.3.2.1 Omission of *bi-*

The indirect object or complement that would regularly follow the verb, precedes it when the subjunctive form lacks *b(i)-* (subjunctive, future, imperative), mostly with *d-* ‘give’:

ez do keyn-er-ā xo [to]-OBL dā
 ‘(then) I FUT daughter my [you] give’;
nē-ver-dā-n-ā [jī-ro] pēš-ī-y-o
 ‘she doesn’t allow (him) [(on her)self] (that) he embrace’.

4.3.2.2 Adverbial complements

Free adverbial complements which regularly precede the verb, follow the copula when the subject is pronominal, in which case the copula may be repeated in final position:

o yo [to heti] mā yē hewnā [tiyā-n-ān-di] nē yē nikā [kotī-dir] ē?
 ‘he is [with you]’; ‘we still are [in this region]’; ‘[where] are they now?’.

4.3.2.3 Topicalization

A subject may be topicalized by following the direct object:

heq-ē rāy-ā to [ez] dā-n-ā
 ‘your travel expense [I] will give’.

4.4 Semantics and use of nominal forms

4.4.1 Use of cases

4.4.1.1 Ergative

Zazaki shows the typical tense-sensitive crosswise ergative pattern of the primary cases, and verb agreeing with the DIR case:

	Subject	Direct Object	Verb Agreement
Present	DIR	OBL	DIR
Past-trans.	OBL	DIR	DIR

Also DIR are the non-specific indefinite direct object in the present, the intransitive subject in the past, and the predicative complement. Secondary syntactic cases are in the OBL. These include the indirect object (“dative”), the local complement of a verb of motion (“allative”); and the temporal nominal adverb.

Inanimate singular nouns, however, tend to show the direct case when in allative function, or when dependent on the primary postpositions *-di(r)* (always), and *-ā* and *-rā* (often). Examples for direct:

ti do žī [hepis-di-DIR_m] bi-mir-ē
 ‘you will also die [in prison]’.

The scale of oblique marking in animate vs. inanimate singular masculine nouns is as follows:

TABLE 9.17: SCALE OF SINGULAR OBLIQUE MARKING

	anim.	inanim.
Subject, patient, indef. object, predicate	–	–
Postposition <i>-di(r)</i>	+	–
Postpositions <i>-ā, -rā, -rē, -ro</i> ; allative	+	+/-
Agent, dative	+	not attested
Def. object., genitive, comparison	+	+
Secondary adpositions, adverb	+	+

4.4.1.2 Verbal rection and transitivity

Some Zazakî verbs govern their complement in a way different from their English equivalent. For example, in verbs such as *hes ker-* ‘love’, the “object” of the emotion is seen as the “source” thus *-rā* ‘from’ (historically ‘for sake of’, beneficiary):

ez o [keynek-ē-rā] hes ke-n-ā
‘I love [a girl]’, with the transitive verb *ker-* ‘do’.

Other Zazakî verbs with postpositions have similarly various ways to govern, whether the basic verb is transitive or not, e.g. *d-* ‘give’ or ‘hit’ + *ro* ‘on, down’ or *rā*:

ti weš dā-n-ē [mi-ro] žew-ser dā-n-ē [rem-ro]
‘you beat [(on) me] well’; ‘immediately they flee’ (‘hit [on flight]’);

poste-y-ē xo dā-n-ā [xo-rā]
‘she puts [on (to) her skin]’.

4.4.2 Subject marking

4.4.2.1 Impersonal subjects and expressions

Impersonal subjects (‘one, they’) may be expressed by the 3p:

mi-rē [vā-n-ē] “Šēr Osmān”
‘[they] call me “Šēr Osmān”’.

Impersonal expressions of the type ‘it is getting dark’, ‘it seems’, etc., are generally with ‘empty’ subject, e.g. *beno zimistān* ‘it becomes winter’:

wini [ās-en-o-3sm], no merdim-ēn-o xāyīn o
‘[it seems] (‘thus’), this man is a traitor’.

But expressions of precipitation require the use the respective nouns:

yāyer ū vew-ri [vār-en-ē]
‘it rains and snows’ (‘rain and snow [are raining]’).

3sm without copula are elliptic expressions like:

[weš ki] to mi-rē vā šikur Ellā-y-rē ki . . .
‘[(it is) good that] you told me’; ‘(may) thanks (be) to god that . . .’.

4.4.2.2 Non-specific past agents

Transitive past tenses occurring without agent may have *passive* sense:

no qundāy qānd-ē čičī [ārd-o]? *māhlā pēro [kerd-o bo]-y-ā*
 ‘what [has] this bundle [been brought] for?’; ‘the whole quarter [has been painted].’

Otherwise, the morphological passive is used:

win-ēn-ē ki āltūn-ī [tir-iy-ā-y-ē]
 ‘they observe that the gold [has been stolen].’

4.4.3 Object marking

4.4.3.1 Implicit direct object

Transitive verbs may occur without object (or one of their objects) if it is understood from the context:

to rind kerd *o vā-n-o: “ti virāz-i!”*
 ‘(this) you have done [it] well’; ‘(the) one says: “you prepare [it]!”’;

ez soz dā-n-ā *yān ti dā-n-ē mi*
 ‘I give (my) word’ ‘either you give (her) to me’.
 (no overt indirect object);

4.4.3.2 Compound verbs

Zazaki compound verbs such as *bār ker-* ‘load’, lit. ‘make load’, often have no simplex pendant, and thus function like simple verbs, and may govern direct objects:

kolī-y-ān [bār ke-n-ā] her-d-ē xo
 ‘she [loads] the firewood upon her donkey’.

Since no sentence can contain two direct objects, these constructions require a special explanation. Either the nominal complement has to be explained as a direct object within the verbal phrase itself, or the compound verb has to be considered a lexical unit that is not further analyzable (cf. Paul 1998c: 131 ff.).

4.5 Semantics and use of verb forms

4.5.1 Possession, ‘have’

Possession is expressed by an *ezāfe* construction together either with the existential verb and forms of *b-* ‘be’, or with the postposition *-rē* (“dative”):

di šārt[-ē mi] est-ē *[ji-rē] lāžek b-en-o*
 ‘I have two conditions’, ‘she has a boy’,
 lit. ‘there are two conditions [of mine]’; lit. ‘[to her] is being’.

4.5.2 Aspectual construction

4.5.2.1 Incipient action

This is expressed by *ker-* here ‘begin, be about to’ and *d- piro* ‘hit on’:

wext-o ki ti [ke-n-ē] bi-zewž-iy-ē
‘when you [make/set about] to get married (subj.);’

nē [dā-n-ē piro], ši-n-ē ki . . .
‘they [hit/set about] (and) go (ind.) to . . .’

4.5.2.2 Instantaneous action

This is expressed by *d-* ‘hit’ (similar to the Persian colloquial use of *zadan* ‘hit’):

žew-ser [dā-n-ē] rem-ro
‘immediately they flee’ (‘[hit on] flight’).

4.5.3 Modal constructions

The main modal auxiliary verbs are: *š-/šā-* ‘be able, possible’, *wāz-/wāšt-* ‘want’, and *wetār-/wetārd-* ‘dare’. Necessity is expressed by *lāzim* ‘necessary’ and *mājbur* ‘obliged’. They are followed by the subjunctive:

hewtāy dew-ān werd-i, finā žī [nē-šā] bi-qed-ēn-o
‘70 villages have eaten (of it), still though it could not be finished up’;
ez [wāz-en-ā] ey-ā bi-zewž-iy-ā
‘[I want] to marry him’;

[lāzim o] ki hešt sehāt-ī šir-ē mā [mejbur ē] to bi-riš-ē
‘[it is necessary] that they walk eight hours’; ‘we [have to] send you (there)’.

4.6 Sentence types and word order

4.6.1 Nominal sentences

Predicative sentences with a noun and copulative verb form a type of their own. These verbs may be intransitive (copula, *b-* ‘be, become’, *mān-* ‘remain, stay’) or transitive (*ker-* ‘make, render’, *silās-n-* ‘recognize as’). The nominal usually precedes the verb, but is postverbal when change of status is implied, specifically with *b-* in the sense of ‘become’ and *ker-* ‘do’:

(1) Preverbal:

no [bes] o hewn-ē ‘eyšā [rāšt] viž-y-ā
‘this is [enough]; ‘Ayše’s dream turned out (to be) [true];’

mērde-y-ē mi [šēx] qebūl kerd-o
‘he has accepted my husband [(as a) Shaikh]’.

(2) Postverbal:

zeřt̄-y-ā ji b-en-ā [nermi] lāžek-ī ke-n-o [pādīšāh]
 ‘his heart becomes [soft]’; ‘he makes the boy [king].’

4.6.2 Other types

Besides declarative sentences, there are interrogative, imperative, optative, and exclamatory sentences for all of which, in principle, the rules established above are valid.

4.6.2.1 Interrogative sentences

These may be distinguished from sentences of statement only by way of intonation:

to ey-rē žt̄ si,žil dā?
 ‘did you give him also a register?’

Interrogative pronouns are not in sentence-initial position (as in English), but in the position of the part of speech asked for:

ti do ji-rē [čičī] vāž-ē?
 ‘[what] will you say to him?’

4.6.2.2 Imperative sentences

These have either a simple subjunctive of the 1st person (exhortative), or an imperative that often (but not always) contains the subject:

(1) Subjunctive

mā [šir-ē] sūk-ē-d-ā gird-i
 ‘[let’s go] to a big city’.

(2) Imperative

xeber-i [bi-d-i] mi [ti] mi-rē heb-ē wext [bi-d-i]
 ‘inform me!’ (‘[give] me information’); ‘[you] [give] me a little time!’

4.6.2.3 Optative sentences

These have a subjunctive verb and one of the optative particles *wā*, or (less frequently) *bokā*. Any of the three persons may be the subject; with the present or past subjunctive, the desire is real or unreal respectively. *Wā*, but not *bokā*, usually stands in sentence initial position:

[wā] veyšān bi-mān-o ez [bokā] mudir b-ā
 ‘may (that) she stay hungry’; ‘may (that) I become director’;

[wā] ellā-y ez nē-dā-y-ē (past subj. -ē)
 would God I-DIR-PAT had not created-1s
 ‘would that God had not created (‘given’) me’.

4.6.2.4 Exclamative sentences

These are often introduced by *či* ‘what (a) N’:

[*či*] *pādišā-y-ē-d-o aḥmāq o!*
 ‘[what a] foolish king he is!’

5 COMPLEX SENTENCES

Zazaki complex sentences are composed of two (or more) constituent clauses that are either coordinate clauses, or main and subordinate clauses to which further clauses may be subordinated.

5.1 Coordinate clauses

A complex sentence consisting of two main clauses is not basically different from a sequence of two independent main clauses. Those two clauses may be connected without or with overt conjunction such as *ū* ‘and’, *yān* ‘or’. Semantically, there may be a logical, temporal, and other dependence between the two main clauses, which is either implicit:

ez do birā-r-d-ē xo-rē vāž-ā, ey ez dā-y-ā ney
 I FUT brother- my tell he-AG-3s I-PAT-1s has given this
 ‘I shall tell my brother, (because) he has given me to this one’;

or is made explicit by one of a series of coordinating conjunctions:

nē pēš-iy-ē-nē lāžek-ī-ro, bāhd-o žī šī-n-ē
 they-AG3p embrace the boy then also they go
 ‘they embrace the boy, **then** they go’;

ez do šir-ā, emā šārt-ē-d-ē mi est-o
 I FUT go but condition of me is
 ‘I shall go, **but** I have a condition (*šārt*)’.

In complex sentences, the subject/agent or direct object/patient of the second main clause may generally be unexpressed if it is identical with that of the first main clause:

mi bilifkerd, ū ē xāpey-n-ā-y
 I-AG1s bluff made **and** them-DIR-PAT-3p cheated-3p
 ‘I made a bluff, and cheated them’.

If one of the two main clauses is constructed ergatively, and the other one intransitive, this is true for 3rd person subjects/agents only:

ez xāpey-n-ā-y-ā, ū rem-ā
 I-DIR-PAT seduced-1s and fled-INTR-3s
 ‘(he) seduced me, and fled’;

dēw ām-e-y-o, finā (āw-ā bīr-d-ē dew-i) birn-ā-y-ā
 diw has come-INTR again (water of well of village)-PAT-f interrupted-PRET-f
 ‘the diw has come, and again (*finā*) interrupted (the village well’s water) (*āw*, f)’.

However, 1st or 2nd person subjects must be expressed:

[to] girot-i, ū [ti] ā-geyr-ē
 you-OBL-AG took-f and you-DIR came back-INTR
 ‘you took her, and you came back’.

5.2 Subordinate clauses

A subordinate clause may be subordinated to a main clause, or to another subordinate clause. Most types of subordinate clauses precede the main clause they qualify. Zazaki subordinate clauses may have nominal, attributive, or adverbial function. Even though not all of these clause types of traditional grammar constitute a distinct formal category in Zazaki, they will be so distinguished here.

Typically, most types of subordinate clauses are introduced by a characteristic clause-initial conjunction or conjunctive phrase, or a relative pronoun or the relative particle *ki*. Clause-final conjunctives are the enclitic particle *-se* ‘if’ (from Turkish) in certain conditional clauses, and the adverb *tepiyā* ‘after’ in certain posterior temporal clauses. Some types of subordinate clause have a variant without conjunction. Most frequent are: the conjunction/particle *ki* which has a broad range of uses in various clause types, and occurs as the (optional) second element of some bipartite conjunctions; further, *eger* ‘if’, *wext-o ki* ‘when’, *hetān (ki)* ‘as long as, until’, *madem (ki)* ‘because’, *çimkî* ‘because’, *hendik(i)* ‘as much/far as’.

5.3 Relative clauses

The relative clauses (RC) of Zazaki may be classified into those that are *independent* and function as subject, object, or adverbial complement, and those that are *attributive* and are governed by a head noun. Independent RCs are introduced by an indefinite relative pronoun or adverb, attributive RCs require an *ezāfe* construction with adjectival *ezāfe* particles connecting the head noun with the relative particle *ki*, or pronoun, dir. *ki*, obl. *kē*: independent *kām šî-n-o* ‘(he) who goes’ vs. attributive *merdim-o ki šî-n-o* ‘a/the man who goes’.

An intermediate position between the nominal *ezāfe* and the attributive RC is held by two constructions, RC with gapped copula:

mār-o ki keynek-i pîz-e-di
 ‘the snake **which** in the girl’s belly’,

and EZ-construction with particles such as *zey* ‘like’:

jinî-y-ā zey to
 ‘a woman like you’.

5.3.1 Functions of RC in the main clause

In functions other than subject, the relative pronoun ‘who’ is in the oblique, *kē*, but not the head noun:

<i>kē tepš-ē</i> ‘ whom they (may) catch’;	<i>kē-di nāmūs est-o</i> ‘at whom is honor’ (‘who(ever) has’);	<i>jewāb-ē kē dihā weš-d-ē mi šîr-o</i> ‘the answer of whom I like more’ (‘whose answer’);
--	---	---

but

girwey-o ki mā nē-zā-n-ē
‘a work **that** we don’t know’.

5.3.2 Resumptive pronouns

Clause-internally, the co-referent to the head noun is resumed by the general 3rd person pronoun *ji* if it functions other than as a subject or object:

qulā ki bin-ē ji nī-y-ās-en-o
‘a hole **that** bottom **of it** is not visible’ (‘whose ground’).

In the main clause, a RC is resumed by a pronoun that disambiguates the syntactic function. This is obligatorily the case of indirect object, and postpositional and *ezāfe* complements:

zūr-ā kē werdī b-o-se, o do virāz-o
‘**whose** lie is smaller, **he** FUT prepare it’ (‘the lie of whom’)
(*b-o* = 3sm subj. of ‘be’ + conditional *-se*);

kām pē-rā bol hes ke-n-o, inān-rē vā-n-ē
‘**who** loves each other (*pē-rā*) very much (*bol*), **to them** they say’.

As part of the main clause, the RC may take the regular position of the constituent as which it functions:

ši-n-o o ode-y-o ki āltūn-ī miyān-dī bī
‘he goes to the room **that** gold in (it) was’ (‘in which was’).

5.3.3 Topicalized relative clauses

A RC may be topicalized and take the sentence initial position. In that case, its regular position remains either vacant:

girwe-y-o ki mā nē-zā-n-ē, mā nē-ke-n-ē
work **that** we not know we not do
‘we don’t do a work that we don’t know’;

or it is resumed by a pronoun that makes his syntactic function explicit:

mā kē-rā qāhr-ēn-ē, mā ey ā-n-ē
we **whom** we are angry we **him** bring
‘about whom we are angry, him we bring’.

5.3.4 Head incorporation

The incorporation of the head noun into the relative clause is a noteworthy option (also known in Persian and other Iranian languages):

por-ē mi kew-n-o kē dest, wā bīr-o
hair of me falls **whose hand** may he come
‘in whose hand my hair (*por*) falls, (he) shall come’.

5.4 Nominal clauses

Subject and object clauses paraphrase subjects and objects but do not occur in their regular position, instead they follow the predicate, often introduced by *ki* ‘that’. Indirect questions also, and sometimes conditional or comparative clauses, may be nominal clauses.

5.4.1 Subject clauses

Subject clauses follow impersonal expressions such as ‘it is necessary, clear, a pity’, etc., whose subject they replace:

b-en-o ki ez bi-mir-ā
‘maybe that I (shall) die’.

They can also resume a preceding demonstrative pronoun that takes the subject position:

sebeb-ē ji no yo ki kes-ī čī-y-ē-d-o winā qebūl nē-kerd-o
reason of it this is that anybody thing such accepted has not made
‘the reason for it is that nobody has (ever) accepted such a thing’.

5.4.2 Object clauses

Object clauses follow verbs of perception and communication and paraphrase the direct object or complement that is usually governed by these verbs:

ez zā-n-ā ki o yo ters-en-o
‘I know that he is fearing’ (‘frightened’).

5.4.3 Quotations

The two verbs of communication *vāž-* ‘say’ and *pers ker-* ‘ask’ (‘to make question’), however, usually govern direct speech instead of an object clause:

pēyember vā-n-o: nē-b-en-o *ji-rā pers ke-n-o: wā-y-ē, ti kām ā?*
‘the prophet says: “it doesn’t work” ’; ‘he asks her: “Sister, who are you?” ’

5.5 Adverbial clauses

5.5.1 Temporal and local clauses

Zazaki temporal clauses always precede the main clause. Formally, they cannot always be clearly distinguished from other subordinate clauses, e.g. temporal clauses introduced by *wext-o ki* ‘when’ (lit. ‘(at) the time that’) are similar to relative clauses introduced by, e.g. *rož-ā ki* ‘(on) the day that’. Here, temporal clauses will be distinguished, according to their temporal relation to the main clause, in anterior, simultaneous, and posterior clauses.

5.5.1.1 Anterior temporal clauses

These are formed with the adverb *tepeyā* ‘after’ following the verb of the temporal clause. This verb is usually in the preterite, but may be in the present tense if the verb of the main clause is also in the present:

Γezāli ši-n-ā tep-ey-ā, Āḥmed weyn dā-n-o činā dā xo-rā tep-e-yā,
 Γ. goes-PRES after A., call hits-PRES clothes put on own **after**
 ‘after Γezāli goes Ahmad calls’; ‘he goes after having put on (his)
ši-n-o
 he goes-PRES
 clothes’.

5.5.1.2 Simultaneous temporal clauses

These are introduced by *wext-o ki* ‘when’, rarely also *ki*, following the subject of the main clause. The verb of the temporal clause is usually in the same tense as that of the main clause:

wext-o ki keko merd, mi zā-n-ā
 ‘when the older brother died-PRET, I-AG knew-PRET . . .’;
keynek-i ki b-en-ā teyšān, ye-n-ā
 girl when she becomes-PRES thirsty she comes-PRES
 ‘when the girl gets thirsty, she comes’.

If the temporal sentence is in the preterite and the main clause in a present tense, *wext-o ki* may correspond to English ‘as soon as’; more often, however, *senī (ki)* occurs in this sense, then causing the same tense in the temporal and main clause:

wext-o ki K. āme, wā Ā. ji ber-o heš-i senī ez dī-y-ā,
 when K. came-PRET, A him bring-SUBJ; bear as soon as I-PAT saw-PRET-1S
rem-ē
 he fled-PRET
 ‘as soon as K. comes, A. shall bring him’; ‘as soon as the bear saw me, he fled’.

5.5.1.3 Posterior temporal clauses

These are introduced by *hetā(n) (ki)* ‘until; as long as’. When a past action is reported, the verb of the temporal clause is in the present indicative or in a past tense, and *hetā(n) (ki)* corresponds to English ‘until’:

hetān xo res-n-en-ē-PRES wede-dē pādīšāy, xeylē māḥkūm-ī mir-en-ē-PRES
 ‘until they proceed to the king’s room, many criminals die’;
hetān amē-PRET res-ē-pret to het, bī newāy ū new
 ‘until it came (and) reached to you, it was-past 99’.

When the reported action is not yet past, the verb is negated and in the present subjunctive or preterite, and *hetā(n) (ki)* corresponds to the English ‘as long as’:

hetān ez ey nē-kiš-ā, mā-rē reḥāt-ey čī-nī-y-ā
 as long as I him not kill-subj to us peace is not
 ‘as long as I don’t kill him, we have no peace’;

hetān ē nī-y-āme-y-ē, mā bi-rem-ē
 as long as they have not come-pret we let us flee
 'let us flee, as long as they haven't come'.

5.5.1.4 Local clauses

Local clauses are introduced by *kotī* '(to) where', *kotī-di* 'where', or *kotī-rā* 'from where' and may be resumed through the local adverb *wižā* 'there' (or *wižā-di*, *wižā-rā* respectively):

kotī-rā ye-n-ē, wā b-īr-ē
 'from where they come, may they come';

kotī-di šān b-en-o, ez wižā-di rā-kew-n-ā
 where dark it becomes I to sleep fall to
 'where(ever) it gets dark, there I go to sleep'.

5.5.2 Conditional clauses

Zazaki conditional clauses are mostly introduced by the conjunctions *eger* or (less often) *ki*, the enclitic particle *-se* following the verb, or by a combination of both (*eger . . . -se*), all meaning 'if'. They always precede the main clause ('apodosis'). The conditional clauses may be distinguished in those with possible and those with impossible condition.

5.5.2.1 Possible conditions

5.5.2.1a Present and subjunctive

With possible conditions, the verb of both the protasis and apodosis often has a present tense form. The mode of the verb of the apodosis is, in principle, independent of that of the protasis.

(1) Present

eger ti wāz-en-ā-PRES, mā piyā bi-zewē-iy-ē-SUBJ
 'if you want, let us marry';

ki ti dā-n-ē-PRES, ti finā ded-ē mā yē-PRES
 'if you give (it), you are again (*finā*) our uncle'.

(2) Subjunctive

eger bi-zān-o, ez do e-y šēx-o rāstā-y ker-ā
 if he knows-SUBJ I FUT him true sheikh I make-SUBJ
 'if he knows (it), I will make him a true sheikh';

bi-kew-o-se, dāḥā weš nē-b-en-o
 he falls-SUBJ if, again well he won't become-PRES
 'if he falls, he won't get well again'.

5.5.2.1b Preterite

If the action of the apodosis must, of necessity, follow that of the protasis in time, the past tense may appear in the protasis:

eger wihēr-ē ji ām-e, mā do heq-ē ji bi-d-ē
 if owner of it came-PRET we FUT fee of him give-SUBJ
 ‘if his owner comes, we will (certainly) pay its fee’;

ti merd-ī-se, mā žī to-y-ā mir-en-ē
 you died-PRET if we also with you die-PRES
 ‘if you die, we also will die with you’.

5.5.2.1c Perfect

Conditions that have already been fulfilled or not, but about whose fulfillment the speaker has no knowledge, have a verb in the perfect tense:

eger čī-y-ē bi Γezāli kerd-o, mā gān-ē xo nān-ē ser
 if anything to Γ. done-PERF-3sm we life of own shall put-SUBJ up
 ‘if they have done anything to Γ., we shall give our lives’;

čīnā-y-ē mā berd-o-se, mā do šir-ē key-d-ē milā-y
 clothes of us taken-PERF3sm if we FUT go-SUBJ house of Molla
 ‘if he has taken away our clothes, we will go to the Molla’s house’.

5.5.2.2 Impossible conditions, -ē

With *impossible past conditions* in the protasis the past subjunctive or past optative in -ē is used, and in the apodosis the conditional:

eger ē nī-y-āme-y-ē, mi do kotī-rā bi-zānā-y-ē
 if they had not come-PASTOPT I-AG FUT wherefrom have known-COND
 ‘if they hadn’t come, from where should I have known’.

Most types of conditional clause may also occur *without conjunction*, and then be recognized as such by the context or a subjunctive verb form only:

wā-y-ē, ti čī-y-ē-rā fāhun ke-n-ā, mā b-īr-ē to hetī
 sister you anything understand-PRES we let come-SUBJ you to
 ‘sister, if you understand anything, let us come to you’;

mi poste-y-ē to nē-veš-n-ā-y-ē, diḥā ti do bī-y-ā-y-ē
 I-AG skin of you had not burnt-PAST-SUBJ further you FUT had eaten-COND
 ‘if I hadn’t burnt your skin, you would have furthermore been . . .’

5.5.3 Concessive clauses

A concessive relation is often expressed through two main clauses that are connected through the coordinating conjunction *finā žī* ‘nevertheless, yet’:

ḥewtāy dew-ān werd-i, finā žī nē-šā bi-qed-ēn-o
 70 villages-AG have eaten again also could not be finished-SUBJ
 ‘70 villages have eaten (of it), still it could not be finished up’.

A conditional clause with *žī* ‘also’ may correspond to an English concessive clause:

mā bi-zān-ē ki to kiš-en-o žī, mā mejbūr ē to bi-riš-ē
 we know-SUBJ that you he is killing **though** we **forced are** you we send-SUBJ
 ‘**even if** we know he is going to kill you, we must send you (there)’.

Sometimes, an optative clause introduced by *wā*, following a main clause, may express a concessive relation:

mā žī šop-tā īn-ān-rā šī-n-ē, wā mā žī bi-kiš-ē
 we also track of them-DIR-OBJ go **may be that** us also they kill-SUBJ
 ‘and we follow their track, even if (‘may that’) they kill us’.

5.5.4 Causal clauses

Causal clauses are introduced by *čimkī* or *mādem (ki)* ‘because, since’ and stand with *čimkī* before or after, with *mādem (ki)* always before the main clause:

čimkī Ābdirreḥmān Beg merd, ē rey-ā-y
 ‘**as** A. Beg died, they got free’;

dew-iž-ī qānd-ē ji šīn ke-n-ē, čimkī her-kes ji-rā ḥes k-en-o
 the farmers sake of him mourning make **because** everyone him liking makes
 ‘the farmers mourn for him, because everybody loves him’;

mādem ti šī, merāq me-k-i
since you went-PRET concern don’t make
 ‘since you have gone, don’t be surprised!’

More often, however, a causal connection is expressed by two main clauses and the coordinating conjunction *qānd-ē joy* ‘therefore’:

čī-y-ē ām-e ey sere-di, qānd-ē joy o nī-y-ām-e
 ‘something has happened (*āme* ‘came’) to him, **therefore** he hasn’t come’.

5.5.5 Final clauses

Final clauses are introduced by *ki*, more seldom by *qānd-ē ki*, both ‘(in order) to’, and require, irrespective of the tense of the main clause verb, a verb in the present subjunctive. With *ki* they follow, with *qānd-ē ki* they precede the main clause:

lāžek šī-y-o ki poste-y-ē ā-y bi-vīn-o
 boy has gone **that** skin of her find-SUBJ
 ‘the boy has gone to find her skin’;

qānd-ē ki ti nē-mir-ē, ti do šop-d-ā ā-y-rā šir-ē
in order that you not die-SUBJ you FUT track of her go-SUBJ
 ‘so as not to die, you will follow her track’.

In addition, a final clause may often also follow the main clause asyndetically, or an optative (main) clause with *wā* may correspond to an English final clause:

ez āme-y-ā gān-ē to bi-gīr-ā
 I have come life of you I take-SUBJ
 ‘I have come to take your life’;

‘ējele ne-y ber-ē, wā čim-ē mi ne-y nē-vīn-ē
 quickly him take **so that** eyes of me him not see-SUBJ
 ‘take him away quickly, so that my eyes may not see him (any longer)’.

5.5.6 Consecutive clauses

Consecutive clauses are built with one of the conjunctions *hendi/winī* . . . *ki* ‘so . . . that’. The consecutive clause, introduced by *ki*, follows the main clause:

hendi rind b-en-ā, ki kes nē-š-en-o wesf-ān-ē jī bi-d-o
 so pretty she is being **that** anybody cannot description of her make-SUBJ
 ‘she is so pretty that nobody can describe her’;

ez do winī rezīl-ey-ā gān-ē to bi-gīr-ān, ki . . .
 I FUT **such** meanness life of you I take-SUBJ **that**
 ‘I shall take your life with such a meanness that . . .’

Certain clauses that follow a double *winī*, or a question clause asyndetically (ex.2), may also correspond to an English consecutive clause:

winī jī-rā hes ke-n-ē, winī hes ke-n-ē, jī-rē gān-ē xo dā-n-ē
 so it liking make **so** liking make for her life of self they give
 ‘they love it so much, so much, (that) they lay down for it their life’;

to čī čāre dī, ti remā-y?
 you-AG **what** remedy saw you have fled
 ‘what remedy have you found, (that) you have fled?’

5.5.7 Comparative clauses

Comparative clauses of *equality* are introduced by *zeki*, *senī* or (rarely) *hendi ki*, and followed by a main clause that is often introduced by *winī*. The verb of the comparative clause is often indicative, but may also (after *zeki*) be subjunctive:

zeki qertel-ī-leš-ī sero ārē-b-en-ē, winī xidir āyā-y sero ārē-b-en-ē
 like vulture corpse upon gather **so** X.Agha upon they gather
 ‘like vultures gather on a carcass, so they gather on X. Agha’;

zeki jīp-i bi-rām-o, dīreksīyon-ī winī ke-n-o xo dest
 like a jeep he may drive-SUBJ wheel **so** he makes his hand
 ‘as if he would drive a jeep, so he takes the wheel in his hand’;

esto-ri senī rem-en-ā, wā bi-rem-o
 ‘like the horse flees, **so** it shall flee’;

hendi ki mi dest-rā b-īr-o, ez do bi-girwe-y-ā
 as much as my hand-from may come-SUBJ I FUT shall work-SUBJ
 ‘as much as I can, I will work’.

6 LEXIS AND REGISTERS

6.1 Word formation

In the fairy tales and folkloristic texts that have been used for this study and that continue to predominate in Zazaki literary production, the possibilities of using derivation and composition for the formation of complex lexical units seem to be rather limited. This may be due to the fact that Zazaki still lacks most registers of official and “higher” literal usage, for which complex lexical units are usually formed.

6.1.1 Derivation of nouns and adjectives

The most important derivational suffixes for the formation of nominals that continue to be productive in Zazaki are:

- *-iž*, the nisba suffix that forms nouns of origin and certain other nouns: *Mālmīsān-iž* ‘(a person) from (the mountain of) Malmisan’, *dew-iž* ‘farmer’ (< *dew* ‘village’);
- *-ey*, attached to adjectives or agent nouns, to form abstract nouns (that are feminine): *rind-ey* ‘goodness’, *dizd-ey* ‘theft’ (< *rind* ‘good’, *dizd* ‘thief’) (in certain dialects, *-ey* is replaced by *-in* or *-en*);
- *-ēr*, attached to names of fruits, to form tree names (that are feminine): *sā-y-ēr* ‘apple tree’ (< *sā* ‘apple’);
- *-en* (or in certain dialects, *-in*), attached to nouns or adverbs, to form adjectives: *jēr-en* ‘(being) below’, *reng-en* ‘coloured’, *gošt-en* ‘fleshy, consisting of meat’ (< *jēr* ‘below’, *reng* ‘colour’, *gošt* ‘meat’).

6.1.2 Compounding

The most important compounding strategies in Zazaki are (Paul 2002):

- (1) right-headed determinative compounds occur seldom, and mostly as ‘frozen’ historical forms: *kē-ber* ‘house-door’ (the regular syntactic *Ezāfe* construction seems to be much more productive to express concepts that correspond to determinative compounds in other languages);
- (2) more frequent are verbal compounds that employ a verbal stem (usually the subjunctive one) as a second element: *merdim-wer* ‘cannibal(istic)’ (lit. ‘man-eating’);
- (3) the most productive compound type that seems to occur in the texts that have been used here are left-headed possessive compounds: *gān-rihāt* ‘comfortable’, lit. ‘(having a) relaxed soul’ < *gān* ‘soul’, *rihāt* ‘comfortable’.

6.2 Sociolinguistic features

6.2.1 Standard and colloquial registers

Since Zazaki still lacks most registers of official and “higher” literal usage and the majority of publications still belong to folk language and everyday speech, it is of no use to talk about the “social” strata of Zazaki. As for the question of a possible standard dialect, there have been various discussions since the 1990s (mostly in the exile press) about which dialect, or combination of dialects, should be preferred to create such a standard dialect,

but (at least until 2005) these discussions seem to have been purely academic, and every Zazaki writer still seems to use his own (village or town) dialect when he (or she) writes.

6.2.2 Kinship terminology

Some of the most important items of kinship terminology that have been met in the texts include the following: *mā(y)* ‘mother’, *pī* ‘father’, *wā(y)* ‘sister’, *birā* ‘brother’, *xāl* ‘(maternal) uncle’, *vistirī* ‘mother-in-law’. Interestingly, some human nouns like ‘boy, girl’, etc., are derived by the suffix *-ek* from their corresponding kinship term, e.g.:

lāž ‘son’ — *lāž-ek* ‘boy’ *keyn(ā)* daughter’ — *keyn-ek* ‘girl’
jinī ‘wife’ — *jinēk* ‘woman’ *mērd* ‘husband’ — *mērd-ek* ‘man’

6.2.3 Terms of address

Many Zazaki terms of address correspond to (Kurmanci) Kurdish equivalents and thus attest to the close ethnic and historical bond that connects the Zaza to the Kurds:

address: *çim-ān ser* ‘goodbye’ (lit. ‘on the eyes’) ~ Kd. *ser çāv-ān* ‘id.’;
 response: *ser-ān ser* ‘goodbye’ (lit. ‘on the heads’) ~ Kd. *ser ser-ān*.

Other Zazaki terms of address include:

address: *ti xeyr āmeye!* ‘welcome!’ (lit. ‘you have come well’);
 response: *xeyr mi yān-di b-o!* (lit. ‘may welfare be in-between’);

address: *xātir bi to!* ‘goodbye’ (lit. ‘goodbye (be) to you’);
 response: *oyur b-o!* (lit. ‘may good luck be!’);
Ellāh rāzī b-o! ‘thank you!’ (lit. ‘may god be content’).

7 SAMPLE TEXT

‘*Ārwēš-ī*’ (‘The Hares’), in Koyo Berz 1988.

ĵjin ū ĵināvīr-ī miyān-di, en tersinok-ē ĵi ārwēš-ī b-en-ē.
 Animals-OBL among-in, most fearful-EZ them-OBL hare-up are
 ‘Among the animals, the most fearful of them are the hares.’

mār xuš-en-o, ārwēš-ī rem-en-ē; her zīř-en-o, ārwēš-ī rem-en-ē,
 snake hisses hare-p flee-3p donkey bray-sm hares-p flee-3p
 ‘The snake hisses, the hares flee; the donkey brays, the hares flee’;

gā qoř-en-o, ārwēš-ī rem-en-ē; kutik lāw-en-o, ārwēš-ī rem-en-ē . . .
 ox bellow-sm hare-p flee-3p dog bark-sm hare-p flee-3p
 ‘the ox bellows, the hares flee; the dog barks, the hares flee . . .’

çičī xo lūn-en-o, ārwēš-ī rem-en-ē.
 what self move-sm hare-p flee-3p
 ‘what(ever) jumps about, the hares flee’.

řož-ē ārweš-ī pēro ārē-b-en-ē pē ser, vān-ē:
 day-a hare-p all gather-3p each-other on, say-3p
 ‘One day, all hares gather around themselves and say.’

no dinyā-di mā-rā tersinok kes čī-nī-y-o,
 this world-in us-from fearful someone not-exists
 ‘In this world, there is no-one more fearful than we are.’

en tersinok-ē nē dinyā-y mā yē.
 most fearful-EZ-p this-OBL world-OBL we are
 ‘we are the most fearful (ones) of this world.’

mā hendā-y-ē tersinok-ī yē,
 we so-much fearful-p are-1p
 ‘We are so fearful.’

ḥādirē! mā pēro šir-ē, xo če-ker-ē dol-ī miyān,
 ahead! we all go-SBJ-1p self inside-do-SBJ-1p lake-OBL into
 ‘(let’s go) ahead, (and) all go (and) throw ourselves into the lake.’

mā pēro bi-fetis-īy-ē ū nē ters-ī-rā bire-yē!
 we all drown-SBJ-1p and this-OBL fear-OBL-from escape-SBJ-1p
 ‘let’s all be drowned, and escape from this fright!’

nē dā-n-ē piro šī-n-ē ki
 these set-3p about go-3p that
 ‘They set about (and) go to’

xo če-ker-ē dol-ē-d-ē līmin-ī miyān.
 self inside-do-SUBJ-3p lake-one-EZ-OBL dirty-OBL into’.
 ‘throw themselves into a dirty lake.’

ye-n-ē nezdī-d-ē dol-ī ki
 come-3p near-EZ-OBL lake-OBL that
 ‘They come close to the lake so that’

xo če-ker-ē dol-ī miyān,
 self inside-do-SBJ-3p lake-OBL into
 ‘throw themselves inside the lake.’

win-ēn-ē kese-y-ē ki
 notice-3p turtles-EZ-p that’
 ‘they notice turtles that’

teber-d-ē dol-ī-di xo dā-y-o ver-ā-rož.
 outside-EZ-OBL lake-OBL-in self give-PERF-3sm sun-place
 ‘have exposed themselves to the sun outside the lake.’

ē kese-y pēro finē-rā
 those turtles-p all suddenly
 ‘All those turtles all of a sudden.’

ters-ān-d-ē ārweš-ān-rā rāp rāp rāp
 fear-OBL-p-EZ-OBL hare-OBL-p-from rap rap rap
 ‘out of fear of the hares, rap rap rap’

xo erz-en-ē dol-ī miyān
 self throw-3p lake-OBL into'
 'throw themselves into the lake'

ū ters-ān-rā xo dol-ī miyān-di nimn-en-ē.
 and fear-OBL-p-from self lake-OBL midst-in hide-3p
 'out of fear, and hide themselves, in the lake.'

ārweš-ī vinder-en-ē, vān-ē:
 hare-p stop-3p say-3p
 'The hares stop, (and) say.'

nē kese-y mā-rā tersā-y
 these turtle-p us-from fear-ed-3p
 'These turtles were afraid of us.'

xo ešt dol-ī miyān ū xo dol-di nimit.
 self threw-3sm lake-OBL inside and self lake-in hid-3sm
 'they threw themselves inside the lake and hid themselves in the lake.'

zey mā tersinok-ī dinyā-di bol-ī yē,
 like us fearful-p world-in many-p are
 'In the world, there are many (who are) fearful like us.'

mā či-rē xo çe-ker-ē dol-ī miyān
 we why self inside-do-SBJ-1p lake inside
 'why should we throw ourselves into the lake'

ū xo bi-fetis-n-ē?
 and self drown-SBJ-1p
 'and drown ourselves?'

ḥādir-ē! mā pey-di šir-ē,
 ahead we back go-SBJ-1p
 '(Let's go) ahead, let's go back(wards).'

mā xo nē-y-erz-en-ē dol-ī miyān'.
 we self NEG-throw-1p lake inside'.
 'we don't throw ourselves into the lake'.

pēro dān-ē piro finā pey-di ā-geyr-en-ē
 all set-1p about again back return-1p
 'All set about (and) return back again.'

ū ši-n-ē, 'erd-ān miyān-rā b-en-ē vilā-y.
 and go-1p soil-OBL-p among-from become-1p scattered-p
 'and go (and) get scattered among the earths.'

xeyr-d-ē kes-ān,
 benefit-EZ-OBL turtles-OBL-p
 'Thanks to the turtles.'

ārweš-ī kišt-en-d-ā xo-rā rey-ēn-ē.
 hares-p killing-INF-EZ-OBL self-from escape-1p
 'the hares escape from their suicide.'

BIBLIOGRAPHY AND REFERENCES

- Koyo Berz (1988) *Na xumxum a . . .* (Collection of stories and tales) (Weşanên Jina Nû 13), Uppsala: Jina Nû Yayınları.
- Lerch, Peter I. A. (1857–58) *Forschungen über die Kurden und die iranischen Nordchaldäer*, 2 vols., St. Petersburg: Eggers.
- Malmisanij (1987) *Zazaca-Türkçe Sözlük/Ferhengê Dimilki-Tirki*, Uppsala: Jina Nû Yayınları.
- Mann, Oskar (1932) *Mundarten der Zāzā, hauptsächlich aus Siwerek und Kor* (= Kurdisch-Persische Forschungen, Abt. III, Bd. IV), Hadank, K. (ed.), Berlin: Verlag der Preußischen Akademie der Wissenschaften, in Kommission bei W. de Gruyter.
- Paul, Ludwig (1998a) 'The position of Zazaki among West Iranian Languages', in Sims-Williams, N. (ed.) *Proceedings of the Third European Conference of Iranian Studies held in Cambridge, 11th to 15th September 1995, Part I, Old and Middle Iranian Studies*, Wiesbaden: L. Reichert, pp. 163–176.
- (1998b) 'Zazaki – Dialekt, Sprache, Nation?', in Köhler, B. (ed.) *Religion und Wahrheit. Religionsgeschichtliche Studien. Festschrift für Gernot Wießner zum 65. Geburtstag*, Wiesbaden: O. Harrassowitz, pp. 385–399.
- (1998c) *Zazaki. Grammatik und Versuch einer Dialektologie*, Wiesbaden: L. Reichert.
- (2002) 'The constituent order of nominal compounds in Zazaki and other Iranian languages', *STUF* 55.1: 64–79.
- Selcan, Zülfü (1998) *Grammatik der Zaza-Sprache. Nord-Dialekt*, Berlin: Wissenschaft & Technik.
- Todd, Terry L. (1985) *A Grammar of Dimilî (also known as Zaza)*, Ph.D. dissertation, University of Michigan, Ann Arbor.

CHAPTER TEN

KURDISH

Ernest N. McCarus

1 INTRODUCTION

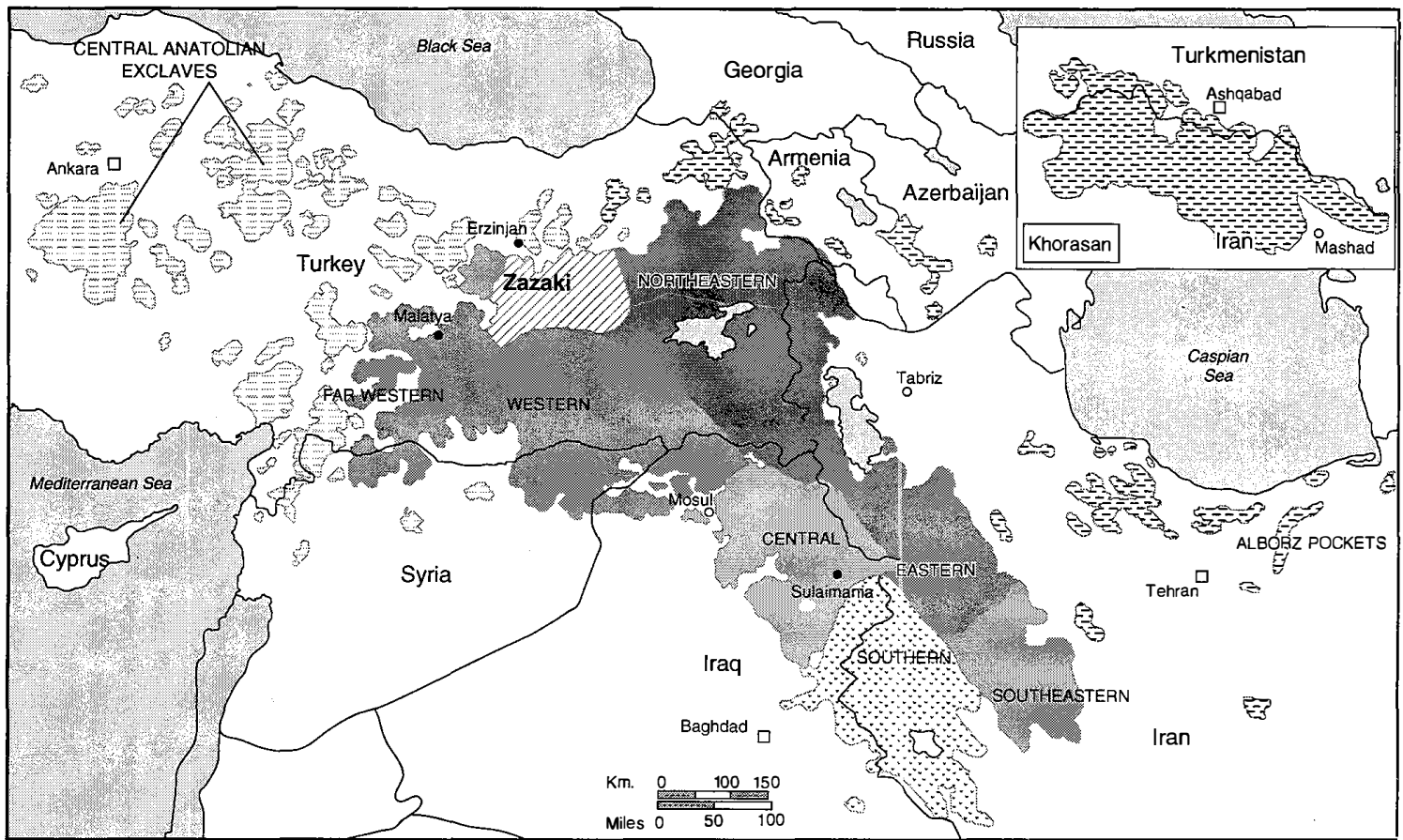
1.1 Overview

Kurdish is a cover term for the largest group of closely-related Western Iranian dialects, spoken in a large contiguous area that extends from Turkey into Iraq, Iran, and Armenia, with smaller groups in other countries. This cover term also reflects the linguistic and ethnic unity perceived by the speakers in spite of considerable linguistic differences. There are three main sub-groups: (1) Northern Kurdish is the most widely spoken variety of Kurdish (ca. 20 million), also known as *Kurmānjī*, which is the autonym used by most speakers, besides *Bahdīnān* in north Iraq. This variety is roughly delimited in the west by the Turkish town of Malatya, in the south by the north-eastern tip of Syria and by Mosul in northern Iraq, and it reaches up to the Armenian capital of Yerevan in the north-east. Relatively large groups are also found in Syria and Lebanon, as well as in Azerbaijan and Georgia in the Caucasus (ca. 500,000 in the former Soviet republics). There are also considerable emigrant groups in urban centers in the Middle East (e.g. ca. 2 million or more in Istanbul) and in Western Europe (ca. 700,000, mostly Germany). (2) Central Kurdish (total ca. 5 million) has two main subgroups, *Sōrānī* in northern Iraq up to the Little Zab river, and *Mukrī* in adjacent Iranian province of Kordestan, with pockets (by earlier deportations) elsewhere in Iran, mostly the Caspian provinces and the north-eastern province of Khorasan. (3) Southern Kurdish (ca. 3 million) is found in the abutting areas of Iraq and Iran, from *Khaneqīn* in Iraq over to *Kermanshah* in Iran and down to north of *Al-Amara*, Iraq, as well as in the *Bijar* region of Iran. Best known are the *Laki* tribes in the *Pish-e Kuh* region of the Zagros mountains between *Kermanshah*, *Khorramabad*, and *Kangavar* (cf. the monumental study by *Fattah*, 2000).

Linguistically, Kurdish as a whole occupies an intermediate position between North-Western and South-Western Iranian dialects. Internally, the three main groups of Kurdish are quite distinct from each other, and Northern Kurdish in particular is not mutually intelligible with the other groups (see section 7 *Dialectology*).

Earliest written documents date from the sixteenth and seventeenth century, prominently the epic poem *Mem û Zîn* by *Ahmadī Khānī* (1650–1707). It is only in the early twentieth century that both Northern Kurdish and *Sōrānī* have developed written standards and alphabets, beginning with the Kurdish alphabet for *Kurmānjī* by *Emir Djeladet Bedir Khan* in the 1930s and 1940s and the newspaper *Zhîn* for *Sōrānī* from the 1920s in *Sulaimania*, Iraq.

In Iraq after the Second World War Kurdish scholars initiated a movement to establish a Modern Standard Kurdish (MSK; see *Ahmad* 1986). It is mainly based on the dialect



MAP 10.1 KURDISH (note that hatched areas = scattered settlements)

of Sulaimania which was considered by most Kurds to be the best literary model. MSK is now used in most Kurdish publications in Iraq and Iran. Accordingly, Sulaimani Kurdish is the focus of the following discussion.

Linguistic scholarship on Kurdish has a long tradition. More recently, in consequence of the massive emigration, paralleling the literary, journalistic, and other cultural activities by Kurds, scholarship has been flourishing in Europe, mainly on the Kurmānjī of Turkey, and prominently from the typological point of view, as have literary studies.

It should be noted that three other Iranian dialect groups which happen to be symbiotic or adjacent to Kurdish-speaking areas are sometimes considered varieties of Kurdish: Zāzākī/Dimilī found in central-eastern Turkey north-west of the present Kurdish area, Gōrānī found in the Awramān region of the middle Zagros in Iran (Hawramānī) and near Mosul in Iraq (mainly Bājilānī), and Lorī-Bakhtiārī in the southern Zagros. However, while these share, to various degrees, a common history and culture with the Kurds, and while there has been considerable convergence on all linguistic levels, in terms of Iranian historical dialectology, they are West Iranian language groups distinct from Kurdish.

1.2 Writing Systems

Kurmānjī Kurdish has been written in a variety of alphabets, from Armenian to Cyrillic to Latin. Today the Kurds of Turkey and Syria use a modified Turkish script, developed originally by Bedir Khan in the 1930s and 1940s, while those of the former Soviet Union use a modified Cyrillic.

Sōrānī Kurdish has always been written, like Persian, in a modified Arabic alphabet. The present-day Kurdish alphabet used in Iraq and Iran was developed mainly after the Second World War, with diacritic marks for Kurdish phonemes not found in either of the other two. It has no symbol for the vowel *i*, but is otherwise a phonemic script, achieved through the use of diacritic marks for non-Arabic consonants and vowels and by reassigning some letter values. The representation of Kurdish phonemes not represented by the Perso-Arabic alphabet, *ŋ* *v* *l* *ř* and *ē* *ō*, is as follows:

ŋ is written by the digraph <ng>; *v* is represented by three dots, instead of one, over <f>; *l* and *ř* are marked by a hachek above <l> and <r>. In addition, the glottal stop is written over the hook of <y> as a seat, except word-finally where it is written aloof, while gemination is represented by doubling the consonant letters.

The long mid-vowels *ō* and *ē* are both indicated by a hachek over <w> and <y>, respectively. *ū* is represented by double <ww>, and *u* by a single <w>; *ī* and *ı* are likewise represented by double <yy>, unless *ı* occurs in a word that retains its Arabic spelling. Long *ā* is represented by the letter *alif*, and short *a* by <h> in postconsonantal position.

TABLE 10.1: KURDISH ALPHABET

<i>Position in Letter Group</i>				Transcription	
last	mid	first	alone		
ـ	ـ	ـ	ـ	ʾ	
ـ	ـ	ـ	ـ	ā	
ب	ب	ب	ب	b	
پ	پ	پ	پ	p	Persian
ت	ت	ت	ت	t	
ث	ث	ث	ث	θ	Arabic
ج	ج	ج	ج	j	
چ	چ	چ	چ	č	Persian
ح	ح	ح	ح	h	
خ	خ	خ	خ	x	
د	ـ	ـ	ـ	d	
ذ	ـ	ـ	ـ	ð	Arabic
ر	ـ	ـ	ـ	r	
ڕ	ـ	ـ	ـ	ř	Kurdish
ز	ـ	ـ	ـ	z	
ژ	ـ	ـ	ـ	ž	Persian
س	س	س	س	s	
ش	ش	ش	ش	š	
ص	ص	ص	ص	ṣ	Arabic
ض	ض	ض	ض	ḍ	Arabic
ط	ط	ط	ط	t̤	Arabic
ظ	ظ	ظ	ظ	ḏ	Arabic
ع	ع	ع	ع	ʿ	
غ	غ	غ	غ	ɣ	
ف	ف	ف	ف	f	
ڤ	ڤ	ڤ	ڤ	v	Kurdish
ق	ق	ق	ق	q	
ک	ک	ک	ک	k	
گ	گ	گ	گ	g	Persian
ل	ل	ل	ل	l	
ڵ	ڵ	ڵ	ڵ	l̥	Kurdish
م	م	م	م	m	
ن	ن	ن	ن	n	
ڤڤ	ڤڤ	ـ	ـ	ŋ	
ه ¹	ه	ه	ه	h	
ه ²	ـ	ـ	ـ	a	Kurdish
و ¹	ـ	ـ	و	w	
و ²	ـ	ـ	و	ō	Kurdish
و ²	ـ	ـ	و	u	Kurdish
وو ²	ـ	ـ	وو	ū	Kurdish
ی ¹	ی	ی	ی	y	
ی ²	ی	ـ	ی	ī	Kurdish
ی ²	ی	ـ	ی	ē	Kurdish

1 = after vowel; 2 = after consonant.

Notes: There is no letter to represent phonemic short *i*.

'Persian' and 'Kurdish' refer to innovating language sound, 'Arabic' to letters used in Arabic names.

2 PHONOLOGY

A characteristic feature of Kurdish in comparison with other Iranian languages is the high degree of conditioned morphophonemic alternation that affects vowels and consonants alike, in particular the intervocalic lenition of *-d-* (the “Zagros-*d*”).

2.1 Inventory

2.1.1 Vowels

Sulaimania Kurdish has nine vowels, five long, \bar{i} \bar{e} \bar{a} \bar{o} \bar{u} , and four short, contrasting in length, height and tenseness, ι i u a .

TABLE 10.2: VOWELS

	Front	Central	Back
High close	\bar{i}		\bar{u}
open	ι	i	u
Mid close	\bar{e}		\bar{o}
open		a	
Low		\bar{a}	

There are two central vowels, both short, a and i . The open mid front vowel a has schwa [ə] as its principal allophone. It fluctuates over the area delineated by [ə], low front [æ] and mid front [ɛ]. The allophone [ə] is obligatory (1) before w , as in aw [əw] ‘he’, $aw\bar{a}n$ [əwɑ:n] ‘they’; (2) before y in the same syllable, as in $ayb\bar{i}n\bar{i}m$ [əy.bi.nim] ‘I see him’. The allophone [ɛ] occurs before y in the following syllable, as in $haya$ [hɛ.yə] ‘there is’. The low central allophone [a] occurs adjacent to pharyngealized \bar{s} as in $\bar{s}a\bar{s}t$ [šaʃt] ‘sixty’.

As shown in Table 10.2, there is a three-way contrast among high unrounded vowels: long front close \bar{i} , short front open ι , and central unrounded open i . ι occurs mostly in Arabic loans, as in $umt\bar{i}h\bar{a}n$ ‘examination’, $mumk\bar{i}n$ ‘possible’, or as an alternant of high front \bar{i} before consonant clusters (see section 2.3.1.5). It contrasts with \bar{i} and i as illustrated by $mumk\bar{i}n$ ‘possible’ versus $a\bar{c}\bar{i}n$ ‘we go’ and $a\bar{c}\bar{i}n$ ‘you-2p go’, in all of which the final vowels are stressed.

Natively, the high open short vowels ι and u occur only as reduced allophones of \bar{i} and \bar{u} (see section 2.3.1), but have attained phonemic status by virtue of their occurrence in the many Arabic loans.

2.1.2 Consonants

2.1.2.1 Overview

Sulaimani Kurdish has 31 consonants as shown in the following table:

TABLE 10.3: CONSONANTS

	Labial	Dental	Phar.	Palatal	Velar	Uv.	Phar.	Glottal
Stops/Affricates								
voiceless	<i>p</i>	<i>t</i>	—	<i>č</i>	<i>k</i>	<i>q</i>	—	ʔ
voiced	<i>b</i>	<i>d</i>	—	<i>j</i>	<i>g</i>	—	—	—
Fricatives								
voiceless	<i>f</i>	<i>s</i>	<i>ʃ</i>	<i>š</i>	<i>x</i>	—	<i>ħ</i>	<i>h</i>
voiced	(<i>v</i>)	<i>z</i>	—	<i>ž</i>	<i>ɣ</i>	—	ʕ	—
Glides	<i>w</i>	—	—	<i>y</i>	—	—	—	—
Nasals	<i>m</i>	<i>n</i>	—	—	<i>ŋ</i>	—	—	—
Flap/Lateral								
plain		<i>r l</i>						
trill/velarized		<i>ř l</i>						

Note that the non-pharyngeal glottal stop, particularly in word-initial position, will not usually be marked in this description.

2.1.2.2 Pharyngeals

There are three, all fricatives: (1) the pharyngealized alveolar sibilant *ʃ*; (2) the voiceless faucalized pharyngeal *ħ*; and (3) the voiced faucalized fricative approximant *ʕ*. These three, together with *ɣ*, occur in Arabic loans, but also in a number of words of Kurdish origin: (1) *šāliḥ* ‘Salih’ (masc. name); in high-frequency Kurdish words such as *šag* ‘dog’, *šad* ‘hundred’, *šāl* ‘year’; (2) *ḥāl* ‘condition, state’, *ḥaywān* ‘animal’, *siftaḥ* ‘first sale of the day’; Kurdish *ḥawt* ‘seven’, *aḥēlēnē* ‘it neighs’; (3) *ʔdat* ‘custom, tradition’, *ʔmr* ‘age (years)’, *saʔt* ‘hour; clock’, *maʔnā* ‘meaning’; Kurdish *āsmān* ~ *ʔsmān* ‘sky’; (4) *pōya* ‘bull calf’.

2.1.2.3 Labio-dental *v*

This voiced labio-dental fricative is a marginal phoneme (indicated by the parenthesis in the table). It occurs: (1) in onomatopoeias: *gīvagīv* (sound of bullets); or (2) in loans from other Kurdish dialects, such as *mirōv* ‘man’ and *gōvār* ‘magazine’, and in (3) *ḥavda* ~ *ḥavva* ‘seventeen’.

2.1.2.4 *l* and *l*

These are voiced, and distinguish a plain dental and a velarized alveolar lateral, articulated by the retraction of the tongue body and the tensing of the tongue dorsum. *l* does not occur word-initially: (1) *lāzim* ‘necessary’, *salāmat* ‘safety’, *kalla* ‘skull’, *gul* ‘leper’; (2) *gul* ‘rose’, *barallā* ‘loose’, *māl* ‘house’ (mostly, *l* < **rd*, **rz*, and spontaneous in loans).

2.1.2.5 *r* and *ř*

These distinguish a voiced single alveolar/apico-laminal flap and an alveolar trill. Geminate flaps equal the trill: (1) *birīn* ‘wound’, *anērē* ‘he will send’, *anērrē* [əne:ře] ‘it will be sent’, *kar-a* [kəɾə] ‘it is a donkey’, *har* ‘only’; (2) *kař-a* [kəřə] ‘he is deaf’, *nař-a-nař* ‘bawling’, *a-bř-ē* [əbře] ‘he cuts’, *abiř-rē* [əbiře:] ‘it will be cut’, *tař* ‘wet’, *řāst* ‘correct’ (mostly, *ř* < **rn*, **rr*, in loans, and initial *r-*) covered in 2.1.2.9.

2.1.2.6 Nasals

These are voiced and show a three-way contrast, *m*, *n*, *ŋ*: *mām* ‘(paternal) uncle’, *ama* ‘this’, *sammūn* ‘bread rolls’, *kām* ‘which?’; *nān* ‘bread’, *zānīn* ‘know’, *dānsāz* ‘dentist’, *ganmasāmī* ‘maize’; and *māŋ* ‘month’, *daŋūbās* ‘information, data’, *baŋ kirdin* ‘call, invite’.

2.1.2.7 Gemination

All consonants except affricates may occur geminate. These are phonetically long and occur only word-medially. Gemination by contact is frequent: *a-yān-nās-im* ‘I know them’, where *-yān* is the 3p suffix.

2.1.2.8 Diphthongs and glides

The glides *w* and *y* combine to form diphthongs freely with all vowels, e.g. *nīw* ‘half’, *sēw* ‘apple’, *āw* ‘water’, *baraw* ‘below’; *kay* ‘when?’.

They are often the regular result of the contraction of final vowels with following unstressed long high vowels: *V + 2s -i> -Vy* in *a-lē-y* ‘you say’, *čū-y* ‘you went’, *a-řū-y* ‘you go’, *āzā-y* ‘you are brave’; *-ō + ū* ‘and’ > *-ōw* in *nō-w yak* ‘nine and one’ (see also section 2.3.1.1 Glide insertion).

y tends to be inserted before initial *ē* when preceded by an affix: *ēs* ‘pain, ache’, but *a-m-y-ēs-ēt* ‘it hurts me’.

2.1.2.9 Distribution

The following continuants do not occur word-initially: the single flap *r*, the velarized lateral *l̥*, and the velar nasal *ŋ*. The short high vowels, *i i u* do not occur in word-final position (but see discussion of *d > i* below). Vowel sequences do not occur (see section 2.3.1.1).

2.1.3 Syllable structure and consonant clusters

Syllable division precedes an intervocalic consonant: *gē.řā.na.wa* ‘narrate’. In syllable division two-consonant clusters are generally divided, as in *sar.gar.dān* ‘confused’.

The minimum syllable is consonant plus vowel, CV, as in *ta.la.ba* ‘student’. Given certain morphophonemic rules (see section 2.3), any two consonants may cluster word-initially, i.e. CCV, except that a stop can only follow another stop or an affricate, glides cannot be first in the cluster, and stops are not followed by nasals: *ktēb* ‘book’, *řfaŋ* ‘rifle’, *řfūr* ‘unveiled’ (woman), *xrāp* ‘bad’. Morphophonemically, however, there are no word-initial consonant clusters. Thus, in the examples above the clusters result from the elision of the non-tense central vowel. Note that in the present description a morphophonemic writing is generally followed.

Phonemic three-consonant clusters CCC do not occur initially but may occur word-medially, in which case they must contain at least one continuant: *qar-z-dār* ‘debtor’, or word-finally, in which case they consist of glide-continuant-stop: *řōyřt* ‘he went’.

2.2 Non-segmental features

Word stress falls on the last vowel of the word unless otherwise marked. The following generalizations, however, can be made: inflectional suffixes, including the suffixed pronouns and the suffixed copula, are not stressed; thus, word stress falls on stem-final vowels.

The secondary noun plural marker *-ān* and the noun marker of definiteness *-akā*, however, are exceptions to this rule. Thus, the stem of the verb *gērā-m-awa* 'I told, related (a story)' is *gērā-* whereas the stem of the noun is *gērā-n-awa* 'relating, telling'.

Vocatives take word stress on the first vowel of the word: *birā* ~ *brā* 'brother', but vocative *birā* 'Brother!'.

On the *phrase or clause level* some words are uttered more loudly than others, whose word stress is then reduced to secondary stress (indicated here by small undercircle) or is suppressed, as in *tāzaya* 'it is new', but *zōr tāzàya* 'it is very new'. Nouns are generally stressed:

řōžēk la řōžān čūyn bō sayrān 'one day we went on a picnic';

so are interrogatives and negatives:

bō a-čīta awē 'why are you going there?'
but *bō nāčīta awē* 'why aren't you going there?'

In a *noun phrase* stress is as follows:

- (1) Dependent attribute. In *izāfa* phrases (those containing the liaison morpheme *-ī*) the final word is stressed: *řōžēkī bāš* 'a nice day'.
- (2) Pre-head modifiers tend to attract stress:
 - (a) interrogatives, like *čī* 'which?, what?', *kām* 'which?', *čāḡ* 'how many?', and the bracket *čī-* . . . (*ēk*) 'what, which?':

čī-pyāw(ēk) hāt? kām pyāw hāt?
'which man came? what man came?';
čāḡ pyāw(ēk) hāt? 'how many men came?';

- (b) numerals and quantifiers: *dū pyāw* 'two men'; *hamū* . . . *ēk* 'each, every'; and *zōr* 'much, many, very':

zōr pyāw hāt 'many men came'.

- (3) Prepositions are generally unstressed, although the preposition *bō* 'for, to' usually does receive sentence stress:

ḡaz akam bičūm bō bāzār 'I'd like to go to the marketplace'.

On the *sentence level*, the following classes of words attract sentence stress:

- demonstratives: *aw darsa saxt bū* 'that lesson was hard';
- adverbials: *umřō zū hal stām* 'I got up early today';
- preverbals: *kay hāt stāy?* 'when did you get up?';
- interrogatives: *kay hātī bō ēra?* 'when did you come here?';
aw kiča kēya? 'who is that girl?';
- negatives: *min azānim, balām aw nāzānē* 'I know, but he doesn't';
na', zōr bāš m-ya 'no, it's not very good'.

2.3 Conditioned and morphophonemic changes

In this section stress will be marked when it plays a significant role in the process; inherently-stressed affixes will also be so marked. The hyphen denotes a morphological boundary and the period marks syllable division.

2.3.1 Vowels

In the following rules it will be noted that the lower a vowel is the more morphophonologically powerful it is. That is, in contact between low and mid vowels, the low vowel is stable and the mid vowel undergoes change, while mid vowels overpower high vowels. Further, long vowels are stronger than short ones, and stressed long high vowels are more stable than unstressed high ones. That is, there is a hierarchy of susceptibility to conditioned change. It is defined by three parameters: (1) articulatory highness; that is, of two vowels in adjacent syllables or in contact, the low vowel is more stable than the mid vowel, and the mid vowel is more stable than the high vowel; (2) relative length; and (3) relative stress. Overall, the vowels of Sul. Kurdish fall into six sets of stability priority, with *ā* being the highest, and the short high vowels the lowest:

TABLE 10.4: HIERARCHY OF VOCALIC SETS

1. long low	<i>lā</i>
2. long mid	<i>lē ō</i>
3. stressed long high	<i>lī ū</i>
4. short low-mid	<i>l a</i>
5. unstressed high	<i>lī ū</i>
6. short high	<i>l i u</i>

Two behaviors are evinced when identical vowels come together: identical high vowels merge whereas identical non-high vowels are separated by glides. The behavior of short *a* is exceptional here in that it behaves like the high rather than the non-high vowels. Phonologically, all long vowels are shortened when unstressed, and high and mid long vowels are shortened and lowered before consonant clusters or before homorganic glides.

2.3.1.1 Glide insertion, *V-G-V*

Two adjacent vowels are connected by an unstable glide (for exceptions see below): *črā-y-ān* ‘lamp-s’, *drō-y-ān* ‘lie-s’, *āzā-y-ī* ‘brave-ry’, *xwē-y-a* ‘it is salt’, *zū-w-a* ‘it is early’.

2.3.1.2 Conditioned shortening

All long vowels, *ī ē ō ū ā*, are phonologically shortened and lowered when unstressed: *ī* > *i* [i]: *šāyī* [ša:yi:] ‘ceremony’, but *hāt-ī* [hɑ:ti] ‘you-2s came’; *ē* > *e*: *a-č-ē* [əčɛ:] ‘he goes’, but *jūr-ē* [jɑ:re] ‘one time, once’; *ō* > *o* [ɔ]: *nō* [no:] ‘nine’, but *bābō* [bɑ:bɔ] ‘father!’ (initial vocative stress); *ā* > *a* [ɑ]: *bābā* [ba:bɑ:] ‘Qalandar dervish’, but *bābā* [ba:ba] ‘indeed, to be sure’.

2.3.1.3 Contraction

Sequences of high vowels and glides contract, (1) to semivowels, *y/w*, when unstressed, and (2) to single long vowel when homorganic:

- (1) *kānī-y-ān* > *kāny-ān* ‘(water) springs’; *nō-w-ū* > *nō-w* ‘nine and . . .’, *čū-y-ī(t)* > *čū-y(t)* ‘you-2s went’, *a-lē-y-ī(t)* > *alē-y(t)* ‘you say’, *āzā-y-ī(t)* > *āzā-y(t)* ‘you are brave’;
- (2) *dī-y-ī(t)* > *dī-(t)* ‘you saw’, *amarakī-y-ūt* > *amarakī-t* ‘you are an American’; *čū-w-ū* > *čū* ‘has gone’.

Exceptions are (1) definite *-aka*; (2) present stems in *-a(h)*; (3) central vowel *i*:

- (1) The definite suffix *-aka* contracts with word-final *-a*, *-ā*, and with the plural ending *-ān*: *talaba + -aka* > *talaba-ka* ‘the student’, *čirā-ka* ‘the lamp’, *čirā-k-ān* ‘the lamps’;
- (2) The 3s ending *-at* contracts to *-ā-t* when attached to the following sets of stems:
- (a) stems in *Ca-*: *a-da + at* > *a-dā-t* ‘gives’; *a-ka + -at* > *a-kā-t* ‘does’; *a-ba + -at* > *a-bā-t* ‘takes, carries’;
- (b) the stems *xwa-* and *řaw-*: *a-xwa- + -at* > *a-xwā-t* ‘eats’; *a-řaw- + -at* > *a-řwā-t* ‘goes’; however, *xwa-*, *řaw-* > *xō-*, *řō-* before the other personal endings or pause, e.g. *a-xō-yn* ‘we eat’, *a-řō-y(t)* ‘2s go’, *bi-xō* ‘eat!’, *bi-řō* ‘go!’.
- (3) Central short *i* is (a) rounded in contact with *w*, and is (b) elided after vowel: *nāw-im* > *nā-w-um* ‘my name’; *čū- + -im* > *čū-m* ‘I went’, *a-xō-in* > *a-xō-n* ‘they eat’; *farmū-i bū* > *farmū bū* ‘had ordered’ (but *xwārd-i bū* ‘had eaten’).

2.3.1.4 Unstressed *i*

Short central vowel *i* is elided in unstressed syllables, except in deliberate speech: *žin-ān* (deliberate register) ~ *žn-ān* (normal register) ‘women’, *a-č-im-a* > *a-č-m-a* ‘I go to’, *dirik-ū-dāl* > *dirk-ū-dāl* ‘thorns and weeds’, *birā* ‘brother’ in vocative, but elsewhere *brā*.

2.3.1.5 Clusters and homorganic glides

The high and mid vowels *ī ō ū*, but not *ē*, are shortened and lowered (1) before consonant clusters and (2) before homorganic glides. Pre-cluster shortening is most notable in present and past stems:

- (1) *dā a-nīs-impldā ništ-im* ‘I sit/sat down’, *a-frōš-impl frōšt-im* [frōšt-im] ‘I sell/sold’, *a-nū-m/nust-im* ‘I sleep/slept’;
- (2) *čī-y-a* > *čī-y-a* ‘what is it?’, *nī-y-a* > *nī-y-a* ‘is not’, *zū-w-a* > *zu-w-a* ‘is early’.

2.3.2 Consonants

Conditioned variation increases from deliberate to normal to rapid speech. Consonants may undergo assimilation, occur intrusively, or be elided, depending on the environment.

2.3.2.1 Voice assimilation

Stops are fully released in word-final position. Voiced stops and fricatives may be partially or fully unvoiced in final position; *sag* [sæg ~ sæk] ‘dog’, *sig* ~ *sik* ‘stomach’, *xālūs* ~ *xālūs* ‘charcoal’.

Voiced and voiceless stops (but not *q*) and sibilants undergo partial regressive assimilation in any position: stop, *bičūk* > *p-čūk* ‘little’; *čī bi-ka-m* ~ *čī p-ka-m* ‘what shall I do?’; sibilant, *xōš bū* > *xōž bū* ‘it was fine’; *pāš nīw-a-řō* > *pāž nīw-a-řō* ‘after noon’.

Voiceless stops (but not *q*) are slightly aspirated in word-initial position: *pān* [p^ha:n] ‘wide’, *talaba* [t^hələbə] ‘student’, *kitēb* [k^hte:b] ‘book’, and optionally so in word-final position.

2.3.2.2 Palatalization and fronting

In informal speech, the velar stops and the labials except *b* and *m* are fronted before front vowels and *y*.

(1) *k g* become the dorso-palatal affricates [tʃ] [dʒ] in contrast with the lamino-postalveolar affricates *č* [tʃ] and *j* [dʒ]. To the uninitiated ear the palatalized stops are easily confused with their corresponding affricates: *kē* [k^hē:] ‘who?’, *gīyān* [g^hi:yā:n] ‘soul’, *agīnā* [ag^hi:nā] ‘otherwise’, *čāk-ī* [tʃa:k^hi:] ‘are you well?’, *mumkīn* [mumk^hi:n] ‘possible’;

(2) *p, w, f*: *pēnj* [p^he:nʃ] ‘five’, *fēnik* [f^he:n^hk] ‘cool’.

2.3.2.3 Fronting of *w*

wērān [w^he:ra:n] ‘ruin’; *āw-ī* [a:w^hi] ‘his water’, *k^hwē* [k^hw^he:] ‘where?’; *k^hwer* [k^hw^he:r] ‘blind’, *g^hwē* [g^hw^he:] ‘eye’.

2.3.2.4 Dentals

Before pause the final *-t* of the personal endings 2s *-ī(t)*, 3s *-at*, *-ēt* and the personal enclitic 2s *-it* may be elided: *a-č-īt* > *a-č-ī* ‘you go’, but *a-č-īt-awa* ‘you go back’; *dast-it* - *dast-i* ‘your hand’; *b-ē(t)* ‘may 3s be’, but *a-b-ēt-awa* ‘becomes’, *hāt-ū-w-a* ‘has come’, but *hāt-ū-w-at-awa* ‘has come back’.

2.3.2.5 The “Zagros *d*”

As a widespread regional feature, termed the “Zagros *d*” (Windfuhr), postvocalic *d* is softened to glide-like *-i-*, or *-w-*, and contracts with adjacent high vowels: *a-da-m* ~ *a-ia-m* ‘I give’; *nādir* ~ *nāir* ‘Nadir’ (masc. proper name); *bad* ~ *bai* ‘bad’; *xwā* ~ *xuwā* ‘God’ < **xudā*. Note that all vocalic past stems originate in *-Vd*: *farmī-* < **farmūd-* ‘ordered’ (see section 3.2.1 Stem formation).

2.3.2.6 Nasal assimilation

Elsewhere, *d* assimilates to a preceding nasal: *qalam-dān* ~ *qalam-mān* ‘pencil box’; *lē-m dā-yt* ~ *lē-m mā-yt* ‘I hit you’ (past tense); *řasm-aka-y pišān dā-m* ~ *řasm-aka-y pišān nā-m* ‘he showed me the picture’.

The sequence *nd* may alternate with contracted *n*, or with contracted *ŋ*: *afandī* ~ *afanī* ‘gentleman’; *māndū* ~ *māŋū* ‘tired’, *čand* ~ *čan* ~ *caŋ* ‘how many; several’, *band-a-xwēn* ~ *banaxwēn* ~ *baŋaxwēn* ‘belt-cord’.

The dental may be lost before *z* and the vowel nasalized, most prominently in the number terms for teens: *bānzīn* [bānzīn] ‘benzine, gasoline’, *yānzā* [yānzā] ‘eleven’, *dwānzā* [dwānzā] ‘twelve’, *sēyānzā* [seyānzā] ‘thirteen’, *pānzā* [pānzā] ‘fifteen’, *šānzā* [šānzā] ‘sixteen’.

2.3.2.7 Intrusive *d*

d is often intrusive in sequences of *-nr-*: the passive perfect participles *čēn-rā-w* ~ *čēn-d-rā-w* ‘sown’; *diřen-rā-w* ~ *dřen-d-rā-w* ‘torn’.

2.3.2.8 Deletion of *h*

h may be lost in non-initial position: *řōž-hal-hāt* ~ *řōž-al-āt* 'sunrise; east'; *na-hāt* ~ *na-āt* 'he didn't come'; *bi-hēn-a* ~ *b-ēn-a* 'bring!'.

3 MORPHOLOGY

Kurdish has the morphologically defined word classes nouns, adjectives, pronouns, verbs, and particles. Particles, which are indeclinable, are subclassed syntactically into interjections, interrogatives, conjunctions, prepositions, adverbs, and preverbals; a closed class, all particles except interjections and conjunctions may receive suffixed pronouns. Particles are found in Syntax in conjunction with the syntactically relevant feature.

3.1 Nominal morphology

In Sulaimania Kurdish grammatical gender has been lost. So have the case distinctions in nouns and pronouns, except for a few remnant traces, and have been replaced by a complex system of person marking. In turn, definiteness has evolved as a major category.

3.1.1 Nouns

3.1.1.1 Gender, number, and definiteness

Gender of animate nouns may be shown lexically in male/female pairs: *gā* 'ox' and *māŋ-gā* 'cow' (< **mādiyān* 'female').

Nouns may be inflected for number and definiteness; some nouns also may show vocative case or relic locative case endings. The bare stem may signify indefinite non-specific singular or a generic plural; *-ēk* signals indefinite specific singular, and *-akā* marks definiteness; *-ān* denotes plural, and follows the definite suffix, if any. The deictic envelopes *amlaw . . . a* 'this/that', inherently definite, may cover a noun with or without the plural suffix. These are all illustrated with *pyāw* 'man' in the following synoptic table (arrangement suggested by Windfuhr):

TABLE 10.5: SPECIFICITY AND DEFINITENESS

	Indefinite	Indefinite Specific	Definite	Near/Far Deictic
Singular	<i>pyāw</i> 'man'	<i>pyāw-ēk</i> 'a man'	<i>pyāw-akā</i> 'the man'	<i>amlaw-pyāw-a</i> 'this/that man'
Plural	<i>pyāw</i> 'some men'	<i>pyāw-ān</i> 'men'	<i>pyāw-ak-ān</i> 'the men'	<i>amlaw pyāw-ān-a</i> 'these/those men'

3.1.1.2 Vocative

Singular names may receive the unstressed vocative suffix *-a* (*-wa* after vowels) and, after plural nouns, *-in-a*, in direct address; polysyllables, especially those of Iranian origin, take word-initial stress: *hēro-w-a*, *wāi-a bō ēra!* 'Hero (fem.), come here!' *xalk-in-a!* 'People! Everybody!'

3.1.1.3 Locative

The relic locative suffix *-ē* is found on a few nouns of place or time: *šār-ē* ‘in the city’, *šaw-ē* ‘in the evening’, with the form *-n-ē* after a vowel: *sibay-n-ē* ‘tomorrow’ (cf. *sibay* ‘morning’, *dū-sbay* ‘day after tomorrow’).

3.1.2 Adjectives and degree

Adjectives, like nouns, can receive the definite suffix, the indefinite suffix and the plural suffix and, additionally, can be made comparative or superlative: *čāk* ‘good’; *čāk-ēk* ‘a good one’; *čāk-aka* ‘the good one’; *čāk-ak-ān* ‘the good ones’.

Degree is indicated by the comparative suffix *-tīr*, with *la* ‘from, than’ introducing the compared item; the superlative form in *-trīn* precedes the noun head:

aḥmad zīrak-tīr-a la kāwā; la hamī-yān zīrak-trīn talaba-y-a.

‘Ahmad is smarter than Kawa; he is the smartest student of (*la*) all of them.’

3.1.3 Pronouns and deixis

3.1.3.1 Personal and demonstrative pronouns

In addition to the personal pronoun, which may be (1) independent, or (2) clitic, there are also reflexive pronouns; for interrogative pronouns see section 3.1.3.4. Independent and clitic suffixed pronouns show distinction of person and number, but are otherwise invariable. Suffixes lose *-i* and *-ī* > *y* in postvocalic contractions.

TABLE 10.6: INDEPENDENT AND DEMONSTRATIVE PRONOUNS

	1st pers.	2nd pers.	3rd pers.	‘that’	‘this’
Indep. sg.	<i>mīn</i>	<i>tō</i>	<i>aw</i>	<i>aw-a</i>	<i>am-a</i>
Indep. pl.	<i>ēma</i>	<i>ēwa</i>	<i>aw-ān</i>	<i>aw-ān-a</i>	<i>am-ān-a</i>
Enclitic sg.	<i>-iml-m</i>	<i>-itl-t</i>	<i>-īl-y</i>		
Enclitic pl.	<i>-mān</i>	<i>-tān</i>	<i>-yān</i>		

The 2s may receive the nominal emphatic vocative clitic: *ma-ka*, *tō-w-a* ‘don’t do (that), you!’.

- (1) The independent pronouns are mostly used for emphasis or contrastive focus. They occur as:

(a) subjects; (b) objects of prepositions; and (c) possessors in *izāfa* noun phrases.

- (2) The clitic pronouns may be suffixed to nouns to denote possessor: *nāw-im* ‘my name’. They may be infixes between predicate and copula as experiencer or beneficiary:

[bō čī]-t-a? ‘what is that for you?’ = ‘what do you want with that?’;

[bas]-t-a? ‘is (that) enough for you?’.

They also serve as pronoun objects of transitive verbs in the present tense, as in *a-yān-nās-im* ‘I know them’ (*-yān*); and as the subject markers of past tense transitive verbs (see section 4.4.4 Past agent suffixes).

- (3) The *demonstrative pronouns* manifest a one-level contrast, proximal *am-a* ‘this, this one’ and distal *aw-a* ‘that, that one’. The separability of *-a* appears in their attributive functions, when they become the discontinuous envelopes *am . . . a* ‘this’ and *aw . . . a* ‘that’:

[*am-a*] *bāš-a-w* [*aw-ān-a*] *bāš nī-n*
 ‘this one is good and those are not good’,
 but [*am*] *qalam-[a]* *bāš-a-w aw-qalam-ān-a zōr bāš nī-n*
 ‘this pen (*qalam*) is okay; those pens are not very good.’

3.1.3.2 Possessive forms

Independent pronominal possession is expressed by the *izāfa*-construction *hī* + the personal enclitics, or possessor: *hī min* ‘mine’, *hī dāyk-im* ‘my mother’s’, etc. (see 4.2.2.1).

3.1.3.3 Reflexive-emphatic pronouns

Reflexive-emphatic pronouns consist of *xō* ‘self’ plus a clitic pronoun: *xō-m* ‘myself’, *xō-t* ‘yourself’, etc.: reflexive, *xō-tān āmāda ka-n* ‘get yourselves ready!’; emphatic: *xō-t a-zān-ī* ‘you know’. – *yak-tir* ‘each other’ serves as a *reciprocal pronoun*.

3.1.3.4 Interrogatives and related adverbials

Interrogatives: *kē* ‘who?’, *čī* ‘what?’, *kām* ‘which?’, *čand* ‘how many?’, *kwē* ‘where?’, *kay* ‘when?’, *čōn* ‘how?’, *bō čī ~ bō?* ‘for what, why?’, *kwā* ‘where is . . .’ typically receives a copula suffix as subject, as in *kwā-n?* ‘where are they?’, *kwā kuř-aka-m?* ‘where is my son?’ (see also sections 3.1.5 Adverbs and 3.1.6.3 Quantifiers).

3.1.4 Adpositions

3.1.4.1 Pre- and postpositions

The basic adpositions include:

- (1) Two prepositions:
 - (a) the multivalent *ba* ‘in, at; by; to’; and
 - (b) the bipolar *la* ‘in’ or ‘from’: *ba kurdī* ‘in Kurdish’, *la karkūk* ‘in Kirkuk’, *la čāk-ak-ān* ‘from the good ones’.
- (2) The postverbal allative clitic *-a* ‘to’: *a-č-īn-a karkūk* ‘we are going to Kirkuk’;
- (3) Two postpositions:
 - (a) *-dā* ‘in’, and
 - (b) *-awa* ‘from’, only combined with prepositions.
- (4) Other adpositions include:
 - (a) *bō* ‘for’, also indirect object; goal, *bō min ~ bō-m* ‘for me’: *a-č-im bō bayā* ‘I’m going to Baghdad’, *a-y-da-m bō tō* ‘I’m giving it to you’;
 - (b) *tā*, *hatā*, *hatākū* ‘till, as far as’: *goyšt-in hatā karkūk* ‘they arrived as far as Kirkuk’;

- (c) *bē, ba-bē . . . (-awa)* ‘without’: *bē īs* ‘without work, unemployed’;
- (d) *la gal . . . (-dā)* ‘together with’, with *d-* often elided or assimilated, e.g. *kē hāt lagat-t-ā* ‘who came with you’;
- (e) *la-bar* ‘on; because of’;
- (f) *bar-awa* ‘towards’;
- (g) *wak, wakū* ‘like’: *wakū min* ‘like me’.

3.1.4.2 Absolute forms

The first four adpositions listed above, *ba, la, -a, dā*, have absolute forms marked by *-ē* and devoicing: *pē, lē, -ē ~ -r-ē* (postvocalic), *tē*. While *ba* and *la* occur with following nouns, independent pronouns or with other prepositions, *pē* and *lē* occur elsewhere, and may take the personal enclitics, thus *ba min* > *pē-m*, *la awān* > *lē-yān*, e.g. *lē-t a-d-ā* ‘he hits (at) you’. It also occurs in infinitives, e.g. *pē kan-īn* ‘to laugh’.

bō ‘for, to’ corresponds to *-a* + noun, and postverbal *-(r)ē*:

- kay a-ē-īt-a bayālbō bayā?* ‘when are you going to Baghdad?’;
- bō min bi-da* ‘give me (some)!’ vs.
- bi-m-da-r-ē-n* ‘give them (*-in*) to (*-r-ē*) me!’ (*-im*);
- dinār-ēk-it a-da-m-ē* ‘I’ll give to (*-ē*) you (*-it*) a dinar’.

3.1.4.3 Circumpositions

ba and *la* form circumpositions with *-dā* and *-awa*. Note that *balla* lose their vowel before the initial vowel of the demonstratives pronouns, *aw-lam-*, and some adverbs, *awē* ‘there’, *ēra* ‘here’, e.g. *l-ēra* ‘in this place, here’. Similarly *-dā* > *-ā* after consonants:

(1) *ba* and *la*:

- (a) *ba . . . awa* ‘with (instrument)’: *ba xwē-y-awa* ‘with salt’; *ba qalam-awa b-ī-nūs-a* ‘write it with a pen (*qalam*)!’;
- (b) *la . . . dā* ‘in’: *la karkūk-(d)ā* ‘in Kirkuk’;
- (c) *la . . . awa* ‘from’, *la karkūk-awa bō bayā* ‘from Kirkuk to Baghdad’, *l-aw-lā, awa* ‘on (“from”) that (*aw*) side (*lā*)’.

(2) Other circumpositions include:

- (a) *ba . . . dā* ‘through’: *ba bāzār-ā čū-yn* ‘we went by way of the bazaar’;
- (b) *la-gal . . . (dā)* ‘together with’;
- (c) *bō . . . awa* ‘through’.

These circumpositions thus also serve to disambiguate the bipolar *la* ‘in’ and ‘from’: *la karkūk* ‘in/from Kirkuk’, but *la karkūk-ā* ‘in Kirkuk’ (< *-dā*) vs. *la karkūk-awa* ‘from Kirkuk’.

Note that *la . . . dā* may function to express *sensations*, with or without temporal-locative *-ān*: *la tīnū-dā* ‘from thirst’, *la tirs-ān-ā* ‘from, with fear’.

3.1.4.4 Compound absolute forms

While *-ē* (< *-a*) is always enclitic to the verb, there are compound absolute forms with *-awa* and *-dā*; thus: *ba-awa* > *pēwa* ‘in, at’; *ba-dā* > *pēdā* > *pyā*: *pyā-m-ā* ‘at, on me’; *dā-awa* > *tēwa*; *dā-dā* > *tēdā* > *tyā* ‘inside’.

Combined with *-ēk*, there are: *pēk*, *pēkawa* ‘together’; *lēk* ‘from each other’; *tēk* ‘together’.

These, in turn, may be attributive, like *tyā* in *naxša-y-ēk-ī-kurdustān[-ī-tyā]-y-a* ‘there’s a map (*naxša*) of Kurdistan in it’.

Both simple and derived absolute forms may also function as preverbs (see section 3.2.1.3).

3.1.5 Adverbs

3.1.5.1 Lexical adverbs

Adverbs as a word class are essentially lexical items, without general marking patterns. Some adverbs have opaque deictic markers: *ēra* ‘here’, *awē* ‘there’, *ēstā* ‘now’, with appositional forms such as *l-ēra* ‘in this place, here’, *l-ēra-wa* ‘from here, hence’, *bō ēra* ‘hither’, *l-awē* ‘in that place, there’.

Some are derived, such as *dwā-y-ī* ‘later’. Others are compounds with or without adpositions, often Persian in origin, such as *yak-sar* ‘immediately’; *damdā* ‘from time to time’, *dar-hāl* ‘immediately’, or Arabic loans such as *xāsat-an* ‘especially’. Others include: *injā* ‘then; in that case’, *hēstā* ‘still, yet’ (with negative verb), *tanhā* ‘only, alone’, *zōr* ‘very’.

A specific subset is terms of time relations, such as *im-īō* ‘today’, *im-šaw* ‘this evening’, *pār* ‘last year’, *pāš-a-īōz* ‘in the future’, *sibhay-n-ē* ‘tomorrow’.

har ‘just, only’ is selective-restrictive: *har l-awē* ‘only over there’; *har ēstā* ‘right now’; *har kabāb a-xō-m* ‘I eat only kabob’ and *min kabāb har a-xō-m* ‘I do nothing but eat kabob’.

wā ‘thus, like this’, (< OIr. **awa-θā*) functions as the manner adverb, with regular adverbial stress: *wā hāt-im* ‘I came like this’, *wā a-xō-m* ‘I eat thus, this is how I eat’. It is distinct from the unstressed locative-temporal verbal particle *wā* ‘here, now’ (< OIr. **awa-dā*): *wā hāt-im* ‘I’m coming, I’m on my way’, *wā a-xō-m* ‘I am now eating’, *wā-n la ērān-dā* ‘they are (-in) now (*wā*) in Iran’.

3.1.5.2 Adverbial phrases

Prepositional phrases and nouns with inherent adverbial meaning, with or without adpositions, may function adverbially:

(1) Prepositional phrases:

pēš sa’āt dū ‘before 2:00’; *la pēš nān xwārd-in* ‘before eating’; *pāš nīw-a-īō* ‘after noon’, and *la pāš awa* ‘(from) after that’, *dwāy* ‘after (time)’: *dwāy awa čū-n bō bāzār* ‘after that they went to the market’.

(2) Nouns:

ba sar . . . *dā* ‘over’ (*sar* ‘head, top’); *la nāw* . . . (*dā*) ‘within, inside of’ (*nāw* ‘inside’); *la nāw* . . . (*awa*) ‘from out of’; *la bābat* . . . (*awa*) ‘about, concerning’ (*bābat* ‘item, topic’); *witār-ēk la [bābat šīr-ī-kurdī-y]-awa* ‘a talk on Kurdish poetry’.

3.1.5.3 *Izāfa* construction

While such phrases precede their dependent directly, they themselves may follow nouns or adverbials, in which case the dependent is attached by regular *izāfa* (see Section 4.2

Noun phrase structure): *la-bar dam-ī X* ‘in front of X’ (*dam* ‘mouth’); *la mā beyn-ī* ‘in between’.

3.1.6 Numerals

3.1.6.1 Cardinal numbers

TABLE 10.7: CARDINAL NUMBERS

1–9	11–19	10–90	100–900	10 ⁿ
<i>yak, -ēk</i>	<i>yānza</i>	<i>da</i>	<i>şad</i>	<i>hazār</i> ‘thousand’
<i>dū</i>	<i>dwānzda</i>	<i>bīst</i>	<i>dū-şad</i>	<i>milyōn</i> ‘million’
<i>sē</i>	<i>syānzda</i>	<i>sī</i>	<i>sē-şad</i>	
<i>čwār</i>	<i>čwārda</i>	<i>čil</i>	<i>čwār-şad</i>	
<i>pēnj</i>	<i>pānza</i>	<i>panjā</i>	<i>pēnj-şad</i>	
<i>şaş</i>	<i>şānza</i>	<i>şaşt</i>	<i>şaş-şad</i>	
<i>hawt</i>	<i>havda</i>	<i>haftā</i>	<i>haw-şad</i>	
<i>haşt</i>	<i>hažda</i>	<i>haştā</i>	<i>haş-şad</i>	
<i>nō</i>	<i>nōzda</i>	<i>nawad</i>	<i>nō-şad</i>	

Numbers higher than ‘1’ are followed by the singular nominal; however, the plural is marked in the verb, e.g. with interrogative-indefinite *čand* ~ *čaŋ*: *čaŋ pyāw (-ēk) hāt-in?* ‘how many men came?’ and *dū pyāw hāt-in* ‘two men came’.

In time telling, the noun *saāt* ‘hour; clock, watch’ precedes the number without connective: *saāt dū* ‘two o’clock’ vs. *dū saāt* ‘two hours’ or ‘two watches’, etc.

Numbers may be nominalized by *-ān*, e.g. *dū-y-ān hāt-in* ‘two (of them) came’, *sē-y-ān* ‘three of them’.

3.1.6.2 Ordinal numbers

Ordinal numbers are marked by stressed *-am*: *yak-am* ‘first’, *čwār-am* ‘fourth’. The selective ordinal adds *-īn*: *yak-am-īn* ‘the very first’, *čwār-am-īn* ‘the fourth’, and precedes the head: *am-a sē-y-am(-īn) kōŋra-y-a* ‘this is the third congress’.

3.1.6.3 Quantifiers

Quantifiers are uninflected nominals that signify amount or number, and may function independently or as pre-head noun modifiers. They typically are followed by a singular noun which has the indefinite suffix *-ē(k)*; they receive phrasal stress, with the following noun receiving secondary stress. They include *hamū* . . . *-ē(k)* ‘each, every’: *hamū řōž-ēk* ‘every day’; *čand* ~ *čaŋ* ~ *čan* . . . *-ē(k)* ‘some, a few’: *čaŋ pyāw-ēk* ‘a few men’; *gal-ē(k)* ‘a great number of’, e.g. *gal-ē jār* ‘many times, often’; *har* . . . *-ē(k)* ‘each, any’: *har yēk-ēk* ‘each one, any one’; *hīč* ‘any’, with negative verb ‘no’: *hīč kas na-hāt* ‘nobody came’. *l-am čand-ān-ā-dā* ‘in the past few days, recently’ illustrates independent usage.

The quantifier *zōr* ‘many’ may be linked to a preceding noun head by *izāfa* but is usually preposed without *izāfa*, e.g. *zōr pyāw* ~ *pyāw-ī zōr* ‘many men’.

There is a miscellaneous group of nouns that function as numerals or with numerals.

- (1) The classifier *sar* ‘head’ is used in a numeral phrase as a pre-head nominal modifier: *panjā sar mar* ‘fifty head of sheep’.
- (2) *dāna* ‘item, thing’ (cf. *dān* ‘grain’) functions with *yak* ‘one’ as a pronoun: *čand-yān hāt-in?* ‘how many (of them) came? – *yak dāna* ‘one.’
- (3) *tā* ‘one of a pair; half’ functions as a noun: *tā-ī dargā* ‘leaf of a double door’.

3.2 Verb morphology

Verbs are inflected for aspect, mood, tense, person and number, and show distinctions of transitivity and voice. All verbs may be made negative, and transitive verbs may receive an enclitic pronominal object, while intransitive verbs may be made causative.

3.2.1 Stem formation and modifiers

3.2.1.1 Present and past stems

Verbs are based on two stems, present and past. Past stems are distinguished typically by the dental stop: *frōš-/frōšt-* ‘sell’, *mir-/mird-* ‘die’, sometimes accompanied by difference in voicing: *kuž-/kušt-* ‘kill’; or additional segments: *rō-/rōyšt-* ‘go’, *awē-/wist-* ‘want’; or ablaut: *bižēr-/bižārd-* ‘choose’. Others may show no overt distinction, such as *past-/past-* ‘press’; or are suppletive, such as *tē-/wt-* ‘say’, *ē-/hāt-* ‘come’.

Typical for Kurdish are also vocalic past stems in *-ī, -ū, -ā* (< **-īd, *-ūd, *-ād*), where the inherent dental *-d* is obscured by postvocalic softening and contraction (“Zagros-*d*”): *larz-/larzī-* ‘tremble’, *farmū-/farmū-* ‘command’, *tirs-/tirsā-* ‘fear’. “Regular” verbs have *-ē-/ā-*: *garē-/garā-* ‘stroll’.

3.2.1.2 Morphological passive and causative

(1) Passive stems are derived from present stems by *-r-ē/r-ā* (< **-r-ād*): *kuž-rē-/kuž-rā-* ‘be killed’. They are inflected with intransitive subject markers: *a-kuž-r-ē-m* ‘I am, will be killed’; *kuž-r-ā-w-im* ‘I have been killed’; *agar bi-kuž-r-ē-m* ‘if I am killed’; *agar bi-kuž-r-ā-m-ā-y-a* ‘if I had been killed’.

Stems in *Ca- > Ci-r/Cu-r*; thus: *ka-* ‘do’ > *ki-r-ē-/kir-ā-* ‘be done’, *xwa-/xō-* ‘eat’ > *xu-r-ē-/xu-r-ā-* ‘be eaten’, *da-* ‘give; to hit’ > *di-r-ē-/di-r-ā-* ‘be given; be hit’.

(2) Causative stems are derived by *-ēn/-ān(d)-*: *tirs-* ‘be afraid’ > *tirs-ēn-/tirs-ānd-* ‘make fear, frighten’, *ga-* ‘arrive’ > *ga-y-ēn-/ga-y-ānd-* ‘make arrive, bring to’; *žī-* ‘live’ > *žī-ēn-/žī-y-ānd-* ‘bring to life, give birth’. Adpositions are attached regularly: *žī-y-ānd-in-awa* ‘revive’.

3.2.1.3 Preverbs and postverbs

Preverbs modify verbal meanings or create new lexical items. They include two main sets: inherited prefixes and absolute prepositional forms (see section 3.1.4 Adpositions). They form a closed class, numbering a dozen or so at most. In the following listing, illustrative combinations are mostly given with the verbs *čū-n* ‘go’, *hāt-in* ‘come’, *hēn-ān* ‘bring’:

TABLE 10.8: PREVERBS AND PREFIXES

Preverb	<i>çû-n</i> 'go'	<i>hât-in</i> 'come'	<i>hênâ-n</i> 'bring'
<i>hal</i> 'up'	'boil over'; 'jump over'	'rise (sun); 'run away'	'produce'; 'hatch (egg)'
<i>dâ</i> 'down'	'sag'; 'lose prestige'	'come down'; 'be new'	'bring down'; 'invent'
<i>rô</i> 'down'	'go down'; 'sink down'		
<i>râ</i> 'forth, away'	'go down'; 'leak'	'become used to'	'train, tame'
<i>dar</i> 'out'	'go out, appear'	'come out, emerge'	'bring out, produce'
<i>war</i> 'from' + <i>girt-in</i>	'get from, receive'		

TABLE 10.9: PREVERBS: EXTENDED PREPOSITIONS

Preposition	<i>çû-n</i> 'go'	<i>hât-in</i> 'come'	<i>hênâ-n</i> 'bring'
<i>pêdâlpÿâ</i> 'across'	'penetrate into'	'come through, traverse'	'rub on'
<i>pêwa</i> 'in, on, with'	'fit into'	'bring luck to'	
<i>pêk</i> 'together'	'quarrel with each other'	'come together, agree, be formed'	'bring together, unite, create'
<i>tê</i> 'in, into, on'	'enter, pierce'		
<i>têk</i> 'together'	'fall together'; 'lose composure'		

Both subsets may combine: [*tê hal*] *dâ-n* 'kick (someone)', [*têk war*] *dâ-n* 'stir' (lit. 'hit together', cf. Persian *be ham zadan*).

3.2.1.4 Adpositions

Prepositions as well as postpositions may similarly modify verbal meanings; note that with infinitives the absolute form is used:

lallê: *lê dâ-n* 'hit (at), play (instrument)'; *lê girt-in* 'receive, get from'; *lê sand-in* 'take from'; *balpê*: *pê dâ-n* 'give to'; *pê kanî-n* 'laugh at'; *pê wit-in* 'say to, tell'; *bû-n ba* 'become, come to be';

-*al-ê* postvocalic -*r-ê*: *çû-n-a* 'go to'; *dâ-n-ê* 'give to';

-*awa*, postvocalic -*r-awa*: *dâ-n-awa* 'give back, return (s.th.)'; *xwârd-in-awa* 'drink' vs. *xwârd-in* 'eat'; *kird-in-awa* 'do again; to open'; *çû-n-awa* 'return, go back';

hât-in-awa 'come back/again'; *bû-n-awa* 'become, happen to'.

Note that *ba*, *la*, *-a* > *pê*, *lê-*, *-ê* occur with personal enclitics (see section 3.1.4 Adpositions).

There is correspondence between the prepositions *bô* and *-a*: *-a* is found on verbs of movement: *bôlba* + *çû-n* 'go to' vs. postverbal *çû-n bô* ~ *çû-n-a* as in *kay a-ç-î [bô bayâ]? ~ kay a-ç-ît[-a bayâ]? 'when will you go to Baghdad?'*; cf. *hât-in[-a dî]* 'come into view, be realized'.

The absolute form of $-(r)-\bar{e}$ ($< -a$) is found with verbs of giving such as $d\bar{a}-n$ ‘give’: $b\bar{o}$ $min\ bi-da-[r-\bar{e}] \sim bi-m-da-[r-\bar{e}]$ ‘give it to me!’

3.2.1.5 Compound verbs

A large number of verbs are compound, consisting of a nominal with a small subset of function verbs, mostly $kird-in$ ‘do, make’ and $b\bar{u}-n$ ‘become’: $\check{c}\bar{a}k\ kird-in$ ‘make good, improve’.

Object complements may be introduced by: (1) Correlated prepositional phrases, such as $gw\bar{e}\ la\ X\ girt-in$ ‘take ear to X, listen to’; or (2) $iz\bar{a}fa$ constructions: $b\bar{a}s-\bar{i}\ X\ kird-in$ ‘make discussion of, discuss X’; $\check{c}\bar{a}wa\check{r}\bar{e}-\bar{i}\ X\ kird-in$ ‘make expectation of, wait for X’.

Lexical intransitive vs. transitive function is found with certain pairs of verbs, such as $h\bar{a}t-in/h\bar{e}n\bar{a}-n$ ‘come’/‘bring’: $h\bar{a}t-in-a\ d\bar{i}$ ‘come into being, be realized’ and $h\bar{e}n\bar{a}-n-a\ d\bar{i}$ ‘bring into being, create’.

3.2.2 Nominal forms

The two most prominent nominal forms are the infinitive (“verbal noun”) and the past participle (“verbal adjective”), both of which are derived from the past stem. Morphologically, they are nouns or adjectives respectively, and syntactically they have the force of verbs as well as that of nouns or adjectives. Semantically, infinitives name the notion of the verb from which they are derived and serve as the citation form of the verb. Participles denote the resultant state on the completion of the action of the verb.

3.2.2.1 Infinitive

Infinitives are marked by $-in$: $kird-in$ ‘make, do’, or $-n$ after postvocalic contraction: $ga\bar{r}\bar{a}-n$ ‘stroll’ ($< *ga\check{r}-\bar{a}d-in$).

3.2.2.2 Past participles

Past participles are marked by $-\bar{u}$, which forms diphthongs or contracts with preceding vowels: $xw\bar{a}rd-\bar{u}$ ‘(having) eaten’, $kawt-\bar{u}$ ‘(having) fallen’, $b\bar{i}n\bar{i}-w$ ‘having seen’ ($< *b\bar{i}n-\bar{i}d-\bar{u}$), $b\bar{u}$ ‘(having) been’ ($< *b\bar{u}d-\bar{u}$).

3.2.3 Person marking and ‘to be, become’

The following is an overview of person marking. It combines the existential verb and the copula, with the personal endings and agent markers, and the verb ‘to be, become’ in the various tenses. Note that the markers of the 2p and 3p have merged except in the transitive past tenses, where the agent markers are in fact the personal enclitics.

TABLE 10.10: PERSON MARKING AND ‘TO BE, BECOME’

	1s	2s	3s	1p	2p	3p
Existential	<i>ha-m</i>	<i>ha-y(t)</i>	<i>ha-y-a</i>	<i>ha-yn</i>	<i>ha-n</i>	<i>ha-n</i>
Copula	<i>-im</i>	<i>-i(t)</i>	<i>-a(t)l-\bar{e}t\bar{i}</i>	<i>-in</i>	<i>-in</i>	<i>-in</i>
Present, Subjunctive	<i>-im</i>	<i>-i(t)</i>	<i>-\bar{e}(t)l-\bar{a}(t)</i>	<i>-in</i>	<i>-in</i>	<i>-in</i>
Imperative		<i>-a</i>			<i>-in</i>	

	1s	2s	3s	1p	2p	3p
Past intrans.	<i>-im</i>	<i>-î(t)</i>	<i>-ø</i>	<i>-în</i>	<i>-in</i>	<i>-in</i>
Past trans.	<i>-im</i>	<i>-it</i>	<i>-î</i>	<i>-mân</i>	<i>-tân</i>	<i>-yân</i>
'be, become'	1s	2s	3s	1p	2p	3p
Subjunctive	<i>b-im</i>	<i>b-î(t)</i>	<i>bê(t)</i>	<i>b-în</i>	<i>b-in</i>	<i>b-in</i>
Preterite	<i>bû-m</i>	<i>bû-y(t)</i>	<i>bû</i>	<i>bû-yn</i>	<i>bû-n</i>	<i>bû-n</i>

Regarding *(t)* in the 2s, 3s endings, it appears before vocalic clitic in 2s *-î(t)* and in 3s *-a(t)*, *-ê(t)*, *-â(t)*.

In the 3s, the copula *-a(t)* 'is' and a preceding 3s suffix *-î* combine to *-êîî*: *pyâw-î + a (-î* as possessor) *> pyâw-y-êîî* 'he is his (*-î*) man'.

Similarly, in present perfect forms the 3s copula *-a(t)* appears as *-a* when in final position: *xwârd-û-y[-a]* 'he has eaten', but *-t* appears when followed by a postposition: *hâtû-w-im[-at]-awa* 'I have returned', *xwârdû-y-[-êî]-î* 'he has eaten it' (*y < 3s agent -î*).

The 3s ending *-â(t)* occurs after present stems in *-a*: *ka-at > k-ât* 'does', *xwa-at > xw-ât* 'eats' (see section 2.3.1.3 Contractions).

The imperative 2s *-a* is absorbed by the stem vowel: *bên-a* 'bring!', but *hî-îô* 'go!'. The verb *ê-lhât-in* 'come' has a suppletive stem in the positive, *war-a!*, but prohibitive *ma-y-ê* 'don't come!'.

The change-of-state forms of *b-* 'be, become' regularly take the imperfective prefix *a-* and the subjunctive-imperative *bi-*:

- (1) 'exist, be there':

êmâ-š ha-yn! 'we also exist' (don't forget us!); but
gal-ê kas ha-bû, *balam êstâ hîc kas nî-y-a*
 'there was a lot of people, but now there is no one';

- (2) 'be, become':

bâš a-b-ê '(that) will be fine';
çî a-b-êî-awa?
 'what will become of it?' (lit. 'what will it become').

The forms of the present perfect of *b-* 'be, become' are identical with the preterite forms due to contraction: *bû-* 'was/were' and 'has/have been' (< pret. **bûd*, < perf. **bûd-û*); the past perfect has *bû-bû-* 'had been' (< **bûd-i bûd*).

3.2.4 Aspectual-modal and negative prefixes

There are two aspectual-modal prefixes, as well as a negative and a prohibitive prefix:

a-, nâ- imperfective (< *na-a-*): *a-îô-y* 'you go' – *nâ-îô-y* 'you don't go';
bî-, na- subjunctive, conditional: *bî-îô-y* '(if) you go' – *na-îô-y* '(if) you don't go';
bî-, ma- imperative, prohibitive: *bî-îô* 'go!' – *ma-îô* 'don't go!'

In the present indicative of 'to come', the imperfective marker and the stem *ê-* merge to 'to *yê-*': *yê-m* 'I come'.

bî- may be omitted, particularly with nominal preverbs: *lâ çô* 'go away!' for *lâ bî-çô*, and tends to show vocalic contraction: *b-ê-t* 'that he come' < *bî-ê-êî*; but *bî-* marks the

distinction between state and process with ‘be’; thus, *b-ēt* ‘(if) 3s be’ vs. *bi-b-ēt* ‘(if) 3s become’; *pyāw-a čāk b-a* ‘be a good fellow!’

The copula is negated by *nī-* with the post-vowel endings: *nī-y-a* ‘is not’. *gal-ī kas ha-bū balām ēstā hīč kas nī-y-a* ‘there were a lot of people, but now there is no one’.

The personal enclitics are inserted between these prefixes and the verb stem to denote the direct object: *a-y-bīn-im* ‘I see him (-i)’; *bi-yān-xō* ‘eat them!’; *nā-yān-nās-īt* ‘you don’t know them’.

3.2.5 System of aspect, mood, and tense

Kurdish has two basic *aspects*, imperfective and perfective. Indicative verbs are perfective in aspect unless marked by the imperfective aspect prefix *a-*. *Perfective* forms depict a single completed event or series of events in a narrative, whereas *imperfective* verbs depict anything else. Specifically, the imperfective aspect encompasses progressive, habitual, predictive (“future”), generic and, in a narrational context, narrative Aktionsarten; these sub-aspects are generally distinguished by accompanying adverbials such as *hamū řōž-ēk* ‘every day’, *wā* (unstressed) ‘right now’, *sibay-nē* ‘tomorrow’, etc.

There are four moods, the *indicative* and non-indicative *imperative*, *subjunctive*, and *counterfactual*. They show the attitude of the speaker toward the truth-value of the utterance: The *indicative* denotes propositions uttered as true or factual; the *subjunctive* denotes potentiality; the *counterfactual* denotes abstractions that are hypothetical or unrealized (“irrealis”); and the *imperative* issues a direct command. The indicative is the default mood, the subjunctive and the imperative being triggered by specified preceding expressions or contexts, such as indefinite antecedents which require a subjunctive verb.

There are two basic *tenses*, present and past, the latter with four forms, preterite, imperfect, present perfect, and past perfect.

3.2.6 Transitivity and ergativity

Transitivity is marked distinctly in past forms in a tense-split ergative pattern. With intransitive verbs, the subject is marked by the personal endings in all tenses and moods. With transitive verbs, the subject/agent is marked by the personal endings in the present, but by the personal enclitics in the past. These suffixes are mobile, attaching to the leftmost component of the clause after the subject position: the direct object/patient noun or noun phrase, the nominal parts of the verb phrase and in the absence of any other component to the verb form itself (see also section 4.4.3 Subject and object marking in Syntax).

This applies also to compound verbs with transitive verbs, such as *kird-in* ‘to do’, even though semantically they may be intransitive: *imřō bayānī zū pyāsa [-m kird]* ‘I went for a walk early this morning’ (‘did a walk’); *ēmā xōš [-mān řā b-wārd]* *tā ēwār-a* ‘we had a good time until evening’ (‘spent [time]’).

3.2.7 Individual verb forms

Altogether there are five indicative morphological tenses: (1) the present, which depicts activities (progressive, habitual, predictive [“future”], generic) or states in present time, that is, as of the moment of speaking; (2) the imperfect, which depicts the same in past time prior to the moment of speaking; (3) the preterite, which denotes a completed event

or series of events in a narrative; (4) the present perfect, which denotes a completed event with present time relevance; and (5) the past perfect, which recounts an event completed in the past prior to the completion of a subsequent event also in the past.

Regarding subject and object marking, with intransitive verbs the subject is marked by the personal endings in all tenses and moods. With transitive verbs, while the subject/agent is marked by the personal endings in the present tense, it is marked by the personal enclitics in the four past forms. These suffixes are mobile, attaching to the leftmost component of the clause after the subject position: (1) the direct object/patient noun or noun phrase, (2) the nominal parts of the verb phrase, and (3) in the absence of any other component to the verb form itself (see also section 4.4.3 Subject and object marking under Syntax).

3.2.7.1 Present tense

Transitive and intransitive verbs take the same subject markers in the present tense:

TABLE 10.11: PRESENT-FUTURE

	'I come', etc.	'I eat', etc.
1s	<i>yē-m</i>	<i>a-xō-m</i>
2s	<i>yē-yt</i>	<i>a-xō-yt</i>
3s	<i>yē-t</i>	<i>a-xwā-t</i>
1p	<i>yē-yn</i>	<i>a-xō-yn</i>
2p	<i>yē-n</i>	<i>a-xō-n</i>
3p	<i>yē-n</i>	<i>a-xō-n</i>

3.2.7.2 Preterite

The preterite ("simple past") consists of the plain past stem and the intransitive subject/transitive agent markers. Stress is placed on the stem vowel:

TABLE 10.12: PRETERITE

	'I came', etc.	'I ate', etc.
1s	<i>hāt-im</i>	<i>xwārd-im</i>
2s	<i>hāt-ī(t)</i>	<i>xwārd-it</i>
3s	<i>hāt</i>	<i>xwārd-ī</i>
1p	<i>hāt-īn</i>	<i>xwārd-mān</i>
2p	<i>hāt-in</i>	<i>xwārd-tān</i>
3p	<i>hāt-in</i>	<i>xwārd-yān</i>

The preterite is a neutral perfective in aspect; it denotes a single event completed before the time of the utterance. It may denote more than one occurrence in a narrative but does not denote progressive or habitual action: *kay hāt-ī bō ēra?* 'when did you come here?'; *sē jār hāt-im bō ēra* 'I came here three times'.

3.2.7.2a “Aorist” function

The preterite is also used to denote the certainty of fulfillment of an action, as well as anticipated certainty: *daŋ-it čū bō amērikā* ‘your voice went (‘has already gone’) to America’ (said to someone whose voice had just been tape-recorded). That is, the “preterite” is not confined to past contexts; hence it may be called an “aorist” in the same sense that the term is applied to similar phenomena in Persian and other Iranian languages (cf. Windfuhr 1986).

3.2.7.3 Imperfect

The imperfect denotes past progressive and habitual action (“was doing, used to do, would do”), and past future (“he said *he would go* and tell her the next day”):

a-hāt ‘3s was coming’ *a-y-xwārd* ‘3s was eating’
a-hāt-in ‘they were coming’ *a-yān-xwārd* ‘they were eating’.

The distinction between the preterite and the imperfect is best shown with inherently past forms of the stative verbs such as ‘I knew’. Here, the preterite denotes a change of state or entering upon a state, while the imperfect denotes a continuous state or repeated action; thus preterite *zānī-m* ‘I knew’ in the sense ‘found out, learned, recognized then and there’ vs. imperfect *a-m-zānī* ‘I knew’ (all along).

3.2.7.4 Present perfect

The present perfect is based on the perfect participle in *-ū/-w*, the person marker, and with transitives the copula *-a*:

hāt-ū-w-a ‘he has come’ *xwārd-ū-y-a* ‘he has eaten’
hāt-ū-n ‘they have come’ *xwārd-ū-yān-a* ‘they have eaten’.

The present perfect denotes an event completed prior to the moment of speaking but with present time relevance. It can often be translated into English as a present state or condition resulting from a completed event; it thus contrasts with the present in that it denotes a completed event in present time whereas the present denotes an action which is repeated or in the process of happening:

hāt-ū-m ‘I have come, I’m here now’;
hāt-im ‘I came’, ‘here I am’;
yē-m ‘I come, I’m coming, I’ll come’.

wastā-w-im ‘I have (now) stood up, I am standing’;
a-wastā-im ‘I was getting up’ or ‘I always stood (there)’;
a-wast-im ‘I am getting up/will stand up’ or ‘I always stand here’.

3.2.7.5 Past perfect

The past perfect is formed of the reduced perfect participle in *-ī* + the past of ‘be’, and the appropriate person markers, which attach to the entire complex verb form:

hāt-i bū ‘3s had come’ *xwārd-i bū-y* ‘3s had eaten’
hāt-i bū-n ‘they had come’ *xwārd-i bū-yān* ‘they had eaten’.

The past perfect denotes an event completed before the completion of a subsequent event in past time, or simply a remote event.

3.2.7.6 Imperative

The imperative is based on the present stem, with the elidable prefix *bi-* and the endings 2s *-ø*, *-a* after consonants, and 2p *-in*. Verbal pronoun objects are inserted before the stem: *bi-řō* ‘go!’, *bi-yān-řrōš-in* ‘sell them!’, *wař-* imperative stem of ‘come’: *wař-in-a ēra!* ‘come-2p here!’, *wař-a*, *bā bi-č-īn bō ēra* ‘come, let’s go over here’; *da . . . (-day)* asseverative, invoking attention: *da bi-řō*, *da bi-řō day!* ‘go ahead!, go on!’; *da wař-a* ‘come here!’.

3.2.7.7–8 Present and perfect subjunctives

The present subjunctive is based on the present stem, and the perfect subjunctive on the past stem + *-i* followed by the subjunctive of the copula. They express imperfective and perfective potential action or state, respectively:

- (1) Present, hortative, *bāš-a*, *čī bi-ka-m?* ‘OK, what shall I do?’; *bāš-a*, *bā bič-īn* ‘good, let’s go’; in dependent clause, *bi-č-ē(t)* ‘(that) 3s go’; *bi-xw-ūt* ‘(that) 3s eat’;
- (2) Perfect, *hāt-i b-ē(t)* ‘3s may/might have come’; *xwārd-i b-ē(t)* ‘3s may/might have eaten’; *řaŋ ha-y-a halparkē-š [krā b-ēt]* ‘perhaps there was folk-dancing, too’ (lit. ‘may have been done’).

3.2.7.9–10 Imperfective and perfective counterfactual

There are two counterfactual forms marked by the modal prefix *bi-* and the enclitic *-ā* + 3s copula *-a*, attached to the past and the reduced perfect stems. They are restricted in use, occurring mostly in conditional clauses, and express imperfective and perfective hypothetical or irreal action or state (see section 5.5.2 Conditional clauses):

- (1) *bi-hāt-ā-y-a* ‘had 3s come’ *bi-xwārd-ā-y-a* ‘had 3s eaten’
- (2) *bi-hāt-i bw-ā-y-a* ‘3s would have come’ *bi-xwārd-i bw-ā-y-a* ‘3s would have eaten’.

3.2.7.11 Overview of system of aspect, tense, and mood

The basic pattern of the system of tense, aspect, and mood may be shown as follows, using traditional terms for the individual paradigmatic forms in the 1s except for the 2s imperative (following suggestions by Windfuhr):

TABLE 10.13: SYSTEM OF ASPECT, MOOD, AND TENSE

PR = present stem; PT = past stem, intransitive			
Imperfective	Indicative	Non-Indicative	
Present	<i>a-PR-im</i>	<i>bi-PR-a</i>	Imperative
Imperfect	<i>a-PT-im</i>	<i>bi-PR-im</i>	Present Subjunctive
		<i>bi-PT-im-ā-y-a</i>	Past Conditional
Perfective			
Aorist “Preterite”	<i>PT-im</i>		
Resultative			
Present perfect	<i>PT-ū-m</i>	<i>PT-i b-im</i>	Perfect Subjunctive
Past perfect	<i>PT-i bū-m-a</i>	<i>bi-PT-i b-ū-m-ā-y-a</i>	Perfect Conditional

4 SYNTAX AND USE OF FORMS

4.1 Coordinating conjunctions and clitics

The two basic coordinating conjunctions are *wa* ~ *-ū/-w* ‘and’, which usually cliticizes, and *yān* ‘or’. Paired coordinators include *wa . . . wa* ‘both . . . and’, *yān . . . (wa) yān* ‘either . . . or’, *na . . . (-ū, -w) na* ‘neither . . . nor’, and others.

In addition, there is the enclitic *-īš* ‘also, too’ which attaches to the first component of the clause, even inside the first main component before other clitics:

mīn-[īš] yē-m ‘I’m coming, too’; *xō-[š]-ī wā wīt* ‘he himself (*xō*) also said (*-ī wīt*) so (*wā*)’; *lē-[š]-it dā-m* ‘you (*-it*) also hit me’ (*lē dā-n* ‘hit’); combined in: *nā-xō-m-ava-w nā-[š]-kēš-im* ‘I don’t drink nor do I smoke’ (*-w < -ū* ‘and’).

4.2 Noun phrase structure

4.2.1 Post-head modifiers

Post-head nominal modifiers are linked by either of two liaison morphemes depending on whether the noun is definite or indefinite, where definite is defined as having the definite suffix *-aka* ‘the’ or a demonstrative envelope, such as *am . . . -a* ‘this’, whereas the indefinite noun has neither of these.

4.2.2 Dependent nominals

Dependent nominals, including adjectives, nouns, and nominal phrases, follow the head noun, and are linked by either of two liaison morphemes, *-ī* or *-a*.

4.2.2.1 Definite nominal phrases

The definite markers *-aka* ‘the’ and the deictic envelope *am/aw . . . -a* ‘this/that’ require the liaison morpheme *-a-*. In all other cases, including non-specific *-ēk* ‘a’, *-ī-* occurs: *kič-ēk-ī-jwān* ‘a pretty girl’ and *kič-a-jwān-aka* ‘the pretty girl’.

Note that *nouns of inalienable possession* (some kinship and body-part names and the noun *nāw* ‘name’) are implicitly definite and do not receive the definite suffix: *dāyk-im* ‘my mother’; *nāw-im* ‘my name’.

4.2.2.2 Indefinite nominal phrases

The simple nominal phrase consists minimally of an unmarked indefinite noun or other nominal head with optional pre- or, most commonly, post-head modifiers. Numbers and quantifiers precede the head and take phasal stress: *dū pyāw* ‘two men’, *galē jār* ‘many times, *čānd šāl?* ‘how many years?’

4.2.3 *Izāfa* construction

The *izāfa* construction with *-ī* is the unmarked type of subordination. The dependent nominal may be qualifying or limiting (with adjectives, nouns, particles as modifiers), possessive (with nouns, pronouns, interrogatives), appositive (nouns), naming (proper nouns), or infinitival phrases.

While the *izāfa* treats the head and the dependent nominal as separate semantic units, the construction with *-a* essentially treats the two components as forming a closer semantic and syntactic relationship. In schematic form the two types are as follows:

Type 1	Type 2
[N]-ī X	[N-a X] (compounds)
[N-ēk]-ī X	
[N-ak <u>a</u>]-ī X	[N-a X]-ak <u>a</u>
[amlaw N-a]-ī X	amlaw [N-a X]- <u>a</u> .

Type 1

[N]-ī X: *hēlka-ī tāza* ‘fresh eggs’; *pyāw-ī zōr* ‘many men’; *dū pyāw-ī čāk* ‘two good men’;
jāsūs-ī šad šāl ‘a spy of a hundred years’; *šaqām-ī wā* ‘streets like this’;
 [N-ēk]-ī X: [*pyāw-ēk*]-ī čāk ‘a good man’;
 [(aw) N-aka]-ī X: [*aw šaw-aka*]-ī *sarjō* ‘that night in Serjo’;
 [amlaw N-a]-ī X: [*am řš-a*]-ī *tō* ‘this/that deed of yours’.

Type 2

[N-a X], compounds: *kilk-a miška* ‘rat-tail’;
 [N-a X]-aka: [*kič-a jwān*]-aka ‘the pretty girl’; [*pyāw-a čāk*]-aka ‘the good man’;
 aw [N-a X]-a: aw [*pyāw-a-čāk*]-ak-*ān*]-a ‘those good men’.

4.2.4 Specific dependents

kitēb-ī kē? ‘whose book?’; *nāw-ī min* ‘my name’;
ganj-ī am-šār-a ‘the young (people) of this city’;
sa āt dū-ī pāš nīwar *ō* ‘two o’clock in the afternoon’;
nāw-ī mudīr-ī [ma’ūrif] ‘the name of the director of education’;
bās-ī mēžū-ī [kurdustān] ‘discussion of the history of Kurdistan’.

4.2.5 Naming, apposition

šār-ī silāmānī ‘Sulaimania City, the city of Sulaimania’; *māḡ-ī galāwēž* ‘the month of Gelawezh’ (July–August); *Aḡmad-ī birā-m* ‘my brother Ahmad’.

4.2.6 Infinitival head or dependent

kišt-in-ī [kāwā] ‘the killing of Kawa, Kawa’s murder’; *xarīk-ī [dars-xwēnd-in]* ‘busy at studying, about to study’.

4.3 Clause structure

4.3.1 Word order

In a declarative sentence the basic sequence of the clause is SUBJECT – OBJECT – VERB. Modifying adverbial phrases are inserted variously: clause-initially, after the subject, and before or after the verb. The clause as a whole may be introduced by one or more

introductory expressions, such as interjections and coordinating connectors. The minimal clause consists of an inflected verb. The full sequence is:

(I) (C) (S) (M1) (DO) (M2) (P) V (M3)

General rules include the following:

- (1) Introductory expressions (I) are typically interjections, conventional expressions of emotions:

ay ‘well, now; as for’ (signaling a change of topic in conversation); *ay tō cōn-ī?* ‘and you, how are you?’; – ‘O’; *ay adīb-ak-ān-ī kurdustān!* ‘O Writers of Kurdistan!’; *hā!* and *xō!* ‘aha!’; *xō* ‘I say!’ (accosting); *tūxwā – txwā* ‘indeed!’ (incredulity, < **tō ū xudā* ‘thou and God’); *marḥaba* ‘hello’; *āfērīm ~ āfērīn* ‘bravo! well done!’; *nōš* ‘cheers!’ (a toast); *amān* ‘alas; woe’, *ōxay* ‘ah’ (delight); *balē, ā, bā, ay* (informal); *arē, arē-walla* (polite) ‘yes’; *na, na-xēr* (polite) ‘no’.

- (2) Connectors (C) include conjunctions such as *çūjka* ‘(that is) because’, *balām* ‘but’ (see section 5 Complex Sentences).
 (3) Direct object phrases (DO) as a rule follow subjects (S), but are not differentially marked.
 (4) Modifying adverbial phrases (M), when they co-occur in the same position, usually follow the sequence: Time – Location – Manner/Instrument:

dwēnē la māl-awa ba dast mār-ēk-ī kušt
 yesterday at home by hand snake-a-he killed
 ‘he killed a snake by hand at home yesterday’.

- (5) In a sequence of two time adverbials the more specific follows; thus: *bayān-ī zū* ‘early (zū) tomorrow’; *dwēnē šaw* ‘yesterday evening’.

4.3.2 Verbal modifiers

Modifiers of the verb are typically indirect objects or expressions of manner:

kitēb-aka [ba dārā] a-frōš-im ‘I’ll sell the book [to Dara]’;
tō kurdī [bāš] a-zān-ī ‘you know Kurdish [well]’;
rašōl [wā] dars a-xwēn-ē ‘[that’s how] Rashol studies (reads lesson)’.

Modifiers are often placed postverbally. This includes typically allative expressions:

ēstā a-ē-īn [bō māl-awa] ‘we’re going to go [home] now’;
çapōk a-kēš-in [ba sar-ī] ‘they slap him [on his head] (in disapproval)’.

Other parts of speech may occupy this position, such as subject clauses and post-copula locative expressions. Otherwise, the postverbal position marks special focus:

bāš bū hāt-ī [la-gal-m-ā] ‘it’s good that you came [with me]’ (lit. ‘it was good’);
am-a tāza-tir-īn šaqām-a [l-am-šār-a] ‘this is the newest street [in this city]’;
a-lē-m [bō tō] . . . ‘[to you] I say’.

4.3.3 Questions

Questions do not change word order. The question words appear where the parts of speech they represent are normally located: *im'ō sar-ī kwē a-da-yn?* 'what place are we visiting today?', *ču-n bo kwē?* 'where did they go?'.
 Yes-no questions, mostly in literary Kurdish, are introduced by *āyā* 'whether', i.e. 'is it the case that': *āyā witār-ēk-ī dirēž-ī xwēnd-awa?* 'did he deliver a long speech?'.
 Questions are marked by a terminal interrogative intonation contour, e.g. *tō kurdī a-zān-ī?* 'do you know Kurdish?'. Briefly, it is a contour that is falling but does not fall as far as the level of a terminal statement contour.

4.3.4 Topicalization

In topic – comment constructions, the topic is extraposed before its clause and a reference to it, typically a resumptive pronoun or verbal agreement, is contained in the clause; thus:

nā-zān-im [tō] īš[-it] čī-y-a
 I don't know you work-your what-is
 'I don't know what *your* work is'.

4.4 Syntax and use of nominal forms

4.4.1 Pronouns

The independent pronouns are mostly used for emphasis or contrastive focus. They occur as: (1) subjects; (2) objects of prepositions; and (3) possessors in *izāfa* noun phrases.

4.4.2 Personal enclitics

The enclitic pronouns serve as: (1) direct object in the present tense; (2) subject markers on transitive verbs in the past tenses and objects of past tense verbs; (3) objects of prepositions and of present tense verbs; (4) possessor in noun phrases; (5) experiencer or beneficiary after adjectives.

As a general rule, the direct object suffixes in the transitive present tense verb phrases (as opposed to single-word verbs), on the one hand, and those of the subject/agent suffixes in transitive past tenses, on the other hand, follow the rule of leftmost attachment. The hierarchy in both present and past is as follows: (1) the nominal parts of the verb phrase; (2) the preverbs; (3) the prefixes. In their absence they are attached to the verb form itself. All of these are superseded in the transitive past by the direct object/patient noun or noun phrase.

The past intransitive person markers (see Table 10.10) are suffixed to the past tense transitive verbs to indicate pronominal direct objects.

4.4.3 Subject and object marking

(1) Nominal components, present tense object:

- (a) Adjective: *kay pān[-yān] a-ka-n?* ‘when will they widen them?’ (*pān* ‘wide’);
 (b) Noun: *bō bās[-ī] a-ka-n?* ‘why are you-2p discussing it?’ (‘doing its discussion’);

tašakkur [-it] a-ka-m ‘I thank you very much’ (‘make gratitude of/to you’);

- (2) Preverb: *bā war[-yān] bi-gr-īn* ‘let’s (*bā*) take them’ (*war-* ‘up’);
 (3) Prefix: *bi[-yān]-xō* ‘eat them!’.

4.4.4 Past agent suffixes

Examples of leftmost attachment (3s patients = zero ending, DO = direct object/patient phrase; AG = agent suffix):

- (1) DO + [AG] + VB:

tanhā yak kitēb [-im] kiṛī ‘[I] bought only one book’;

čī[-tān] kird l-awē? ‘what did [you-2p] do there?’;

ustāz-ak-ān-yān aw-řōž-a bās-ī mēžū-ī kurdustān[-yān] na-kird.

‘on that day their professors ([they]) did not discuss the history of Kurdistan’, lit. ‘did not make the discussion of K.’.

- (2) DO + [AG] + Locative + VB:

bō čī am-wiša-y-a-[t] la-sar taxta-ka nūsī? ‘why did [you] write this word on the blackboard?’;

- (3) DO + [AG] + Nominal + VB:

māl-aka[-mān] čāk kird/la-kird/kird-ū-wa

‘[we] improved/were improving/have improved our house’.

- (4) DO + [AG] + Preverb + VB:

āw[-ī] tēk war dāla-dāldā-w-a ‘[he] stirred/was stirring/has stirred the water’.

4.4.5 Particles *čōn* and *wā*

Exceptionally, the suffixes may be attached to the pair of manner adverbs *čōn* ‘how?’, *wā*, ‘thus’: *čōn[-it] zān-ī?* ‘how did you come to know that?’; *wā[-m] zānī* ‘I came to know it thus’ = ‘as far as I know, I believe’; *wā[-yān₁] nāv-yān₂ nūsī* ‘they (-yān₁) wrote their (-yān₂, ‘the others’) names like this’.

4.4.6 Verb-only phrases

There are specific rules for sequencing past agent suffixes and patient endings after simple verb forms. The basic rule requires that the agent suffixes be attached directly to the verb form, followed by the personal ending marking the patient: VB + agent suffix + patient ending:

TABLE 10.14: PRETERITE AGENT AND PATIENT

Agent/Enclitic	Patient/Ending
1s Stem <i>-im</i>	<i>-im</i>
2s <i>-it</i>	<i>-it</i>
3s <i>-ī</i>	<i>-ø</i>
1p <i>-mān</i>	<i>-īn</i>
2p <i>-tān</i>	<i>-in</i>
3p <i>-yān</i>	<i>-in</i>

Examples:

bīnī-m-īt 'I saw you-2s'; *bīnī-t-im* 'you-2s saw me';
bīnī-mān-in 'we saw you-2p'; *bīnī-tān-īn* 'you-2p saw us';
bīnī-tān-in 'you-2p saw them'; *bīnī-yān-in* 'they saw you-2p'.

However, there are two exceptions to the basic rule: (1) the 3s agent suffix *-ī* is always second; (2) the 1s *-im*, whether agent suffix or patient ending, precedes any *plural* subject or object:

3s agent second	1s before plural patient and agent
<i>bīnī-m-ī</i> 'he saw me';	<i>bīnī-m-in</i> 'I saw you-2p';
<i>bīnī-t-ī</i> 'he saw you-2s (-ī(t))';	<i>bīnī-m-in</i> 'I saw them';
<i>bīnī-n-ī</i> 'he saw us/you-2p/them';	<i>bīnī-m-tān</i> 'you-2p saw me';
(<i>bīnī-n</i> < <i>bīnī-īn</i> , <i>-in</i> , <i>-in</i>);	<i>bīnī-m-yān</i> 'they saw me'.

The rule of 1s + plural agents may not apply to all varieties of Kurdish, as other varieties have different rules.

In all other combinations, the agent comes before patient.

4.4.7 Preposing of adpositional suffixes

When an adpositional phrase follows a direct object, the personal enclitic precedes its adposition and is directly attached to the object, and the adposition assumes its absolute form; thus, *bō min* 'to, for me' > *-im bō*, *ba min* 'to me' > *-im pē*, *la tō* 'of, from you' > *-it lē*, e.g.:

čā[-m bō] b-ēn-a! 'bring me some tea!';
ūd[-it bō] lē a-da-m 'I'll play the lute for you';
a-m-awē čay su'āt-ēk[-it lē] bi-ka-m 'I want to ask you some questions';
gwē [-y lē] a-gr-im 'I'll listen to him';
ērā [-m pē] xōš-a 'I like it here' ('here is nice to me').

In sequences of two adpositional phrases the attachment of the suffixes may result in complex intertwined patterns: with *gwē lē girt-in* 'listen ('take ear') to':

gwē [-t]-ī[bō] [lē] a-gr-im
 'I'll listen to him (-ī . . . lē) for you (-t . . . bō)'.
 However, in transitive past tenses, the agent suffixes precede the adpositional phrase, and block the inversion:

su' āl-ēk-im [lē-y] kird

'I (-im) asked him (-y) a question' (*su'āl la N kird-in* 'ask something of a person').

4.4.8 Past tense object and possessor raising

In the absence of direct object phrases, the beneficiary expressed by indirect objects or objects of adpositions may behave similarly to direct objects in past tenses. That is, as an alternative to the regular ergative construction, they may be syntactically "raised" and marked by the personal endings. Similarly, the possessor of a direct object/patient phrase may be "raised":

(1) Beneficiary:

salām-yān [lē] kird [-īn]

'they (-yān) greeted us (*lē -īn*)', lit. 'made greeting to us';

[pē]-m wit [-in]

'I (-im) said to you-2s' (*pē -in*).

(2) Possessor:

[bačkak-ān]-ī a-xwārd-im

'it (-ī) used to eat my (-im) children' (MacKenzie 1961: 115).

4.4.9 Three-place verb 'to give to'

This verb typically has double valence, the direct and the indirect object, which is expressed in Kurdish by the postposition *-a*, absolute form *-ē*. The latter follows the inflected verb forms (preceded by *-r-* after vocalic ending), but its object suffix is attached according to the leftmost rule (MK = MacKenzie 1961):

(1) Present:

dīnār-ēk[-it] a-da[-m]-ē

dinar-one you-2s IMPF-give-I-to

'I'll give you (-it -ē) a dinar';

a[-y]-d-[-āt]-ē

IMPF-him-he gives-to

'he gives it (-ø) to him (-ī -ē)';

(2) Preterite and perfect, indirect object raised:

kič-ēk[-yān] na-dā[-yn]-ē

girl-one-they-AG-not-gave-we are-to

'they did not give us (-īn -ē) a girl' MK 80;

hīč[-yān] na-dā-w[-im] -at-ē

any-they-AG-not given-PF-I am-PF-to

'they have given nothing to me'

MK 116.

In the absence of a preceding noun phrase, there may be a *triple* sequence, such that the indirect object is expressed by a second personal ending. The relative order is essentially that of transitive verbs:

(1) Present:

a[-t]-da[-m]-ē[-n]

IMP-you-2s-give-1s-to-they are

'I'll give them to you';

(2) Preterite:

dā[-m] [-īt] [-in]-ē

gave-1s Ag-you-2s are-they are-to

'I gave you to them' MK 116.

4.5 Semantics and use of verb forms

4.5.1 Possession

Possession is expressed by the 3s of the existential verb, *ha-y-a* 'there is, exists' and the personal enclitics in their function as indirect object/beneficiary. They are attached to the object possessed: *-m ha-y-a* 'to me is' = 'I have', exceptionally to *ha-* in the absence of an object: *da dīnār [-im ha-bū]*, *balām ēstā hīč pāra[-m nī-y-a]* 'I had ten dinars, but now I don't have any money at all'.

4.5.2 Aspectual constructions

4.5.2.1 Virtual accomplishment

Virtual accomplishment or completion of an action is expressed by the locative particle *wā* 'here, now' and the preterite ("aorist"): *wā hāt-im* 'I'm coming! I'm on my way!'; cf. *wā-n la ērān-dā* 'they are (-in) now (*wā*) in Iran (*la . . . dā*)'.

4.5.2.2 Progressive action

Progressive action may be highlighted by the locative infinitival construction: *la* + [infinitive] + *-(d)ā* + copula:

la [rōyšt-in]-ā-yn 'we are going' vs. simple present:
a-rō-yn 'we go, are going'.

4.5.2.3 Continued intensive action

Continued intensive action may be expressed by an infinitival *izāfa* construction depending on semantically fitting nominals such as *xarīk* 'busy' + infinitive:

māḡ-a *xarīk-ī [īš kird-in]-in*
 (it) month is (that) busy-IZ [work-doing] they are.
 'for months they (-in) have been hard at work'

4.5.3 Modal constructions

The modal verbs in Sōrānī expressing ability, desire, etc., are realized mostly as canonical verbs triggering a subjunctive verb; necessity is typically expressed by verb or adjective plus subjunctive.

4.5.3.1 Ability

Ability is expressed by the transitive verb *twān-/twānī-* 'be able, can'; it may take a direct object:

nā-twān-im bi-rō-m 'I can't go';
a-t-twān-im 'I can manage you (-t)!'.

4.5.3.2 Desire, necessity, and assumption

Desire, necessity, and assumption are expressed by the modals (1) *awē(t)/wist-* ‘want’, (2) *a-b-ē* ‘(it) must (be)’, and (3) *pē wist-a* ‘is necessary’. These are impersonal constructions; with *awē-/wist* ‘want’ the experiencer must be expressed by personal enclitic (‘to me is desire’):

(1) ‘want’:

<i>ama[-m] a-wē</i>		‘I (-m) want this one’;
<i>hīč-yān [-mān] nā-wē</i>		‘we (-mān) don’t want any of them (-yān)’;
<i>a[-m]-awē</i>	<i>bi-zān-ī</i>	‘I want you to know’;
<i>a[-mān]-awē</i>	<i>bi-řō-yn</i>	‘we want to go’;
<i>a[-mān]-a-wist</i>	<i>bi-řō-yn</i>	‘we wanted to go’;

(2) ‘must’:

<i>a-b-ē</i>	<i>bi-řō-yn</i>	‘we must go, we have to go’;
<i>nā-b-ē</i>	<i>bi-řō-yn</i>	‘we mustn’t go’;
assumption,		
<i>a-b-ē</i>	<i>řōyšt-i b-ēt</i>	‘he must have gone’ (perfect subjunctive);

(3) ‘necessary’:

<i>pē wist-a</i>	<i>bi-řō-yn</i>	‘it is necessary that we go, we have to go’;
<i>pē wist-a</i>	<i>zū b-ī-kay</i>	‘you must do it (-ī) right away’ (‘it is necessary that’).

5 COMPLEX SENTENCES

5.1 Coordinate clauses

The canonical simple sentence consists of a single independent clause; it may also consist of an interjection or, through ellipsis, a single word or phrase.

Statements are signaled by a sentence-final intonation contour. In written Kurdish they are signaled by Western punctuation marks and practices.

Compound sentences may be coordinated or subordinated, with or without overt connectors, and may be quite complex.

The connectors may be simple conjunctions or complex phrases.

Coordination conjunctions include *wa* ~ *-ū/w* ‘and’, *yā* ‘or’; *balām* ‘but’, *bal-kū* ‘but’: *wa la-pāš aw-a, čī-t kird?* ‘and after that, what did you do?’ (see also the coordinating conjunctions and clitics discussed in section 4.1).

5.2 Subordinate clauses

Matrix verbs and other expressions that govern assertions are followed by verbs in the indicative, whereas verbs and expressions that require action or a change of state by the embedded subject or express doubt or other such emotions are followed by embedded subjunctive verbs.

Subordinating conjunctions include: *ka* ‘that; when’; *agar* ‘if’; *agī na* ‘otherwise’; *magar* ‘unless, if not’; *ha-tā* ~ *tā* ‘until’, ‘in order that’: temporal, purpose, and conditional clauses; *čūŋka* ‘because’; *wak, wakū* ‘as’.

In addition, there are complex conjunctive phrases with prepositions of varying syntactic complexity, both coordinating and subordinating. They are marked by adpositions, often based on the neutral demonstrative *awa* ‘that’ which may be followed by the relative connector *-ī*, with or without the conjunction *ka* ‘that’. These phrases mostly originate in adverbial constructions, and are particularly frequent in literary registers.

Clauses introduced by conjunctive phrases such as ‘before, without, instead of’ also trigger the subjunctive:

har gwē a-gr-in bē aw-a-ī hawbašī bi-ka-n la qiša kird-in-ā
 ‘they only listen **without** taking part in the discussion’
 (lit., ‘without (that) they may take part in the discussion’).

5.3 Relative clauses

As a rule, relative clauses are introduced by: (1) the conjunction *ka* ‘who’ if definite and specific; (2) *-ī* if definite, restrictive; (3) no overt connector, or *-ēk*.

Clause-internally, the head (antecedent) is resumed by the personal ending if subject, but by the personal enclitic in all other functions.

The degree of definiteness also determines the mood of the verb in the relative clause. It is: (1) indicative if the antecedent is definite, whether specific or not; (2) subjunctive if the antecedent is indefinite:

5.3.1 Subject

- (1) Definite antecedent:

aw pyāw-a ka hāt la-gal-t-ā
 ‘the (**that**) man **who** came with you’;

am-a sē-y-am koŋgra-y-a ka l-am bābat-a-wa a-gīr-ē
this is the third congress that on this topic will be held’;

aw-a-ī ka hāt la-gal-t-ā
 ‘**the one that** came with you’.

- (2) Indefinite antecedent:

awa-ī bi-ē-ēt-a šār-ī kwēr-ān, a-bē dast ba čāw-awa bi-gr-ēt
 ‘**he who** goes to the city of the blind must cover his eyes with his hand’
 (“when in Rome do as the Romans do”, proverb);

kas ha-y-a l-awē kurdī bi-zān-ē?
 ‘is there **anyone** there (that) knows Kurdish?’.

5.3.2 Pronoun other than subject

- (1) Specific:

aw pyāw-a ka l-awē a-y-bīnī
 ‘**the man that** you see him there’;

(2) Indefinite, *-ēk*:

duktōr-ēk a-nās-im, a-č-īn-a lā-y a-mān-awē kas-ēk bi-zān-ē kurdī
 'I know a doctor (that) we can go to him'; 'we want **someone** (that) he
 knows Kurdish';

hīč tīp-ēk nī-y-a, bi-twān-ē ba sar tīp-aka-mān sar kaw-ē
 'there is no (any one) team (that) can beat (fall ahead of) our team'.

5.4 Subject and object clauses

Subject and object clauses as a rule follow the verb phrase, and may be introduced by the conjunction *ka*.

5.4.1 *Subject clauses*

bāš-bū hāt-ī la-gal-m-ā
 'it's good (that) you came with me'.

5.4.2 *Object clauses*(1) No *ka*:

kē a-zān-ē ma nā-y am-wiša-y-a čī-y-a?
 'who knows **what** the meaning of this word is?';

(2) with *ka* 'that':

bīst-im ka la saratā-ī mān-ēk-ī tīr sarōk jamhūriyat-ī amarikā a-č-ē bō faransa
 'I heard **that** at the beginning of next month the American President will go to
 France';

(3) Sequential clauses:

wā bi-zān-im hamū kas-ēk a-zān-ē
 'I believe everybody knows

ka xanjar-ī-kurdustān la hamū xanjar-ēk bāš-tīr-a
that the daggers of Kurdistan are better than any other daggers';

(4) Asyndetic:

bīnī-m-īt a-hāt-ī 'I saw you coming' ('I saw you, you were coming').

5.4.3 *Quotations*

Quotations are in direct speech:

ba tō a-lē-m nā-zān-im 'to you I say, "I don't know"';

pē-y wīt-im nā-m-a-wē bi-č-īm la-gal-tān
 'he told me he didn't want to go with us' (lit. 'I don't want to go with **you-2p**').

5.4.4 Predicative clauses

awa ‘that’ occupies the predicate position; the appositive predicate clause follows the copula:

čāk-tir-īn šit l-am-kitēb-a awa-y-a naxšay-ēk-ī kurdustān-ī tyā-y-a
 ‘the best thing about this book **is this (that)** there is a map of Kurdistan in it’;

yak-ēk l-avāna awa bū
 ‘one of them [news items] **was this**
māmwastā-yān-ī kurd . . . la baḡā bās-ī zimān-ī kurdī a-ka-n
 (that) the Kurdish teachers will discuss the Kurdish language in Baghdad’.

5.4.5 Appositive complement clauses

In constructions like *la birwā-dā bū-n* ‘be of opinion, believe’, the appositive clauses are in apposition with *aw-birwā-y-a* ‘that opinion’:

mīn l-aw-birwā-y-a-dā-m,
 ‘I am of that opinion
kobūn-awa-ī sarōkwazīr-ān-a wilāt-a-gawra-k-ān šitēk-ī zōr ba kalk-a
 ‘(that) meetings of the prime ministers of the major powers are a very beneficial thing’.

Note that in this and similar constructions the appositive clause occurs without the conjunction *ka* ‘that’.

5.5 Adverbial clauses

5.5.1 Temporal clauses

Temporal clauses are introduced by conjunction *ka* ‘when’, and as a rule precede the main clause. A particular use of tense is that of the preterite verb form in its function of indicating completed action or condition, in either past or present time (“aorist”). Anticipated action before another, *pēš awa-ī* ‘before’, requires the subjunctive:

ka bū ba sa’āt dwānza, nān-mān xwārd
 ‘when it was (became) twelve o’clock we ate’;

ka gayšt-in-a awē, kāḡaz-ēk-im bō bi-nūs-in
 ‘when you-2p get (‘got’) there send me a letter’;

pēš awa-ī b-ē-yn bō sinif ‘before we come to class’.

5.5.2 Conditional clauses

5.5.2.1 Real, possible, and unreal conditions

Conditional clauses precede the main clause, and are normally, but not necessarily, introduced by *agar* ‘if’.

Real conditions, or conditions presumed as fact, are in the indicative; potential or probable conditions are in the subjunctive; and counterfactual or irreal conditions are in the “conditional” mood.

There are various patterns of the tense-mood sequences in protasis and apodosis, which will be indicated in the examples below.

(1) Factual conditions:

(a) present + present:

agar a-zān-ī imrō-š har bē tāqat-ī, duktōr-ēk a-nās-im a-č-īn-a lā-y
 ‘if **you think** you are still indisposed, **I know** a doctor to whom we will go’;

(b) present + imperative:

hāz a-ka-yt, tō-š waš-a pyāsa bi-ka-yn
 ‘(if) **you feel** like (it), you too come for a stroll’ (‘that we stroll’).

(2) Possible conditions:

Subjunctive + present,

agar dā nāš-im, diraj a-b-ē
 ‘if **I sit down** **I will be** late’;

wā bi-zān-im, min-iš kabāb bi-xō-m bāš-a

‘thus I should think, “(if) I too should eat some kabob, it’s fine”’, i.e. ‘I think I should have some kabob too’.

(3) Counterfactual conditions:

The basic pattern in counterfactual conditional clauses is that the condition in the protasis is marked by the conditional forms in *-ā-y-a*, followed in the apodosis by either the imperfect or the past perfect in their counterfactual function:

	Protasis	Apodosis
Imperfect:	<i>bi-kird- . . . ā-y-a</i>	<i>ā- . . . kird</i>
Past perfect:	<i>bi-kird-i bw- . . . ā-y-a</i>	<i>kird-i bū . . .</i>

Of these, the imperfect form may refer to both present and past contexts, but the past perfect form only to anterior contexts. Occasionally, the imperfect and past perfect may occur in the protasis as well:

imperfect + imperfect

agar xanjar-aka-t-mān na-dā-yt-ā-y-awa, čī-t a-kird?
 if dagger- you-2s (-īt)-we-AG had not given-it-COND what-you-AG would do-IMPF
 ‘if we hadn’t given you back your dagger, what would you have done?’.

Here the verb is *dā-n-awa* ‘give back’, whose positive conditional form would be *bi-dā[-ā-y-a]-awa*.

5.5.2.2 *agīna* ‘if not, otherwise’

This explicit apodosis conjunction is regularly followed by the imperfect in its modal function:

na-m-zānī na-xōš bû-y, agīna na-a-čû-m
‘I didn’t know you were sick, otherwise I wouldn’t have gone’:

agīna dwēne har sar-im a-dā-y la ütēl-aka.
‘Otherwise I **would** certainly **have** visited you at the hotel yesterday’
(*sar dā-n la* ‘to drop in on, visit’).

5.5.2.3 Wishes

Wishes are introduced by the particle *xōzga* ‘would that’ (< **xwaz ka* ‘wish that’) and parallel conditional clauses. Fulfillable wishes are in the subjunctive, unfulfillable ones in the imperfect:

xōzga lē-m bi-tirs-in ‘I hope **they will (may)** be afraid of me’:

xōzga jār-ēk-ī tir-iš a-hāt-īt-awa nāw-mān
‘I wish **you-2s would come** again (and live) among us’.

5.5.3 Causal clauses

Causal clauses that imply a reason or condition precede the main clause, and are introduced by demonstrative relative phrases such as *la-bar awa-ī* ‘on that’ + *ka*. Those that are explanatory follow the main clause, usually introduced by *čūŋ-ka*:

la-bar awa-y-ka a-č-in bō bayā bi-xwēn-in,
‘**since** you are going to Baghdad to study (‘that you study’),

kāyaz-ēk a-nūs-im bō řašōl
I’ll write a letter to Rashol’;

ba řastī, ħaz nā-ka-m dā nīš-im-SUBJ,
‘to tell the ruth, I don’t want to sit down,
čūŋ-ka a-m-awē tōž-ēk pyāsa bi-ka-m-SUBJ
because I want to take a little stroll’.

5.5.4 Purpose clauses

Clauses expressing purpose generally follow the main clause, with the verb properly in the subjunctive. They may be introduced by the adverbial phrase *bō awa-ī* ‘in order that’, or *hatā, tā* ‘so that’, but often occur without overt conjunctive. The latter is also the typical pattern for potential action after verbs and phrases implying purpose:

5.5.4.1 Potential action

dā nīš-a, tōz-ēk qiša bi-ka-yn
‘sit down (*dā*), (so) we can talk a bit’;

kāyaz-ēk a-nūs-im bō bawk-im
 ‘I’ll write my father a letter,
bō awa-ī yārmātī-tān bi-d-ā ka gayšt-in-a awē
 so that he will help you when you get there’;

hatā sar bi-kaw-in ‘that you-2p fall ahead’, i.e.
 ‘in order for you to succeed’.

5.5.4.2 Non-factive phrases

zōr hūz a-ka-m šāl-ēk-ī tir bi-č-im bō amarīka bi-xwēn-im
 ‘I very much want to go to America next year to study’ (‘that I study’);

hūwadār-im la xōšī-w baxtyārī-dā b-in
 ‘I hope (that) you-2p will live (‘be’) in happiness and prosperity’;

biryār-mān dā bi-č-īn bō sayrān
 ‘we decided to go on a picnic’ (‘that we go’);

mīn āmāda-m bō-t yārmātī bi-ka-m
 ‘I am ready to help you’ (‘that I do help’);

řaŋ ha-y-a halparkē-š krā b-ēt
 ‘perhaps there was folk-dancing, too’ (lit. ‘may have been done’).

5.5.4.3 Suggestions

Independent subjunctive clauses express proposed actions and questions about actions:

bō kām lā bi-č-īn? ‘which way shall we go?’ (*lā* ‘side’).

Usually there is an introductory hortative *bā* ‘let’ (< **b-ā-d* ‘may it be’) for suggestions:

bā bi-č-īn! ‘let’s go!’; *bā b-ēt!* ‘have him come!’;
bā har qša na-k-am ‘I’d just better not say anything.’

6 LEXIS

6.1 Word formation

6.1.1 Derivation

6.1.1.1 Nouns

-čī, occupation: *bō yāčī* ‘bootblack’ (*bō yā* ‘shoe shine’);
 -ōka, diminutive: *minālōka* ‘(dear) little child’ (*mināl* ‘child’).

6.1.1.2 Nouns or adjectives

-ī, abstractions, quality or state, ‘-ness, -hood’: *dāykī* ‘motherhood’; *jwānī* ‘beauty’;
 -ā-y-at-ī, abstractions: *sarōkāyatī* ‘leadership’ (*sarōk* ‘leader’);
 -ī, gentilic, origin: *řāyī* ‘Christian’;

- gā, place: *kārgā* ‘workplace’ (*kār* ‘work’);
- xāna, building, room: *čāyxāna* ‘tea-house’ (*čāy* ‘tea’);
- ā, quality or condition: *sārdā* ‘coldness’.

6.1.1.3 Adjectives

- ā-w-ī, covered with: *qōrāwī* ‘muddy’;
- bāw, of the shade of: *sūrbāw* ‘reddish’.

6.1.1.4 Negation

na-, *nā-*, unstressed; *nā-* seems to be used more than *na-* for figurative meanings; with nouns, adjectives, verb stems:

- (1) *xōš* ‘good, fine’: *naxōš* ‘ill, sick’; *nāxōš* ‘unpleasant’; *fām* ‘comprehension’: *nafām* ‘innocent (child)’; *nāfām* ‘ignorant’;
- (2) others: *nākas* ‘baseborn’ (nobody); *nabū* ‘non-existent’; *namir* ‘undying, immortal’.

6.1.2 Compounding

6.1.2.1 With or without a composition vowel:

čwārpē ‘quadruped’ (‘four-foot’); *mārmāsī* ‘eel’ (snake-fish’); *gīyānlabar* ‘animal’ (‘soul-upon’), *čāwařē* ‘expecting, in wait’ (‘eye-to-road’); *qişaxōš* ‘witty’ (‘speech-pleasant’); *dūrūdrēž* ‘prolonged, in detail’ (‘far-and-long’).

6.1.2.2 Verbal component

noun + verb: *řasmgīr* ‘photographer’ (‘picture-taker’), *dilgīr* ‘sad’ (‘heart-taken’), *māřšxōr* ‘employee’ (‘salary-eater’, cf. Persian *-xor*); *řōžhālhāt* ‘sunrise; east’ (‘sun-up-coming’); adjective + verb: *dūrbīn* ‘telescope’ (‘far-seeing’).

6.1.2.3 Reduplication

Reduplicated nominals are most frequently nouns and adjectives, with either distributive meaning, like *damdam* ‘from time to time’, from *dam* ‘period of time, a while’, or intensified meaning, like *palapal* ‘great haste’ from *pala* ‘haste’. The latter example illustrates the elision of word-final *-a*.

6.2 Sociolinguistic aspects

6.2.1 Registers

Modern Standard Kurdish constitutes the literary register of Sulaimani Kurdish. It is characterized by primarily literary expressions, e.g. *tanānat* ‘then’, and by none of the consonantal morphophonemic changes discussed under Phonology. The imperfective aspect marker in MSK is *da-* rather than *a-*; it is the prefix found in most Sōrānī outside of Sulaimania. There is also a spoken Formal register, a deliberate variation in which, for example, the palatalization of *k* and *g* may be blocked.

6.2.2 Modes of address

Polite forms consist of second person plural verb forms or pronouns used in addressing an individual, and in the use of specialized honorific lexical items. Illustrations of the latter are the verb *farmū-n* ‘to order, command’ which in the imperative means “please (do the honor of) . . .” as in *farmū dā nîš-a* ‘please sit down’ and in the indicative means ‘speak, talk’ as in *wakū a-farmū-y* ‘as you say [polite]’.

The expression *ōyir* ‘blessed goings and comings’ occurs in the polite expression *ōyir-bē* ‘where are you going?’ and its response *ōyir-tān xēr-bē* ‘may your goings be blessed’ (*bē < b-ēt* ‘may it be’), and in the verbal phrase *ōyir kird-in* ‘to depart’. *kāk-a* ‘sir!’, vocative of *kāk* ‘older brother’, is commonly used as a term of respect for older men.

7 DIALECTOLOGY

The major work on comparative Kurdish dialectology remains MacKenzie (1961), together with his other comparative publications. More recently numerous scholars have contributed both comprehensive studies, particularly Fattah (2000) for Southern Kurdish, and typologically oriented studies. A notable recent contribution is the succinct overview of Kurmānjī by Haig and Paul (2001). The following highlights the major features in dialectal variation.

The major break, defined by convergent bundles of isoglossic lines, is that between Northern Kurdish and Sōrānī. Northern Kurdish itself shows a western and eastern division (not discussed here).

7.1 Phonology

The basic differences in the phonological systems between Northern and non-Northern Kurdish are shown in the following synoptic tables. In the vocalic system, typical for the latter is the retention of *ō*, and the phonemic status of *ɪ* in Sulaimania. A feature that is common to all dialects of Kurdish is that the most common allophone of *a* is [ə].

In the consonantal system, the most salient Northern feature is the phonemic opposition between aspirated and non-aspirated stops, undoubtedly due to contact with Armenian in Eastern Anatolia. In turn, non-Northern are the lenization of intervocalic *-d-* to a glide (Zagros-*d*), and the phonemic status of velar *ŋ* and of velarized *l*, as found in Sulaimania.

TABLE 10.15: COMBINED KURDISH VOWEL PHONEMES

[] = only Sul.; () = only North.; *a* mostly /ə/

	Front	Central	Back
High long	<i>ī</i> (<i>ū</i>)		<i>ū</i>
short	[<i>ɪ</i>]	<i>i</i>	<i>u</i>
Mid long	<i>ē</i>		[<i>ō</i>]
short		<i>a</i>	
Low		<i>ā</i>	

TABLE 10.16: COMBINED KURDISH CONSONANT PHONEMES

[] = only Sul., [ʔ] = Zagros-d; () = only North

	Labial	Dental	Phar.	Palatal	Velar	Uv.	Phar.	Glottal
Stops/Affricates								
aspirate	(p ^h)	(t ^h)			(k ^h)			
voiceless	p	t	(t)	ç	k	q		ʔ
voiced	b	d		j	g			
Fricatives								
voiceless	f	s	ʃ	ʃ̣	x		ħ	h
voiced	v	z	(z)	ž	ɣ		ʕ	
Glides	w	[ʔ]		y				
Nasals	m	n			[ŋ]			
Flap/Lateral								
plain		r l						
trill/velarized		ʀ [l]						

7.2 Nominal inflection

The typical system of nominal inflection in northern Kurdish has binary distinction in number, case, as well as gender, but confined to the singular. Number and gender are also marked in the Kurdish *izāfa* construction. The direct plural is typically zero vs. the oblique *-ān*. The following examples show the pattern in the northern variety of Cizre-Botan (Jizre-Bohtān) spoken south of Lake Van (adapted from Haig and Paul 2001).

TABLE 10.17: PERSONAL PRONOUNS IN CIZRE-BOTAN

	1s	2s	3s	1p	2p	3p
DIR	<i>ez</i>	<i>tu</i>	<i>ew</i>	<i>em</i>	<i>hūn</i>	<i>ew</i>
OBL	<i>min</i>	<i>te</i>	m. <i>wī</i> f. <i>wē</i>	<i>me</i>	<i>we</i>	<i>w-ān</i>

TABLE 10.18: NOMINAL INFLECTION IN CIZRE-BOTAN

	‘man’	plural	‘woman’	plural
<i>Definite</i>				
DIR	<i>mēr</i>	<i>mēr</i>	<i>jin</i>	<i>jin</i>
OBL	<i>mēr-ī</i>	<i>mēr-ān</i>	<i>jin-ē</i>	<i>jin-ān</i>
<i>Indef.</i>				
DIR	<i>mēr-ek</i>	<i>mēr-in</i>	<i>jin-ek</i>	<i>jin-in?</i>
OBL	<i>mēr-ek-ī</i>	<i>mēr-in-ān</i>	<i>jin-ek-e</i>	<i>jin-in-ān?</i>
<i>Izāfa</i>				
Def.	<i>mēr-ē</i>	<i>mēr-ēn</i>	<i>jin-a</i>	<i>jin-ēn</i>
Indef.	<i>mēr-ek-ī</i>	<i>mēr-in-e</i>	<i>jin-ek-e</i>	<i>jin-in-e?</i>

This table includes indefinite plural forms in *-in*, which are cited in certain grammatical descriptions, but may be doubtful (here marked by “?”; cf. Standard Persian indefinite plurals such as *mard-hā-ī* ‘certain men’ in specific contexts).

Sulaimani Kurdish and the Wārmāwa variety to its south represent a distinct island within the Kurdish expanse. They have lost nominal inflection (except for the reflexes in the vocative and a few locative/temporal forms mentioned above), and have generalized the oblique plural ending *-ān* as plural marker, and the indefinite masculine marker of the *izāfa*, *-ī*.

The northern dialects lack the definite marker *-aka*, while *-a* with preceding demonstratives is confined to the proximate in a subset of those dialects.

The retention of gender and case distinction in Northern Kurdish may well be due to the Zāzākī dialects where an even more complex system is found, particularly in the *izāfa* construction, both adjectival and nominal, and in combination (see also section 4.2.1 *Ezāfe* constructions in the Zāzākī Chapter 9 in this volume). The northern pattern is shown in the following examples (adapted from Schroeder 2002):

adjectival: *šav-ā čū-y-ī* = night-EZ-F gone’ = ‘last night’,
 indefinite: *gund-ak-ī xwašik* = village-EZ-INDEF-EZ-M nice = ‘a nice village’;
 nominal: *dast-ē kačik-ē* = hand-EZ-M girl-OBL-F = ‘the hand of the girl’, ‘the girl’s hand’.

When combined, the adjective qualifying the head noun is postponed after the inserted possessor, e.g.:

hawš-ā w-ān-ā marmarī
 courtyard-EZ-F they-OBL-3p-EZ of-marble = ‘their marble courtyard’.

Northern Kurdish also shares with Zāzākī the loss of the personal enclitics. These dialects therefore differ remarkably from the others where particularly in transitive past tenses one finds complex patterns of postverbal sequencing of personal agent and patient marking, as shown above for Sulaimanī.

7.3 Verb system

Northern Kurdish has a modal particle for future, *(d)ē*, followed by the present tense (cf. Zāzākī). It has a periphrastic passive construction, and lacks the morphologically derived passives. In turn, it has retained the ergative construction with past transitive verbs, where the agent is in the oblique, and the patient in the direct case and marked by the verbal ending: *min* (DIR) *tō* (DIR) *ʔafō kir-ī* = I thou forgiveness made-2s = ‘I have forgiven you’ (Akre variety). This contrasts with the agential construction, where the agent is marked by the personal enclitic, as described for Sulaimanī above.

8 SAMPLE SŌRĀNĪ TEXT

This excerpt is from a short story “Kwā kuř-aka-m?” by Mistifā Sāih Karīm about a man who is in the hospital after having saved a boy from a raging flood, thinking that he was saving his own son who has in fact perished in the flood. Here he is talking to his wife Galawezh (Abdulla and McCarus 1967: 49).

kwā kuř-aka-m?

where-is son-the-my
'Where is my son?'

.....

ma-gir-ī, galawēž, ma-gir-ī.
not-cry-2s Galawezh not-cry-2s
'Don't cry, Galawezh, don't cry.'

ta pē-m bi-lē sizā la kwē-y-a? la māl-awa-y-a.
you to-me IMPV-say Siza in where-is in house-is
'Now, tell me where Siza is.' 'He's at home.'

ay, bō na-t-hēnā la-gal xō-t?
hey why not-you-brought with self-your?
'Well, why didn't you bring him with you?'

řāstē-ka-y dīl-im na-a-hāt (< řāstī-aka-ī)
truth-the-its heart-to me not was-coming
'The truth of the matter is that I didn't have the heart'

pē-y bi-lē-m tō la naxōš-xāna-y. (-y < -ī(t))
to-him SUBJ-say-I you in hospital-are
to tell him that you are in the hospital.'

āxīr nā-bē, min a-m-awē b-ī-bīn-im. (3s nā-bē < na-a-b-ē(t))
after all! not-will-be I IMPF-me-want SUBJ-him-see-I
'Come, now, that won't do. I want to see him.'

la dwāyī-y-ā ka čāk a-b-īt-awa (a-b-īt < a-ba-ī(t))
in later-in when good IMPF-be-you-again
'Later, when you are well again,

d-ē-t-a dar-ē a-y-bīn-ī. (d- < dā, dial. var. of a-)
come-you-to outside-at IMPF-him-see-you
you will come out and see him.'

galawēž min mināl nī-m, a-zān-im
Galawezh, I child not-am I IMPF-know-I
'Galawezh, I'm not a child; I know

ka l-ēra ba-zīndūwī nā-y-ē-m-a dar-awa.
that from-here with-being alive not-come-I-to outside
that I will not get out of here alive.'

*a-mir-īm. tū-xwā agar wā bi-lē-y!**
IMPF-die-I by God if thus SUBJ-say-you
'I am dying.' 'Please, don't talk like that!'

da bāš-a am jār-a b-ī-hēn-a la-gal xō-t
EMPH good-is this-time IMPV-him-bring-! with self-you
'That's enough! Next time bring him with you.'

**tû-xwā*, *agar wā bi-lē-y!* implies: 'if you should talk like that, beware of God' (*tû-xwā* < *tō ū xudā* '(between) you and God').

ACKNOWLEDGEMENT

I would like to express my gratitude to Gernot Windfuhr for many suggestions, additions, and improvements.

BIBLIOGRAPHY AND REFERENCES

- Abdulla, J. J. and McCarus, E. (1967) *Kurdish Readers, III. Kurdish Short Stories*, Ann Arbor: University of Michigan Press.
- Ahmad, A-M. R. (1986) *The Phonemic System of Modern Standard Kurdish*, Ph.D. dissertation, University of Michigan, Ann Arbor.
- Amindarov, A. (1994) *Kurdish-English, English-Kurdish Dictionary*, New York: Hippocrene.
- Bakaeu, Ch. Kh. (1962) *Govar kurdov Turkmenii* (The dialect of the Kurds of Turkmenistan), Moskva: Izd-vo Akademii nauk SSSR.
- (1965) *Iazyk azerbaidzhanskikh kurdov* (The language of Azarbaijani Kurds), Moskva: Nauka.
- (1983) *Osnovy kurdskoj orfografii* (Principles of Kurdish orthography), Moskva: Nauka.
- Bedir Khan, D. and Lescot, R. (1970) *Grammaire kurde (dialecte kurmandji)*, Paris: Adrien Maisonneuve and Bonn: Kurdisches Institut.
- Blau, J. (1965) *Dictionnaire kurde-français-anglais*, Bruxelles: Centre pour l'Étude des Problèmes du Monde Musulman Contemporain.
- (1975) *Le kurde de 'Amādiya et de Djabal Sindjār: analyse linguistique, textes folkloriques, glossaries*, Paris: C. Klincksieck.
- (1980) *Manuel de kurde (dialecte sorani): grammaire, textes de lecture, vocabulaire kurde français et français-kurde*, Paris: C. Klincksieck.
- (1989) 'Le kurde', in CL1, pp. 327–335.
- (2000) *Méthode de kurde: Sorani*, new rev. ed., Paris: Harmattan.
- Bulut, C. (2000) 'Indirectivity in Kurmanji', in Johanson, L. and Utas, B. (eds). *Evidentials: Turkic, Iranian and Neighbouring Languages*, Berlin and New York: Mouton de Gruyter, pp. 147–184.
- Chyet, M. L. (2003) *Kurdish-English Dictionary. Ferhenga Kurmanci-Ingliz*, New Haven, CT: Yale University Press.
- Dorleijn, M. (1996) *The Decay of Ergativity in Kurmanji*, Tilburg: University Press.
- Fattah, I. K. (2000) *Les dialectes kurdes meridionaux. Étude linguistique et dialectologique*, (= Acta Iranica 37), Leuven: N.V. Peeters.
- Ferhadi, A. (1990) *Some Morphological and Morphological Features of Arbili Kurdish*, Ph.D. dissertation, University of Michigan, Ann Arbor.
- Friend, R. C. (1985) *Some Syntactic and Morphological Features of Suleimaniye Kurdish*, Ph.D. dissertation, University of California, Los Angeles.
- Haig, G. and Paul, L. (2001) 'Kurmanji Kurdish', in Garry, J. and Rubino, C. (eds) *Facts about the World's Languages*, New York, Dublin: The H. W. Wilson Company, pp. 398–403.
- Hassanpour, A. (1992) *Nationalism and Language in Kurdistan, 1918–1985*, San Francisco: Mellen Research University Press.
- Kahn, M. (1976) *Borrowing and Variation in a Phonological Description of Kurdish*, Ph.D. dissertation, University of Michigan, Ann Arbor.

- Kalbāsi, I. (1983/1362) *Guyeš-e Kordi-ye Mahābād* (The Kurdish dialect of Mahābād), Tehran: Cultural Research and Studies Institute.
- Kurdoev, Kanat K. (1978) *Grammatika kurdkogo iazyka na materiale dialektov kurmandzhi i sorani* (Kurdish grammar based on materials of the Kurmanji and Sorani dialects), Moskva: Nauka.
- Leezenberg, M. (1993) 'Gorani influence on Central Kurdish: Substratum or prestige borrowing?', X-93-03, ILLC Research and Technical Notes Series, The Netherlands. <http://www.cogsci.ed.ac.uk/~siamakr/Kurdish/Papers/Leezenberg93/text.html>
- MacKenzie, D. N. (1961) *Kurdish Dialect Studies*, 2 vols, London and New York: Oxford University Press.
- Matras, Y. (1997) 'Clause combining, ergativity, and coreferent deletion in Kurmanji', *Studies in Language* 21(3): 513–563.
- McCarus, E. N. (1958) *A Kurdish Grammar. Descriptive Analysis of the Kurdish of Sulaimania. Iraq*, New York: American Council of Learned Societies.
- (1997) 'Kurdish phonology', in Kaye, A. S. (ed.) *Phonologies of Asia and Africa*, Winona Lake, IN: Eisenbrauns, vol. 2, pp. 691–706.
- (2007) 'Kurdish morphology', in Kaye, A. S. (ed.) *Morphologies of Asia and Africa*, Vol. 2: pp. 1021–1049.
- Meho, L. I. and Maglaughlin, K. L. (2001) *Kurdish Culture and Society. An Annotated Bibliography*, Westport, CT: Greenwood Press.
- Omar, F. F. (1992) *Kurdisch-deutsches Wörterbuch*, Berlin: Kurdische Studien Berlin im VEB-Verlag für Wissenschaft und Bildung.
- Qazzaz, S. (2000) *The Sharezoor Kurdish-English Dictionary*, Arbil: Aras Press and Publishers.
- Rizgar, B. (1996) *Learn Kurdish. Dersên Kurdî*, London: Lithosphere Print & Production Network.
- Schroeder, C. (2002) 'Zur Nominalphrasenstruktur des Kurmancî', in Bublitz, W., von Roncador, M., und Vater, H. (eds) *Philologie, Typologie and Sprachstruktur, Festschrift für Winfried Boeder zum 65. Geburtstag*, Frankfurt am Main: Peter Lang, pp. 191–210.
- Tsukerman, I. I. (1986) *Khorasanskii kurmandzhi: issledovanie i teksty* (Khorasanian Kurmānji: introduction and texts), Moskva: Nauka.
- Wahby, T. (1956) *Qawā'id al-luyah al-Kurdiyah. The Kurdish Grammar*, 2 parts, Beirut: Al-Bayan Press.
- Wahby, T. and C. J. Edmonds (1966) *A Kurdish-English Dictionary*, Oxford: Clarendon Press.
- Windfuhr, G. (1986) 'A spatial model for tense, aspect and mood,' *Folia Linguistica* 19.3: 415–461 (verb systems of Persian, French, Russian, Arabic, Turkish, Creole).
- (1989) 'New Iranian languages: Overview', in *CLI*, pp. 246–262.
- Wurzel, P. (1997) *Einführung in die kurdische Sprache* (Introduction to the Kurdish language), 2 vols, Wiesbaden: L. Reichert.

CHAPTER ELEVEN

BALOCHI

Carina Jahani and Agnes Korn

1 INTRODUCTION

1.1 Overview

Balochi (Bal.) is spoken in south-western Pakistan, in the province of Balochistan as well as by smaller populations in Punjab and Sindh, and by a large number of people in Karachi. It is also spoken in south-eastern Iran, in the province of Sistan and Baluchistan, and by Baloch who have settled in the north-eastern provinces of Khorasan and Golestan. It is furthermore spoken by smaller communities in Afghanistan (particularly in the province of Nimruz), in the Gulf States (especially in Oman and the United Arab Emirates), in the Marw / Mari region in Turkmenistan, in India, East Africa, and nowadays also by a considerable number of Baloch in North America, Europe and Australia.

It is difficult to estimate the total number of Balochi speakers. Central authorities readily underestimate ethnic minorities, while members of ethnicities sometimes do the opposite. Censuses generally ignore the bi- or multilingual situation of most speakers. Moreover, large numbers of those who identify as Baloch do not speak the language any more, particularly in the areas bordering Indian languages in Punjab and Sindh, on the one hand, and in Khorasan and Golestan, on the other hand, as well as in East Africa and in the Gulf States. In contrast, Balochi has been retained quite well in Turkmenistan due to the adherence to a traditional rural lifestyle and the generally low level of education. The total number of speakers of Balochi has been estimated as being between 5–8 million (Jahani 2001: 59), but might also be somewhat higher than that.

From a historical point of view, Balochi is classed as a North-Western Iranian language (see also Korn 2003, 2005a: 328–330), although it is spoken in the south-eastern corner of the Iranian linguistic area today. It is likely that the original habitat of the Baloch was in the north-western part of the Iranian linguistic area, and that they migrated south-eastwards under pressure from the Arabic and Turkic invasions of the Iranian plateau (from mid-seventh century AD onwards). It also appears that tribes and groups of various ethnic origins, including Indo-European, Semitic, Dravidic, Turkic, Bantu and others, have been incorporated into the very heterogeneous ethnic group today known as the Baloch (see e.g. Spooner 1989: 599–600, 606–607, Korn 2005a: 43–51). On the other hand, the Balochi epic tradition provides them with a ‘true Islamic’ genealogy, according to which they are of Arabic origin and migrated from Aleppo in Syria after the battle of Karbala, where, despite the fact that the majority of the Baloch today are Sunni Muslims, the epic tradition has it that they fought on the side of the Shi’ite Imam Hussein against his enemy, the Umayyad caliph Yazid (Dames 1907, I: 1–2).

Balochi is surrounded by languages belonging to at least five language families. In the

Balochi mainland it stands in contact with other Iranian languages, Persian (Farsi and Dari) in the west and north-west, and Pashto in the north and north-east, as well as with Indic languages: Urdu, Panjabi, Lahnda and Sindhi in the north-east and east. In the Gulf States, Balochi is spoken alongside Arabic, and in East Africa with languages such as Swahili. In the central parts of Pakistani Balochistan, the Dravidian language Brahui has lived in symbiosis with Balochi for centuries, and in Turkmenistan, Balochi is in contact with Turkmen (a Turkic language). In the diaspora, Balochi is meeting new languages, mainly of the Indo-European family. It is only natural that Balochi has been influenced by all these languages, some of which are official languages and languages of education in the various states where the Baloch live. The influence of neighbouring languages will be evident in many of the sections to follow.

Balochi is neither an official language nor a language of education in any of the countries where it is spoken. This is reflected, for example, in the lack of a standard written norm for Balochi (see 1.3). There is also a dispute about which dialect, or dialects, ought to be the basis of the literary language (on the importance of the dialects, see 1.2). On the whole, writing and reading Balochi is at the moment an exclusive activity carried out by a limited number of persons belonging to the Balochi literary elite. Balochi is, as a minority language, largely restricted to traditional and informal domains such as home, neighbourhood, and traditional occupations (e.g. pastoralism and agriculture), and a career outside these traditional sectors is to a great extent linked to higher education and a good mastery of the national language. Efforts to preserve and promote the Balochi language are mainly of an unofficial character, based on private initiatives. However, there is a growing concern among the Baloch that their language may well be lost within a few generations if it does not develop a written standard.

Thus, a number of educated Baloch, mainly from Pakistan, where conditions for a long time were less unfavourable than in Iran, have since the 1950s actively attempted to preserve and promote Balochi by creating a literature in this language, establishing 'language academies' and literary societies, running periodicals, publishing dictionaries, oral poetry and folk-tales, etc. They have also tried, although with little success, to promote Balochi as a language of education. Quetta and Karachi are the main centres of these activities, although this effort has in later years spread to other places inside and outside of Balochistan where culturally active Baloch reside.

There is a Balochi Academy in Quetta, founded in 1961, the most important activities of which are the publication of books (mainly in Balochi), and arranging literary meetings. A number of periodicals have been published in the language for varying lengths of time (see Dashtyari 2003). There have also been some attempts at starting primary education in Balochi. In 1991 a state programme for mother tongue education in the Province of Balochistan, Pakistan, was launched, but it did not last long; neither did it result in any official decision on matters of language standardization. In the higher education system it is at the moment possible to study Balochi at the University of Balochistan in Quetta. Private initiatives have also been taken to teach Balochi. Radio broadcasts in Balochi are taking place in Iran and Pakistan, and there are also TV programmes in Balochi broadcast from Quetta.

Balochi oral literature remains rich when it comes to both prose and poetry. The poetic tradition consists of a large number of heroic and romantic epic poems, mainly centred on the tribal wars of the Baloch, where the two main tribes contesting for power were the Rind and Lashari tribes, and on the contests between the Baloch and the Europeans. Some of the famous Baloch heroes, who are still referred to with great pride, are the tribal chiefs Mir Chakar and Mir Gwahram as well as Hammal-i Jiand, who fought against the

Portuguese. There are also romantic ballads of a more common Islamic origin such as those of the famous couples Dosten and Shiren, and Hani and Shih Murid. The first Balochi poet known by name is Jam Durrak, court poet of Nasir Khan I of Kalat (1749–1794). As for prose, there is a rich tradition of fairy tales, tales of the life of the Prophet and his companions, and other stories told for entertainment and to teach moral lessons. Some attempts at collecting and publishing these stories have been made. The most active editors of Balochi oral literature, apart from the Baloch themselves, are M. L. Dames and Josef Elfenbein. Modern short story and novel writing is still a new genre in Balochi, although some good novelists and short story writers are emerging, particularly in Pakistan. Some factual prose writing also takes place.

1.2 Dialects and other variations

Owing to the fact that Balochi does not have a standard language (nor a standardised orthography) or standardised variants within the individual dialects, it is impossible to discuss any phenomenon of the Balochi language without making constant reference to its dialects. For a synchronic description of Balochi, it is not a viable method to present one variant as the standard and describe the others in terms of deviations from that system. Rather, the Balochi dialects need to be presented as grammatical systems in their own right (thus Rossi 1979: 179 for the Bal. phonemic systems).

The main dialect split is between Western (WBal.), Southern (SBal.) and Eastern (EBal.) Balochi. This is a very broad dialect division, within which further dialect demarcations can be made. Some dialects do not easily fit any of these groups. This is true, for example, of the dialect spoken in Iranian Sarawan, which shows transitional features between Western and Southern Balochi.

The fact that the Balochi speakers are separated into several countries with various official languages (see 1.1) is another factor which increases dialect differentiation in Balochi. The dialects in Pakistan are not only influenced by neighbouring languages, but also by the national language Urdu, and by English, the official language in British India and also in Pakistan, where it is the most important language both in administration and in higher education.

There are also sociolectal variations, often coinciding with generatiolects, since it is mainly the younger generation that has received education and thereby also a better command of the national language.

It is access to data and linguistic consultants rather than the relative importance of the dialect in question that has determined which dialects have been better described than others. Among the Pakistani dialects, the most comprehensive modern description is that of Noshke Balochi (Barker/Mengal 1969). Earlier works from Pakistan were mostly on Eastern Balochi: Dames (1891, 1907) and Gilbertson (1923, 1925). Farrell (1989, 1990, 2003) has described the Southern Balochi dialect of Karachi and Collett (1986) that of Oman. The Western Balochi dialect of Turkmenistan has attracted much interest from Russian linguists, e.g. Sokolova (1953), Sokolov (1956), Zarubin (1932, 1949) and Axenov (2006). The most comprehensive description of the dialect spoken in Afghanistan is based on data from only one speaker (Buddruss 1988). Nawata (1981) provides another brief description of this dialect. As for Iranian Balochi, descriptions are available for the Central Sarawani dialect (Baranzehi 2003) and the Lashari dialect (Yusefian 1992).

1.2.1 *Western Balochi*

The Western dialects are spoken in:

- (a) the Mari region of Turkmenistan (TBal.);
- (b) eastern Iran (IrBal.), in pockets in Khorasan and Golestan, then further south in Sistan and in the northern parts of Balochistan around Zahedan and Khash, with varieties often referred to as Sistani (not to be confused with the symbiotic cross-border dialect of Persian) and Sarhaddi (IrBalSarh.);
- (c) southern Afghanistan (AfBal.);
- (d) south-western Pakistan, including the dialects spoken in the districts of Chagai, Kharan and Kalat where the dialect around Noshke / Nushki (WBalNosh.) is of particular importance, because it is spoken by several well-known poets and writers.

WBal. dialects have sometimes collectively been referred to as Rakhshani, and sometimes this term is used for WBal. dialects spoken in the valley of the river Rakhshan in Pakistani Balochistan.

Transitional dialects between Western and Southern Balochi include Sarawani in Iran (IrBalSarw.) and Panjguri in Pakistan.

1.2.2 *Southern Balochi*

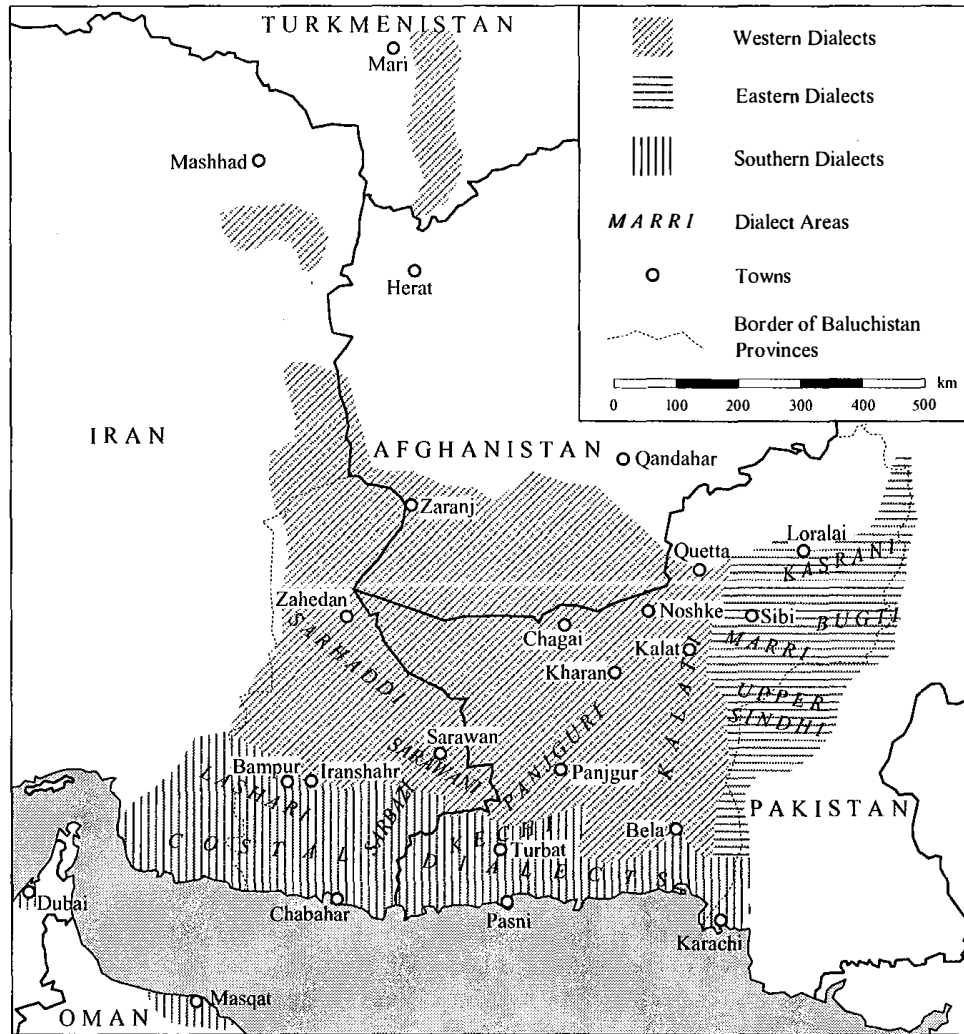
The Southern dialects are spoken in:

- (a) the southern parts of Iranian Balochistan, up to Iranshahr, dialects including Lashari, Bampuri and Sarbazi;
- (b) south-western Pakistan, up to the Kech Valley;
- (d) the Kech Valley itself;
- (e) the coastal areas and in the Gulf States;
- (f) the city of Karachi. Karachi Balochi (SBalKar.) is mainly based on SBal. dialects spoken in Iran (from where there was a heavy influx of Baloch especially after 1928) but is today also under heavy influence from surrounding Indic languages, particularly Urdu.

In particular, the Sarbazi and Kechi dialects have been important vehicles of Balochi poetry.

1.2.3 *Eastern Balochi*

The Eastern dialects constitute a conglomerate of dialects often referred to by their respective tribal names, and include (a) Maḥḥi; (b) Buḡḡi; (c) Leghari; (d) Mazari (both not on the map), (e) Upper Sindhi, and (f) Kasrani. There is too little material available to make well-motivated dialect divisions within this group, even though some of the early descriptions of Balochi (e.g. those by Dames and Gilbertson) were based on Eastern Balochi. Not only is this area largely inaccessible, particularly to foreigners, but there is also very little modern literary output.



MAP 11.1 BALOCHI

Note that the dialect areas shown are only tentative. Large areas of land are uninhabited and dialect areas may overlap due to migration, nomadism, etc.

1.3 Writing systems

Although there have been many suggestions concerning a suitable orthography for Balochi, based on Arabic/Persian/Urdu (see e.g. Jahani 1989: 134–143), Roman and Cyrillic scripts (see e.g. Axenov 2000), no one orthography has won general acceptance among the Balochi cultural elite.

When Balochi is written by the Baloch themselves, the default script employed is an Arabic-based script because it is the script used for the national languages in the countries where Balochi is predominantly spoken.

Questions concerning the pronunciation and the rendering of borrowed phonemes in the Arabic script are likewise hotly debated among the Baloch (cf. Jahani 1989: 150ff.). The typically Arabic sounds found in loans (six pharyngeals and two interdental) are

TABLE 11.1: ALPHABET COMMONLY USED FOR BALOCHI IN PAKISTAN

Letter	Represents the following phoneme(s)	Comments
ا	<i>ā</i>	also for V in vocalic onset
ب	<i>b</i>	
پ	<i>p</i>	
ت	<i>t</i>	
ث	<i>t</i>	
س	<i>s</i>	in Pers.-Ar. loanwords, <i>ṣ</i> in Eastern Balochi
ج	<i>j</i>	
چ	<i>č</i>	
ه	<i>h</i>	
خ	<i>x, k, h</i>	see 2.1.2.1 (2)
د	<i>d</i>	
ڊ	<i>d</i>	
ذ	<i>z</i>	in Pers.-Ar. loanwords, <i>ḏ</i> in Eastern Balochi
ر	<i>r</i>	
ړ	<i>r</i>	
ز	<i>z</i>	
ژ	<i>ž</i>	
س	<i>s</i>	
ش	<i>š</i>	
س	<i>s</i>	in Pers.-Ar. loanwords
ز	<i>z</i>	in Pers.-Ar. loanwords
ت	<i>t</i>	in Pers.-Ar. loanwords
ز	<i>z</i>	in Pers.-Ar. loanwords
-	-	in Pers.-Ar. loanwords (see 2.1.2.2)
گ	<i>ğ, g</i>	see 2.1.2.1 (2)
ف	<i>f, p</i>	see 2.1.2.1 (2)
ک	<i>k</i> (also: <i>x, h, ġ</i>)	see 2.1.2.1 (2)
ک	<i>k</i>	
گ	<i>g</i>	
ل	<i>l</i>	
م	<i>m</i>	
ن	<i>n</i>	
ڻ	<i>ŋ</i>	see 2.1.2.1 (1)
ن		nasal element of a nasalized vowel
و	<i>w, u, ō</i>	
و	<i>h</i>	
ی	<i>y, ī, ē</i>	for <i>ē</i> in word-final position, see below
ی	<i>ē</i>	only in word-final position

pronounced without those features, as in New Persian (Pers.), i.e. *t̤* is rendered as *t*; *ʃ* and *ʒ* are rendered as *s*; *z̤*, *ʒ̤*, *ʒ̇* as *z*; *h̤* as *h* (see also 2.1.2.1 (2), 2.1.2.2). These letters, as well as those for the peripheral and borrowed phonemes /f/, /x/, /ǧ/, and /q/ (see also 2.1.2.1 (2) and 2.1.2.2) are treated in various ways in the Balochi orthographies. While some writers use the Pers.-Ar. spelling of loanwords, others prefer to indicate the actual Balochi pronunciation. Mixed strategies are also found. The problem is rendered even more difficult by the use of some ‘Arabic letters’ for certain phonemes of the Eastern Balochi dialects (see 2.1.2.3 (2)).

Another phenomenon that has entered the Balochi writing system is ‘morpho-phonemic writings’ which means that a particular morpheme is represented by a special sign. Various such writings have been suggested for Balochi (see Jahani 1989: 153–155), but the ones that are the most commonly accepted are *é* for the oblique singular case ending *-ā* (in some dialects pronounced *-a*) and *é* for the genitive singular case ending *-ay* (also pronounced *-ē*, or even shorter) and *é* for the word *u* ‘and’.

In Turkmenistan, Balochi has been written in Cyrillic and Roman script (cf. Axenov 2000), if at all (cf. Rzehak 1998: 152).

Also among the literary elite in Pakistan there has been a debate for and against Roman script (cf. Jahani 1989: 144–149 and Elfenbein 2000) and in the 21st century, the ‘lay’ use of Roman letters for representing Balochi in emails, SMS and on Internet pages has again stirred up this discussion.

1.4 Descriptive approach

The text below describes grammatical phenomena for a wide range of Balochi dialects. The description is not exhaustive, though, since not all variants have been investigated yet, and the sources for some dialects do not contain information about all features to be discussed here. While phonology and morphology have been reasonably well investigated for a number of dialects, there are until now few phonetic descriptions for specific Balochi dialects. There are also few descriptions of sentence level syntax, particularly hypotaxis, and none as yet of discourse structure.

Particularly in the field of syntax there is a considerable degree of divergence between the written and the spoken language. However, no systematic studies have yet been made of the different registers of Balochi. Therefore the main variants of Balochi that will be described are those based on geographical distribution (geolects) rather than those based on literary register or variations that are to be attributed to social or other factors (sociolects, etc.).

Since a mere listing of the various systems might be confusing to the reader, the text will occasionally make use of the concept ‘Common Balochi’ (ComBal.), which is meant to refer to an abstract (reconstructed) entity which does not show the characteristics of a particular dialect. The term Common Balochi is thus used to illustrate which elements are common to all or most variants of Balochi.

For the sake of brevity, the description makes only few references to the sources used in the language description. All the examples will, however, contain a reference specifying the dialect and/or source. They are here somewhat unified in transcription and the (rather literal) translation, glosses are mostly added.

2 PHONOLOGY

Owing to the fact that the Balochi dialects do not share a common phonemic system (nor a common morphology and syntax), in principle, a separate description would need to be done for each variant (cf. Rossi 1979, Elfenbein 1997). More details about the words mentioned may be found in Korn (2005a: 54–70, 171–286), on which this section is largely based.

2.1 Inventory

2.1.1 Vowels

2.1.1.1 Common Balochi vowel phonemes

Phonologically, Balochi has five vowels and a quality contrast (short vs. long) for the high vowels and for the low one (*a*, *i*, *u*), while the middle vowels (*ē*, *ō*) do not have short counterparts.

The sequences *ay*, *aw* have generally been analysed as diphthongs in descriptions of Balochi, but the analysis as V + C (*a + y*, *a + w*, see 2.1.3.2) is to be preferred in view of the syllable patterns in Balochi (see 2.1.3.1). The sequences *ay*, *aw* are nevertheless included in the vowel charts since they do not pattern the same way in some variants as would *a* plus other consonants.

The ComBal. system of vowel phonemes can be posited as in Table 11.2.

TABLE 11.2: VOWEL PHONEMES OF COMMON BALOCHI

Vowels	<i>i</i> , <i>ī</i>	<i>ē</i>	<i>ō</i>	<i>u</i> , <i>ū</i>
		<i>a</i> , <i>ā</i>		
V+G	<i>ay</i>		<i>aw</i>	

Nasal vowels are found in the vast majority of Balochi dialects. In most cases, they can be analysed as allophonic sequences of V + *n* (for which see 2.3.1), but there are dialects where at least some nasal vowels seem to have acquired phonemic status (cf. Jahani 1989: 80f.; Farrell 2003: 173f.).

In some SBal. and EBal. dialects, a fronting of *ū* > *ī* has taken place (e.g. SBal. *bīl* ‘was’ vs. other dialects *būt*). Owing to adoption of loanwords it seems that there is no Balochi variant from which *ū* is entirely absent.

Several Balochi variants (independent of their affiliation to one of the major dialect groups) show a loss of quantity contrast for word-final vowels. This chiefly applies to word-final *a*, but in some dialects also to other vowels (see Korn 2005a: 251, 260).

2.1.1.2 Vowel systems of Balochi dialects

Phonetically, the Balochi dialects vary in their pronunciation of /i/, /u/ and of /ay/ and /aw/. As in many languages, the short high vowels tend to be pronounced more open than the long ones. Still, the vowel system of the dialects spoken in Afghanistan (AfBal.) and Turkmenistan (TBal.), as well as of most dialects spoken in Pakistan, is close to the system shown in Table 11.2.

Persian influence has effected a pronunciation of /i/, /u/ as *e*, *o* in most Balochi dialects spoken in Iran, and of *ay* and *aw* as approaching *ey*, *ow* in some of them. Also, /ā/ moves to an open back position under the influence of Persian in some dialects spoken in Iran. Some of these also show the break up of *ē*, *ō* to the falling diphthongs *ie*, *ue*, so that the resulting system has three short vowels (*e*, *a*, *o*), three long ones (*ī*, *ā*, *ū*), plus *ie*, *ue*. The diphthongs *ie*, *ue* differ from *ey*, *ow* in terms of syllable weight (see 2.1.3). These dialects show the vowel system in Table 11.3.

TABLE 11.3: VOWEL SYSTEMS OCCURRING IN BALOCHI DIALECTS IN IRAN

Vowels	<i>ī</i>	<i>e</i>	<i>ē</i>	<i>ō</i>	<i>ū</i>	<i>ī</i>	<i>e</i>	<i>u</i>	<i>ū</i>
			<i>a</i> , <i>ā</i>					<i>a</i> , <i>ā</i>	
Diphthongs						<i>ie</i>			<i>ue</i>
V+G		<i>ey</i>			<i>ow</i>	<i>ey</i>		<i>ow</i>	

2.1.2 Consonants

(1) Common Balochi consonant phonemes

While the Balochi vowel system is essentially in line with that of other Iranian languages, the ComBal. consonant system is simpler than one might expect from a comparative Iranian perspective (Table 11.4).

TABLE 11.4: CONSONANT PHONEMES OF COMMON BALOCHI

Stops and affricates	voiceless	<i>p</i>	<i>t</i>	<i>č</i>	<i>k</i>
	voiced	<i>b</i>	<i>d</i>	<i>ǰ</i>	<i>g</i>
Fricatives	voiceless		<i>s</i>	<i>š</i>	<i>h</i>
	voiced		<i>z</i>	<i>ž</i>	
Nasals		<i>m</i>	<i>n</i>		
Tap			<i>r</i>		
Lateral			<i>l</i>		
Glides		<i>w</i>		<i>y</i>	

(2) The Eastern Balochi consonant shift

TABLE 11.5: EASTERN BALOCHI CONSONANT SHIFT

ComBal.	EBal.	
	word-initial and postconsonantal position	postvocalic position
<i>p</i> , <i>t</i> , <i>k</i>	aspiration: <i>p^h</i> , <i>t^h</i> , <i>k^h</i>	fricatives: <i>f</i> , <i>θ</i> , <i>x</i>
<i>b</i> , <i>d</i> , <i>g</i>	(no change: <i>b</i> , <i>d</i> , <i>g</i>)	fricatives: <i>β</i> , <i>ð</i> , <i>ǰ</i>
<i>č</i> , <i>ǰ</i>	aspiration: <i>č^h</i> (no change: <i>ǰ</i>)	fricatives: <i>š</i> , <i>ž</i>
<i>w</i>	aspiration: <i>w^h</i>	(no change: <i>w</i>)

The main feature distinguishing the dialects of the EBal. group from the Southern and Western ones (collectively noted SWBal.) is a lenition of consonants in postvocalic position, which shifts voiceless and voiced stops to the corresponding fricatives and effects a loss of the dental element in the affricates (see Table 11.5). Retroflex consonants (see 2.1.2.1 (1)) are not affected by this shift.

2.1.2.1 Specific consonant classes

Contemporary Balochi dialects diverge from the system shown in Table 11.4 in several ways, but chiefly by the adoption of retroflex consonants and by the addition of fricatives.

(1) Retroflex consonants

The phonemic systems of the contemporary Balochi dialects include the retroflex phonemes *ʃ*, *ɖ* and *ʀ*, which mainly occur in loanwords from Indic (Ind.) languages, but are also found as results of developments within the language. Retroflex consonants are rather frequent, indicating that they are an integral part of the phonemic systems of most Balochi variants (for a discussion of *ʀ* see Rzehak 1998: 152). There are, however, a small number of dialects spoken in the westernmost parts of Iranian Balochistan without retroflex consonants.

In SBal., *ɳ* is attested only in a limited number of loanwords and it seems uncertain whether a majority of speakers pronounce it, or rather replace it by *n* (Collett 1986 does not include *ɳ*). However, *ɳ* seems to have acquired phonemic status in Eastern Balochi (Bashir 2008: 58).

The retroflex consonants of Indic loanwords are pronounced as such in Balochi. */t/* and */d/* in loanwords from English (which has alveolar rather than dental */t/* and */d/*) are often replaced by retroflex consonants in Pakistani Balochi as they are in Modern Indic languages, e.g. *ʃikaʃʃ* ‘ticket’, *kabāʃ* ‘cupboard’, *fuʃbāl*, *puʃbāl* ‘football’; *ɖākʃar* ‘doctor’, *ɖīs* ‘(satellite) dish’.

(2) Fricatives and *q*

While the ComBal. phonemic system (see Table 11.4) does not include fricatives other than the sibilants and *h*, the fricatives are integrated to varying degrees in the various contemporary dialects.

The complete set of *f*, *ʃ*, *x*, (*β*?), *ð*, *ǰ* is present as regular phonemes in EBal. (see 2.1.2.3 (2)). In SWBal., the fricatives *f*, *x*, *ǰ* (which chiefly occur in Persian and Arabic loanwords, but also in loanwords from other languages and potentially from Eastern Balochi) are found as peripheral phonemes (see 2.1.2.3 (1)). The tendency to replace them by sounds found in the core of the phonemic system is strong in SBal., where the fricatives are absent from the repertoire of most speakers, but weaker in the WBal. dialects. In IrBal., the pronunciation of *f*, *x* and *ǰ* in loanwords has become more usual in recent times because of the increasing influence of Persian via television and the educational system. Also, speakers of Balochi with higher education may use a ‘learned pronunciation’, specifically for Pers.-Ar. words from the sphere of religion. So there are doublets of ‘common’ and ‘learned’ pronunciation of the same word: *k(a)rāb*, (*h*)*arāb* ‘bad’ (Pers.-Ar. *xarāb* ‘out of order’); *kudā*, *xudā*, *hudā* ‘God’ (Pers. *xudā*); *kīāl*, *xayāl*, *hayāl* ‘thought, opinion’ (Pers.-Ar. *xayāl*, *xiyāl* ‘imagination, idea’); *kēriat*, (*h*)*ēriat*, *xayriat* ‘health, well-being’ (Pers.-Ar. *xayriat* ‘welfare’).

When the fricatives are not retained, Pers.-Ar. *f* is replaced by *p*: *pāida*, *fāida*

‘advantage’ (Pers.-Ar. *fā’ida*), *pilm* (English *film*). For *x*, the most common substitution is that by *h*: *hālīg* ‘empty’ (Pers.-Ar. *xālī*), *halās* ‘finished’ (Pers.-Ar. *xalās*), *halk* ‘village, people’ (Pers.-Ar. *xalq*). Since *h* disappears in some WBal. dialects (see 2.1.2.3 (1)), *x* in a loanword may appear as zero (cf. e.g. Buddruss 1988: 44; Axenov 2006: 35–36): *alās* ‘finished’, *alk* ‘village, people’, *ayrān* ‘surprised’, (*h*)*abar* ‘news, talk’ (Pers.-Ar. *xabar*). The substitution of *k* for *x* appears to be a more recent phenomenon and has been attested, for example, in Karachi Balochi (Farrell 2003: 176): *kāmōš*, *xāmūs* ‘quiet’ (Pers. *xāmōš*), *kat*, *xat(t)* ‘letter’ (Pers.-Ar. *xatt* ‘line’). The sound *ġ* is replaced by *g* in most cases: *garīb* ‘poor’ (Pers.-Ar. *ġarīb*), *ġalat* ‘wrong’ (Pers.-Ar. *ġalat* ‘mistake’), *gam(m)* ‘grief’ (Pers.-Ar. *ġam(m)*); *augān* ‘Afghan’ (Pers. *afġān*). In Afghanistan and Turkmenistan Balochi, the pronunciation of *x* and *ġ* is common (Buddruss 1988: 44; Axenov 2006: 43).

The uvular plosive *q* is as a rule replaced by *k*; substitution by *x* and *h* also occurs: *ākibat*, *ākubat*, *āxibat* ‘future’ (Pers.-Ar. ‘*āqibat*’), *kismat* ‘destiny’ (Pers.-Ar. *qismat*), *kulp*, *kulf*, *kubl* ‘(door-)lock’ (Pers.-Ar. *qufl*), *pakīr* ‘beggar’ (Pers.-Ar. *faqīr*), *arax* ‘sweat’ (Pers.-Ar. ‘*araq*’), *wahd* (Ar. *waqt*, Pers. *vaxt*). As *q* and *ġ* fall together in some variants of Modern Persian, borrowings in Balochi may also have *ġ* or *g*: *alġa-gōš* ‘earring’ (Pers.-Ar. *ħalqa* ‘ring’), *āšīġ*, *āšix*, *āšik* ‘lover’ (Pers.-Ar. ‘*āšiq*’), *āġā* ‘lord’ (Pers.-Turkic *āqā*).

(3) Aspiration

As a rule, plosives and affricates are unaspirated in Western and Southern Balochi. However, aspiration is found in several SWBal. variants.

The EBal. consonant shift, which brings about aspirated voiceless stops and *č^h* in word-initial position (see Table 11.5), has been used as a criterion to differentiate Balochi dialects. However, the EBal. aspirated voiceless stops and *č* are still the word-initial allophones of the unaspirated counterparts. Also, it seems that the significance of EBal. aspiration has been slightly overestimated. Aspiration of word-initial stop followed by a sonorant seems to be weak (if there is any) so that unaspirated word-initial stops have probably always been present to some degree in Eastern Balochi. In fact, aspiration of word-initial voiceless stops has been observed in other dialects as well (Farrell 2003: 179). Aspirated voiced stops (*b^h*, *d^h*, *ġ^h*, *g^h*) and the affricate *ġ^h* likewise occur in Indic loanwords in Eastern Balochi, but insofar as these sounds occur at all, they may be considered borrowed phonemes. However, the EBal. dialect from Kohlu district described by Bashir (2008: 58–59) appears to have contrastive aspiration for voiceless and voiced stops and sonorants.

Apart from this, aspiration does not seem to be phonemic in Balochi. In this sense, the only reason to note aspiration of voiceless stops in Eastern Balochi is that EBal. words have been cited this way so far.

(4) Geminaton

With the exception of *y*, *h* and the peripheral phonemes, all consonants may be geminated. Geminate only occur under certain conditions, though (see 2.1.3.3). They are mainly found in loanwords, but are also the result of sound changes in Balochi words.

Dialects and sources vary significantly in the occurrence (or notation) of geminates. The details remain to be studied. So far, it seems that geminates are frequent in some Western dialects and less frequent elsewhere, and gemination also appears to depend on the subdialect or other sort of variant involved.

In what follows, gemination is noted as C(C): *čam(m)* ‘eye’, *duz(z)* ‘thief’, meaning that one or several sources show gemination whereas others do not.

2.1.2.2 Pronunciation of loanwords

'Foreign' consonants other than those mentioned in 2.1.2.1 are rendered as follows in Balochi:

The majority of the Arabic loans are likely to have entered Balochi via Persian, so it is not surprising that they show features also exhibited by their form in Persian. However, Arabic words may also have been adopted via Indic languages or via direct contact. Since the source of a particular word may be difficult to determine, these words will be labelled 'Pers.-Ar.'.

Arabic 'ayn (ع) disappears in word-initial position (as it does in Persian, for the treatment of Arabic 'ayn in Persian, see Jahani 2005a): *adālat(t)* 'court' (Pers.-Ar. 'idālat 'justice'), *izzat* 'honour' (Pers.-Ar. 'izzat), *uzr* 'excuse' (Pers.-Ar. 'uzr), *akl, agl* 'intelligence' (Pers.-Ar. 'aql), *ayb* 'fault' (Pers.-Ar. 'ayb). See also 2.1.3.1.

In postvocalic position, 'ayn lengthens the preceding vowel (as in Urdu and in spoken Persian): *bād* 'later' (Pers.-Ar. *ba'd*), *tārīp* 'praise' (Pers.-Ar. *tārīf*), *tām* 'taste' (Pers.-Ar. *ta'm*), *dāwā* 'lawsuit' (Pers.-Ar. *da'wā*). Postvocalic word-final 'ayn seems to be lost: *tama* 'desire' (Pers.-Ar. *tama*).

The loss of intervocalic 'ayn or hamza (ء) may result in a sequence V + V: *pāida, fāida* 'advantage' (Pers.-Ar. *fā'ida*). A glide may then be inserted (see 2.3.2.1): *sāhat, sahat* 'hour, time' (Pers.-Ar. *sā'at*).

Sindhi *ḍ* is replaced by *ḍ* or *d*. The other implosive consonants are replaced by the corresponding non-implosive consonants.

Although *v* in loanwords (see also 2.1.2.3 (2)) is as a rule replaced by *w*, some cases of word-initial *v*- show up as *g*- or *b*- (and in Iran, also *v*-) when taken over into Balochi (see Korn 2005a: 279–280).

2.1.2.3 Consonant systems of Balochi dialects

(1) Southern and Western Balochi

As the divergences between the SWBal. consonant systems are rather minor, it is convenient to present the overall system as shown in Table 11.6.

TABLE 11.6: CONSONANTS IN SOUTHERN AND WESTERN BALOCHI

Stops and affricates	voiceless	<i>p</i>	<i>t</i>	<i>t̪</i>	<i>č</i>	<i>k</i>
	voiced	<i>b</i>	<i>d</i>	<i>ḍ</i>	<i>ǰ</i>	<i>g</i>
Fricatives	voiceless	<i>(f)</i>	<i>s</i>		<i>š</i>	<i>(x) h</i>
	voiced		<i>z</i>		<i>ž</i>	<i>(ǧ)</i>
Nasals		<i>m</i>	<i>n</i>			
Taps			<i>r</i>	<i>ɾ</i>		
Lateral			<i>l</i>			
Glides		<i>w</i>			<i>y</i>	

ž is rare in SWBal. (comparable to its status in Persian), but as there are minimal pairs, it should be considered a separate phoneme.

h is absent from the WBal. dialect of Turkmenistan and weak in the other WBal. dialects (thence noted as optional e.g. in Barker/Mengal 1969).

f, *x*, and *ǧ* are peripheral phonemes (see 2.1.2.1 (2)).

(2) *Eastern Balochi*

The EBal. consonant system (Table 11.7) differs markedly from that of the other dialects by the integration of a series of voiceless and voiced fricatives. For the status of aspiration, see 2.1.2.1 (3).

Historically, the EBal. fricatives started out as allophones of the corresponding stops, but when the consonant shift (see 2.1.2 (2)) had ceased to operate, the adoption of loanwords containing postvocalic stops and fricatives not only in postvocalic position created a new phonemic opposition.

The number of labial fricative / glide phonemes in Eastern Balochi is not entirely clear (see Korn 2005a: 61–62 for discussion). Some authors have assumed the presence of a glide *w* corresponding to ComBal. *w* (with a word-initial aspirated, or maybe rather devoiced, allophone *w^h*), and of a fricative *β* (noted *β* here for systematic reasons, without implications as to its pronunciation) resulting from postvocalic ComBal. *b* via the EBal. consonant shift. Other authors assume only one phoneme *w*, with or without an allophone *w^h*. It might also be possible that there is a phonemic opposition between *w* and *β* in some EBal. dialect(s) and none in others.

On a synchronic level, the EBal. palatals correspond to the set seen in the other dialects. However, owing to the sound change described in Table 11.5, the status of *ž* within the phonemic system is much stronger than in the other dialects. The affricates *č*, *ǰ* have been reintroduced in Eastern Balochi in postvocalic position via borrowings from other languages and dialects.

TABLE 11.7: CONSONANTS IN EASTERN BALOCHI

Stops and affricates	voiceless	<i>p^(h)</i>	<i>t^(h)</i>		<i>t^(h)</i>	<i>č^(h)</i>	<i>k^(h)</i>	
	voiced	<i>b</i>	<i>d</i>		<i>d̪</i>	<i>ǰ</i>	<i>g</i>	
Fricatives	voiceless	<i>f</i>	<i>θ</i>	<i>s</i>		<i>š</i>	<i>x</i>	<i>h</i>
	voiced	<i>β(?)</i>	<i>ð</i>	<i>z</i>		<i>ž</i>	<i>ǰ</i>	
Nasals		<i>m</i>	<i>n</i>		<i>(n)</i>			
Taps			<i>l</i>	<i>r</i>	<i>ɾ</i>			
Glides		<i>w^(h)</i>				<i>y</i>		

In some EBal. dialects spoken in Sindh the dental fricatives *θ*, *ð* change to the corresponding sibilants *s*, *z* (e.g. *č^hās* ‘spring, fountain’ vs. *č^hāθ*, SWBal. *čār*; *šuz* ‘hunger’ vs. *šud*, SWBal. *šud*, cf. Geiger 1891: 428, Grierson 1921: 414), yielding a simpler system of consonant phonemes (Table 11.8).

TABLE 11.8: CONSONANTS IN EASTERN BALOCHI DIALECTS OF SINDH

Stops and affricates	voiceless	<i>p^(h)</i>	<i>t^(h)</i>	<i>t^(h)</i>	<i>č^(h)</i>	<i>k^(h)</i>	
	voiced	<i>b</i>	<i>d</i>	<i>d̪</i>	<i>ǰ</i>	<i>g</i>	
Fricatives	voiceless	<i>f</i>	<i>s</i>		<i>š</i>	<i>x</i>	<i>h</i>
	voiced	<i>β(?)</i>	<i>z</i>		<i>ž</i>	<i>ǰ</i>	
Nasals		<i>m</i>	<i>n</i>	<i>(n)(?)</i>			
Taps			<i>r</i>	<i>ɾ</i>			
Lateral			<i>l</i>				
Glides		<i>w</i>			<i>y</i>		

2.1.3 Syllable structure and consonant clusters

The points described below instantiate a common principle: namely, they restrict the maximum weight of a syllable.

2.1.3.1 Syllable patterns

There are nine possible syllable patterns in Balochi (see Table 11.9). In all these syllable patterns, the vowel may be either long or short.

TABLE 11.9: SYLLABLE PATTERNS IN BALOCHI

V	<i>u</i> ‘and’, <i>ā</i> ‘that, he, she’, <i>ē</i> ‘this’
VC	<i>aš</i> ‘from’, <i>āp</i> ‘water’, <i>ēr</i> ‘down, below’
VCC	<i>ars</i> ‘tear (n.)’, <i>umr</i> ‘life’, <i>ārt</i> ‘flour’
CV	<i>na</i> ‘no’, <i>bō</i> ‘smell’, <i>tū</i> ‘big’
CVC	<i>man</i> ‘I’, <i>nōd</i> ‘cloud’, <i>šēr</i> ‘lion’
CVCC	<i>mark</i> ‘death’, <i>garm</i> ‘warm’, <i>bāsk</i> ‘upper arm’
CCV	<i>bra</i> ‘go (imper. 2S)’, <i>trū</i> ‘aunt’, <i>brā</i> ‘brightness’
CCVC	<i>gwar</i> ‘beside’, <i>srēn</i> ‘loins’, <i>brās</i> ‘brother’
CCVCC	<i>bramš</i> ‘twinkle’, <i>dračk</i> ‘tree’, <i>prōšt</i> ‘break’ (pret. 3S)

The pronunciation of a syllable with an initial V after a pause normally starts with a non-phonemic consonantal feature (glottal plosive or glottal fricative) (cf. also Jahani 2005a). In the northern parts of the WBal. area, the glottal plosive is normally encountered. Word-initial *h* is also normally dropped (see also 2.1.2.3). In some dialects, e.g. IRBal. of Sarbaz, the common consonantal onset feature is *h*, e.g. *[h]asp* / *[h]aps* ‘horse’, *[h]ošter* ‘camel’. In the SBal. variant of Chabahar the two onset features are employed side by side, even by the same speaker, e.g. *[h]awr* / *[ʔ]awr* ‘rain’, *[h]astē* / *[ʔ]astē* ‘he/she/it is’.

Onset clusters found in a large variety of Balochi dialects are:

- any plosive (except the retroflex ones) and *s*, *z* + *r*: *prušag* ‘to break (itr.)’, *bramš* ‘twinkle’, *tramp* ‘drip of water’, *drāj* ‘long’, *krinč* ‘wrinkle’, *grand* ‘thunder’, *srēn*;
- *b*, *d*, *g*, *ǰ*, *n*, *s*, *š* + *w*: *dwārag* ‘again’, *gwarā* ‘beside’, *ǰwān* ‘good, well’, *mwāsag* ‘grand-child’, *swār* ‘riding, mounting’, *šwānag* ‘shepherd’, *zring* ‘well’;
- any plosive (except the retroflex ones) and *m*, *n*, *s*, *z* + *y*: *pyālag* ‘cup’, *tyāb* ‘seashore’, *dyār* ‘land’, *kyāmat(t)* ‘resurrection’, *gyābān* ‘desert’, *myān* ‘middle’, *nyām* ‘middle’, *syā(h)* ‘black’, *zyāt* ‘much’;
- *sp*, *st*, *sk*: *spēt* ‘white’, *stār/stāl* ‘star’, *skūl* ‘school’ (these clusters are, however, often pronounced with a prothetic *i*:- *ispēt*, *istār* / *istāl*, *iskūl*).
- For consonant clusters involving the modal prefix *bi*-, see 2.3.2.2.

The number of possible coda clusters found in Balochi is much greater than the onset clusters. (Intrasyllabic clusters are not discussed here.) Some examples of coda clusters are:

- C plus plosive: *-pt* (*(h)apt* ‘seven’), *-tk* (*izbōtk* ‘lovage (ligusticum ajowan)’); *-čk* (*dračk* ‘tree’), *-ǰg* (*maǰg* ‘marrow, bra in’); *-šp*, *-št*, *-šk* (*mušk* ‘mouse’, *trušp* ‘sour’, *gōšt* ‘meat’); *-sk*, *-sp*, *-st* (*āsk* ‘gazelle’, *bwasp* ‘sleep’, *bust* ‘stand up (both imper. 2S)’); *-hd* (*wahd* ‘time’); *-mp* (*tramp* ‘drip of water’), *-nt*, *-nǰ*, *-nd*, *-nǰ*, *-nk*, *-ng* (*pant* ‘advice’, *lunǰ* ‘lip’),

- rand* ‘after’, *gurānd* ‘ram’, *tēlānk* ‘pushing’, *jang* ‘fight’); *-rp*, *-rt*, *-rd*, *-rk*, *-rg* (*barp* ‘snow’, *ārt* ‘flour’, *mard* ‘man’, *mark* ‘death’, *murg* ‘bird’), *-lp*, *-lk*, *-lg* (*kulp* ‘lock’, (*h*)*alk* ‘settlement, encampment, village’, *balg* ‘leaf’); *-wt* (*rawt* ‘he/she/it goes’), *-yp*, *-yb*, *-yk*, *-yt* (*kayp* ‘pleasure’, *ayb* ‘fault’, *hayk* ‘egg’, *kayt* ‘he/she/it comes’);
- C plus sibilant or affricate: *-nč*, *-nǰ*, *-nz* (*panč* ‘five’, *brinj* ‘rice’, *pōnz* ‘nose’); *-mš* (*bramš* ‘twinkle’), *-bz* (*labz* ‘word’); *-rč*, *-rs*, *-rz* (*kārč* ‘knife’, *ars* ‘tear (n.)’, *burz* ‘high’); *-wš*, *-wz* (*kawš* ‘shoe’, *jawz* ‘walnut’), *-yš* (*ayš* ‘luxury’);
 - combinations of stop plus sonorant: *-br*, *-tr*, *-dr*, *-kr* (*zabr* ‘good’, *putr* ‘enter, sneak in (imper. 2S)’, *nadr* (an interjection of appreciation), *pikr* ‘thought’); *-kl*, *-km* (*akl* ‘reason’, (*h*)*ukm* ‘order’);
 - other combinations: *-šn*, *-žn* (*gušn*, *gužn* ‘hunger’); *-sl*, *-sm* (*gusl* ‘washing’, *kism* ‘kind, sort’), *-zm*, *-zr* (*izm* ‘permission (to leave)’, *uzr* ‘excuse’); *-hl*, *-hm* (*tahl* ‘bitter’, *tuhm* ‘seed’); *-lm* (*zulm* ‘oppression’), *-rm*, *-rn* (*garm* ‘warm’, *karn* ‘century’); *-wl*, *-wm*, *-wn*, *-wr* (*kawl* ‘promise’, *kawm* ‘people’, *sawn* ‘divorce’, *hawr* ‘rain’); *-yl*, *-yr* (*sayl* ‘looking’, *hayr* ‘well’).

2.1.3.2 Diphthongs

The sequences *ay* and *aw* (see 2.1.1.1) often referred to as ‘diphthongs’ are biphonemic combinations of *a* plus a glide, which implies they are never followed by another consonant cluster or a geminate. Conversely, *ē* and *ō* (including the IrBal. variants *ie* and *ue*, see 2.1.1.2) may be followed by two consonants, but not by a geminate.

2.1.3.3 Geminated consonants

As a rule, the geminated consonants (see 2.1.2.1 (4)) are limited to the position after short vowels. Gemination is also limited to word-final or intervocalic position. According to Elfenbein (1991: 115), a word-final geminate is shortened when the next word begins with a consonant. So from the perspective of the sentence as a whole, geminates only occur in intervocalic position. Geminates count as consonant clusters in metrical texts.

Some SBal. dialects appear to geminate consonants after *ī* and *ū*. Other cases of geminates after long vowels are specific to some sources, and all of them are likely to be loanwords, e.g. *šādđī* ‘joy’ (WBalNosh., BMC).

2.2 Non-segmental features

2.2.1 Word stress

The scarce information available on word stress in Balochi suggests that it differs considerably among the dialects, and even in the very principles it follows. The diverging stress patterns appear to mirror language contact (see Korn 2005a: 67–70).

Western Balochi

For the Western Balochi dialect of Turkmenistan (cf. Axenov 2006: 48–50, partly revising Rastorgueva 1966: 328f., 334–336; Moshkalo 1991: 40, 67ff.) and Afghanistan (Buddruss 1988: 45ff.), a system of stress has been noted which is similar to that of New Persian. The stress is on the last syllable of the word as a rule, except for the individuation marker (see 4.2) *-ē*, the pronominal suffixes and the present copula, which are enclitic. The verbal

stem is also stressed on the last syllable, but the endings of the present tense are usually stressed. The negative prefix *na-* and the prohibitive prefix *ma-* attract the stress. Verb forms with the prefix *b(i)-* (see 3.2.5.2) likewise have the stress on the first syllable (thus on the first syllable of the stem if the vowel of the prefix is omitted, see 2.3.2.2), and the nominal part of complex predicates takes the phrase stress. Case endings added to monosyllabic nouns always bear the stress (on their first syllable if the ending has more than one syllable), whereas polysyllabic nouns are stressed on the last syllable of the stem, all case endings being enclitic. Inflected forms of the interrogative pronouns *čē* ‘what’ are stressed on the first syllable, however. There are also a few adverbs that are stressed on the first syllable.

Southern Balochi

According to available data (which, however, would need careful checking), the stress patterns in the SBal. dialect of Karachi differs markedly from the WBal. system in that the place of the stress depends on the weight of the syllables. A heavy syllable is one which contains a long vowel or diphthong or a short vowel followed by more than one consonant. The nasal element of nasal vowels does not count as a consonant. A word-final long vowel and a word-final short vowel followed by one consonant count as light syllables. The domain of the stress is the last three syllables of a word. Counting from the end of the word, the stress is on the first syllable that is heavy. A secondary stress may fall on the initial syllable of longer words, and some morphologically conditioned rules may override these principles.

In the SBal. dialect of Chabahar (Iran), however, stress is on the final syllable irrespective of syllable quantity.

Eastern Balochi

Information on the EBal. stress is particularly scarce. The description by Gilbertson (1925/II: 797), although presented in categories different from those used here, seems to suggest rules very similar to (if not identical with) those observed for Karachi Balochi: i.e. essentially, the last heavy syllable of a word is stressed. Complex verbs and preposition + noun are treated as a unit as far as stress is concerned.

2.2.2 *Intonation*

There is contrastive intonation in Balochi. Declarative sentences and sentences with question words are characterized by a falling pitch on the final syllable of the sentence. Interrogative sentences without a question word have a gradually rising pitch towards the end while an added tag question has a falling tone.

2.3 **Conditioned morphophonemic changes**

2.3.1 *Nasal vowels*

Nasal vowels are widespread in Balochi, specifically in the Eastern and Southern dialects; some WBal. dialects do not have nasal vowels at all. In most instances nasalisation is not phonemic, and nasal vowels can be analysed as allophonic sequences of V + *n*. There is often a pronunciation *Vn* when a vowel follows: OBL.P *-ān* + V vs. *-ā̃* + C, adj. suffix *-ēn* + V vs. *-ē̃* + C.

Nasalisation mostly affects *ān* and *ēn*; the number of dialects that show nasalisation of other vowels is smaller. In some dialects, *Vnt* may yield a nasal vowel as well: *zānt* vs. *zā* ‘knows’, *int* vs. *ē* (3S copula) in Karachi Balochi.

In some dialects, certain nasal vowels have acquired phonemic status. This applies to *ā̃* and *ē̃* in the EBal. dialect described by Bashir (2008: 60), to the vowels just mentioned and to *ō̃* in Karachi Balochi (Farrell 2003:174), and to all vowels in Sarawani (Baranzehi 2003: 81).

In the following sections, nasal vowels (e.g. *ā̃*) will be noted when the given dialect(s) do not have a non-nasalised variant; otherwise the notation will be *Vn* (e.g. *ān*).

2.3.2 Consonants

2.3.2.1 Hiatus

Adjacent vowels may be separated by a hiatus-deleting consonant (HI, see Korn 2005a: 162, 235, 242, 249, 259). This applies to hiatus caused by absence of C (e.g. *šā(h)ir* ‘poet’ (Pers.-Ar. *šā’ir*), *mahī* ‘May’) and to the addition of suffixes to stems in vowels (e.g. *ā-h-ān* / *ā-y-ān* / *ā-w-ān* DEM-HI-POBL. *čā-h-ān* ‘tea (POBL)’, *zā(h)-* / *zāy-* ‘give birth’, *grē(w)-* ‘weep’). In other cases, the vowel hiatus remains (see 2.1.2.2).

To some extent, the use and the choice of the consonant depend on the dialect involved. The use of *h* is widespread, with WBal. dialects also making use of *w* and *y*.

2.3.2.2 The modal prefix *bi-*

The vowel of the modal prefix *bi-* (see 3.2.5.2) may be rounded before back vowels in the following syllable (e.g. *bu-pōš* ‘put on (clothes)’), or it may be omitted. In this case it often assimilates to the following consonant (cf. Buddruss 1988: 59): *p-kašš* ‘pull!’ (Baranzehi 2003: 96), *p-kan* ‘do’, *b-gind* ‘see’, *b-čār* / *p-čār* ‘look’, *b-dōč* ‘sew’, *b-šōd* / *p-šōd* ‘wash’, *b-nind* / *m-nind* ‘sit’. It may merge with *p-* and *b-*, e.g. *putr* ‘enter’, *bar* ‘carry’. Before vowels, it displays the variants *b-* (before back vowels) e.g. *b-ōšt* ‘stand’, *b-ust* ‘stand up’, or *by-*, e.g. *by-ār* ‘bring’ (all examples are imperative 2S).

2.3.2.3 Verbal suffixes

Parallel to the nasalisation of vowels (see 2.3.1), the form of some verbal suffixes may depend on the following sound. In some dialects, there is a tendency for the *-t* of the 3S and the 3P (see 3.2.4) to be limited to the position before vowel while there is a reduced or nasalised form in other contexts (Baranzehi 2003: 88 note 22, cf. also the frequent alternation *farmāinī* ‘(God) commands’ vs. *farmāinīt-ī* (with 3S pronominal suffix) in Farrell 2008: 130ff.).

Similarly, the perfect participle (see 3.2.3.2) is mostly *-ag* before a suffix with vowel, but often *-a* otherwise.

2.3.2.4 Metathesis

There are several instances of metathesis of consonant clusters (see Korn 2005a: 176–178, 234, 240–241, 265–267). One type involves stops or a stop plus *s*: *waps-* vs. *wasp-* ‘sleep’, *bašk-* ‘forgive’ (also *bakš-*) (vs. Pers. *baxš-*); SBal. *gētk* vs. WBal. *gēht* < **gēxt-*, past stem of *gēj-* ‘throw’), others arrange syllables according to the preferred sonority pattern (*kulp* vs. Pers.-Ar. *qufl* ‘lock’, *tahl* vs. Pers.-Ar. *talx* ‘bitter’).

3 MORPHOLOGY

Balochi dialects differ considerably in many aspects, and the morphology is no exception. For the notation of nasal vowels, see 2.3.1.

3.1 Nominal morphology

3.1.1 Nouns

3.1.1.1 Nominal categories

There is no grammatical gender in any dialect of Balochi.

The categories relevant for the Balochi nominal system are case and number (singular (S) and plural (P), see also 4.5.1).

Mass nouns like *āp* ‘water’, *čā* ‘tea’, *nān* ‘bread’, *gōšt* ‘meat’, *galla(g)* ‘wheat’, etc. are treated as plural nouns and used with a 3P verb when they denote a certain quantity:

ē čammag-ay yaxx-ēn u pāk-ēn āp-ānā-a war-ant
 DEM spring-GEN ice-ATTR and pure-ATTR water-P.OBJ-V.EL drink.PR-3P
 ‘they drink the cold and pure water of this spring’. (TBal., ABT: 68)

In the generic meaning, they are treated as singular:

āp-a war-ant
 water-V.EL drink.PR-3P
 ‘they drink water’. (WBal.)

3.1.1.2 Case system

(1) There is no agreement in grammatical descriptions of Balochi on the number of cases and what they should be called (see Korn 2005b, 2008b). Essentially there are the following cases (see also 3.2.7, 4.1 and 4.4 for discussion of case marking):

- The direct (DIR) case (also called nominative by some authors) denotes the subject in sentences constructed nominatively and the patient in sentences constructed ergatively.
- The oblique (OBL) case is used for the agent in ergative sentences, with prepositions and in locative and adverbial functions.
- The object (OBJ) case is employed for direct and indirect objects in sentences with nominative-accusative alignment. The oblique case is also used in this function.
- The genitive (GEN) case is used for possessors and with postpositions.
- The vocative (VOC) case is used in direct address.

Points of much variation are the use of the object case suffix *-rā*, and of the genitive case. The WBal. dialects have *-ay* for the GEN.S and *-ārā* or *-ānā* for the OBJ.P. The SBal. system has genitive singular *-ē* and the OBJ.P is marked with *-ānā(rā)*, *-ānā* or *-ānrā*. In EBal., the suffix *-ān* may optionally be used in the DIR.P. The object case has the suffixes *-ār(ā)* (S), *-ār(ā)* (P). The GEN.S may be marked with *-ē*, *-a* or zero.

(2) Table 11.10 shows the case system common in dialects spoken in Pakistan.

(3) In the Balochi dialects of Afghanistan and Turkmenistan (Table 11.11), there is no separate form of the object case in the singular and the marker of the GEN.S is *-ay* for nouns while *-ī* is used on *xudā* ‘God’ and on names. Objects are found in the OBJ case.

TABLE 11.10: CASE SYSTEM OF PAKISTANI BALOCHI

	Direct	Oblique	Object	Genitive	Vocative
Singular	$-\emptyset$	$-\bar{a}$	$-\bar{a}r\bar{a}$	$-ay, -\bar{e}, -\bar{i}$	$-\emptyset$
Plural		$-\bar{a}n$	$-\bar{a}n\bar{a}, -\bar{a}nr\bar{a}$	$-\bar{a}n\bar{i}$	$-\bar{a}n$

TABLE 11.11: CASE SYSTEM OF TURKMENISTAN AND AFGHANISTAN BALOCHI

	Direct	Oblique	Object	Genitive	Locative	Vocative
Singular	$-\emptyset$		$-\bar{a}$	$-ay, \bar{i}$	$-ay\bar{a}, -\bar{i}\bar{a}$	$-\emptyset$
Plural		$-\bar{a}n$	$-\bar{a}n\bar{a}$	$-\bar{a}n\bar{i}$	$-\bar{a}n\bar{i}\bar{a}$	$-\bar{a}n$

The element $-r\bar{a}$ is not systematically used for nouns, but occurs in the pronominal paradigms as marker of the object case.

More noteworthy is a case called ‘oblique II’ by Buddruss (1988: 45–52) and ‘locative’ by Axenov (2006: 80–83, 104), which is used in local/directional function (cf. Korn 2008a). Its marker is composed of the oblique case ending $-\bar{a}$, which is suffixed to that of the genitive, thence $-ay-\bar{a}$ or $-\bar{i}-\bar{a}$, pl. $-\bar{a}n-\bar{i}-\bar{a}$.

(Note that the case ending $-\bar{a}$ in these dialects is glossed in two different ways, OBJ when it is an object (direct or indirect) and OBL elsewhere.)

(4) In IrBal., there is a restructuring in progress towards a primary contrast between singular and plural, whereas the opposition between direct vs. oblique case tends to get lost (cf. Jahani 2003). The suffix $-\bar{a}n$ is used for the functions that DIR and OBL case have in the other dialects, while the OBL.S is mostly not marked in functions other than objects (i.e. $-\emptyset$ after prepositions, for adverbial use and for the agent in ergative constructions), yielding a system that may be described as in Table 11.12. As in AfBal. and TBal., the element $-r\bar{a}$ is not systematically used for nouns, but occurs in the pronominal paradigms as case marker of the OBJ.

TABLE 11.12: CASE SYSTEM OF IRANIAN BALOCHI

	Direct	Object	Genitive	Vocative
Singular	$-\emptyset$	$-\bar{a}(r\bar{a})$	$-ey$	$-\emptyset$
Plural	$-\bar{a}n$	$-\bar{a}n\bar{a}$	$-\bar{a}n\bar{i}$	$-\bar{a}n$

Note, however, that in the southernmost IrBal. dialects (spoken in, for example, Sarbaz, Kaserkand and Chabahar) the case system is similar to that of Pakistani Balochi, insofar as the DIR.P is unmarked and the OBL.P takes the ending $-\bar{a}n$, but follows the IrBal. system for OBJ/OBL.S and OBJ.P. There are also intermediary dialects with $-\bar{a}n$ for the DIR.P.

IrBal. dialects also use the *ezāfe* construction (see Jahani 1994, 2003: 128, cf. also 3.1.3 (1)). In the dialect of the Central Sarawan valley, the GEN is replaced by the *ezāfe* construction (Baranzehi 2003: 81, see Table 11.13).

TABLE 11.13: CASE SYSTEM OF CENTRAL SARAWANI

	Direct	Object	Vocative
Singular	-Ø	-ā(rā)	-Ø
Plural	-ān	-ānā	-ān

3.1.2 Pronouns and deixis

Balochi pronouns exhibit a variation in number and function of cases similar to the one observed for nouns in 3.1.1.2. As the pronouns present more complex systems than the nouns in several respects (cf. also Korn 2005b), they are kept as one unit (i.e. without hyphens) in the morphological analysis.

3.1.2.1 Personal pronouns

Tables 11.14 and 11.15 display the pronouns for the 1st and 2nd persons. They show the same form for the DIR and OBL in several dialects. It is not clear whether the contrast DIR vs. OBL noted in the earlier sources for the EBal. IS pronoun is (still) valid, and if so, in which dialect(s). Bashir (2008: 48) has *mā*, *mā̃* for both direct and oblique case.

TABLE 11.14: PRONOUNS FOR THE 1ST AND 2ND SINGULAR

		Direct, Oblique	Object	Genitive	Locative	
1S	EBal.	<i>ma, mā, mā̃, mā</i>	<i>manā, manā̃</i>	<i>maī, maī̃</i>	—	
	SBal.	<i>man</i>	<i>manā, manārā</i>	<i>m(a)nī</i>		
	WBal.	Pakistan	<i>man</i>		<i>m(a)nā</i>	<i>mnā̃</i>
		Afgh.+Turkm.				
	IrBal.		<i>man</i>			
Sarawani		<i>mon</i>	<i>mona</i>	—		
2S	EBal.	<i>t^haw, t^ha</i>	<i>t^har(ā)</i>	<i>t^haī, t^hī̃</i>	—	
	SBal.	<i>taw, tō</i>	<i>t(a)rā, tarārā</i>	<i>taī</i>		
	WBal.	Pakistan	<i>taw</i>		<i>t(a)rā</i>	<i>tī̃</i>
		Afgh.+Turkm.	<i>ta</i>			
	IrBal.		<i>taw, ta</i>	<i>tarā, torā</i>	<i>taī, tī̃</i>	<i>tīā</i>
Sarawani		<i>ta</i>	<i>tara</i>	—		

TABLE 11.15: PRONOUNS FOR THE 1ST AND 2ND PLURAL

		Direct, Oblique	Object	Genitive	Locative	
1P	EBal.	<i>mā</i>	<i>mār(ā)</i>	<i>māī, māī</i>	—	
	SBal.	<i>mā</i>	<i>mārā</i>	<i>maē, mē</i>		
	WBal.	Pakistan	<i>mā</i>	<i>mārā</i>		<i>may</i>
		AfBal.+TBal.	<i>am(m)ā</i>	<i>am(m)ārā</i>	<i>am(m)ay</i>	<i>am(m)ayā</i>
	IrBal.	<i>mā</i>	<i>mārā</i>	<i>may</i>	—	
Sarawani	—					
2P	EBal.	<i>š(a)wā, šā</i>	<i>š(a)wār, šār</i>	<i>š(a)wāī, šāī</i>	—	
	SBal.	<i>šumā</i>	<i>šumārā</i>	<i>šumē</i>		
	WBal.	Pakistan	<i>š(u)mā</i>	<i>š(u)mārā</i>		<i>š(u)may</i>
		AfBal.+TBal.				
	IrBal.	<i>š(o)mā</i>	<i>š(o)mārā</i>	<i>š(o)mey</i>	—	
Sarawani	—					

For the functions of the non-direct cases, pronominal suffixes are also used (see 3.1.2.2), and possessive adjectives may be used instead of the genitive (see 6.1).

Demonstrative pronouns are employed for the 3rd person (see 3.1.2.3).

A separate form for the 1P inclusive (*mā-š(u)mā*) is found in the Balochi dialects of Afghanistan (Buddruss 1988: 51), Turkmenistan (Sokolov 1956: 69) and some IrBal. dialects, e.g. Sarawani; the latter also has an additional 1P exclusive (*mā-wat*, lit. 'we ourselves'), and dual (*mā-taw* 'I and you.S') (Baranzehi 2003: 85).

3.1.2.2 Pronominal suffixes

In addition to the independent pronouns, there are pronominal suffixes, also called enclitic pronouns (ENCL.PRO, Table 11.16), which are used in all functions of the oblique case, i.e. agent of ergative constructions (see 3.2.7, 4.4.1), direct and indirect objects (4.4.2), and as possessive pronouns.

TABLE 11.16: PRONOMINAL SUFFIXES

	1st			2nd			3rd		
	EBal.	WBal.	IrBal.	EBal.	WBal.	IrBal.	EBal.	SBal.	WBal., IrBal.
Sg.	<i>-ā, -ū</i>	<i>-un</i>	<i>-on, -om</i>	<i>-ē</i>	<i>-it</i>	<i>-et</i>	<i>-ī</i>	<i>-ī, -ē, -iš</i>	<i>-ē, -ī</i>
Pl.		<i>-in</i>	<i>-en</i>		<i>-ū</i>	<i>-ō, -iš</i>	<i>-iš</i>	<i>-iš, -ē</i>	<i>-iš, -ēš</i>

The pronominal suffixes are used to markedly varying degrees. Balochi dialects in Pakistan, Afghanistan and Turkmenistan use them mainly for the 3rd person, although in the literary language (e.g. oral and written poetry, modern short stories and novels) other persons are also encountered. IrBal. dialects use them frequently for all persons.

3.1.2.3 Demonstrative pronouns

The demonstrative pronouns are inflected essentially in the same way as the nouns. The distal demonstratives are also used as personal pronouns for the 3rd person.

The attested stems differ across the dialects (Table 11.17). Some demonstrative pronouns only occur in adjectival function, that is, in attributive position. Frequently, demonstratives take the prefix *ham* for emphasis (*(h)am-ā*, *(h)am-ē*, EBal. *haw-*). In some dialects, these forms have become the usual forms of the demonstratives.

TABLE 11.17: DEMONSTRATIVE PRONOMINAL STEMS

	Stem	Use	Dialects
proximal	<i>ē</i>	attributive, in direct case also substantive	all dialects
	<i>ēš</i>	inflected	
	<i>īn</i>		several dialects
	<i>ī</i>	attributive, in direct case also substantive	EBal.
	<i>(i)š-</i>		several dialects
distal	<i>ā</i>		SWBal.
	<i>ān</i>		EBal.

3.1.2.4 Other pronouns

The pronouns *kay* ‘who’ and *čē / čī* ‘what, which’ are inflected like demonstrative pronouns. The pronoun *wat* ‘self’ is used as a reflexive pronoun for functions entirely parallel to that of Persian *xod* (both < **l/wat-*). Its case forms are, depending on the dialect, OBL *wat(ā)*, OBJ *watā / wat(ā)rā*, GEN *watī*. When the possessor coincides with the subject, *watī* must be used for all persons, for example

man watī pādīšcū u taxt-ā tark-a kan-īn
 I self.GEN kingdom and throne-OBJ abandoning-V.EL do. PR-1S
 ‘I will leave my (lit. own) kingdom and throne’. (TBal.)

3.1.3 Adjectives

(1) Attributive adjectives usually precede the head noun. In this position, they are used with the suffix *-ēn*: *warnā-ēn bačak(k)* ‘the young boy’, *waš(š)-ēn hāl* ‘good news’. Adjectives can also be used as nouns, in which case they take case endings: *(h)am-ē warnā-ā* ‘that young (man) (OBL/OBJ)’. When postposed to a noun for emphasis and clarification, the adjective is treated as a noun:

gunākārī-ē mazan-ēn-ā
 sin-INDEF big-ATTR-OBJ
 'a big sin' (TBal., ABT: 88)

In fixed phrases (and in IrBal. also elsewhere), *ezāfe* constructions are used. Here, the adjective is found with or without the attributive suffix:

awlād-e nēk-ēn o swāleh-ēn
 children-EZ good-ATTR and pious-ATTR
 'good and pious children' (IrBalSarw., BSD: 84)

pa izzat-i tamām
 for honour-EZ complete
 'with great honour' (TBal., ABT: 92)

(2) The comparative is formed by adding *-tir*, and the superlative by adding the attributive ending *-ēn* to the comparative: *šar(r)-tir* 'better', *šar(r)-tir-ēn* 'better, the best' (*šar(r)* 'good'). This means that *šarrtir-ēn kitāb-ē* (with the individuation suffix *-ē*) means 'a better book' and without this suffix 'the best book'. Noteworthy irregular forms are *mas-tir* and *kas-tir* (for *mazan* 'big' and *kasān* 'small'). The adposition used in comparisons is *ša / čē* (etc., see 3.1.5) 'from, than'.

(3) A number of dialects also have analytic expressions with *ša / čē*:

āā čē mazan
 DEM.OBL from big
 'bigger than he/she', 'lit. 'from him/her/it big' (SBal., Karachi dialect, Farrell 2003: 197)

man ša zāg-ā ũ at-un
 I from child-OBL big COP.PT-1S
 'I was bigger than the child'. (AfBal., BLB: 50)

The superlative form occurs in attributive position and when the adjective functions as a noun. In predicative position the formulation 'than all' is added to the comparative to give the superlative meaning.

tājir bi swār-ānī mastir-ēn-ā gušt
 merchant to rider-P.GEN big.SUP-OBL say.PT.3S
ki pādīšā-ay zāmās am-ēš int
 CL king-GEN son-in-law EMPH-DEM COP.PR.3S
 'the merchant told the eldest (lit. biggest) of the riders that this was the king's son-in-law'. (TBal.)

mastir-ēn gušt mnī piss nājōr at
 big.SUP say.PT.3S I.GEN father ill COP.PT.3S
 'the eldest one said: my father was ill'. (TBal.)

ašpswār-ay zām ša mučč-ān ũ-tir at
 rider-GEN sword from all-P.OBL big-COMP COP.PT.3S
 'the sword of the rider was the biggest'. (TBal., ABT: 90)

3.1.4 Adverbs

Adjectives normally also function as adverbs in Balochi: *ǰwān int* ‘it is good’ vs. *ǰwān kurt* ‘he/she did well’ (WBal.). Certain adjectives in the OBL may function as adverbs: *dūr-ā* ‘hard’ (adv.), *ǰahl-ā* ‘deep down’ (see also 3.1.1.2 (1)).

Further adverbial expressions can be formed with the help of the postposition *sarā* ‘on’ (see 3.1.5), e.g. *zōr-ē sarā* ‘powerfully’ (SBalKar., FBB: 68), or the preposition *pa* ‘for’, e.g. *pa drōg* ‘lyingly’ (TBal., ABT: 97).

3.1.5 Adpositions

Prepositions, postpositions as well as circumpositions are found in Balochi. Table 11.18 shows some frequently encountered adpositions.

TABLE 11.18: BALOCHI ADPOSITIONS (SELECTION)

Prepositions	Postpositions	Circumpositions
<i>ša</i> / <i>ča</i> / <i>aš</i> / <i>ač</i> ‘from’	<i>čē</i> ‘from’	<i>ša</i> . . . <i>puštā</i> (etc.) ‘from behind (etc.)’
<i>pa(r)</i> ‘for’	<i>puštā</i> ‘behind’	<i>pa</i> . . . <i>xātir-ā</i> ‘for, for the sake of’
<i>bi</i> ‘to’	<i>tahā</i> ‘inside’	<i>bi</i> . . . <i>tahā</i> ‘into’
<i>gōn</i> ‘with’	<i>sarā</i> ‘on (top of)’	
	<i>dēmā</i> ‘in front of, before’	
	<i>čērā</i> ‘under’	
	<i>nēmagā</i> ‘towards’	
	<i>nyāmā</i> ‘between, among’	

- (1) Many postpositions are originally nouns in the OBL (literally ‘at the front/back/top of’ . . .), so the noun or pronoun is in the genitive: *čukk-ē dēm-ā* ‘in front (lit. face) of the child’ (SBal.); *ēs-ānī pad-ā* ‘behind them’, lit. ‘in the footstep of them’ (WBal.).
- (2) Prepositions are for the most part used with the OBL of nouns and pronouns, e.g. *pa zāg-ā* ‘for the child’, *ša man* ‘from me’, *gōn āw-ān* ‘with them’ (WBal.).
- (3) Circumpositions occur in most dialects; they are composed of one of the most common prepositions, i.e. mostly *ač*, *ča* (etc.) ‘from’, *pa* ‘for’ and to some degree also *bi* ‘to(wards)’, and one of the postpositions. The noun or pronoun normally follows the rule for postpositions, i.e. is in the GEN: *pattī pad-ā* ‘after you’ (TBal., ABT: 150), *pa wāntin-ay xātir-ā* ‘for the sake of studying’ (TBal., ABT: 150), *až* . . . *p^had-ā* ‘behind’ (EBal., Dames 1891: grammar: 37).
- (4) The dialects heavily influenced by Persian prefer prepositions, but those in direct contact with Indic languages chiefly use postpositions (cf. Jahani 2003: 127; Farrell 2003: 195ff.). The difference can conveniently be exemplified with two of the most common adpositions, namely the preposition *ša* (with variants) ‘from’, which is a postposition in Karachi Balochi (*čē*), and the postposition *tahā* ‘inside’, which is often encountered as a preposition in IrBal.

3.1.6 Numerals

Table 11.19 gives the ordinals found in the various dialects and sources presenting the most common variant(s) first. ° means that the given item is found as second member of a compound (like *nine* in *thirty-nine*).

TABLE 11.19: CARDINAL NUMBERS

1–9	11–19
<i>yak(k)</i> <i>dō, du</i> <i>sai, saē, sē, sah</i> <i>čār, EBal. čyār</i> <i>panč, panĵ</i> <i>šaš(š)</i> <i>(h)apt, (h)afī</i> <i>(h)ašt</i> <i>nō, nuh, nu, °nau</i>	<i>yāzda(h), yānzda(g), EBal. yāzdah</i> <i>d(u)wāzda(h), dwānzda(g) EBal. dwāzdah</i> <i>sēzda(h), sēnzda(h)</i> <i>čārda(h), BMC čārda(g), EBal. čyārdah</i> <i>pānzda(h)</i> <i>šānzda(h)</i> <i>(h)abda(h), hawdah</i> <i>(h)ažda(h)</i> <i>nōzda(h), nōzda(g), nūzdah</i>
10–90	100–
<i>da(h), FBB dā, BMC °da(g)</i> <i>bīst, EBal. gīst</i> <i>sī</i> <i>čil(l), čihil</i> <i>panĵā(h)</i> <i>šast, šašt, EBal. sai-gīst, sah-gīst</i> <i>(h)aptād, (h)afīād, EBal. sai-gīst-u-dah</i> <i>(h)aštād, EBal. čyār-gīst</i> <i>nawad, nuwad, EBal. čyār-gīst-u-dah</i>	<i>sad</i> ‘100’ <i>hazār</i> ‘1,000’ <i>lak(k)</i> ‘100,000’ <i>kurōr</i> ‘10,000,000’ <i>arab</i> ‘1,000,000,000’

Interestingly, there is a system of counting by 20s in EBal. (For more details about the attestation and history of the numerals, see Korn 2006).

arab is 1,000 million in FBB and Sayad Hashmi 2000, but 100 million in BMC (which is indeed the meaning of Urdu *arab*). Numbers like ‘21’ are compounds made in the same way as in Persian (thus *bīst-u-yak*, EBal. *gīst-u-yak*). 200 etc. is *dōldu-sad* etc.

According to Dames (1913:13ff.), subtraction strategies also occur in EBal. (e.g. *say k^ham yāzdaġ-gīst* ‘3 less than 11 times 20’ = 217).

3.2 Verb morphology

3.2.1 Simple and complex verbs

Balochi verbs are either simple or complex. Complex verbs may have a noun, adjective or preposition (preverb) as their first element (e.g. *sar kan-* ‘bring, deliver’, *tarr kan-* ‘wet’, *dar kan-* ‘take out, remove’). Borrowed verbs are often incorporated in this way, e.g. *pōn kan-* (SBal. of Karachi, Farrell 2003: 185), *zang ĵan-* (WBal. of Iran) both meaning ‘phone’, *tačīt kan-* ‘sharpen’ (TBal., Axenov 2003: 250), *ġarašakk kan-* ‘wait’ (TBal., Axenov 2003: 251), but may also be incorporated as simple verbs, e.g. *dār-* ‘hold, have’, *likk-* ‘write’, *čūtī-* ‘be freed’ (Korn 2001, A. 1).

3.2.2 Stem formation

3.2.2.1 Present and past stems

As in most Iranian languages, Balochi verbal morphology is based on the dichotomy of present stem (PR) and past stem (PT).

Most past stems are formed from the corresponding present stems by adding the suffix *-it* (which is productive and the most common suffix, also for borrowed words and denominatives), *-t*, or, more rarely, *-āt*, e.g. PR *gwār-* ‘rain’ → PT *gwār-t*, *čār-* ‘look’ → *čār-it*, *bašk-* ‘forgive’ → *bašk-āt*.

Some verbs have synchronically unpredictable past stems, reflecting Old Iranian formations (e.g. *kan-* → *kurt* ‘do’, *gind-* → *dīst* ‘see’). Others show retention of historically regular, but synchronically irregular forms: *rud-* → *rust* ‘grow’, *waps-*, *wasp-* → *wapt* ‘sleep’. For a list of verb stems, see e.g. Moshkalo (1991: 53–62) or Korn (2005a: 312–322, 344–416).

3.2.2.2 Causatives and double causatives

Causatives are formed by suffixing *-ēn-* to the present stem. Double causatives occur in many dialects. The suffix for double causatives is SWBal. *-āēn-* / EBal. *-ain-*:

man čuk-ā *ras-ēn-ā*
I child-POBL arrive.PR-CAUS-1S
‘I transport the children’, lit. ‘cause the children to arrive’.

man āyā *čē* *čuk-ā* *ras-āēn-ā*
I DEM.OBL from child-POBL arrive.PR-DBL.CAUS-1S
‘I make him/her transport the children’, lit. ‘I cause the children to be transported by him/her’. (both SBalKar., FBB: 49).

3.2.3 Non-finite forms

3.2.3.1 Infinitive

The formation of the infinitive depends on the dialect. Most dialects use *-ag* added to the present stem: *kan-ag* ‘to do’. In a number of dialects, all from the WBal. group, the infinitive is formed by adding *-in* to the past stem; thus from *kan-* → *kurt-* ‘do’: *kan-ag* / *kurt-in* ‘to do’. Infinitives may be used and inflected as a verbal noun, e.g. *pa wānt-in-ay xātirā* ‘for reasons of the studying’ (TBal., ABT: 291).

3.2.3.2 Participles

A present participle is formed by adding *-ān* to the present stem, e.g. *kan-ān* ‘doing’, *raw-ān* ‘going’, *k-āy-ān* ‘coming’.

The perfect participle is derived from the past stem by the suffix *-ag* / *-a* (see 2.3.2.3), e.g. *kurt-a(g)* ‘done’, *šut-a(g)* ‘gone’, *likkit-a(g)* ‘written’.

3.2.3.3 Gerundive

The gerundive is formed by adding *-ī* to the infinitive, e.g. *man guš-ag-ī ā* / *un* ‘I am about to say’ (SBal./WBal).

tī *piss* *trā* *wat bimman dāt-in-ī* *būt-a*
 you.S.GEN father you.S.OBJ self to.I give.PT-INF-GERV become.PT-PF.3S
 'your father himself was going to give you to me'. (TBal., ABT: 211)

3.2.3.4 Agent noun

There is an agent noun with *-ōk* added to the present stem: *kuš-ōk* 'killer', *nivīs-ōk* 'writer', which to a certain extent may also be used as a present participle.

3.2.4 Person marking and auxiliaries

The verbal endings are largely identical with the corresponding forms of the copula. Where no separate forms are given in Table 11.20, the quoted forms apply for both.

TABLE 11.20: VERBAL ENDINGS AND COPULA

		EBal.	SBal.	WBal.	Sarawani
1S	ending present	<i>-ān, -ūn</i>	<i>-ān, -un, -ō</i>	<i>-īn, -ān</i>	<i>-ān</i>
	copula			<i>un, ān</i>	
	ending past	<i>-ān</i>		<i>-un</i>	
2S	endings & copula	<i>-ē</i>		<i>-ay</i>	<i>-ey</i>
3S	ending present	<i>-ī ʔ, -t</i>	<i>-īl, -ī, -t</i>		<i>-et, -t</i>
	copula	<i>ē</i>		<i>int</i>	<i>o</i>
	ending past	<i>-∅</i>			
1P	ending present, copula	<i>-ūn</i>	<i>-ēn, -an, -in</i>	<i>-an, -ēn</i>	<i>-ēn</i>
	ending past				
2P		<i>-ē ʔ, -ē</i>	<i>-ēt, -ē, -it</i>	<i>-it</i>	<i>-ēt, -ē</i>
3P		<i>-ant, -ā</i>		<i>-ant</i>	<i>-ent, -ē</i>

The 3S ending *-t* is confined to certain verbs, and some of these forms lengthen the preceding vowel. Some of these verbs even have an alternative reduced form: *dan-t* 'gives', *jan-t* 'hits', *kan-t* 'does', *zān-t* 'knows', *raw-t* 'goes'; with lengthening: *gīr-t* 'takes', *wār-t*, *wā* 'eats'; *bār-t*, *bā* 'carries'.

The imperative has the ending zero in the 2S: *b-ra* 'go!', *p-kan* 'do!', *b-gir* 'take!' (see 3.2.6.3).

Note that in Balochi dialects that pattern ergatively (see 3.2.7), the past endings in Table 11.20 only apply to intransitive verbs, while transitive ones do not agree with the agent in the verb forms from the past stem.

Most Balochi dialects have a past tense for the copula which is composed of a stem *at-* (SBal. sometimes *it-*) and the copula forms from Table 11.20 (the 3S is without ending: *at*). Some SBal. dialects use the past stem of 'become' plus verbal endings. The copula is also used for the compound tenses (see 3.2.8.2) and in some periphrastic aktionsart and modal constructions (see 4.5.3).

The present stem of ‘become’ is (depending on the dialect) *bay-*, *bē(y)-*, *bī-*, *bū-*, *b-*, its past stem is *būt* / *bīt* (see 2.1.1.1). This verb is also used as an auxiliary for the passive voice (see 3.2.6.4) and in some of the aktionsart and modal constructions (see 4.5.3).

The *verbum existentiae* (‘to be, to exist’) is 3S (*h*)*ast* ‘there is’ (negative *nēst* ‘there is not’), 3P (*h*)*ast-ant* (see also 3.2.8.4). Its other TAM-forms are supplied by the verb ‘become’.

Other verbs used in an auxiliary-like function in some aktionsart and modal constructions include *dār-* ‘hold’, *kan-* ‘do’ and *lōt-* ‘want’ (see also 3.2.8.4 and 4.5.3).

3.2.5 Aspectual-modal affixes and negation

3.2.5.1 Imperfective markers

(1) In many dialects, some or most verbs with a word-initial vowel show a prefix *k-* in the present indicative (e.g. *k-āy-īn* ‘I come’, WBalNosh.), the present participle (*k-āy-ān* ‘coming’) and (where applicable) in the imperfective aspect of the past tense. In most IrBal. dialects (except in the very south), the verb stem is often found with a prothetic *y-* instead (*y-āy-īn* ‘I come’, WBal. of Iran).

bāndā āhinkār-ē-ā k-ār-an
tomorrow blacksmith-INDEF-OBL V.PREF-bring.PR-1P
‘tomorrow we will bring a blacksmith’. (WBalNosh.)

mučč mardum ša ham-ōdā āp-a k-āwurt-ant
all human from EMPH-there water-V.EL V.PREF-bring.PT-3P
‘all the people used to bring water from there’. (TBal.)

(2) In several Balochi dialects, a ‘verbal element’ *-a* marks the imperfective aspect (see Buddruss 1977: 9–13, 1988: 62–65; Axenov 2006: 166–170); in the non-past tense the verbal element also marks the indicative mood as opposed to the subjunctive mood (see 3.2.6.2 and 4.5.2.6).

While morphologically, the element belongs to the verb, it is enclitic to the word preceding the verb phonologically (and thus generally written as if it were a suffix). It can therefore not occur in clause initial position or after a pause. Other restrictions apply as well (for details, see Axenov 2006: 168–170).

3.2.5.2 Modal markers

The prefix *bi-* (see 2.3.2.2 for its variants and 3.2.6.3 for its use) marks the imperative and the subjunctive: *bi-ka(n)* ‘do!’, *bi-kan-it* ‘that he/she do’, *bi-kurt-ēn* ‘he/she would do.’ The prefix is often omitted in complex verbs, e.g. *bir gard* ‘return!’, *sōj kan* ‘ask!’ (TBal., ABT: 165), *dūr (p-)kan* ‘remove!’.

The past subjunctive adds *-ēn-* to the past stem, usually also employing *bi-* at the same time.

The optative is marked by the suffix *-āt-* (see 3.2.6.3).

3.2.5.3 Negation

The negative prefix is *na-*; the prohibitive prefix is *ma-* (occasionally *na-* in IrBal.), which is used for all non-indicative forms, i.e. those that also use positive *bi-*, with which both are mutually exclusive (see 3.2.5.2). In compound TAM-forms, complex verbs and periphrastic aktionsart and modal constructions, *na-* and *ma-* are added to the main verb.

3.2.6 System of tense, aspect, mood and voice

3.2.6.1 Tense

Balochi dialects essentially contrast non-past (present/future) and past tense. There is no separate construction to express the future tense.

3.2.6.2 Aspect

Several Balochi dialects show an opposition of perfective vs. imperfective aspect in the indicative mood (see also 4.5.2.5 and 4.5.2.6). In some of these dialects (WBal. of Iran, Afghanistan and Turkmenistan, Sarawani) the ‘verbal element’ *-a* (see 3.2.5.1 (2)) is employed to mark the imperfective aspect in the past tense:

dars-a wānt ‘he/she used to study (imperfective aspect)’, as opposed to
dars wānt ‘he/she studied (perfective aspect)’.

3.2.6.3 Mood

Balochi dialects show indicative, imperative, subjunctive and optative mood. The optative mood has a limited distribution (see 4.5.2.4), but is attested in, for example, IrBal. of Lashar and Pakistani WBal.

(1) In many Balochi dialects, the prefix *bi-* (see 3.2.5.2) is employed in a way parallel to Persian to generate the imperative and subjunctive. In EBal. dialects, the corresponding forms are more often found without *bi-* (cf. Bashir 2008: 75–77).

bi- plus the present stem yields the imperative sg.; the pl. is identical with the 2P subjunctive (see 4.5.2.3). *bi-* is used with the forms of the present tense to form a subjunctive (see 4.5.2.2).

bi- with the past stem (for this pattern see 3.2.6.3 and Tables 11.10 and 11.11) and the suffix *-ēn-* (according to Farrell 1990: 75 also without *-ēn-*) generates a form which denotes counterfactual modality (see 4.5.2.9 and e.g. Barker/Mengal 1969/I: 459–460; Baranzehi 2003: 99; Axenov 2006: 197–199).

(2) The optative adds the suffix *-āt-* plus the copula to the present stem, e.g. *nend-āt-en(t)* ‘may they sit’ (IrBal., Yusefian 1992: 66). The 3S has either the copula 3S or zero suffix:

ahmad yāy-āt-e
 PN come.PR-OPT-COP.PR.3S
 ‘would that Ahmad came’. (IrBal., Yusefian 1992: 59)

(3) In WBal. dialects spoken in Iran, Afghanistan and Turkmenistan the verbal element *-a* is found in the indicative mood of the non-past tense (parallel to the Modern Persian use of the prefix *mī-*), so for the non-past tense, it marks indicative mood rather than aspect (see also 3.2.6.2).

3.2.6.4 Voice

Notwithstanding the existence of ergativity in many Balochi dialects (see 3.2.7), a passive voice is also found. However, owing to the predominantly oral style of Balochi texts, it is not frequently used, and rare with an explicit agent.

It is constructed by (1) the past participle or (2) the infinitive (see 3.2.3.1) with *bay-* ‘become’ as auxiliary:

ar dukēn-ānī sar burrit-ag biit
 each two.ATTR-P.GEN head cut.PT-PF become.PT.3S
 ‘the heads of both of them were cut off’. (TBal., ABT: 200)

ē mardom-ān jat-en bīt-ē
 DEM human-P.DIR hit.PT-INF become.PT-3P
 ‘these people were beaten’. (IrBalSarw., BSD: 100)

maskat-ē tōkā lānč gir-ag bīt-ag-it-ā
 PN-GEN in boat take.PR-INF become.PT-PF-COP.PT-3P
 ‘boats had been caught in Muscat’. (SBal.Kar, Farrell 1989: 37)

(3) Eastern Balochi has acquired a morphological passive, which is formed by affixing *-īj-* to the present stem of the verb (see Bashir 2008: 60–64 for discussion).

aḡar ta āhī bārava tavk-ē kan-ay
 if you.S DEM.GEN about talk-INDEF do.PR-2S
guṛā jan-īj-a bay
 then hit.PR-PASS-V.EL become.PR.2S
 ‘if you make a complaint about him you will be beaten’. (EBal., Bashir 2008: 63)

3.2.7 Transitivity and ergativity

Transitivity is an important category in Balochi, since the construction of verb forms and the marking of arguments depend on this feature. Many dialects of Balochi show split ergativity (cf. Farrell 1989, 1995, Korn 2008c and forthc.): while an intransitive verb is inflected according to the nominative pattern throughout (namely, the endings of the verb indexing subject), transitive verbs follow a different pattern in all verb forms based on the past stem (see 3.2.2.1), including those formed from the perfect participle (see 3.2.3.2) and the past subjunctive (see 3.2.6.3 and 4.5.2.9).

Balochi ergative constructions show the logical subject (i.e. agent) in the OBL case and the object (patient) in the DIR case by default. It is common to use pronominal suffixes (see 3.1.2.2) for the agent (particularly in IrBal. dialects); they may occur in addition to an overt agent, to which they are never suffixed (see also 4.4 for further discussion).

There is no person (or gender, see 3.1.1.1) agreement of the verb with the patient in any dialect of Balochi, but there is an optional agreement in number for 3rd person patients, i.e. the 3P ending is optionally added if the patient is understood to be plural. This marking of plurality for patients is independent of animacy and definiteness.

janēn-ā mard-ē bi bādšāh-ay nēmagā dēm dāt
 woman-OBL man-INDEF to king-GEN beside face give.PT
 ‘the wife sent a man to the king’. (WBalNosh.)

bānuk-ā zahm kaššit drust jat-ant
 lady-OBL sword pull.PT all hit.PT-3P
 ‘the lady drew the sword and struck [them] all’. (SBal., Elfenbein 1983: 14)

ahmad-ā ē ketāb zort-ant
 PN-OBL DEM book take.PT-3P
 ‘Ahmad bought these books’. (IrBal. of Sarbaz)

haw-ē čāḡ kʰay-ā jaḡ-a
 EMPH-DEM well who-OBL hit.PT-PF
 ‘who has dug that well?’ (EBal., Gilbertson 1923: 121)

Balochi dialects differ to a considerable degree in their use of the ergative construction, ranging from a rather consistent ergative pattern to a complete substitution by the nominative construction in some WBal. dialects, particularly AfBal. and TBal. In these variants, the category of transitivity is not relevant:

mā āyrā bāz ·kāgad likkit-an
 we DEM.OBJ much letter write.PT-1P
 ‘we wrote many letters to him/her’. (WBalNosh., BMC I: 282)

In those IrBal. dialects that show a restructured case system (see 3.1.1.2 (4)), both agent and object of ergative constructions are in the DIR case. The agent is often also referred to by a pronominal suffix.

alī jan-i-ī košt-a o jēst-a
 PN wife-ENCL.PRO.3S-ENCL.PRO kill.PT-PF and flee.PT-PF
 ‘Ali has killed his wife and run away’. (IrBal. of Lashar, Mahmoodi Bakhtiari 2003: 143)

dars-en-a wā ā zamān-ā
 lesson-ENCL.PRO.1P-V.EL read.PT DEM time-OBL
 ‘we used to study in those days’. (IrBal. of Khash, Jahani 2003: 125)

3.2.8 Individual verb forms

3.2.8.1 Simple tenses

The endings shown in Table 11.20 are joined to the present stem to build the simple present (sometimes also called aorist). The preterite (simple past) is built by suffixing the verbal endings to the past stem. However, in the Balochi dialects with ergative patterning (see 3.2.7), this applies only to intransitive verbs whereas the bare past stem is used with transitive verbs. In EBal. dialects, the present perfect is used instead of the preterite (see 4.5.2.7).

With the affixes shown in 3.2.5, the present and preterite can be converted into modal and aspectual forms. For examples, see 4.5.2.

3.2.8.2 Compound tenses

Compound tenses are based on the perfect participle (see 3.2.3.2). They are constructed ergatively in dialects that pattern ergatively if the main verb (that is, the one represented by the perfect participle) is transitive (note the contrast to the periphrastic aktionsart and modal constructions mentioned in 4.5.3, where it is the transitivity of the finite verb that determines the type of construction).

The present perfect uses the perfect participle with the present copula, but omitting the copula in the 3S:

man kapt-ag-ō ā kapt-a man gušt-a
 ‘I have fallen’; ‘he/she has fallen’; ‘I have said . . .’
 (all SBalKar., FBB: 73)

The past perfect uses the perfect participle or the past stem with the past copula or the auxiliary ‘become’ (see 3.2.4):

<i>kalāt-ē</i>	<i>bast-ag-at-ant</i>	<i>rapt-at-ant</i>
fort-INDEF build.PT-PF-COP.PT-3P		go.PT-COP.PT-3P
‘they had built a fort’		‘they had gone’
(WBalNosh., BMC I:348)		(IrBalSarh., ASG: 17)
intransitive		transitive
<i>man kapt-a</i>	<i>bīt-ag-ō</i>	<i>āyā</i>
I fall.PT-PF	become.PT-PF-COP.PR.1S	DEM.P.OBL say.PT-PF
‘I had fallen’.		‘they had said . . .’
(both SBalKar., FBB: 73–74)		

3.2.8.3 Basic Balochi TAM-forms

The following tables show the basic Balochi TAM-forms, from which individual dialects may diverge to a certain extent.

Table 11.21 shows the forms which apply (with varying personal endings, here demonstrated with those of WBalNosh.) to WBal. of Pakistan, and, with the modifications noted in 3.2.4 and 3.2.8.2 also to SBal., and with those noted in 3.2.4 and 3.2.8.1 to EBal. Table 11.22 (with TBal. endings) presents those applying for systems with the verbal element *-a*. In these dialects the past perfect form with past stem + past copula is predominant.

TABLE 11.21: BASIC BALOCHI TAM-FORMS: SYSTEM I

Present			
indicative	subjunctive	imperative	optative
<i>kap-īn</i>	<i>bi-kap-īn</i>		<i>bi-kap-āt-un</i>
<i>kap-ay</i>	<i>bi-kap-ay</i>	<i>bi-kap-Ø</i>	<i>bi-kap-āt-ay</i>
<i>kap-īt</i>	<i>bi-kap-īt</i>		<i>bi-kap-āt-Ø</i>
<i>kap-an</i>	<i>bi-kap-an</i>		<i>bi-kap-āt-an</i>
<i>kap-it</i>	<i>bi-kap-it</i>	<i>bi-kap-it</i>	<i>bi-kap-āt-it</i>
<i>kap-ant</i>	<i>bi-kap-ant</i>		<i>bi-kap-āt-ant</i>
Past			
preterite	subjunctive	perfect	past perfect
<i>kapt-un</i>	<i>bi-kapt-ēn-un</i>	<i>kapta(g)-un</i>	<i>kaptag-at-un</i>
<i>kapt-ay</i>	<i>bi-kapt-ēn-ay</i>	<i>kapta(g)-ay</i>	<i>kaptag-at-ay</i>
<i>kapt-Ø</i>	<i>bi-kapt-ēn-Ø</i>	<i>kapta(g)-Ø</i>	<i>kaptag-at-Ø</i>
<i>kapt-an</i>	<i>bi-kapt-ēn-an</i>	<i>kapta(g)-an</i>	<i>kaptag-at-an</i>
<i>kapt-it</i>	<i>bi-kapt-ēn-it</i>	<i>kapta(g)-it</i>	<i>kaptag-at-it</i>
<i>kapt-ant</i>	<i>bi-kapt-ēn-ant</i>	<i>kapta(g)-ant</i>	<i>kaptag-at-ant</i>

TABLE 11.22: BASIC BALOCHI TAM-FORMS: SYSTEM II

Present				
indicative		subjunctive	imperative	optative
-a kap-īn		(bi-)kap-īn		(bi-)kap-āt-un
-a kap-ay		(bi-)kap-ay	(bi-)kap-Ø	(bi-)kap-āt-ay
-a kap-īt		(bi-)kap-īt		(bi-)kap-āt-Ø
-a kap-an		(bi-)kap-an		(bi-)kap-āt-ēn
-a kap-it		(bi-)kap-it	(bi-)kap-it	(bi-)kap-āt-it
-a kap-ant		(bi-)kap-ant		(bi-)kap-āt-ant
Past				
preterite	imperfect	subjunctive	perfect	past perfect
kapt-un	-a kapt-un	bi-kapt-ēn-un	kapta(g)-un	kapt-at-un
kapt-ay	-a kapt-ay	bi-kapt-ēn-ay	kapta(g)-ay	kapt-at-ay
kapt-Ø	-a kapt-Ø	bi-kapt-ēn-Ø	kapta(g)-Ø	kapt-at-Ø
kapt-an	-a kapt-an	bi-kapt-ēn-an	kapta(g)-an	kapt-at-an
kapt-it	-a kapt-it	bi-kapt-ēn-it	kapta(g)-it	kapt-at-it
kapt-ant	-a kapt-ant	bi-kapt-ēn-ant	kapta(g)-ant	kapt-at-ant

3.2.8.4 Possession

There are various constructions for the concept ‘to have’ in Balochi (see Jahani 1995):

(1) A construction corresponding to the Latin *mihi est* ‘to me there is’ is used to express permanent ownership (in EBal., ownership in general). Forms that are used for this purpose are the 3S of the copula or *verbum existentiae* (when existence or non-existence is in focus) (see 3.2.4), with or without agreement in number with the items possessed. The possessor is in the OBJ or GEN case, for example

<i>manā dagār ast</i> I.OBJ land V.EX.3S ‘I have land’. (WBalNosh., BMC I: 69)	<i>tarā brās nēst</i> you.S.OBJ brother NEG.V.EX.3S ‘you don’t have brothers’. (WBalNosh., BMC I: 73)
<i>manī bāz brās ant</i> I.GEN much brother COP.PR.3P ‘I have many brothers’. (WBalNosh., BMC I: 74)	<i>nī t^hara mahal ē</i> now you.S.OBJ leisure COP.PR.3S ‘are you now at leisure?’ (EBal., Gilbertson 1923: 187)

(2) Adpositions are used to express temporary ownership:

<i>taī kitāb gō manī brās-ā int</i> you.S.GEN book with I.GEN brother-OBL COP.PR.3S ‘my brother has your book’, lit. ‘your book is with my brother’. (WBalNosh., BMC I: 73)
--

gōn man bāz kitāb ant
 with I much book COP.PR.3P
 ‘I have many books (in the sense of: I have them with me, without necessarily owning them)’. (WBalNosh., BMC I: 73)

taī kīrrā dān ast-ē
 you.S.GEN beside rice V.EX.3S-COP.PR.3S
 ‘do you have rice?’,
 lit. ‘is there rice at your side?’ (SBalKar., FBB: 27)

(3) In some dialects (e.g. TBal., AfBal. and some IrBal. variants) the verb *dār- / dāšt* (which otherwise means ‘hold, keep, protect’) is used for ‘have’:

yē dēb mazan-ēn xazānag-ē dār-īt
 DEM demon big-ATTR locker-INDEF hold.PR-3S
 ‘this demon has a big locker’. (TBal.)

čākar u hānī sey zahg dār-ant
 PN and PN three child hold.PR-3P
 ‘Chakar and Hani have three children’. (IrBal. of Zahedan)

4 SYNTAX

4.1 Noun phrase structure

(1) As a rule, the Balochi noun phrase is head-final (left-branching): genitives and attributive adjectives are placed before the head noun. The same applies to adpositions (see 3.1.5) since, on the whole, postpositions dominate.

mnī brās-ay kitāb mazan-ēn asp gis-ay puštā
 I.GEN brother-GEN book big-ATTR horse house-GEN behind
 ‘my brother’s book(s)’ ‘big horse(s)’ ‘behind the house’. (all WBal.)

(2) When two semantically similar nouns are enumerated, nominal endings are used only once, at the end of the phrase (‘group inflection’):

bādšāh u wazīr-ay jīnikk āht-ant
 king and minister-GEN girl come.PT-3P
 ‘the king’s and the minister’s daughters came’. (WBalNosh.)

4.2 Definiteness

There is no definite article in Balochi. The term ‘indefinite article’ has been used for the suffix *-ē*, whose main function is to mark an indefinite specific form in the singular. Except for its function of marking antecedents of restrictive relative clauses (see 5.2.2), it is not used with plural nouns in Balochi. So *-ē* might be better termed ‘singularity marker’ or ‘individuation marker’, e.g. *kitāb-ē* ‘a [certain] book’ (not ‘any [one] book’).

Definiteness also plays a role in word order and in the marking and placement of objects (see 4.3.1 and 4.4.2.1), but not for verbal agreement in ergative constructions (see 3.2.7).

4.3 Clause structure and word order

4.3.1 Word order

The default word order is SOV (Subject – Object – Verb). Adverbial phrases generally follow the verb. Balochi is a pro-drop language. The order between the direct and indirect object is not fixed. A definite direct object tends to precede the indirect object, whereas an indefinite or generic one usually follows it. Alternative constituent orders may occur for pragmatic reasons (see 4.3.2).

yakk pīr čilla-kašš-ē manā damī-ē sōp-ē
 one holy.man forty-pull.PR-INDEF I.OBJ blessed-ATTR apple-INDEF
 dāt-ag
 give.PT-PF.3S
 ‘a holy man observing a period of fasting and praying gave me a blessed apple’.
 (WBalNosh.)

man digar-ay rōzī-ā bitta dāt-ag-a na-kan-īn
 I other-GEN daily.bread-OBJ to.you.S give.PT-PF-V.EL NEG-do.PR-1S
 ‘I cannot give someone else’s daily bread to you’. (AfBal.)

šut-ant bādšāhī māṛī-ā
 go.PT-3P royal mansion-P.OBL
 ‘they went to the royal mansions’. (WBalNosh.)

bakī-y-ā dāt bi rā-y-i xudā
 rest-HI-OBJ give.PT.3S to way-HI-EZ God
 ‘he gave the remainder as charity’. (AfBal.)

4.3.2 Focus and prominence

A constituent can be placed in the postverbal or preverbal position, or be left-dislocated to the beginning of the sentence for prominence.

sundūk-ā mōlid-ā baḍḍ dāt-ant
 chest-OBL maidservant-P.OBL back give.PT-3P
 ‘the maidservants carried the chest’. (WBalNosh.)

putr am-ē sandūx-ay tā
 enter.PR.SBJ EMPH-DEM chest-GEN in
 ‘get inside this chest!’ (TBal., ABT: 233)

ē kār-ā man kurt-un
 DEM work-OBL I do.PT-1S
 ‘it is I who has done this’. (WBalNosh.)

Cleft constructions are another strategy for marking focus:

ā čīz-ē ki šmā dīst-ag-it
 DEM thing-INDEF CL you.P see.PT-PF-COP.PR.2P
ā mardum-ē būt-a
 DEM human-INDEF become.PT-PF.3S
 ‘what you have seen turned out to be a man’. (TBal., ABT: 240)

4.3.3 Nominal clauses

Balochi also employs nominal clauses, but only for the 3S of the non-past indicative. These sentences have a nominal predicate followed by the individuation marker *-ē* (see 4.2), e.g.:

<i>murg-ē</i> Ø	‘it is a bird’, but
<i>murg-ē at</i>	‘it was a bird’, and
<i>bačakk-ē ay</i>	‘are you a boy?’

4.4 Case marking

4.4.1 Subject marking

In nominative constructions, the subject is in the direct case, which has a zero suffix in both the singular and the plural (except in IrBal., see 3.1.1.2 (4)). This applies to the transitive agent of verb forms based on the non-past stem, to the intransitive subject in all dialects, and to the transitive agent in non-ergative dialects (see 3.2.7).

In ergative constructions, the agent is in the oblique case. This applies to the transitive agent of verb forms based on the past stem in ergative dialects. Since pronominal suffixes (see 3.1.2.2) function as unstressed variants of the oblique case pronouns, they may also be used for ergative agents.

<i>wāja ālī gopt-ī</i>	<i>baḡd-ī</i>	<i>ko</i>	<i>karbās-ānā</i>
mister PN say.PT-ENCL.PRO.3S	back-ENCL.PRO.3S	do.PT	cotton-P.OBJ
‘Mr. Ali said . . .’	‘she carried the cotton’.		
(IrBalSarw., BSD: 105)	(IrBalSarw., BSD: 106)		

4.4.2 Object marking

4.4.2.1 Direct object

The marking of the direct object by the oblique and object case is governed by the principles of DOM (differentiated object marking). A direct object that is generic/indefinite, non-specific and inanimate is in the direct case (see also Farrell 1990: 65) whereas definite direct objects are in the oblique or object case (see 3.1.1.2 (1)).

In sentences constructed ergatively, the direct object is usually in the direct case, but it may also be in the oblique or object case if it is definite. This is even the predominant pattern for the 1st and 2nd person pronouns (cf. Farrell 1995, Korn, *forthc.*), but the DIR of these pronouns is occasionally found when a pronominal clitic follows.

<i>nām-ā</i>	<i>har kas-ā</i>	<i>zānt</i>	<i>dīt-iš</i>	<i>mard-ārā</i>
name-OBL	each person-OBL	know.PT	see.PT-ENCL.PRO.3P	man-OBJ
‘everyone knew the name’.			‘they saw the man’.	
(SBal., Elfenbein 1983: 30)			(SBal., Elfenbein 1983: 16)	

<i>ēšān</i>	<i>ki wāja panč-kuš-ā</i>	<i>dīst</i>	<i>sakk waššāl būt-ant</i>
DEM.P.OBL	CL mister five-kill.PR-OBL	see.PT	very happy become.PT-3P
	‘when they saw Mr. Five-slayer, they became very happy’.		
	(WBalNosh., BMC II: 178)		

čī-ā man-et weyl kot-ag
 what-OBL I-ENCL.PRO.2S abandoned do.PT-PF
 ‘why have you left me?’ (IrBal. of Sarbaz)

rāh-ā mnā tunn-ā jat-a
 way-OBL I.OBJ thirst-OBL hit.PT-PF
 ‘on the way, I was overcome by thirst’,
 lit. ‘thirst struck me’. (AfBal., Rzehak 1998: 178)

4.4.2.2 Indirect object

Indirect objects are marked with OBL or OBJ. OBL is used only for nouns, and in dialects following the case system outlined in Table 11.10. OBJ can also be used to mark nouns as indirect objects in these dialects.

mō dāt-ē ramazān-a ke ra
 I give.PT-3P PN-OBJ CL go.PT.3S
 ‘I gave [them] to Ramazan, who [then] went’. (IrBalSarw., BSD: 83)

gallapān-ā bādšāh-ārā hāl dāt
 horse.herder-OBL king-OBJ news give.PT
 ‘the horse herder gave news to the king’. (WBalNosh., BMC II: 174)

manā lōt-īt ki har čō bi-būt
 I.OBJ want.PR-3S CL each how SBJ-become.PR.3S
man ā mazār-ā bi-gir-īn u zindag-ā bādšāh-ā pēš
 I DEM tiger-OBL SBJ-take.PR-1S and alive-OBL king-OBL before
bi-dār-īn
 SBJ-hold.PR-1S
 ‘by all means he wants me to catch that tiger and bring (lit. show) it alive to the king’.
 (WBalNosh., BMC II: 175)

In TBal., AfBal. and IrBal. indirect objects constructed with the preposition *bi* ‘to’ (dative drift) are frequently encountered:

ta be čākar manā bakšet
 you.S to PN I.OBJ give.PT
 ‘you gave me to Chakar’. (IrBal. of Khash, Jahani 2003: 126)

4.5 Semantics and use of verb forms

4.5.1 Concord

Except for ergative constructions (see 3.2.7 and 4.4), verbs most often agree with the subject in number and person (see also 3.1.1). However, there are cases of 3P subject with singular verb:

ar dukēn-ānī sar burrit-ag būt
 each two.ATTR-P.GEN head cut.PT-PF become.PT.3S
 ‘the heads of both of them were cut off’. (TBal., ABT: 200)

4.5.2 Use of TAM-forms

4.5.2.1 Present indicative

The present indicative basically denotes continuous and habitual actions and states. It may also imply present continuous and future time reference. It is thus used for habitual actions or states and ongoing actions in the non-past temporal field, as well as for general truths and future events. It is, furthermore, used for past events in narrative style (historical present). See 3.2.5.1 (1).

man ham-ē rang-ē mard-ē un
I EMPH-DEM kind-ATTR man-INDEF COP.PR.1S
u ē dawl-ē sitāh-ānā k-arz-īn
and DEM kind-ATTR praise-P.OBJ V.PREF-deserve.PR-1S
'I am such a man and I am worthy of this kind of praise'. (WBalNosh., BMC II: 173)

man watī laṭṭ u kawš-ānī badal-ā gir-īn
I self.GEN stick and shoe-P.GEN exchange-OBL take.PR-1S
'I will take revenge for having been beaten up'. (WBalNosh.)

šumā šut-ag-it mā k-ā-h-ān
you.P go.PT-PF-COP.PR.2P we V.PREF-come.PR-HI-1P
'you have gone [to the afterworld already], we will come [as well]'. (Badalkhan 2008: 302)

gwast-ag-ēn dawr u zamān-ān yakk warnā-ē
pass.PT-PF-ATTR period and time-P.OBL one young.man-INDEF
pa dukkāndār-ay jinēn-ā āšig-a bīt
for shopkeeper-GEN woman-OBL enamoured-V.EL become.PR.3S
'in ancient times a young man fell in love with the shopkeeper's wife'. (TBal., ABT: 183)

4.5.2.2 Present subjunctive

The present subjunctive denotes uncertain actions and states (irrealis) in the non-past temporal field. It is employed in modal constructions (see 3.2.6.3 and 4.5.3) and various types of subordinate clauses (see 5.2.3) as well as in exhortations, admonitions, wishes, etc.

marōčī madrasag-ē-ā bi-rav-īn
today school-INDEF-OBL SBJ-go.PR-1S
balkin čī-ē by-ār-īn
perhaps thing-INDEF SBJ-bring.PR-1S
'let me go to school today, maybe I will bring something [from there]'. (WBalNosh.)

aga da b-day-ay sad-a dant
if ten SBJ-give.PR-2S hundred-V.EL give.PR.3S
'if you give ten, he will give a hundred'. (AfBal.)

4.5.2.3 Imperative

The imperative is used for commands.

pīr gušt šarr tav ē laṭṭ-ā b-zūr
 holy.man say.PT.3S good you.S DEM stick-OBL SBJ-take.PR
b-ra āšk-ay sōp-ay dračk-ā b-ĵan
 SBJ-go.PR that.side-GEN apple-GEN tree-OBL SBJ-hit.PR
 ‘the holy man said: Well, take this stick, [and] go and hit the apple tree over there’.
 (WBalNosh.)

4.5.2.4 Optative

The optative is used in literary style and formalized phrases to communicate wishes. Most dialects (in non-literary style all dialects) replace it with the subjunctive.

gul bi-rič-āt
 flower SBJ-pour.PR-OPT.3S
 ‘may flowers shower [down]’. (Badalkhan 2008: 302)

durāh bāt-ay
 healthy become.PR.OPT-2S
 ‘farewell’, lit. ‘may you be healthy’. (WBalNosh.)

manī zird-ay sirāb-ēn wāhag int ki
 I.GEN heart-GEN burning-ATTR desire COP.PR.3S CL
gam-ay garm-ēn gwāt tāi dēm-ā ma-kašš-āt
 sorrow-GEN warm-ATTR wind you.S.GEN face-OBL PROH-pull.PR-OPT.3S
 ‘the burning desire of my heart is that the hot wind of sorrow should not blow upon
 your face’. (WBal.)

4.5.2.5 Preterite

The preterite (simple past) indicates past tense with perfective aspect, that is, events completed in the past and viewed as a whole with no internal time structure. It is the most common verb form in narratives. It is additionally used as a relative tense referring to future events in temporal and conditional clauses (see 5.2.3.1 and 5.2.3.6) when the subordinate clause is backgrounded and/or assumed to take place before the main clause and for events which are about to happen.

pōryāgir paṭṭiṭ dīst šinikk-ā gipt
 labourer search.PT.3S see.PT.3S goat.kid-OBL take.PT.3S
dast-ā burt šinikk-ay nukkk-ā
 hand-OBL take.PT.3S goat.kid-GEN mouth-OBL
inčukk-ē trišk-ē kaššit ĵan-ā dāt-ē
 such-ATTR piece-INDEF pull.PT.3S wife-OBL give.PT.3S-ENCL.PRO.3S
 ‘the labourer searched and saw the goat kid, caught it and put his hand in the goat
 kid’s mouth, pulled out a small (lit. such a) piece and gave it to the wife’.
 (WBalNosh.)

yakk sōp-ē kapt kapp-ē watī ĵan-ā b-day
 one apple-INDEF fall.PT.3S half-INDEF self.GEN wife-OBL SBJ-give.PR
 ‘[if] one apple falls, give half [of it] to your wife’. (WBalNosh.)

ešo yaht-ē
 look! come.PT-3P
 ‘here they come (lit. came)’. (IrBalSar., BSD: 92)

4.5.2.6 Imperfect

In those dialects that have a distinct imperfective past it is used for ongoing and habitual states or events in the past temporal field. Dialects without this verb form employ the preterite in these instances.

iškārī pa watī kār-ā bāz bāz pašōmānī-a kaššit
 hunter to self.GEN work-OBL much much regret-V.EL pull.PT.3S
 ‘the hunter was (being) very regretful of his deed’. (Tbal., ABT: 190)

4.5.2.7 Present perfect

The present perfect expresses events in the past which are relevant to the present. For transformative verbs (‘lie down’, ‘stand up’, ‘sit down’, etc.) it has a stative meaning. In EBal. the perfect form is also used where other dialects have the preterite (and imperfective past), see 3.2.8.1.

dīst-ī ki āā mazār pa radī-ā
 see.PT-ENCL.PRO.3S CL DEM.OBJ tiger for mistake-OBL
gipt-a u bast-a
 take.PT-PF and tie.PT-PF
 ‘he saw that he had caught and tied up the tiger by mistake’. (WBalNosh., BMC II: 177)

yakk ballī-ē ništ-a
 one grandmother-INDEF sit.down.PT-PF.3S
šīr u šīlānč kan-ag-ā int
 milk and whey do.PR-INF-OBL COP.PR.3S
 ‘an old woman is sitting and making whey (out of milk)’. (WBalNosh.)

mā zahm ārḏ-ag-ant
 I.OBL sword bring.PT-PF-3P
 ‘I have brought the swords’. (EBal., Gilbertson 1923:113)

There also seems to be an inferential use of the perfect (for assumptions, indirect report, etc., see also Rossi 1989).

wat dōšī bē-wāb būt-a u atk na-kan-t
 self last.night without-sleep become.PT-PF.3S and come.PT NEG-do.PR-3S
 ‘as for him, he was sleepless last night and can’t come’. (WBalNosh., BMC II: 177).
 Actually, he was not sleepless according to the story, but the wife reports him to have been sleepless as an excuse.

4.5.2.8 Past perfect

The past perfect indicates that the situation in question is located prior to another reference point in the past. For transformative verbs (see 4.5.2.7) it expresses a past state.

naryān ki hičč swārī na-dīst-ag-at,
 stallion CL any rider NEG-see.PT-PF-COP.PT.3S
traḡḡ-ān u gwānz war-ān-ā dēm pa laškar-ay
 skip.PR-PART and swing eat.PR-PART-OBL face to army-GEN
nēmagā šut
 towards go.PT.3S
 ‘the stallion, which had not carried (lit. seen) any rider [before], went skipping and swinging towards the army’. (WBalNosh., BMC II: 180)

pūll-ē tamām kort-at
 money-ENCL.PRO.1P finished do.PT-COP.PT
zang-ē jat ke pūll be-y-ār-ēt
 phone-ENCL.PRO.1P hit.PT CL money SBJ-HI-bring.PR-2P
 ‘we had run out of money, [so] we phoned [and told them]: bring money’.
 (IrBalSarw., BSD: 95)

yak rōč-ē sardār bi watī gis-ā ništ-at u
 one day-INDEF chief in self.GEN house-OBL sit.PT-COP.PT.3S and
gwarag-ay kabāb wārt-ī
 lamb-GEN roast.meat eat.PT.3S-ENCL.PRO.3S
 ‘one day the chief was sitting in his house and ate roast lamb’. (TBal., ABT: 197)

4.5.2.9 Past subjunctive

The past subjunctive expresses counterfactual events or actions such as doubts, unreal conditions and wishes in the past and non-past temporal fields.

na-ke ōda be-rapt-ēn-ā āhīā kār kuḡ-ē
 NEG-CL there SBJ-go.PT-SBJ-IS DEM.OBL work do.PT-SBJ
 ‘I wish I had gone there’. ‘if he had worked’
 (IrBalSarw., BSD: 99) (EBal., Bashir 1991: 105)

aga ē jāga-ay badalā taw diga bādšāhī-ā būt-ēn-ay,
 if DEM place.GEN instead you.S other kingdom-OBL become.PT-SBJ-2S
taī bāz qadr u wāhag bīt
 you.S.GEN much recognition and desire become.PT.3S
 ‘if you had been in another kingdom instead of in this place you would have been given due honour’. (WBalNosh., BMC II: 172)

It can also be used in counterarguments:

mnī piss mnā arčōn dōst dāšt-ēn ham,
 I.GEN father I.OBJ although friend hold.PT-SBJ.3S also
pa wānt-in-ay xātir-ā mnā-a gipt u jat
 for study.PT-INF-GEN sake I.OBJ-V.EL take.PT.3S and hit.PT.3S
 ‘even though my father loved me very much, he took and beat me because I did not study’. (TBal., ABT: 265)

In subordinate clauses it may denote iterative past:

puč wa nān wa masrap-ē ki
 clothes and bread and consumption-INDEF CL

man bi-dāšt-ēn-un mnā dāt
 I SBJ-hold.PT-SBJ-1S I.OBJ give.PT.3S
 'he gave me clothes and food and [other] consumables that I needed'. (AfBal.,
 BLB: 22)

4.5.3 Aktionsart and modal constructions

There are various periphrastic constructions that express both aktionsarten and modalities. These are either nominal constructions or, in Persian style, double finite constructions (for the nominal forms, see 3.2.3), as shown in Table 11.23 and 11.24. For a (not necessarily exhaustive) list of the dialects in which these constructions are found, see the examples.

4.5.3.1 Aktionsart constructions

TABLE 11.23: AKTIONSART CONSTRUCTIONS

nominal constructions	Persian-style constructions
(1) Progressive / Ingressive INF-OBL + COP PR.PART + COP INF-OBL + <i>lagg-</i> / <i>laggit</i> 'begin' (bi) INF-OBL + <i>binā kan-</i> / <i>kurt</i> 'begin' or <i>binā kan-</i> / <i>kurt</i> + <i>bi</i> + INF-OBL	<i>golāyīš</i> + COP + SBJ <i>dār-</i> / <i>dāšt</i> 'hold' + SBJ
(2) Terminative INF-OBL + <i>ill-</i> / <i>išt</i> 'leave'	(3) Immediate future <i>āy-</i> / <i>āt</i> 'come' <i>ki</i> + SBJ

(1) Progressive / Ingressive

Several Balochi dialects have a syntactic construction comparable to the English 'continuous form', employing the infinitive (see 3.2.3.1) in the OBL (thus a kind of locative expression) with the copula to express the progressive.

man guš-ag-ā un man guš-ag-ā it-ō
 I say.PR-INF-OBL COP.PR-1S I say.PR-INF-OBL COPPT-1S
 'I am saying' (WBal.) 'I was saying' (SBalKar., FBB: 74)

In Pakistani SWBal. there is a continuous construction based on the present participle + copula:

wāḷa panč-kuš bi har-ā swār rav-ān at
 mister five-kill.PR to donkey-OBL riding go.PR-PART COPPT.3S
 'Mr. Five-slayer kept moving forward riding on his donkey'. (WBalNosh., BMC
 II: 178)

Central Sarawani has a special progressive construction consisting of *golāyīš* + copula employed with the present form of the verb:

golāyīš-ē zapt-a kan-ē
 PROG-COP.PR.1P recording-V.EL do.PR.SBJ-1P
 'we are busy recording'. (IrBalSarw., BSD: 90)

golāyīš-at-ē dar-a yāy-ē
 PROG-COP.PT-1P out-V.EL come.PR-1P
 'we were coming out'. (IrBalSarw., BSD: 93)

Some variants of IrBal. (e.g. Sarhaddi Balochi of Granchin, described by Ahangar 2007) have acquired a progressive/ingressive construction modelled on the Persian construction with the verb *dār- / dāšt* 'hold, have':

dāšt-on rapt-on ke ta yaht-e
 hold.PT-1S go.PT-1S CL you.S come.PT-2S
 'I was about to go when you came'. (IrBalSarh., ASG: 14–15)

kessa dār-īt gehter o gehter-a bīt
 story hold.PR-3S better and better-V.EL become.PR.3S
 'the story is getting better and better'. (IrBalSarh., ASG: 11–12)

Ingressives may be expressed by the infinitive in the oblique with the verb *lagg- / laggit* 'begin'. In TBal., *binā kan- / kurt* 'to start' may also be used in this sense.

man giš-ag-ā lagg-ā
 I say.PR-INF-OBL start.PR-1S
 'I begin to say'. (SBal.Kar., FBB: 76)

čārit-in-ā laggit binā kurt bi zulm-ay
 look.PT-INF-OBL start.PT.3S building do.PT.3S to cruelty-GEN
kurt-in-ā
 do.PT-INF-OBL
 'it started to graze'. 'he started to do cruel things (lit. cruelty)'.
 (both TBal., ABT: 218)

nān u āp-ay jōr kurt-in-ā binā-a kan-t
 bread and water-GEN ready do.PT-INF-OBL building-V.EL do.PR-3S
 'she starts to prepare the meal'. (TBal.)

(2) Terminative

TBal. also has a terminative construction with *ill- / išt* 'to leave':

watī jang kurt-in-ā išt
 self.GEN fight do.PT-INF-OBL leave.PT.3S
 'he stopped quarrelling'. (TBal., ABT: 218–219)

(3) Immediate future

TBal. also has a construction for immediate future, in the present or past, employing the motion verb *āy- / āt* 'come' followed by a purpose clause in the subjunctive:

pādišā k-ayt ki āyrā maland u masxara b-kan-t
 king V.PREF-come.PR.3S CL DEM.OBJ joke and ridicule SBJ-do.PR-3S
 'the king is going to ridicule him'. (TBal.)

āt ki dār-ānā baḡḡ kan-t
 come.PT.3S CL wood-P.OBJ back do.PR.SBJ-3S
 'he was just going to load up the wood on his back'. (TBal.)

4.5.3.2 Modal constructions

TABLE 11.24: MODAL CONSTRUCTIONS

nominal constructions	Persian-style constructions
(1) Ability / possibility PT(-PF) + <i>kan-</i> / <i>kurt</i> 'do'	<i>t(o)wān-</i> / <i>t(o)wānt</i> 'be able' + SBJ
(2) Intention / ability GERV + COP	
(3) Necessity / intention noun-OBL GERV + <i>int</i> noun-OBL INF + <i>lōtī(t)</i>	
(4) Desiderativity INF + <i>lōt-</i> / <i>lōtīt</i> 'want'	

(1) Ability / possibility

Ability may be expressed with *kan-* / *kurt* 'do' as auxiliary:

dīst-ag-a *kan-ay* *šut* *na-kurt*
see.PT-PF-V.EL do.PR-2S go.PT NEG-do.PT.3S
'you can see' (TBal., ABT: 223) 'he could not go'. (WBal.)

kut *kan-ē* *āyā* *kapt* *kut-ag-ā*
do.PT do.PR-2S DEM.P.OBL fall.PT do.PT-PF-3P
'you can do' 'they could have fallen'. (both SBalKar., FBB: 58, 75)

In IrBal. this construction is less frequently found. Instead, the construction modelled on Persian with *t(o)wān-* / *t(o)wānt* 'be able to' + subjunctive of the main verb is normally used to express ability:

mō šerā *na-towān-ā* *čest* *kan-ā*
I DEM.OBJ NEG-be.able.PR-1S lifting do.PR.SBJ-1S
'I cannot lift this up'. (IrBalSarw., BSD: 97)

(2) Intention / ability

Pakistani SWBal. also has constructions expressing intent, ability and necessity based on the gerundive (see 3.2.3.3) + copula.

man raw-ag-ī *un*
I go.PR-INF-GERV COP.PR.1S
'I intend to / will be able to / have to go'. (WBalNosh., BMC I: 238) / the same construction means 'I am about to go' in (SBalKar., FBB: 76)

(3) Necessity / intention

manā raw-ag-ī *int*
I.OBJ go.PR-INF-GERV COP.PR.3S
'I have to go'. (WBalNosh., BMC I: 240; same construction SBalKar., FBB: 76)

Another construction expressing necessity, which consists of infinitive + the verb *lōt-* / *lōtīt* 'want, wish', is found in SBal. of Karachi:

manā likk-ag lōṭ-ī
 I.OBJ write.PR-INF want.PR-3S
 'I should write', lit. 'to me it needs to write'. (SBalKar., FBB: 45)

(4) Desiderativity

The verb *lōṭ- / lōṭīt* 'want, wish' is also used to express desiderativity. In some Balochi dialects the construction employs the infinitive + *lōṭ-* (*raw-ag lōṭ-īn*), in others it is constructed in the Persian way with the subjunctive of the main verb, e.g. *lōṭ-īn b-rīn* 'I want to go'. This is particularly common in Balochi dialects in close contact with Persian (IrBal., TBal., AfBal.).

lōṭīt ki mnā watī zāmās b-kan-t
 want.PT.3S CL I.OBJ self.GEN son-in-law SBJ-do.PR-3S
 'he wanted to make me his son-in-law'. (TBal.)

5 COMPLEX SENTENCES

5.1 Coordinate clauses

The most common coordinate conjunctions are *u*, *aw*, *wa* 'and', *balē*, *walē*, *ammā* 'but', *yā* 'or'. Coordination may also be accomplished by the reduplicated conjunctions (*h*)*am* . . . (*h*)*am* 'both . . . and', *yā* . . . *yā* 'either . . . or', and *na* . . . *na* 'neither . . . nor'.

Examples:

watī čunt rōč-ay āp u nān-ā zurt u
 self.GEN some day-GEN water and bread-OBJ take.PT.3S and
ham-ā šār-ay rāh-ā watī dēmā gipt
 EMPH-DEM town-GEN way-OBJ self.GEN before take.PT.3S
 'he took [enough] bread and water for a few days and set out for that very town'.
 (TBal.)

bādšāh gōn watī laškar-ā bi šahr-ā atk balē wāḡa panč-kuš
 king with self.GEN army-OBL to town-OBL come.PT.3S but mister five-kill.PR
ča āhā pēsar bi watī lōg-ā atk u rast
 from DEM.POBL before to self.GEN house-OBL come.PT.3S and arrive.PT.3S
 'the king came to the town with his army but Mr. Five-slayer had reached his house
 before them'. (WBalNosh., BMC II: 178)

am grēt u am gal-a kurt-ī
 also cry.PT.3S and also joy-V.EL do.PT.3S-ENCL.PRO.3S
 'he both cried and rejoiced'. (TBal.)

5.2 Subordination

There are few descriptions of subordination in Balochi. The most comprehensive one so far is that of Axenov (2006: 243–267) for TBal., on which this section is largely based (examples are mainly from TBal. and other WBal. dialects). As Balochi is primarily a spoken language, the syntax is usually fairly simple. Subordination closely follows the basic pattern of Persian, and several conjunctions have been borrowed from neighbouring languages (e.g. *agar* 'if', *lēkin* 'but').

Subordinate clauses sometimes precede and sometimes follow the matrix clause, into which they can also be embedded. The particle *ki* serves as a general subordinator: it introduces complement and relative clauses as well as quoted speech. It may also introduce adverbial clauses. Other subordinating conjunctions are mainly formed with *ki* as their basis, e.g. *wahdē ki* ‘when’, *pa(r)čē ki* ‘because’, *tān ki* ‘in order that’.

Sometimes a subordinate argument is asyndetically connected to the matrix clause (i.e. without a conjunction):

mūsā [] ta xudā-ayā raw-ay ta allā-ā b-guš ki
 PN you.S God-LOC go.PR-2S you.S God-OBJ SBJ-say.PR CL
am-ē yakk rōč-ē mnī lāp-ā ša nān-ā
 EMPH-DEM one day-INDEF 1.GEN belly-OBJ from bread-OBL
sēr kan
 full do.PR.SBJ
 ‘Moses, [when] you go to God, tell him to fill my belly with food even if only for one day’. (AfBal.)

kass-ē hast [] manā kamm-ēn āp bi-dant
 person-INDEF V.EX.3S 1.OBJ little-ATTR water SBJ-give.PR.3S
 ‘is there anyone [who could] give me a little water?’ (WBal.)

5.2.1 Complement clauses

Complement clauses are normally introduced by the subordinator *ki*. Complements may constitute either the subject or the direct object of the main clause.

tārīk at u zabr mālūm-a na-būt
 dark COP.PT.3S and good evident-V.EL NEG-become.PT.3S
ki mardum-ē yā diga čīz-ē
 CL human-INDEF or other thing-INDEF
 ‘it was dark and it was not quite clear whether it was a man or something else’.
 (TBal., ABT: 247)

na-zān-ay ki manī nām wāḡa pañč-kuš int
 NEG-know.PR-2S CL 1.GEN name mister five-kill.PR COP.PR.3S
 ‘don’t you know that my name is Mr. Five-slayer?’ (WBalNosh., BMC II: 172)

Quoted speech can be considered a subtype of object complement.

ḡanēn-ā gwašt ki nēt gapp čē int
 woman-OBL say.PT.3S CL now talk what COP.PR.3S
 ‘the woman said (that:) what is the matter (lit. talk) now?’ (WBalNosh., BMC II: 175)

Non-finite complements employing infinitives are also common.

mēmān-ay xizmatt kurt-in čō-m-ē asal-ayā wašš int
 guest-GEN service do.PT-INF like-EMPH-DEM honey-LOC good COP.PR.3S
 ‘to pay honour to a guest is sweet like honey (lit. this honey)’. (TBal., ABT: 246)

5.2.2 *Relative clauses*

The subordinator *ki* also introduces relative clauses. Head nouns of restrictive relative clauses are introduced by a demonstrative or indefinite pronoun (functioning as an adjective) and/or marked by the individuation suffix *-ē* (see also Jahani 2008).

If the antecedent has the syntactic function of subject or direct object in the relative clause, it can be resumed either by zero (more common) or by a pronoun (less common). If the antecedent carries another syntactic function, which is very rare in Balochi, there must be a resumptive pronoun in the relative clause.

(Temporal and locational clauses are structurally relative clauses, but are used as adverbial complements and do not require a resumptive pronoun, see 5.2.3.1 and 5.2.3.2).

čīz-ē-rā ki dīn guš-īt wāgīyat dār-īt
 thing-INDEF-OBJ CL religion say.PR-3S truth hold.PR-3S
 ‘what religion says holds true’. (AfBal., Jahani 2008: 147)

ta p-ē kār-ān ki man-a kan-īn baxt na-dār-ay
 you.S for-DEM work-P.OBL CL I-V.EL do.PR-1S luck NEG-hold.PR-2S
 ‘you don’t have any success in the things that I occupy myself with’. (TBal., Jahani 2008: 148)

gis-ē ki mnī balluk bi āi tahā zindagī-a kurt
 house-INDEF CL I.GEN grandmother in DEM.GEN inside life-V.EL do.PT.3S
annūn bēxī wayrān int
 now totally destroyed COP.PR.3S
 ‘the house in which my grandmother lived is now totally in ruins’. (TBal., ABT: 251)

mard-ē ki ča panč mazār-ay kuš-ag-ā bajj na-wārt
 man-INDEF CL from five tiger-GEN kill.PR-INF-OBL fear NEG-eat.PR.3S
ē yakk-ē mazār-ay kuš-ag pa āi
 DEM one-ATTR tiger-GEN kill.PR-INF for DEM.OBL
hičč čī-ē būt na-kan-t
 any thing-INDEF become.PT NEG-do.PR-3S
 ‘to kill this one tiger cannot be any problem for a man who is not afraid of killing five tigers’. (WBalNosh., BMC II: 174–175)

čōš-ē mard-ē bī ki gunāh čī na-k^huθ-a
 such-ATTR man-INDEF become.PR.3S CL sin thing NEG-do.PT-PF
 ‘it should be such a man who has not committed any sin’. (EBal., Dames 1891: Part I:5)

pa-čē am-ā lunka-ē ki āt-ay ki b-ōr-ay
 for-what EMPH-DEM mouthful-INDEF CL come.PT-2S CL SBJ-eat.PR-2S
am-āyrā um padā išt-ay
 EMPH-DEM.OBJ also afterwards leave.PT-2S
 ‘why did you leave the mouthful that you were about to eat?’ (TBal.)

5.2.3 Adverbial clauses

5.2.3.1 Temporal clauses

Temporal clauses are introduced by one of the conjunctions *wahd-ē ki* / *waxt-ē ki*, *(h)ar (ka)dēn ki* ‘when’, *an-čō ki*, *(h)am-ē ki* ‘as soon as’ and *tā(n) ki*, *tā wahd-ē ki* ‘until’, or by the subordinator *ki*, which in this function is placed after the first word or phrase of the clause. In some dialects the main clause is frequently resumed by *ta*, *tō*, *guḏā* ‘then’.

In narratives, the time frame is generally the past, and the preterite is used in the temporal clause. In the non-past temporal field, the indicative denotes certainty and the subjunctive uncertainty. Also, the preterite can be used as a relative tense in non-past temporal clauses with the matrix clause as the deictic centre: it puts the matrix clause in the foreground and the temporal clause is construed as background information.

yakk rōč-ē hardē ki ā bēwass
 one day-INDEF when CL DEM poor.one
ča pōryāt-ā bi gis-ā atk u rast
 from work-OBL to house-OBL come.PT.3S and arrive.PT.3S
ta āi čamm bi panč mušk-ānī sarā kapt
 then DEM.GEN eye to five mouse-P.GEN on fall.PT.3S
 ‘one day when that poor one came home from work his eye fell on five mice’.
 (WBalNosh., BMC II: 172)

jinēnzāg ki ša ḏannā k-ayt
 woman CL from outside V.PREF-come.PR.3S
pādišā kitāb-ā by-āi-a dant
 king book-OBJ to-DEM.OBL-V.EL give.PR.3S
 ‘when the woman comes from outside, the king gives her the book’. (TBal., ABT: 254)

arkadēn ki watī čiz-ānā bā b-kan-īn
 when CL self.GEN thing-P.OBJ sale SBJ-do.PR-1S
pa wat čiz b-zin-īn ra-īn
 for self thing SBJ-buy.PR-1S go.PR-1S
 ‘whenever I sell my things and buy something for myself, I will go’. (TBal., ABT: 254)

waxtē ki ta mardum āwurt-ay am-ē gis-ay tā
 when CL you.S human bring.PT-2S EMPH-DEM house-GEN in
ša mučē-ān dēmā ta wat putr
 from all-POBL before you.S self enter.PR.SBJ
 ‘when you bring people, go into this house, yourself, before everyone [else]’. (TBal., ABT: 255)

5.2.3.2 Locational clauses

In locational clauses, *ki* is preceded by an adverb or a noun indicating place, e.g. *idā* ‘here’, *ōdā* ‘there’, *(h)ar jā* / *(h)ar kuḏā* ‘wherever’. Depending on the prominence of the two clauses and the degree of realis of the locational clause, the preterite, past subjunctive, or non-past indicative or subjunctive may be used in this clause.

arjā ki oštāt, zān-ay
 wherever CL stand.PT.3S know.PR-2S
ki tī brās am-ōdān ant
 CL you.S.GEN brother EMPH-there COP.PR.3P
 ‘know that your brothers are wherever it stops’. (TBal.)

ēšānā āwurt am-ōdā
 DEM.POBJ bring.PT.3S EMPH-there
ki wazīr-ay jīnikk-ā-a šōd-ant
 CL minister-GEN girl-OBJ-V.EL wash.PR-3P
 ‘he brought them to the place where they were washing the minister’s daughter’.
 (TBal., ABT: 257)

5.2.3.3 Clauses of manner

Clauses of manner are linked to the matrix clause by, for example, *čōn ki* ‘like’, and the matrix clause may be resumed by an adverb of manner such as *amā rang* ‘in that way’.

nōkar um čōn ki šā abbās gušt-at-ī
 servant also like CL king PN say.PT-COP.PT.3S-ENCL.PRO.3S
am-ā rang kurt
 EMPH-DEM kind do.PT.3S
 ‘and the servant did it the way that Shah Abbas had told him’. (TBal., ABT: 262)

5.2.3.4 Causal clauses

Causal clauses are most frequently introduced by *pa(r)-čē ki* or *š-am-ā ki* ‘because’. The subordinator *ki* and other conjunctions also have this function. Since the causal clauses express factual circumstances, the verb is in the indicative mood.

bē-šakk tāi nām panč-kuš int
 without-doubt you.S.GEN name five-kill.PR COP.PR.3S
parčē ki ā panč-ē mušk tāi janōk-ē ant
 because CL DEM five-ATTR mouse you.S.GEN slayer-ATTR COP.PR.3P
 ‘undoubtedly your name is Mr. Five-slayer because you have slain those five mice’,
 lit. ‘those five mice are your slain’. (WBalNosh., BMC II: 172)

šamā ki mnī gīs-ā-a k-āy-ay
 because CL I.GEN house-OBL-V.EL V.PREF-come.PR-2S
mnī mēmān am bāz-a bay-ant
 I.GEN guest also much-V.EL become.PR-3P
 ‘I will have many guests because you come to my house’,
 lit. ‘my guests will be many, too’. (TBal., ABT: 259)

mnā yakk rūpī-ay brinj b-day
 I.OBJ one rupee-GEN rice SBJ-give.PR
ki mnī mard-ay dil pa palāv-ā mant-a
 CL I.GEN man-GEN heart for rice.dish-OBL remain.PT-PF.3S
 ‘give me one rupee[’s worth] of rice because my husband’s heart is longing for *pilaw*’.
 (TBal., ABT: 259)

5.2.3.5 Consecutive and final clauses

Consecutive (resultative) and final (purposive) clauses are both introduced by the subordinator *ki*. The only difference between these two types is that the indicative mood is used in consecutive clauses, since they describe real events, whereas final clauses always exhibit the subjunctive mood, owing to the fact that they state a desired purpose.

Result (indicative)

ē dēb-ā yakk zām-ē ĵan-t
DEM demon-OBJ one sword-INDEF hit.PR-3S
ki sarag-ī ādā-a kap-īt
CL head-ENCL.PRO.3S there-V.EL fall.PR-3S

‘he strikes the demon with a sword once so that his head falls far away’. (TBal., ABT: 260)

Purpose (subjunctive)

sakk gō zamzīl-ān bi-band-ī
hard with chain-P.OBL SBJ-tie.PR-ENCL.PRO.3S
ki wat-ā pačč kurt ma-kan-t u ma-rawt
CL self-OBL open do.PT PROH-do.PR-3S and PROH-go.PR.3S

‘tie it tight with chains so that it can not untie itself and get away’. (WBalNosh., BMC II: 176)

A non-finite construction with the infinitive in the oblique case is also common to express purpose.

šut kū-ā āp kašš-ag-ā
go.PT.3S well-OBL water pull.PR-INF-OBL
‘he went to the well to pull up water’. (WBalNosh.)

5.2.3.6 Conditional clauses

Conditional clauses are introduced by the conjunction *aga(r) (ki)*. The matrix clause is sometimes resumed by the particle *tō / ta*.

Hypothetical conditions have either a preterite (relative tense) or a non-past subjunctive form of the verb depending on focus.

age be bāzār rapt-ō čīz-ē ger-ō
if to market go.PT-1S thing-INDEF take.PR-1S
‘I will buy [a little] something if I go to the market’. (IrBal. of Lashar, Mahmoodi Bakhtiari 2003: 138)

agar ki ēla-ī ma-kan-t mnī dil patta-a
if CL free-ENCL.PRO.3S PROH-do.PR-3S I.GEN heart for.you.S-V.EL
suč-īt
burn.PR-3S

‘if he does not set him free, I will feel sorry for you’. (TBal., ABT: 264)

Counterfactual conditions have a past subjunctive form in the conditional clause and an imperfect form in the main clause in dialects with the verbal element *-a* (see 3.2.5.1 (2)). In other dialects the verb form in the main clause is preterite.

aga ta w manī nām wāja pañč-kuš ēr ma-kurt-ēn
 if you.S I.GEN name mister five-kill.PR down PROH-do.PT-SBJ
u manī sitā-ā ma-kurt-ēn
 and I.GEN praise-OBL PROH-do.PT-SBJ
ta par čē marōčī ē balāh manī sarā kapt
 then for what today DEM calamity I.GEN on fall.PT.3S
 ‘if you had not called me Mr. Five-slayer and praised me, why then would this
 calamity have befallen me’. (WBalNosh., BMC II: 175)

5.2.3.7 Concessive clauses

Concessive clauses are marked by *(h)ar-činkas*, *(h)ar-čī*, *(h)ar-čōn* ‘even if, although’ (+ *ki*). The matrix clause is sometimes resumed by *balē* ‘but’. The verb in the concessive clause is found both in the indicative and in the subjunctive (see also 4.5.2.9 for an additional example).

harčinkas ki janēn-ā diljamī dāt
 although CL woman-OBL comforting give.PT
balē wāja pañč-kuš na-mannit
 but mister five-kill.PR NEG-obey.PT.3S
 ‘even though Mr. Five-slayer’s wife comforted [him], he did not calm down’ (i.e. did not listen to her). (WBalNosh., BMC II:177)

6 WORD FORMATION

6.1 Nominal word formation

Common nominal suffixes are *-ag*, *-k*, *-ak(k)*, *-ik*, *-uk* (see Korn 2005a: 163–169), which in many cases do not change the meaning of the word, and are also used interchangeably, e.g. *ādēn* vs. *ādēnk*, *ādēnag*, all ‘mirror’.

Several of these suffixes are also used to derive adjectives from nouns, as are the suffixes *-ūg* and *-īg*. Here as well, there is a certain degree of interchangeability of suffixes, e.g. *šud* ‘hunger’ → *šudag*, *šudīg*, *šudīk* ‘hungry’. However, the most productive adjective suffix is *-ī*, e.g. *nōšī* ‘tasty’ (from *nōš* ‘drink(ing)’), *waxtī* ‘on time, timely’ (from *waxt* ‘time’).

Adjectives derived from personal pronouns with the suffix *-īg* are used in the function of predicative possessives in some dialects, e.g. *manīg* ‘mine’, *taīg* ‘yours’, *āīg* ‘his, hers, its’, *mayg* ‘ours’, *šumayg* ‘yours’, *āwānīg* / *āyānīg* / *āhānīg* ‘theirs’.

The suffixes *-ik(k)*, *-uk* and *-luk* (to a certain extent also *-ak(k)*) (also) have a diminutive function; *-uk* is particularly productive (also on names, cf. Badalkhan 2003: 297).

Another suffix *-ī* derives abstract nouns from nouns and adjectives (e.g. *waš(š)ī* ‘happiness’ from *waš(š)* ‘good, nice’; *duz(z)ī* ‘theft’ from *duz(z)* ‘thief’), and *-ōk* forms agent nouns from the present stem of verbs (see also 3.2.3.4). Both suffixes are productive.

6.2 Verbal word formation

(1) Denominative verbs can be formed from inherited as well as borrowed nominals without adding a suffix, that is, by using the noun as the present stem, e.g. *čōp-* ‘hit’ (from

čōp ‘stick’), *duz(z)*- ‘steal’ (from *duz(z)* ‘thief’), *šōm*- ‘plough’ (from *šōm* ‘ploughing’), *ǰall*- ‘contain’ (from *ǰall* ‘box’), *g(a)rand*- ‘thunder’ (from *g(a)rand* ‘thunder(n.)’), *badl*- ‘change’ (from *badal* ‘change (n.)’).

Very frequently, inherited and borrowed nominals are converted into verbs by combining them with a light verb to give a complex verb (see 3.2.1). Common light verbs are *kan*- ‘do’ and *ǰan*- ‘hit’ for actions with the feature [+ control], and *war*- ‘eat’ (see Farrell 2008: 126–127 for a list of examples), *gir*- ‘take’ and *bēy-lbay*- ‘become’ for [– control].

(2) The formation of causative verbs is productive. Causatives are formed by suffixing *-ēn*- to the present stem; double causatives use the suffix *-āēn*- (see 3.2.2.2).

7 SAMPLE TEXTS

The two sample texts below are intended to illustrate the great variation among the Balochi dialects. Karachi Balochi is located in the very south-east of the Balochi linguistic area and Turkmenistan Balochi in the very north. The first sample text is from Farrell 2003: 207–209, and the second is an excerpt from the text A 1.5 published by Axenov 2006: 291–293 (with glosses added), both with a few minor changes in text and/or translation.

Karachi Balochi: Fishing

matlab iš ē, ǰō ast-ē,
 meaning DEM COP.PR.3S REL V.EX.3S-COP.PR.3S
 ‘It means that, you know’,

idā ǰab dar k-āy-ant-ē,
 here when out V.PREF-come.PR-3P-COP.PR.3S
 ‘when they leave here’,

išānī takrīban tēm bīt-ē dā baǰ-ē tēm-ā
 DEM.P.GEN about time become.PR.3S-COP.PR.3S ten hour-GEN time-OBL
 ‘it will be about ten o’clock for them’.

dā baǰ-ē tēm-ā, ǰō ast-ē
 ten hour-GEN time-OBL REL V.EX.3S-COP.PR.3S
 ‘At ten o’clock, you know’,

ē māhōr-ā dawr dē-ant-ē,
 DEM net-OBL throwing give.PR-3P-COP.PR.3S
 ‘they will throw out the net’.

dā baǰ-ē tēm-ā, na?
 ten hour-GEN time-OBL NEG
 ‘At ten o’clock, right?’

dā-ē tēm-ā māhōr-ā dawr dē-ā.
 ten-GEN time-OBL net-OBL throwing give.PR-3P
 ‘They throw out the net at ten’.

māhōr-ā dawr dē-ā, dāki saē baǰ tak, ǰō ast-ē,
 net-OBL throwing give.PR-3P until three hour until REL V.EX.3S-COP.PR.3S
 ‘They throw out the net until three o’clock, you know’,

ē māhōr-ā gō sarpuṭ bē-ā.
 DEM net-P.OBL with busy become.PR-3P
 'they will be busy with the nets'.

anču rēč-ant išā, padā čār baḡ tak
 so pour.PR-3P DEM.P.OBL then four hour until
 'So they throw them out, and then, until four o'clock',

ē kammuk nind-ā, ārām kan-ant-ē.
 DEM little sit.PR-3P quiet do.PR-3P-COP.PR.3S
 'they sit a bit and rest'.

saḡḡ-ē šap dā . . . māhōr, ḡō ast-ē,
 whole-ATTR night until net REL V.EX.3S-COP.PR.3S
 'All night until . . . the nets, you know',

āp-ē tahā bīt-ē.
 water-GEN inside become.PR.3S-COP.PR.3S
 '[they] are in the water'.

ham-ā sōb-ē sarā takriban pañč baḡ-ē ṭēm-ā,
 EMPH-DEM morning-GEN on about five hour-GEN time-OBL
 'Early that morning at about five o'clock',

pañč-ē ṭēm-ā, na?
 five-GEN time-OBL NEG
 '-- at five o'clock, right?'

nī māhōr-ē kašš-ag-ā šurū kan-ant-ē,
 now net-GEN pull.PR-INF-OBL start do.PR-3P-COP.PR.3S
 'Now they begin to pull in the net'.

dāki šap-ē, matlab rōč-ē bār baḡ tak ē māhōr
 until night-GEN meaning day-GEN twelve hour until DEM net
kašš-ant-ē.
 pull.PR-3P-COP.PR.3S
 'Until twelve at night, I mean day, they pull in the net'.

māhōr-ā kašš-ant-ē maččī-ān alag kan-ā
 net-OBL pull.PR-3P-COP.PR.3S fish-P.OBL separate do.PR-3P
 'They pull in the net [and] separate the fish'.

maččī-ān alag kan-ā, barf-marḡ kuṭṭ-ant-ē, lāč-ē sarā
 fish-P.OBL separate do.PR-3P ice-DBL cut.up.PR-3P-COP.PR.3S boat-GEN on
 'They separate the fish, they cut up ice, on the boat'.

sāḡā bār baḡ, yak baḡ lagg-ī, fārig bē-ā,
 half twelve hour one hour take.PR-3S free become.PR-3P
 'When it turns half past twelve or one o'clock, then they are free [for that day]',

padā warag-marḡ war-ant-ē, lāč-ē sarā
 then food-DBL eat.PR-3P-COP.PR.3S boat-GEN on
 'and eat some food, [right there] on the boat'.

ara ce bād watī mähōr-ā sāiḡ-ā kan-ant-ē,
 DEM.OBL from after self.GEN net-OBL side-OBL do.PR-3P-COP.PR.3S
 ‘After that they [again] put their net to the side’

padā mähōr-ā rēč-ā āp-ā
 then net-OBL pour.PR-3P water-OBL
 ‘and throw the net into the water’.

Turkmenistan Balochi: My Childhood

gwandakkī-ā man watī jīnd-ay gis-ay māl-ānā-a
 childhood-OBL I self.GEN self-GEN house-GEN cattle-P.OBJ-V.EL
čārēnt-un.
 graze.CAUS.PT-1S

‘When I was a child, I used to graze the sheep of our household’.

ammā nazzīk-i sad pas dāšt-an.
 we near-EZ hundred sheep hold.PT-1P
 ‘We had almost one hundred sheep’.

man u mnī brās usmān u sapor wārī-pa-bērī māl-ānā-a
 I and I.GEN brother PN and PN turn-by-turn cattle-P.OBJ-V.EL
čārēnt-an.
 graze.CAUS.PT-1P

‘I and my brothers Usman and Sapor took turns grazing the sheep’.

ammā pa watī māl-ān yakkyakkī nām išt-at-an,
 we for self.GEN cattle-P.OBL one.by.one name leave.PT-COP.PT-1P
 ‘We had given names to each one of our sheep’.

āwānā bāz dōst dāšt-an u zabr drust-a kurt-an.
 DEM.P.OBJ much friend hold.PT-1P and good recognized-V.EL do.PT-1P
 ‘We loved them very much and would recognize [each one] right away’.

gis-ay dapā bāg u pālēz ham kišt-an.
 house-GEN in.front.of garden and melon.field also plant.PT-1P
 ‘We planted a garden and a melon field in front of our house as well’.

man da-dwāzda sālagī-ā gō ar-ā alap-a k-āwurt-un.
 I ten-twelve year.old-OBL with donkey-OBL grass-V.EL V.PREF-bring.PT-1S
 ‘When I was about ten or twelve years old, I used to fetch grass with the donkey’.

ham-ā waxt-ā yakk spēt-ēn ar-ē dāšt-an,
 EMPH-DEM time-OBL one white-ATTR donkey-INDEF hold.PT-1P
 ‘At that time we had a white donkey’.

nām-ī dabbū at.
 name-ENCL.PRO.3S PN COP.PT-3S
 ‘its name was Shorty’.

mnī piss mnā ša panč-šaš sālagī-ā
 I.GEN father I.OBJ from five-six year.old-OBL
 ‘From the age of five or six, my father’

qurān-ay wānt-in-ā ēl dāt.
 Quran-GEN read.PT-INF-OBL learning give.PT.3S
 ‘taught me [*mnā*] to read the Quran’.

man gāvaxt-a na-wānt-un, ša āī dēmā-a ǰist-un.
 I sometimes-V.EL NEG-read.PT-1S from DEM.GEN before-V.EL flee.PT-1S
 ‘Sometimes I did not study, I ran away from him’.

mnī piss mnā arčōn dōst dāšt-ēn ham,
 I.GEN father I.OBJ ‘even.though friend hold.PT-SBJ.3S also
 ‘[And] even though my father loved me very much’,

pa wānt-in-ay xātirā mnā gipt u ǰat.
 for read.PT-INF-GEN sake I.OBJ take.PT.3S and hit.PT.3S
 ‘when it came to studying, he used to take me and hit me’.

man zār-a būt-un, tā čunt rōč
 I angry-V.EL become.PT-1S until some day
 ‘[And then] I got angry, [and] for several days’

gō ičkass gis-ay tā abar-a na-dāt-un.
 with nobody house-GEN in word-V.EL NEG-give.PT-1S
 ‘I did not talk to anyone in the house’.

nūn, waxt-ē ki tū būt-un, sarpad-a bān
 now time-INDEF CL big become.PT-1S aware-V.EL become.PR.1S
 ‘Now that I have grown up, I understand

ki mnī piss zabr kurt-a ki mnā ǰat-a.
 CL I.GEN father good do.PT-PF.3S CL I.OBJ hit.PT-PF.3S
 ‘that my father did a good thing to beat me.’

agar ma-ǰat-ēn, man-a na-wānt-un.
 if PROH-hit.PT-SBJ.3S I-V.EL NEG-read.PT-1S
 ‘if he hadn’t beaten [me], I wouldn’t have studied’.

ACKNOWLEDGEMENTS

The authors wish to thank Tim Farrell and Serge Axenov for permission to use texts they recorded and published, as well as for access to Serge Axenov’s digital corpus on TBal. We also thank Christian Rammer for making the map, Erik Anonby for checking English accuracy, and the Swedish Research Council for providing necessary funding for this checking. Our thanks are also due to Erik Anonby and Gernot Windfuhr for their comments on previous versions of this chapter. Any remaining mistakes are, of course, our own responsibility.

LIST OF ABBREVIATIONS

ABT	Axenov 2006 (TBal.)
AfBal.	Balochi spoken in Afghanistan
ASG	Ahangar 2007 (IrBalSarh.)
Bal.	Balochi
BLB	Buddruss 1988 (AfBal.)
BMC	Barker/Mengal 1969 (WBalNosh.)
BSD	Baranzehi 2003 (IrBalSarw.)
ComBal.	Common Balochi (see 1.4)
EBal.	Eastern Balochi dialect(s)
FBB	Farrell 1990 (SBalKar.)
IrBal.	Balochi spoken in Iran
IrBalSarh.	Iranian Balochi of Sarhadd
IrBalSarw.	Iranian Balochi of Sarawan
SBal.	Southern Balochi dialect(s)
SBalKar.	Southern Balochi of Karachi
SWBal.	Southern and Western Balochi
TBal.	Balochi spoken in Turkmenistan
WBal.	Western Balochi dialect(s)
WBalNosh.	Western Balochi of Noshke, Pakistan

BIBLIOGRAPHY AND REFERENCES

- Ahangar, Abbas Ali (2007) 'A description of the verbal system in Sarhaddi Balochi of Granchin', *Orientalia Suecana* 56: 5–24.
- al Ameer, Saeed Mohammad (2003) 'The Baloch in the Arabian Gulf States', in Jahani and Korn (eds) (2003), pp. 237–243.
- Axenov, Serge (2000) 'Balochi Orthography in Turkmenistan', in Jahani (ed.) (2000), pp. 71–78.
- (2003) 'The Balochi Language in Turkmenistan', in Jahani and Korn (eds) (2003), pp. 245–258.
- (2006) *The Balochi Language of Turkmenistan (Studia Iranica Upsaliensia 10)*, Uppsala: Uppsala University.
- Badalkhan, Sabir (2003) 'Language Contact in Balochistan and Its Impact on Balochi Personal Names', in Jahani and Korn (eds) (2003), pp. 279–311.
- (2008) 'Zikri Dilemmas: Origins, Religious Practices, and Political Constraints', in Jahani, Korn and Titus (eds) (2008), pp. 293–326.
- Baranzehi, Adam Nader (2003) 'The Sarawani Dialect of Balochi and Persian Influence on It', in Jahani and Korn (eds) (2003), pp. 75–111.
- Barker, Muhammad Abd-al-Rahman, and Aqil Khan Mengal (1969) *A Course in Baluchi*, Montreal: McGill University Press, 2 vols.
- Bashir, Elena (1991) *A Contrastive Analysis of Balochi and Urdu*, Peshawar: Directorate of Primary Education, Washington, D.C.: Academy for Educational Development.
- (2008) 'Some transitional features of Eastern Balochi: An areal and diachronic perspective', in Jahani, Korn and Titus (eds) (2008), pp. 45–82.
- Buddruss, Georg (1977) 'Buttern in Balochistan', in Benzing, Brigitta *et al.* (eds) (1977) *Wort und Wirklichkeit. Studien zur Afrikanistik und Orientalistik, Eugen Rapp zum 70. Geburtstag*, vol. II, pp. 1–16.
- (1988) *Aus dem Leben eines jungen Balutschen, von ihm selbst erzählt*, Stuttgart: Steiner.

- Collett, Nigel A. (1986²) *A Grammar, Phrase Book and Vocabulary of Baluchi (as Spoken in the Sultanate of Oman)*, Abingdon.
- Dames, M. Longworth (1891) *A Text Book of the Balochi Language, consisting of Miscellaneous Stories, Legends, Poems, and a Balochi-English Vocabulary*, Lahore.
- (1907) *Popular Poetry of the Baloches*, London: David Nutt, 2 vols.
- (1913) 'Balōčistān', in *Enzyklopaedie des Islam* 1, pp. 650–666.
- Dashtyari, Saba (2003) 'Periodicals in Balochi: A Brief Description of Balochi Printed Media', in Jahani and Korn (eds) (2003), pp. 321–342.
- Elfenbein, Josef (1963) *A Vocabulary of Marw Baluchi (Quaderni della sezione linguistica degli annali 2)*, Naples: Istituto Universitario Orientale.
- (1983) *A Baluchi Miscellany of Erotica and Poetry: Codex Additional 24048 of the British Library*, Naples: Istituto Universitario Orientale.
- (1989a) 'Balōčī', in *CLI*, pp. 350–362.
- (1989b) 'Baluchistan 3: Baluchi language and literature', in *EnIr* 3, pp. 633–644.
- (1990) *An Anthology of Classical and Modern Balochi Literature*, Wiesbaden: Harrassowitz, 2 vols.
- (1991) 'Weakness and Thirst: Unpublished poems in Balochi by Gul Khān Nasīr and 'Atā Sād', *Iran* 29: 109–115.
- (1997) 'Balochi phonology', in Kaye, Alan S. (ed.) (1997) *Phonologies of Asia and Africa*, Winona Lake: Eisenbrauns, pp. 761–776.
- (2000) 'Unofficial and Official Efforts to Promote Balochi in Roman Script', in Jahani (ed.) (2000), pp. 79–88.
- Farrell, Tim (1989) *A Study of Ergativity in Balochi* (unpublished MA thesis, SOAS, London).
- (1990) *Basic Balochi. An introductory course (Baluchistan Monograph Series 1)*, Naples: Istituto Universitario Orientale.
- (1995) 'Fading ergativity? A study of ergativity in Balochi', in Bennett, David C., Bynon, Theodora and Hewitt, B. Georges (eds) (1995) *Subject, Voice and Ergativity: Selected Essays*, London: SOAS, pp. 218–243.
- (2003) 'Linguistic influences on the Balochi spoken in Karachi', in Jahani and Korn (eds) (2003), pp. 169–211.
- (2008) 'The Sweet Tongue: Metaphor in Balochi', in Jahani, Korn and Titus (eds) (2008), pp. 101–138.
- Filippone, Ela (1996) *Spatial Models and Locative Expressions in Balochi (Baluchistan Monograph Series 4)*, Naples: Istituto Universitario Orientale.
- Foxton, William (1985) 'Arabic/Baluchi Bilingualism in Oman', *Newsletter of Baluchistan Studies*, 2, Naples: Istituto Universitario Orientale, pp. 31–39.
- Geiger, Wilhelm (1891) 'Lautlehre des Balūčī', *Abhandlungen der I. Classe der Königlich-Bayerischen Akademie der Wissenschaften* 19/2: 397–464.
- Geiger, Wilhelm, and Kuhn, Ernst (eds) (1895–1901) *Grundriß der iranischen Philologie* 1, Straßburg: Trübner, 2 vols.
- Gilbertson, George (1923) *The Balochi Language, A grammar and manual*, Hertford.
- (1925) *English-Balochi Colloquial Dictionary*, Hertford, 2 vols.
- Gladstone, C. E. (1874) *Biluchi Handbook*, Lahore: Rukn-ud-din.
- Grierson, George A. (1921) 'Balōchī', *LSI* X, pp. 327–451.
- Jahani, Carina (1989) *Standardization and Orthography in the Balochi Language (Studia Iranica Upsaliensia 1)*, Uppsala: Uppsala University.
- (1994) 'Notes on the use of genitive construction versus *izāfa* construction in Iranian Balochi', *StIr* 23: 285–298.
- (1995) 'Notes on how to express 'to have' in Iranian Balochi', *StIr* 24: 125–130.
- (1999) 'Persian influence on some verbal constructions in Iranian Balochi', *StIr* 28: 123–143.

- (ed.) (2000) *Language in Society – Eight Sociolinguistic Essays on Balochi* (*Studia Iranica Upsaliensia* 3), Uppsala: Uppsala University.
- (2001) ‘The Balochi language’, in Garry, Jane and Rubino, Carl (eds) *Facts about the World's Languages*, New York, Dublin: H. W. Wilson, pp. 59–64.
- (2003) ‘The case system in Iranian Balochi in a contact linguistic perspective’, in Jahani and Korn (eds) (2003), pp. 113–132.
- (2005a) ‘The glottal plosive – a phoneme in spoken Modern Persian or not?’, in Csató, Éva A., Isaksson, Bo and Jahani, Carina (eds) (2005) *Linguistic Convergence and Areal Diffusion. Case Studies from Iranian, Semitic and Turkic*, London and New York: Routledge, pp. 79–96.
- (2005b) ‘State control and its impact on language in Balochistan’, in Rabo, Annika and Utas, Bo (eds) (2005) *The Role of the State in West Asia*, Istanbul: Swedish Research Institute, pp. 151–163.
- (2008) ‘Relative clauses in Balochi and the marking of the antecedent’, in Jahani, Korn and Titus (eds) (2008), pp. 139–166.
- Jahani, Carina, and Agnes Korn (eds) (2003) *The Baloch and Their Neighbours: Ethnic and Linguistic Contact in Balochistan in Historical and Modern Times*, Wiesbaden: Reichert.
- Jahani, Carina, Agnes Korn, and Paul Titus (eds) (2008) *The Baloch and Others: Linguistic, historical and socio-political perspectives on pluralism in Balochistan*, Wiesbaden: Reichert.
- Korn, Agnes (2001) ‘Archaismus und Innovation im Verbalsystem des Balochi’, in <http://www.uni-bamberg.de/split/iranistik/@online.htm>; <http://titus.uni-frankfurt.de/personal/agnes/bamberg.pdf>.
- (2003) ‘Balochi and the concept of North-Western Iranian’, in Jahani and Korn (eds), pp. 49–60.
- (2005a) *Towards a Historical Grammar of Balochi. Studies in Balochi Historical Phonology and Vocabulary* (*Beiträge zur Iranistik* 26), Wiesbaden: Reichert.
- (2005b) ‘Das Nominalsystem des Balochi, mitteliranisch betrachtet’, in Schweiger, Günter (ed.) (2005) *Indogermanica: Festschrift Gert Klingenschmitt. Indische, iranische und indogermanische Studien dem verehrten Jubilar dargebracht zu seinem fünfundsechzigsten Geburtstag*, Tübingen: VWT-Verlag, pp. 289–302.
- (2006) ‘Counting sheep and camels in Balochi’, in Bogoljubov, Mixail N. et al. (eds) (2006) *Indoiranskoe iazykoznanie i tipologija iazykovykh situatsii. Sbornik statei k 75-letiiu professora A. L. Griunberga*, St. Petersburg: Nauka, pp. 201–212. <http://titus.uni-frankfurt.de/personal/agnes/numbers.pdf>.
- (2008a) ‘A New Locative Case in Turkmenistan Balochi’, *Iran and the Caucasus* 12: 83–99.
- (2008b) ‘The nominal systems of Balochi: How many grammars?’, in Jahani, Korn and Titus (eds) (2008), pp. 167–194.
- (2008c) ‘Marking of arguments in Balochi ergative and mixed constructions’, in Karimi, Simin, Samiian, Vida and Stilo, Donald (eds) (2008) *Aspects of Iranian Linguistics*, Newcastle: Cambridge Scholars Publishing, pp. 249–276.
- (forthc.) ‘The ergative system in Balochi from a typological perspective’, <http://titus.uni-frankfurt.de/personal/agnes/ergativ.pdf>.
- Lodhi, Abdulaziz Y. (2000) ‘A Note on the Baloch in East Africa’, in Jahani (ed.) (2000), pp. 91–95.
- Mahmoodi Bakhtiari, Behrooz (2003) ‘Notes on the Tense System in Balochi and Standard Persian’, in Jahani and Korn (eds) (2003), pp. 133–145.
- Mockler, M. E. (1877) *A Grammar of the Baloochee Language, as it is spoken in Makrân (ancient Gedrosia), in the Persi-Arabic and Roman characters*, London: Henry S. King & Co.
- Morgenstierne, Georg (1932) ‘Notes on Balochi etymology’, *NTS* 5: 37–53 (= *Irano-Dardica*, pp. 148–164).

- Moshkalo, Vjacheslav V. (1991) 'Beludzhskii iazyk' (The Balochi Language), in *Osnovy I*, pp. 5–90.
- Nawata, Tetsuo (1981) *Baluchi (Asian and African Grammaticals Manuals 17b)*, Tokyo: Tokyo University of Foreign Studies.
- Pierce, E. (1874) 'A description of the Mekranee-Beloochee dialect', *JRAS* (Bombay) 11: 1–98.
- Rastorgueva, Vera S. (1966) 'Beludzhskii iazyk' (The Balochi Language), in *Iazyki narodov SSSR I: Indoevropskie iazyki*, Moscow: Nauka, pp. 323–341.
- Rossi, Adriano (1979) 'Phonemics in Balōchī and modern dialectology', in Gnoli, Gherardo and Rossi, Adriano (eds) (1979) *Iranica*, Naples: Istituto Universitario Orientale, pp. 161–232.
- (1989) 'L'inferenziale in Baluci', in de Fouchécour, Charles-Henri and Gignoux, Philippe (eds) (1989) *Études irano-aryennes offertes à Gilbert Lazard (Studia Iranica Cahier 7)*, pp. 283–291.
- Rzehak, Lutz (1998) 'Das belutschische Epos von Šay Murūd und Āno (Materialien aus dem Nachlaß von Alexander L. Grünberg)', in Kushev, V. V. et al. (eds) (1998) *Strany i narody vostoka XXX: Tsentral'naia Aziia. Vostochnyi Gindukush (Countries and peoples of the East XXX: Central Asia. Eastern Hindukush)*, St. Petersburg: Peterburgskoe Vostokovedenie, pp. 149–186.
- Sayad Hashmi (Sayad Hāšmī) (2000) *Sayad Ganj: The First Balochi Dictionary / Sayad Ganj: Balōchī awwalī bzānt balad*, Karachi.
- Sokolov, Sergej N. (1956) 'Grammaticheskii ocherk iazyka beludzhei Sovetskogo Soiuza' (A Grammatical Description of the Balochi Language of the Soviet Union), in *Trudy Instituta Iazykoznaniiia 6 (Proceedings of the Linguistic Institute)*, Moscow: Nauka, pp. 57–91.
- Sokolova, Valentina P. (1953) 'Beludzhskii iazyk' (The Balochi Language), in Sokolova, Valentina P., *Ocherki po fonetike iranskikh iazykov (Descriptions of Phonetics of Iranian Languages)*, Moscow/Leningrad: Nauka, pp. 7–77.
- Spooner, Brian (1967) 'Notes on the Balochī Spoken in Persian Baluchistan', *Iran* 5: 51–71.
- (1989) 'Baluchistan 1: Geography, history, and ethnography', in *EnIr* 3, pp. 598–632.
- Yusefian (Yūsefiān, Pākzād) (1992) *Goruh-e esmi, fe'li va qeydi-ye zabān-e baluči (guyeš-e lāšāri) (The Noun, Verb and Adverb Classes of the Balochi Language (Lashari Dialect))*, Esfahan (MA thesis, manuscript) 1371 h. š.
- Zarubin, Ivan I. (1932) *Beludzhskie skazki I (Balochi Stories I)*, Leningrad: Nauka.
- (1949) *Beludzhskie skazki II (Balochi Stories II)*, Moscow/Leningrad: Nauka.

CHAPTER TWELVE

PARACHI

Charles M. Kieffer

1 INTRODUCTION

Parachi (Parāčī) is spoken north-east of Kabul in three valleys in the southern foothills of the western Hindukosh by a total of approximately 3,500 speakers (as of 1981): in the Shotol valley, north of Golbahār, with approx. 1,500 speakers; in the Pacheghān branch of the Nejraw valley, north-east of Golbahār, with approx. 1,200 speakers; and in the Ghochūlān branch of the same valley, with approx. 700 speakers (see Map 14a.1). These represent remnant linguistic pockets of an earlier dialect continuum (cf. Morgenstierne 1970: 350–2; Kieffer 1979a: 45–47; 1983: 512).

The earliest references to the Parachi date to the sixteenth century. One is by the emperor Babur (Beveridge 1922: 207) in his memoirs, where they are mentioned in a list of tribes in the region of Kabul. The other is by the Ottoman admiral Ali Reīs (1899: 65) in his travelogue to the eastern regions. They are next mentioned in the nineteenth century by the travelers Elphinstone (1815: 316) and Masson (1842 I, 219: 221–22). The earliest reference to their specific location, though without mention of their name, also dates to the sixteenth century, and is found in an eyewitness report on an Islamic jihad in Kafiristan by Qadi Muhammad Salim (Scarcia 1965: 73), who cites “the pagan populations of Nejraw and Pacheghān”.

Georg Morgenstierne is the true discoverer of Parachi. Having collected extensive materials in 1924, he established the dialectological position of Parachi, together with Ormuṛi, as a South-East Iranian group which is historically closely related to North-West Iranian (1932: 18–39), and three years later published his masterly grammatical and philological description with extensive texts and glossary (1929 [1973]; dialect position: 8–17), which was followed some 50 years later by a succinct summary (Morgenstierne 1983). E. Benveniste collected materials in 1947 in the valleys of Shotol and Nejraw, which remain unpublished (Kohzad 1954; 1957). Charles Kieffer conducted field research in all Parachi-speaking locations between 1961 and 1973, and in 1978; in 1981, he worked with refugees in Kabul (1977a; 1979a, 1980a; 1977b, 1977c, 1978, 1979b, 1980b, 1981; 1983; 1986; 1989). Four questionnaires collected by him in 1965 for the Linguistic Atlas of Afghanistan are in the archives of the latter (Kieffer 1989: 445).

Kieffer established that Parachi is still spoken in the same separate locations as in 1924, but is confined to use among members of extended families, who are bi- and multilingual, and even individually show considerable variation. Nevertheless, he found relatively few substantial changes in phonology and grammar between 1924 and today, except for an increase in Persian patterns as could be expected. It should also be noted that the medium of data collection then and now was the Persian of Afghanistan.

2 PHONOLOGY

2.1 Inventory and distribution

2.1.1 Vowels

The system of vowel phonemes has not been established with certainty, and may be shown as follows:

TABLE 12.1: PARACHI VOWELS

Qualitative Characteristics			Quantitative Characteristics					
	Front	Back						
High	<i>ī</i>	<i>ū</i>	Long	<i>ī</i>	<i>ē</i>	<i>ā</i>	<i>ō</i>	<i>ū</i>
Mid close	<i>e</i>	<i>o</i>	Short	<i>e</i>	<i>a</i>		<i>o</i>	
Mid open	<i>ē</i>	<i>ō</i>						
Low	<i>a</i>	<i>ā</i>						

The long back rounded vowels tend to be fronted: *ū* > *ū̄*, and in Shotol particularly long *ō* > *ō̄* [ø] in most instances. Low back *ā* tends to be rounded, and raised pre-nasally. Mid front *ē* is tense and tends to be raised.

The long vowels are fairly stable, but tend to be shortened in unstressed position. Thus, the unstressed mid vowels *ē* *ō* tend to merge with *e* *o*, and *ī* > *i*, *ū* > *u*.

The short vowels show considerable variation: *e* *o* vary between high and mid-high position, [*e* ~ *i*, *o* ~ *u*]; *e* is reduced to [ə] in unstressed position; *a* is often fronted and raised to [ä].

Length distinction appears to be upheld only in Persian loans. While there is strong systemic interference from Kaboli Persian, the system of Parachi appears to be more determined by vocalic quality than length, and thus tends toward a system similar to Pashto, except for phonemic ə in the latter.

Note: For practical purposes, in the following discussion vowels will be represented in a broad transcription, and umlaut will only rarely be indicated.

2.1.1.1 Diphthongs

Diphthongs do not appear to have phonemic status, and are best analyzed as vowel + semivowel. They tend to be reduced particularly in rapid speech, thus final *aw* > *a*, and final *ow*, *ōw* > *ō*: personal enclitics 2s *-aw* ~ *-a*, 2p *-ōw* ~ *-ō*. Medial *ay* and *aw*, *āw*, which only occur in loans, are mostly reduced to *e* ~ *ī* and *ō*, respectively. The causative infix *ēw* + vowel often contracts to *e*: *dal-ēw-em* ~ *dal-e-m* 'that I tear'.

2.1.2 Consonants

Like its neighbor languages, Parachi has developed a series of retroflex phonemes. What distinguishes Parachi most is the frequency of aspiration, which is confined to initial and medial position. It is found in loans from Indian: *phōr* 'fruit', *dhār* 'hill', but also originates in Iranian word-initial voiceless fricatives, mostly *x* > *kh*: *khān-* 'to laugh', *khar* 'donkey'. It is prominently the result of the metathesis of earlier medial *h*: *čhē* 'to

go' < *čēh- < OIr. *čiyuta-, but may also be quite recent: *mhetar* 'groom' < Persian *mehtar*. There is also secondary aspiration. Even though the phonemic status of the aspirates remains to be determined, they are listed in the following chart. Two of the phonemes, *q* and *y*, are confined to loanwords, where *q* alternates with *k*.

TABLE 12.2: CONSONANTS

		Labial	Dental	Retro.	Palatal	Velar	Uv.	Glott.
Stops	voiceless	<i>p</i>	<i>t</i>	<i>ʈ</i>	<i>č</i>	<i>k</i>	<i>(q)</i>	
	aspirate	<i>ph</i>	<i>th</i>	<i>ʈh</i>	<i>čh</i>	<i>kh</i>		
Stops	voiced	<i>b</i>	<i>d</i>	<i>ɖ</i>	<i>ǰ</i>	<i>g</i>		
	aspirate	<i>bh</i>	<i>dh</i>		<i>ǰh</i>	<i>gh</i>		
Fricatives, voiceless		<i>f</i>	<i>s</i>		<i>š</i>	<i>x</i>		<i>h</i>
	aspirate		<i>sh</i>					
	voiced	<i>(v)</i>	<i>z</i>		<i>ž</i>	<i>ɣ</i>		
Nasals		<i>m</i>	<i>n</i>	<i>ɳ</i>		<i>(ŋ)</i>		
	aspirate	<i>mh</i>	<i>nh</i>					
Trills			<i>r</i>	<i>ʀ</i>				
	aspirate		<i>rh</i>					
Laterals			<i>l</i>					
	aspirate		<i>lh</i>					
Semivowels		<i>w</i>			<i>(y)</i>			

2.2 Non-segmental features, stress

The accent in Parachi is one of intensity, and is free to the extent that it may fall on any one of the last three syllables of a word. This system is obscured by two main factors, coupled with the different chronological levels: One is the large loan component from the Kohestani Persian koine, with its different system of stress, in particular final stress in nominals, and the liberal switch of these bilinguals between the two languages, and thus their stress patterns. The other is that fact that during the elicitation of materials, researchers tend to pay less attention to suprasegmental aspects. For practical purposes, in the following accent is indicated by the underlining of the respective vowel.

Stress alternation is particularly evident with nominal inflectional endings, which appear to have been unstressed. Instructive examples are the plural marker and the singular ablative ending *-ī*: The unstressed general plural ending *-ān* alternates with the stressed ending *-ā̄n*, on the model of the Persian plural ending *-ān*. The optional stress on ablative *-ī* is most likely due to the stressed *-ī* in Persian adverbial noun phrases (*be zīd-ī* 'in quickness, quick', *az khoshhāl-ī* 'from ~ out of happiness').

In general, it appears that word stress is superseded by a yet unexplored indigenous pattern of rhythmic, if not prosodic, sequences of stressed and unstressed syllables.

2.3 Morphophonemic changes

Like other Iranian languages, Parachi inherited the morphophonological changes between present and past stems of many verbs. Singular-plural ablaut is found in remnants. The only regular morphophonemic change is that of *a* > *ō* in closed syllables. It is

most prominent in the 3s forms of the past tense: *nhašt-em* ‘I sat down’ vs. *nhōšt* ‘he, she sat down’.

3 MORPHOLOGY

3.1 Nominal morphology

3.1.1 Nouns

3.1.1.1 Gender

There is no morphological distinction of gender. As elsewhere, natural gender may be expressed lexically by different etyma, and by classifiers such as *nar* ‘male’, *māda* ‘female’ for animals, either as a regular N₂-N₁ compound, or *ezāfe* construction: *nar-gu* vs. *māda-gu*, *māda-γasō* or *γasō-e māda* (*gu* ‘cow, bull’, *γasō* ‘calf’).

3.1.1.2 Number

The plural marker is *-ān* (*ā* strongly rounded, [â]) for both animates and inanimates. It is mostly stressed (after Pers. *-ān*), but may alternate with unstressed *-ān*, which appears to be the original pattern: *γos-ān* ‘houses’, but *ketāb-ān* ‘books’, *biyār-ān* ‘brothers’ (sg. *biyār*). The rare ending *-a* occurs only after numerals, including indefinites, and is probably a numerative (cf. Sogdian): *hodī biyār-a* ‘both brothers’. Paired nouns may inflect as a group: *dūst o rafiq-ān* ‘friends and comrades’.

The marker *-hā* (unstressed) is found in the elliptic duals *bāw-ehā* ‘father and son’, *māw-ehā* ‘mother and daughter’.

Ablaut plural is found in two loans: sg. *phōr*, pl. *phar* ‘fruit, grain’ (cf. Pashai *phāl*), also itemizer (Pers. *dāne*, *tā*), and the progressive participial suffix sg. *-tōn*, pl. *tan* (cf. Pashai *-tō*).

3.1.1.3 Declension

There are three inflectional cases in singular and plural: direct (DIR), oblique (OBL), and ablative (ABL). The singular oblique has the variants *-eka*, *-ikī*, but *-ān* for personal names (and pronouns, cf. below). The oblique *-āna* and ablative *-ān* are rarely distinguished. For the numerative *-a* see section 3.1.6 Numerals, specifically 3.1.6.5a.

TABLE 12.3: NOMINAL DECLENSION

	Singular	Plural
DIR	<i>ōsp</i>	<i>ōsp-ān</i>
OBL	<i>ōsp-eka</i> , <i>Alī-ān</i>	<i>ōsp-āna</i> ~ <i>ōsp-ān</i>
ABL	<i>ōsp-ī</i>	<i>ōsp-ān</i>
NUM		<i>ōsp-a</i>
Adpositional Cases (singular)		
ACC, spec.	<i>ma ōsp</i>	
DAT	<i>ōsp-kon</i>	
LOC/DIR	<i>ōsp-tar</i>	
INST/COM	<i>ōsp-pen</i>	

Other case relations are expressed by unstressed adpositions. These include prominently (1) the preposition *ma* for specific direct object (ACC), (2) the postpositions *kon* for dative (DAT), *tar* for locative-directive (LOC/DIR), and (3) *pen* for instrumental-comitative (INST/COM). The basic paradigms are shown above, with *ōsp* ‘horse’.

The vocative may be indicated by particles such as *ai*, *wō* preceding the direct case.

3.1.2 Adjectives

Adjectives are morphologically unmarked, unless functioning as a noun. There is no comparative or superlative marker: *biyā-e ghaṇḍ* ‘elder brother’. The referent is marked by *-tar* ‘from, than’, and *kol-tar* ‘than all’, respectively: *kaštē-tar ghaṇḍ-a* ‘he is bigger than the girl’. Note *godar-* ‘to pass, surpass’ in *ān [tō]-tar ham godar-em*, [*Māmad Hanīfa*]-*tar na* ‘I (*ān*) surpass even (*ham*) thee, but not M.H.’.

3.1.3 Pronouns and deixis

3.1.3.1 Personal pronouns

The independent personal pronouns distinguish direct and oblique case, and the direct object. The personal enclitics of the 1p, 3p have mostly coalesced. In synopsis, the inflectional forms of personal and demonstratives pronouns are as follows:

TABLE 12.4: PERSONAL PRONOUNS

	1s	2s	3s	1p	2p	3p
DIR	<i>ān</i>	<i>tū</i> (<i>tō</i>)	<i>ō</i>	<i>mā</i>	<i>wā</i>	<i>ō-ān</i>
OBL	<i>mon</i> , <i>mun</i>	<i>t-ō</i> (<i>tū</i>)	<i>ōd-ē</i>	<i>mā</i>	<i>wā</i>	<i>od-ān-ān</i>
DAT	<i>mon kon</i>	<i>t-ō kon</i>	<i>ōd-ē kon</i>	<i>mā kon</i>	<i>wā kon</i>	<i>od-ān-ān kon</i>
ACC	<i>ma mun</i>	<i>ma tō</i>	<i>m-ond-ē</i>	<i>ma mā</i>	<i>ma wā</i>	<i>m-ond-ān-ān</i>
POSS	<i>man-ān</i>	<i>t-ān</i>	<i>od-ān</i>	<i>māk-ān</i>	<i>wāk-ān</i>	<i>od-ān-ān</i>
ENCLITIC	<i>-(o)m</i>	<i>-(w)a(w)</i>	<i>-(w)-ē</i>	<i>-(w-)an</i> ~ <i>V-n</i>	<i>-(w)ō(w)</i> ~ <i>V-wlu</i>	<i>-(w-)an*</i> ~ <i>V-n</i>

* *-(w)-* = postvocalic glide

3.1.3.1a Exclusive forms

A singular pronoun may combine with another to form quasi-exclusives: *dāl [mon o tō]* ‘with us’ (we and thee); [*mā o tū*] *par-a-man* ‘let us go’ (we and thou).

3.1.3.2 Demonstrative pronouns

The demonstratives show binary deixis. Of the two, the far member has largely become the independent 3s pronoun, while the near member has become the general, deictically neutral referent, often with article-like function in noun phrases (see examples in section 8 Sample text). A typologically prominent feature is various degrees of focus, which may re-combine. The basic deictics and two of the combinatory ones are shown here.

TABLE 12.5: DEMONSTRATIVE PRONOUNS

Demonstrative and Deictic Patterns						
	Independent		Attributive		Location-Direction	
	far	near	far	near	far	near
Unmarked	<i>ō</i>	<i>ē</i>	<i>hō</i>	<i>hē</i>	<i>ōk</i>	<i>ēk</i>
Selective	<i>hōwyak</i>	<i>hēwyak</i>	<i>hōwī</i>	<i>hawī</i>	<i>ōnhāk</i>	<i>ēnhāk</i>
Emphatic	<i>ok-wīyak</i>	<i>ek-wīyak</i>	<i>ōk-ō</i>	<i>ēk-ē</i>	<i>ōk-enhāk</i>	<i>ēk-enhāk</i>
Note also <i>pū</i> , <i>pī</i> ‘that, this side’.						
Independent Demonstrative Inflection						
	‘that’	‘this’	‘those’	‘these’		
DIR	<i>ō</i>	<i>ē</i>	<i>ō-ān</i>	<i>ē-ān</i>		
OBL	<i>ōd-ē</i>	<i>ēd-ē</i>	<i>od-ān-ān</i>	<i>ed-ān-ān</i>		
DAT	<i>ōd-ē kon</i>	<i>ēd-ē kon</i>	<i>od-ān-ān kon</i>	<i>ed-ān-ān kon</i>		
ACC	<i>m-ond-ē</i>	<i>m-end-ē</i>	<i>m-ond-ān-ān</i>	<i>m-end-ān-ān</i>		
POSS	<i>od-ān</i>	<i>ed-ān</i>	<i>od-ān-ān</i>	<i>ed-ān-ān</i>		
Demonstrative Inflection, Selective Focus						
	‘that same’	‘this same’	‘those same’	‘these same’		
DIR	<i>hōwyak</i>	<i>hēwyak</i>	<i>hōwyakān</i>	<i>hēwyakān</i>		
OBL	<i>odhēk</i>	<i>edhēk</i>	<i>odhēkān</i>	<i>edhēkān</i>		
DAT	<i>odhēk kon</i>	<i>edhēk kon</i>	<i>odhēkān kon</i>	<i>edhēkān kon</i>		
ACC	<i>m-ondhēk</i>	<i>m-endhēk</i>	<i>m-ondhēkān</i>	<i>m-endhēkān</i>		
POSS	<i>odhēk(-ān)</i>	<i>edhēk(-ān)</i>	<i>odhēkān(-ā)</i>	<i>edhēkān(-ā)</i>		

3.1.3.3 Reflexive pronoun

The reflexive pronoun is *xō* ‘self’, obl. *xō-k-ān*; emphatic form: *xōxo*.

3.1.3.4 Interrogative-indefinite pronouns and related adverbials

Interrogatives mostly function also as indefinites, either by themselves or in combination with others:

- pronoun: animate, *kī*, obl. *k-ān* ‘who’ (sg. and pl.); *ka* + copula; inanimate, *čī*, *če* + N ‘what’;
- selective: *khāin* (+ N), obl. *-eka* ‘which one, some one’; *khāin-zā* ‘which thing, something’ (*zā* ‘thing’); *khān-ja-ī* ‘from what place, from some place’;
- location: *kū* ‘where’; *kučend* ‘from where, somewhere’; *kā-wanō* ‘where to, somewhere to’;
- time: *kabī* ‘when’;
- manner & measure: *čā* ‘how, why’; *čā(war)* + N ‘how many’; *čeqa* ‘how much’ (see also 3.1.6 Numerals);
- kind: *zanēng* ‘what kind’; *zanēng-ī* ‘how, what manner’.
- inclusive: *har* ‘every, each’ + interrogative: *har kī* ‘every one’; *har khān* ‘every which one’; *har kū* ‘everywhere’; *har kābī* ‘anytime’;
- exclusive: *hēč* ‘any’ + interrogative-indefinite: *hēč kābī* ‘anywhere’.

3.1.4 Adpositions

3.1.4.1 Prepositions

The preposition *ma* marks (1) the specific direct object, and specific location, time and goal (cf. Ormuri *ku-*): *ma dur* ‘at the river’; *ma dōwās* ‘at 12 o’clock’; *ma [žū qāter] suwār nhōšt* ‘he sat down on/mounted [a mule]’. It also marks (2) the recipient, or dative, with ‘to give’ and ‘to say’ (similar to *kon*): [*ma gu-γ-ōn*] *āo dhay-m* ‘I shall give water [to the cows (*gu-*)]’.

Other prepositions include: *men* ‘in, among’, *men yanir* ‘in the middle of the terrace’, *men āwγān* ‘among (the) Afghans’; *dāl* ‘in front of, before, with’: *dāl mon* ‘in front of me, with me’ (Kb. *pēš-e man*).

3.1.4.2 Ablative -ī and postpositions

Ablative *-ī*, direction ‘from’ (singular): *mhawr-ī* ‘out of the garden’; *qala-ī narγ-ō* ‘he left the farm’; *kābol-ī āγ-ō* ‘he has come from Kabul’; *xōm-ī čīd-ī* ‘he woke up from sleep’; *hawī waxt-ī* ‘from that time on’.

kon, direction ‘to, towards; till’: *sāt-kon* ‘to the village’, *žā sar-kon* ‘till next year’; *tū M-kon āšeq čh-ē* ‘thou becamest a lover to/of M.’; [*mon-kun*]-*e yār-aw* ‘for me, your friend’ (*ezāfe*); *gu-γ-ōn-kun-ē dā* ‘he (-ē.AG) gave to the cows’.

tar, bi-directional general location and direction, ‘in, to’ and ‘from’: *tar-ē* ‘before him’; *dārū-tar hos moγ* ‘all (*hos*) died from the medicine (*dārū*)’.

pen, instrumental, comitative: DIR, *rafīq-pen-ē* ‘with his (-ē) friend’; *pašō-pen* ‘with the axe’; OBL *-ān* of person: [*Māmad Hanīf*] *-ān-pen* ‘with M. H.’.

The locative adpositions frequently combine with each other, and may form circumpositions (for prepositional adverbs with *ezāfe*, see 4.2 Noun phrase structure): [*men*] *xō [-tar]* ‘in, among themselves’; [*ma*] *γōs [-ī] čhī* ‘he left/went from, the house’; [*dāl mā[-ī] čhī*] ‘he went away from (before) me’ (often in allegro form: *dāmāi* ‘from us’); *ē [sōr ōsp]-ī čhar-ī [dharma]-tar* ‘she fell from the top of the horse to the ground’; [*dōs sar]-tar-ī* ‘ten years ago, earlier’ (= Kb. *pēš az da sāl*).

3.1.5 Adverbs

Adverbs are not morphologically marked, except for focus *-k*: *bete* ‘again’, *badē* ‘now’, *nī* ‘now, today’, *nī-hēk* ‘right now’, *nī xawān* ‘tonight’; *aze* ‘yesterday’, *sabā* ‘tomorrow’, *āsoγ* ‘this year’; *wakhē*, *wačhanē* ‘up’; *pastō* ‘down’ (for further locative and temporal adverbs, cf. pronouns above).

3.1.6 Numerals

3.1.6.1 Cardinal numbers

The decades 60–90 are based on the vigesimal system, *γōšt-ak* ‘unit of twenty’; note Indic *lal*:

TABLE 12.6: NUMERALS

1–9	11–19	10–90	100+	
<i>žū</i>	<i>žūwōs</i>	<i>dōs</i>	<i>sō</i>	hundred
<i>dī</i>	<i>dowās</i>	<i>γōšt</i>	<i>hazār</i>	thousand
<i>šī</i>	<i>šidōs</i>	<i>šos</i>	<i>lak</i>	hundred-thousand
<i>čōr</i>	<i>čadōs</i>	<i>čhel</i>	<i>dōs lak</i>	ten lak = one million
<i>pōnč</i>	<i>paes</i>	<i>penjā</i>		
<i>xī</i>	<i>xodōs</i>	<i>šīγōštak</i>		
<i>hōt</i>	<i>hatōs</i>	<i>šī nīm γōštak</i>		
<i>ōšt</i>	<i>aštōs</i>	<i>čōr γōštak</i>		
<i>nō</i>	<i>nams</i>	<i>čōr nīm γōštak</i>		

Digits follow the decades: *γōšt o žū* ‘21’; hundreds have *-sōy*, or Kaboli *-sat*: *dī-sōy*, *čōr-sōy*; *dōs-sōy* ‘ten-hundred’ = ‘one thousand’, or *dī-sat* ‘200’, *šī sat* ‘300’, etc.

3.1.6.2 Ordinal numbers

The ordinal numbers include: *avvalī* ‘first’, *āxer* ‘last’; others have the suffix *-om-ī*: *šī-om-ī* ‘third’; substantive *-gī*: *čōr-om-gī* ‘the fourth one’.

3.1.6.3 Distribution

Distribution is indicated by repetition: *žū žū* ‘by one, one each’.

3.1.6.4 Other numerals and measure

These include (1) preposed: *hodī*, pl. *hodī-n-ān* ‘both’; *hos*, pl. *hosi-n-ān* ‘all’; *kol* ‘all’; *γalaba* ‘many, very’; *γolū* ‘many, very’; *čāvar*, *čā* ‘several’; and (2) postpositional *dal* ‘a crowd, many’; and others. Examples: *kol-e wā-k-ān* ‘all of you’, *čāvar mazdur* ‘several servants’, *γalaba sāl* ‘many years’; *mā dal* ‘we all’.

3.1.6.5 Numerals and case marking

3.1.6.5a Numerative *-a*

Numerals larger than one may take what appears to be a numerative ending *-a*: *γalaba biyār-a* ‘many brothers’, *dī bāl-a bēn* ‘there were two boys’ (sg. *bāl-ō*).

3.1.6.5b Genitive of measure

The partitive genitive is used with measure: *γōšt sīr bīz-eka* ‘20 sirs of grain’; *hazār nafar-eka* ‘one thousand persons’; *hos-e seb-eka* ‘all of the apples’.

3.1.6.5c Plural of specificity

Numerals implying specificity as a rule take the plural: *hodī puš-ān* ‘both sons’; *hawī γulū* *ōsp-ān* ‘these many horses’; *hos dōš-ān* ‘all the hair’, *[hosi-n-ān]-an-a* ‘it is (-a) all (pl.) of theirs (-an)’.

3.2 Verb morphology

3.2.1 Stem formation

The verbal system is based on two stems, present and past. Past stems originate in a dental (except for two verbs: *pēc-lp(h)ak-* ‘to cook’, *?lhušk-* ‘to dry’), which may be retained, or lost (examples are cited as: present/past).

Regular verbs (‘weak’) have past stems in *-ī* (< **-īd*): *berkh-lberkhī-* ‘to fear’.

Irregular (‘strong’) verbs mostly show morphophonemic change. The dental may be retained, contracted to *r* < **urd*, or post-vocally lost: (a) *pān-lpānt-* ‘to know’, *bež-lbast-* ‘to bind’, *nhīn-lnhašt-* ‘to sit down’; (b) *kan-lkor-* ‘to do’ (< **kurd* < **krtá-*); (c) *ošt-l oštā-* ‘to stand (up)’ (cf. Pers. *īstād*); *dah-ldā(-w)-* ‘to give’ (cf. Persian *dād*) vs. regular *deh-l dehī-* ‘to hit’. Note that short *a* is rounded before clusters in pause: *nhašt-am* ‘I sat down’, but *nhōšt* ‘he sat down’.

Suppletive stems are *žē-lāya-* ‘to come’ (cf. Pers. *ā-lāmad*), while *čh-lčhī-* ‘to go, to become’ takes the present stem of *par-a-* ‘to go, walk, wander’ as a suppletive in the present-future.

Causative stems are regular, and marked by *-ēw/-ēw-ī-*: *berkh-ēw-lberkh-ēw-ī-* ‘to frighten’; in context: pres. *ān-e.SBJ dal-ēw-em* ‘that I (*ān*) tear’, past *mun dal-ēw-ē* ‘I (*mun*.AG) tore’.

3.2.2 Nominal forms

There are a good number of nominal forms (for their individual use, see discussions below):

3.2.2.1 Forms based on the present stem

- (1) The progressive participle in sg. *-tōn*, pl. *-tan* (cf. Pashai *-tō*): *kan-tōn* ‘be doing’.
- (2) The verbal noun, or infinitive, in *-en*, only with *nar-* ‘to be able’: *kan-en* ‘doing, to do’.
- (3) The passive participle in *-en* in the passive construction (cf. Pashai *-en*): *kan-en* ‘be done’.
- (4) The present conjunctive participle in *-en*, *-īn*, mainly with verbs of movement (cf. Pashai *-mana*).
- (5) The absolutive in *-amān*.
- (6) The attributive gerundive in *-ānē*: *čhem-ānē* ‘to be done’ (*čhēm-lčhemī-* ‘to go, move, wander, become’).
- (7) The deverbal noun in *-ō*: *mer-ō* ‘death’ (*mer-lmur-* ‘to die’).

3.2.2.2 Forms based on the past stem

- (1) The perfect participle in *-ō*: *nhašt-ō* ‘having sat down’ = ‘seated’.
- (2) The verbal noun in *-ō*: *rhīnt-ō* ‘weeping’ (*ruh-lrhīnt-* ‘to weep’).
- (3) The perfect adjective in *-ō-ī*.

3.2.3 Person marking and ‘to be’

The three persons in the singular and plural are indicated by distinct endings. The endings of the 2s, 3s show some irregularity, as shown in the following synoptic table:

TABLE 12.7: PERSON MARKING

Existential Verb						
	1s	2s	3s	1p	2p	3p
PR	(h)ē-m	(h)ē	(h)ā	(h)ē-man	(h)ē-r	(h)ē-n
PT	(h)ast-am	(h)ast-a	(h)ōst	(h)ast-aman	(h)ast-ar	(h)ast-an
PR inan.			sī			sī (cf. PT)
PT inan.			sō bō-n			sē bē-n (cf. PP)
Copula and 'to be'						
	1s	2s	3s	1p	2p	3p
PR	-eml-ēm	-ī	-al-ā	-eman	-ēr	-an
PT	bē-m	bē	bī-n	bē-man	bē-r	bē-n
SBJ	bi-m	b-ī	b-ē	bi-man	b ^r -ōr	b-ēn
Personal Endings, intransitive						
	1s	2s	3s	1p	2p	3p
PR	-em	-ē	-a	-eman	-ēr	-an
PT	-T-am	-T-ē	-T	-T-eman	-T-ēr	-T-an
PP	-T bē-m	-T bē	-T bō-n	-T bē-man	-T bē-r	-T bē-n
PT	-ē-m	-ē	-ī	-ē-man	-ē-r	-ē-n
PP	-ē bē-m	-ē bē	-ō bō-n	-ē bē-man	-ē bē-r	-ē bē-n

The imperative 2s is generally the bare present stem, while the ending of the 2p is *-ōr*: *da(h), dāh-ōr* 'give (2s, 2p)!'.¹

3.2.4 Negation

There is only one general marker *na*: *na xoṛ-om* 'I did not eat'; *na hēm* 'am not', *nā* 'is not'. The negation usually precedes the verb form immediately, but the copula and the personal ending may be attached to it: *na-m boč-etōn* 'I am not seeing (it)', *na-m narī* 'I could not'. Persian prohibitive *ma* is rarely used. Note that apprehension, such as 'to fear', requires the negation:

ān-em berkhi-tōn-em, če mēri-en [na] par-ī
 I-am fearing-am that killed not he go
 'I am afraid, that he may be killed';

ḡgar-om xūn koṛ, [na]-ī če xešem koṛ-ō b-ē
 liver-my blood made not-he.AG that anger he have done-SUBJ
 'I am distressed, that he may have gotten angry',

3.2.5 System of tenses and moods

The basic system is defined by the opposition between present and past tense, and continuous (imperfective) and perfect (perfective) aspect. The perfect also distinguishes temporary state. There are corresponding non-indicative moods. Transitive past tenses have the ergative construction.

3.2.5.1 General present and past

The simple present and past forms are based on the bare present and past stems. The subject is either marked by the personal ending, or, in the ergative transitive past, by the agent enclitic:

<p>Intransitive pres., <i>ān nhīn[-em]</i> ‘I (shall) sit down’ past, <i>nhašt[-em]</i> ‘I sat down’;</p>	<p>Transitive pres., <i>kan[-em]</i> ‘I (shall) do’ past, <i>koṛ[-om], [-om] koṛ</i> ‘I did’.</p>
---	---

3.2.5.1a Past immediacy marker

In familiar speech, the enclitic *-(n)a*, attached to the past of the 3s appears to indicate immediacy: *āy-a-na* ‘he arrived, there he is’ (cf. Kb. *āmad-ak* ‘the “guy” is here’).

3.2.5.1b Continuous present and past

The continuous, i.e. ongoing or habitual, present and past is a participial construction. It is marked by the formant sg. *-tōn*, pl. *-tan*, added to the present stem. The copula marking the subject may be attached to the participle, or precede it:

<i>kan-tōn[-em]</i> or <i>ān[-em] kan-tōn</i>	‘I am doing’;
<i>ān fārsi [hast-am] xān-tōn</i>	‘I was studying, used to study, Persian’.

While the continuous present expresses ongoing and habitual action, the general present may express: (1) unspecific present and future action, including suggestions and intention, as well as (2) dependent potential action, i.e. the subjunctive.

The unmarked past expresses the completion of an action, which may include repeated action seen as a whole. A distinctive use of this form is to express completion or anticipated certainty of an action in dependent clauses, most conspicuously in conditional clauses (see section 5.5.2).

3.2.5.1c Temporary state present and past

State is expressed by the perfect participle with the copula in the present, and *hasta-* in the past. It is mostly found with change-of-state verbs: *nhašt-ō (h)ēn*, *nhašt-ō (h)asta-n* ‘they are, were sitting’ (‘to sit (down)’). The distinction between temporary state and the perfect forms is clearly seen in the past perfect *čh[-ē] bē-n* ‘they have sat down’ vs. *čh[-ō] bē-n* ‘they were seated’.

3.2.5.2 Present perfect and past perfect

The forms of the present and past perfect are based on the perfect participle and the present and past forms of the copula with intransitive verbs, and the agent enclitic with transitive verbs:

Intransitive,

perfect, *ō āy[-ō]* ‘he has arrived’, *ō dal[-ō]* ‘he has been torn, he burst’;

Ergative transitive,

perfect, *[-om] koṛ-ō* ‘I have done’;
past perf., *[-om] kōr-a bōn* ‘I have done, I had done’.

3.2.5.3 Dubitative

3.2.5.3a Dubitative-future, *xō*

Dubitative is marked by the particle *xō* preceding the unmarked present and past forms:

present, *m-ond-ē [xō] mēr-a* 'he may, might kill him (*m-ond-ē*)'
 (cf. Kb. *xāt ū-rā be-koš-eš*, *xāt košt-eš*);
 past, *bālō [xō] bin* 'it may, could be the boy (*bālō*)' (cf. Kb. *xāt būd*).

3.2.5.3b Dubitative future perfect, *xāhāt*

An apparent future perfect form, with *xāhāt* followed by the past stem, is found in the following:

ma mā ham [xāhāt dehī] 'you will indeed (*ham*) have beaten us'.

3.2.5.4 Subjunctive and conditional

3.2.5.4a Present subjunctive and conditional

The only verb to distinguish the subjunctive is 'to be' (see section 3.2.3 Person marking). With other verbs the subjunctive is identical with the unmarked present-future, used prominently in dependent clauses: *bōčh(ē)-em* 'that I see' (cf. Kb. *be-bin-am*). To it corresponds the rare conditional with the invariant conditional of the copula, *bā*:

na-m koṛ [bā] 'if only I had not done (it)' (cf. Kb. *na-mē-kad-om*).

3.2.5.4b Continuous subjunctive

A further, apparently continuous, subjunctive is constructed with the present participle in *-en*:

xoxo-e.SBJ raḥīq-pen-ē aiš kan[-en] b-ē
 'may she be enjoying herself (*xoxo*) with her friend'.

3.2.5.4c Perfect subjunctive and conditional

(1) The perfective subjunctive is marked by the subjunctive forms of 'to be':

magam mardum-ikā nhāmōṛ [čh-ō b-ē]
 perhaps people-EZ memory it will have gone
 'people will probably have forgotten (me)'.

(2) The perfect conditional is expressed by the past perfect:

aze āya b-ē, mon žū čhaṭak-ī dā bē
 yesterday hadst thou come I.AG one rupee had given-us

'had thou come yesterday, I would have given thee a rupee (*čhaṭak*)',

where *bē* is 2s past of 'to be' (< *bē-ē*).

3.2.5.5 Note on contracted past and perfect forms

Verbs with past stems in *-ī*, prominently the regular verbs, distinguish past and perfect forms clearly only in the 3s. These stems derive from **-īd*, e.g. **čhīd* ‘gone’ (< **čiyuta-*), and show contraction with the personal endings due to the loss of intervocalic **-d-* as follows: **-īd* before vowel > *-ē-*, but > *-ī* in pause.

The perfect forms of these verbs show further contraction due to the loss of intervocalic *-g-* in the participial form **-īd-ag-* as follows: before vowel > *-ē-ē-* > *-ē-*, but > *-ō/ō* in pause. Thus, in the past, **čhīd* with the personal endings *-am*, *-ē*, *-Ø*, *-aman*, *-ēr*, *-an*, and in the perfect **čhīd-ag-* (< **čiyutaka-*) with the personal endings, have the following forms:

‘to go’	1s	2s	3s	1p	2p	3p
Past	<i>čhēm</i>	<i>čhē</i>	<i>čhī</i>	<i>čhēman</i>	<i>čhēr</i>	<i>čhēn</i>
Perfect	<i>čhēm</i>	<i>čhē</i>	<i>čhōlčhō</i>	<i>čhēman</i>	<i>čhēr</i>	<i>čhēn</i>

3.2.5.6 Overview of past and perfect forms

The past perfect forms are subject to similar contractions. The following table shows the patterns of the various strong verbs ending in a dental or vowel, and of the weak verbs:

TABLE 12.8: PAST AND PERFECT FORMS

Past					
	‘to sit down’ strong	‘to die’ strong	‘to stand up’ strong	‘to come’ strong	‘to go’ weak
<i>*-tld-</i>	3s <i>nhošt</i> 1s <i>nhašt-am</i>	<i>moř</i> <i>moř-om</i>	<i>oštā</i> <i>oštā-ē-m</i>	<i>āγ-a</i> <i>āγ-ē-m</i>	<i>čh-ī</i> <i>čh-ē-m</i>
Perfect					
<i>*-tld-ag-</i>	3s <i>nhašt-ō</i> 1s <i>nhašt-ē-m</i>	<i>moř-ō</i> <i>moř-ē-m</i>	<i>oštā-w-ō</i> <i>oštā-w-ē-m</i>	<i>āγ-ō</i> <i>āγ-ē-m</i>	<i>čh-ō</i> <i>čh-ē-m</i>
Past Perfect					
<i>*-tld-ag-</i>	3s <i>nhašt-a bōn</i> 1s <i>nhašt-a bē-m</i>	<i>moř-a bōn</i> <i>moř-a bē-m</i>	<i>oštā bōn</i> <i>oštā bē-m</i>	<i>āγ-a bōn</i> <i>āγ-a bē-m</i>	<i>čh-ō bōn</i> <i>čh-ē bē-m</i>

3.2.5.7 Overview of tenses, moods, and aspect

The following synoptic table shows 1s forms of the transitive verb *kan-lkoř-* ‘to do’ for the basic pattern here ergative (see section 4.3.3 Subject and object marking, and 3s forms of the change-of-state verb *nhīn-lnhašt-* ‘to sit down’ for temporary state.

TABLE 12.9: SYSTEM OF TENSE, MOOD, AND ASPECT

Indicative		Non-Indicative	
PR <i>kan-em</i>	‘(will) do’	<i>kan-em</i>	‘that I do’
PT <i>-om koṛ</i>	‘did’		
PF <i>-om koṛ-ō</i>	‘have done’	<i>-om koṛ-ō b-ē</i>	‘may have done’
PP <i>-om koṛ-a bōn</i>	‘had done’	<i>-om koṛ-a bōn</i>	‘would have done’
Temporary State			
PR <i>našt-ō a</i>	‘is sitting’		
PT <i>našt-al-ō (h)ost</i>	‘was sitting’		
Continuous			
PR <i>kan-tōn-em</i>	‘am doing’	<i>kan-en bi-m</i>	‘may be doing’
PT <i>kan-tōn (h)ast-am</i>	‘was doing’		
Dubitative			
PR <i>xō kan-em</i>	‘may do’		
PT <i>-om xō koṛ</i>	‘might do’		

3.2.6 Aktionsart and modal constructions

3.2.6.1 Incipient action

Incipient action is expressed by the perfective verbal noun in *-ō + da(r) into* ‘and verbs of movement, such as *čh-* ‘to go, become’, *čhaṛ-* ‘to fall’: [*da naγōn xoṛ-ō*] *čh-ē-n* ‘they began to eat (bread)’; [*da deh-ō*] *čhaṛ-ī* ‘he fell to = started striking’ (*deh-ldehī* ‘to hit’).

3.2.6.2 Intention

Intention is expressed by phrases such as *benā kan-* ‘to begin, intend’ and *xāeš dēr-* ‘to have the wish’ in *ezāife* constructions with the oblique of the verbal nouns in *-ō*:

benā [-e udhēk-ān māt-ō-eka] koṛ
‘he has the intention [of killing that one];’

xāeš [-e jīnč bar-ō-eka] dēr-a
‘he has the wish [of taking a wife].’

3.2.6.3 Ability and possibility

Ability or possibility is expressed by the verb *nar-/narī* ‘be able, be possible’ and the verbal noun in *-en*. The construction may be (1) personal, marked by the subject endings, or (2) impersonal, marked by the personal enclitic in experiencer (‘dative’) function:

Personal ending,

[*jang kan-en-ē*] *nar-ē* ‘canst thou fight him (*kan-en-ē*)?’;
[*xar-en-ē*] *te na nar-em* ‘I cannot eat it (-ē)’.

Personal enclitic,

nā-m nar-ē xar-en ‘I cannot eat it (-ē)’, where *nā-m* < *na-om*, lit. ‘not-to me’.

3.2.6.4 Necessity and obligation

Necessity is expressed by the impersonal *bāyat* ‘is necessary, must’, also *bāyat as(t)* ‘should’ followed by the subjunctive (cf. local Persian *bāyat*; *me-bāyest*):

[*bāyat as*] [*če xoxo feker kan-eman*]
‘it should be that self thought we make’, i.e. ‘think ourselves’.

Obligation may be expressed by the verbal noun in *-ō* (past stem) together with terms implying appropriateness, such as *rawā* ‘agreeable’, *bakār* ‘appropriate’:

[*čh-ō*]-*aw rawā nā*
‘thy going is not appropriate’, i.e. ‘thou shouldst not go’;

[*malāmat koṛ-ō*] *bakār nā*
‘(thy) reproaching (him) is not good’, i.e. ‘appropriate, necessary’.

3.2.7 Passive and causative formations

3.2.7.1 Passive and gerundive

3.2.7.1a Passive

Passive is expressed periphrastically by the passive participle in *-en* and the verb *čh- / čhī* ‘to go, become’ (present-future *par-a-*), alternating with *čhēm- / č(l)ēmi-* ‘to go, move, wander’:

[*mēr-en*] *čhī*
‘he was killed’;

ān-em berkhi-tōn-em če [mēri-en na parī]
‘I am afraid that he may be killed’].

3.2.7.1b Potential passive

Passive with potential connotation: *howī ker-a [čhēm-tōn]* ‘that work can be done’ (*-a* ‘is’, continuous present). Note that both uses (and the etymology) correspond to that of *šodan* in Persian (*košte mišavad, mišavad kard*).

3.2.7.1c Adjectival gerundive

The gerundive, present stem + *-āne*, is only found in adjectival position:

ker-e [čēm-āne] bin, čhī
‘there was work [to be done], (and) it was (done)’.

3.2.7.2 Causation

Direct causation is indicated by the marker *-ēiv* (see section 3.2.1 Stem formation). Other-agent causation may be expressed by the verb *dah- ba* ‘to give to’, and the imperfective verbal noun in *-ō*:

[*ba mer-ō*]-*aw te dah-em*
to be killed-thee will I give
‘I will give thee over to be killed, will have thee killed’.

Together with inherited intransitives and transitive-causatives, there thus may be four forms:

Intransitive: *mer-lmoř* ‘to die’; *mēr-en čhčhī* ‘to be killed’;
 Transitive: *mēr-lmāt* ‘to kill’; *ba mer-ō dah-l-dā* ‘to have s.o. killed’.

4 SYNTAX AND USE OF FORMS

4.1 Conjunctions and clitics

The coordinating conjunction is *o*, postvocalic *wo*. The enclitic connector is *ham* ‘also, even’. Alternates include *na-na* ‘neither-nor’. An emphatic particle is *xō, xu*. There are two modal particles *te* and *e* used with the general present (see section 4.3.5). The general subordinating conjunction is *če*.

4.2 Noun phrase structure

Parachi belongs to those languages which are basically left branching, that is, the head noun follows dependent nominals. However, due to interference from Persian, the inverse order is equally found.

4.2.1 Noun and adjective

- (1) Adjectives may immediately precede the noun, ADJ – N: *činō puš* ‘little boy’.
- (2) More frequently they follow the noun in an *ezāfe* construction, N-*e* ADJ: *puš-e činō*. In that case, the nominal case marker is attached to the adjective as a phrase clitic: (*ōsp-e čhatō*)-*eka zīn* ‘[the white horse]’s saddle (*zīn*)’.
- (3) The inverse *ezāfe* is also found, Adj-*e* N: *aga bad-e gonākār ham čh-ēm* ‘even if I become an evil (*bad*) sinner’.

4.2.2 Noun and dependent noun

There are three options: The dependent noun in the oblique case:

- (1) immediately precedes the head, OBL – N: [*hōwī ādam-eka*] *ōsp* ‘[that man’s] horse’;
or
- (2) immediately follows the head, N – OBL: *ōsp* [*hōwī ādam-eka*]; or
- (3) the dependent noun follows the head by *ezāfe*. In this case, it is usually in the oblique case, N-*e* OBL: *pā-n-ān* [*-e bāw-eka*] ‘the feet of father’s’; but it may also be in the direct case, N1 -*e* N2: *ōsp* [*-e hōwī ādam*] ‘the horse of that man’.
- (4) There is also a focused construction marked by an anaphoric enclitic: [*žū-eka*] *nām-[ē]* *Air bīn*, lit. ‘one’s, his name was Air’.

These options may lead to ambiguity: the sequence

ma zīn ōsp-eka pōř-tar lam da may be interpreted as
[ma zīn ōsp-eka] pōř-tar lam da ‘put the saddle of the horse on its back’, or
ma zīn [ōsp-eka pōř-tar] lam da ‘put the saddle on the horse’s back’.

4.2.3 Functions of the oblique

In overview, the functions of the oblique in noun phrases include the following:

- (1) possession, genitive: [*Māmad Hanīfa Sāheb-eka*] *ōsp* ‘M. H. Saheb’s horse’;
- (2) genitive of kind and measure: [*naḡōn-eka*] *thōī-ān* ‘burnt pieces of bread’;
- (3) object of adpositions, alternating with the direct case: [*Māmad Hanīf-ān*]-*pen* ‘with M. H.’; but [*rafiq*]-*pen-ē* ‘with his (-ē) friend’;
- (4) object of adverbial *ezāfe* constructions: *nazdik* [-*e ma ḡos-eka*] *āya* ‘he came near the house’; *pešt-e* [*Māmad Hanīfa Sāheb-eka* *ōsp-eka*] ‘behind M. H. Saheb’s horse’; *pošt-[-e pā-n-ān -e bāw-eka]* ‘behind father’s feet’; *mōx-tar* [-*e Mīrzā, biyā-e ghaṇḍ-eka*]-*ī* ‘on the face (*mox*) of M., his elder brother’s’.

4.3 Clause structure and word order

4.3.1 Order of constituents

Parachi is basically a SOV language. Generally, any part of speech that is determined by the semantic range of the verb occurs in immediate preverbal position. However, adverbial phrases giving a particular direction, as well as direct objects, and more rarely subjects, may be found in postverbal position.

4.3.2 Semantics and use of nominal forms

4.3.2.1 Animacy

While there is no distinction of masculine and feminine gender, there is a distinction of animacy. This includes the following:

- (1) the interrogative-indefinite pronoun, animate *ki*, inanimate *čē*;
- (2) the existential verb, general and animate sg. *hā*, pl. *hēn*, inanimate pres. *sī*, past sg. *sō bōn*, pl. *se bē-n*;
- (3) the singular oblique *-ān*, with high agency pronouns and named persons as opposed to *-eka* with all others.

4.3.2.2 Plurality and agreement

Plural is as a rule not marked after numerals and numbers: *žū, dī* *ōsp* ‘one, two horses’, *čā phar sēb* ‘several items apples’ (but see 3.1.6.5a Numerative *-a*). It is, however, mostly marked after the implicitly indefinite specific *hos* ‘all’, *hodī* ‘both’ (see 3.1.6 Numerals).

Bare nouns may function as collectives. In that case, the plural is generally marked in the verb, if subject:

ganem-ōw āsor bakār chē-n
wheat-2p this year good became-3p
‘your wheat has become good this year’.

Otherwise, it is understood by context:

man-ān xat ko xān-en na-nar-a
‘no one (*ko na-*) can read my letters (sg. *xat*)’.

Conversely, plural items or persons may have singular personal endings or pronominal referents, apparently when seen as a set:

juwān-ān-e kārī-a ‘they are (-a ‘is’) good young men’;
hō čōr-ē čhē-n ‘those four of them (-ē ‘its, of it’) went’.

In turn, singular items or persons may be expressed by plurals:

mā-tar gur-ē ‘take it from us = me’.

4.3.2.3 Definiteness and specificity

The unmarked noun may be definite or indefinite, depending on the context. Similarly, the plural may be indefinite: [*jalāt-ān*]-ē *dhēwī* ‘he (-ē.AG) called for (some) executioners’.

Specific indefiniteness may be indicated by the unstressed numeral *žū* ‘one’: *žū ošp* ‘one, a (certain) horse’. Also found is the Persian indefinite -ī: *čhil suwār-ī* ‘some forty horse-men’. Its attachment is still rather free: *jangal-tar* ~ *jangal-ī tar* ~ [*jangal-tar*]-ī *zahī* ‘he came to a (certain) forest’.

Definiteness may be indicated by the stressed demonstrative, usually the near demonstrative adjective: *hē ošp* ‘this/the horse’.

Specificity is obligatorily marked in case of the direct object, mostly with the preposition *ma*.

4.3.3 Subject and object marking ergative

Parachi is an ergative language. The subject in the present, and in the past of intransitives, is in the direct case, but the agent/subject of past transitive verbs is in the oblique case if pronominal, either independent or attributive: *ōd[-ē.AG] pānt če* . . . ‘he understood that. . .’; *ēd[-ē.AG] mīrža jāri* ‘this scribe said’; else it is unmarked:

dukāndār *žāri ma kaṭṭō zāif*
 ‘the shopkeeper.AG said to the old woman’.

The direct object of transitive verbs is marked by the oblique case in both the present and past. This may be the bare oblique, particularly with personal pronouns, but usually the object is marked by *ma*: present, [*mōn*] *poṭ kan* ‘hide me’; *ān zanēng [ma xā-m] gom kan-em* (subj.) ‘how can I lose my husband?’; past, [*ma mōn*] *dhōr* ‘he saw me’; *mon [ma wā] dhor* ‘I saw you.p’, with agent oblique *mon*.

4.3.3.1 Patient ending

As seen in the examples above, the ergative in Parachi has “decayed”: patient marking in past tenses is the same as in the present, and the verb ending is usually that of the unmarked 3s. However, the materials contain several instances of the full ergative construction; in particular, the patient is expressed in the verb by the personal ending, even though the direct object/patient is marked by the oblique and *ma*:

žū espō [*ma mon*] *šoṇḍ-pen* *danān koṭ[-am]*
 one dog.AG me muzzle-with tooth did-Is
 ‘a dog bit me with his muzzle’.

4.3.3.2 Indirect object

The indirect object as experiencer and beneficiary may be expressed by the plain oblique case, i.e. without adposition:

(a) with *čh-* ‘to happen’:

[*puš-eka*] *zā na čhō bō*
‘nothing had happened [to the son]’;

(b) emotions and sensations:

feyān oštā [kašt-eka] *ešq sōr-tar-ē āya [šekār-eka]*
‘lament (*feyān*) rose [from the girl]; fancy (*ešq*) [for a hunt] came to his head’;

[*mon o tān*] *jang-ahē ketāb [hōwī ādam-ikā]*
‘there is quarrel to/between [us]’; ‘this book is [that man’s]’,
lit. ‘to me and thee’.

4.3.3.3 Indirect object raising

There are likewise instances where the indirect object (‘indirect affectee’) may be raised and be marked by the personal ending as is shown in the following example:

aze āy-a b-ē, *mon žū čhaṭak-ī dā bē*
yesterday come be-2s (perf. subj.), I.AG one rupee-a given be-2s (past perf.)
‘hadst thou come yesterday, I would have given thee a rupee’.

4.3.3.4 Attachment of personal enclitics

The personal enclitics function as general oblique case markers. This includes the genitive, dative, direct object, and object of adpositions, and the agent in past tenses of transitive verbs.

Their leftmost attachment includes interrogatives or conjunctions:

č-ē jar-tōn ‘what art (-ē) thou saying’;
če-om qasam xoṛ ‘that I (-om) have sworn an oath’.

Sequential person marking occurs frequently:

general present, *mēr[-an-om] te* ‘they (-an) will kill me (-om)’;
imperfective present, *larz-ēw-tōn[-om-a]* ‘he (-a ‘is’) making me (-om) tremble’;
perfect, *lam-ēw-ō[-ī-a]* ‘thou hast (-a) hung it (-ī) up’.

In the ergative, and in the absence of any other parts of speech, the agent enclitic is attached to the verb form: *xoṛ-[a]* ‘thou (-a) atest’. Otherwise, it is mostly attached to the preverbal phrase, but may be attached to any other phrase for selective emphasis. For example, in the following sentence the parentheses indicate possible positions of 2s -a ~ aw:

tū nī xawān(-a) nāgōn(-a) čē-pen(-a) xoṛ
thou to-night bread what-with ate
‘with what atest thou the bread tonight?’

4.3.4 *Types of possession*

- (1) The copula with the oblique marks possession: *man-ān-a* ‘it’s mine’; *hē ketāb ālī-ān a* ‘that book is Ali’s’; *hē ketāb [hōwī ādam-ikā]* ‘this book is that man’s’; plural, *hē [ādam-āna]-en* ‘these are the men’s’.
- (2) In addition, there is a locative construction with *ba* ‘with’, probably indicating temporary possession: [*ba bāw-om*] *čāvar mazdūr hēn*, lit. ‘with my father are several servants’.
- (3) Possession may also be expressed by the verb *dēr-ldērē b-*, subj. *dērē b-* ‘to have, hold’: *aze ker dērē bīm* ‘I had work (to do) yesterday’.

4.3.5 *Future and modal clause clitics te and e*

The clause enclitics *te* and *e* disambiguate the future and potential future function of the unmarked present (cf. Northern Kurdish *dē, ē*, MacKenzie 1961, I: 180–182). Both occur mostly adjacent to pronominal markers, or in clause final position, which points to their focusing function.

4.3.5.1 *Future clitic te*

The future function of *te* is shown in the following sentence:

γos-tar ĵar-em dāl xā-m [te], če žū xī-γoγōk-om mhēmān-om āγō
 house-in I will tell to husband-my [will] that one sister-son guest-my has come
 ‘In the house I will tell my husband, that my sister-son has come as my guest’.

This sentence is part of an extended direct speech, where a cheating wife tells a Molla what she is going to do, and where every main clause is likewise marked by *te* (text: Morgenstierne 1929: 209).

Similar to English ‘will’ as in ‘that will be Harry’, implying probability, *te* may also connote supposition:

khāin-zā-m [te] gas-a ‘something [seems] to be biting (*gas-*) me (*-m*)’.

4.3.5.2 *Assertive clitic e*

The function of the enclitic *e* (< 3s opt. **hē* ‘may be’) is similar to that of *te*, but implies assertion (cf. section 4.5.3.2 Factual enclitic *ta* in Chapter 14):

žē, če ma tō [-e] gap ĵan-em ēr-em [-e] ān [-e]
 come that to thee [let] talk I hit ‘I [shall] kill (thee)’;
 ‘come, let me say a word to thee’;

The uncertainty implied may be emphasized by the addition of dubitative *xō*:

bī rafiq-om ma šār na [(-e)] (xō) par-a-m
 without friend-my to town not [shall perhaps] I go
 ‘perhaps I will not go to the town (*šār*) without my friend’.

5 COMPLEX SENTENCES

5.1 Coordinate clauses

Coordination of clauses is by far the dominant way of combining two or more clauses, with or without overt coordinating conjunctions, clitics, or conjunctive phrases. Such patterns are seen in the Sample Text below. In both coordination and subordination there is a strong tendency to follow the Persian patterns.

5.2. Subordinate clauses

The indigenous pattern appears to have been that the conjunction is placed immediately before the verb phrase. This is still found, as in the following example from the sample text below:

tar-e ē žū-ika žu bālō če bin,
 from-EZ this one-OBL one boy CONJ he-was
 ‘that one who was someone’s only child . . .’

5.3 Relative clauses

Following the Persian pattern, the head noun phrase of the restrictive relative clause is often preceded by *-ī*, with or without demonstrative: *dāl howī zaīf-ī, če* ‘near that woman, who’; *hawī māneš (-ī) če* ‘this man, who’. The relative *-ī* follows postpositions: *žangal-tar-ī če* ‘to the forest which’.

če may be absent:

ker-a koṛ-ō, xūb ker-ā
 ‘the work thou (*-a ~ -aw*, agent) did, is good work’.

5.3.1 Generalized clauses

Such is often the case with generalized clauses:

harkī-an γont, mēr-tan-en
 ‘whomever they (*-an*) found, they used to kill’,
 lit. ‘they are killing’ (note the present tense).

5.3.2 Anaphoric pronouns

The syntactic relationships are indicated by anaphoric pronouns and similar means, and may be so even if in subject function:

molā nasroddīn žū māneš bin, če ē ma mhawṛ riz-ō hōst
 ‘Molla Nasroddin was a certain man, that he was sleeping in a garden’.

Similarly, an anaphoric pronoun often follows in the main clause:

ma šēr, če wā dhoṛ-ō, ō ham tāb-e manān-a
 ‘the lion that you saw, **it** also is my subject’.

5.3.3 Case attraction

Case attraction, that is copying the syntactic function of the head within the relative clause onto the head, is not infrequent:

havī kašt-eka če dōst-ān-ē bastō bōn, thārī če . . .,
lit. 'the girl's **that** you tied her hands, said that . . .';

ma šēr če wā dhor-ō, ō ham tāb-e manān-a
'the lion (dir. obj.) that you saw, **it** also is my subject' (cf. Payne 1982: 351 ff.).

5.4 Object clauses

Object clauses follow the main clause, with or without *če*. The dependent verb is as a rule in the present or present perfect of the observed action:

dhor-an, če onhāk žū māneš našt-ō a
'they saw, **that** a man (*māneš*) was sitting there', lit. 'is sitting'.

5.5 Adverbial clauses

5.5.1 Temporal clauses

Temporal clauses precede the main clause and are marked by the enclitic *če* attached to the first part of speech. The verb may be in the indicative or non-indicative, depending on the degree of reality:

sabā če čhī, wāda barābar čhī
'when the next day came, the marriage union (*wāda*) became due';

ma maktab če hast-am, fārsī hast-am xān-tōn
'when I was in school, I was studying Persian'.

5.5.2 Conditional clauses

Conditional clauses are mostly introduced by *aga(r)* 'if'. In the protasis, most frequently found is the general present or the unmarked past, followed in the apodosis by the present-future or imperative, with or without *te* and *e*. The apodosis may be introduced by the particle *xō*:

aga ēn-en-ē na nar-em, xō xar-em te
'if I cannot bring it, I will eat it'.

In the following example, the verb *narī-m* is in the unmarked past, implying real possibility and anticipated completion, followed by the unmarked present-future *žī-m*:

agar narī-m, sabā žī-m
'if I can, I go tomorrow'.

Unfulfilled conditions are indicated by the past perfect:

aze āy-a bē, mon žū čhaṭak-ī dā bē
'hadst thou come yesterday, I would have given thee a rupee'.

5.5.3 Causal clauses

Causal clauses usually follow the main clause, and are introduced by *če*:

na čaṭak-ōr, če ē šēr moṭ
‘don’t flee-1p, because that lion has died’.

Another frequent conjunctions is *čo-kon (če)*, lit. ‘what for’ (cf. Persian *zī-rā (ke), cūn-ke*).

5.5.4 Final clauses

Final clauses, implying intended or potential action, are in the general present in its subjunctive function. They may or may not be introduced by *če*:

γax kan-em, (če) mēr-an-aw
‘I will call, that they kill thee (-aw)’;

ān zanēng ma xā-m gom kan-em-SUBJ, če kā-wanō parī-a
‘how (*zanēng*) may I get rid of my husband, so that he goes somewhere else (*kā-wanō*)?’

5.5.5 Other clauses and conjunctions

Other conjunctions include *tā, tā če* ‘till, until’, and complex conjunctive relative phrases such as *ēdē az xāter-ī če* ‘in the thought that’, ‘in order to’; or topicalized *ēdhēk-tar . . . če*, ‘from this that’, ‘because’:

ēdhēk-tar-om pānt, če tamām-e jān-eka-ē tā zānū šor ōsp larzī
‘I understood (-om) it for that reason, that all of her body (*jān*) to the knees (*zānū*) trembled on the horse (*ōsp*)’.

More complex subordination is occasionally found. Thus, the following example includes a relative clause embedding direct speech:

molā-ē (če jaṭ-ō “bōn če xī-γoṭōk-om-a”) alāhedī yos-tar nhānt
‘she (-ē, agent) put the Molla that she had said that “he is my sister-son” in another house’.

5.6 Nominalized clauses

Nominalized clauses are participial or infinitival, and prominently express temporal relations between two actions.

5.6.1 Coordinating participle in -en

This construction appears to express simultaneous or immediately successive actions:

bī-γam nhīn (khān-en)
‘sit down without worry, laughing’;
(*m-endān-ān gurī-n*) *hā dhār-tar whē-tōn*
‘taking these with him, he is going to the mountains (*dhār*)’.

5.6.2 Subordinating participle in *-amān*

This construction expresses prior action if unmarked, or subsequent action indicated by the preposition *tā*:

(*naγōn xar-amān*) *ošta-eman*
 ‘having eaten the bread, we got up’,
 cf. the finite clause
naγōn-an čē xor ‘when we (-an.AG) had eaten (lit. ‘ate’);

tā dāda-m *jōr čē-amān*
 ‘until my father having gotten better’.

An example of embedded subordination is found in

(*čēm-amān*), (*jar-amān*) *xō-pen-om oštā*
 ‘having told (*jar-*) him, upon leaving (*čēm-*), he got up together with me’.

5.6.3 Verbal noun + *tar* ‘from’

The ablative of the verbal noun in *-ō* may function similarly:

(*hē ker kor-ō*)-*tar* *m-onde-m dhōk*
 ‘after doing this work, I (-M.AG) said to him’.

6 LEXIS AND SOCIOLINGUISTIC ASPECTS

6.1 Word formation

Word formation in Parachi is similar to that of Persian and other Iranian languages. Both derivation and composition are well documented in the Shotol data. Prefixation is rare, with the exception of Persian *bī*, *bē* ‘without’: *bī-bān* ‘dumb’, lit. ‘without tongue’. Typical derivative suffixes are the adjectival *-ī* and the abstract *-ī*: adjective from noun, *dhār* ‘mountain’ > *dhār-ī* ‘mountain-’, and noun from adjective, *γorča* ‘hungry’ > *γorčagī* ‘hunger’. Other suffixes include diminutive *-ōk*: *biyā-k-ōk* ‘younger brother’, *bārīk-ōk* ‘very slender’; and *-ak*: *kamān* ‘gun’ > *kamān-ak* ‘forearm’; note, *γōšt-ak*, the vigesimal unit for the higher decades.

Composition is prominently found with kinship terms: *biyā-γorōk* ‘brother-son’, and plant names. It includes largely obscured old compounds such as *gū-mān* ‘stable’ (*gū* ‘cattle’, and the obsolete *mān* ‘dwelling’), and loaned neologisms such as *alī-kāfstar* ‘helicopter’, lit. ‘Ali’s pigeon’, which includes sound-play with the English term, with probable reference to the white doves of Mazār-e Sharif.

6.2 Sociolinguistic aspects and loan component

The loan component consists of two major historical layers (cf. the overview in Morgenstierne 1983: 524): Indic, specifically the neighboring Dardic language Pashai, which can be traced back to antiquity; and Persian, specifically its Kaboli variety, the main impact of which began with the Islamization of the Parachi-speaking areas in the sixteenth century. Both thoroughly affected not only the lexicon, but also phonology and grammatical structure. The Indic impact resulted in the integration of retroflex and

aspirated consonants in phonology, and in grammatical restructuring, particularly in the verb system, including loan morphemes, such as the progressive *-tōn*, and participial *-en*. The impact from Persian is evident throughout, affecting the vowel system and stress patterns, and syntax on all levels. This impact gained new impetus during the efforts of Amir Abdur Rahman at unification between 1885 and 1901, and the subsequent increasing centralization and modernization, including the building of roads into the valleys. More recently, impact comes from the expanding regional variety of Persian, known as Kohestani. The impact from Pashto, as one of the two official languages, is fairly recent and has been relatively limited. Smaller components include loans from Western languages, in particular English.

These processes of progressive acculturation are reflected throughout the lexical domains. They include both the spiritual-religious domain: *ḥarām* 'forbidden', *ḡen(d)* 'jinn', *nemāz* 'prayer' *parda* 'curtain, chastity, virtue', and the material domain, with doublets such as Kb. *čōb* 'wood for carpentry' vs. Par. *žōx* 'firewood'. A distinct sub-type are partially integrated loans and calques: *khōr-e-gū* 'hare' < Par. *khōr* 'donkey' and *gū* 'ear', patterned after Persian *xar-gūš*; *tī-e bē-dāna* < Par. *tī* 'mulberry' + *ezāfe* + Kb. *bē-dāna* 'seedless'; and poetic *ezāfe*-compounds such as *āwe-e dāda-eka* 'tear' < Par. *āwe* 'water' and Persian *dāda* 'eye', with Par. oblique suffix *-eka*, after Persian *āb-e dāda* 'water of the eye'.

On the whole, the loan component is freely used, often without phonetic or morphological modification, while indigenous neologisms either in form or meaning are rare, and increasingly unlikely.

7 DIALECTOLOGY AND LANGUAGE DEVELOPMENT

Internally, the sub-dialects of Parachi differ relatively little from each other (cf. Phonology). Given their relative locations, Shotol is in a Perso-phone environment and thus more exposed to Persian than are the others. As indicated, the number of speakers of Parachi is approximately 3,500, and is decreasing under the impact of the steadily increasing dominance of the regional and national languages, which is fostered by obligatory military service and other factors contributing to population mixture, and the impact of the media. It appears that the younger generation, already thoroughly multilingual like the older generation, has ceased to use Parachi as their preferred language of communication, even locally, and no longer regards it as part of their ethnic identity (Kieffer 1977, 1979, 1980).

8 SAMPLE TEXT

The following is a tentative broad phonemic transcription of the Pacheghan text in Kieffer 1977c: 251–252 (CONJ = conjunction; EZ = *ezāfe* construction; INAN = inanimate; INDEF = indefinite; NUM = numerative case; OBL = oblique case).

ē žu qesa a če e di bāl-a bē-n;
 this one story is CONJ this two boy-NUM they-were
 'This is a story, that there were the two boys;'

ma maktab hast-an.
 in school they-used-to-be
 'they used to be in school.'

ē di bāl-a žu-y hōdi ham-senf bē-n, žu senf.
 this two boys-NUM one-EZ both same-class they-were, one class
 'The two boys were both classmates, in the same class.'

tar-e ē ži-ika žu bālō če bin,
 from-EZ this one-OBL one boy CONJ he-was
 'Of these one who was someone's only child.'

ēd-ān bāb-ān γolu hasti zyāt sē bōn.
 this-OBL father-OBL very wealth much was-INAN
 'his father had great wealth' (lit., 'to the father was')

e hēžā(-m) leṛē če bin, ēd-e γolu bi-čāra bin.
 this other-also lad CONJ was, this-OBL very poor was
 'and the other lad, who was there, as to him, he was very poor.'

e hōdihewi aṇḍiwāl bē-n.
 this both comrade they-were
 'Both were comrades.'

waxt-e če ehewi γarib hōst,
 time-OBL CONJ this-same unfortunate was-being
 'While he who was very unfortunate,'

hēwi bi-čāra hōst če γolu kušēš kan-ta,
 this-same poor was-being CONJ very effort making
 'the one who was poor, he was working very hard,'

zyāda tar e hežāy če γolu pōl-dār bin.
 more than this other CONJ very wealthy was
 'more than the other one who was very rich.'

ē wōs kušēš na-kant-ta. (note *wōs* < *hōst* and *-ta* < *-ton*)
 this was effort not-making
 'That one made no efforts.'

ē wōs sor ēd-e khaṇ-ta u maskaragi
 this was on this-OBL laughing and ridicule
 'He used to laugh at and ridicule the other one,'

wos khaṇ-ta sor hēwi bi-čāra-tar
 was laughing on this-same poor-to
 'laughed at the poor one.'

žu waxt-e čhī če ē bi-čāra awal-nōmra čhī.
 one time-INDEF came CONJ this poor first-number became
 'A time came when the poor one became the first of the class.'

āxer ē hewi edānek-ān če hasti sē bon,
 finally this same this-very-OBL CONJ wealth was-INAN
 'In the end, the one who had great wealth,'

e edānek-ān hāzer-bāš moqarar čhī.
 this this-very-OBL attendant appointed became
 'he was appointed attendant of the other one (in the army).'

e bi-čāra lerē feker kor
 this poor lad thought made
 ‘The poor lad thought’

če ē hasti u ē zendagi edānek-ān
 CONJ this wealth and this life this-very-OBL
 thus: ‘The wealth and (good) life for him’

če ba kār āya-n-a?
 what-Q good has-come?
 ‘what good has come?’

aw mun če zyāmat nōt, aw zyāmat-om kašēyi.
 and me-OBL CONJ trouble endured and trouble-me drawn
 ‘And because I have pained myself, and have endured pain,’

mon kon e če-qadar fāyda-e zyāt čh!
 me-OBL for this how-much profit-EZ plenty came
 ‘so much good has happened to me!’

če e tō edh-ān hasti u zendagi,
 CONJ this thee this-OBL wealth and life
 ‘i.e. this thy wealth and (good) life’

u zendagi u zendagi tō seyl kan
 and life and life thou, look make
 ‘and at thy life and wealth look’

če-ē bay āy-ō
 CONJ-him now come-has:
 ‘what has happened to him now?’

man-ān nafar-xesmati yā kan-ta.
 my-OBL person service he is doing
 ‘He is my personal attendant.’

bāyat as če mā xōxō feker kan-eman
 necessary is CONJ we self thought we-make
 ‘We ourselves must think’

m-endhēk xo-k-ān gap.
 this-OBL own-OBL talk
 ‘talk about this our story.’

BIBLIOGRAPHY AND REFERENCES

- Atlas linguistique de l’Afghanistan* (under the direction of Georges Redard), Bern.
 Beveridge, A. S. (1922) *The Bābur-nāma in English (Memoirs of Bābur)*, translated from the original Turki text of Žahiru’d-dīn Muhammad Bābur Pādshāh Ghāzī, London (repr. 1969).
 Elphinstone, M. (1815) *An Account of the Kingdom of Caubul and its Dependencies in Persia, Tartary, and India*, London: Longman, Hurst, etc. (repr. Graz: Akadem. Druck- u. Verlagsanstalt, 1969).

- Faiz, A. (1994) *The Parachgan*, Lahore: Frontier Post Publications.
- Kieffer, C. M. (1977a) 'The approaching end of the relict Southeast Iranian languages Örmürī and Parāčī in Afghanistan', *Linguistics* 191: 71–100; also: *International Journal of the Sociology of Language* 116: 175–185.
- (1977b) 'Études parāčī, I. Les Parāci et leur environnement', *StIr* 6.1: 97–125.
- (1977c) 'Études parāčī, II. Textes, phraséologie et lexicographie', *StIr* 6.2: 249–281.
- (1978) 'Études parāčī, III. Notes sur la phonologie et la morphosyntaxe du parāčī', *StIr* 7: 81–107; 269–277.
- (1979a) 'La fin proche des langues iraniennes résiduelles du Sud-Est, Örmürī et Parāčī, en Afghanistan, I-II', *Langage et Société* (Paris) 10: 37–71.
- (1979b) 'Études parāčī, IV. Le lexique parāčī: Glossaire', *StIr* 81: 67–106; 245–267.
- (1980a) 'La fin proche des langues iraniennes résiduelles du Sud-Est, Örmürī et Parāčī, en Afghanistan, II', *Langage et Société* (Paris) 11: 3–18.
- (1980b) 'Études parāčī, IV. Le lexique parāčī: Glossaire', *StIr* 9: 99–119; 233–249.
- (1981) 'Études parāčī, IV. Le lexique parāčī: Glossaire', *StIr* 10: 283–306.
- (1983) 'Afghanistan v. Languages', in *EnIr* 1, pp. 501–516.
- (1986) 'La maintenance de l'identité ethnique chez les Arabes arabophones, les Örmür et les Parāčī en Afghanistan' in Orywal, E. (ed.) *Die ethnischen Gruppen Afghanistans*, Wiesbaden: L. Reichert, pp. 101–164.
- (1989) 'Le Parāčī, et le groupe des langues iraniennes du Sud-Est', in *CLI*, pp. 445–455.
- Kohzad, M. N. (1954) 'Un coup d'oeil sur la vallée de Nedjrau', *Afghanistan* 9.2: 60–62.
- (1957) 'The distribution of the Parachi language', *Afghanistan* 12.2: 39–41.
- MacKenzie, D. N. (1961) *Kurdish Dialect Studies*, 2 vols, London, New York: Oxford University Press.
- Masson, Ch. (1842) *Narrative of Various Journeys in Balochistan, Afghanistan, and the Panjab, Including a Residence in those Countries from 1826 to 1838*, 3 vols, London: R. Bentley (repr. Graz, Akademische Druck- u. Verlagsanstalt, 1975).
- Morgenstierne, G. (1926) *Report on a Linguistic Mission to Afghanistan*, Oslo: H. Aschehoug & co.; Cambridge, Mass.: Harvard University Press, pp. 14–39.
- (1970) 'Istālif and other place-names of Afghanistan', *BSOAS* 33: 350–363.
- (1973). 'Parachi', in Morgenstierne, *IIFL* 1, pp. 1–304; Addenda, pp. 417–428 (1st edition 1929).
- (1983) Afghanistan vii. Parāčī, in *EnIr* 1, pp. 522–525.
- Nawata, T. (1983) *Parachi*, Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Payne, J. (1982) 'Relativization in the Iranian languages of the U.S.S.R.', *Folia Slavica* 5.1–3: 344–363.
- Scarcia, G. (1965) *Sifat-nāma-yi Darvūš Muḥammad Ḥan-i Gāzī. Cronica di una crociata musulmana contro i Kafiri di Laḡmān nell' anno 1582*. Manoscritto persiano-kābuli edito e tradotto, con introduzione e indici dei nomi, da G. Scarcia, Roma: Istituto italiano per il Medio ed Estremo Oriente.
- Sidi Ali Reīs (1899) *The Travels and Adventures of the Turkish Admiral Sidi Ali Reīs in India, Afghanistan, Central Asia, and Persia, during the years 1553–1556*, translated by A. Vambéry, London: Luzac & Co. (repr. Lahore: al-Biruni, 1975).

CHAPTER THIRTEEN

PASHTO

Barbara Robson and Habibullah Tegey†

1 INTRODUCTION

1.1 Overview

Pashto is a principal language in Afghanistan and Pakistan. It is spoken natively by over half the population of Afghanistan, an estimated 7.5 million, and an estimated 14 million (ca. 90 percent of the population) in the Northwest Frontier Province of Pakistan. It is also spoken natively in Baluchistan, the province of Pakistan which lies directly south of central Afghanistan, and by a community of over two million native Pashto speakers in Karachi. There are also about 50,000 native Pashto speakers in Iran.

The name of the Pashto language is Romanized in several ways, such as Pashto and Pakhto which reflect dialectal differences. In the following the transcription is based on the pronunciation of the Central dialect, and most examples are from the authors' *Pashto Reference Grammar* of 1996.

Pashto is the national and an official language of Afghanistan, while in Pakistan, Pashto has no official status (for Pashto identity in Pakistan, see Rahman 1995). There Pathans, who receive an education, do so in Urdu or English, or both. In fact, it is in Pakistan where, as a result of the British presence in the North west Frontier Province, there are more Pashto publishing houses, and more Pashto books are published in Peshawar than in Afghanistan.

The earliest known example of written Pashto is a multilingual book of religious verses. The text, in Pashto and Arabic, dates from the end of the sixteenth century and was written by the Afghan anti-Mughal mystic and poet Bāyazid Ansāri of Waziristan (ca. 1525–ca. 1573), known as Pir Roshan and founder of the Roshaniya movement. While making religious writing available to people in their own language, the Pashto in that pioneering work is a stilted, unnatural kind of rhymed prose that reflects the style of Arabic found in the Koran.

The earliest Pashto literature was written by the Khattak clan the foremost writer of which was the pre-eminent poet Khoshal Khan Khattak (1613–1689), whose adaptation of the Persian alphabet laid the foundations for the modern Pashto writing system. The Khattaks spoke the Kandahar dialect of Pashto and the spelling system still reflects the Kandahar dialect more than it does the Central or Eastern dialects.

Since Pashto is not standardized, there is no universal agreement among Pashto speakers as to what constitutes “correct” Pashto, either oral or written. One of the results of this lack of standardization is that individual Pashtun writers vary widely, both in spelling and punctuation, not only from one writer to another, but often with the same writer, and within the same document. Pashtuns in Pakistan frequently adapt the conventions of written Urdu in writing Pashto, including spelling conventions. Finally, many

Pashto texts are written by writers who have only a few years of formal schooling, and most probably not in Pashto, and their writing may reflect both lack of extensive education, and the lack of language standardization.

1.2 Writing systems

The following table lists the letters of the Pashto alphabet in their standard order (underdot = Pashto retroflex phonemes.—A = letters found mostly in Arabic loanwords; P = letters found in Persian and Pashto.—W = Western Pashto dialects; C = Central Pashto dialects; E = Eastern Pashto dialects).

TABLE 13.1 PASHTO ALPHABET

Position in Letter Group				Name	Pronunciation
End	Middle	Begin	ALONE		
ا			ا	'alef	ʾ
آ			آ		â
ب	ب	ب	ب	be	b
پ	پ	پ	پ	P pe	p
ت	ت	ت	ت	te	t
ټ	ټ	ټ	ټ	ټe	ټ Pashto
س	س	س	س	A se	s
ج	ج	ج	ج	jim	j
چ	چ	چ	چ	P che	ch
ح	ح	ح	ح	A he	h
خ	خ	خ	خ	khe	kh
ځ			ځ	tse	ts
ښ			ښ	dze	dz
د			د	dāl	d
ډ			ډ	ḍāl	ḍ Pashto
ز			ز	A zāl	z
ر			ر	re	r
ړ			ړ	re	ړ Pashto
ز			ز	ze	z
ژ			ژ	P zhe	W zh, C z, E j
ږ			ږ	zhe, ge	W zh, C g, E g Pashto
س	س	س	س	sin	s
ش	ش	ش	ش	šin	sh
ښ			ښ	šin, xin	W šh, C x, E kh Pashto
ص	ص	ص	ص	A sād	s
ض	ض	ض	ض	A zād	z
ط	ط	ط	ط	A tā	t
ظ	ظ	ظ	ظ	A zā	z
ع	ع	ع	ع	A ʿayn	ʿ
غ	غ	غ	غ	ghayn	gh
ف	ف	ف	ف	fe	f
ق	ق	ق	ق	A qāf	q
ک	ک	ک	ک	kāf	k
گ	گ	گ	گ	P gāf	g
ل	ل	ل	ل	lām	l
م	م	م	م	mim	m

ن	ن	ذ	ن	nun	n
ښ	ښ	ښ	ښ	ṇun	ṇ Pashto
و	و	و	و	wāw	w, u, o
ه	ه	ه	ه	he	h; a in final position
ی	ی	ی	ی	ye	y, i
ی	ی	ی	ی		e
			ی		əy in final position, masculine
			ی		əy in final position, feminine

2 PHONOLOGY

2.1 Inventory

2.1.1 Vowels

The Central dialect of Pashto has nine vowels, of which *î* and *û* are confined to borrowings, indicated by parentheses.

TABLE 13.2: VOWELS

	Front (unrounded)	Mid	Back (rounded)
High	<i>i</i> (<i>î</i>)		(<i>û</i>) <i>u</i>
Mid	<i>e</i>	<i>ə</i>	<i>o</i>
Low		<i>a</i>	<i>â</i>

Semivowels. There are not phonemic diphthongs. Phonemically, the postvocalic semivowels *w* and *y* are consonants.

2.1.1.1 Conditioned variation

The vowel *a* tends to alternate with *ə* in unstressed position. In all dialects, but particularly the Central dialect, *e* and *o* frequently become *i* and *u*, respectively, if a following vowel is *i* or *u*: *keḡəṃ* ‘I do’ as opposed to *keḡi* or *kigi* ‘he does’. *a* changes to *â* when the following final syllable ends in *ə*: *taṛəṃ* ‘I tie’ as opposed to *tâṛə* ‘he ties’.

2.1.1.2 Distribution

Initial occurrence of *î*, *e*, *o*, and *û* is confined to borrowings, e.g. *eyjansey* ‘agency’.

2.1.2 Consonants

2.1.2.1 Inventory

There are 32 consonants in the Central dialect.

TABLE 13.3: CONSONANTS

	Labial	Dental	Retr.	Palatal	Velar	Back Velar	Glottal	Phar.
VI Stops	<i>p</i>	<i>t</i>	<i>ʈ</i>		<i>k</i>	<i>q</i>	ʔ	
Vd Stops	<i>b</i>	<i>d</i>	<i>ɖ</i>		<i>g</i>			
VI Fricatives	<i>f</i>	<i>s</i>		<i>sh</i>	<i>x</i>	<i>kh</i>	<i>h</i>	<i>H</i>
Vd Fricatives		<i>z</i>		<i>zh</i>	<i>gh</i>			<i>ʒ</i>
Nasals	<i>m</i>	<i>n</i>	<i>ɳ</i>					
VI Affricates		<i>ts</i>		<i>ch</i>				
Vd Affricates		<i>dz</i>		<i>j</i> [dzh]				
Trill		<i>r</i>	<i>ɽ</i>					
Lateral		<i>l</i>						
Semivowels	<i>w</i>			<i>y</i>				

Examples

p: voiceless bilabial stop, *parda* ‘curtain’, *chap* ‘left’, *khapa* ‘unhappy’;

b: voiced bilabial stop, *bəl* ‘other’, *tartīb* ‘order’, *bābā* ‘grandfather’;

t: voiceless dental stop, *tashakūr* ‘thank’, *tut* ‘mulberry’, *atə* ‘eight’;

d: voiced dental stop, *da* ‘this’, *tod* ‘warm’, *judā* ‘separate’;

ʈ: voiceless retroflex stop, *ʈol* ‘all’, *tsəʈ* ‘nape’, *moʈar* ‘car’;

ɖ: voiced retroflex stop, *ɖer* ‘very’, *ɖand* ‘valley’, *doɖəy* ‘bread’;

k: voiceless velar stop, *kār* ‘work’, *bārīk* ‘sensitive’, *akā* ‘uncle’;

g: voiced velar stop, *gwəl* ‘flower’, *jəg* ‘tall’, *bāzingar* ‘dancer’;

q: voiced aspirated uvular stop, *qār* ‘anger’, *sandūq* ‘trunk’, *daqīqa* ‘minute’;

ʔ: glottal stop, *majmūʔ* ‘collection’;

m: voiced bilabial nasal, *mor* ‘mother’, *salām* ‘hello’, *māmā* ‘maternal uncle’;

n: voiced dental nasal, *nə* ‘not’, *tānvān* ‘burden’, *karwanda* ‘field’;

ɳ: voiced retroflex nasal, *lune* ‘daughters’, *atan* ‘dance’;

f: voiceless labio-dental fricative, *fārsī* ‘Farsi’, *talafūz* ‘pronunciation’;

s: voiceless dental fricative, *sur* ‘red’, *lās* ‘hand’, *asān* ‘easy’;

z: voiced dental fricative, *zoy* ‘son’, *andāz* ‘style’, *əvāza* ‘rumor’;

sh: voiceless palatal fricative, *shpag* ‘six’, *shā* ‘back’, *ashnā* ‘familiar’;

x: voiceless palatal/front velar fricative, *xa* ‘good’, *partugāx* ‘drawstring’, *paxto* ‘Pashto’;

kh: voiceless velar fricative, *khvdāy* ‘God’, *tārīkh* ‘date’, *wakht* ‘time’;

gh: voiced velar fricative, *ghāx* ‘tooth’, *rogh* ‘healthy’, *āgha* ‘that’;

h: voiceless glottal fricative, *həm* ‘also’, *nəhə* ‘nine’.

H: voiceless pharyngeal fricative, *Hamal* ‘March–April’, *Habīb* ‘dear’;

ʒ: voiced pharyngeal fricative, *ʒaqəl* ‘wisdom’, *tābʒy* ‘subject’, *be-ʒamala* ‘without study’ (Arabic *ʿaql*, *tābīʿ*, *ʿamal*, respectively);

ts: voiceless dental affricate, *tso* ‘how many’, *hets* ‘no’, *nâtsɛgi* ‘they dance’;
dz: voiced dental affricate, *dzân* ‘self’, *termāndz* ‘between’ *xədzə* ‘woman’;

ch: voiceless palatal affricate, *chera* ‘where’, *mrəch* ‘pepper’, *achawi* ‘they throw’;
j: voiced palatal affricate, *joɾ* ‘good’, *khârij* ‘outside’, *tâjək* ‘Tajik’;

l: voiced dental lateral, *lās* ‘hand’, *khpəl* ‘his own’, *dəltə* ‘here’;
r: voiced dental tap or trill, *rogh* ‘healthy’, *tashakūr* ‘thank’, *chera* ‘where’;
ɾ: voiced retroflex flap, *ɾund* ‘blind’, *zoɾ* ‘old’, *stəɾay* ‘tired’;

w: voiced bilabial semi-vowel, *wāza* ‘open’, *yaw* ‘one’, *wəwāya* ‘say’;
y: voiced palatal semi-vowel, *yaw* ‘one’, *khwdāy* ‘God’, *aryāyi* ‘Aryan’.

Of these, *f*, *q*, *H*, *ɣ* occur in borrowings. In the informal speech of educated Pashtuns, and in the formal and informal speech of uneducated Pashtuns, (1) *f* is pronounced as *p*; (2) *q* is pronounced as *k* (although Pashtuns growing up in areas where there is extensive interaction with Dari speakers will often have *q* exclusively in the relevant words), and (3) *H* and *ɣ* (reflecting Arabic *h* and ‘(‘ *ayn*) are dropped altogether.

The consonantal system as a whole is characterized by: (1) dental vs. palatal affricates: *ts*, *dz* and *ch j* [dzh]; (2) retroflex dental-alveolars: *ɬ*, *ɖ*, *ɾ*, *ɳ* (but not *l*); (3) palatal vs. palatal/front velar vs. velar voiceless fricatives: *sh*, *x*, *kh* (see section 6.1 Sub-groupings of dialects).

2.1.2.2 Distribution

h and *ɳ*. In native Pashto words, *h* is found only in syllable-initial position; *ɳ* only in syllable-final position.

2.1.3 Clusters and syllable structure

Pashto is unique among Iranian languages in the extent of permitted clustering. Binary initial clusters include stops with same-voice fricatives and continuants in first or second position (including *xk* and *khɸ*), and continuants with continuants, with relatively few exceptions; *n* does not occur in second position. Triple initial clusters are rare, and only include: *skw*, *skhw*, *shkhw*, *khkw*, *khvd*, *khvl*, and *ndr*. The most common final clusters in native Pashto words include: *sk*, *st*, *xt*, *kht*, *ndr*, *nd*, *ɾkh*, and *rg*.

2.2 Stress

Pashto is the only Iranian language to have phonemic stress:

<i>ṭānga</i> ‘pear tree’	<i>ghuṭa</i> ‘knot’	<i>aspa</i> ‘mare’
<i>ṭāngə</i> ‘cart’	<i>ghuṭə</i> ‘dive!’	<i>aspa</i> ‘spotted fever’

<i>joɾə</i> ‘well (f)’	<i>goɾə</i> ‘look’	<i>kenastə</i> ‘he sat’
<i>joɾə</i> ‘pair’	<i>goɾə</i> ‘fair-skinned’	<i>kenastə</i> ‘he’s sitting’.

While stress is not predictable, as a general rule, stress is on the last syllable when ending in a consonant, and on the penultimate syllable if the last syllable ends in a vowel *randzūr* ‘sick’ (m) vs. *randzura* ‘sick’ (f). This rule-of-thumb, however, is contradicted

by the very name of the language, *paxto*. The sole cases of predictable stress are the prohibitive prefix *mə*, the negative prefix *nə*, and the perfective prefix *wə*, all of which retain stress even in combination:

daredəm 'I was standing' *nədaredəm* 'I wasn't standing'
wədaredəm 'I stood' *wənədaredəm* 'I didn't stand'.

See also Bečka (1969), and Morgenstierne (1973) for a comparative-historical study.

3 MORPHOLOGY

3.1 Nominal morphology

Nouns, adjectives, and pronouns in Pashto distinguish masculine and feminine gender (m, f), singular and plural number (s, p); and direct and oblique case (D, O). Within each gender, there are distinct inflectional classes.

The inflections of the adjectival and nominal classes correspond to each other, one of the distinctive differences being the absence of the nominal animate and inanimate plural markers *-ân* vs. *-un* in adjectives.

A major distinction between pronouns and nouns, in addition to their morphology, is functional: pronouns as direct objects in present tenses appear in the oblique case, as opposed to the direct case of nouns.

A specific feature of number-gender assignment is the fact that the abstract categories of (1) infinitives (verbal nouns) and (2) subject, object, and complement clauses, are grammatically marked as masculine plurals.

3.1.1 Nouns

3.1.1.1 Gender

The most basic division of nouns is gender. In addition to natural gender, each noun is either masculine or feminine gender, which determines agreement within the noun phrase and within the sentence. Overt gender distinction is marked by:

(1) the feminine ending *-a*:

parastâr 'nurse' (m) *parastâr-a* 'nurse' (f)
shâgârd 'student' (m) *shâgârd-a* 'student' (f)
shpun 'shepherd' (m) *shpan-a* 'shepherdess, shepherd's wife'.

(2) the feminine ending *-əy* in opposition to masculine *-ay*:

sp-ay 'dog' *sp-əy* 'bitch'
wur-ay 'lamb' (m) *wur-əy* 'lamb' (f)
kâbul-ay 'man from Kabul' *kâbul-əy* 'woman from Kabul'.

In general, nouns ending in a consonant, *-ay* and *-u* are masculine, and nouns in unstressed *-a*, *-əy*, and *-o* are feminine, while nouns ending in other vowels may be either gender.

3.1.1.2 Inflectional classes

The grammatical study of Pashto has not progressed to the point where there is widespread agreement on such basic topics as the number of noun classes. Writers differ in their analyses, and use different criteria for choosing the number and characteristics of classes.

The classification suggested here is largely based on the forms of plural endings. There are three classes of masculine nouns, including two sub-classes, and three classes of feminine nouns. In addition, there are irregular masculine and feminine nouns.

Masculine nouns

- (1) M_{1a} and M_{1b} nouns end in a consonant; of these,
 - (a) M_{1a} nouns are mostly animate, with animate plural marker $-\hat{a}n$, while
 - (b) M_{1b} nouns are mostly inanimate, with plural marker $-un-$;
- (2) M_{2a} and M_{2b} nouns end in unstressed or stressed $-ay$, respectively;
- (3) M_3 nouns end in a stressed vowel, and are usually animate.

Feminine nouns

- (1) F_1 nouns end in a vowel, usually $-a$;
- (2) F_2 nouns mostly end in $-\underline{ə}y$;
- (3) F_3 nouns end in $-a$ or $-\hat{a}$, sometimes in $-o$ or $-i$.

Morphologically, the three masculine and three feminine classes are correlated, M_1 with F_1 , M_2 with F_2 , and M_3 with F_3 .

There are two basic patterns of case marking. In most the direct singular is identical with the oblique singular, $Ds = Os$, but in two, F_1 and M_2 , the oblique singular is identical with the direct plural, $Os = Dp$. In this analysis, these alternate patterns determine the paired masculine and feminine classes, as follows:

M_1	$Ds = Os$
F_1	$Os = Dp$
M_2	$Os = Dp$
F_2	$Ds = Os$
M_3	$Ds = Os$
F_3	$Ds = Os$

These patterns are shown in the following table and examples.

TABLE 13.4: NOMINAL ENDINGS

M = final consonant -C, F -a					
	Ds	Os	Dp	Op	
M _{1a}	-C	= -C	-â _n	-â _n -o	mostly animate
M _{1b}	-C	= -C	-un-a	-un-o	mostly inanimate
F ₁	-a	= -e	= -e	-o	unstressed ending
M -ay, F -əy					
	Ds	Os	Dp	Op	
M _{2a}	-ay	= -i	= -i	-o	stressed ending
M _{2b}	-ay	= -i	= -il-y-â _n	-ol -y-â _n -ol -əy-o	
F ₂	-əy	= -əy	= -əy	-əy-ol -əyol -y-â _n -el -əy-g-â _n -e	stressed ending
M. â i u ə, F. â i o ə a					
	Ds	Os	Dp	Op	
M ₃	-â	= -â	-âg-â _n	-âg-â _n -o	
F ₃	-â	= -â	-âw-el -âg-â _n -e	-âw-ol -âg-â _n -o	

TABLE 13.5: EXAMPLES OF DECLENSIONS

	Ds	Os	Dp	Op	
M _{1a}	'student'	<i>shâgərd</i>	= <i>shâgərd</i>	<i>shâgərd-â_n</i>	<i>shâgərd-â_n-o</i>
M _{1b}	'house'	<i>kor</i>	= <i>kor</i>	<i>kor-un-a</i>	<i>kor-un-o</i>
F ₁	'woman'	<i>xədz-a</i>	<i>xədz-e</i>	= <i>xədz-e</i>	<i>xədz-o</i>
	Ds	Os	Dp	Op	
M _{2a}	'friend'	<i>məlgə_r-ay</i>	<i>məlgə_r-i</i>	= <i>məlgə_r-i</i>	<i>məlgə_r-o</i>
M _{2b}	'man'	<i>sar-ay</i>	<i>sar-i</i>	~ <i>sar-il</i> <i>sar-y-â_n</i>	<i>sar-ol</i> <i>sar-y-â_n-o</i> <i>sar-əy-o</i>
F ₂	'family'	<i>koran-əy</i>	= <i>koran-əy</i>	~ <i>koran-əy/ol</i> <i>koran-əy-g-â_n-el</i> <i>koran-y-â_n-e</i>	<i>koran-əy-ol</i> <i>koran-əy-g-â_n-ol</i> <i>koran-y-â_n-o</i>
	Ds	Os	Dp	Op	
M ₃	'uncle'	<i>mâmâ</i>	= <i>mâmâ</i>	<i>mâmâ-g-â_n</i>	<i>mâmâ-g-â_n-o</i>
F ₃	'salary'	<i>tankhâ</i>	= <i>tankhâ</i>	<i>tankhâ-w-el</i> <i>tankhâ-g-â_n-e</i>	<i>tankhâ-w-ol</i> <i>tankhâ-g-â_n-o</i>

3.1.1.2a Irregular nouns and ablaut

Pashto has a considerable number of irregular nouns, each of which is clearly masculine or feminine, but beyond that, so varied in their oblique and plural forms that they cannot readily be grouped into any of the more regular noun classes described above.

A distinctive feature of these inflections is various types of ablaut, frequently correlated with final stress. They are mostly found with masculines, and mostly follow the pattern of Os = Dp of M₂. Feminine forms are derived from the weakest stem. Some examples are:

	Ds	Os	= Dp	Op
<i>CuloC > CâlâC</i>				
'Pashtun'	<i>paxt-un</i>	<i>paxt-ân-ə</i>	<i>paxt-ân-ə</i>	<i>paxt-an-o</i>
'rider' (m)	<i>swor</i>	<i>swâr-ə</i>	<i>swâr-ə</i>	<i>swâr-o</i>
<i>CaliluC > CC</i>				
'mountain' (m)	<i>ghar</i>	<i>ghr-ə/una</i>	<i>ghr-ə/una</i>	<i>ghr-o</i>

Irregular nouns include some of the most common kinship terms, such as the following:

	Ds	= Os	Dp	Op
'mother'	<i>mor</i>	<i>mor</i>	<i>maynd-e</i>	<i>maynd-o</i>
'sister'	<i>khôr</i>	<i>khôr</i>	<i>khwâynd-e</i>	<i>khwâynd-o</i>
'aunt'	<i>tror</i>	<i>tror</i>	<i>trâynd-e</i>	<i>trâynd-o</i>
'daughter-in-law'	<i>ngor</i>	<i>ngor</i>	<i>ngâynd-e</i>	<i>ngâynd-o</i>
'daughter'	<i>lur</i>	<i>lur</i>	<i>luñ-e</i>	<i>luñ-o</i>
'son'	<i>zoy</i>	<i>zoy</i>	<i>zâmən</i>	<i>zâmən-o</i>
'brother's son'	<i>wrârə</i>	<i>wrârə</i>	<i>wrer-un-a</i>	<i>wrer-ol/wrer-un-o</i>
'husband's brother'	<i>lewar</i>	<i>lewar</i>	<i>lewr-un-a</i>	<i>lewr-un-o</i>

In addition, masculine stems in *-C*, mostly inanimate, take "numerative" *-a* after cardinal numbers larger than one or other quantifiers (e.g. the terms for '100' and '1000', see section 3.1.6.1 below).

3.1.1.2b Dialectal variation

Irregular nouns vary widely from dialect to dialect, from speaker to speaker, and even within the same speaker at different times. For the most part, this variation arises when irregular nouns become regularized, e.g. *as* 'horse', which is M_{1b} in the Central dialect, and as such violates the "rule" that animate nouns are M_{1a}. In the Western dialect, however, the word occurs with both M_{1a} and M_{1b} endings. Other examples are M_{1a} nouns denoting inanimates, like *nuk* 'finger-, toenail' and *shay* 'thing', which frequently occur with M_{1b} endings.

3.1.1.2c Singular-only and plural-only nouns

Pashto *mass nouns* occur either only in singular or only in the plural. These nouns tend to denote such things as grain, food, liquid, or other elements that are not practically countable. Examples of singular-only mass nouns are: *chalaw* '(cooked) rice', *kabâb* 'kebob', *palaw* 'pilaf', *xurwa* 'soup'; examples of plural-only mass nouns are: *wrə* 'flour', *wobə* 'water', *khwârə* 'food', *sharâb* 'wine', and *dâl* 'lentils'.

The plural-only nouns have no overt plural endings, but are determined by their plural oblique ending *-o*, and by their co-occurrence with plural adjectives and verbs, as the following sentences illustrate:

<i>wob[-ə]</i>	<i>khwag[-e]</i>	<i>d[-i]</i>
water[-Dpf]	sweet[-Dpf]	be.PR[-3p]
'the water is sweet';		

de kâbəl de sin de wob[-o] kwand nə sta
 of Kabul of river of water[-OBLp] taste NEG be.EXIST
 'water from the Kabul River is not drinkable, 'lit. 'there is no taste to ...'

3.1.1.2d Inflection of infinitives and of nominal clauses

The infinitive (verbal noun), as well as subject, object, and complement clauses belong to this masculine plural-only class:

wayəl[-o] ta ye mə gor-a
 saying-INF[-OBLpm] to his PROH look.PR.IMP-2s
 'don't look at his saying', i.e. 'what he says'.

3.1.1.3 Use of cases

(1) The oblique case of nouns is used for:

- (a) the subject with past transitive verbs (past agent, see 3.2.6.2 Transitivity);
- (b) the object of adpositions;
- (c) in vocatives.

(2) The direct case is used in all other functions, including:

- (a) the subject with present verbs;
- (b) the subject with past intransitive verbs;
- (c) the direct object with present transitive verbs;
- (d) the direct object with past transitive verbs.

Usage of the two cases is illustrated in the sentences below with the irregular noun *paxtun* (AG = agent; PAT = patient; D = direct; O = oblique; PR = present, PT = past; IPFV = imperfective; PFV = perfective).

3.1.1.3a Direct case, subject in present and intransitive past tense

[paxt-un] pə kâbəl ke wos-eg-il wos-ed[-ə]
 [Pashtun]-Dsm in Kabul in live.PR.IPFV-3s/ live.PT.IPFV-3s
 'the Pashtun] lives/lived in Kabul';

[paxt-ân-ə] pə kâbəl ke wos-eg-il wos-ed[-ə]
 [Pashtuns]-Dpm in Kabul in live.PR.IPFV-3p/live.PT.IPFV-3pm
 'the Pashtuns] live/lived in Kabul.'

3.1.1.3b Direct case, direct object in all tenses

ahmad [paxtun] win-i
 Ahmad [Pashtun]-Dsm see.PR.IPFV-3s
 'Ahmad is seeing [the Pashtun]';

ahmad [paxt-ân-ə] lidəl
 Ahmad [Pashtuns]-Dpm see.PT.PFV-3pm
 'Ahmad saw [the Pashtuns]'.

3.1.1.3c Oblique case, subject in transitive past tenses

[paxt-ân-ə] topak râ-wor
 [Pashtun]-OBLsm rifle here-bring.PT.PFV-3s
 'the Pashtun] brought the rifle';

[*paxt-an-ə*] *topak rā-wor*
 [Pashtuns]-Opm rifle here-bring.PT.PFV.3s
 ‘[the Pashtuns] brought the rifle’.

3.1.1.3d Oblique case, object of adpositions

[*de paxt-ân-ə*] *zāmən pə kâbəl ke wos-i*
 [of Pashtun]-Osm sons in Kabul in live.PR.IPFV-3p
 ‘[the Pashtun’s] sons live in Kabul’;

[*de paxt-an-ə*] *zāmən pə kâbəl ke wos-i*
 [of Pashtuns]-Osm sons in Kabul in live.PR.IPFV-3p
 ‘[the Pashtuns’] sons live in Kabul’.

3.1.1.3e Oblique case, vocatives

ey nik-ə! ‘O grandfather!’, *ey xədz-o!* ‘O women!’.

Masculine nouns ending in consonants, including *-y*, form the vocative by adding the suffix *-a* to the direct singular form:

alək-a! ‘boy!’, *malgər-y-a!* ‘friend!’ (D *malgər-ay*).

3.1.2 Adjectives

3.1.2.1 Classes

As in the nominal inflection, there are three classes of adjectives in this analysis, two of which have variants, as well as a number of irregular adjectives.

- (1) Adj_{1a} (final syllable unstressed) and Adj_{1b} (final syllable stressed) form parallel M₁ and F₁ nouns, masculine Adj_{1b} forms adding *-ə* in the Os/Dp.
- (2) Adj_{2a,b} (unstressed, stressed) forms parallel M₂ and F₂ nouns. The Adj_{2a} class includes the present and perfect participles (described in section 3.2.3).
- (3) Adj₃ forms are invariable, but may have an optional plural oblique *-o*.

TABLE 13.6: ADJECTIVAL CLASSES

	Ds	Os	=	Dp	Op	
Adj _{1a} (m)	<i>tit</i>	<i>tit</i>		<i>tit</i>	<i>tīto</i>	‘scattered’
Adj _{1a} (f)	<i>tīta</i>	<i>tīte</i>		<i>tīte</i>	<i>tīto</i>	
Adj _{1b} (m)	<i>zig</i>	<i>zigə</i>		<i>zigə</i>	<i>zigo</i>	‘rough’
Adj _{1b} (f)	<i>ziga</i>	<i>zigə</i>		<i>zigə</i>	<i>zigo</i>	
Adj _{2a} (m)	<i>təgay</i>	<i>təgi</i>		<i>təgi</i>	<i>təgo</i>	‘thirsty’
Adj _{2a} (f)	<i>təge</i>	<i>təge</i>		<i>təge</i>	<i>təgo</i>	
Adj _{2b} (m)	<i>naray</i>	<i>naril</i>		<i>naril</i>	<i>narol</i>	‘thin’
		<i>narəy</i>		<i>narəy</i>	<i>narəyo</i>	
Adj _{2b} (f)	<i>narəy</i>	<i>narəy</i>		<i>narəy</i>	<i>narəyo</i>	
Adj ₃ (m/f)	<i>xâysta</i>	<i>xâysta</i>		<i>xâysta</i>	<i>xâystal</i> <i>xâystəwo</i>	‘pretty’

3.1.2.2 Irregular adjectives

As is the case with nouns, there are irregular adjectives, in which the feminine form is derived from the weakest ablaut stem forms; the patterns are *oluC* ~ *â/aC* and *CaliluC* ~ *CC-*.

TABLE 13.7: IRREGULAR ADJECTIVES

	Ds	Os	=	Dp	Op	
Adj. (m)	<i>pokh</i>	<i>pâkhə</i>		<i>pâkhə</i>	<i>pakhə</i>	'ripe, cooked'
Adj. (f)	<i>pakhā</i>	<i>pakhe</i>		<i>pakhe</i>	<i>pakhə</i>	
Adj. (m)	<i>sor</i>	<i>sâre</i>		<i>sâre</i>	<i>sarə</i>	'cold'
Adj. (f)	<i>sarā</i>	<i>sare</i>		<i>sare</i>	<i>sarə</i>	
Adj. (m)	<i>shin</i>	<i>shnə</i>		<i>shnə</i>	<i>shnə</i>	'green'
Adj. (f)	<i>shnā</i>	<i>shne</i>		<i>shne</i>	<i>shnə</i>	

3.1.2.3 Comparison

Pashto does not have comparative or superlative suffixes.

(1) Comparative

Simple comparison is expressed by the ambipositions *lə . . . na* 'from':

asad [lə amân na] jəg da
Asad [from Aman from] tall be.PR.3s
'Asad is taller than Aman';

de amân plâr de asad [lə plâr na] ləg bâgh-un-a lar-i
of Aman father of Asad [from father from] few orchards-Op have.PR.IPFV-3s
Aman's father has fewer orchards than Asad's father'.

(2) Superlatives

Superlatives are expressed in three ways:

(a) with *lə . . . na* 'from', in which the object of the ambiposition is all-inclusive:

khoshâl pə sinf ke [lə ɬol-o alək-ân-o na] jəg da
Khoshal in class in [from all boys-Op from] tall be.PR.3s
'Khoshal is taller than all the boys in the class'.

(b) A second way to express superlatives is with the adverb *ɖer* 'very', in which case the sentence is ambiguous:

khoshâl pə sinf ke [ɖer jəg] alək da
Khoshal in class in [very tall] boy be.PR.3s
'Khoshal is the tallest boy in the class', lit. 'a very tall boy in the class'.

(c) The third way is with *ɖer* 'very', but with the subject moved out of position to just before the verb, disambiguating the sentence:

pə sinf ke [ɖer jəg] alək khoshâl da
in class in [very tall] boy Khoshal be.PR.3s
'Khoshal is the tallest boy in the class'.

3.1.3 *Pronouns and deixis*

3.1.3.1 Personal pronouns

Pashto has enclitic, proclitic and independent personal pronouns.

3.1.3.1a *Enclitic pronouns*

TABLE 13.8: CLITIC PRONOUNS

<i>Enclitic</i>	1 pers.	2 pers.	3 pers.
Sing.	<i>me</i>	<i>de</i>	<i>ye</i>
Plur.	<i>molam</i>	<i>molam</i>	<i>ye</i>
<i>Proclitic</i>			
Sg./pl.	<i>râ</i>	<i>dər</i>	<i>wər</i>

The enclitic pronouns function as subjects/agents in past transitive sentences, and in possessive constructions, as do the nominal oblique cases; but different from those, they also function as direct objects in present tense sentences.

The proclitic pronouns are phonologically identical with the directional prefixes that form the first syllables of prefixed verbs described in section 3.2.2.3. This identity and the correlation between persons and directions suggests that at some earlier stage of the language the two functions were identical as well.

(1) Direct object in the present tense, and as subject/agent in the past tenses of transitive verbs:

mach-aw-i [me]
kiss.PR.IPFV-3s [me]
'he is kissing me';

mach-aw-əl-əm [ye]
kiss.PT.IPFV-1s [he-AG]
'he was kissing me'.

(2) Possessor:

kitâb [me], *kitâb* [ye]
book [my], book [his]
'my book', 'his book'.

(3) Proclitic with adpositions:

Asad as [râ ta] *akhl-i*
Asad horse [me to] buy.PR.IPFV-3s
'Asad is buying a horse for me';

dâ [dər pəre] *khând-i*
she [you at] laugh.PR.IPFV-3s
'she's laughing at you';

amân nə [wər sara] *dz-i*
Aman not [him with] go.PR.IPFV-3s
'Aman isn't going with him';

laylâ [wər ta] *wə-wel-e*
Layla [them to] PFV.tell.PT.PFV-3sf
'Layla told them'.

3.1.3.1b Independent pronouns

TABLE 13.9: INDEPENDENT PRONOUNS

<i>Singular</i>	1s	2s	3sm	3sf	3sm	3sf
Subject	<i>zə</i>	<i>tə</i>	<i>day</i>	<i>dâ</i>	<i>agha</i>	<i>agha</i>
Object/Prepos.	<i>mâ</i>	<i>tâ</i>	<i>day</i>	<i>dâ</i>	<i>agha</i>	<i>agha</i>
Possessive	<i>di mâl</i>	<i>di tâl</i>	<i>di də</i>	<i>di də</i>		
	<i>zmâ*</i>	<i>stâ*</i>				
<i>Plural</i>	1p	2p	3pm	3pf	3pm	3pf
Subj./Obj./Prep.	<i>mung</i>	<i>tâse</i>	<i>duy</i>	<i>duy</i>	<i>aghu</i>	<i>aghu</i>
Possessive	<i>zung</i>	<i>stâse</i>				

* Non-Central dialects

The independent pronouns are used for emphasis:

[*zə*] *ahmaq nə* *γ-əmm*; [*day*] *ahmaq da*
 [I] stupid NEG be.PR-1s [he] stupid be.PR.3sm
 'I'm not stupid; he is'.

Their oblique cases function both as direct objects in the present tenses and as subjects in transitive past tenses, and occur with adpositions. In the Central dialect, the possessive forms 1s *di mâ* and 2s *di tâ* are used; in other dialects, the forms *zmâ* and *stâ* are preferred, while the 1p and 2p forms are *zung* and *stâse* in all dialects.

The 3rd person pronouns distinguish near and far deixis.

3.1.3.2 Demonstrative pronouns

Demonstrative pronouns in Pashto function as both as independent pronouns and as attributive adjectives. As pronouns, they have initial stress, as adjectives final stress. There is a three-way contrast, here called near, middle, and far.

TABLE 13.10: DEMONSTRATIVE PRONOUNS

<i>Masc./fem.</i>	Dsmf	Osmf	Dp (= sg.)	Op (= sg.)
Near	<i>dâ</i>	<i>de</i>	<i>dâ</i>	<i>de</i>
<i>Masculine</i>	Dsm	Osm	Dpm (= sg.)	Opm
Near	<i>dagh-a</i>	<i>dagh-ə</i>	<i>dagh-a</i>	<i>dagh-o</i>
Mid	<i>agh-a</i>	<i>agh-ə</i>	<i>agh-a</i>	<i>agh-o</i>
Far	<i>ugh-a</i>	<i>ugh-ə</i>	<i>ugh-a</i>	<i>ugh-o</i>
<i>Feminine</i>	Dsf (= m)	[Opf =	Dpf]	Opf (= m)
Near	<i>dagh-a</i>	<i>dagh-e</i>	<i>dagh-e</i>	<i>dagh-o</i>
Mid	<i>agh-a</i>	<i>agh-e</i>	<i>agh-e</i>	<i>agh-o</i>
Far	<i>ugh-a</i>	<i>ugh-e</i>	<i>ugh-e</i>	<i>ugh-o</i>

Focus and identity is marked by the prefix *ham-* 'same' with bisyllabic demonstratives, e.g. *ham-dagha* 'this very one'.

3.1.3.3 Reflexive and possessive pronouns

Reference to self is expressed by the adjective *khpəl* (regular Adj₁ inflection) as well as the noun *dzân* (regular M₂ noun inflection):

ahmad [khpəl] kitâb râ-wor
 Ahmad [own] book here-bring.PT.PFV.3s
 ‘Ahmad brought his (own) book’;

ahmad [ye] kitâb râ-wor
 Ahmad [his] book here-bring.PT.PFV.3s
 ‘Ahmad brought his (someone else’s);

kitâb [dzân] ta akhl-əm
 book [self] for buy.PR.IPFV-1s
 ‘I’m buying the book for myself’.

3.1.3.4 Indefinite and interrogative pronouns and related adverbials

Some basic indefinite pronouns, which also serve as interrogatives, are:

Pronominal

<i>tsok</i> , obl. <i>châ</i>	‘someone; who’
<i>tsə</i>	‘something; what’ (no inflection)
<i>tsəwəm</i>	‘which (number)’
<i>tso</i>	‘several; how many/much’ (plural only)
<i>tsumra</i>	‘that/how much’
<i>kum</i>	‘(a) certain’ (Adj _{1a})

Adverbial

<i>kəla</i>	‘when’	<i>tsə wakht</i>	‘when’
<i>chera, cherta</i>	‘where’	<i>tsə ranga</i>	‘how’
<i>də kəm-a</i>	‘where from’	<i>wali</i>	‘why’

Other inflected pronominal adjectives include: *har* ‘each’, *tol* ‘all’, *bəl* ‘other, else’, *nor* ‘others’, *dzinil-o*.

These pronouns may be combined: *yaw*, *yaw tsok*, *yaw tso*, *yaw-tsə* ‘somebody, some (people), something’; *har yaw* ‘each one’, *har-tsok* ‘each one’; *bəl-tsok* ‘somebody else’. The negative pronoun is *hets* ‘any’ with negated verb, e.g. *hetsok* ‘nobody’; also *hechera* ‘nowhere (to)’, etc.

3.1.4 Adpositions

Pashto adpositions include prepositions, postpositions, and ambipositions which consist of pre- and post-nominal elements. Nominal objects of adpositions are in the oblique case, except in some literary or formal usages.

3.1.4.1 Prepositions and postposition

There are only two prepositions, which are of high frequency, and only one common postposition that we know of:

- *de* ‘of’ forms possessive phrases with nouns or strong pronouns: *de asad plâr* ‘Asad’s father’;
- *pə* has a large range of meanings, the most common of which are ‘with/by means of’: *pə chârə* ‘with/by means of a knife’, and ‘at’ in time expressions: *pə yawā baja* ‘at one o’clock’;
- *ta* ‘to’ typically expresses the indirect object or direction, or dative: *bâzâr ta* ‘to the market’, *asad ta* ‘to/for Asad’.

3.1.4.2 Ambipositions

Most adpositions have two elements, bracketing the nominal object. The first element of these is either *pə*, *lə*, or *tər*. The second element is likely to be one of the following particles: *ke*, *bânde*, *pəse*, *na*, *tsəkha*, *sara*, *lânde*, *pore*. Some of the commonest ambipositions are the following:

<i>pə . . . ke</i> ‘in, at’:	<i>pə kâbəl ke</i>	‘in Kabul’;
<i>pə . . . bânde</i> ‘on/to’:	<i>pə plə bânde</i>	‘on the bridge’;
<i>pə . . . pəse</i> ‘after’:	<i>pə d̄rs pəse</i>	‘after class’;
<i>lə . . . na</i> ‘from’:	<i>lə laylâ na</i>	‘from Layla’;
<i>lə . . . tsəkha</i> ‘from’:	<i>lə logar tsəkha</i>	‘from Logar’;
<i>lə . . . lânde</i> ‘under’:	<i>lə plə lânde</i>	‘under the ridge’;
<i>lə . . . sara</i> ‘with’:	<i>lə asad sara</i>	‘with Asad’;
<i>tər . . . lânde</i> ‘under’:	<i>tər mez lânde</i>	‘under the table’;
<i>tər . . . -a pore</i> ‘up to’:	<i>tər sahâr-a pore</i>	‘until morning’.

Either element of an ambiposition, most often the first, may be deleted: *lə asad na* > *asad na* ‘from Asad’, *lə asad sara* > *asad sara* ‘with Asad’.

The first element of an ambiposition is mutually exclusive with the enclitic pronouns (described in 3.1.3.1a): *wər sara* ‘with him’, *dər na* ‘from you’, and *wər bânde* ‘on it’.

In the ambiposition *lə . . . na* ‘from’, *-n* is frequently elided after consonants as in *lə asad na* > *lə asad-a* ‘from Asad’ in the example above.

The alternation just cited, with *na* > *-a*—which we have analyzed as an optional phonological process—is considered by other scholars to be a second oblique case in other analyses, notably that in Penzl (1955; in general, see also Heston 1987).

3.1.5 Adverbs

See section 3.1.1.3 Use of cases.

3.1.6 Numerals

3.1.6.1 Cardinal numbers

Pashto numbers are all indeclinable adjectives of the Adj_j class, with optional plural

oblique *-o*. The exception is ‘one’, which distinguishes gender and case according to class Adj_{1b}, thus *yaw* (m), *yaw-a* (f).

In the mostly irregular numbers from 11 to 99 the unit precedes the decade: *dre atyâ* ‘3–80’ = 83. Hundreds above 100, and thousands above 1000, follow the unit in their distinct oblique plural forms in *-a*: (*yaw*) *səl* ‘100’, *shpəg səw-a* ‘600’, (*yaw*) *zər* ‘1000’, *dre zər-a* ‘3000’. Larger compound numbers have the connective *aw* ‘and’: *dre zər-a aw shpəg sew-a aw dre atyâ* ‘3683’.

Nouns after numbers larger than one require the plural.

TABLE 13.11: BASIC CARDINAL NUMBERS

1–9	11–19	10–90	100+
<i>yaw*</i>	<i>yawo-las</i>	<i>las</i>	<i>səl</i> ‘100’
<i>dwa</i>	<i>do-las</i>	<i>shəl</i>	<i>zər</i> ‘1000’
<i>dre</i>	<i>dyaŕ-las</i>	<i>ders</i>	
<i>tsalor</i>	<i>tswar-las</i>	<i>tsalwext</i>	
<i>pindzə</i>	<i>pindzə-las</i>	<i>pəndzəs</i>	
<i>shpəg</i>	<i>shpəŕ-as</i>	<i>shpetə</i>	
<i>wə</i>	<i>wə-las</i>	<i>awyâ</i>	
<i>atə</i>	<i>atə-las</i>	<i>atyâ</i>	
<i>nəhə</i>	<i>nun-as</i>	<i>nawî</i>	

* m. *yaw*, f. *yaw-a*, Adj_{1b}.

TABLE 13.12: DECADES + SINGLE DIGITS

21–29	31–69	71–99
<i>yawist</i>	<i>yaw-ders</i>	<i>yaw-awyâ</i>
<i>dw-a-wist</i>	<i>d-u-ders</i>	<i>dw-a-awyâ</i>
<i>dər-wist</i>	<i>dr-i-ders</i>	<i>dr-e-awyâ</i>
<i>tsaler-ist</i>	<i>tsalor-ders</i>	<i>tsalor-awyâ</i>
<i>pindzə-wist</i>	<i>pindzə-ders</i>	<i>pindzə-awyâ</i>
<i>shpəg-wist</i>	<i>shpəg-ders</i>	<i>shpəg-awyâ</i>
<i>wə-wist</i>	<i>wə-ders</i>	<i>wə-awyâ</i>
<i>atə-wist</i>	<i>atə-ders</i>	<i>atə-awyâ</i>
<i>nəhə-wist</i>	<i>nəhə-ders</i>	<i>nəhə-awyâ</i>

3.1.6.2 Ordinal numbers

Ordinal numbers are formed from the cardinal numbers by adding the suffix. *-am* (m), *-əma* (f). The resulting adjectival inflection has the pattern Adj_{1b} [Os = Dp], e.g.:

‘fourth’ (cardinal ends in consonant)		‘seventh’ (cardinal ends in vowel)	
masc.	fem.	masc.	fem.
Ds <i>tsalor-am</i>	<i>tsalor-əm-a</i>	<i>w-am</i>	<i>w-əm-a</i>
Os <i>tsalor-əm</i>	<i>tsalor-əm-e</i>	<i>w-əm</i>	<i>w-əm-e</i>
Dp <i>tsalor-əm</i>	<i>tsalor-əm-e</i>	<i>w-əm</i>	<i>w-əm-e</i>
Op <i>tsalor-əm-o</i>	<i>tsalor-əm-o</i>	<i>w-əm-o</i>	<i>w-am-o</i>

3.2 Verb morphology

The verb system of Pashto is based on two stems: present/past, and two aspects: perfective/imperfective, and distinguishes gender and number in the third person. Mood is distinguished by inflection and particles. All verbs are anchored on the following four verb formations:

PR	PT
Imperfective Perfective	Imperfective Perfective.

The imperfective/perfective distinction appears to relate mostly to whether the action denoted by a verb has been completed or is incomplete. In some analyses, various occurrences of the perfective have been associated with subjunctive functions. These semantic distinctions are not consistent, however, and it seems most prudent—at least until further inquiry is completed—to consider the perfective a grammatical form only.

3.2.1 Person marking and 'be, become'

3.2.1.1 Basic forms

Functionally, endings agree in number person, number, and gender with either subjects or objects depending on the tense and particular construction. The forms of the verb 'be, become' exemplify the inflection of regular verbs. (Note that elements in parentheses indicate that the element *-ə* is optional.)

TABLE 13.13: PERSON MARKING AND 'BE, BECOME'

	Endings		'be' (imperfective)		'become, be' (perfective)	
	PR	PT	PR	PT	PR	PT
1s	-ə <i>m</i>	-(ə)-ə <i>m</i>	<i>y</i> -ə <i>m</i>	<i>w</i> -ə <i>m</i>	<i>s</i> -ə <i>m</i>	<i>sw</i> (-ə)-ə <i>m</i>
2s	- <i>e</i>	-(ə)- <i>e</i>	<i>y</i> - <i>e</i>	<i>w</i> - <i>e</i>	<i>s</i> - <i>e</i>	<i>sw</i> (-ə)- <i>e</i>
Imper.	- <i>a</i>				<i>s</i> - <i>a</i>	
3sm	- <i>i</i>	-ə	<i>d</i> - <i>ald</i> - <i>a</i> <i>y</i>	<i>w</i> -ə	<i>s</i> - <i>i</i>	<i>sw</i> -ə
3sf	- <i>i</i>	-(ə)- <i>a</i>	<i>d</i> - <i>a</i>	<i>w</i> - <i>a</i>	<i>s</i> - <i>i</i>	<i>sw</i> (-ə)- <i>a</i>
1p	- <i>u</i>	-(ə)- <i>u</i>	<i>y</i> - <i>u</i>	<i>w</i> - <i>u</i>	<i>s</i> - <i>u</i>	<i>sw</i> (-ə)- <i>u</i>
2p	-ə <i>y</i>	-(ə)-ə <i>y</i>	<i>y</i> - <i>āst</i> ə <i>y</i>	<i>w</i> -ə <i>y</i>	<i>s</i> -ə <i>y</i>	<i>sw</i> (-ə)-ə <i>y</i>
Imper.	-ə <i>y</i>				<i>s</i> - <i>e</i> <i>y</i>	
3pm	- <i>i</i>	-ə(-ə)	<i>d</i> - <i>i</i>	<i>w</i> -ə	<i>s</i> - <i>i</i>	<i>sw</i> -ə(-ə)
3pf	- <i>i</i>	-(ə)- <i>el</i> -ə	<i>d</i> - <i>i</i>	<i>w</i> - <i>e</i>	<i>s</i> - <i>i</i>	<i>sw</i> (-ə)- <i>e</i>

3.2.1.2 Third person existential and modal forms

In the 3rd person the copula has three additional forms which are undifferentiated as to person, number, and gender: *sta*, *wi*, and *way*; they are negated with *nə*.

- (1) *sta* is the existential verb: 'there is/are'; negated *nə-sta*.
- (2) *wi* (also referred to as subjunctive) is conjunctural, used when an assumption or given fact is being discussed, in contrast to *da* and *di* which are used when reporting an

observation. The contrast between them is clear, for example, in the following sentences referring to a public picnic table:

khalək dālta nāst [d-i] ‘people are sitting here’ (we can see them);
khalək dālta nāst [w-i] ‘people (may) sit here’ (we can use this table).

- (3) *way* (also referred to as conjectural and optative) is used in irreal clauses (*wāy* in non-Central dialects).

Note that the present and past forms of the copula, including *wi* and *way*, are also used to form periphrastic forms of the verb.

3.2.2 Stem formation and tenselaspect

Verbs may be classified according to the way in which their forms are constructed. There are three main classes of verbs: (1) simple verbs; (2) derivative verbs; and (3) prefixed stress-shifting verbs.

3.2.2.1 Simple verbs

Simple verbs form their perfective stems by adding the prefix *wə-* to the imperfective present and past stems.

This class includes: (1) regular verbs where the past stem is derived from the unexpanded present stem by the regular past morpheme *-əl-*; (2) a great number of irregular verbs where the morphological interrelationship between the two stems is opaque; (3) verbs where there is stem suppletion.

In the following examples the forms are in the 1s; forms in parentheses indicate optional omission of the element.

TABLE 13.14: REGULAR SIMPLE VERBS

	PR.IPFV	PT.IPFV	PR.PFV	PT.PFV
‘drink’	<i>tsk-ə̆m</i>	<i>tsk-əl-ə̆m</i>	<i>wə-tsk-ə̆m</i>	<i>wətsk-əl-ə̆m</i>
‘throw’	<i>shind-ə̆m</i>	<i>shind-əl-ə̆m</i>	<i>wə-shind-ə̆m</i>	<i>wə-shind(-əl)-ə̆m</i>
‘have’	<i>lār-ə̆m</i>	<i>lār-əl-ə̆m</i>	<i>wə-lār-ə̆m</i>	<i>wə-lār-əl-ə̆m</i>
‘write’	<i>lik-ə̆m</i>	<i>lik-əl-ə̆m</i>	<i>wə-lik-ə̆m</i>	<i>wə-lik-əl-ə̆m</i>
‘dance’	<i>gād-eg-ə̆m</i>	<i>gād-ed-ə̆m</i>	<i>wə-gād-egə̆m</i>	<i>wə-gād-ed-ə̆m</i>
‘tie’	<i>taṛ-eg-ə̆m</i>	<i>taṛ-ed-ə̆m</i>	<i>wə-taṛ-eg-ə̆m</i>	<i>wə-taṛ-ed-ə̆m</i>

TABLE 13.15: IRREGULAR SIMPLE AND SUPPLETIVE VERBS

	PR.IPFV	PT.IPFV		PR.IPFV	PT.IPFV
‘take, buy’	<i>akhl-</i>	<i>akhist(-əl)-</i>	‘read’	<i>lwan-</i>	<i>lwist(-əl)-</i>
‘wear’	<i>aghund-</i>	<i>aglūst(-əl)-</i>	‘show’	<i>xəy-</i>	<i>xod(-əl)-</i>
‘kill’	<i>wāzn-</i>	<i>wāz(-əl)-</i>	‘see’	<i>gor-</i>	<i>kat(-əl)-</i>

3.2.2.2 Derivative verbs

Derivative verbs are transparently based on adjectives or nouns, and constitute the largest and only open class of verbs. They occur with auxiliaries.

Intransitive and transitive

- (1a) the transitive PR *(k)eg-*, PT *(k)ed(əl)-* related to the full verb meaning ‘be done, become’, and
 (1b) the transitive formants PR *(k)aw-*, PT *(k)aw(əl)-* which when used as a full verb means ‘do’.
 (2a) If the noun or adjective ends in a vowel, the auxiliary is a separate word, and has forms as listed below, but
 (2b) when the noun or adjective ends with a consonant, the *k-* of the auxiliary is dropped, and the rest of the auxiliary is added to the noun or adjective to form a single word.

Note that the imperfective forms of these verbs are indistinguishable from the imperfective forms of many of the simple verbs.

Imperfective and perfective aspects

- (1) Both (a) *(k)eg-*, *(k)ed(əl)* and (b) *(k)aw-*, *(k)awəl-* represent the imperfective forms of these verbs.
 (2) The corresponding perfective forms are expressed by (a) intransitive PR *s-*, PT *sw(əl)-* ‘become, occur’ and by (b) transitive PR *k-*, PT *kr(əl)-*.

The pattern of these distinctions is shown in the following, again with examples in first person singular:

TABLE 13.16: DERIVATIVE VERBS: INTRANSITIVE-TRANSITIVE

	PR.IPFV	PT.IPFV	PR.PFV	PT.PFV
Basic auxiliary verb				
‘be done, become’	<i>k_{eg}-ə_m</i>	<i>ked(ə_l)-ə_m</i>	<i>s-ə_m</i>	<i>sw(-ə_l)-ə_m</i>
‘do, make’	<i>kaw-ə_m</i>	<i>kaw-ə_l-ə_m</i>	<i>wə_l-k(r)-ə_m</i>	<i>wə_l-kr(-ə_l)-ə_m</i>
Derivative verbs				
	PR.IPFV	PT.IPFV	PR.PFV	PT.PFV
‘be injured’				
<i>zobl-eg-</i>	<i>zobl-eg-ə_m</i>	<i>zobl-ed(-ə_l)-ə_m</i>	<i>zobəl s-ə_m</i>	<i>zobəl sw(-ə_l)-ə_m</i>
‘injure’				
<i>zobl-aw-</i>	<i>zobl-aw-ə_m</i>	<i>zobl-aw-ə_l-ə_m</i>	<i>zobəl k-ə_m</i>	<i>zobəl kr(-ə_l)-ə_m</i>
‘become pretty’				
<i>xâysta k-eg-</i>	<i>x. k-eg-ə_m</i>	<i>x. k-ed(-ə_l)-ə_m</i>	<i>x. s-ə_m</i>	<i>x. sw(-ə_l)-ə_m</i>
‘beautify’				
<i>xâysta k-aw-</i>	<i>x. k-aw-ə_m</i>	<i>x. k-aw-ə_l-ə_m</i>	<i>x. k-ə_m</i>	<i>x. kr(-ə_l)-ə_m</i>

3.2.2.3 Prefixed stress-shifting verbs

Prefixed verbs are stress-shifting verbs where aspect is determined by the position of the stress. In the perfective forms, stress is on the nominal component or the prefix, or the first syllable of the verb, while in the imperfective forms it is on the last or penultimate syllable.

The prefixes include forms phonologically identical to the proclitic forms of the pronouns, and carry rough approximations of the meanings of the proclitics (see Proclitic pronouns under 3.1.3.1a):

- râ* ‘here’ (ich-deictic)
dər ‘in your direction’ (du-deictic)
wər ‘there’ (er-deictic)

Other prefixes include directional *ke* ‘down’, *pre* ‘away’, and *nəna* ‘inside’ (the meanings are approximate and inconsistent), plus a number of syllables for which the meaning is not immediately apparent.

This class of verbs is very small and highly irregular, involving both morphological change and shifts in stress, in addition to idiosyncratic third person forms in the past. The complexity of these verbs is demonstrated in the following list:

TABLE 13.17: PREFIXED VERBS: *râ-*, *dar-*, *war-*

	PR.IPFV	PT.IPFV	PR.PFV	PT.PFV
‘go’	<i>dz-əm</i>	<i>tl(-əl)-əm</i>	<i>lâr s-əm</i>	<i>lâr(-əl)-əm</i>
‘come here’	<i>râ-dz-əm</i>	<i>râ-tl(-əl)-əm</i>	<i>râ-s-əm</i>	<i>râ-ghl(-əl)-əm</i>
‘go to you’	<i>dar-dz-əm</i>	<i>dar-tl(-əl)-əm</i>	<i>dar-s-əm</i>	<i>dar-ghl(-əl)-əm</i>
‘go there’	<i>war-dz-əm</i>	<i>war-tl(-əl)-əm</i>	<i>war-s-əm</i>	<i>war-ghl(-əl)-əm</i>
‘give here’	<i>râ-kaw-əm</i>	<i>râ-kaw(-əl)-əm</i>	<i>râ-kɾ-əm</i>	<i>râ-kɾ(-əl)-əm</i>
‘give to you’	<i>dar-kaw-əm</i>	<i>dar-kaw-əl-əm</i>	<i>dar-kɾ-əm</i>	<i>dar-kɾ(-əl)-əm</i>
‘give to him’	<i>war-k-aw-əm</i>	<i>war-kaw-əl-əm</i>	<i>war-kɾ-əm</i>	<i>war-kɾ(-əl)-əm</i>
‘take, carry’	<i>wɾ-əm</i>	<i>wɾ-əl-əm</i>	<i>wes-əm</i>	<i>wɛ-wɾ-əm</i>
‘transport here’	<i>râ-wɾ-əm</i>	<i>râ-wɾ-əl-əm</i>	<i>râ-wɾ-əm</i>	<i>râ-wɾ-əl-əm</i>
‘bring here’	<i>râ-wast-əm</i>	<i>râ-wast-əl-əm</i>	<i>râ-wast-əm</i>	<i>râ-wast-əl-əm</i>
‘take there’	<i>war-wɾ-əm</i>	<i>war-wɾ-əl-əm</i>	<i>war-wɾ-əm</i>	<i>war-wɾ-əl-əm</i>

TABLE 13.18: PREFIXED VERBS: *ke-*, *pre-*, *nəna-*; NOUN + VERB

	PR.IPFV	PT.IPFV	PR.PFV	PT.PFV
‘sit down’	<i>ke-n-əm</i>	<i>ke-nâst(-əl)-əm</i>	<i>ke-n-əm</i>	<i>ke-nâst(-əl)-əm</i>
‘plant’	<i>ke-naw-əm</i>	<i>ke-naw-əl-əm</i>	<i>ke-naw-əm</i>	<i>ke-naw-əl-əm</i>
‘squeeze’	<i>ke-kâg-əm</i>	<i>ke-xkod(əl)-əm</i>	<i>ke-kâg-əm</i>	<i>ke-xkod(əl)-əm</i>
‘put’	<i>gd-əm</i>	<i>ke-xod(-əl)-əm</i>	<i>ke-gd-əm</i>	<i>ke-xod(-əl)-əm</i>
‘allow’	<i>pre-gd-əm</i>	<i>pre-xod(-əl)-əm</i>	<i>pre-gd-əm</i>	<i>pre-xod(-əl)-əm</i>
‘lie down’	<i>pre-wəz-əm</i>	<i>pre-wat(-əl)-əm</i>	<i>pre-wəz-əm</i>	<i>pre-wat(-əl)-əm</i>
‘arrive, enter’	<i>nəna-wəz-əm</i>	<i>nəna-wat(-əl)-əm</i>	<i>nəna-wz-əm</i>	<i>nəna-wat(-əl)-əm</i>
‘cut’	<i>pre-kaw-əm</i>	<i>pre-kaw-əl-əm</i>	<i>pre-kɾ-əm</i>	<i>pre-kɾ(-əl)-əm</i>
‘wash’	<i>pre-mindz-əm</i>	<i>pre-mindz-əl-əm/pre-wəl-əm</i>	<i>pre-mindz-əm</i>	<i>pre-mindz-əl-əm/pre-wəl-əm</i>
<i>khatā</i> ‘error’				
‘deceive’	<i>khatā</i> <i>bâs-əm</i>	<i>khatā</i> <i>yest(-əl)-əm</i>	<i>khatā</i> <i>bâs-əm</i>	<i>khatā</i> <i>yest(-əl)-əm</i>
‘err’	<i>khatā</i> <i>wəz-əm</i>	<i>khatā</i> <i>wat(-əl)-əm</i>	<i>khatā</i> <i>wəz-əm</i>	<i>khatā</i> <i>wat(-əl)-əm</i>

TABLE 13.19: VERBS WITH FUSED PREFIXES

	PR.IPFV	PT.IPFV	PR.PFV	PT.PFV
'lose'	<i>bâyl-ə̃m</i>	<i>bâylođ(-ə̃l)-ə̃m</i>	<i>bâył-ə̃m</i>	<i>bâylođ(-ə̃l)-ə̃m</i>
'take'	<i>byây-ə̃m</i>	<i>bow(-ə̃l)-ə̃m</i>	<i>boz-ə̃m</i>	<i>bołl(-ə̃l)-ə̃m</i>
'get up'	<i>pâts-eg-ə̃m</i>	<i>pâts-ed(-ə̃l)-ə̃m</i>	<i>pâts-eg-ə̃m</i>	<i>pâts-ed(-ə̃l)-ə̃m</i>
'open'	<i>prâ-niz-ə̃m</i>	<i>prâ-nəst(-ə̃l)-ə̃m</i>	<i>prâ-niz-ə̃m</i>	<i>prâ-nəst(-ə̃l)-ə̃m</i>

3.2.3 Nominal forms

3.2.3.1 Present participle

Present participles are adjectives in *-ay* of the Adj_{2a} class, which are derived from present stems by the formant *-unk*, e.g. *đâraw-* 'threaten' > *đâraw-unk-ay* 'threatening'. These participles are artificial constructs, created by the Pashto Academy in the 1940s/1950s. They never caught on, and there are only a handful of words, certainly not a productive class, that made it into common usage, e.g. *xow-unk-ay* 'teacher'.

3.2.3.2 Perfect participle

Perfect participles are formed by affixing *-ay* to the imperfective or perfective past stems of the verbs. They are of the Adj_{1b} class, as shown with the example of the verb *dar-eg-l* *dar-ed-ə̃l* 'stand, stop':

	masculine	feminine
Ds	<i>dar-ed-ə̃l-ay</i>	<i>dar-ed-ə̃l-e</i>
Os	<i>dar-ed-ə̃l-i</i>	<i>dar-ed-ə̃l-e</i>
Dp	<i>dar-ed-ə̃l-i</i>	<i>dar-ed-ə̃l-e</i>
Op	<i>dar-ed-ə̃l-o</i>	<i>dar-ed-ə̃l-o</i>

They function in a number of verb constructions: (a) the "perfect" tenses; (b) ability statements; (c) the passive construction; and (d) conditional sentences.

3.2.3.3 Conditional participle

In the central dialect described here, the conditional participle is formally identical to the perfect participle: *dar-ed-ə̃l-ay*. It is marked for number and gender only when used with the past conditional form of the copula, *way*:

... *ke-nâst-ə̃l-e ba way*
 '... I would have sat-sf down' (female speaker).

3.2.3.4 Infinitive

Infinitives consist of the past imperfective stem + *-ə̃l*. They belong to the masculine plural-only class, Dp *-Ø*, Op *-o*:

<i>[bow-ə̃l]</i> <i>der grân fîl day</i>	<i>de laylâ [kat-ə̃l] saray wazhn-i</i>
[taking] very difficult verb is	of Layla looking man kill.PR.IPFV-3p
'bowə̃l is a very difficult verb';	'Layla's gaze kills a man'.

3.2.4 Aspectual-modal particles and negation

3.2.4.1 Future, necessity, and obligation

There are three aspectual-modal particles, in addition to perfective *wə*:

- (1) definite future is marked by *ba*; except for the infinitive, this particle co-occurs with each of the imperfective-perfective pairs;
- (2) necessity is marked by *de*;
- (3) obligation is marked by *bāyād*.

3.2.4.2 Negation

The prohibitive prefix is *mə*, and the general negative particle *nə*. Both always occur immediately before the inflected verb form, and attract primary stress. Thus with simple, derivative, and prefixed stress-shifting verbs these two particles are inserted after: (1) the perfective *wə*; (2) the prefix; and (3) the nominal component, respectively. (4) In the periphrastic perfect tenses (see 3.2.7.3), where the inflected component is the copular auxiliary at the end of the verb form, insertion entails reordering, by which the negated auxiliary component is shifted to the head of the construction.

- (1) Simple verb, after perfective *wə*

zə ba lik wə [nə] leg-əm
 I FUT letter PFV NEG send.PR.PFV-1s
 'I won't send the letter'.

- (2) Prefixed stress-shifting verb, after prefix

zə ba ke [nə] n-əm
 I FUT pref NEG sit.PR.PFV-1s
 'I won't sit down' (*ke-n-lkenāst*).

- (3) Derivative verb, after nominal component

zə ba chalaw pokh [nə] kr-əm
 I FUT rice-sm cooked-sm NEG do.PR.PFV-1s
 'I won't cook the rice'.

- (4) Periphrastic perfect forms

tor tsawk-əy [nə] da jor-a kər-a
 Tor.AG chair NEG COP build-sf became-3sf
 'Tor has not built the chair'.

3.2.5 Overview of tense, mood, and aspect

Pashto verb forms and verbal constructions involve one or the other of the binary tense distinction and of the equally binary aspectual distinction, as well as various auxiliaries and particles.

The system has been described in various ways. The following analysis, which is based on the fundamental opposition between imperfective and perfective, shows the symmetry of the system. This system has the following subsets: (1) the imperative forms; (2) the

simple present and past forms; (3) the periphrastic perfect forms. In terms of non-indicative modality, there are two: (1) future forms in the present and present perfect; (2) conjectural subjunctive forms, also in the present and present perfect, but mostly confined to use in dependent clauses. In synopsis, this system may be shown as follows (suggested by the Editor of this volume). The verb chosen here is the regular intransitive verb *ras-ed-lras-ed-əl-* 'arrive' with 2s ending.

TABLE 13.20: PASHTO BASIC VERB SYSTEM

	IMPERFECTIVE	PERFECTIVE	
IMP	<i>rasedg-a</i>	<i>wə-rasedg-a</i>	IMP
PR	<i>rasedg-e</i>	<i>ba wə-rasedg-e</i> <i>wə-rasedg-e</i>	FUT SUBJ
PT	<i>rased-e</i>	<i>wə-rased-e</i>	PT
PRPF	<i>rasedəl-ay ye</i>	<i>ba rasedəl-ay ye</i> <i>rasedəl-ay wi</i>	FUT SUBJ
PTPF	<i>rasedəl-ay we</i>		

This table excludes two periphrastic forms which only occur with irreal function in conditional constructions (see 5.5.2 Conditional clauses).

3.2.6 Preliminary notes on inflectional pattern

3.2.6.1 Inflection of verb classes

The individual tenses and verb constructions will be discussed in sections 3.2.7–10. The example sentences given there include verb forms of the three classes of verbs: simple, derived, and prefixed stress-shifting. It is particularly the derived verbs where the inflection is distinctly more complex than that of the simple verb *rasedg-lrasedəl-* shown in Table 13.20. This is because the verb forms in certain tenses are split into (a) an inflected nominal component and (b) the auxiliary. The following table provides an overview of those patterns. See also section 3.2.3.2 Perfect participle, and Tables 13.6 and section 13.7 Adjectival classes.

Note that verbs derived from ablauting adjectives retain the ablaut as well as number and gender marking, e.g.:

zə ba chalow [pokh] kɾ-əm 'I will cook rice' (masc. sg. mass noun), but
tarisā ba āshak [pākh-ə] kɾ-i 'Theresa will cook aushak' (masc. pl. mass noun).

TABLE 13.21: OVERVIEW OF VERB CLASSES AND INFLECTION

	'dance'	'build'	'be built'	'go'	'take'
CLASS	regular	derivative tr.	derivative intr.	suppletive	fused prefix
PR.IPFV	<i>gaḍ-eg-PR</i>	<i>joṛ-aw-PR</i>	<i>joṛ-eg-PR</i>	<i>dz-PR</i>	<i>byây-PR</i>
PR.PFV	<i>wə gaḍ-eg-PR</i>	<i>joṛ-A₁ k-PR</i>	<i>joṛ-A₁ s-PR</i>	<i>lâṛ s-PR</i>	<i>boz-PR</i>
PT.IPFV	<i>gaḍ-ed-PT</i>	<i>joṛ-aw-əl-PT</i>	<i>joṛ-ed-əl-PT</i>	<i>tl(-əl)-PT</i>	<i>bow(-əl)-PT</i>
PT.PFV	<i>wə gaḍ-ed-PT</i>	<i>joṛ-A₁ kṛ(-əl)-PT</i>	<i>joṛ-A₁ sw(-əl)-PT</i>	<i>lâṛ(-əl)-PT</i>	<i>boṭl(-əl)-PT</i>
PF	<i>gaḍ-ed-əl-ay</i> <i>y-lḍ-</i>	<i>joṛ-A₁ kəṛ(-əl)-ay</i> <i>-lḍ-</i>	<i>joṛ-A₁ sw(-əl)-ay</i> <i>y-lḍ</i>	<i>tl-əl-ay</i> <i>y-lḍ</i>	<i>boṭl-ay</i> <i>y-lḍ</i>
PP	<i>gaḍ-ed-əl-ay</i> <i>w-</i>	<i>joṛ-A₁ kəṛ(-əl)-ay</i> <i>w-</i>	<i>joṛ-A₁ sw(-əl)-ay</i> <i>w-</i>	<i>tl-əl-ay</i> <i>w-</i>	<i>boṭl-ay</i> <i>w-</i>

PR and PT = present and past endings
s- and *sw-* = present and past endings of perfective 'be, become'
y-lḍ- = present imperfective of 'be'
w- = past imperfective of 'be'
A₁ = inflected as adjectival class *A_{1a}*
-ay = inflected as adjectival class *A_{2a}*

3.2.6.2 Transitivity and ergative construction

Transitivity is a crucial category in Pashto (Tegey 1979). It is a split-ergative language similar to a good number of modern Iranian languages where verb forms of transitive verbs based on the past stem morphologically change the arguments from the direct-oblique pattern to the oblique-direct pattern, with a concomitant shift of person, number, and gender agreement to the patient. In the examples sentences below: AG = agent oblique case, and D = direct patient case.

3.2.6.3 Order of presentation

The sequence of constructions discussed in the following will be:

- (1) basic tenses and moods (3.2.7);
- (2) modal constructions (3.2.8);
- (3) passive constructions (3.2.9);
- (4) infinitives as subject and object (3.2.10).

Inside each section, the sequence will be present, past, and future forms. For ease of reference, the canonical forms of the verb *rased-ṛased-əl-*, all 2s, found in Table 13.20 will be given first.

The verbs in the examples sentences are mostly taken from the following list to show differences in construction from one verb class to the other.

Intransitive		Transitive
Simple:	<i>gaḍeg-</i> 'dance'	<i>leg-</i> 'send'
Derivative:	<i>joṛeg-</i> 'get better, be built'	<i>joṛaw-</i> 'build'
Prefixing:	<i>dz-</i> 'go'	<i>byây-</i> 'take'

Note that the modal markers of future *ba*, necessity *de*, and obligation *bâyad* are clause clitics and attached to the first simple or complex component of the clause, which may result in the change of word order.

3.2.7 Basic tense constructions

3.2.7.1 Present time constructions

3.2.7.1a Imperative

raseg-ahwə-raseg-a

Note that positive commands usually require the perfective present stem, prohibitive command marked by prohibitive *mə* require the imperfective present stem.

(1) Positive ('go')

<i>wə-gad-eg-əy</i>	<i>tsawk-əy jor-a k-a</i>
PFV.dance.PR.PFV-IMP.2p	chair-sf build-sf AUX.PR.PFV-IMP.2s
'dance (everyone)!';	'build the chair!';

māshumân dər sara bəz-a
 children you with take.PR.PFV-IMP.2s
 'take the children with you'.

(2) Prohibitive

mə + present imperfective stem

<i>maktub wos mə leg-a</i>	<i>wos mə dz-a</i>
letter now PROH send.PR.IPFV-IMP.2s	now PROH go.PR.IPFV-IMP.2s
'don't send the letter now!';	'don't go now!';

kor sarək ta nizde mə jor-əw-əy
 house road to close PROH build.PR.IPFV-IMP.2p
 'don't build the house so close to the road';

laylâ kor ta mə byây-əy
 Layla house to NEG take.PR.IPFV-IMP.2p
 'don't take Layla home yet'.

3.2.7.1b Imperfective present

raseg-e

The imperfective present expresses general, habitual, and ongoing action in the present, and contextually in the future.

(0) 'be'

<i>dukân-dâr y-əm</i>	<i>wəg-ay y-e</i>
shop-keeper be.PR.IPFV-1s	hungry be.PR.IPFV-2s
'I am a shopkeeper';	'you're hungry'.

(1a) Intransitive

gaḍ-eg-əm *pohantun ta dz-u*
 dance.PR.IPFV-1s university to go.PR.IPFV-1p
 ‘I’m dancing/I dance’; ‘we go/are going to the university’.

(1b) Derivative class

tsapl-əy pə pexawər ke joṛ-eg-i
 sandal in Peshawarin **be made**.PR.IPFV-3p
 ‘sandals are made in Peshawar’.

(2) Transitive

laylā khpəl-e koran-əy ta pays-e leg-i
 Layla own family to money send.PR.IPFV-3s
 ‘Layla sends/is sending money to her family’;

mājət joṛ-aw-i
 mosque build.PR.IPFV-3p
 ‘they are building a mosque’.

3.2.7.1c *Perfective present*

wə-raseg-e

The perfective present mostly functions as a subjunctive. It expresses admonition, doubt, or similar modalities in independent clauses, and potential action in dependent clauses (see also 3.2.8 Modal constructions).

asad che maktab khlas k-i, maaləm ba s-i
 Asad **when** school finish-3s do.PR.PFV teacher **FUT** be.PR.PFV-3s
 ‘Asad will be a teacher when he finishes school’;

de asad plār ghwâ-i tse bâgh w-âkhl-i
 of Asad father want.PR.IPFV-3s **CONJ** orchard PFV.buy.PR.PFV-3s
 ‘Asad’s father wants to buy an orchard’.

3.2.7.1d *Present future*

ba wə-raseg-e

(0) ‘be’

tâse ba zər tæg-i səy day ba jæg s-i
 you.p **BUT** soon thirsty-pm be.PR.PFV-2s he **FUT** tall.sm be.PR.PFV-3s
 ‘you all will be thirsty soon’ ‘he will be tall’.

(1) Intransitive

ahmad ba pə wâdə ke wə-gaḍ-eg-i
 Ahmad **FUT** at wedding at PFV.dance.PR.PFV-3s
 ‘Ahmad will dance at the wedding’;

laylā ba joṛ-a s-i
 Layla **FUT** healthy-sf AUX.PR.PFV-3s
 ‘Layla will get better’;

asād ba melmastyâ ta lâr s-i
 Asad **FUT** party to go.sm AUX.PR.PFV-3s
 'Asad will go to the party'.

(2) Transitive

asād ba dewâl jor k-i
 Asad **FUT** wall build.sm AUX.PR.PFV-3s
 'Asad will build a wall';

tor ba ye boz-i
 Tor **FUT** him take.PR.PFV-3s
 'Tor will take him'.

3.2.7.2 Past time constructions

3.2.7.2a Imperfective past

rased-e

The imperfective past has two functions:

- (a) Similar to the imperfective present, the imperfective past expresses general, habitual, and ongoing action in the present and contextually future in the past (prospective past);
- (b) It expresses unreal action in expressions such as wishes, and in conditional clauses (see section 5.5.2).

(0) 'be'

<i>stār-ay w-əm</i>	<i>stār-ay sw-əm</i>
tired be.PT.IPFV-1s	tired be.PT.PFV-1s
'I was tired';	'I became tired';

<i>xowunk-ay w-ə</i>	<i>xowunk-ay s-ə</i>
teacher be.PT.IPFV-3sm	teacher be.PT.PFV-3sm
'he was a teacher';	'he became a teacher'.

(1a) Intransitive

mung wâdə ta tl-u
 we wedding to go.PT.IPFV-1p
 'we were going to the wedding'.

(1b) Derivative class

tsapl-əy pə pexawər ke jor-ed-əl-e
 sandals-pf in Peshawar in **be made**.PT.IPFV-3pf
 'sandals used to be made in Peshawar'.

(2) Transitive

laylâ khpəl-e koran-əy ta pays-e leg-əl
 Layla.AG own family-O to money-Dpf send.PT.IPFV.3pf
 'Layla was sending money to her family';

mâjət ye jor-aw-ə (< -əl)
 mosque.Dsm they.AG build.PT.IPFV.3sm
 ‘they were building a mosque’;

tâse as-una bâzar ta bow-əl?
 you.p.AG horses-Dpm bazaar to take.PT.IPFV.3pm
 ‘were you taking the horses to the market?’

3.2.7.2b Perfective past

wə-rased-e

The perfective past expresses the action or event as a completed whole. See also Syntax for its use in 5.5.1 temporal and 5.5.2 conditional clauses.

(1) Intransitive

xədz-e jor-e sw-e pohantun ta lâ-r-u
 women-pf recover-pf AUX.PT.PFV-3pf university to go.PT.PFV-1p
 ‘the women recovered’; ‘we went to the university’;

asad wə-gad-ed-ə
 Asad.sm PFV.dance.PT.PFV-3sm
 ‘Asad danced’.

(2) Transitive

laylâ khpəl-e koran-əy ta pays-e wə-leg-əl-e
 Layla.AG own family to money-Dpf PFV.send.PT.PFV-3pf
 ‘Layla sent money to her family’;

mâjət ye jor kr
 mosque.Dsm they.AG build.sm AUX.PT.PFV.3sm
 ‘they built a mosque’;

tâse as-una bâzar ta botl-ə?
 you.p.AG horses-Dpm bazaar to take.PT.PFV-3pm
 ‘did you take the horses to the market?’.

3.2.7.3 Perfect constructions

3.2.7.3a Present perfect

rased-əl-ay ye

The present perfect expresses an action the result of which continues in the present.

(1) Intransitive

de asad plâr pakistân ta tl-əl-ay da
 of Asad father.sm Pakistan to go- PART-IPFV-3sm be.PR.IPFV.3s
 ‘Asad’s father has gone to Pakistan’.

(2) Transitive

laylâ lik leg-əl-ay da
 Layla.AG letter.Dsm send.PART-IPFV-sm be.PR.IPFV.3s
 ‘Layla has sent the letter’;

tor tsawk-əy jor-a kər-e da
 Tor.AG chair.Dsf build-sf PART.PFV-3sf be.PR.IPFV.3s
 ‘Tor has built a chair’.

3.2.7.3b *Past perfect*

rasedəl-ay we

The past perfect expresses a remote action, or an anterior action in the past.

(1) Intransitive

de asad plâr pakistân ta tl-əl-ay w-ə
 of Asad father.Dsm Pakistan to go- PART.PFV-sm be.PT.IPFV-3sm
 ‘Asad’s father had gone to Pakistan’.

(2) Transitive

Tor kor jor kər-ay w-ə
 Tor.AG house.Dsm build.sm PART.PFV-sm be.PT.IPFV-3s
 ‘Tor had built a house’.

3.2.7.3c *Future perfect*

ba rasedəl-ay ye ‘will have gone’

(1) Intransitive

de asad plâr ba pakistân ta tl-əl-ay w-ə
 of Asad father.Dsm FUT Pakistan to go.PART.PFV-3sm be.PT.IPFV-3sm
 ‘Asad’s father will have gone to Pakistan’.

(2) Transitive

tor ba tsawk-əy jor-a kər-e w-a
 Tor.AG FUT chair.Dsf build-sf PART.PFV-sf be.PT.IPFV-3sf
 ‘Tor will have built a chair’.

3.2.7.3d *Conjectural perfect*

rasedəl-ay wi

Similar to the perfective present as a subjunctive, the perfect subjunctive expresses conjecture and assumption, but of a completed action or event. While occurring mostly in dependent clauses, it may be used in main clauses as well, as in this example:

khalək dālta nâst [w-i]
 people here seated be.SUBJ
 ‘people sit here’, lit. ‘people may sit/be sitting here’ (we can use this table).

Compare the use of *wi* in the following (see section 7 Sample Text):

aw sar ye pə larg-ə wər-wə-pox-ə,
 and head its with wood there.PFV.cover.PT.PFV.3s
che lə bâd aw bârân-a pək-e pə amân [wi]
 that from wind and rain blow in safety be.SUBJ

‘and he covered it over with wood planking, so that he could live there protected from the wind and rain’.

For the use of the present and past conditional forms, see section 5.5.2 Conditional clauses.

3.2.8 Modal constructions

3.2.8.1 Ability ‘can’

3.2.8.1a Present ability

ras-ed-əl-ay + *s-*, all tenses and moods.

Note that the auxiliary is identical with the perfective forms of ‘be, become’.

(1) Intransitive

asad xə gaḏedəl-ay s-i
Asad well dance-PART-IPFV be.AUX.PR.PFV-3s
‘Asad can dance very well’;

pohantun ta tləl-ay s-əy
university to go-PART-IPFV be.AUX.PR.PFV-2p
‘you can go to the university’.

(2) Transitive

kor alta jor-awəl-ay s-e
house there build-PART-IPFV be.AUX.PR.PFV-2s
‘you can build a house there’;

kor ta me bowəl-ay s-e?
home to me take-PART-IPFV be.AUX.PR.PFV-2s
‘can you take me home?’

3.2.8.1b Past ability

(1) Intransitive

mung nən sahâr pə bâgh ke-har-chera tləl-ay s-w-u
we today morning in orchard in any-where go.PART.IPFV be.AUX.PT.PFV-1p
‘we could go anywhere in the orchard this morning’.

(2) Transitive

laylâ lik wə-legəl-ay s-ə
Layla.AG letter.Dsm PFV.send.PART.PFV be.AUX.PT.PFV-3sm
‘Layla could send the letter’.

3.2.8.1c Future ability

(1) Intransitive

asad ha xə wə-gaḏedəl-ay s-i
Asad FUT well PFV.dance-PART.PFV be.AUX.PR.PFV-3s
‘Asad will be able to dance very well’.

(2) Transitive

laylâ ba lik sabâ ta wə-legəl-ay s-i
 Layla **FUT** letter tomorrow to PFV.send-PART.PFV be.AUX.PR.PFV-3s
 'Layla will be able to send the letter tomorrow';

alta ba kor jor kəɹ-ay s-e
 there **FUT** house build PART.PFV be.AUX.PR.PFV-2s
 'you will be able to build a house there';

kor ta ba me botl-ay s-e?
 house to **FUT** me take.PART.PFV be.AUX.PR.PFV-2s
 'will you be able to take me home?'

3.2.8.2 Necessity and assumption, 'must'

de + perfective present stem, and related modal words or expressions.

The particle *de* expresses both strong and weak necessity as well as assumption, and accordingly governs the potential verb form, i.e. the perfective present. In addition, adjectives and nouns with similar semantic connotations are used.

(1) *de*

(1a) Intransitive

day de wə-gaɖ-eg-i
 he **must** dance.PFV.PR.PFV-3s
 'he must dance';

laylâ de kor ta lâɹ-a s-i
 Layla **must** house to go-sf be.AUX.PR.PFV-3s
 'Layla must go home'.

(1b) Transitive

duy de kor sarək tanizde jor k-i
 they **must** house.sm road toclose build.sm do.AUX.PR.PFV-3p
 'they must build the house close to the road';

asad me de kor ta bəz-i
 Asad me **must** house to take.PR.PFV-3s
 'Asad must take me home'.

(2) *zarur* 'necessary' + 'be'

pə paxt-an-o ke dâ zarur nə da che melmə wə-bal-ə
 in Pashtuns in this **necessary** NEG is **CONJ** guest PFV.invite.INF
s-i
 be.PR.PFV-3p
 'among Pashtuns it is not necessary that a guest be invited'.

3.2.8.3 Obligation, 'should'

Obligation is expressed by *bâyad* + perfective present.

(1) Intransitive

mung hám báyad wə-gəḏ-eg-u
we also **should** PFV.dance.PR.PFV-1p
'we should dance too';

laylá báyad kor ta lâṛ-a s-i
Layla **should** house to go-sf be.AUX.PR.PFV-3s
'Layla should go home'.

(2) Transitive

duy báyad nəw-ay kor jor k-i
they **should** new house,sm build.sm do.AUX.PR.PFV-3p
'they should build a new house';

mashum-ân báyad tsok mel-e ta bəz-i
children **should** someone picnic to take.PR.PFV-3s
'someone should take the children to the picnic'.

3.2.9 Passive constructions

Infinitive + *keg-lked-əl-* in all tenses and moods

This construction is formed with the infinitive plus the tenses and moods of the auxiliary *keg-lked-əl-*, which is derived from the verb 'make, do'. See also the passive-like examples of the derivative class under 3.2.7.1b and 3.2.7.2a.

(1) Present

day waz-əl k-eg-i
he kill.INF become.AUX.PR.IPFV-3s
'he is getting killed';

as məlgər-i ta bax-əl k-eg-i
horse friend to give.INF become.AUX.PR.IPFV-3s
'one's horse is given only to friends';

sar-ay pə bad-a wradz ke pəzand-əl k-eg-i
man on bad day on know.INF become.AUX.PR.IPFV-3s
'a true man makes himself known in difficult times'(proverb),
lit. 'a true man becomes known . . .'

(2) Past

pə pəxawər ke njun-e təxt-aw-əl k-ed-e
in Peshwar in girls-3pf kidnap.INF become.AUX.PT.IPFV-3p
'girls used to be kidnapped in Peshawar'.

3.2.10 *Infinitives as subject and object*

Infinitives may function as nominalized short clauses. As indicated under 3.1.1.1 Gender, they count as masculine in gender, and as plural in number.

(1) Subject

bowəl der grân fil day
take.INF very difficult verb be.PR.3pm
'*bowəl* is a very difficult verb';

de laylâ katəl sar-ay wazhn-i
of Layla look.INF man kill.PR.IPFV-3pm
'Layla's *gaze* kills a man';

tləl me stər-ay kaw-i
go.INF me tired-D3sm do.PR.IPFV-3pm
'*traveling* makes me tired'.

(2) Object

wayəl-o ta ye mə gor-a, kawəl-o ta ye gor-a
talk.INF-Opm to his PROH look-IMP2s do.INF-Opm to his look-IMP2s
'don't look at his *words*, look at his *actions*'.

4 SYNTAX AND USE OF FORMS

4.1 Noun phrase structure

The regular sequence of the constituent parts in the noun phrase is left-branching:

- (1) Adjectives precede the noun, ADJ – N: *spîn moṭar* 'white car'.
- (2) In the extended noun phrase, demonstratives, quantifiers, and adjectives, precede the noun, in that order, and adverbs precede the adjective; the normal sequence is thus:
DEM – QUANT – (ADV – ADJ) – N.
- (3) Adpositions precede and/or follow the noun phrase.

4.1.1 *Agreement*

Adjectives and demonstrative pronouns in a noun phrase, and predicate adjectives, agree with the nouns they modify in gender, number, and case. If the noun phrase consists of both masculine and feminine genders, the adjective is masculine.

4.1.2 *Dependent noun phrases*

- (1) Dependent possessive noun phrases, marked by the preposition *de*, precede the head noun phrase, NP₂ – NP₁: *de asad kitâb-un-a* 'Asad's books'.
- (2) Demonstratives and adjectives remain with their noun phrase:

[*de asad (agha loy kitâb-un-a)*]
 of Asad those big books
 ‘those big books of Asad’, lit. ‘Asad’s those big books’.

- (3) Similarly, prepositions similarly remain within their noun phrase:

[*de asad de plâr (lâ tsalor-o ðer-o xâysta luṅ-o sara)*]
 of Asad of father with four very pretty daughters with)
 ‘with Asad’s father’s four very pretty daughters’.

4.1.3 Topicalization

The unmarked order may be violated when the speaker wishes to focus on one or the other of the modifiers, in which case the emphasized element tends to be in first position. For example, in the following extended noun phrase, with multiple embeddings, the locative phrase [*pə kəli ke*] ‘in the village’ precedes the first possessive phrase [*de tor*] ‘of Tor’, and is also far removed from its referent *kor* ‘house’ at the end:

[*(pə kəli ke) de tor de tarikh de xowunk-i de məshər wror*
 in village in of Tor of history of teacher of older brother
xâysta (kor)
 beautiful [house]
 ‘Tor’s history teacher’s older brother’s beautiful house in the village.’

4.2 Clause structure and word order

4.2.1 Basic word order

The basic word order in independent, coordinate, and dependent clauses is SUBJECT–OBJECT–VERB (S–O–V):

S	O	V	
<i>Asad</i>	<i>lik</i>	<i>leg-i</i>	‘Asad is sending the letter’
<i>tə</i>	<i>tsə kâr</i>	<i>k-aw-e?</i>	‘what work do you do?’

Adverbs and adverbial phrases of time are usually clause initial, followed by locational and directional phrases, which may be close to the verb depending on the semantics of the latter. However, if in a prepositional phrase the object is a proclitic pronoun, the phrase immediately precedes the verb, as shown in the following examples:

(*parun*) *me (lâ məlgər-o sara) pə rasturân ke kabâb*
 yesterday I.AG with friends with in restaurant in kabob-sm
wə-khwŕ-ə
 PFV.eat.PT.PFV-3sm
 ‘I ate kebob at the restaurant with my friends yesterday’.

Proclitic

(*parun*) *me pə rasturân ke kabâb (wər sara) wə-khwŕ-ə*
 yesterday I.AG in restaurant in kabob-sm them with PFV.eat.PT.PFV-3sm
 ‘I ate kebob at the restaurant [with them] yesterday’.

4.2.2 *Split-ergativity and agreement*

Whenever person is marked by personal endings, enclitic pronouns must drop.

With intransitive verbs, the personal endings agree with the subject in person, number, and gender in both present and past tense, and agree also with the gender of subjects in past tenses.

With transitive verbs, the pattern differs from the intransitive in past tenses, where the personal endings agree with the object in person, number, and also in gender with 3rd person objects.

(1) Nominal subject and object

<p>Present: <i>saɾ-ay maŋ-a khwr-i</i> man-Dsm apple-Dsf eat.PR.IPFV-3s 'the man is eating the apple';</p> <p><i>xədz-a ʔikr-ay akhl-i</i> woman-Dfs scarf-sm buy. PR.IPFV-3s 'the woman is buying the scarf';</p>	<p>Past: <i>saɾ-i maŋ-a khwar-əl-a</i> man-OAGsm apple-Dsf eat.PT.IPFV -3sf 'the man was eating the apple';</p> <p><i>xədz-e ʔikr-ay akhist-ə</i> woman-OAGsf scarf-Dsm buy. PT.IPFV-3sm 'the woman was buying the scarf'.</p>
--	--

(2) Nominal subject, enclitic object

<p>Present: <i>zmar-ay ye rā-wal-i</i> Zmaray-Dsm 3s bring.PR.IPFV-3s 'Zmaray is bringing her';</p>	<p>Past: <i>zmar-i rā-wast-əl-a</i> Zmaray-OAG-sm bring.PT.IPFV-3sf 'Zmaray was bringing her'.</p>
--	---

(3) Ending and enclitic as subject and object, post-verbal

<p>Present: <i>khartsaw-u ye</i> sell.PR.IPFV-1p 3s 'we're selling it';</p>	<p>Past: <i>khartsawəl-a mo</i> sell. PT.IPFV -3sf we.AG 'we were selling it'.</p>
--	---

(4) Derivative verbs

In past tenses of derivative verbs whose first element is an adjective, the adjective agrees with the object of the verb if there is one, and with the subject if there is no object, in number and gender.

In the tenses formed with the perfect participle, the participle agrees in gender and number with the object of the sentence if there is one, and with the subject otherwise.

If the noun phrase consists of both masculine and feminine genders, the verb is masculine.

4.3 Rules of cliticization

Besides the ergative construction, the rules for the position of the enclitic pronouns and of particles, and their order, are probably the most interesting aspect of Pashto grammar (Tegey 1975a, 1975b, 1976, 1978).

Enclitic pronouns in any syntactic function are always attached after the first stressed part in the noun phrase: (*plâr*) *me* ‘my father’.

4.3.1 Sentence

The same rule holds for the sentence, where the first element may be one of a number of grammatical elements, ranging from short or extended noun phrases, to particles, to successive verbal prefixes in the absence of other elements, and to a single verb form, as is shown in the following examples:

after the subject:	<i>khoshal khân me plâr da</i>	‘Khoshal Khan is my father’;
after a perfective marker:	<i>wə me pezând-ə</i>	‘I recognized him’;
after a negative particle:	<i>nə me pezân-i</i>	‘he doesn’t know me’;
after verb:	<i>râ-leg-ə me</i>	‘I was sending it here’.

By this rule, the pronoun may be considerably separated from the referent noun:

(a) after an adverb:

(*wrusta lo plâr na*) *me râ-gh-ay*
 after from father from **my** arrive.PT.PFV-3sm
 ‘he arrived after **my** father’.

(b) after a direct object:

(*de plâr bâgh*) *me akhl-i*
 of father orchard **my** buy.PR.IPFV-3sm
 ‘he is buying **my** father’s orchard’.

4.3.2 Order of particles

The rule of enclitic attachment to the first part of the sentence also applies to various particles. These include *kho* ‘but, well, then, at least, maybe’, the future marker *ba*, and the necessitative *de* ‘must, should’.

The various particles and enclitics appear in the following order, with 1st person pronoun preceding the general 3rd person pronoun:

kho + ba + 1p melmo + de + 3rd ye.

This rigid ordering can lead to ambiguity. For example, the sentence below may be understood in two ways: *me* is either (a) possessive, or it is (b) the direct object, in which case *de* is the possessive, attached to the first stressed part of the sentence, while *me* is the object. In most cases the context or use of independent pronouns solves the problem, here with the pronominal clitics 1s *me* and 2s *de* in hierarchical sequence:

plâr kho me de leg-i
 father maybe my/me you/your send.PR.IPFV-3sm
 (a) ‘maybe [my] father is sending [you]’; or
 (b) ‘maybe [your] father is sending [me]’.

4.4 Conjunctions

4.4.1 Agreement

Nouns are conjoined in Pashto most simply by means of the conjunctions *aw* ‘and’ or *yâ* ‘or’.

- (1) When one of the conjoined elements is a pronoun, the pronoun ordinarily comes first.
- (2) When nouns conjoined with *aw* ‘and’ or *yâ* ‘or’ have the same gender, adjectives modifying both nouns agree with them in whatever the gender is, and are in the plural:

xkwəl-e njəl-əy aw xədz-a
pretty-pf girl-sf and woman-sf
‘pretty girl and woman’.

- (3) When the nouns are of different genders, the adjective must be repeated, with endings in agreement with each noun:

zɾawər-a xədz-a aw zɾawər sər-ay
brave-sf woman-sf and brave-sm man-sm
‘brave woman and man’.

- (4) Verbs in sentences with conjoined subjects, or objects are in the plural when coordinated by *aw* ‘and’. When the nouns are of different genders, the verb is masculine. This is certainly the case in the Central dialect, even though there has apparently been a prescriptive tradition in the language that the verb agrees with gender of the last nominal:

amân khoshâl aw laylâ mel-e ta lâŕ-əl
Aman.sm Khoshal.sm and Layla.sf picnic went.PT.PFV.3pm
‘Aman, Khoshal and Layla went on the picnic’.

- (5) When nouns are connected by *yâ* ‘or’, the verb agrees with the noun closest to it:

amân yâ laylâ wə-gađ-ed-a
Aman.sm or Layla.sf danced.PT.PFV-3sf
‘Aman or Layla danced’.

4.4.2 Reduplicative conjunctions

Words and phrases may be conjoined with reduplicative conjunctions, for example *yâ* . . . *yâ* ‘either . . . or’, *nə* . . . *nə* ‘neither . . . nor’ and *həm* . . . *aw həm* ‘both . . . and’. In this case, the second occurrence of the verb is avoided (“gapped”):

amân həm târikh poh day aw həm likwal
Aman also historian COP.3s and also writer
‘Aman is both a historian and a writer’;

nə *âsha* *x-a* *da*, *nə* *gwəla*;
 not Asha good be.PR.IPFV.3s NEG Gwela;
duy *de* *wə-khwr-i* *yaw* *tər* *bəl-a!*
 they MAY PFV.eat.PR.PFV-3p one toward other
 ‘neither Asha nor Gwela is good; may they destroy each other!’ (proverb)
 other.

4.5 Questions

Yes–no questions are differentiated from statements solely by the rise of the voice at the end of the sentence.

statement: *patang dâktar da*↓ ‘Patang is a doctor’;
 question: *patang dâktar da*↑ ‘is Patang a doctor?’

Questions pertaining to particular parts of the sentence retain the regular word order, and the question word occurs in the position of the respective part of speech (for question words, see 3.1.3.4 Indefinite Pronouns):

asad kâr *k-aw-i* Asad work do.PR.IPFV-3s ‘Asad is doing work’;
asad tsə *k-aw-i?* Asad what do.PR.IPFV-3s ‘what is Asad doing?’

5 COMPLEX SENTENCES

Noun clauses in Pashto are syntactically identical with corresponding independent sentences, and are usually introduced by the conjunctions *tse* or *che*. They function as subjects, direct objects, complements, or adpositional adverbial phrases.

Grammatically they are masculine plural, which reflects their status as abstract subjects and objects, which is evident from the verb forms preceding subject and object clauses. Especially longer noun clauses are often post-posed to the end of the sentence, with a demonstrative marker left in their place.

Adverbial clauses include causal, temporal, and conditional types. Combinations of tense and aspect frequently play a major role in these clauses.

5.1 Relative clauses

Relative clauses usually follow the head noun they modify and are introduced by *che* or *tse*. The definite head noun must be marked by a demonstrative.

Within the relative clause, the head noun is referred to either by the personal ending or the enclitic pronoun if subject or object, depending on the transitivity of the verb and the tense, and in prepositional phrases by the enclitic pronoun. This includes the possessive relationship:

(1) Possessor

agh-a njəl-əy tse *kamis ye* . . .
that girl that dress her . . .
 ‘that girl whose dress . . .’, lit. ‘that girl that her dress . . .’.

(2) Subject

agh-a njəl-əy tse melmastyâ ta râ-gh-əl-e w-a, de râbyâ khor da
that girl **that** party to come.PART-3sf was-3sf, of Rabya sister is
 ‘the girl **who** came to the party is Râbyâ’s sister’, lit. ‘had come’.

(3) Ergative patient

agh-a njəl-əy tse pə melmâstyâ ke me wə-lid-a, de râbyâ khor da
that girl **that** at party at I. AG PFV.see.PT.PFV-3sf, of Rabyasister is
 ‘the girl **whom** (3sf -a) I saw at the party is Râbyâ’s sister’.

(4) Adpositional

agh-a njəl-əy tse wər sara nâst w-əm, de râbyâ khor da
that girl **that** her with sitting was-1s, of Rabya sister is
 ‘the girl **with whom** I was sitting is Râbyâ’s sister’.

5.2 Noun clauses*5.2.1 Subject clauses*

The verb of the main clause is (1) in the 3rd person masculine plural, or (2) in the masculine singular in the presence of a demonstrative. The subject clause is usually post-posed to the end of the sentence, either with or without a demonstrative marker.

(1) Without demonstrative

x-ə sw-əl tse wakht-a râ-ghl-e
 good-3pm be.PT.PFV.3pm **that** early come.PT.PFV-2s
 ‘it was good that you arrived early’.

(2) With topical demonstrative

dâ x-ə swə tse wakht-a râ-ghl-e
 that.3sm good-3sm be.PT.PFV.3sm **that** early come.PT.PFV-2s
 ‘it was good that you arrived early’.

5.2.2 Object clauses

In past tenses, the verb of the main clause is in the masculine plural, agreeing with the clause in the normal ergative construction.

(1) Present

wây-i tse de asad plâr x-ə sar-ay da
 say.PR.IPFV-3p **that** of Asad father good-sm man-m be.PR.IPFV.3s
 ‘they say that Asad’s father is a good man’.

(2) Past ergative

mâ kat-əl che asad râ-ghl-ay
 I.AG saw.PT.PFV.3pm **that** Asad here-come.PART-3sm
 ‘I saw that Asad had arrived’, lit. ‘has arrived’.

5.2.2.1 Quotations as object clauses

Quotations as object clauses are mostly direct speech. The main clause verb agrees with the clause in past tenses, and the subject in present tenses.

(1) Present

laylâ amân ta way-i che mel-e ta ba âshak wes-əm
 Layla Aman to say.PR.IPFV-3s **that** picnic to FUT aushak take.PR.PFV-1s
 ‘Layla is telling Aman that she will take aushak to the picnic’, lit. ‘I will take’.

(2) Past

laylâ amân ta wə-way-əl che mel-e ta ba âshak wes-əm
 Layla Aman to say.PT.PFV.3pm **that** picnic to FUT aushak take.PR.PFV-1s
 ‘Layla told Aman that she would take aushak to the picnic’, lit. ‘I will take’;

amân wə-way-əl tse bəla myâst ba pakistân ta lar
 Aman PFV-say.PT.PFV.3pm **that** next month FUT Pakistan to go.PFV
 s-əm
 be.PR.PFV-1s
 ‘Amân said that he would go to Pakistan next month’, lit. ‘I will go’.

5.2.2.2 Infinitival direct object clauses

Many direct object noun clauses translate into other languages as infinitives. In fact, a sentence like the following can be heard in the Pashto variety spoken in Pakistan, possibly due to influence from Urdu or English.

khoshal kâr kawəl nə ghwâr-i
 Khoshal work do.INF not want.PR.IPFV-3s
 ‘Khoshal doesn’t want to work’.

In the Central dialect, however, such clauses must be finite noun clauses with *tse* and a perfective verb (subjunctive):

khoshal nə ghwâr-i tse kâr wə-k-i
 Khoshal not want.PR.IPFV-3s **that** work PFV.do.PR.PFV-3s
 ‘Khoshal doesn’t want to work’, lit. ‘that he work’;

khoshâl ghwâr-i tse plâr ye as w-âkhl-i
 Khoshal want.PR.PFV-3s **that** father his horse PFV.buy.PR.PFV-3s
 ‘Khoshal wants his father to buy a horse’ lit. ‘that his father buy a horse’;

tarisâ koshish wə-kṛ-ə
 Theresa.AG attempt PFV.make.PT.PFV.3pm
tse âshak pâkh-ə kṛ-i
that aushak cook AUX.PR.PFV-3p
 ‘Theresa tried to cook aushak’ lit. ‘that she cook’, note 3p endings.

5.3 Complement clauses

Complement clauses may be pointed to by a demonstrative pronoun in the main clause.

- (1) Without demonstrative

omed day che sol-a ba rā-s-i
hope is **that** peace FUT here-come.PR.PFV-3s
'the hope is that peace will come'.

- (2) With demonstrative

pə paxt-an-o ke dā zarur nə da che melm-a wə-bal-əl
in Pashtuns in **this** necessary not is **that** guest invited-INF
s-i
be.PR.PFV-3p (PASS)
'among Pashtuns it is not necessary that a guest be invited'.

5.3.1 Clauses as objects of adpositions

These clauses are almost always positioned after the adposition, but occasionally at the end of the sentence, with demonstrative markers left in place.

- (1) 'about', *di de . . . pə bâra ke*

mung di de pə bâra ke che tə kâr rā sara wə-kɾ-e,
we **of this on subject on that** you work me with PFV.do.PR.PFV-2s,
wə-ghag-ed-u
PFV.talk.PT.PFV-1p
'we talked about your working with me', lit. 'that you work'.

- (2) 'instead of . . . ' *di de . . . pə dzây + PR.PFV/(subjunctive)*

di de pə dzây che lə ahmad sara kâr wə-kɾ-e,
of this in place that with Ahmad with work PFV.do.PR.PFV-2s
mâ sara ye wə-kɾ-a
me with it PFV.do.IMP-2s
'instead of doing the work with Ahmad, do it with me', lit. 'that you work'.

5.4 Adverbial clauses

5.4.1 Causal clauses

Causal relationship between clauses is overtly coordinate, with the causal clause unmarked in the Central dialects.

The result clause is frequently unmarked as well in the Central dialect. It can also be introduced by: (a) *no* 'then, so', which is also often found in the apodosis of conditional clauses, or (b) by *dzəka* 'therefore', or (c) combined *dzəka no*, *no dzəka*, or (d) *dzəka che* particularly in literary, and in Western Pashto, calqued after Persian patterns.

day bur-a nə khwr-i, tâwân wər ta kaw-i
he sugar NEG eat.PT.PFV loss him to do.PR.IPFV-3
'he doesn't eat sugar [because] it's not good for him';

de asad garmi ked-a *no wob-ə ye* *wə-tskəl-e*
of Asad heat become.PT.IPFV-3sm **so** water he.AG. PFV3 drink.PT.PFV-3pf
‘Asad felt hot, so he drank some water’;

tāse *wos de ta tsə zyân nə-sh-əy* *rasawəl-ay*
you-p now him to any harm NEG-can.PR.IPFV-2p, bring.IPFV.PART
dzəka che *day zmā porawər-ay day*
because that he my debtor be.PR.IPFV.3s
aw mâ ta ye *panâ rā-wər-e* *da*
and me to he.AG asylum here-bring.PART.PFV-3pf be.PR.IPFV.3s
‘none of you can harm him because he owes me and has come to me for asylum’
(Ulfat); with transitive verb *panâ wər-* ‘take one’s shelter to’.

5.4.2 Temporal clauses

Temporal clauses are among the most common subordinate clauses in Pashto. They are introduced by *kəla che* or *che* ‘when’, or by an enclitic *che*.

- (1) ‘when’, *kəla che*

kəla che *laylā de asad pə aksident khabər-a* *siv-a,*
when Layla of Asad with accident informed-sf become.AUX.PT.PFV-sf
samdasti *roghtun ta lâṛ-a*
immediately hospital to go.PT.PFV-3sf
‘when Layla heard about Asad’s accident, she immediately went to the hospital’;

laylā che *rā-ghl-a,* *ḍer wakht wər sara* *ke-nâst-əm*
Layla **when** come.PT.PFV-sf, much time her with down-sit.PT.PFV-1s
‘when Layla came I spent a lot of time with her’.

- (2) ‘as soon as’, *tsənga che* + PR.PFV

tsənga che *kâṛ ta lâṛ* *s-e,*
as soon as city to go.PFV be.AUX.PR.PFV-2s,
wər sara wə-ghag-eg-a
him with PFV.talk.PR.IMP-2s
‘talk to him as soon as you go to the city!’

- (3) ‘until’, *tər tso che* + PR.IPFV

tər tso che *asad nəw-ay moṭar akhl-i,*
until Asad new car buy.PR.IPFV-3s,
de âmân moṭar gərdz-aw-əl-ay s-i
of Aman car use.PART-3sm can.PR.PFV-3s
‘until Asad buys a new car, he can use Aman’s’.

- (4) ‘after’, *wrusta lə* + PT.PFV

wrusta lə *də che âshak mo* *wə-khwər-əl,* *wə-gaḍ-ed-u*
after **from that that** aushak we.AG PFV.eat.PT.PFV-3p, PFV.dance.PT.PFV-1p
‘after we ate the aushak, we danced’.

5.4.3 Conditional clauses

5.4.3.1 Overview

The protasis of conditional clauses is introduced by *ka* 'if'; the apodosis may be introduced by *no* 'then, so'. Tense and aspect determine the degree of reality.

The *conditions* in the protasis are expressed by two sets of conditions:

- (1) potential, either (a) as a supposition or (b) assumed to be completed before the result; They are expressed by the perfective present and the perfective past: *wə-raseg-e* and *wə-rasedəl-e*, respectively.
- (2) irreal, referring either to (a) the present, or (b) the past. These two are expressed by two periphrastic perfect forms (see section 3.2.3.3 and Table 13.20). Their vocalism in the Central dialect differs slightly from that in the other dialects:

	Irreal in present	Irreal in past
Central dialects:	<i>rasedəlay</i>	<i>rasedəlay way</i>
Other dialects:	<i>rasedəlay</i>	<i>rasedəlay wāy</i> (invariant)

In the *apodosis*, the results are expressed by three sets, two of them marked by the future particle *ba*:

- (1) the present result of the two potential conditions is expressed by the present future *ba wə-raseg-e*;
- (2) the present result of the present and past irreal conditions is expressed by the imperfective past form, *rasedəl-e*;
- (3) the past result of the same two irreal conditions is expressed by the future past irreal form: *ba rasedəlay way*.

TABLE 13.22: PROTASIS AND APODOSIS IN CONDITIONAL CLAUSES

Real/potential				
‘if’	Protasis	Apodosis	Condition	Result
<i>ka</i>	<i>wə-raseg-e</i>	<i>ba wə-raseg-e</i>	Present	Present
<i>ka</i>	<i>wə-rasedəl-e</i>	<i>ba wə-raseg-e</i>	Completed	Present
Unreal, result in present				
	Protasis	Apodosis	Condition	Result
<i>ka</i>	<i>rasedəlay</i>	<i>rasedəl-e</i>	Present	Present
<i>ka</i>	<i>rasedəlay way</i>	<i>rasedəl-e</i>	Past	Present
Unreal, result in past				
	Protasis	Apodosis	Condition	Result
<i>ka</i>	<i>rasedəlay</i>	<i>ba rasedəlay way</i>	Present	Past
<i>ka</i>	<i>rasedəlay way</i>	<i>ba rasedəlay way</i>	Past	Past

5.4.3.2 Examples of conditional clauses

Note that past constructions require the ergative construction with past transitive verbs.

(1) Potential conditions, result in present

(a) supposed condition:

Present perfective	Present perfective
<i>ka day pays-e paydâ k-i,</i>	<i>samdastî ba ye wə-lagaw-i</i>
if he money-p find AUX	immediately FUT it spend.
PR.PFV-3s	PFV.PR.PFV-3s

‘if he finds some money, he will spend it immediately’.

(b) assumed completed condition:

Past perfective	Present perfective
<i>ka də pays-e paydâ kṛ-e,</i>	<i>samdastî ba ye wə-lagaw-i</i>
if he.AG money-pf find AUX	immediately FUT it spend.
PT.PFV-3pf-ERG	PFV.PR.PFV-3s

‘if he finds (as soon as he finds) some money, he will spend it immediately’.

(2) Unreal conditions, result in present

(a) present condition:

<i>ka amân aw laylâ kor ta rā-tl-əl-ay,</i>	<i>pəs-ə mo akhist-ə</i>
if Aman and Layla house to come.	lamb-sm we.AG buy.
PART.IPFV	PT.IPFV-3sm (ergative)

‘if Aman and Layla came to the house, we would buy a lamb’.

(b) past condition:

<i>ka amân aw laylâ har wakht kor ta rā-tl-əl-ay,</i>	
if Aman and Layla each time house to come.	PART.IPFV
<i>kili ba mo war-kəṛ-e way</i>	
key-sf FUT we.AG give would	
PART.IPFV-sf	AUX

‘if Aman and Layla often came to the house, we would have given them a key’.

(3) Unreal condition, result in past

(a) present condition:

<i>ka zə nārōgh-a nə way, nən sīnf ta tl-əl-əm</i>
if I sick NEG would be today class to go.
PT.IPFV-1s

‘if I weren’t sick, I would go to class today’.

(b) past condition:

<i>ka tâ way-əl-i way, ke-nâst-əl-e ba way</i>	
if you.s.AG ask would, down-sit. FUT would	
PART.IPFV-pm-ERG	PART.IPFV-sf

‘if you had asked, I would have sat down’ (female speaker).

Note that the 3s patient ending *-i* in this ergatively constructed conditional clause refers to the implicit object ‘(ask) a question’.

5.4.4 Consecutive clauses

Consecutive clauses frequently occur with anticipatory adverbs like *dumra* 'so' and *dāse* 'such', and are introduced by *che* or *tse*.

(1) *che* 'so that, in order to'

zə dā kitāb khpəl wror ta war-k-aw-əm, che wə ye hwal-i
I this book own brother to give.PR.IPFV-1s that PFV it read.PR-3s
'I'm giving this book to my brother to read', lit. 'that he read';

dā kitāb me khpəl wror ta war-kr-ə che wə ye hwal-i
this book I.AG own brother to give.PR.PFV-sm, that PFV it read.PFV-3s
'I gave this book to my brother to read'.

(2) *dumra* . . . *che* 'so much . . . that'

kəla kəla dumra zyāt-a wāwr-a wor-eg-i,
sometimes so much heavy snow fall.PR.IPFV-3sf,
che sarak-ın-a bikhi band-eg-i
that streets completely block.PR.IPFV-3p
'sometimes it snows so much that the streets are completely blocked'.

(3) *dāse* . . . *che* 'such . . . that'

dā dāse kitāb day, che hwast-əl ye grān di
this such book is.m, that read.INF it difficult be.PR.IPFV.3p
'this book is such that it is hard to read'.

6 DIALECTOLOGY, LEXIS AND SOCIOLINGUISTIC ASPECTS

6.1 Sub-groupings of dialects

The dialects and sub-dialects of Pashto have not been classified or studied to any great extent, and most Pashtuns themselves are sensitive only to the obvious differences in pronunciation and vocabulary. There are three major dialects of Pashto:

- the Western (Kandahar) dialect;
- the Central (Kabul) dialect;
- the Eastern (Ningrahar) dialect;

(for a four-fold division see Kieffer 1974; Skjærvø 1989: 387).

Speakers of the Western dialect are found mostly in south-west Afghanistan and in Baluchistan; speakers of the Central dialect mostly in the Kabul, Logar, Ghazni and Parwan provinces; speakers of the Eastern dialect in the north-east sections of Afghanistan, and in the Northwest Frontier Province of Pakistan.

The dialectal differences among these major dialect groups of Pashto are relatively minor, and the speakers of most dialects of Pashto are readily understood by almost all speakers of other dialects, as many Pashtuns have large passive vocabularies of words from other dialects. The most distinctive isoglosses are the following which reflect different diachronic development and the present phonetic realization of three central phonemes (note that Central *x* is a voiceless palatal/front velar fricative):

	Western	Central	Eastern
	<i>zh</i>	<i>z</i>	<i>j</i> [dzh]
retrofl.	<i>ḡh</i>	<i>g</i>	<i>g</i>
retrofl.	<i>ṣh</i>	<i>x</i>	<i>kh</i>

The variation between *ṣh* ~ *x* ~ *kh* is the source of the different Western spellings of the name of the language.

6.2 Contacts

In some respects, Pashto is perhaps the most conservative among the Iranian languages. Thus, it has retained phonemic stress, the ergative construction, and gender distinction. But Pashto has also innovated both internally, and by contact, for which the retroflex consonants due to contact with Indian languages are a prominent example.

6.3 Lexis

Within Iranian, Pashto shares great numbers of words with Dari-Persian, including shared borrowings from Arabic, not only because the genetic relationship of the two languages, but also because Pashtuns and Dari speakers have been neighbors for centuries, and more recently have been citizens of the same country, where both are official languages. By geographic contact, the Western and Central dialects (spoken in Afghanistan) are most likely to borrow vocabulary from Persian (and its Arabic component), whereas the Eastern dialect (spoken mostly in Pakistan) is more likely to borrow vocabulary from Urdu (and its English component).

7 SAMPLE TEXT

Note that the Pashto variant of this sample text by the author M. Pâsanay has *sh-lshw-* vs. the Central auxiliary *s-lsw-* ‘be, become’ found in the variant described in the present chapter.

de wux pəx-a
of camel.m leg-f
The Camel’s Legs

kəla-che (de ārchi pə daxta) wəb-ə
time-that (of Archi on plain) water-pf
rā-spar-e shw-e,
here-distribute.PART-pf be.PT.PFV-3pf
‘When the Archi Plain was irrigated,’

aw de dzmək-o wesh payl shu
and of land-Opf distribution started become.PT.PFV.3sm
‘and land distribution began,’

no hāji sāyib həm (ders jərib-a dzmək-a)
then Haji Sayib.AG also (thirty jarib land-f)
tər lās-a kr-a.
in hand-to do.PT.PFV-3sf
‘Haji Sayib acquired thirty jaribs of land.’

kaða ye wər-wastəl-a
 belongings his there-take.PT.PFV-3sf
aw (pə khpəl-a dzmək-a) ye wārawəl.
 and (to own-f land-f) he.AG move.PT.PFV.3sm
 'He moved his household and belongings onto his land.'

tsəngə-che (lə yaw-e khwa) wakht ləg w-ə
 since (from one-Osf side) time short be.PT-3sm
 'Because on the one hand time was short,'

aw (lə bəl-e khwa)
 and (from other-Osf side)
 'and on the other hand'

(pə ārchi ke) (de khatgər-o pata) nə laged-a,
 (on Archi on) (of masons-Opm trace) NEG find.PT.IPFV-3sf
 'there were no masons available in Archi,'

no (lə nāchār-əy) ye
 then (from desperation-Osf) he.AG
(de nor-o nāqel-in-o) pə tser
 (of other-Opm settlers-Opm) in keeping
 'he, like the other settlers, in desperation'

samlāsi dzmək-a zhawər-a kr-a
 immediate ground-PAT-sf deep-sf make.PT.PFV-3sf
 'immediately dug a hole in the ground'

aw sar ye (pə larg-o) wər-wə-pox-ə
 and head its (with wood-Opm) there-PFV-cover.PFV-3sm
 'and covered it over with wood planking,'

che (lə bād aw bārân-a) pək-e pə amân w-i.
 that (from wind and rain-from) blow-Osf in safety be.SBJ-3sm
 'so that he could live there protected from the wind and rain.'

hâji sâyib kisa kawəl-a
 Haji Sayib.AG story.PAT tell.PT.IPFV-3sf
 'Haji Sayib told this story,'

aw wayəl ye,
 and say.PT.IPFV.3pm he.AG
 saying,'

yaw-a wradz (pə khpəl-e dəgh-e khun-e ke)
 one-f day-f (in own-f that-f room-f in)
nâst w-əm.
 sitting.3sm] be.PT-1s
 'One day I was sitting in my room.'

yaw-nâtsâ-pa (lə pâsa) (khâwr-e aw lut-e)
 suddenly (from overhead) (dust-pf and clay-pf)
 'Suddenly dust and clods of earth'

râ-toy-e *shw-e*
 here-pour.PART.PFV-3pf become.PT.PFV-3pf
 ‘poured down from overhead.’

bar me wə-katəl *tsə (wugd-ə shay-ân)*
 up I.AG PFV-see.PT.PFV-pm that (long-pm things-pm)
 ‘I looked up to see long things’

râ-wə-dzaredəl *khwadzed-ə*
 here-PFV-hangdown.3pm move.PT.IPFV.3pm
 ‘hanging down, moving,’

aw (də khun-e bâm) (ye lə khwadzed-o sara)
 and (of room-Osf roof) (their with movement-Op with)
wâr-khatâ
 move.PT.IPFV.3pm
 ‘and as they moved, the roof of the room moved.’

wâr-wə-wat-əm
 there-PFV-get.out-1s
katəl me che yaw ghaṭ wux
 see.PT.PFV.3pm I.AG that one.m huge camel
 ‘I hastily got out of there, and on looking saw that a huge camel’,

(zung pə bâm) (de awâr-e dzmək-e) pə gumân
 (our on roof) (of flat-Osf ground-Osf) by guess
 ‘apparently thinking our roof was solid ground,’

râ-barâbar shəw-ay day
 here.to reach.PART-3sm be.PR.3sm
aw (de chat wâr-ə larg-i)
 and (of ceiling small-pm wood-pm)
 ‘had wandered onto it,’

(ye tər pəx-o lānde)
 (his under feet-Op under)
mât shəw-i di
 broken become.PART-pm be.PR.3pm
 ‘and the slats of the roof had broken under his feet.’

(tsalor wâr-a pəx-e ye)
 (four all-pf feet-pf his)
(tər bâm-a) watəl-i
 (through roof from) hang.PT-IPFV-3pf
 ‘All four of his legs were hanging down through the roof,’

aw nor (de khun-e pə shâtir)
 and other (of house-Osf with beam)
tam shəw-ay day
 supported become.PART-sm be.PR.3sm
 ‘but his body was supported by the roof beam.’

wux che wos har-tsumra dzân khvadzaw-i,
 camel that now however-many-times body moved-PR.IPFV-3sm
 ‘However the camel struggles,’

nə (yə pəx-e) (dzmæk-e ta) raseg-i;
 not (his feet-pf) (ground-Osf to) reach.PR.IPFV-3pf
 ‘his feet do not reach ground;’

nə (khpəl be-wâka leng-i)
 not (own un-controlled legs)
wər-ʔolawəl-ay sh-i
 there-collect.PART can.PR-3sm
 ‘he can neither collect his legs’

aw nə shâtir mâteg-i.
 and not beam break.PR-3sm
 ‘nor does he break the beam.’

(yaw-tsə shib-a) yə pə nandâraw-əm;
 (one-some moment) it at watch.PT.IPFV-1s
 ‘I watched for a while,’

tso che (nor de shâw-khwâ khalək) həm
 until (other of neighborhood people) also
khabar shw-əl
 news become.PT.PFV.3pm
 ‘then others in the neighborhood also got the news,’

aw (pə ǧer taklif)
 and (with great difficulty)
 ‘and with great difficulty’

mo wux khlâs kəʔ-ay shu.
 we.AG camel released make.PART can.PT.PFV.3sm
 ‘we managed to free the camel’.

ACKNOWLEDGEMENT

Our dear colleague Dr Habibullah Tegey died in 2005. I would like to acknowledge with gratitude the patience of his co-author, Dr Barbara Robson, with my suggestions for the arrangement of several sections.

[The Editor]

BIBLIOGRAPHY AND REFERENCES

- Bečka, J. (1969) *A Study in Pashto Stress*, Prague: Oriental Institute in Academia.
- Bellew, H. W. (1863) *A Grammar of the Pukhhto or Pukshto Language*; reprint Peshawar: Daeed Book Band and Subscription Agency.
- Elfenbein, J. (1997) 'Pashto phonology', in Kaye, A. S. (ed.) *Phonologies of Asia and Africa: Including the Caucasus 2*, Winona Lake, Indiana: Eisenbrauns, pp. 733–760.
- Ganiev, A. (1985) 'K probleme slozhnogo glagola va sovremennom afganskom iazyke (Pashto) v sviazi s issledovaniem glagol'noi frazeologii' (On the problem of the compound verbs in contemporary Pashto in connection with an investigation of the verbal phraseology), in *Iranskoe iazykoznanie. Ezhegodnik 1981*, Moskva, pp. 98–107.
- Heston, W. L. (1987) 'Pashto ambipositions and historical antecedents', in Bashir, E., Deshpande, M. M. and Hook, P. E. (eds.) *Select Papers from SALA-7: South Asian Languages Roundtable Conference held in Ann Arbor, Michigan, May 17–19, 1985*; Bloomington, Indiana: Indiana University Linguistics Club, pp. 163–181.
- Ijaz, M., 'Phonemic inventory of Pashto', http://crulp.org/Publication%5CCrulp_report%5CCR03_15E.pdf
- Kieffer, C. (1974) 'L'établissement des cartes phonétiques: premiers résultats', in *L'atlas linguistique des parlers iraniens: Atlas d'Afghanistan*, Bern: Universität Bern. Institut für Sprachwissenschaft. Arbeitspapiere 13, pp. 21–91.
- Lorimer, D. L. R. (1915). *Pashtu, Part 1: Syntax of Colloquial Pashtu, with Chapters on the Persian and Indian Elements in the Modern Language*, Oxford: Clarendon Press.
- Mackenzie, D. N. (1987) 'Pashto', in Comrie, B. (ed.) *The World's Major Languages*, London and Sydney: Oxford University Press, pp. 547–565; 2nd ed 2009, pp. 460–475.
- Meyer-Ingwersen, J.C. (1966) *Untersuchungen zum Satzbau des Pashto*, Ph.D. dissertation, Universität Hamburg.
- Morgenstierne, G. (1973) 'Traces of Indo-European accentuation in Pashto?', *NTS* 27: 61–65.
- (2003) *A New Etymological Vocabulary of Pashto*, compiled and ed. by Elfenbein, J., MacKenzie, D. N., and Sims-Williams, N., Wiesbaden: Reichert.
- Penzl, H. (1955) *A Grammar of Pashto: A Descriptive Study of the Dialect of Kandahar; Afghanistan*, Washington, D.C: American Council of Learned Societies.
- Rahman, T. (1995) 'The Pashto language movement and identity formation in Pakistan', *Journal of Contemporary Asia* 4.26: 151–70.
- Redard, G. (1974) 'État des travaux et publications. Quelques cartes onomasiologiques', in *L'atlas linguistique des parlers iraniens: Atlas d'Afghanistan*, Bern: Universität Bern. Institut für Sprachwissenschaft. Arbeitspapiere 13, pp. 7–19.
- Roberts, T. (1997a) 'The optimal second position in Pashto', in Booij, G. E. and Weijer, J. van de (eds.) *Phonology in Progress – Progress in Phonology HIL phonology papers III*, The Hague: Holland Academic Graphics, pp. 367–401.
- (1997b) 'Pashto free relatives and triply-filled Comp: evidence for a headed analysis', *Lingua* 102.2–3: 77–85; 21–53.
- Septfonds, D. (1994) 'La personne en pashto', *Actances* 8: 187–201.
- (1997) 'Constructions anti-impersonnelles en pashto', *SILTA* 26.2: 271–306.
- (2001) 'L'aspect en pashto: perfectivité et "éventuel"', *Actances* 11: 131–144.
- (2003) 'Marques personnelles et structures d'actances en pashto', *Actances* 12: 95–108.
- Shafeev, D. A. (1964) *A Short Grammatical Outline of Pashto*, translated from the Russian by H. H. Paper, The Hague: Mouton & Co.

- Skalmowski, W. (1996) 'Dative counterparts in Pashto', in Belle, W. van and Langendonck, W. van (eds) *The Dative*, Vol. 1. *Descriptive Studies*, Amsterdam: Benjamins, pp. 395–406.
- Skjaervo, P. O. (1989) 'Modern East Iranian languages', in *CLI*, pp. 370–83.
- (1989) 'Pashto,' in *CLI*, pp. 363–410.
- Tegey, H. (1975a) 'A study of Pashto clitics and implications for linguistic theory', *Studies in the Linguistic Sciences* (University of Illinois) 5.1, pp. 154–198.
- (1975b) 'The interaction of phonological and syntactic processes: examples from Pashto', in *Papers from the Eleventh Regional Meeting, Chicago Linguistic Society 1975*, pp. 571–582.
- (1976) 'Some problems concerning surface structure constraints and clitic placement rules', in *Papers from the Twelfth Regional Meeting, Chicago Linguistic Society*, pp. 628–639.
- (1978) *The Grammar of Clitics*, Kabul: International Centre for Pashto Studies.
- (1979) 'Ergativity in Pushto Afghan', in Rauch, I. and Caen, G. F. (eds) *Linguistic Method: Essays in Honor of Herbert Penzl*, The Hague: Mouton, pp. 369–418.
- Tegey, H. and Robson, B. (1990) *Beginning Pashto*, Washington, D.C.: Center for Applied Linguistics.
- (1991) *Intermediate Pashto*, Washington, D.C.: Center for Applied Linguistics.
- (1992) *A Pashto Reader*, Washington, D.C.: Center for Applied Linguistics.
- (1993) *Pashto Conversation Manual. Pushto Khabaray*, Washington, D.C.: Center for Applied Linguistics.
- (1996) *A Reference Grammar of Pashto*, Washington, D.C.: Center for Applied Linguistics.

Note: Dr. Habibullah Tegey has further numerous articles in Pashto and Dari on topics in Pashto language and literature, in such journals as *Kabul* (Journal of the Pashto Academy), *Wazhma* and *Adab* (published by the Faculty of Letters at Kabul University).

THE PAMIR LANGUAGES

D. (Joy) I. Edelman and Leila R. Dodykhudoeva

1 INTRODUCTION

1.1 Overview

“Pamir languages” is the generalized conventional term for a group of languages that belong to the eastern branch of the Iranian language family, and are spoken in the valleys of the western and southern Pamirs and adjacent regions: the Mountainous Badakhshan Autonomous Region (Tajik *Viloyati Kuhistoni Badakhshon*) of the Republic Tajikistan; the Badakhshan province in Afghanistan; parts of northern Pakistan (Chitral, Gilgit, Hunza); and parts of the Xinjiang-Uygur Autonomous Region of China.

The Pamir languages constitute four distinct genetic sub-groups that derive from several distinct proto-dialects of East Iranian origin (see also Chapters 14b and 15 on Shughn(an)i and Wakhi):

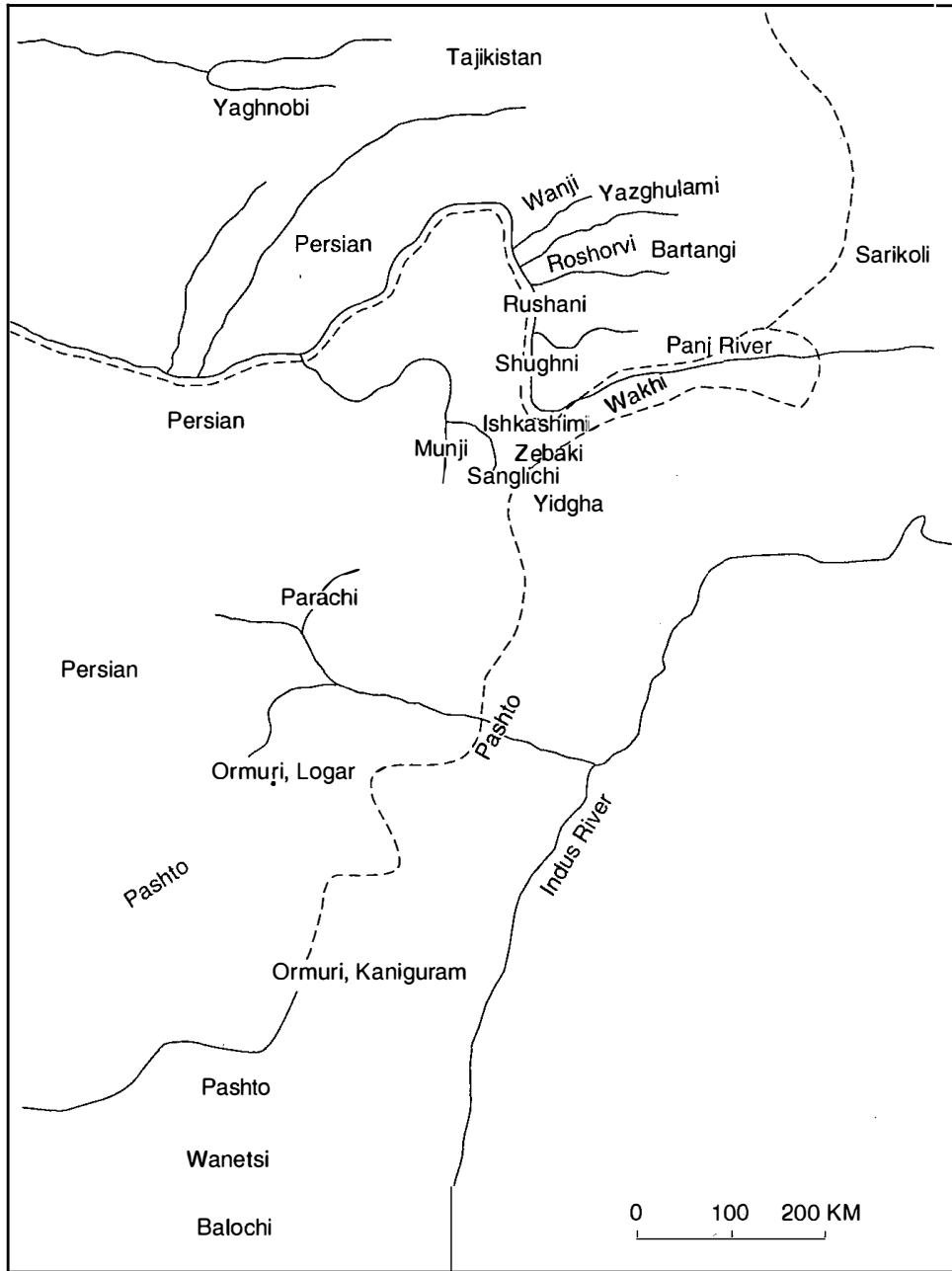
1. “North Pamir” group
 - (a) Old Wanji (extinct), relatively close to (b) Yazghulami, and (c) the Shughni-Rushani group to the south of it (see Chapter 14b).
2. Ishkashimi group
 - (a) Ishkashimi proper, (b) Sanglichi, (c) Zebaki (extinct).
3. Wakhi.
4. Also, owing to a series of features
 - (a) Munji, (b) Yidgha.

Extinct Sarghulami in Afghan Badakhshan is usually included. However, the very existence of this particular vernacular is doubtful. The material, described by Prof. I. I. Zarubin in the 1920s, could never be verified. It is based on the information from a speaker of one of the neighboring villages of Sarghulam, who called it *lavz-i mazor* ‘the speech of *mazār*’, presumably referring to the Afghan village of Sarghulam, which had such a shrine.

1.2 Social status and survival

The various groups of speakers of the Pamir languages consider themselves ethnic minorities. In official statistics and the census of Tajikistan they are not recognized as such, but identified as Tajiks of the Mountainous Badakhshan Autonomous Region.

The official language policy until 1989 did not recognize the independent status of these languages. It was only in 1989, with the “Tajikistan Republic Law on language”, that the existence and the special status of Pamir languages was acknowledged *de jure*.



MAP 14A.1 PAMIR LANGUAGES, PARACHI AND ORMURI

It stipulated special actions for “the independent development and usage of the Mountainous Badakhshanian (Pamir) languages”, but did not include official permission for writing in these languages. Thus, de facto the Pamir languages remain spoken languages only. In fact, one among them has gained the status of lingua franca, Shughn(an)i, at least in less official situations, beside official Tajik.

All Pamir languages can be called endangered languages to a certain extent. Some of these languages were included in the “Red Book of endangered languages of the world” (UNESCO 1995), *viz.* Rushani, Yazghulami and Ishkashimi. However, other languages, particularly those with fewer speakers, like Roshorvi, Bartangi, Sarikoli, Khufi and Sanglichi, are not included.

Moreover, a number of vernaculars are not well known to scholarship, or not described at all such as the vernaculars of Barwoz in Shughnan, Sawnob, and Bardara in Bartang. While some are still spoken, like Bajuwi and Sanglichi, others, like Barwoz, are apparently no longer used, and still others have been completely lost. Only in some cases, such as Old Wanji and Zebaki, there exist short preliminary descriptions.

The official figures given for the Mountainous Badakhshan Autonomous Region in the census of 1989 (last Soviet census) and the census of 1999 (the first sovereign Republic Tajikistan census after the migration caused by civil war) are as follows: 1989, total 154,554 (21,000, Khorog, Taj. Khorugh, administrative center of the region; 20,150, Wanj region including Yazghulam; 17,868, Rushan; 48,806, Shughnan; 19,287, Ishkashim; 11,843, Murghab). – 1999, total 211,292 (25,000, Khorog; 29,987, Wanj region including Yazghulam; 25,157, Rushan; 37,574, Shughnan; 23,915, Roshtkala; 27,265, Ishkashim; 15,885, Murghab; Darwaz, 25,952, not included).

In rough approximation, then, the total number of speakers of Pamir languages proper in Badakhshan can be estimated to be about 120,000 for 1989 and 170,000 for 1999.

2 LINGUISTIC SETTING AND DYNAMICS

2.1 Substrate and superstrate

The Pamir languages are the successors of several distinct ancient East Iranian dialects that penetrated the region, though probably not simultaneously. It is rather difficult to ascertain the time of divergence of this group from the other Iranian languages, but it most likely coincided with the period of the disintegration of the common East Iranian proto-language some time before the end of the first millennium BC.

Spreading to the Pamir valleys and adjacent regions, speakers of these emerging Iranian dialects formed layers over a common substratum of pre-Indo-European Pamir languages, or groups of languages. This substrate layer gave the immigrating East Iranian groups as a whole a set of shared characteristics on all linguistic levels, be it phonology, morphology, syntax, or lexical inventory. Significantly, the substrate features include, in comparison with other Iranian languages, shared shifts in the semantics not only of certain sets of terms, but also of the content of some grammatical categories. Thus, numerous structural parallelisms and shared loans provide the evidence for the continuous substratal coherence of this group, even though in the course of time, there developed internal divergences in phonetics and other characteristics due to the problem of communication between the speakers of these language groups in these mountainous regions.

In turn, during the last several centuries a new unifying layer was superimposed over all languages in the Pamir region by dialects of Fārsi, specifically Fārsi-ye Dari of Afghanistan, and Tajiki and its regional dialects. These gradually came to function as the language of oral communication, having been first introduced in the Middle Ages through the missionary activities of Islam. Today, Modern Literary Tajik is the official language in Tajik Badakhshan.

That the original area of the Pamir languages was considerably larger than today, particularly towards the north and north-west, can be demonstrated on the basis of documented but now extinct East Iranian languages as well as the toponymy and the areal and substrate lexicon. It is apparent that the expansion of the Tajik dialects occurred not only on the margins, but also into its inner parts due to forced migrations, as shown by the dialect of Ghoron and some Tajik dialects of Ishkashim. The various Tajik dialects with which closest contacts and connections developed include Darwaz, Wanj, Vakhiyo, Qarategin, Rogh, and Badakhshan in the North and West, while various Dari dialects provided contacts in the West.

Other contact languages in the larger Pamir area include Pashto, Burushaski and the Nuristani and Dardic languages in the south, and south-east, as well as Urdu, Kyrgyz and Uzbek dialects in the east and north-east along with Uyghur and Chinese.

2.2 Genetic relationships and Sprachbund

Comparative-historical research, prominently expounded in the works of Georg Morgenstierne in the 1930s, has demonstrated that the Pamir languages show divergent historical developments. They represent several groups each with its own distinct sets of characteristic features, including shared early innovations. Significantly, the features typical for each group individually are not shared with other groups inside the Pamir region proper, but with other East Iranian groups both within or outside of the larger region. That is, the Pamir languages constitute distinct genetic sub-groups that derive from several distinct proto-dialects of East Iranian origin. These facts contradict the assumption, still sometimes found, that the Pamir languages represent a unified family of related languages that derive from a common Iranian “pre-Pamirian” proto-language.

The degree of genetic relationship among the Pamir languages varies considerably. There is only one group that exhibits a genetic relationship on the basis of their historical phonology, morphology and shared innovations that is sufficiently close to allow for the reconstruction of a common inner-Pamir proto-language, or a group of very close dialects. This is the group which by some is called the “North Pamir” group, with three members: (1) extinct Old Wanji, relatively close to Yazghulami; (2) Yazghulami; and (3) the Shughni-Rushani group to the south of it. The reconstruction of their shared proto-language was partly undertaken by Valentina S. Sokolova (1967) and D. (Joy) I. Edelman (1980, 1986, 1987, 2009).

In turn, the Shughn(an)i-Rushani group itself (see Chapter 14b) consists of several local members that genetically constitute sub-groups, with various degrees of relationship: (a) Shughn(an)i-Bajuwi-Barwozi; (b) Rushani; (c) Khufi, (d) Bartangi-Roshorvi; and (e) Sarikoli. Further, each of these has its own internal subdialects and varieties, with clear geographical boundaries and mutually intelligible.

As indicated above, genetically related to the North Pamir group are: (4) Ishkashimi with Sanglichi and extinct Zebaki; further (5) Munji with Yidgha, which constitute closely related, but relatively independent, groups that share a set of features with the

others, but show sufficient number of substantial differences that makes a close relationship doubtful; finally, (6) Wakhi, which is genetically more distant, originating in a North-East Iranian dialect (see chapters on Khotanese and Tumshuqese, and Wakhi).

At the same time, there is convergence among the Pamir groups. It is evidenced by the fact that, typologically, there is a well-defined resemblance in structure (somewhat less so for Munji and Yidgha among the six groups) as well as shared vocabulary. However, those shared features are not due to close genetic relationship, but to the effect of convergence in the region. It is a process that was enhanced by the influence first of the common early substrate and then in the last few centuries by the penetration of Tajik. We are therefore justified in considering this region to be a “linguistic area” (Sprachbund).

Table 14a.1 combines the Cyrillic and Latin transcriptions developed for the alphabets of the individual Pamir languages.

3 MAIN LINGUISTIC CHARACTERISTICS AND TYPOLOGY

The main structural and typological characteristics (especially the more recent ones) as a whole are similar for all Pamir languages.

3.1 Phonology

3.1.1 *Vocalic systems*

The vocalic systems show various types of qualitative and quantitative opposition. The inherited Iranian opposition of length, $\bar{i} \bar{e} \bar{a} \bar{o} \bar{u}$ vs. $i a u$, is not retained as such. In general, length distinction is lost on the “margins”, but retained in the inner area. Thus, Shughni has a pair of low vowels \bar{a} and a , and also two series of three long vowels each which are apposed to a single short vowel, i and u , respectively, which have a wide range of phonetic realizations. Similar systems are found in the other members of the Shughni-Rushani group (except distant Sarikoli), and in Munji.

In Yazghulami, Ishkashimi and Sarikoli length distinctions are essentially lost. Distinctions in vowel quality are replaced by opposition of relative stability; thus, unstable central vowels are in opposition to all the other, stable vowels. Yazghulami, though, does retain length distinction in the low pair $\bar{a} a$. Wakhi has no quantitative gradation, and length distinction is not phonemic but conditioned by position, differing by the subdialect. The various systems are shown in Table 14a.2:

3.1.2 *Consonantal systems*

A common typical feature of the Pamir languages is the absence of an independent phoneme $/h/$, and the conditioned, non-phonemic “rough breathing” sound $[h]$ before word- or syllable-initial vowel.

3.1.2.1 Velar and uvular fricatives

In Yazghulami, the Shughni-Rushani group and in Wakhi the phonemic pair x and y has shifted to uvular position (like in Tajik), and the emptied velar place was occupied by the new phonemic velar pair \check{x} and \check{y} . In Ishkashimi, however, such velar phonemes did not develop. In Munji, x remained velar, and a new pair \check{x} and \check{y} developed from other sources.

TABLE 14a.1: TRANSCRIPTIONS FOR THE PAMIR LANGUAGES

C = Cyrillic, L = Latin

C	L	C	L	C	L	C	L
a	a	ж	ž	o	o	ц	c (ts)
ā	ā	ж	ž (W, I)	ō	ō	ч	č (tš)
б	b	з	z	ö	ö (B, Rsh)	ч	č (tš) (W, I)
в	v	з	j, ʒ (dz)	п	p	ч	č (dž)
б̄	w	и	i	р	r	ч	č (dž) (W)
г	g	й̄	ī	с	s	щ	š
ѓ	ǰ	й̄	y	т	t	щ	š (I, W)
г°	g° (Y)	к	k	ћ	θ	ь	ʔ (I i)
ѓ	ǰ	ќ	ќ (Y)	т	ʔ (I, W)	ы	ʔ (i) (I i)
ѓ°	ǰ° (Y)	к°	k° (Y)	у	u		
ђ	ǰ̄	к	q	ӯ	ū		
д	d	к°	q° (Y)	џ	ǰ		
ď	ǰ̄	л	l	ф	f		
д	ǰ (W, I)	л	l (I)	х	x		
е	e (Y, I)	м	m	х°	x° (Y)		
ē	ē	н	n	х̄	x̄		
ê	ê (Sh)			х̄°	x̄° (Y)		
э	ə (Y, W)			х	h		

B = Bartangi; I = Ishkashimi; Rsh = Rushani; S = Shughni; W = Wakhi; Y = Yazghulami.

TABLE 14a.2: PAMIR VOCALIC SYSTEMS

Long vs. Short

ī ē	ā	ō ū ū	Rushani (Shughnani-Rushani group)
i e	a	o u	
ī ē œ	ā	ō ū ū	Khufi (Shughnani-Rushani group)
i e	a	o u	
ī ē ê	ā	ō ū ū	Shughnani (Shughnani-Rushani group)
i e	a	u	
ī ē	ā	ō ō ū	Bartangi-Roshorvi (Shughnani-Rushani group)
i e	a	u	
ī ē	ā	ō ū ū	Munji
ə	a	u	
ī ē	ā	ō ū ū	Sanglichi
i e ə	a	o u	

Stable vs. Unstable*

i e *ə	ā	o ū u	Yazghulami
	a		
i e *b	a	o ū u	Ishkashimi
*i e *bl	a	o ū u	Sarikoli (Shughnani-Rushani group)
i *bl ə	*a	o *u	Wakhi of Wakhan proper

(For Wakhi, see also Chapter 15, section 2.1.1.)

3.1.2.2 Palatalized affricates and fricatives

In the North-Pamir languages **k* and **g* of the proto-language were conditionally palatalized. These variants developed into phonemic palatals in Yazghulami, and into phonemic *č* *ǰ* *ž* in the Shughn(an)i-Rushani group.

3.1.2.3 Labialized velars and uvulars

In Yazghulami, a series of six labialized phonemes developed as correlates of non-labialized phonemes. As a result, there is a triple post-palatal opposition (reminiscent of, but certainly distinct from, the triple Indo-European series): *k-k-k̰* and *g-g-g̰*.

3.1.2.4 Cerebrals

In Munji (with Yidgha), Ishkashimi (with Sanglichi and Zebaki), and Wakhi there developed sets of cerebral phonemes opposed to the unmarked sets varying by language group. These included the dental plosives stops, palatal affricates, and the sonorants. It is not only the very inventory of these phonemes, but more importantly the frequency of their use and their distribution (different frequency and use of classes) that show that they were not borrowed from the Indo-Aryan languages, but from a language like Burushaski. The conclusion that cerebralization in an intrusive process is further supported by the absence of cerebrals in the North-Pamir languages, i.e. Yazghulami and the Shughni-Rushani group. The various systems are shown in the following table, using digraphs for affricates.

TABLE 14a.3: PAMIR CONSONANTAL SYSTEMS

	Labial		Dental		Palatal		Velar		Uvular	Sonorant
Yazghulami	<i>p b</i>	<i>t d</i>	<i>ts</i>	<i>dz</i>	<i>tš</i>	<i>dž</i>	<i>k̰ g̰</i>	<i>k g</i>	<i>q</i>	
	<i>f v</i>	<i>θ ð</i>	<i>s</i>	<i>z</i>	<i>š</i>	<i>ž</i>	<i>k° g°</i>	<i>q°</i>		
							<i>č ǰ</i>	<i>x γ</i>	<i>x° γ°</i>	<i>w y m n r l</i>
Shughn(an)i-Rushani	<i>p b</i>	<i>t d</i>	<i>ts</i>	<i>dz</i>	<i>tš</i>	<i>dž</i>	<i>k</i>	<i>g</i>	<i>q</i>	
	<i>f v</i>	<i>θ ð</i>	<i>s</i>	<i>z</i>	<i>š</i>	<i>ž</i>	<i>č</i>	<i>ǰ</i>	<i>x γ</i>	<i>w y m n r l</i>
Ishkashimi	<i>p b</i>	<i>t d</i>	<i>ts</i>	<i>dz</i>	<i>tš</i>	<i>dž</i>	<i>k</i>	<i>g</i>	<i>q</i>	
		<i>t̪ d̪</i>			<i>(tš)</i>					
	<i>f v</i>		<i>s</i>	<i>z</i>	<i>š</i>	<i>ž</i>	<i>x</i>	<i>γ</i>		<i>w y m n r l</i>
					<i>ṣ̌</i>					<i>(l)</i>
Sanglichi	<i>p b</i>	<i>t d</i>	<i>ts</i>	<i>dz</i>	<i>tš</i>	<i>dž</i>	<i>k</i>	<i>g</i>	<i>q</i>	
		<i>t̪ d̪</i>			<i>tṣ̌</i>					
	<i>f v</i>	<i>θ ð</i>	<i>s</i>	<i>z</i>	<i>š</i>	<i>ž</i>	<i>č</i>	<i>ǰ</i>	<i>x γ</i>	<i>w y m n r l</i>
										<i>ŋ !</i>
										<i>ŋ</i>

Continued

TABLE 14a.3: *Continued*

	Labial Dental				Palatal		Velar		Uvular	Sonorant
Munji	<i>p b</i>	<i>t d</i> (<i>ʔ d</i>)	<i>ts</i>	<i>dz</i>	<i>tʃ</i> <i>tʃ̣</i>	<i>dʒ</i> <i>dʒ̣</i>	<i>k</i> <i>ḳ</i>	<i>g</i> <i>g̣</i>	(<i>q</i>)	
	<i>f v</i>		<i>s</i>	<i>z</i>	<i>ʃ</i> <i>ʃ̣</i>	<i>ʒ</i> <i>ʒ̣</i>	<i>x</i> <i>x̣</i>	<i>χ</i> <i>χ̣</i>	<i>x γ</i>	<i>w y m n r l</i> (<i>ŋ</i>)(<i>ʔ</i>)
Wakhi	<i>p b</i>	<i>t d</i> <i>ʔ d</i>	<i>ts</i>	<i>dz</i>	<i>tʃ</i> <i>tʃ̣</i>	<i>dʒ</i> <i>dʒ̣</i>	<i>k</i>	<i>g</i>	<i>q</i>	
	<i>f v</i>	<i>θ ð</i>	<i>s</i>	<i>z</i>	<i>ʃ</i> <i>ʃ̣</i>	<i>ʒ</i> <i>ʒ̣</i>	<i>x̣</i>	<i>ɣ̣</i>	<i>x γ</i>	<i>w y m n r l</i>

3.1.3 Suprasegmental features

The pattern of stress in nominals tends to be word-final, while that of the verb forms varies according to language, except for obligatory stress on the negative prefix. In general word stress is rather weak and subordinate to a sentence stress.

3.2 Morphology

For nominal morphology, the absence of distinct morphological boundaries between substantive, adjective and adverb is characteristic, e.g. Bartangi *tor* 'top; upwards; upper' may function in all three categories. That is, the function is largely determined on the basis of semantic and syntactic criteria.

3.2.1 Nominal morphology and categories

3.2.1.1 Gender

The category of gender, distinguishing masculine and feminine, is preserved in a rather reduced form in Yazghulami, the Shughni-Rushani group (with the exception of Sarikoli), and Munji (with Yidgha). In Yazghulami gender is exposed only by correlation with gender-marked 3s pronouns in the oblique case which derive from demonstrative pronouns. In the Shughn(an)i-Rushani group and Munji (with Yidgha) gender is formally expressed only in a small set of nouns. Otherwise, the gender of a noun is indicated by demonstrative pronouns, as well as by those adjectives and verb forms that have retained gender distinction. In Ishkashimi, Sanglichi, Wakhi and Sarikoli the category of gender is lost.

Semantically, one observes the tendency towards transformation of the category of gender to a system that is defined by semantic classes, evidencing a shift that is due to the substratum. The process is practically completed in Yazghulami. In the Shughni-Rushani group, the category of gender is partially shifted to the category of concreteness, concrete/general (abstract). Thus, irrespective of natural gender, names of inanimate objects, of animals (when sex is not specified), and objects appear in the masculine gender when they indicate the general, categorical idea, or the totality of objects, while a concrete object is in the feminine gender. For example, in Rushani *māwn* 'apple' has feminine

gender when referring to a single apple, but masculine gender when the totality of apples is referred to.

3.2.1.2 Number

The category of number in most Pamir languages is expressed by the opposition of the bare, unmarked noun (as an object, concept, or totality) and the noun with plural markers (distributed multitude). In most of the Pamir languages the markers are agglutinative suffixes. Munji, Wakhi, and Sarikoli distinguish plural direct and oblique forms. In Bartangi plural is indicated syntactically by the combination of the bare noun with demonstrative pronoun and the verbal number markers. In all Pamir languages, with the exception of Munji, the singular is required after numbers higher than two, which thus implies totality.

3.2.1.3 Case marking

Case marking is partially preserved in Munji (and Yidgha), Wakhi, and Sarikoli, in form of two main cases: direct and oblique, sometimes with the addition of adpositions. In the Shughn(an)i-Rushani group (excepting Sarikoli) the case (as well as gender and number) of the substantive is denoted by the case of a demonstrative pronoun that serves as a definite article. In Yazghulami and Ishkashimi nouns are not inflected, and their syntactic function is determined by position within the clause, as well as adpositions, adpositional phrases, and other markers such as adverbs.

3.2.1.4 Definiteness

The category of definiteness is marked morphologically by articles that have largely lost the semantic distinctions of their origin. The indefinite article is derived from the numeral 'one', usually in truncated form. Definiteness is indicated by articles that are inflectionally reduced sets of the demonstrative pronouns. The Shughni-Rushani group is noteworthy for having retained gender, number, and case distinctions in these demonstrative pronouns, which thereby serve not only as definite articles, but also to identify noun and noun phrases grammatically.

3.2.2 Adjectives

The inflection of adjectives is even more reduced than that of substantives. Gender is preserved only in the Shughn(an)i-Rushani group (with the exception of Sarikoli) and in Munji (with Yidgha); number is preserved in Munji (with Yidgha); case is lost everywhere.

Comparison is marked by reflexes of OIr. **-tar-a*, e.g. Sh. *-di*, Rsh. *-dōr*, Y. *-dūr*). There are also degrees of intensity, 'a little more', etc. expressed by either old or recent means such as reduplication, e.g. Sh. *-dar-di*, Y. *-dar-dūr* < **-tara-tara*. – The superlative is expressed descriptively.

3.2.3 Pronouns

In most of the languages the 3rd person pronoun is not developed and the demonstrative pronouns serve in this function. A special case is Yazghulamii, where historically

demonstrative sets were redistributed into personal pronoun function and demonstrative function. In general, the 1st and 2nd personal pronouns reflect the ancient personal pronouns, albeit with considerable and varying assimilations between the forms. The exception is found with the second person plural pronouns. They were innovated, on the model of substrate system, by copying initial *t-* from the 2s pronoun, thus, Shughni *māš* ‘we’ – *tama* ‘you’, Yazgh. *mox* – *təməx*. This is not the case in Munji (with Yidgha) and Wakhi, even though there the 2p personal pronouns are likewise innovative.

The demonstrative pronouns preserve (with the exception of Yazghulami) the ancient Indo-European system of triple deixis: (1) *ich*-deixis, nearest to speaker ‘this’ < **ima-*; (2) *du*-deixis, nearest to addressee (interlocutor) and not that far, ‘this, that’ < **aita-*, also used as emphatic and anaphoric pronouns ‘that very’, ‘that one that’, and (3) *er*-deixis, either distant or neutral, ‘that; he’ < **awa-*.

The declension of pronouns is relatively archaic, and is preserved even in those languages that have lost the declension of other nominal classes, as is the case in Yazghulami, the Shughn(an)i-Rushani group and Ishkāshimi.

3.2.4 Numerals

Numerals up to ten are usually indigenous. Higher numbers are often loaned from Tajik (in Sarikoli from Uygur), but the older generation tends to preserve indigenous composite numbers. The teens are additive, ‘ten (and) digit’. From 40 and above, Yazghulami, Munji (with Yidgha) and Wakhi have a vigesimal system; thus ‘twice 20’ = 40, ‘twice 20 and ten and five’ = 55, ‘four times 20’ = 80. By contrast, the system of the Shughni-Rushani group is decimal; thus ‘four times ten’ = 40, ‘five times ten and five’ = 55.

3.3 Verb morphology and system

3.3.1 Past tenses

The basic two subsystems are present and past, typical for most Iranian, and three stems, typical for this group. Present forms, present-future, and imperative, are derived from the present stems. In the past tenses, forms are based on the past stem, usually originating in the OIr. perfect participle in **-ta*, very rarely in **-wa*, in Wakhi also in **-na*. The perfect and pluperfect forms are based on the extended participle masc. **-taka*, fem. **-tačī* (with *i*-umlaut).

In the present tenses, person and number are marked by endings that reflect, on the whole, the OIr. “primary” active endings. In the past and perfect tenses, participles of intransitive verbs have gender and number agreement. Person and number in all Pamir languages are typically marked by enclitics that resulted from the partial merger of the OIr. enclitic pronouns with copula forms. By that merger the earlier ergative construction, where agents of past transitive verbs were expressed by the personal clitics, was mostly lost. However, the markers still show traces of that construction, e.g. in the Shughni-Rushani group (with the exception of Sarikoli), the 3s past marker *-i* with transitive verbs originates in the 3s personal enclitic **-hai*.

The various patterns of present and past person markers are shown in the following table. Parentheses indicate optional marking; note that past tense markers are clause clitics.

TABLE 14a.4: PAMIR PERSON MARKING IN PRESENT AND PAST

	1s	2s	3s	1p	2p	3p
<i>Yazghulami</i>						
Present	-in	-ay	-t/-	-əm	-it	-an
Past	əm	at	ay	an	əf	an
Trans.			(ay)			
<i>Shughni</i>						
Present	-um	-i	-t/-d	-ām	-ēt	-ēn
Past	um	i	∅	ām	ēt	ēn
Trans.			i			
<i>Rushani</i>						
Present	-um	-i	-t/-d	-am	-at/-af	-an
Past	um	at	∅	am	af	an
Trans.			i			
<i>Ishkashimi</i>						
Present	-bm	-i	-u	-on	-bv	-on
Past	bm	bt	i	on	bv	on
Sanglichi has the same endings, except for pres. 3s -∅						
<i>Munji</i>						
Present	-əm	-əy	-t/-d	-ām	-āf	-āt
Past	ām	āy	∅	ām	āf	āt
Trans.	əm	ət	a	ām	āf	āt
<i>Wakhi</i>						
Present	-əm	-i	-t/-d	-ən	-əv	-ən
Past	əm	ət	∅	ən	əv	ən
<i>Sarikōli</i>						
Present	-am	-∅	-t/-d	-an	-it	-in
Past	am	at	i	an	af	af

4 LEXIS

4.1 Word formation and semantic sets

All Pamir languages share the same word-building models. A noteworthy phenomenon is the contextually restricted use of certain sets of words and of acts, including taboos, e.g. the name of the wolf or mouse in different situations, as well as different kinds of acts of prohibition protected by taboo, such as Ishkashimi *pbčbrbm* which is the ritual prohibition to enter the house at Nawruz. In response to the need of new vocabulary items, new words with descriptive semantics or old words that already exist in some Pamir languages are used in new senses, or words are borrowed.

In general, the lexicon continues the inherited one, but also includes innovations and semantic shifts at various stages, dating from Indo-European to relatively recent periods. Thus, ancient is Yazghulami *doyd* ‘daughter’ < IE. **dhugō₂ter-*, whereas Rushni *razēn* ‘daughter’ < **fra-žanyā-* (fem.) ‘born’ is an early innovation.

4.2 Loans

The largest number of loan words in all Pamir languages comes from Tajik. Of interest are loans that have become obsolete in Tajik, as well as in Persian and Dari, though preserved in all or several Pamir languages, and sometimes show semantic shifts. There are also numerous terms loaned from Arabic via Tajik, etc. mostly in the religious and cultural spheres. Loans from Indo-Aryan and Burushaski generally are terms of material culture. There are also a small number of loans from Turkic, mainly from Kyrgyz, and from Uygur and Chinese in Sarikoli. More recent are loans from European languages, mainly Russian and English that were incorporated during a rather long period, most intensively from the end of nineteenth century onward.

There have also been some mutual borrowings among the Pamir languages. Certain lexical strata and sets of words are of substrate origin, or are areal words whose source or origin is unknown.

BIBLIOGRAPHY AND REFERENCES

See also Chapters 14b and 15 on Shughni and Wakhi.

- Bakhtibekov, T. (1979) *Grammatikai zaboni Shughnoni* (Grammar of Shughni language), Dushanbe: Donish.
- Dodykhudoev, R. Kh. (1971–1972) ‘Die Pamir Sprachen Zum Problem der Konvergenz’, *Mitteilungen des Instituts für Orientforschung* 17: 463–470.
- (1975) *Pamirskaiia mikrotoponimiia: Issledovanie i materialy* (Pamir microtoponymy: Introduction and materials), Gertsenberg, L. G. (ed.), Dushanbe: Irfon.
- Dodykhudoeva, L. R. (1997) ‘Sotsio-lingvističeskaia situatsiia v Badakhshane (The socio-linguistic situation in Badakhshan),’ in *Malye iazyki Evrazii. Sotsio-lingvističeskii aspekt* (Sociolinguistic situation in Badakhshan. Minor languages of Eurasia: sociolinguistic aspect), Moskva: MGU, Filologičeskii fakul’tet, pp. 78–89.
- (2000) ‘Pamirskie iazyki’ (The Pamir languages), in *Iazyki mira. Iranskie iazyki* III, pp. 170–174.
- Edelman, D. I. (1966) *Iazghuliamskii iazyk* (The Yazghulami language), Moskva: Nauka.
- (1971) ‘K voprosu o slovoobrazovanii mestoimenii v indoiranskikh iazykakh’ (On the problem of pronoun formation in the Indo-Iranian languages), in *Indiiskaia i iranskaia filologiia. Voprosy leksiki*. Moskva: Nauka, pp. 151–160.
- (1975) ‘K genezisu vigezimal’noi sistemy chislitel’nykh (On the genesis of the vigesimal system of numbers),’ *VIA* 1: 30–37.
- (1976) ‘K istorii iazguliamskikh i shughno-rushanskikh ukazatel’nykh mestoimenii’ (on the history of the Yazghulami and Shughni-Rushani demonstrative pronouns), in Rastorgueva, V. S. (ed.) *Iranskoe iazykoznanie. Istoriiia, etimologiia, tipologiia: k 75-letiiu prof. V.I. Abaeva*, Moskva: Nauka, pp. 85–96.
- (1980a) ‘History of the consonant systems of the North-Pamir languages’, *IJ* 22: 287–310.
- (1980b) ‘K substratnomu naslediiu tsentral’noaziatskogo iazykovogo soiuza’ (On the substratal heritage of the Central-Asian Sprachbund), *VIA* 5: 21–32.
- (1986; 1990; 2009). *Sravnitel’naia grammatika vostochnoiranskikh iazykov*, I. *Fonologiia*, II *Morfologiia*, *Elementy sintaksisa*, III, *Leksika* (Comparative grammar of the East-Iranian languages, I. Phonology; II Morphology, Elements of syntax, III, Lexis), Moskva: Nauka Vost lit-ra RAN.

- (1987a) ‘Shughnanskoi gruppа iazykov’ (The Shughni-Rushani language group) in *Osnovy* III.2, pp. 236–306.
- (1987b) ‘Iazghulamskii iazyk’ (The Yazghulami language), in *Osnovy* II.2, pp. 348–407.
- (2000a) ‘Shughnano-rushanskaia gruppа iazykov’ (The Shughni-Rushani language group) in *Iazyki mira. Iranskіe Iazyki* III, pp. 225–242.
- (2000b) ‘Iazghulamskii iazyk’ (The Yazghulami language) in *Iazyki mira* III, pp. 274–290.
- (2000c) ‘Vandzhskii iazyk’ (The Wanji language) in *Iazyki mira* III, pp. 57–58.
- Edelman, D. I. and Iusufbekov, Sh. P. (2000) ‘Sarikolskii iazyk’ (The Sarikoli language) in *Iazyki mira* III, pp. 269–273.
- Faizov, M. (1963) *Iazyk rushantsev Sovetskogo Pamira* (The language of the Rushanis of the Soviet Pamir), Dushanbe: Akademia nauk Tadzhikskoi SSR, Institut iazyka i literatury.
- Griunberg, A. L. (1972) *Iazyki Vostochnogo Gindukusha: Mundzhanskii iazyk* (The languages of Eastern Hindukush: The Munji language). Leningrad: Nauka.
- (2000) ‘Mundzhanskii iazyk’ (The Munji language) in *Iazyki mira. Iranskіe iazyki* III, pp. 154–169.
- Griunberg, A. L. and Steblin-Kamenskii, I. M. (1976) *Iazyki Vostochnogo Gindukusha: Vakhanskii iazyk* (The languages of Eastern Hindukush: The Wakhi language), Moskva: Nauka.
- Karamshoev, D. (1963) *Badjuvskii dialekt shugnanskogo iazyka* (The Bajuw dialect of the Shughnani language), *Trudy* 16, Dushanbe: Donish.
- Karamkhudoev, N. (1973) *Bartangskii iazyk* (The Bartangi language), Dushanbe: Donish.
- Kurbanov, Kh. (1976) *Roshorvskii iazyk* (The Roshorvi language), Dushanbe: Donish.
- Lentz, W. (1933) *Pamir-Dialekte. I. Materialien zur Kenntnis der Schughni Gruppe*, Göttingen: Vandenhoeck & Ruprecht.
- Mirzabdinova, S. (1983) *Shevai khufii zaboni rushoni* (The Khufi dialect of the Rushani language), Dushanbe: Donish.
- Morgenstierne, G. (1938) *Iranian Pamir Languages* = Morgenstierne, *IIFL* II.
- (1974) *Etymological Vocabulary of the Shughni Group*, Wiesbaden: Reichert.
- Pakhalina, T. N. (1959) *Ishkashimskii iazyk* (The Ishkashimi language), Moskva: Izd-vo Akademii nauk SSSR.
- (1966) *Sarikolskii iazyk* (The Sarikoli language), Moskva: Nauka.
- (1969) *Pamirskie iazyki* (The Pamir languages), Moskva: Nauka.
- (1983) *Issledovanie po sravitel’no-istoricheskoi fonetike pamirskikh iazykov* (Introduction to the comparative-historical phonetics of the Pamir languages), Moskva: Nauka.
- (1989) *Sravitel’no-istoricheskаia morfologіia Pamirskikh iazykov* (Comparative-historical morphology of the Pamir languages), Moskva: Nauka.
- Payne, J. R. (1989) ‘Pamir Languages’, in *CLI*, pp. 417–444.
- (1980) ‘The decay of ergativity in Pamir languages’, *Lingua* 51: 147–186.
- Sokolova, V. S. (1953) *Ocherki po fonetike iranskikh iazykov* (The phonetics of the Iranian languages). II, Moskva, Leningrad: Izd-vo Akadmii nauk SSSR.
- (1966) ‘Shughnano-rushanskaia gruppа iazykov’ (The Shughnani-Rushani language group) in *Iazyki narodov SSSR* (Languages of the people of USSR), I, Leningrad: Nauka, pp. 362–397.
- (1967) *Geneticheskie otnoshenia iazguliamnskogo iazyka i shugnanskoi iazykovoi gruppy* (Genetic relations between the Yazghulami language and the Shughnani language group), Leningrad: Nauka.

- (1973) *Geneticheskie otnoshenia mundž-anskogo iazyka i shughano-iazguliamskoi iazykovoï gruppy* (Genetic relations between the Munji language and the Shughnani-Yazghulami language group), Leningrad: Nauka.
- Yusufbekov, Sh. P., Dodykhudoeva, L. R. (2008) 'Sanglichskii iazyk' (The Sanglichi language) *Osnovy* VII., pp. 110–234.
- Zarubin, I. I. (1918) 'Materialy i zametki po etnografii gornyykh tadzhikov' in *Sbornik Muzeia Antropologii i Etnografii imeni Petra Velikago* (Petrograd), V: 97–148.

SHUGHNI

D. (Joy) I. Edelman and Leila R. Dodykhudoeva

1 INTRODUCTION

1.1 Overview

The Shughni, or Shughnani, ethnic group, ethnonym *xuŋnī*, *xuŋnīnī*, populates the mountain valleys of the West Pamir. Administratively, the Shughni-speaking area is part of the Mountainous Badakhshan Autonomous Region (Tajik *Viloyati Mukhtori Kuhistoni Badakhshon*) of the Republic of Tajikistan, with its major center of Khorog, Taj. Khorogh (37° 30'N, 71° 31'E), and of the adjacent Badakhshan Province of Afghanistan.

In Tajikistan, the Shughn(an)i live along the right bank of the longitudinal stretch of the river Panj from (Zewar) Dasht in the North to Darmorakht in the south, as well as along the valleys of its eastern tributaries, the Ghund (*Gund*, Tajik *Gunt*) and the Shahdara (*Xāšdarā*), which meet at Khorog. They also constitute the major population group in the high mountain valley of Baju(w)dara (*Baĵū(v)darā*) to the north of Khorog. Small, compact groups are also found in central Tajikistan, including Khatlon, Romit, Kofarnikhon, and other regions.

In Afghanistan, the Shughn(an)i have also compact settlements, mainly on the left bank of the river Panj in Badakhshan Province. A sizeable Shughn(an)i-speaking community is also found in Kabul (cf. Nawata 1979) and in Faizabad, the capital of Afghan Badakhshan.

Linguistically, the Shughn(an)i language, endonym (*xuŋn(īn)ī*, *xuŋn(īn)ī ziv*, *xuŋnīn ziv*), belongs to the Shughn(an)i-Rushani sub-group of the North Pamir languages. This group includes, in addition, Yazghulami as well as the now extinct Old Wanji in Tajikistan, relatively close to the former. These languages are genetically closely tied together, and it is quite possible to reconstruct the common North Pamir proto-language, defined by numerous shared innovations in historical phonetics and morphology, a task that has been undertaken by V. S. Sokolova (1967) and by D. (Joy) I. Edelman (1980, 1986, 1987, 1990, 2009). Altogether, the Pamir languages belong to the East Iranian branch of the Iranian language family (see Chapter 14a).

The Shughni-Rushani group itself consists of some seven members. Though varying in types and degree of relationships, four genetic sub-groups can be distinguished (see also 7 Dialectology):

- (1) Shughni, Bajuwi, and Barwozi;
- (2) Rushani with Khufi to the north;
- (3) Bartangi with Roshorvi to the north-east (widely known in its Kyrgyz form Oroshori);

(4) isolate Sarikoli spoken in the Xinjiang-Uygur province of China.

Specifically, the location of the members of the first group and their varieties of speech are as follows:

- (1) Shughni proper, spoken on the banks of the river Panj from Sokhcharv in the north to Darmorakht in the south, and in the provincial center Khorog;
- (2) the micro-dialect Bajuwī (*Bajūwī*) in the high Bajū(w)darā valley, with the micro-dialects of the villages of Bajū(w) and Bajū(w)-pastev north to Khorog;
- (3) the micro-dialects of Barwoz (*Barwoz*) in the highest part the Shahdara valley now confined to a few families and in the process of being replaced by the common variety;
- (4) in addition, there are the micro-dialects of the Ghund valley, and of the lower and upper villages of the Shahdara valley, specifically from Khorog to Tavdem, and from Tavdem to Jawshanghoz. The varieties in these two valleys are to a considerable degree due to the admixture of immigrant population from Rushan, Bartang, Roshorv, Sarez, and the Wakhan-speaking areas.

The speakers of Shughn(an)i consider themselves an ethnic minority in Tajikistan. The official statistics and the census record them simply as Tajiks of the Mountainous-Badakhshan Autonomous Region. The census of 1989 gave the following numbers for that region: total population 154, 554; Shughan district 48,806; the census of 1999 gave: Shughnan 37,574; Roshtqala 23,915; and Bajū(w) (which is included to Rushan region) 4,150; total 65,639; for the town of Khorog, ca. 21,000 and 25,557, respectively. By unofficial indirect estimate, the total number of the Shughn(an)i-speaking residents of historical Shughnan is ca. 80,000 and 100, 000.

Shughn(an)i is used mainly as a spoken language. As elsewhere in Tajikistan, Tajik is the official language. It is the only language that can be legally used for official purposes, and is mandatory in school, the mass media, etc. Accordingly, Tajik dominates even in verbal communication. For less formal purposes, however, oral communication may be conducted in the Shughn(an)i language.

In fact, Shughn(an)i serves as the de facto lingua franca for the entire Pamir linguistic area besides Tajik. In addition, a kind of regional Tajik vernacular has traditionally evolved in Badakhshan, which has at least three major inputs: It is based not only on the local dialects, and on the linguistic norms of official Soviet Tajik, which were formed not later than the 1980s and are now predominant in schools, but also on classical Farsi-Tajiki, which is popular in the region, as well as on contemporary Farsi and Dari norms.

1.2 Writing systems

Shughn(an)i, even today, is regarded as an oral language. However, in the 1930s an alphabet was developed based on Roman characters, soon after which several textbooks as well as a few works of fiction were published. For several years afterward, reading and writing instruction was conducted in Shughn(an)i in the elementary schools of the region. In the 1980s a new alphabet was created based on Cyrillic characters, with diacritics for specific Shughn(an)i phonemes (see Table 14a.1 in Chapter 14a).

2 PHONOLOGY

2.1 Inventory and distribution

2.1.1 Vowels

Vowels are characterized by qualitative and partly quantitative phonological oppositions:

TABLE 14b.1: SHUGHN(AN)I VOWELS

Qualitative Characteristics			Quantitative Characteristics			
	Front	Center	Back			
High	<i>ī</i>		<i>ū</i>	Long	<i>ī ē ê</i>	<i>ā</i> <i>ū ũ ḡ</i>
High Mid	<i>ē i</i>		<i>u ũ</i>	Short	<i>i</i>	<i>a</i> <i>u</i>
Mid	<i>ê</i>		<i>ḡ</i>			
Low	<i>a</i>		<i>ā</i>			

The three long phonemes *ū, ũ, ḡ* are opposed to the single short phoneme *u*, which has a wide range of phonetic variants [u, y, o, ɔ] corresponding in quality to the long three vowels. The same holds for long *ī, ē, ê* as opposed to the single *i* with variants [i, ɪ, e, ε, ɛ, ê]. The low vowels *a, ā* constitute a single pair. In general, long vowels are stable, and the degree of variation is rather low.

The phoneme *ũ*, moved forward compared to *ū*, is the result of, or may occur as a conditioned variant of the following: (1) long *ō* in prenasal position: *pōnd* > *pūnd* ‘road’; (2) *aw* in preconsonantal contraction: *sāw-d* > *sū-d* ‘goes, is going’; (3) short *u* with compensatory lengthening and loss of pharyngeal: *muhlat* > *mūlat* ‘(granted) period’. In addition, *ũ* occurs in expressive vocabulary, as in the interrogative emotive particle *ũ*, and the interjection *ũx̄xa*.

The trend towards monophthongization is well developed, and affects inherited, old and recent loans, such as *nawbat* > *nowbat* > *nūbat* ‘(one’s) turn’.

Similarly, the phoneme *ē* reflects the contraction of the diphthong **ai*: *sipēd* ‘white’. The phoneme *ê*, besides its occurrence in indigeneous words such as *nix̄ēb* ‘to make sleep’, reflects a further level of contraction, that of the diphthong *-ey*: *xeyr* > *xēr* ‘an emotional particle of surprise, also of endorsement’, and is also found in pre-uvular or pharyngeal position in loans: *tēy* ‘razor’.

Prenasal raising. In general, in most of the Shughni local dialects the long mid-high phonemes are raised before nasal to the next level of narrowness: *ê* > *ē* > *ī* and *ḡ* > *ũ* > *ū*: **divēn-* > *divēn-* ‘winnow’, **x̄ēn* > *x̄ūn* ‘blue’ and **nōm* > *nūm* ‘name’, **ȳūnj* > *ȳūnj* ‘hair’. Bajuwī and some Shughn(an)i micro-dialects, however, tend to retain the broader vowels: Sh. *mūn*, Bj. *mūn* ‘apple’. Nevertheless, vowels of one and the same word in Shughni and Bajuwī, while similar in their linguistic characteristics, may form different lexical variants.

Final *ī, i* and *ā, a*. (1) Long *ī* in final position is reduced to short *i*: *tūr* > *tī* ‘upwards, upper part’, *dī* > *di* ‘village’. (2a) Long *ā* in final pre-pausal position serves as an archiphoneme for both *a* and *ā*: *garḡa-y-um zoxt* ‘I took a flat bread’, but *garḡā* ‘flat bread’. (2b) In turn, short *a* may occur in final position as the result of truncation: *sāw* > *sa* ‘go!’, *a yidā* > *a ḡa* ‘address to a young man’.

2.1.2 Consonants

The system of Shughn(an)i consonants may be shown as follows:

TABLE 14b.2: CONSONANTS

	Labial	Dental	Palatal	Velar	Uvular	Glottal
Stops	<i>p</i> <i>b</i>	<i>t</i> <i>d</i>		<i>k</i> <i>g</i>	<i>q</i>	
Affricates		<i>c</i> (<i>ts</i>) <i>j</i> (<i>dz</i>)	<i>č</i> <i>ǰ</i>			
Fricatives	<i>f</i> <i>v</i>	<i>θ s</i> <i>ð z</i>	<i>š</i> <i>ž</i>	<i>x̣</i> <i>ɣ̣</i>	<i>x</i> <i>ɣ</i>	(<i>h</i>)
Glides	<i>w</i>		<i>y</i>			
Nasals	<i>m</i>	<i>n</i>				
Liquids		<i>r, l</i>				

2.1.2.1 Back-lingual *k g* and fronting

In Shughni proper the phonemes *k g* are always back-lingual. However, in the larger Shughn(an)i group, including Bajuwī, these two are palatalized to [k̟, g̟] under two conditions: before front vowels, and in final position after any vowel. In Bajuwī, pre-vocalic fronting occur in some other positions as well: *gāx̣t* [g̟ax̣t, j̟ax̣t] ‘appealed’, *kā* [k̟ā, čā] ‘where’.

2.1.2.2 Uvular *x ɣ* vs. velar *x̣ ɣ̣*

The uvular phonemes *x ɣ* are opposed to the velar pair *x̣ ɣ̣*. These are articulated with the back of the tongue raised high, while the remainder of the tongue remains flat, or its tip is lowered. Therefore, those sounds could be considered single-focus, or double-focus phonemes with pronounced front focus: *x̣ac* ‘water’, *wōx̣* ‘grass’.

2.1.2.3 Palatal *č ǰ* vs. dental *c j*

Among the affricates, the double-focus palatal pair *č ǰ* is opposed to the dental single-focus pair *c j*. The latter tends to soften, or lose, occlusion, and merge with *z*. In turn, the fricative *s* may be occluded to *c* after consonants: *wižafcl/s-* ‘to return’.

2.1.2.4 Labio-dental *v* vs. bilabial *w*

Labio-dental *v* is opposed to bilabial *w*. The latter is groove or slit depending on position. In word- or syllable-final position it appears as a semivowel resulting in phonetic diphthongs: *žōw* ‘cow’, *baǰū(w)* ‘Bajuw’.

2.1.2.5 Non-phonemic *h*

For Shughn(an)i, the absence of phonemic *h* is typical, although [h] occurs in recent loans from Tajik, and as an onset glide before initial vowel, alternating with [w] or [y]. This process thus contrasts with the truncation of final vowels or syllables mentioned above.

2.1.3 Consonant clusters

There are some positional alternations connected with historical processes or with the modern implementation of phonemes.

2.1.3.1 Clusters and structural assimilation

In general, there are no initial clusters. Final clusters in loans which deviate from permitted indigenous clusters are assimilated by release vowels: *umr* > *umri* 'life', *naql* > *naqli* 'narration'.

2.1.3.2 Metathesis

Further, metathesis of phonemes and syllables, which is distinctive for Iranian languages, particularly East Iranian, is an intriguing phenomenon in Shughni: *naḡq* < Tajik *naqš* 'ornament, ornamentation', *ċārḡ* 'wheel' < OIr. *caxra*.

2.1.3.3 Voice assimilation

Consonants show voice assimilation: *bad-qār* > *bat-qār* 'angry', but *tūd-piḡt* > *tūd-biḡt* 'mulberry flour', *tūd-paj* > *tūd-baj* 'the season when mulberry is ripe'.

2.1.3.4 Assimilation of *n*

The nasal *n* may have dental, palatal, and guttural varieties: *rang* [raŋg] 'color', *carang* [caraŋg] 'how'. Before palatalized [k̄, ġ], *n* could occur as post-alveolar palatalized sound.

2.1.3.5 $v > \delta, \theta > f$

In some words, *v* is changed to δ due to noise intensification: *vδidirm* 'besom'. In turn, the voiceless non-strident fricative θ may change to labial: Sh. *ḡirf* 'slippery', Barwozi *ḡirθ*, Sh. *ḡagarf* 'proper name' < *ḡagarθ* < *ḡagarδ* < *ḡogard* < Tajik *ḡogird* 'pupil'.

2.2 Non-segmental features

2.2.1 Word stress

- (1) Enclitics, including person markers, postfixes, postpositional *-i* (*izāfa*), are unstressed, or have no primary stress.
- (2) Otherwise, nominal stress is final, including final derivative suffixes.
- (3) Composite forms may have primary and secondary stress.
- (4) Similarly stress is on the final syllable of verbal stems, and always on the negative and prohibitive markers *na* and *mā*.

2.2.2 Phrasal stress

Word stress is rather weak and is subordinated to that of the sentence or clause. Parts of a syntagm are grouped around a word on which the phrase stress falls. Phrase stress is achieved by increasing the stress of one of the words included into the syntagm, which

has a slight rise on the stressed syllable. On the last syllable of the syntagm the tone rises sharply and simultaneously the vowel is lengthened.

2.2.3 Sentence stress and intonation

- (1) In declarative sentences, a syntagm, especially at the end of a phrase, has a sharp drop in the tone of the syllable on which the phrase stress rests.
- (2) Interrogative sentences have distinct interrogative intonation with rising tone, which is rather prominent in Shughni: *sāv-i-y-o?* 'are you going?'

2.3 Morphophonemic alternations

- (1) Like other Iranian languages, Shughni inherited the morphophonological alternations between present and past stems of the verb. In addition, morphological change is found in perfect and infinitive stems as well as in forms of the 3s.
- (2) A subset of nouns shows both gender and number alternation, and a subset of adjectives shows gender alternation.

3 MORPHOLOGY

Traditionally, the following basic parts of speech are distinguished: noun, pronoun, numerals, verb, preposition, postposition, particle, and conjunction. For nominal morphology, the absence of the morphological distinction between the word classes: noun, adjective, and adverb is typical, so that the identification of some words as belonging to one of these categories requires additional semantic and syntactic data.

Agglutination together with elements of internal inflection for nominal forms appears frequently, as is the case in verbal inflection in the present-future tense and the analytical structure of forms of the past tenses.

Modal and aspectual-temporal constructions are formed analytically.

3.1 Nominal morphology

Nominals possess the category distinctions between masculine and feminine gender, singular and plural number, definite and indefinite, as well as person and non-person, and alienable and inalienable possession (marked syntactically).

3.1.1 Nouns

3.1.1.1 Gender

3.1.1.1a Morphological gender

Nouns are either masculine or feminine. However, morphologically, this distinction is preserved only in a small set of substantives, animate nouns and adjectives, and is marked by ablaut: m. *kud*, f. *kid* 'dog'; m. *vūyd*, f. *vōyd* 'demonic creature'; m. *čux* 'cock', f. *čax* 'hen'.

Similarly, a small set of nouns functions as lexical gender markers in compounds, such as m. *-buc*, f. *-bic* 'child'; m. *-gil*, f. *-gāl* 'head'; m. *-vor*, f. *-vēr* 'a person bringing some-

thing’: m. *têr-gāl*, f. *têr-gāl* ‘with black hair’; m. *žīz-vōr*, f. *žīz-vêr* ‘a person bringing firewood’.

Similarly, natural gender is inherent in derived nouns such as *virod-ēj* ‘stepbrother’, *nān-ēj* ‘stepmother’, *xaray-ēj* ‘inhabitant of Khorog’, either male or female.

Otherwise, animate gender is expressed lexically: *čōrik* ‘man’, *ŷinik* ‘woman’.

3.1.1.1b Gender marking by attributives and verb forms

Nouns are not morphologically marked, but whether animate or inanimate, their gender becomes apparent by gender-marking attributive demonstratives, adjectives, and intransitive forms of the past tenses. Thus, in the adverbial phrase: *wi rūz-at wam xāb* ‘(for) that-OBLm day-and that-OBLf night’, the oblique pronouns masculine *wi* and feminine *wam* indicate the masculine and feminine gender of ‘day’ and ‘night’, respectively.

3.1.1.1c Classificatory gender

There is a tendency to transform the category of gender into a classificatory system according to the principles of semantic classes. Essentially, abstract nouns, such as *mōj* ‘famine’, are masculine, while concrete nouns are classified as masculine or feminine by semantic class, or cognitive sets. Thus, generally feminine are particular parts of the body, parts of clothes, and tools, as well as parts of landscape. The common denominator appears to be “body”, of both human and of earth, with its parts and cover, and tools to shape them. Also feminine are: *mêst* ‘moon’, *xāb* ‘night’, and *xitêrj* ‘star’, as opposed to masculine *xīr* ‘sun’.

Following this tendency, terms for inanimate objects and those for animals appear in the masculine gender when referring to the general category or species in their entirety, irrespective of natural gender. For example, masculine *yu rubcak* ‘fox’ refers to the fox species, while the feminine *yā rubcak* refers to a concrete object, here a particular fox irrespective of natural gender ‘this fox’ (in some situations, even without further indication of sex).

3.1.1.2 Number

Generally the singular is unmarked. The most productive and polyfunctional plural suffix is *-(y)-ēn* (in Shahdara *-jēv*, *-ēv*): *ŷinik-ēn* ‘women’, *čōrik-ēn* ‘men’, *māraka-y-ēn* ‘social gatherings’; [*sūr mēθ*]-*ēn* ‘wedding days’, [*žōw-at šag*]-*ēn* ‘cows-and calves’; with ablaut: sg. *čīd* ‘house’, pl. *čad-ēn*; sg. *puc* ‘son’, pl. *pac-ēn*. The marker *-jēv* is found in Shughni with terms denoting time periods: *tōbistūn-jēv* ‘summers’, *ar sōl buōr-jēv* ‘springs of every year’, *madōr-jēv* ‘afternoons’.

The kinship terminology uses specific suffixes side by side *-ēn*, including *-yūn*, *-gūn*, *-jūn*; and *-êrj*, *-ōrj*, *-ār*: *xolak-ēn* ~ *xolak-yūn* ‘uncles’; *xêr-ēn* ~ *xêr-yūn* ‘nephews, nieces’; *yax-ēn* ~ *yax-jēn* ~ *yax-jin-ēn* ‘sisters’; *abīn-ēn* ~ *abīn-êrj* ‘co-wives’ (Bajuwi *abīn-ōrj*).

The expression for plurality may be indicated by specific markers such as *-xēl* ‘group’, and *-gal(l)ā*: ‘flock’: *bač-galā* ‘children’, *vaz-xēl* ‘goats’.

Indefinite collectiveness may be expressed by using alliterative patterns of the form *c—m/p—*, or suffixed *-adis* (lit. ‘and such’): *tōč-mōč* ‘all sorts of dishes’; *tōč-adis* ‘dishes and such’.

Named human groups may be indicated by the direct and oblique 3p demonstrative pronouns, such as *wād*, *dād*, *mād*, and *wēv*, *dēv*, *mēv* ‘those, they’ added to the name, though usually pronouns of the 1st and 2nd series are used (see 14b.3 Table of pronouns):

Mādbēk-wāδ ‘Madbek and those connected with him, his group, family’ (cf. Persian [*Name*] *in-hā*). Similar meaning is expressed by the suffixes *-yūn* and *-ēn*: *Mādbēk-yūn*, *Mādbēk-ēn* ‘Madbek and those connected with him, his group, family’.

3.1.1.3 Case marking by demonstratives

Nouns do not distinguish case. The syntactic direct and oblique case in singular or plural of a noun or noun phrase is indicated by the case of demonstrative pronoun.

3.1.1.4 Definiteness

Definiteness is formally expressed by preposing forms of the remote demonstrative pronoun, serving as definite articles: m. *yu*, f. *yā*, and their corresponding oblique and plural forms: *yā yāc* ‘that girl’, *yu yīδā* ‘that boy’.

Indefiniteness is formally expressed by the cardinal number (*y*)*īw* ‘one’, usually in its reduced form (*y*)*i*, serving as an indefinite article: *tar yi jīngāl yō puχtā* ‘to a forest or a mountain plain’; cf. its use as cardinal number in: *yi mēθ . . .*, *wi yi-ga mēθ-ard* ‘for (-ard) one day . . ., (and) the other day . . .’

3.1.2 Adjectives and degree

Gender is still distinguished by ablaut in a small subset: m. *rūšt*, f. *rōšt* ‘red’; m. *tuχp*, f. *taχp* ‘sour’. There is no number distinction.

The *comparative* of qualitative adjectives is marked by *-di*. In addition, the comparison may be differentiated by degree: *-di-di*, approximately ‘much more’; and *-dar-di*, approximately ‘some, yet more’: *jāld* ‘fast’, *jāld-di* ‘faster’, *jāld-(d)ar-di* ‘more fast, faster yet’. The *elative* is expressed lexically by *lap* ‘very’: *lap xušrūy* ‘very beautiful’. Note *as kor yu gāp lap-di δī-d* ‘he talks more than he works’ (*lap δāδ-* ‘to chat, talk without result’).

The *superlative* is expressed adverbially by words and phrases such as *sar* ‘top; over’, *as fuk* ‘than all’, *bar fuk* ‘upon all’, added to the comparative form: *sar jāld-di* ‘most fast one’, *as fuk bašand-di* ‘best of all’, *bar fuk xušrūy-di* ‘most beautiful’.

3.1.3 Pronouns and deixis

Shughni distinguishes personal, demonstrative, reflexive-possessive, interrogative, and indefinite pronouns, which may combine with emphatic and negative components.

3.1.3.1 Personal pronouns

The personal pronouns occur for the 1st and the 2nd person; the 3rd person is indicated by demonstrative pronouns. Only the 1s has retained case distinction, *wuz/mu*. The 2p Sh. *tama* (Bartangi, Rushani, Sarikoli *tamaš*) is a historical innovation with initial *t-* copied from 2s *tu*, built on a substrate model.

3.1.3.2 Demonstrative pronouns and deixis

The demonstrative pronouns distinguish gender, and direct and oblique case in singular and plural. They preserve the ancient Indo-European triple deixis based on bicentric system that is defined by degree of distance relative to the object:

- 1) *ich*-deixis, proximate, nearest to the speaker, ‘this’;
- 2) *du*-deixis, referring to the sphere of communicator, ‘not that far’, ‘this’ ~ ‘that’;
- 3) *jener*-deixis, remote ‘that’. Similarly adverbials have triple deixis.

Du-deixis also serves as an emphatic and anaphoric: ‘that very’, ‘that which’, while *jener*-deixis serves as a definite article, as well as the 3rd person pronoun.

TABLE 14b.3: PRONOUNS AND DEICTIC ADVERBIALS

Person	1		3/Jener-Deixis		Du-Deixis		Ich-Deixis	
	m.	f.	m.	f.	m.	f.	m.	f.
DIRs	<i>wuz</i>	<i>tu</i>	<i>yu</i>	<i>yā</i>	<i>yid</i>	<i>yid</i>	<i>yam</i>	<i>yam</i>
OBLs	<i>mu</i>	<i>tu</i>	<i>wi</i>	<i>wam</i>	<i>di</i>	<i>dam</i>	<i>mi</i>	<i>mam</i>
DIRp	<i>māš</i>	<i>tama</i>	<i>wāḏ</i>		<i>dāḏ</i>		<i>māḏ</i>	
OBLp	<i>māš</i>	<i>tama</i>	<i>wēv</i>		<i>dēv</i>		<i>mēv</i>	
ADV			<i>yam</i> ‘there’ (far)		<i>yēd</i> ‘there’		<i>yūd</i> ‘here’	
-ard ‘for’			<i>yam-ard</i>		<i>yēd-ard</i>		<i>yūd-ard</i>	
<i>tar</i> ‘to’			<i>tar-am</i>		<i>tar-ēd</i>		<i>tar-ūd</i>	
<i>az</i> ‘from’			<i>az-am</i>		<i>az-ēd</i>		<i>az-ūd</i>	

The plurals of all personal and demonstrative pronouns may have secondary forms with the nominal plural marker *-ēn*: *māš-ēn*, *tamaš-ēn* ~ *tama-y-ēn*; *wāḏ-ēn*, *wēv-ēn*.

Emphatic and anaphoric forms are marked by *yik-* (*ik-*) preceding the demonstrative forms, ‘the very, that which’, etc. Thus, the remote or neutral demonstrative combines as follows: singular, direct case, m. *yik-u*, f. *yik-ā*; oblique, m. *yik-wi*, f. *yik-wam*; and so forth: direct case, *xu vō yik-u bōzī* ‘then, again (there was) the very same entertainment’; oblique, *māš-ām [ik-wi kōr] čūj* ‘we (*-ām*) would do (all) this (very) work’; [*yik-di kōr*] *pōc lūv-ēn* ‘this (very) activity they call *pōc*’. Note the adverb of manner *dis* ‘thus, in such a way’; emphatic *yik-dis* ‘(exactly) this way’.

These forms are prominently found with restrictive relative clauses (see 5.3).

3.1.3.3 Reflexive and possessive pronouns

Reflexive: *xuba0* ‘oneself’, oblique *xu*, note *yu xu-rd-a0-i lūd* ‘he said to (*-rd*) himself’; possessive: *xu(d)* ‘(one’s) own’.

3.1.3.4 Interrogatives and indefinites and related adverbials

Interrogative-relative: human, *čāy*, oblique *či* ‘who’; animal, non-human, *čīz* ‘what (thing)’; *ca* ‘what, which’; *ca* + N: *ca-waxt* ‘what time, when’; *ca-rāng* ‘how’; *čīz-ardlrad* ‘what for, why?’ *čīdīm* ‘which one?’; *cūnd* ‘how much’, *dūnd* ‘so much’.

Generic: *ar* ‘each, every (one)’; *ar-čāy* ‘every, each one’; *fuk(-a0)* ‘all’; *var0* ‘both’.

Indefinite: (*y*)*i-čāy* ‘anybody; somebody’; exclusive: (*y*)*ič(-a0)* — (*na-*) ‘(not) any’: *atā yi-čāy as wēv-andi na ris-t* ‘and not anyone from among them remains’.

Interrogative adverbials follow the triple deictic system, with post- and prepositions: *k-u*, *k-id*, *k-am*; *kā-* ‘where when’: *kād-and* ‘where’.

3.1.4 Adpositions

Numerous syntactic particles and words are commonly used in Shughni as in any other oral language. The adpositions selected here are based on forms that occur frequently and in multiple contexts. They reflect a highly developed system of spatial orientation and movement characteristic for the linguistic area.

3.1.4.1 Prepositions

The main prepositions are:

<i>tar</i>	horizontal movement, 'towards, to';
<i>as, az</i>	source, 'from', 'about (a topic)'; also specific direct object;
<i>pī</i>	upwards, 'on, up(to)';
<i>ar</i>	downwards, 'down, in(to)';
<i>pīs</i>	final and spatial orientation and time, 'after, following';
<i>čī</i>	position, 'on, at, to'; intention;
<i>mī</i>	position of something, 'at, in';
<i>par</i>	means of movement, 'by';
<i>tō</i>	limit in time, space, 'till';
<i>ba</i>	instrumental; manner of action.

3.1.4.2 Postpositions

The main postpositions are:

<i>-(a)nd</i>	constant availability, presence; possession; (< * <i>ana-</i> , contamination with * <i>antar-</i> , cf. Av. <i>ana</i> , Yazgh. <i>na</i>);
<i>-ja</i>	location, possession;
<i>-ti</i>	location on some surface, 'at, on, upwards, for';
<i>-avēn</i>	'for (the sake of)';
<i>-aǰ</i>	direction;
<i>-andīr, -and(i)</i>	definite, exact direction or location (often inside), location both 'in' and 'from' (inside); definite time (< * <i>antar-</i>);
<i>-(a)rd, -ra(d)</i>	indefinite location and time; direction, both 'towards' and 'from'; addressee of the speech, action, aim;
<i>-ēc</i>	limit, period; instrument;
<i>-qati</i>	sociative-instrument.

Postpositions require the oblique case of the pronouns.

Pre- and postpositions may combine, such as *az, as* 'from' with *-and* and directional *-ard*: *az wi xēz-and* 'from his house'; *as yiǰīd-ard* 'out from the sheds'.

3.1.4.3 Case functions

As indicated, three of the postpositions also function as markers of main syntactic cases:

- (1) directional *-ard* for dative and experiencer or beneficiary: *mu-rd dāk* 'give me', *māš-ard lūv* 'tell us';
- (2) definite locational-directional *-(a)nd(i)* as ablative: *wēv-andi* 'from them';

- (3) *-and* for possessive: *yu mu-nd* ‘this is mine’. Note that some are bi-directional, such as *-andi* ‘from’ in the preceding example vs. *-and* ‘in’ in *qišlōq-and* ‘in a village’.

3.1.5 Adverbial demonstratives and adverbs

Similar to the demonstrative and interrogative pronouns, the pronominal adverbs are used concretely with pre- and postpositions, such as the directional suffix *-ard* in *yam-ard* ‘there’ (far), *yēd-ard* ‘there (to)’, *yūd-ard* ‘here (to)’, or the locational suffix *-and* in *kād-and* ‘where’, *yam-and* ‘there’, *yūd-and* ‘here’; *tar kā* ‘to(wards) where’.

3.1.6 Numerals

3.1.6.1 Cardinal numbers

The cardinal numbers up to ten are indigenous: *(y)īw* ‘one’, *diyūn* ‘two’, *aray* ‘three’, *cavōr* ‘four’, *pīnj* ‘five’, *xōj* ‘six’, *wūvd* ‘seven’, *wāxt* ‘eight’, *nōw* ‘nine’, *δīs* ‘ten’. The teens follow an additive pattern, ‘ten’ + N: *δīs-at yīw* ‘ten-and one’, *δīs-at diyūn* ‘ten-and two’.

Higher numbers are loaned from the Tajik. However, the older generation still preserves the indigenous decimal system and compound numbers: The three upper teens are counted by subtraction, e.g.: *yīw kam du δīs* ‘one less of two ten’ = ‘nineteen’. The higher units are multiples of ten: *du δīs* ‘two ten’ = Taj. *bīst* ‘twenty’; *cavōr δīs* ‘four ten’ = Taj. *čil* ‘forty’; *δīs-δīs(-ak)* ‘ten ten’ = Taj. *sad* ‘one hundred’; *δīs-δīs-δīs-ak* = *hazor* ‘one-thousand’.

3.1.6.2 Ordinal numbers

Ordinal numbers are formed by adding the suffix *-(y)um*: *yīw-um*, ‘first’, *aray-um* ‘the third’; but the Tajik ordinal numerals are widely used.

3.1.6.3 Distributives

Distributive numerals are formed by adding the stressed suffix *-(y)j̄*: *yīw-ī̄* (*yīw-ī̄*) ‘one by one’.

3.2 Verb morphology

The verb system distinguishes tense, mood, person, number, transitivity, and voice. A small set of intransitive verbs have gender distinction in past tenses.

3.2.1 Stem formation

3.2.1.1 Present, past, and perfect stems

In terms of morphophonemic alternations, there are regular and irregular verbs. Because of such alternation, a total of five (present, past, perfect, past perfect and infinitive) stems can be distinguished. In addition, the form of the 3s present may be distinct from the other persons, and the mentioned set of past intransitives has different masculine, feminine and plural stems.

- (1) The present stems reflect most of the productive Old Iranian present stem classes.
- (2) Past stems are formed by adding *-t*, *-d*, reflecting the Old Iranian perfect participle in **-ta*. Regular verbs simply add *-t*, *-d* to the present stem: pres *fām-*, past *fām-t-* ‘to understand, know’; irregular verbs show considerable stem alternation: pres. *xāw-*, past *xīw-d-* ‘to eat’.
- (3) Perfect stems are formed by the suffix *č* or *ǰ* (reflecting **-ka-*): regular pres. *palōys-*, past *palōys-t*, perf. *palōys-č-* ‘to work’; irregular pres. *sāw-*, past m. *sūt-t*, f. *sa-t*, perf. m. *su-δ-ǰ*, f. *si-c*, pl. *sa-δ-ǰ* ‘to become’.
- (4) Past perfect stems are derived from the perfect stem by suffix *-at* (Shd. *-it*; originally a past of ‘to be’): *su-δ-ǰ-at* ‘had become’.

3.2.1.2 Causative stems

Causative stems, intersecting with transitivity, partially reflect the Old Iranian patterns, notably in the pattern non-causative *CaC-* vs. causative *Cč/čC-*, mostly with further stem alternation. Newer causative stems are derived by *-ēn*, which is infixed before root labial, *-ē(m)b*; also *-ūn*, *-ūn*:

	Present/Past ‘to fly up’	Present/Past ‘to stick to’	Present/Past ‘to be fried’	Present/Past ‘to crumble’
Intr.	<i>rivāz-irivuxšt</i>	<i>piδafs-lpiδūw-d</i>	<i>sitafs-lsitūw-d</i>	<i>raz-lrišt</i>
Caus.	<i>rivēz-irivēz-d</i>	<i>piδēmb-lpiδēm-t</i>	<i>sitēb-lsitēb-t</i>	<i>raz-ēn-lraz-ēn-t</i>

Note denominative: *tult* ‘rag’ > *tult-ūn-t-ōw* ‘to drag’.

3.2.1.3 Compound verbs

There are numerous denominal compound verbs, formed by a nominal with a common verb serving as auxiliary verb: *vār-δēd-ōw* ‘to be able’, *kōr čīd-ōw* ‘to do work, work’, onomatopoetic *taq-taq* ‘rat-tat’ > *taq-t-ōw* ‘to knock’. These include calques from Tajik: *naqli čīd-ōw* ‘to make a narration, to tell’, Taj. *naql kard-an*.

3.2.2 Nominal forms

The following description is confined to the most frequently found nominal forms.

3.2.2.1 Agent noun

Agent noun, present stem + *-īǰ*: *lūw-īǰ* ‘speaker’.

3.2.2.2 Infinitive

The infinitive corresponds to the past stem, usually with *i*-umlaut (reflecting **-ti*): past *tūyd-*, inf. *tīd* ‘to go’; extended form past stem + *-ōw*: *čēvd-ōw*, beside *čēvd* ‘to sleep’.

3.2.2.3 Perfect participle

The perfect participle consists of the perfect stem + *-ak*: intransitive *nīvǰ-ak* ‘weeping’, transitive *nivīšt-ak* ‘written’.

3.2.2.4 Adjectival participle

The adjectival participle consists of the perfect stem + *-in*: *zinōdǎj-in* ‘washed’; gender is distinguished where retained in the verb: m. *tūyǎj-in*, f. *tīc-in* ‘somebody who has gone away’.

3.2.2.5 Future participle

The future participle consists of the infinitive + *-mēǎj* and connotes intention: *tūd-mēǎj* ‘preparing to go’.

3.2.3 Person marking and ‘to be’

The type of conjugation of the present-future tense differs significantly from that of the past tenses. Thus, in the present-future the three persons in singular and plural are the inherited personal endings added to the present stem. In the three past tenses, however, the three persons in singular and plural are indicated by personal enclitic pronouns (resulting from the merger with the earlier copula). Unlike the personal endings, these markers are as a rule attached to the first constituent part of the clause.

The distinction between copula and present and past person marking is leveled except in the 2s and 3s. The forms of the verb *vi-*, *vud-* ‘to be’ replace the copula in all other tenses and contextual uses.

TABLE 14b.4: COPULA AND PERSONAL ENDINGS

	1s	2s	3s	1p	2p	3p
Copula	<i>-um</i>	<i>-at</i>	<i>-∅</i>	<i>-ām</i>	<i>-ēt</i>	<i>-ēn</i>
Present	<i>-um</i>	<i>-i</i>	<i>-t/-d</i>	<i>-ām</i>	<i>-ēt</i>	<i>-ēn</i>
Past	<i>-um</i>	<i>-at</i> imp. <i>-∅</i>	intr. <i>-∅</i> , tr. <i>-(i)</i>	<i>-ām</i>	<i>-ēt</i> imp. <i>-ēt</i>	<i>-ēn</i>

The existential verb is 3s *yast*, neg. *nist*, which also serves to express possession (see 4.5.1 below).

3.2.4 Negation

The general negative marker is *na*, which precedes the verb, including preverbs: *yatt-ōv-um na-vār-ōd* ‘I could not come’; *nist* ‘is not’. The subjunctive and prohibitive marker is *mā*:

mā-ya(d) ‘don’t come’;

parwōs Sindēv bōyad yu mā-vi-d
last year Sindev must he not to be
‘last year he must not have been in Sindev’.

3.2.5 System of tenses

The basic inflectional system of tenses is four-fold: present-future, past, perfect, past perfect:

Present-Future	Past
Perfect	Past Perfect

The imperative is distinguished by 2s -Ø; the present subjunctive by the use of the negative marker *mā-*: *bōyad yu mā-vi-d* 'he must not be'.

The following exemplifies typical intransitive and transitive conjugations.

TABLE 14b.5: INTRANSITIVE AND TRANSITIVE STEMS AND IMPERATIVE

Intransitive verbs								
Present	Past		Perfect			Infin.	Imper.	
	m.	f./pl.	m.	f.	pl.	•		
<i>vi-</i>	<i>vud-</i>	<i>vad-</i>	<i>vidǰ-</i>	<i>vic-</i>	<i>vaδǰ-</i>	<i>vid</i>	<i>vi</i>	'to be'
<i>sāw-</i>	<i>sut-</i>	<i>sat-</i>	<i>sudǰ-</i>	<i>sic-</i>	<i>saδǰ-</i>	<i>sit</i>	<i>sāw, sa</i>	'to become'
<i>ti-</i>	<i>tūyd-</i>	<i>tōyd-</i>	<i>tūyǰ-</i>	<i>tīc-</i>	<i>tōyǰ-</i>	<i>tīd</i>	<i>tī, tē</i>	'to go'
Transitive verbs								
Present	Past		Perfect			Infin.	Imper.	
<i>kin-</i>	<i>čūd-</i>		<i>čūǰǰ-</i>			<i>čīd</i>	<i>ki(n)</i>	'to do'
<i>palōys-</i>	<i>palōyst-</i>		<i>palōysč-</i>			<i>palōyst</i>	<i>palōys</i>	'to work'
<i>āžēr-</i>	<i>āžért-</i>		<i>āžérč-</i>			<i>āžért</i>	<i>āžér</i>	'to moisten'

3.2.6 Transitivity and voice

3.2.6.1 Transitivity

The differential marking of the 3s is a major indicator of transitive verbs as opposed to intransitives. In turn, the distinction of transitivity is marked in intransitive verbs where in a small, but significant set gender and number are marked by means of ablaut. Both features reflect an earlier ergative system (see section 7 Dialectology).

3.2.6.2 Passive

Passive is expressed by the perfect participle of transitive verbs usually marked by the suffix *-ak*, followed by the auxiliary verb *sitt-ow* 'to become': *yā xat nivīšč-ak sat* 'that (f.) letter (*xat*) written was (f.)'.

4 SYNTAX AND USE OF FORMS

4.1 Coordination

Some of the coordinating conjunctions are enclitic. Both noun phrase and clause clitics are *-at* 'and', *yō* 'or'. Paired conjunctions include *ām-ām* 'both . . . and', *yō . . . yō* 'either . . . or'.

TABLE 14b.6: INTRANSITIVE CONJUGATION

'to be'							
	Present	Past, m.	f./pl.	Perf. m.	f.	Past Perf. m.	f.
1s	<i>vi-y-um</i>	<i>vud-um</i>	<i>vad-um</i>	<i>vuđj-um</i>	<i>vic-um</i>	<i>vuđj-at-um</i>	<i>vic-at-um</i>
2s	<i>vi-y-i</i>	<i>vud-at</i>	<i>vad-at</i>	<i>vuđj-at</i>	<i>vic-at</i>	<i>vuđj-at-at</i>	<i>vic-at-at</i>
3s	<i>vi-d</i>	<i>vud</i>	<i>vad</i>	<i>vuđj</i>	<i>vic</i>	<i>vuđj-at</i>	<i>vic-at</i>
1p	<i>vi-y-ām</i>	<i>vad-ām</i>	<i>vad-ām</i>	<i>vađj-ām</i>	<i>vic-ām</i>	<i>vađj-at-ām</i>	<i>vic-at-ām</i>
2p	<i>vi-y-ēt</i>	<i>vad-ēt</i>	<i>vad-ēt</i>	<i>vađj-ēt</i>	<i>vic-ēt</i>	<i>vađj-at-ēt</i>	<i>vic-at-ēt</i>
3p	<i>vi-y-ēn</i>	<i>vad-ēn</i>	<i>vad-ēn</i>	<i>vađj-ēn</i>	<i>vic-ēn</i>	<i>vađj-at-ēn</i>	<i>vic-at-ēn</i>
'to become'							
	Present	Past, m.	f./pl.	Perf. m.	f.	Past Perf. m.	f.
1s	<i>sāw-um</i>	<i>sut-um</i>	<i>sat-um</i>	<i>suđj-um</i>	<i>sic-um</i>	<i>suđj-at-um</i>	<i>sic-at-um</i>
2s	<i>sāw-iyi</i>	<i>sut-at</i>	<i>sat-at</i>	<i>suđj-at</i>	<i>sic-at</i>	<i>suđj-at-at</i>	<i>sic-at-at</i>
3s	<i>sū-d</i>	<i>sut</i>	<i>sat</i>	<i>suđj</i>	<i>sic</i>	<i>suđj-at</i>	<i>sic-at</i>
1p	<i>sāw-ām</i>	<i>sat-ām</i>	<i>sat-ām</i>	<i>sađj-ām</i>	<i>sic-ām</i>	<i>sađj-at-ām</i>	<i>sic-at-ām</i>
2p	<i>sāw-ēt</i>	<i>sat-ēt</i>	<i>sat-ēt</i>	<i>sađj-ēt</i>	<i>sic-ēt</i>	<i>sađj-at-ēt</i>	<i>sic-at-ēt</i>
3p	<i>sāw-ēn</i>	<i>sat-ēn</i>	<i>sat-ēn</i>	<i>sađj-ēn</i>	<i>sic-ēn</i>	<i>sađj-at-ēn</i>	<i>sic-at-ēn</i>
'to go'							
	Present	Past, m.	f./pl.	Perf. m.	f.	Past Perf. m.	f.
1s	<i>tī-y-um</i>	<i>tūyd-um</i>	<i>tōyd-um</i>	<i>tūyj-um</i>	<i>tīc-um</i>	<i>tūyj-at-um</i>	<i>tīc-at-um</i>
2s	<i>tī-y-i</i>	<i>tūyd-at</i>	<i>tōyd-at</i>	<i>tūyj-at</i>	<i>tīc-at</i>	<i>tūyj-at-at</i>	<i>tīc-at-at</i>
3s	<i>tīz-d</i>	<i>tūyd</i>	<i>tōyd</i>	<i>tūyj</i>	<i>tīc</i>	<i>tūyj-at</i>	<i>tīc-at</i>
1p	<i>tī-y-ām</i>	<i>tōyd-ām</i>	<i>tōyd-ām</i>	<i>tōyj-ām</i>	<i>tīc-ām</i>	<i>tōyj-at-ām</i>	<i>tīc-at-ām</i>
2p	<i>tī-y-ēt</i>	<i>tōyd-ēt</i>	<i>tōyd-ēt</i>	<i>tōyj-ēt</i>	<i>tīc-ēt</i>	<i>tōyj-at-ēt</i>	<i>tīc-at-ēt</i>
3p	<i>tī-y-ēn</i>	<i>tōyd-ēn</i>	<i>tōyd-ēn</i>	<i>tōyj-ēn</i>	<i>tīc-ēn</i>	<i>tōyj-at-ēn</i>	<i>tīc-at-ēn</i>

Clause clitics include: *-atā* (*-ata*, *-at*, *-tā*, *-ta*, *-ā*, *-a*) 'and, but'; *-xu* 'and, then'; *-mis* 'also'; *-vō* 'again'; others include: interrogative *-ō*; referential *-ik*; and intensifying *-aθ*. Particles include: vocative *a/ē*, *ō*; interjections: *ā*, *aγ*, *(w)uč*, *učča*.

4.2 Noun phrase structure

Shughni is a head-final language. The unmarked sequence is:

Determiner – Possessive – Numeral – Adjective – Head Noun, and
Dependent Noun Phrase – Head Noun Phrase.

The attribute may be represented by pronoun, noun, numeral, adjective, participle, or an infinitive and as a rule precedes the head noun.

4.2.1 Attributive adjectives

In general, adjectives precede the head noun by parataxis, and agree in gender with the head where gender is retained, thus ADJ (gender) – NOUN:

xuŷmūn ziv 'Shughni language'; with gender agreement: *rōšt mūn* 'red-(f.) apple',
tīc-in γāc 'walking (f.) girl'; *tēr cimūd* 'black basket', pl. *tēr cimūd-ēn*.

TABLE 14b.7: TRANSITIVE CONJUGATION

'to do'				
	Present	Past	Perfect	Past Perfect
1s	<i>kin-um</i>	<i>čūd-um</i>	<i>čūṣj-um</i>	<i>čūṣj-at-um</i>
2s	<i>kin-i</i>	<i>čūd-at</i>	<i>čūṣj-at</i>	<i>čūṣj-at-at</i>
3s	<i>kin</i>	<i>čūd-(i)</i>	<i>čūṣj(-i)</i>	<i>čūṣj-at(-i)</i>
1p	<i>kin-ām</i>	<i>čūd-ām</i>	<i>čūṣj-ām</i>	<i>čūṣj-at-ām</i>
2p	<i>kin-ēt</i>	<i>čūd-ēt</i>	<i>čūṣj-ēt</i>	<i>čūṣj-at-ēt</i>
3p	<i>kin-ēn</i>	<i>čūd-ēn</i>	<i>čūṣj-ēn</i>	<i>čūṣj-at-ēn</i>
'to work'				
	Present	Past	Perfect	Past Perfect
1s	<i>palōys-um</i>	<i>palōyst-um</i>	<i>palōysč-um</i>	<i>palōysč-at-um</i>
2s	<i>palōys-i</i>	<i>palōyst-at</i>	<i>palōysč-at</i>	<i>palōysč-at-at</i>
3s	<i>palōys-t</i>	<i>palōyst</i>	<i>palōysč</i>	<i>palōysč-at</i>
1p	<i>palōys-ām</i>	<i>palōyst-ām</i>	<i>palōysč-ām</i>	<i>palōysč-at-ām</i>
2p	<i>palōys-ēt</i>	<i>palōyst-ēt</i>	<i>palōysč-ēt</i>	<i>palōysč-at-ēt</i>
3p	<i>palōys-ēn</i>	<i>palōyst-ēn</i>	<i>palōysč-ēn</i>	<i>palōysč-at-ēn</i>
'to moisten'				
	Present	Past	Perfect	Past Perfect
1s	<i>āžêr-um</i>	<i>āžêrt-um</i>	<i>āžêrc-um</i>	<i>āžêrc-at-um</i>
2s	<i>āžêr-i</i>	<i>āžêrt-at</i>	<i>āžêrc-at</i>	<i>āžêrc-at-at</i>
3s	<i>āžêr-t</i>	<i>āžêrt-(i)</i>	<i>āžêrc-(i)</i>	<i>āžêrc-at-(i)</i>
1p	<i>āžêr-ām</i>	<i>āžêrt-ām</i>	<i>āžêrc-ām</i>	<i>āžêrc-at-ām</i>
2p	<i>āžêr-ēt</i>	<i>āžêrt-ēt</i>	<i>āžêrc-ēt</i>	<i>āžêrc-at-ēt</i>
3p	<i>āžêr-ēn</i>	<i>āžêrt-ēn</i>	<i>āžêrc-ēn</i>	<i>āžêrc-at-ēn</i>

4.2.2 Pronominal adjectives

Pronominal adjectives precede the head noun with gender and number agreement: PROADJ (gender/number) – NOUN:

yu rūz 'that (m.) day'; *wam xāb* 'that (f.) night'; *dāδ ṣinik-ēn* 'these women'.

Note that pronouns are also marked for case, here direct *yu* and oblique *wam*, and thus not only indicate the gender of the noun, but also the case of the noun phrase.

4.2.3 Dependent nouns and noun phrases

Dependent nouns and noun phrases paratactically precede the head noun or postposition, and are in the oblique case, which is evident with pronouns: NP2 (Oblique) – NP1:

[*mu bāx*] *mirōs* 'my share of the inheritance', [*wi sūrdōr*] *nūm* 'the name of that wedding-holder'.

Postpositional phrases: [*wi*] *garginūxā* 'around him'; [*wi*] *pāli-ndi* 'on his side, beside him'.

4.2.4 Izāfa

The Tajik-type *izāfa* construction NP1-*i* NP2 is found frequently:

Shughni, *sōl-ō-y-i sōl-ēc* ‘for many years’ (with Shughni temporal *-ēc*), cf. Tajik, *sol-ho-y-i sol*; *kām-i kām* ‘little of little, at least’. Even in this loan construction, the oblique case may be marked: *as nūm-i [wi]* ‘in his name, on his behalf’.

4.3 Clause structure and word order

The standard sentence structure is subject-object-verb, SOV.

The syntactic function of a noun phrase is specified by its place in a sentence, pre- and postpositions, and other markers.

The predicate is usually located at the end of the sentence. However, depending on intonation, reflecting the sense and other circumstances of the utterance, the predicate can be transferred to the beginning of the sentence, thus placing it before the subject:

sāt-ām māš tar sūr
‘went-1p we to the wedding’.

In turn, location and particularly direction which semantically depends on the verb rection are often placed after the verb:

yu naštūyd tar vaš
‘he went out’.

Otherwise, adverbial phrases locating or qualifying the sentence as a whole are often in initial position.

In emotional speech, however, the word order can be disregarded. In oral Shughni speech, intonation is highly important as a means to convey the syntactical connection of words in a phrase.

4.4 Semantics and use of nominal forms

4.4.1 Person, animacy, and alienability

The category of person vs. non-person is conveyed by lexical means, and in reduced form by the interrogative-indefinite pronouns, such as personal *čāy* ‘who’ vs. *čīz* for animal, non-human, and *ca* for inanimate.

The category of *alienable* vs. *inalienable* possession is denoted by lexical means as well, and is overt with inalienable possession, including parts of the body, kinship terms, and others, by preposed personal pronouns: *[mu] yōδ* ‘my memory’, *[mu] δūst* ‘my hand’. This distinction becomes more evident in the presence of a preposition. In that case, the possessive pronoun is focused and precedes the preposition, whereas it precedes the noun directly if alienable:

mu [pi] tanā ‘on my body’; *mu [tar] δūst* ‘in my hand’; *mu [ba] yōδ* ‘in my memory’; as opposed to the unmarked order; *[tar] mu čīd* ‘in my house’.

Note also *tu-(a)t [ca] nūm* ‘what’s your name?’ and the split of *(y)ik-wi* in the prepositional phrase: *pirō yik[az]wi* ‘before (this that)’ (Persian *piš az ān ke*).

4.4.2 Syntactic cases

4.4.2.1 Subject marking

Subjects are unmarked and in the direct case.

4.4.2.2 Direct object marking

Direct objects are in the oblique case, where distinct, and may be preceded by the preposition *as*:

mu nān [wēv] wīnt ‘my mother saw [them]’;
nēδ-ēn [wi]. . . ‘they place [him] on. . .’;
tu [as wēv] na wīnt ‘you have not seen [them]?’.

4.4.2.3 Indirect object marking

Indirect objects are generally marked by postpositions, but may be unmarked with the verb *δēd-ōw* ‘to give’:

qamōč [mu-rd] vār ‘bring [me] some bread’;
šinōwari [wi] δ-ēn ‘they give [him] a washing’.

4.4.2.4 Adverbial phrases

Adverbial phrases may stand as the plain oblique case where distinct:

[wi] sar pirō rūz ‘on [that-m] very first day’;
[wam] šāb ‘(during) [that-f] night’;

or may otherwise be unmarked:

šumnē sūr-ēt ‘tomorrow you (will arrange) a wedding’;
tarū ya ‘come here’.

More often they are more specified by pre- and postposition:

[pi] mazōr sāv-um ‘I am going [to] the shrine’;
[tar] mu xēz ya ‘to my place come!’;
čīd xēz[-and] ‘[at] a house’s location, place’.

4.5 Semantics and use of verb forms

4.5.1 Possession

Possession and ownership is expressed by a locative construction, where the possessor is marked by the postposition *-and*, and the verb is the copula, the existential verb *yast*, or other forms of ‘to be’, i.e. ‘to X is/was’:

mu[-nd] ič-aθ pūl nīst
 ‘I have (at all) no money’, lit. ‘to me’;
tu[-nd] cūnd puc
 ‘how many sons do you have?’, lit. ‘to me’;

wēv[-and] virōdyast
‘they have a brother’, lit. ‘to them’;

mu dōd[-and] wi dūnd xizmatgōr
‘my father has so many servants’ (3s copula regularly zero), lit. ‘to my father’s’;

4.5.2 Person, number, and agreement

4.5.2.1 Person and number

The singular refers to a single object, concept, or objects in their entirety as a class; hence numerals larger than one generally require the singular. The plural expressedly refers to plurality. It is morphologically marked by agglutinative suffixes, together with ablaut in a small subset of nouns, as discussed above.

Indefinite agents are as a rule expressed by the 3p.

4.5.2.2 Subject-verb agreement

The verb is in agreement with the subject in person, number, and gender:

[wuz]-ta tiy[-um] ‘I am going’;

[māš] mūm-ēn tiy[-ēn] ‘our grandmothers are going’.

However, if the subject contains components like *-gal(l)ā* or *-xēl*, which connote collectivity, the predicate may be in the singular:

civinc [-xēl]-ta rivōz[-d] ‘these wasps are flying’.

Similarly, collective nouns like *mardum* ‘people’ and *ōdam* ‘people, humans’ usually show the 3s (for the particle *-ta* see 4.5.3.2)

4.5.2.3 Predicates

- (1) Predicative noun phrases, even when referring to plural subjects, may be singular or plural (*-ēn*):

māš-ām tama ziryōt(-ēn)
‘we are (*-ām*) your child/children (*-ēn*)’.

- (2) The predicate may be expressed by finite verb forms, nouns, or other parts of speech with separable predicative copulas:

māš tama na-wzūn[-ām] *kasal[-um] vud* *dāδ[-ēn] γal nist*
‘we don’t know you’; ‘I was sick’; ‘they have not (come) yet’.

- (3) Nominal sentences in the 3s have zero marking:

tu-t ca nūm ‘what (is) thy name?’, lit. ‘thou-thy what name’.

4.5.3 Use of “tenses”

Given the fact that there are only four morphologically distinct tenses, each form functions in a variety of semantic-cognitive contexts.

The basic functions are temporal. Thus, the present-future refers to ongoing, habitual, or future actions. The past refers to past actions. The present perfect expresses present relevance: *sic* ‘she has gone, left’. The past perfect refers to a past or remote result or state: *šinik-ēn ya0č-at-ēn* ‘the women had come’.

Contextually all four also have modal connotations (see 5.2 Subordinate clauses).

The following notes are confined to the more salient points, here focusing on the present-future.

4.5.3.1 Present-future, modal function

tar kāl-tēxēj sut, xu kāl [tēx-t]
to barber he went, that his head he shaves
‘he went to the barber to shave his head’.

4.5.3.2 Factual enclitic *ta*

The enclitic particle *ta* disambiguates the various functions of present-future, and emphasizes the reality or fact of an action. It is usually attached to the first constituent part of the clause, as in the first sentence in Sample Text 1 below:

ar sōl [ta] buōr-jēv bāt ayūm īd [anj-ēn]
each year *ta* springtime Bat Ayom fest they celebrate
‘every year they celebrate the Bat Ayom festival during springtime’.

The text describes the fact that the ancient customs of the New Year festival continue to be performed to this day. In fact, with its first occurrence *ta* predicates that text as a whole.

Similarly, the factual function of *ta* together with the present-future may contextually connote certainty in the completion of an action:

naw [ta] yu [yōd-d], xu māš ta tar maktab [sāw-ām]
new *ta* he comes, then we *ta* to school we go
‘as soon as he comes, then we to school will go’;

yu [ta] ca waxt [yōd-d], māš [ta] awqōt [xār-ām]
he *ta* what time comes, we *ta* dinner we eat
‘when he comes, we (shall) have dinner.’

The contrast between the presence and absence of *ta* is shown in the following examples with present-future forms, where the second sentence implies that the coming may be an accomplished fact:

mu gumūn yu šič kōr [kič-t]
me assumption (is) he now work does
‘I think he may be working now (*šič*)’;

4.6.2 Inception

The inception of an action may be expressed by replacing ‘to be’ with the action verbs *sāw-*, *sut-* ‘to become’, *kin-*, *čūd-* ‘to do’, or *δāδ-*, *δōd-* ‘to give’. Here *dar*, *daraw*, *dawum* are often used instead of *pay* and *či*:

Present: [Subject] [X *pay* Infinitive] *sāw*-Personal Ending.
 Past: [Subject]-Person [X *pay* Infinitive] *sut*, or
 [Subject] [X *pay* Infinitive]-Person *sut*

(1) Present:

γac-ēn [dar/daraw/pay nīwd] *s-ēn [daraw dēwōl čīdōw] s-ēn*
 ‘girls begin to cry’; ‘(they) begin to build the walls’;
[zamīn dar čērt] kin-ām *dī ųin [pāy vid]*
 ‘(we) begin to plough the land’; ‘his wife is waiting (for a child)’.

(2) Past:

γac-ēn [dar/daraw/pay nīwd] sat *[wi pay-ām daraw mīzd] sat*
 ‘girls began to cry’; ‘that sour milk (we) began to churn’;
yā talaw [pay tīd] sat
 ‘that barrel began to leak’;
tama γal dakō xār-ēt-at, *wuz [sōz dar/daraw lūwd] sāw-um*
 ‘you-2p are still eating, and I begin to sing songs’.

4.6.3 Habitual action

Habitual action is expressed by preposing the adverbial *dōyim(-ik)* ‘continuing’ in the present and past:

atā [dōyim-ik]-ēn bēx-di yō bāt ayum [lūd], yō xidīr ayum
 but [continuing]-AG3p more either Bat Ayom [called], or Xidir Ayom
 ‘but they used to call (*lūd*) it Bat Ayom, or Xidir Ayom’.

4.7 Modal constructions

A broad range of modal meanings can be expressed by constructions with modal words and auxiliaries.

4.7.1 Ability

Ability is expressed by the verb *vār-đi-*, past *vār-đōd-*, perf. *vār-đōđj-* ‘can, to be able’. The form of the dependent verb is either (1) the present-future; (2) the full infinitive form or (3) the infinitive in *-ōw*. The infinitives may precede or follow the auxiliary, e.g.:

- (1) Present-future: *wuz vār-δīm [sāw-um]* ‘I can go’.
 (2) Infinitive: *yu [xīnōwari] na vār-δ-ēd* ‘he can not swim’.

- (3) Infinitives in *-ōw*:
wuz [nivišt-ōw] vār-δ-iy-um ‘I can write’;
wuz-um vār-δōδj [di kōr čīd-ōw] ‘I could do this work’.

biyōr mu mūm [siḥd-ōw] na- vār-δōd
 ‘yesterday my grandmother was unable to get up’.

4.7.2 Wishes

In wishes, the subject-experiencer is marked by the dative *-ard*, with the transitive auxiliary *fōrt*, and the extended infinitive if the action refers to the subject:

- Pres.: [N-*ard*] Infinitive-*ōw fōrt*: *mu-rd īd-ōw fōrt* ‘I want to go (home)’;
 Past: [N-*ard*]-*i* Infinitive-*ōw fōrt*: *mu-rd-i īd-ōw fōrt* ‘I wanted to go (home)’.

biyōr mu-rd fōrt, idi tu-t tarūd yaθč
 yesterday me-for wish, *idi* you-you-2s here have-come-PF
 ‘yesterday I wished you would have come’.

Alternatively the explicit term *xōyiḥ kin-lčūd* ‘make wish, want’ is used (see also Wish clauses):

rēdž-in ta wuz xōyiḥ čūd-um, wūn wēδ-um
 ‘on the remainder (of it) *ta* I want to place wool myself’.

4.7.3 Intention

Intention may be expressed by the participle *-mēj*. This construction is confined to the verbs ‘to go’ and ‘to stay’: *yu īd-mēj* ‘he is going to go’ (zero copula); *yā tīd-mēj na-vad* (fem.) ‘she had no intention to go’.

4.7.4 Necessity and obligation

- (1) Necessity as well as obligation is expressed by *bōyad* ‘must, to have to’, followed by the present-future for present action, and by the perfect for past action; thus:

Present: [Subject] *bōyad* [Present-Personal Ending]
 Past: [Subject]-Person *bōyad* [Perfect]

- (a) Present:

wuz bōyad [nur memūnī sāw-ām]
 ‘I must [go to visit my friends today (*nur*)]’.

- (b) Perfect:

wuz-um bōyad [parāxīb memūnī suδj]
 ‘I had to go to visit my friends the day before yesterday’.

- (2) Suggested obligation may be expressed by *bōyad* followed by the imperative; thus: *bōyad* [X Imperative]:

tu bōyad [namīnā sa]
‘you should [be(come) an example (for others)]’;

tu bōyad [qīw] wi
‘you should [call him]’.

4.7.5 Supposition

Supposition may be expressed by *bōyad* followed by the present-future; negated action requires prohibitive *mā*, rather than *na*:

parwōs Sindēv bōyad [yu mā-vi-d]
last year Sindev must [he not to be]
‘last year he must not have been in Sindev’.

4.7.6 Assumption

Assumption may be expressed by *gumūn* ‘assumption, thought’, extended *mu gumūn* ‘my assumption (is)’, or *wuz gumūn kin-um* ‘I make the assumption’. The dependent verb phrase is in the present-future, with additional *ta* when referring to a possible completion of the action:

ongoing: *gumūn* [X Present-Personal Ending]
completed: *gumūn* [X *ta* Present-Personal Ending]

mu gumūn [yu šič kōr kīx-t]
‘I think [he may be working]’;

wuz gumūn kin-um gumūn [yā ta yōd-d]
‘I make the assumption [she may have come]’.

5 COMPLEX SENTENCES

As Shughni is an oral language, the structure of the clause is usually simple. However, simple clauses may be combined into complex sentences in two ways, by coordination and by subordination, which may be conveyed by suprasegmental means of intonation and by conjunctions.

5.1 Coordinate clauses

The structure of the coordinating sentence may be asyndetic, i.e. with the omission of conjunctions in sentence constructions in which they would usually be used. This is prominently the case in case of simultaneous, sequential, or contrasting events:

yu xēyd, yā xōvd
‘he read, she slept’.

More frequently, clauses are connected by *clitic* coordinating conjunctions conveying sequential or consecutive actions, or actions resulting from another, *atā*, *at* ‘and’ and *xu* ‘then, and so’:

šič qarōr niθ-atā, *wuz luv-um*
‘you now keep quiet, and I (will) speak’;

wuz-um rinūxt-at, *yu-i yōdōwarī na-čūd*
‘I forgot, and he (-i AG) did not remind (me)’;

yu as xu jōy-ti andīyd-at, *māš-ām sat tar wi xēz*
‘he stood up from his place, and we (-ām) went to his house’;

yu as xu jōy andūyd-xu, *māš-ām sat tar wi xēz*
‘he stood up from his place, then we went to his house’;

tar yi-či-aθ-i na čūxt-xu, *ricūst*
‘he (-i AG) did not look at anybody, then he ran away’ (i.e. ‘without looking’).

5.2 Subordinate clauses

Most of the subordinate clauses in a complex sentence perform a function similar to the syntactical functions of the parts of a simple sentence. Since subordinating conjunctions have a broad range of use and functions, they occur in different types of subordinate clauses. The relative position of subordinate clauses and the placement of conjunctions and particles vary. The most frequent conjunctions and their main uses, together with the unmarked sequence of the dependent clauses, are:

Dependent clauses preceding the main clause

<i>ca</i>	preverbal: relative, concessive, temporal, and conditional clauses;
<i>aga</i>	clause initial: conditional clauses, often with <i>ca</i> ;
<i>idi(idē)</i> , <i>di</i> , <i>didi</i>	clitic: causal clauses.

Dependent clauses following the main clause

<i>idi</i> , <i>di</i>	clause initial: object, restrictive relative, consecutive, and final clauses;
<i>xu</i>	clitic to main clause: final clauses.

Conjunctive phrases

These are mostly calqued on Persian, and include:

<i>dūnd-ēc-i idi to</i>	‘until’ (‘so much-TIME that’);
<i>kād-and ca</i>	‘while, if’ (‘where-in that’);
<i>ba jō-i di</i>	‘instead of’ (Persian <i>be jā-ye in ke</i>);
<i>pirō yik az wi di</i>	‘before (this that)’ (Persian <i>piš az ān ke</i>);
<i>dūnd-avēn idi</i>	‘for (the sake of)’ (Persian <i>ba-rā-ye ān ke</i>).

5.3 Relative clauses

5.3.1 Basic structure

Relative clauses are usually marked by preverbal *ca*, which is, however, not obligatory. The pronominal referent in the relative clause takes the appropriate case marker, but is usually absent:

yid tu pūxōk, narm-at safēd ca, dūnd-ga xizmat tu-rd kiḫ-t, yō nay?
 this your dress, soft-and white-ca, such service you-to make, or not
 ‘this dress of yours-2s, which is soft and white, will it serve you this long as well?’.

5.3.2 Generalized relative clauses, ar- ‘each, all’

ar čīz wi-nd ca vud, fuk-a0-i binēst
 every-thing him-POSS what was, all he (-i AG) lost
 ‘he lost everything that he had’.

5.3.3 Head noun incorporation

ar čī-nd nawna vē-d, zinōdž-in ta pinīz-d
 to every whom new not is, washed ta he-puts on
 ‘everyone who has no new clothes puts on clean ones’.

5.3.4 Focused restrictive clauses

In focused restrictive clauses the antecedent is marked by pronominal forms with the particle (y)ik-, followed by the conjunction *idi*, *idē*, with or without *ca*:

ōdam yast, idē ayūm mē0 sad tarmurx yēs-t
 person exists, that fest day 100 egg takes away
 ‘there are people who can win up to one hundred eggs on the day of the fest’;

yid ik-u čōrik idi vēgā-y-um di ar bōzōr ca wīnt
 this-very-he man that yesterday-I him at bazaar ca saw
 ‘this is a man whom I saw at the bazaar yesterday’;

id ik-u čōrik idi dam mūn az wi xēz-and-um zōḫt
 this-very-he man that this-OBLf apple from his-OBLm place-LOC-I took
 ‘this is a man from whom I brought the apple’;

yid ik-ā ḡinik idi wuz ik-wām rizīn-um čī-vīd
 this-very-she woman that I very-her-OBLf daughter-I am going-to marry
 ‘this is the woman whose daughter I am going to marry’.

5.4 Object clauses

Object clauses directly follow the main clause; here with interrogative pronoun:

wuz na fām-um, yu tar kā ravūn
 I not know, he to where goes
 ‘I don’t know where he goes’.

5.5 Adverbial clauses

5.5.1 Temporal clauses

Temporal clauses may precede or follow the main clause, depending on the sequence of action, with or without preverbal *ca*, or with conjunctive phrases, and may be followed by the clitic *xu* ‘then, so’:

yu ta ca waxt yōd-d, māš ta awqōt xār-ām
 he *ta ca* what time comes-PR, we *ta* dinner eat-PR
 ‘when he comes, we (would) have a dinner’;

naw ta yu yōd-d-xu māš ta tar maktab sāv-ām
 newly *ta* he comes-PR-then we *ta* to school we-go-PR
 ‘as soon as he comes, then we will go to school’;

ptrō yik az tar čīd dēd-ōw, ōdam bōyad xu būt wēd-d
before this from to house to enter-INF, person must own shoe removes-PR
 ‘before going inside the house, people should put off their shoes’.

5.5.2 Conditional clauses

Conditional clauses are marked by the preverbal conjunction *ca*, and are often introduced by the conjunction *aga* ‘if’. These clauses are of three main types:

5.5.2.1 Real or likely conditions

Real or likely conditions are expressed by the present-future, followed by the imperative and present-future, or perfect; or by the perfect in both clauses:

5.5.2.1a Present + imperative:

aga yu ca yōd-d, lū wi-rd (idi wuz ta pi wi ni0-um)
if he *ca* comes-PR, say-IMP him-for (that I *ta* for him sit-I-PR)
 ‘if he comes, tell him that I will wait for him’.

5.5.2.1b Present + present:

sūr purrā ca naǰjimb-ēn, kām-i kām array mē0 darkōr
 wedding full *ca* they-perform-PR, little-IZ little three day necessary-PR
 ‘if one wants to, to hold the wedding ceremony in full, at least three days is necessary’.

5.5.2.2 Possible conditions

Possible conditions are expressed by the present-future or perfect in their subjunctive function:

5.5.2.2a Present + perfect:

tu wi ca fām-i, tu-rd yu xuš yaθč
 you-2s him *ca* you-understand-PR, you-2s-for he well has-come-PF
 ‘if you get to know him, you would like him’.

5.5.2.2b *Perfect + perfect:*

aga yu ca yaθč, *māš-ām ik-wi* *kōr čūjǰ*
 if he *ca* has-come-PF, we-we the-same work have-done-PF
 ‘if he would come, we would do this work’.

Here the perfect appears to connote possible or assumed future fact, approximately ‘once he has arrived/will have come, we will certainly have completed’, or ‘we will complete all this work’.

5.5.2.3 Irreal conditions

Irreal or desirable conditions are expressed by the past perfect, followed by the past, or more generally by the past perfect:

5.5.2.3a *Past perfect + past:*

aga yu ca yaθč-at, *wi fāmt* (*idi yi čōrik mūjǰ-at*)
 if he *ca* had-come-PP, he understood-PT (that one man has-died-PP)
 ‘if he would have come, he would have seen that a man had been dying for him’.

5.5.2.3b *Past perfect + past perfect:*

tu-t *wi ca fāmč-at,* *tu-rd yu xuš na yaθč-at*
 you-you-2s him *ca* had-understood-PP, you-for he well not had-come-PP
 ‘if you knew him, you would dislike him’;

aga yu šič-ard ca yaθč-at, *māš-ām ik-wi* *kōr čūjǰ-at*
 if he now-for *ca* had-come-PP, we-AG this-very-OBLm work had-done-PP
 ‘should he have come by now/already, we would do this work’;

aga yu ca yaθč-at, *fuk sōz sudǰ-at*
 if he *ca* had-come-PP, all fitting had-become-PP
 ‘if he had come, everything would be alright’.

5.5.3 *Concessive clauses*

Concessive clauses precede the main clause, and are marked by the enclitic conjunction *ca*:

kōr-i *mis ca čūd, pūl-ēn* *wi-rd na-δōd*
 work-AG3s though *ca* did, money-they him-to not-gave
 ‘although he did his work, he was not given money’;

pi wi ca qīw-um, yu na-yat
 to him *ca* call-I, he not came
 ‘though I called him, he did not come’;

ar cūnd-i *yu qīwd, wuz-um na-yat*
 every how-much-AG3s he called, I-I not came
 ‘however often he called me, I did not come’ (Persian *har čand*).

5.5.4 Consecutive clauses

Consecutive clauses have a cataphoric antecedent and are introduced by *idi*:

wi-nd dis lap kōr vud, idi yi čīz-a0-i uspēt na-čūd
 he-for so much work was, so that one thing-3SAG in time not did.
 ‘he had so much work, that he could not do one thing in time’;

yu dūnd-ēc-i kōr čūd, idi tō ušmand na-suđj
 he so much-TIME-AG3s work did, so that till skilled not has become-PF
 ‘he worked hard, until he finally became skilled enough’.

Note the negation in *tō na suđj*, lit. ‘till not he has become’.

5.5.5 Causal clauses

Causal clauses either precede the main clause, marked by the conjunction *di*, or follow the main clause introduced by an explanatory phrase, such as *di jāt* ‘(for) this reason’ (Pers. *az in jāhat*):

parwōs di kasāl-um vud, yat-ōw-um na-vār-đōd
 last year di ill-I was, to come-INF-I not was able
 ‘because I was ill last year, I could not come’;

šitō sut, dijāt mōl ar yi jīd
 cold became, this reason animals in sheep-cote
 ‘it got cold; this is why the cattle (are) in a sheep-cote’.

5.5.6 Final clauses

Final clauses follow the main clause and may be introduced by the clitic *xu* ‘then, so (that)’. The verb is usually in the present-future in its subjunctive function:

tar kāl-tēxēj sut-xu kāl tēx-t
 to barber he-went, so that head he-shaves-PR
 ‘he went to the barber to shave his head’.

5.5.7 Wish clauses

Wish clauses are introduced by *idi* and are followed by the present-future or the present perfect in their subjunctive function:

mu-rd xōyiš, idi tu ya(d)
 me to wish, that you come
 ‘I wish that you (would) come-PR;

parwōs-um xōyiš vud, idi tu-t tarū yaθč
 Last year-I wish was, that you-you here have come
 ‘last year I wished that you would (have) come here-PP’.

5.5.8 Request clauses

Clauses implying request are usually expressed as direct speech which is introduced by the particles *nā* (*la*), the conjunction (*d*)*i-di*, with or without the verb *lūvd-ōw* ‘to speak’:

yā tar mu gāxt dīdi ti tar čīd
she to me turned that you to house!
‘she turned to me: “go to the house” ’;

yu mu-rd lūd idi tu niθ
he me-to said that you sit!
‘he told me: “sit down/take a seat” ’.

5.5.9 Narration and dependent speech

Constructions of indirect narration are practically not used. The usual way of conveying direct speech are the verbs *lūvd-ōw* ‘to speak’, *qīvd-ōw* ‘to call’, *pēxst-ōw* ‘to ask’: *lūd: vār dēv* ‘(he) said: “bring them”, followed by the conjunction (*d*)*idi*, such as *yu lūd idi* ‘he said that’:

yu yat, (lūd) idi ti-y-ām
‘he came (said) that: “let us go” ’;

aga yu ca yō δ-d, lū wi-rd, idi niθ-um pi wi
‘if he comes, tell him that I wait for him’.

Alternatively, the particles *nā*, *nāla*, with the meaning ‘says that’:

yā nāla, mu-nd mu kōr lap
she says mine my work much
‘she (says): “I have a lot of work” ’;

yu nāla, pi yēl garm sut
‘he (says): “in the mountain pasture it got hot” ’.

6 LEXIS

6.1 Word formation

6.1.1 Derivation

The derivation by affixes is mostly by suffixation, some of which are closely related to Tajik equivalents. Prefixation as a model of formation is rarely used, and mostly loaned from Tajik.

6.1.1.1 Suffixation

- ī abstract nouns: *rūšt-ī* ‘redness’;
- gar agent nouns: *vidōj-gar* ‘irrigation-worker’;
- ēj (a) the geographic origin of a person; *xūf-ēj* ‘from Khuf’, *xaray-ēj* ‘from Khorog’;

- (b) general relationship of things: *vidīrm-ēj* ‘bush used for making brooms’;
 (c) step-kinship for masculine gender: *virōd-ēj* ‘stepbrother’;

-*ēj* step-kinship for feminine gender: *nān-ēj* ‘stepmother’;
 -*ik*, -*ak* diminutives: *julik-ik* ‘very small’, *čib-ak* ‘small spoon’;

- in* (a) adjectives from nouns: *šar0k-in-ak* ‘earthenware toy’ (*šar0k* ‘clay’);
 (b) participles: *tīc-in* ‘having come (f.)’.

6.1.1.2 Prefixation

Adjectival *ba-* ‘with’, *bē nō* ‘without’, form adjectives: *ba-mazā* ‘delicious, taste-ful’, *bē-tamīz* ‘shame-less’.

6.1.2 Compounding

6.1.2.1 Basic types

Juxtaposition, simple reduplicated: *xêx-tabōr* ‘relatives’ (Persian *xiš-tabūr*); *pêx-palak* ‘shoes, lit. (high) boot-foot-cloth’; *dil-dil čīdōw* ‘to rock to sleep (child)’.

Alliteration, with substitution of the initial consonant mainly with *m-* or *p-*: *čōy-pōy* ‘tea’; *bač-kač* ‘children; family’; *cīlīy-pīlīy* ‘noise’; *pūl-mūl* ‘money’; *xōriji-mōriji* ‘foreigner’.

Connective formants are: (a) *-at* ‘and’, e.g. *šin-at čōr* ‘wife and husband, couple’; (b) others include *-o*, *-tar*, *-a*, *-či*, *-ba*, e.g. *γēv-tar-γēv* ‘abuse’, *šab-o-šab* ‘till dawn’.

6.1.2.2 Determinative compounds

Noun-noun: *būy-wōx* ‘aromatic herb’ (*būy* ‘aroma’ + *wōx* ‘herb’), *can-kamūnak* ‘sling’ (*can* ‘gun’ + *kamūnak* ‘bow’), *xidōrj-žīr* ‘millstone’, *gul-yunčā* ‘bud’ (‘flower bud’).

Adjective + noun: *safēd-gāl* ‘an old woman’ (*safēd* ‘white’ + *gāl* ‘head’, fem. form).

Numeral + noun: *du-zivak* (*du* ‘two’ + *zivak* ‘tongue’), *var0-tanā* ‘both persons’.

Noun + verb stem or agent noun: *žīz-vōr* ‘fuel gatherer’, *žīndam-xōr* ‘wheat eater’, *šac-vārīj* ‘water carrier’.

Noun + lexical gender marker: m. *tēr-gīl*, f. *tēr-gāl* ‘with black hair’, m. *žīz-vōr*, f. *žīz-vēr* ‘a person who brought firewood’.

6.2 Sociolinguistic aspects

6.2.1 Naming

Shughn(an)i anthroponyms merged with Iranian Muslim anthroponymics and are partly canonized, but have retained some specific features. Thus the tradition of preserving the name of a deceased ancestor is carried out by a ritual of transferring his name to a descendant. Ritually preserved names consist of two or three components: *Mast-āli*, lit. ‘intoxicated (by love for the shiite champion) Ali’, *Yusuf-āli-šō* < *Yusuf-Ali-Shāh*.

Nicknames are a unique source of language materials, for their anthroponymy is based on appellatives that are rarely used or not used at all in the modern language, e.g. the proper names f. *Kanginā* ‘a type of vessel’, m. *Yūr-ak* ‘(bear) cub’.

6.2.2 *Kinship terms*

The Shughni kinship terminology is mostly indigenous (for plural forms see Section 3.1.1.2 Number).

TABLE 14b.8: KINSHIP TERMS

<i>pid</i>	'father'	<i>nān</i>	'mother'
<i>bōb</i>	'grandfather'	<i>mūm</i>	'grandmother'
<i>puc</i>	'son'	<i>rizīn</i>	'daughter'
<i>nibōs</i>	'grandson'	<i>nibēs</i>	'grandson, -daughter'
<i>virōd</i>	'brother'	<i>yax</i>	'sister'
<i>xisūr</i>	'father-in-law'	<i>xīx</i>	'mother-in-law'
<i>amak</i>	'uncle (paternal)'	<i>ammā</i>	'aunt (paternal)'
<i>xōlak</i>	'uncle (maternal)'	<i>xōlā</i>	'aunt (maternal)'
<i>virōd-ēj</i>	'step-brother'	<i>xēr</i>	'nephew, niece'
<i>nān-ēj</i>	'step-mother'	<i>abīn-ēn</i>	'co-wives'

Note, *pid* also 'father-in-law'; *dōd*, *tāt* 'father' (nursery term).

Members of the family commonly address each other using these terms. To indicate an extended family or a kin, the term *awlōd* is used: *Nazaršō awlōd*. The term for blood relatives is *xēx-tabōr*; others: *xēl*, *zōt*, *qawmiyōt* (*qawm* 'relative'); *tōyfā* 'kinship; ethnic group'.

6.2.3 *Familiar address and phraseology*

Polite phraseology is represented by some words of address, which are used in Shughn(an)i for socialization and friendly small talk:

a-ḡa 'guys' < *a-yiḡā*; *a-rō* 'brothers' < *a-virōd*; *a-čēn* 'children' < *a-bačēn*;
a-lō 'girl', address among girls and women; *a-lā* 'you-all', address to everybody.

When an elder family member addresses a young person (son or daughter, boy or girl), he or she uses the term for his or her own kinship relation:

a-tāt 'dad', father to youngster; *a-nān* 'mom', mother to youngster;
a-mūm 'grandmother', elder female to young; *a-bōb* 'grandfather', elder male to young.

When a son or a daughter is born, the person who brings the news is called *xuš-pōy-i qadam* 'who has easy-going step, brings happiness'. This person greets the happy father with the ritual phrase *Mubōrak-i tirandōz* 'congratulations by a shooter' in case of a son, and *Mubōrak-i (h)alvō-pāz* 'congratulations by a person preparing sweets' if a daughter was born.

Frequent are also such verbal formulae as *salōmōlēk!* 'How do you do; welcome', cf. less official *salūm* greetings, 'hi'. A polite answer for greeting is *wālēk bar salūm*, accompanied by a *das(t)-būsi* 'a gesture of ritual kissing each other's hand'.

7 DIALECTOLOGY AND CONTACTS

7.1 Internal dialectology of Shughni

The Shughn(an)i linguemes that are located in neighboring valleys have rather clear geographical boundaries, but remain mutually intelligible. In their contact zones transitional dialects develop, and their level of linguistic affinity varies.

In vocabulary, there are found series of differentiations, particularly in the most active stratum of cultural lexemes. In the Shahdara subdialect the speech of residents of the lower villages, i.e. those from Tavdem to Khorog, closely resembles that of Shughni, while higher up the Shahdara river, from Tavdem to Jawshanghoz, are found some subdialectal distinctions.

The subdialect of Barwoz, which is located in highest part of the valley, is quite exclusive and can clearly be distinguished. It has developed some specific phonetic, grammatical, and lexical characteristics, which still need to be investigated in detail. This dialect is now limited to use within a few families.

The distinctive criteria applied for dialects and subdialects are usually those of a phonetic and/or grammatical nature. Regarding the consonant system, one can find several distinctions in special features of several phonemes, together with particular lexemes: in the Shughn(an)i proper and Shahdara subdialects the consonants *k g* are postdorsal in all positions. In the Bajuwi and Barwoz subdialects these consonants differ from Shughn(an)i, and by their articulation approaches that of Rushani and Khufi palatalized *k g*. The fricative *ɣ* in the Shughn(an)i proper and Shahdara subdialects also has permanent postdorsal articulation. In the Bajuwi and Barwoz subdialects, fricative *ɣ* in some positions approaches that of Rushani and Khufi palatalized *ɣ*. Thus, for example, in the Bajuwi subdialect the articulation of the fricative *ɣ* in position before frontal *i*, *i* substantially moves forward, and develops into a sonant palatal variant (Karamshoev 1963: 69). Particularly noteworthy is the correspondence between Bajuwi *w* – Shughni *ɣ̣*, which reflects deeper divergent historical developments.

Distinguishing features in grammar include the prevalence of different grammatical markers, such as the plural marker Shughn(an)i *-ēn* vs. Shahdara *-(j)ēv*, mentioned above.

Of major importance for the larger Shughn(an)i-Rushani sub-group is the distinction between transitive and intransitive inflection in past tenses. In particular, the ergative construction is retained in Shughn(an)i only minimally by the marker *-i* for the 3s agent, whereas the ergative construction is relatively better preserved in Rushani and Bartangi.

7.2 Adjacent languages and dialects

The relatively large geographic area where Shughn(an)i is spoken is adjacent to a large variety of linguistic neighbors. In the North, the neighboring languages are Rushani and Khufi, which are closely related to Shughni. In the south, the prevalent languages are Tajik, and then Ishkashmi and Wakhi. In the east and north-east are the Turkic Kyrghyz dialects.

The Shahdara and Ghund valleys were the destination for numerous immigrants from various regions of the West Pamir, including Rushan, Bartang, Roshorv, and even more distant Sarez and Murghab. In the upper reaches of the Shahdara valley, where the almost inaccessible mountain pass to Wakhan is located, the local people trace their grandparents to Wakhan. To this day their relationship has not been interrupted

completely, and close marriages and social contacts with the Wakhi-speaking community are still common there.

The closest contacts and connections with the Tajik language are those with the Badakhshan Tajik subdialects in the south-west of the region, Ghoron, Ishkashim, and with some other Tajik dialects, such as Munji in Shahdara. In the west there are close contacts with the Dari dialects of Afghanistan.

7.3 Loan component

Indigenous forms are preserved in verbs, pronouns, nouns (cultural terms, body parts, crafts and occupations, agricultural, ritual terms). However, numerous terms related to traditional culture, such as the names of the seasons in the folk calendar, house utensils, types of buildings, ritual dishes, clothes, etc. are disappearing, and the older terms tend to be preserved only in bound contexts such as collocations and phraseology.

Most numerous in the contemporary language are loans from Tajik. These include terms relating to modern technology as well as traditional Islam: *šarāt* 'shariat' (the Islamic law), and *šaytūn-arōba* 'bicycle', adopted in the 1920s. Some loans include terms, which have become obsolete in the Tajik itself, and are only preserved in the Pamir languages, sometimes with shift of meanings, such as *jamāt* 'family' (beginning of the twentieth century) in the term *jamōat-xinā* used by the Ismaili community for a place for prayer.

There are some loanwords from Arabic that entered via classical Farsi or Tajik, but have become archaisms in contemporary Tajik. Such Shughni terms include *šitō* 'cold' < A. *šitā* 'winter', *saxrā* 'stubborn fellow, country man', Taj. *saxra* 'rocks; country man (or person who lives in high mountains)', Ar. *saxrā* 'rock; stubborn'; *yalīz* 'crude; inadmissible', *yalīz gāp* 'crude word(s)', Ar. *yalīz* 'coarse; large, fat'.

In turn, numerous Shughni words or their phonetic variants are used in Tajik dialects, such as *yūrš* 'dung', Taj. dial. *yūruš*, f. Taj. *yušā* id.; *yūr-ak* 'unripe fruit', Taj. Badakhshan *yurk*, *yuruk* id., Northern Taj. dial. *yurak* 'fruiting, ovary', *yur* 'nutshell'.

7.4 Pamir languages and Indo-Aryan

There are some loan words from other Pamir languages, including the following loaned via Wakhi: "mountain stream, flow; tributary", Barwoz micro-dialect *zirāv* < Wakhi *īerav* (vs. Shughni proper *šarvidōi* and similar in other varieties); "wolf", Sarikoli *xīθπ* < Wakhi *šapt*, an old taboo term < Dardic Khowar *šapir* < **āpita-*, OInd. *āapta-* 'cursed' (vs. lexemes derived from **wārka-* in most Shughni varieties, e.g. Shughni proper *wūri*). In general, terms borrowed from Indo-Aryan relate to material culture, such as *birūi* 'birch bark'; *čipōs rūgan* 'cotton oil' < MInd. **kappāsa-* (as in Pali), OInd. *karpāsa-* 'cotton'; *kappur* 'vessel made of pumpkin' < MInd. **kappara-*, OInd. *karpāra-* 'bowl'.

7.5 Turkic Kyrghyz

A number of loanwords are from Turkic, especially Kyrghiz, most of which relate to cattle breeding, dairy farming, housekeeping equipment (especially associated with summer pasturing).

7.6 Substrate features

A certain lexical stratum is of substrate origin. There is also a group of “areal words” which are used in different languages of the region (Iranian, Indo-Aryan, Dardic, and non-Indo-European), but whose origins are unknown.

The substratum influence was significant, and not only resulted in supplying new vocabulary items, but also caused a shift in the semantics of words that already existed in Shughni. For example, *dawūm čīd-ōw* ‘to begin’, loaned from Arabic *davām* via Tajik *davōm* ‘continuation’, could change its meaning from ‘continuation’ to ‘beginning’ not spontaneously but only under the influence of a substratum form like Burushaski *dohon* ‘beginning, the beginning’ (with substitution *h* for *w*). Substrate influence is also reflected in the innovative derivation of *t-ama*, the pronoun of the 2p, from the singular pronoun *t-u*.

8 SAMPLE TEXTS

8.1 *Bāt Ayūm* ‘The Spring Equinox Festival’

ar sōl ta buōr-jēv bāt ayūm īd anj-ēn.
each year ta spring-TIME Bat Ayom fest they-set about
‘In the spring of every year they celebrate the Bat Ayom fest.’

yid as fuk-a0 yulla-di ayūm māš-and,
this from all biggest fest us-POSS
‘This is our greatest holiday.’

šič ta bēx mardum dam ayūm nawrūz luv-ēn
now ta more people this-OBLf fest Nawruz they-say
‘Nowadays most people call it Nawruz.’

atā dōyim-ik-ēn bēx-di yō bāt ayūm lūd yō xidīr ayūm.
and continuous-they more this-OBLm either Bat Ayom call-PT or elder Ayom
‘but they used to call it either Bat Ayom or Xidir Ayom.’

ažda-um-i märt yō bīst-u-aft-um-i amal
18-th-iz March or 27-th-iz Amal
‘On March 18, or Amal 27’

atā mardum ta xu kōčōr fuk-a0 as čīd-and ziwēd-ēn-xu
and people ta own thing all from house-LOC they-bring out, then
‘people take all their belongings outside, then’

xu čad-ēn tōzā kin-ēn, sipē-gil yō gāč wēv δ-ēn,
own houses fresh they-make, white-clay or lime those-OBLp they-hit
‘clean their houses, whitening them with white clay or lime’

dōyim-ēn yīnik-ēn xu čīd būrj-ēn gul-buri mis čūd yōyǰ qati.
continuous-AG3p women own house columns decoration also made-PT flour with
‘women usually decorated the walls of their houses with flower ornaments made with flour.’

wi ayūm mē0 ta mardum fuk-a0 naw pūxōk pinīz-d,
that-om fest day ta people all new dress he-puts on
'On the day of the holiday people put on new clothes.'

ar čī-nd naw na vē-d, zinōđj-in ta pinīz-d-at
each who-to new not is, washed ta put on, and
'those who have no new clothes put on clean clothes, then'

bād naḫti-y-ēn tar māraka-yēn,
after they-leave to social gatherings
'afterwards go to social gatherings.'

ḡinik-ēn dišad-ēn ti dāf di-y-ēn,
women roofs on drum they-hit
'women play local drums on the roofs of their houses.'

čōrik-ēn guj tāž-ēn-at qastīn anj-ēn.
men kid they-drag-and wrestling they-take up
'while men compete in goat-dragging and wrestling (and similar sporting events).'

atā as fuk-a0 bēx-di ta tarmurx δēd wēđ-ēn,
and from all more ta egg fighting they-put
'And the best of all is the egg cracking competition.'

ōdam yast idē ayūm mē0 sad tarmurx yēs-t.
person exists that fest day 100 egg takes away
'there are people who can win up to one hundred eggs on the day of the fest.'

8.2 The Prodigal Son, Luke 15, 11–32

atā yēsō dawūm-i čūd: yi ōdam-and-ēn du puc vad.
and Jesus begin-AG3s made: one man-to-3p two son were
'And Jesus began to speak: One man had two sons.'

fištīr puc xu dōd-ard-i lūd: mu bāx mīrōs mu-rd dā-k.
younger son own father-for-AG3s said: my portion inheritance me-for give
'The younger son told his father: Give me my portion of the inheritance.'

atā yu čōrik xu mōl-at ōl-i xu pac-ēn-ard bāx čūd
and that man own property-and wealth-AG3s own sons-for division made
'And that man divided his (own) livelihood between his sons.'

bād as δu-aray mē0 yu fištīr puc fuk-a0-i xu mōl
after from two-three day that younger son all-AG3s own property
jām čūd,
collection made

'After two three day the younger son gathered all his livelihood'

xu pūl-qati tar ḍaraki jōy tūyd. tar-am-i yik-dis ayš čūd
 own money-with to far-off place went there-AG3s such-way enjoyment made
 ‘And went with his money to a far country. There he so enjoyed himself’

idē ar čīz wi-nd ca vud, fuk-a0-i binēst.
 that every thing his-POSS ca was-m all-AG3s lost
 ‘that he lost everything that he had.’

atā yik-di waxt-and yik-wi mumlakat-and mōjĵ sut-at
 and very-this-OBLm time-LOC very-that-OBL-1m country-LOC famine it-became-and
 ‘And just at that time in that country famine arose, and’

yu sōf bē-navō sut.
 he plain poor he-became
 ‘he became completely poor.’

bād as maĵbūri sut tar yi bōy xēz kōr tilāpt,
 after from necessity he-went to one rich-man house work requested
 ‘After that he became forced to ask for work at one rich citizen’s farm,’

atā yu bōy wi-y-i tar xu qišlōq xūg-būni čīd-ōw bōxt.
 and that rich him-AG3s to own village swine-herding making sent
 ‘And that citizen sent him to his village to herd the swine.’

yu amēxā, yik-dis maĵjūnĵ vud idē xūg-ēn xurōk-ā čī-xīd sut,
 he always such-way hungry he-was that swine fodder to-eating he-became
 ‘He was always so hungry that he began to eat the fodder of the swines,’

dūnd-avēn idē wi-rd-ēn yi čīz-a0 na-ḍōd.
 such-sake that him-for-AG3s one thing not gave
 ‘because they gave nothing to him.’

dāḍ wi wuḥ tar wi yat-xu yu xu-rd-a0-i lūd:
 then his conscience to him came-then he self-to-AG3s said
 ‘Then this awareness came to him, and he said to himself:’

mu dōd-and wi dūnd xizmat-gōr
 my father-to that-many servant
 ‘My father has so many servants’

atā yi-čāy as wēv-andi maĵjūnĵ na-ris-t-at,
 and one from them-among hungry not remains-and
 ‘and none of them stays hungry, and’

wēv xurōk bēx mis ḍēd.
 their food more also remained
 ‘there are always leftovers of their food’.

BIBLIOGRAPHY AND REFERENCES

- Bakhtibekov, T. (1979) *Grammatikai zaboni Shughnoni* (A grammar of the Shughni Language), Dushanbe: Donish.
- Dodykhudoev, R. Kh. (1962) *Materialy po istoricheskoi fonetike shugnanskogo iazyka* (Materials on Shughni historical phonetics), Dushanbe: Tadzhiiskii Gos. Univ. im. V.I. Lenina.
- Dodykhudoev, R. H. and Dodykhudoeva, L. R. (2001) (tr., ed.) *Luqo Injil* (The Gospel of Luke), Moskva: Institut tarjuma Kitobi Muqaddas.
- Dodykhudoeva, L. R. (1988) *Shugnanskii glagol v istoricheskoi osveshchenii* (The Shughni verb in historical perspective), Dushanbe: Donish.
- (1997) ‘Sotsio-lingvisticheskaia situatsiia v Badakhshane’ (The sociolinguistic situation in Badakshan) in *Malye iazyki Evrazii – sociolingvisticheskii aspekt*, Moskva: Filologicheskii fakul’tet MGU im. M.V.Lomonosova, pp. 78–89.
- (2003). ‘Shugnanskii iazyk’, in *Iazyki Rossiiskoi Federatsii i sosednikh gosudarstv. Entsiklopedia*, III, Moskva: Nauka.
- Edel’man, D. I. (1980) ‘History of the consonant system of the North-Pamir languages’, *IJL* 22: 287–310.
- (1986; 1990; 2009) *Sravnitel’naia grammatika vostochnoiranskikh iazykov*. I. *Fonologiya*. II. *Morfologiya*, *Elementi sintaksisa*, III. *Leksika* (Comparative grammar of the East-Iranian languages, I. Phonology; II. Morphology. Elements of syntax, III. Lexis), Moskva: Nauka, Vost lit-ra RAN.
- (1987) ‘Shugnano-rushanskaia iazykovaia gruppa’ in *Osnovy* IV, pp. 236–347
- Edel’man, D. I. and Iusufbekov, Sh. P. (2000) ‘Shugnanskii iazyk’ (The Shughni language), in *Iazyki mira* III, pp. 225–42.
- Karamshoev, D. (1963) *Badzhuvskii dialekt shugnanskogo iazyka* (The Bajuw dialect of the Shughni language). *Trudy*, vol. 16, Dushanbe: Akademiia nauk Tadzhiiskoi SSR. Institut iazyka i literatury im. Rudaki.
- (1988; 1991; 1999) *Shugnansko-russkii slovar* (Shughni-Russian dictionary) 3 vols, Moskva: Nauka.
- Lentz, W. (1933) *Pamir-Dialekte*. I. *Materialen zur Kenntnis der Schughni Gruppe*, Göttingen: Vandenhoeck & Ruprecht.
- Morgenstierne, G. (1928) ‘Notes on Shughni’, *NTS* 1: 32–84
- (1938) *Iranian Pamir Languages* = Morgenstierne, *IIFL* II.
- (1974) *Etymological Vocabulary of the Shughni Group*, Wiesbaden: Reichert.
- Nawata, T. (1979) *Shughni* (Asian and African Grammatical Manual, No. 175), Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University.
- Shaw, B. (1876; 1877) ‘On the ‘Galchah Languages (Wakhi and Sarikoli)’; ‘On the Shighni (Ghalchah) Dialect’, *JASB* 45: 139–278; 46: 97–126.
- Sokolova, V. S. (1953) *Ocherki po fonetike iranskikh iazykov*, II (Phonetics of the Iranian languages, II), Moskva: Akademiia Nauk SSSR.
- (1966) ‘Shugnano-rushanskaia gruppa iazykov’ (Shughni-Rushani language group, in: *Iazyki narodov SSSR*, I, Moskva: Nauka, pp. 362–397.
- (1967) *Geneticheskie otnosheniia iazguliamskogo iazyka i shugnanskoi-iazikovoi gruppy* (Genetic relations between Yazghulami and Shughni language group), Leningrad: Nauka.
- (1973) *Geneticheskie otnosheniia mundzhanskogo iazyka i shugnano-iazguliamskoi yazikovoi gruppy* (Genetic relations between the Munji language and Shughni-Yazghulami language group), Leningrad: Nauka.
- Zalemann (Salemman), K. G. 1895. ‘Shugnanskii slovar’ D. L. Ivanova’ (The Shughnani vocabulary of D. L. Ivanov), in *Vostochnye zametki Sankt-Peterburgskogo universiteta* 1895, pp. 269–320.
- Zarubin, I. I. (1960) *Shugnanskii teksty i slovar’* (Shughni texts and dictionary), Moskva, Leningrad: Akademiia Nauk SSSR.

CHAPTER FIFTEEN

WAKHI

Elena Bashir

1 INTRODUCTION

1.1 Overview

Wakhi (Tajik Persian *wakhī*, self-designation *ǰik*; Khowar *xikwar*, Burushaski *guīski*, Shina *guītsaá*) is an Eastern Iranian Pamir language. In Tajikistan and Afghanistan, Wakhi villages are situated on both banks of the Panj River (= upper Amu Darya, Oxus) below the confluence of the Wakhan River. On the Tajikistan side, Wakhi villages extend from Namadgut to Ratm, interspersed with a few Tajiki settlements. On the Afghanistan side, settlements stretch from P̄it̄ur, near Ishkashem, to Sarhad on the upper reaches of the Wakhan River. In Pakistan, the main settlements are in Gojal, including part of the upper Hunza valley and the Shimshal and Chapursan valleys; in Ishkoman; and the upper Yarkhun valley in Chitral. In China, two separate settlement clusters are in Dafdar (Sarikol) and the towns of Kilian and Sanja (Pishan) (Kreutzmann 1996: 137).

Wakhi is spoken (1990–91 data) by a total of around 37,570 people, of whom 12,500 live in Tajikistan, 9500 in Afghanistan, 11,770 in Pakistan, and 3800 in China (Kreutzmann 1996: 137). In the following, Wakhi dialects will be discussed with reference to national boundaries because of the dearth of comparable data from Tajikistan, Afghanistan, China, and Pakistan. Moreover, the national boundaries established during the nineteenth and twentieth centuries have divided and continue to further divide a formerly more cohesive and homogeneous language area (see also 7 Dialectology). The main sources for Wakhan Wakhi (WW) are Pakhalina (1975) [hereafter P] and Griunberg and Steblin-Kamensky (1976) [hereafter GSK]; for Hunza Wakhi (HW), Lorimer (1958), Buddruss (1986), Reinhold (1992a, b, c), Mock (1998), and the present author's field notes from Hunza. All examples the source of which is not otherwise indicated are from Bashir's field notes.

As far as is known, the original home of the Wakhi speakers is the area along the Wakhan River in present-day Tajikistan and Afghanistan. There is no evidence in oral tradition or onomastics that would point to immigration from any other place (Morgentierne 1938: 435). From Wakhan proper, Wakhi speakers have migrated during historical times to Sarikol (China) and several areas in the Northern Areas and Chitral regions of Pakistan (see Kreutzmann 1996: 139 for a schematic of the migration patterns). Mock (1998: 17, 22) reports that the earliest permanent Wakhi settlements in Shimshal and Avgardh were probably established from 250 to 400 years ago, and that Chapursan was settled during the nineteenth century. Wakhi speakers migrated to Ishkoman around 1880 (Lorimer 1958: 8). Apparently there were Wakhi settlements in Chitral along the upper reaches of the Yarkhun River in the early twentieth century (Stein [1921] quoted in Payne 1989: 420).

The position of Wakhi in Tajikistan, Afghanistan, and Pakistan presents an interesting sociolinguistic laboratory. In Tajikistan, the language of communication, writing, and education is Tajik Persian. Wakhi oral tradition is also bilingual in Wakhi and Tajiki and many Wakhis also speak Shughni, the second lingua franca of the Pamir region. Schooling is obligatory for everyone in Tajikistan and the medium of education is Tajiki, leaving only old women and young children monolingual. This diminishes the role of Wakhi in public contexts, though it is still preferred in domestic settings. Until very recently, Wakhi speakers in Tajikistan felt themselves a marginalized group with negative attitudes toward their language; however the poems published in Reinhold (1992a) indicate that there is an incipient awakening of interest in local languages.

In Afghanistan Wakhis also speak Dari Persian, or Pashto. However, literacy is not widespread and the admixture of other elements into Wakhi is less than in Tajikistan.

The situation of Wakhi in Gojal (Pakistan) is unique. Almost 100 percent of Wakhi-speaking children in Hunza now attend school, which exerts greater influence on the language than any single factor in the past. This has resulted in an enormous literacy gap between younger and older generations. All schoolgoing children and many young women up to the age of 20 know Urdu, whereas hardly any of the older generation do (Reinhold 1992b). In addition, Gojal Wakhi is being heavily influenced by both Urdu and English (Reinhold 1992b). Nevertheless, it is an increasingly vital language, and speakers have a very positive attitude toward their language, evidenced by the Wakhi Tajik Cultural Association (Mock 1998: xvi, 38).

1.2 Writing Systems

Wakhi is now being written in both Tajikistan and Gojal. The poems appearing in Reinhold (1992a) were written in modified IPA-based script inspired by the work of Griunberg and Steblin-Kamensky. Some Wakhi literateurs, perceiving IPA as closer to English than the Perso-Arabic script, attribute some prestige to it. A similar script is being used in Gojal (Mock 1998: 36–7), and a prominent Wakhi cultural figure has published a primer for writing Wakhi (Haqiqat Ali n.d.). The symbols used in this primer are: < a, ā, à, b, č, C, ċ, d, D, đ, e, ə, f, g, γ, ħ, h, i, j, ĵ, η, k, l, m, n, o, ō, ò, p, q, θ, r, s, š, S, t, ts, u, ū, ù, v, w, x, ẋ, y, ž, z, Z >. However, others in Pakistan prefer to use a modified Perso-Arabic script. See also the Cyrillic alphabet, Table 14a in Chapter 14a Pamir languages.

2 PHONOLOGY

2.1 Inventory and distribution

2.1.1 Vowels

Wakhi dialects present some difference in their vowel inventories. The system of WW appears to have five vowel phonemes, *i u ʊ ə o*. In Payne (1989), *ʊ* = *ɨ*. Pakhalina (1983: 411) finds this sound somewhat rounded. For HW, Buddruss (1986: 28) gives the phonemes *i e ω a o u*, describing *ω* as a high back unrounded vowel. For Gojal, Mock (1998: 37) gives seven phonemes, *i u ʊ o ə e a*. Mock characterizes *ʊ* as a high central rounded vowel. For Sarikol, Gao (1985: 101–2) gives *i e ω a o u*, and ten diphthongs. In comparison:

TABLE 15.1: VOWELS

Wakhan	<i>i</i>	<i>ɪ ə</i>	<i>o u</i>
Hunza	<i>i e</i>	<i>ɔ</i>	<i>a o u</i>
Gojal	<i>i e</i>	<i>ɛ (ə)</i>	<i>a o u</i>
Sarikol	<i>i e</i>	<i>ɔ</i>	<i>a o u</i>

Hunza *e* corresponds to Wakhan *ə*: HW *yem*, WW *yəm* ‘this’, and is distinct from *i* and *a*: *ɣeʃ* ‘male (of animals)’, *ɣaʃ* ‘mouth’, *ɣiʃ* ‘ear’. Gojal also has [ə], which, however, appears to be an allophone of unstressed *e* or *a*, depending on intonational factors. As in other Iranian languages, the length distinction has been replaced by the distinction between stable vowels with little allophonic variation, and unstable central vowels which show wide allophonic variation, even loss in unstressed syllables (Payne 1981: 166).

The question of vowel quantity remains problematic. All authors note that there is considerable phonetic variation in vocalic length. Pakhalina (1975: 11; 1983: 410–12) finds six basic vowel qualities with long-short pairs for each. Griunberg and Steblin-Kamensky (1976: 542–5), however, present experimental evidence that synchronic vowel length is not phonemic, but correlates with positional and intonational factors. They suggest that there co-exist two separate phonological systems, one for native Wakhi words including old borrowings and one for (more recent) borrowings. Buddruss finds no clear minimal pairs for length (1986: 29).

2.1.2 Consonants

2.1.2.1 Overview

The inventory of consonant phonemes is the same for all WW dialects studied, for HW, and for Sarikol Wakhi:

TABLE 15.2: CONSONANTS

	Labial	Dental	Ret.	Pal.	Velar	Uv.	Gl.
Stops	<i>p</i> <i>b</i>	<i>t</i> <i>d</i>	<i>ʈ</i> <i>ɖ</i>		<i>k</i> <i>g</i>	<i>q</i>	
Affricates		<i>ts</i> <i>dz</i>	<i>ʈʂ</i> <i>ɖʂ</i>	<i>ʧ</i> <i>ʤ</i>			
Fricatives	<i>f</i> <i>v</i>	<i>θ s</i> <i>ð z</i>	<i>ʃ</i> <i>ʒ</i>	<i>ʂ</i> <i>ʐ</i>	<i>x̣</i> <i>ɣ̣</i>	<i>x</i> <i>ɣ</i>	(<i>h</i>)
Glides	<i>w</i>			<i>y</i>			
Nasals	<i>m</i>	<i>n</i>					
Liquids		<i>r, l</i>					

2.1.2.2 Fricatives and affricates

Wakhi is rich in fricatives, contrasting labio-dental, interdental, dental, palatal, retroflex, velar, and uvular positions: *sak* ‘we’, *ʃak* ‘bad’, *ʂak* ‘dew’, *x̣ak* ‘to make’.

Voicing is contrastive at each of these positions. Significant features are the triple distinction of dental, palatal and retroflex affricates and fricatives, and the distinction between the two pairs of velar and uvular fricatives, ṣ ṣ̣ and x γ .

2.1.2.3 Interdental δ

In the Yasin dialect (Pakistan) (Backstrom 1992: 68) and in most dialects of WW, the voiced interdental fricative δ is being replaced by the dental stop d : Langar *pbi δ* , WW *pbid* 'foot' (Lashkarbekov 1975: 75).

2.1.2.4 q and k

With regard to the status of q and k , while q may be indigenous in a few native Wakhi words, it is mostly confined to obvious loanwords, mainly Arabic and Turkic, but also Khovar, which may reinforce its phonemic status. In WW, q in loans alternates with k : *qəmočlkəmoč* 'bread' (GSK: 550).

2.1.3 Syllable structure and consonant clusters

Native Wakhi words consist of one or two syllables, with the exception of some three-syllable verb forms. The predominant syllable type is CV, but VC, VCC, and CVC are also frequent. In initial position, two-consonant clusters, CCV, are common: *sp-*, *st-*, *sk-*, *br-*, *dr-*, *tr-*, *gr-*, *kr-*, *bl-*, *pl-*; and some triple clusters, CCCV, are found as well: *spr-*, *str-*, *stw-*, *xtl-*: *strɪzɔg* 'large broom', *xtlivung* '*Artemisia Tourniortiana*'. In syllable-final position, up to three consonants may occur: *spundr* 'plowing implement' (GSK: 552).

2.2 Non-segmental features

Wakhi stress is strong and in most native Wakhi nominals falls on the last syllable: *kitkan* 'doll', *kitkanək* 'little doll', *kitkan-dun-ək* 'place for dolls' (Pakhalina and Lashkarbekov 2000: 175). This includes nominalizations such as infinitives and participles. Grammatical suffixes and enclitics are unstressed. In verb forms with personal endings, the stress is shifted onto the initial syllable: infinitive *wizm-ak* 'to bring', but finite *wizm-d* 'brings'. Stress is indicated here and in subsequent examples by underlining stressed vowels.

3 MORPHOLOGY

3.1 Nominal morphology

Wakhi nouns and pronouns are marked for number and case, the pronominal paradigm having adopted the nominal markers. Grammatical gender has been lost, although some word forms preserve its historical traces (Pakhalina 1989: 18–20).

3.1.1 Nouns

The major distinction is between direct (DIR) and oblique (OBL). There are two oblique cases, simple oblique (OBL₁) and focused oblique marked by a front or mid vowel (OBL₂). The 1s and 2s pronouns also have a distinct possessive/genitive form (POSS). In addition, the dative/goal case (GL) is marked by $-(V)r(k)$, and the ablative/source case

(SRC) by $-(V)n$. These four cases may co-occur with prepositions to form circumpositions, and also with a variety of postpositional adverbial and nominal elements, which show a continuum of grammaticization from transparent nominal to opaque lexicalization. There are thus several levels of case constructs. The basic case distinctions are as follows; note that the singular OBL_1 has zero-marking:

TABLE 15.3: CASE ENDINGS

	Hunza		Wakhan	
	Sing.	Plural	Sing.	Plural
<i>Level-1</i>				
DIR	—	-išt	—	-išt
OBL_1	—	-ev	—	-əv
<i>Level-2</i>				
OBL_2	-e	-ev-e	-əyl-i	-əv-əyl-i (upper/lower)
Dative/GL	-e-r	-ev-e-r	-ər(k), -(r)ək	-əv-ər(k), -(r)ək
Ablative/SRC	-e-n	-ev-e-n	-ən	-əv-ən

Level-1 forms are the base for the level-2 case markers, e.g. WW *kənd-əyl/kəndəv-əy* ‘the woman/women’; *kənd-ər(k)/kənd-əv-ər(k)* ‘to the woman/women’; *kənd-ən/kənd-əv-ən* ‘from the woman/women’. OBL_2 in both WW and HW marks: (a) some transitive past tense subjects; (b) some direct objects; (c) some genitive relations; and (d) in HW some intransitive subjects. Level-2 case markers appear to be transitional between postposition and suffixal case marker. One indication of this is that the occurrence of OBL_2 varies from dialect to dialect and even speaker to speaker, especially in HW. This area of Wakhi syntax is a prime area in which to observe language change in progress.

3.1.2 Adjectives

Adjectives are invariant. They may also be used substantively, in which case they take the regular nominal endings: *digar-ev-ər* ‘to the others’.

The comparative marker is *-tər*: *lup-tər* ‘bigger’; however, comparison is usually expressed analytically, with the plain adjective and the circumposition *tse NP-en*:

yəm vrət ts-a [vrət]-en zaq tei
 ‘this brother is younger (*zaq* “small”) than that brother’ (Lorimer 1958: I, 139).

The superlative is expressed by *kəlxət* plus the ablative *-ən*, or *kəyl-i* ‘all’:

yəm tsə kəlxət-ən dzəqlay(-tər)
 ‘she is the smallest of all’ (GSK: 564).

3.1.3 Pronouns and deixis

3.1.3.1 Personal and demonstrative pronouns

Only the 1st and 2nd person pronouns distinguish DIR, OBL_1 and OBL_2 , and also have separate possessive/genitive forms. The demonstrative pronouns have triple deixis, far,

mid, and near distance; but they do not distinguish DIR and OBL₁ in the singular. The far demonstrative pronoun often loses its deictic function, and serves as the 3rd person pronoun:

TABLE 15.4: PERSONAL AND DEMONSTRATIVE PRONOUNS, HUNZA

Hunza						
	Singular			Plural		
	DIR	OBL ₁	OBL ₂	DIR	OBL ₁	OBL ₂
1	<i>wuz</i>	<i>maž/ma</i>	<i>maž-e</i>	<i>sak</i>	<i>sak</i>	<i>sak-e</i>
2	<i>tu</i>	<i>tow/taw/ta</i>	<i>tow-el/taw-e</i>	<i>sašt</i>	<i>sav</i>	<i>sav-e</i>
3far	<i>yaw/yal</i> <i>yow</i>	<i>yaw/yal</i> <i>yow</i>	<i>yaw-e</i>	<i>yašt</i>	<i>yav</i>	<i>yav-e</i>
3mid	<i>yet</i>	<i>yet</i>	<i>yet-e</i>	<i>yet-išt</i>	<i>yet-ev</i>	<i>yet-ev-e</i>
3near	<i>yem</i>	<i>yem</i>	<i>yem-e</i>	<i>yem-išt</i>	<i>yem-ev</i>	<i>yem-ev-e</i>
POSS				POSS		
1	<i>žbi</i>			<i>spo</i>		
2	<i>ti</i>			<i>ti</i>		
3far	<i>yaw-e</i>			<i>yav-e</i>		
3mid	<i>yet-e</i>			<i>yet-ev-e</i>		
3near	<i>yem-e</i>			<i>yem-ev-e</i>		

TABLE 15.5: PERSONAL AND DEMONSTRATIVE PRONOUNS, WAKHAN

Wakhan						
	Singular			Plural		
	DIR	OBL ₁	OBL ₂	DIR	OBL ₁	OBL ₂
1	<i>wuz/wəz</i>	<i>maž/ma</i>	<i>maž-əy</i>	<i>sak</i>	<i>sak</i>	<i>sak-əy</i>
2	<i>tu</i>	<i>taw/ta</i>	<i>taw-əy</i>	<i>sa(y)išt</i>	<i>sav</i>	<i>sav-əy</i>
3far	<i>yaw/yal</i> <i>yow</i>	<i>yaw/yal</i> <i>yow</i>	<i>yaw-əy</i>	<i>ya(w)išt</i> <i>yow-išt</i>	<i>yav</i>	<i>yav-əy</i>
3mid	<i>yət</i>	<i>yət</i>	<i>yət-əy</i>	<i>yət-išt</i>	<i>yət-əv</i>	<i>yət-əv-əy</i>
3near	<i>yəm</i>	<i>yəm</i>	<i>yəm-əy</i>	<i>yəm-išt</i>	<i>yəm-əv</i>	<i>yəm-əv-əy</i>
	POSS			POSS		
1	<i>žbilžə</i>			<i>spo</i>		
2	<i>ti</i>			<i>*sav/savišt</i>		
3far	<i>yaw-əy</i>			<i>yav-əy</i>		
3mid	<i>yət-əy</i>			<i>yət-əv-əy</i>		
3near	<i>yəm-əy</i>			<i>yəm-əv-əy</i>		

*In upper WW, the DIR plural ending *-išt* is extended marginally to possessive *sa(v)išt*.

All demonstratives may take emphatic *(h)a-* (cf. Persian *ham-* ‘same’): *ha-ya vaxt* ‘that same, that very time’. *(h)a* also occurs with similar functions in the Dardic languages Pashai, Kalasha, Khowar, and Palula (Bashir 2003). Demonstrative adjectives are invariant before nouns.

3.1.3.2 Possessive forms

The predicative adjectival forms of the 1st and 2nd pers. possessive pronouns consist of the ablative of the possessive/genitive: *ṣ̣bi-n(-ən)* ‘mine’, *ti-n-ən* ‘yours’, *spo-ts(-ən)* ‘ours’.

3.1.3.3 Reflexive-emphatic and reciprocal pronouns

The reflexive pronoun is *ṣ̣at* ‘self’; the attributive adjective *ṣ̣bi*; and the predicative adjective *ṣ̣i-n-ən*. The reflexive also functions as an emphatic. The reciprocals are indigenous WW *yoman*, HW *alamān*, and borrowed *am-digar* ‘each other’.

3.1.3.4 Interrogative and indefinite pronouns and related adverbials

Interrogative pronominals include: *kuy* ‘who’, *čiz* ‘what’, *čiz-ər*, *čir* ‘why’, *kum(d)* ‘which, what kind of (of finite set)’, *tsum* ‘how much, many’, *star* ‘what (substantive)’. These forms also function as indefinites: *kuy* ‘some(one), any(one)’, *kum-yor* ‘someone’, *čiz* ‘some’ (non-specific), *čiz-i* ‘some’ (specific), *tsum* ‘some, several’, *iḡun* ‘a, some, any’ (non-specific): *iḡun kitob ma-r rand* ‘give me a book’ (GSK: 573).

3.1.4 Adpositions

3.1.4.1 Prepositions

The basic indigenous prepositions express various locative relations, usually with reference to the parameters of horizontality or verticality:

<i>sək/skə</i>	location or direction above the reference object;
<i>pə</i>	direction or location above or horizontal to reference object;
<i>tər</i>	location horizontal to (or above) reference object;
<i>dəltə</i>	direction or location (usually) horizontal to the reference object;
	location in time;
<i>(V)r(V)</i>	location below the reference object;
<i>tsə</i>	direction away from object, usually with <i>-ən</i> .

They take the OBL₁ case: *pə maž* ‘on me’, *pə daraxt* ‘up the tree’. They often fuse with their pronominal host element: *sk-av* ‘on them’ < *skə-yav*.

3.1.4.2 Postpositions

Postpositions take the possessive/genitive form of the personal pronouns: *ti mis* ‘in front of you’, *spo šiḵn āl* ‘stay with us’.

3.1.4.3 Circumpositions

Circumpositions (ambipositions) are centrally important in Wakhi, as in other Iranian languages of the region (Stilo 1987). For example, combined with the SRC marker *ə*, *pə* can mark instrument: *pə xingər-ən* ‘with the sword’ (GSK: 581), a usage found mostly in upper WW (P: 99). Another WW instrumental usage is the simple preposition *də*: *də θin yupk də sobyūn* ‘with hot water and soap’ (GSK: 639–40). In HW instrument is expressed either with *-en* (SRC) alone: *lup yār-ən* ‘with big stones’, or with the circumposition *tsə-NP-ən*: *ts(ə) tambuk-ən* ‘with a pellet bow’ (Lorimer 1958: I, 107, 116).

3.1.5 Adverbial demonstratives and adverbs

A central aspect of Wakhi nominal morphology is the fusion of the basic locative prepositions expressing the parameters of verticality and horizontality with the demonstratives to yield a complex series of locative adverbials (see GSK: 582–7 for full paradigms). For example:

TABLE 15.6: ADVERBIAL DEMONSTRATIVES (SINGULAR FORMS)

	Near	Mid	Far
Above	<i>(a)-sk-əm</i> ‘here’	<i>(a)-sk-ət</i> ‘there’	<i>(a)-sk-a</i> ‘there’
Horizontal	<i>(a)-dr-əm</i> ‘here’	<i>(a)-dr-ət</i> ‘there’	<i>(a)-dr-a</i> ‘there’
Below	<i>(a)-r-əm</i> ‘here’	<i>(a)-r-ət</i> ‘there’	<i>(a)-r-a</i> ‘there’

In general, the emphatic *(h)a-* precedes the preposition, *ha*-PREP-DEM: *(h)a-ts-ə* ‘from that/his/there < *ha-tsə-ya(w)*. In the case of reduced *yaw* > *-a*, *-n-* (or *-r-* in WW) is inserted intervocalically: *ha-ts-a-n-ən* ‘from that/him/there’.

Other adverbs are treated as lexical items, having no special morphological shape or rules of formation.

3.1.6 Numerals

3.1.6.1 Cardinal numbers

Two notable features of the original Wakhi numeral system are the remnants of a vigesimal system and the (10 + n) structure for teen numerals, both characteristics shared by a number of neighboring languages (Bashir 1988; Tikkanen 1995). However, the teens and decads larger than 20 are now only used by older speakers, and except for the single digits, Tajiki numerals are used. The original Wakhi numerals are the following (GSK: 645):

TABLE 15.7: CARDINAL NUMBERS

1–9	11–19	10–90	100+
<i>(y)i(w)</i>	<i>δas-yiw</i>	<i>δas</i>	<i>pandz-bist</i> ‘one hundred’
<i>bu(y)</i>	<i>δas-bu(y)</i>	<i>bist</i>	
<i>tru(y)</i>	<i>δas-tru(y)</i>	<i>bist-ət δas</i>	
<i>tsəbɨr</i>	<i>δas-tsəbɨr</i>	<i>bu-bist</i>	
<i>pandz</i>	<i>δas-pandz</i>	<i>bu-bist-ət-δas</i>	
<i>šaδ</i>	<i>δas-šaδ</i>	<i>tru(y)-bist</i>	
<i>(h)ɨb</i>	<i>δas-(h)ɨb</i>	<i>tru(y)-bist-ət-δas</i>	
<i>(h)at</i>	<i>δas-(h)at</i>	<i>tsəbɨr-bist</i>	
<i>naw</i>	<i>δas-naw</i>	<i>tsəbɨr-bist-ət-δas</i>	

With the numerals for ‘1’, ‘2’, and ‘3’, the consonant-final forms (-w, -y) are substantive and the vowels-final forms adjectival. Nouns usually appear in the singular after cardinal numerals.

3.1.6.2 Ordinal numbers and distributives

Ordinal numbers are formed by suffixing *-ɨng*: *buɨ-ɨng* ‘second’, and distributives by suffixing *-i*: *buɨ-i* ‘by twos’. A distributive suffix *-gōn* is attested with ‘one’: *i-gōn i-gōn* ‘one each’ (Morgenstierne 1938: 489).

3.1.6.3 Classifiers

Wakhi makes some use of nominal classifiers, loaned from Tajik: *dona* ‘thing’ (for objects), *nafar* ‘person’ (for persons), *sar* ‘head’ (for animals): *bist sar tuɨ* ‘twenty goats’.

3.2 Verbal morphology

Finite verb forms mark person, number, tense, mood, and evidentiality, which is a regular development of the perfect form. There is no morphological *voice* opposition; an action with an unspecified agent is expressed with a 3rd person plural verbal ending: *yark-ep tsar-en* ‘the work will be done’ (lit. ‘they will do the work’). Wakhi does *not* mark *aspect* inflectionally; many of the functions associated with imperfective forms in other languages are performed by the particle *-ǰ*.

3.2.1 Stem formation

3.2.1.1 Overview

There are only about 300 simple verbs in Wakhi (GSK: 582). The rest are conjuncts of nominal elements and a small set of simple verbs (“light verbs”), which define the transitivity or intransitivity of the conjunct. The most frequent of these are *tsar-* ‘do’, *goɨ-* ‘do, make’, *kaɨ-* ‘put’, *xaɨ-* ‘pull’, *di-* ‘beat, hit’, *δɨr-* ‘have’ for transitives; and *wots-*, past *vit-* ‘become’ for intransitives.

Wakhi verb forms are based on three stems: present, past, and perfect. Based on the morphological relation between present and past stems, there are regular and irregular verbs.

3.2.1.2 Present and past stems

In past stems, the most numerous group of regular verbs simply suffix *-t/d* to the present stem: pres. *čok-*, past *čok-t-* ‘cut’; pres. *al-*, past *al-d-* ‘remain’. Another group changes *ɪ* to *o*: pres. *bɪn-*, past *bon-d-* ‘throw’; pres. *δiw(ɪ)y-*, past *δowoy-d-* ‘steal’. These include derived causatives. The remaining verbs show various types of stem alternation: pres. *yund-*, past *yut-* ‘take away’; pres. *θaw-*, past *θət-* ‘burn’.

Some irregular verbs have past stems in *-n*: pres. *δits-* ‘milk’, past *δəy-n-*.

3.2.1.3 Perfect stems

Perfect stems are formed in two ways: (1) past stem + *-k/g*: *puv-* ‘drink’, past *pit-*, perf. *pit-k-*; (2) present stem + *-ətk*: *bɪn-* ‘throw’, past *bon-d-*, perf. *bɪn-ətk*; *wɪdɪw-* ‘take, seize’, past *wodort-*, perf. *wɪd(ɪ)r-ətk-*. Class 2 includes most of the derived transitives and causatives.

3.2.1.4 Causative stems

Causative stems are formed regularly: pres. + *-ɪw/ɪw*, past + *-owd/owd*: *čərm-* ‘enter’ > *čərm-ɪw-*, past *čərm-owd* ‘lead into’; *puv-* ‘drink’ > pres. *puw-ɪw-*, past *puw-owd-* ‘cause/give to drink’.

3.2.2 Non-finite forms

3.2.2.1 Verbal nouns or infinitives

There are two verbal nouns or infinitives: (1) present stem + *ak*, (2) present stem + *-n* (or *-g*, with *n*-final stems): *kəmy-ak* and *kəməy-n* ‘to desire’; *wɪn-ak* and *wɪn-g* ‘to see’. Most verbs have both infinitives, but some only one or the other. Forms in *-n* predominate in upper Wakhan, forms in *-k* elsewhere in Wakhan. In contemporary HW, *n*-forms predominate, perhaps partially due to recent influence from Shina and Urdu. Where both forms exist, it appears that the *ak*-forms are more agent-focussed, as distinguished from the activity focus of *n*-forms. Griunberg and Steblin-Kamensky call WW *ak*-forms “infinitives”, and *n*-forms “action nouns”. The extent to which this distinction is maintained or valid for Hunza remains to be investigated. As nouns, they may assume any syntactic function, and also serve in various types of nominalized clauses (see 5.3–6).

3.2.2.2 Participles

3.2.2.2a Present participle

A present participle is derived from the verbal noun by the suffix *-kɪzɡ*. With *k*-infinitives, the suffixal *-k* is lost: *lədown-kɪzɡ* ‘giving, giver’ < *lədown* ‘to give’, *šak-ɪzɡ* ‘working, worker’ < *šak* ‘to do’ (P: 88). The present participle also functions substantively as an agent noun.

3.2.2.2b Perfect participles

There are two perfect participles: one is identical with the perfect stem and functions either adjectivally: *de- $\dot{t}k$ $\dot{s}apik$* ‘burned bread’, *mer- k fuk s* ‘dead snake’, or substantively: *no- $\dot{w}in$ - $\dot{e}tk$* ‘one who hasn’t seen’; the other consists of the perfect stem plus the adjectival ending *- $\dot{b}ing$* (upper WW), *- $\dot{a}ng$* (lower WW), *- $\dot{i}n$* (middle and lower WW) (GSK: 601): *yak \dot{s} - $\dot{b}w$ - $\dot{a}tk$ - $\dot{i}n$ $yupk$* ‘boiled water’ (P: 89).

3.2.3 Person marking and the verb ‘to be’

3.2.3.1 Overview

Person and number are marked by personal endings in the present-future. Pronominal clitics mark the subject in all past and perfect tenses of both transitive and intransitive verbs. These clitics are homophonous in most cases with the copula and the personal desinences (Table 15.8).

TABLE 15.8: COPULA, PERSONAL ENDINGS AND CLITICS

	Hunza			Wakhan		
	Copula	Pres./Fut.	Past (clitic)	Copula	Pres./Fut.	Past (clitic)
1s	<i>-em</i>	<i>-em</i>	<i>-em</i>	<i>-am</i>	<i>-am</i>	<i>-am</i>
2s	<i>-et</i>	<i>(-o, -i)</i>	<i>-et</i>	<i>-at</i>	<i>(-ayli)</i>	<i>-at</i>
3s	<i>-o, -it</i>	<i>-t/d</i>	<i>-o, -i</i>	<i>-o, -it</i>	<i>-t</i>	<i>-o, -ayli</i>
1p	<i>-en</i>	<i>-en</i>	<i>-en</i>	<i>-an</i>	<i>-an</i>	<i>-an</i>
2p	<i>-ev</i>	<i>-it</i>	<i>-ev</i>	<i>-av</i>	<i>-it/-av</i>	<i>-av</i>
3p	<i>-ev</i>	<i>-en</i>	<i>-ev</i>	<i>-av</i>	<i>-an</i> (upper/lower)	<i>-av</i>

Except for the 3s, which is optional and cannot attach to the verb, the pronominal clitics may attach to any constituent in the sentence, but are most often found attached to the first constituent of the clause. They may appear more than once in a clause.

3.2.3.2 Existential verb ‘to be’

There are four indicative tense forms: pres. *tei*, past *tu*, perf. *tuvetk*, past perf. *tuvetu*. These forms are unmarked for person or number, which are indicated by the pronominal clitics. The suppletive subjunctive base is HW *h $\dot{u}mi$ -/WW (y/h) $\dot{b}m$ ($\dot{b}l$) \dot{y} -*, of which only the present/future forms are attested (see Table 15.9).

TABLE 15.9: EXISTENTIAL VERB 'TO BE'

	Hunza			Wakhan		
	Present	Past	Subjunctive	Present	Past	Subjunctive
1s	<i>tey-em</i>	<i>tu-m</i>	<i>humi-em</i>	<i>təy-əm</i>	<i>tu-m</i>	<i>ыты-эм</i>
2s	<i>tey-et</i>	<i>tu-et</i>	<i>humi</i>	<i>təy-ət</i>	<i>tu-t</i>	<i>ыты-у</i>
3s	<i>tey(-it)</i>	<i>tu</i>	<i>hum(ыli)-t</i>	<i>təy(-it)</i>	<i>tu</i>	<i>ыты-t</i>
1p	<i>tey-en</i>	<i>tu-n</i>	<i>humi-en</i>	<i>təy-ən</i>	<i>tu-n</i>	<i>ыты-эн</i>
2p	<i>tey-ev</i>	<i>tu-v</i>	<i>humi-it</i>	<i>təy-əv</i>	<i>tu-v</i>	<i>ыты-эв</i>
3p	<i>tey-ev</i>	<i>tu-v</i>	<i>humi-en</i>	<i>təy-.: v</i>	<i>tu-v</i>	<i>ыты-эн</i>

In copular function, the personal clitics can also attach to another constituent: *-em tei*, *-et tei*, *-en tei*, *-ev tei*.

The corresponding non-stative verb 'to become' has the following basic forms: pres. *wots-*, past *vit-*, perf. *vit-k*, and the two infinitives *wots-ak* and *wots-n*.

3.2.4 Aspectual-modal affixes and negation

3.2.4.1 Aspectual clitic -š

The aspectual clitic *-š* appears in numerous variants which may be summarized as *-(V)š(V)(k)*. Also, with consonant-final stems, a vowel harmonizing with the stem vowel of the host word appears; with vowel-final stems, *-š* is either directly attached: *tu-š* or is connected by a homorganic glide and vowel: *tu-w-uš* 'you'. *-š* combines with the present-future and with the past to express specificity and immediacy, either present relevance or immediate future, and various imperfective meanings. It also expresses real conditions with the present-future, and irrealis conditions with the past and with the distant perfect (see the discussion of the individual tenses in section 3.3.5). *-š* is usually enclitic to the first sentence constituent; however, it may attach to other constituents including the verb, and may also occur more than once in a clause:

yiw-iš ma-r kət-ə, nəy, tow-əš be yow-em ti zəman-əš be yow-em
 'throw one to me or else I am going to eat (*yow-*) both you and your child' (Mock 1998: 451).

3.2.4.2 Particle *-(V)p*

The particle *-(V)p* specifies non-past situations as non-immediate, i.e. as future or as tenseless/generic (see section 3.2.5.2 Present-future).

3.2.4.3 Negative particles

The negative particles are pre-verbal *nə* for indicative and *mə* for non-indicative, *nəy* and *məy* post-verbally. The negated copula is *nast* 'is not'. All nominal forms of the verb take *no-*: infinitive *no-diš-ak* 'not to know', perfect participle *no-diš-etk*, adjectival *no-diš-etk-in*.

3.2.5 System of tense, mood, and aspect

3.2.5.1 Overview

The fundamental tense distinction in Wakhi is between past and non-past, i.e. present/future. Perfect forms indicate not tense but resultativity, stativity, or inferentiality. Non-indicative mood is morphologically distinct only in the imperative, which is zero-marked, and in the forms of the subjunctive of ‘be’, HW *humi-/WW (y/h) бытыу-*. Other aspectual and modal functions are expressed contextually by the tense forms described below and/or by the aspectual clitics.

3.2.5.2 Present-future

The present-future tense consists of the present stem with the personal endings. In contrast to the pronominal clitics, these attach only to the verb, and appear only once in the clause. This tense refers to all non-past events and also functions as a historical present.

3.2.5.2a Subjunctive function without particle

Without a specifying aspectual particle, the present-future conveys various subjunctive meanings involving potential action, including suggestion, hortation, and a positive or negative wish:

- wuz kuy-e rem-im*
‘whom should I send?’;
salim [kum-ər] er gilt me wizi-t
‘[hopefully] Salim won’t come (down) to Gilgit’;
i wuz ыmy-əm, i kəmpir
‘would that there would be only me and the old woman’ (GSK: 635);
zman ya-r kam hūmī-t
‘she should have few children’ (lit. ‘let there be to her’) (Reinhold 1998: 345).

With 2nd person subjects, this form constitutes an imperative:

- tr-a mə-reč ʃaq-ep wots*
‘don’t go there, you will slip’.

3.2.5.2b Specificity and immediacy with -ʃ,

With the particle -ʃ, the present-future expresses specificity or immediacy, either present relevance or immediate future:

- wuz-eʃ coy pev-em*
‘I am drinking tea’;
a-ya xalg [naql-əʃ tsə tsar-əm]
‘that man [(whose) story I am telling]’ (GSK: 637);
wuz-eʃ [angrezi nivis-en]-iʃ diš-em
‘I know [(how) to write English]’.

3.2.5.2c Future and tenseless/generic function with -(V)p

With the particle -(V)p, the present-future specifies future or tenseless/generic. Pakhalina (1975:75) describes it as specifying the completion of an action in future time. GSK do

not find the future meaning in their materials and give only one example of it, in which the reading is of a timeless generic situation:

xalg-əp a-zī nə gox̄-t
 ‘a person never does like that (*a-zī*)’ (GSK: 654).

-(*V*)*p* appears more frequently in HW than in WW. Lorimer’s materials show frequent use of it to indicate both generic and future actions. In Mock’s texts and my materials, sentences referring to specific future actions appear consistently with -(*e*)*p*. It occurs in both indicative and non-indicative contexts:

jald tuy-p mə tsar-ən
 ‘they should/will not marry soon’ (Reinhold 1998: 344);
tu-p žī xun-e batken reč
 ‘come/you will come to my house!’ (*batken* ‘up to’).

3.2.5.3 Past

The (simple) past tense consists of the past stem, with the subject indexed by the pronominal clitics. When no other particle or personal marker is attached to it, the past stem ends optionally in *-i/ə(y)* (upper WW/lower WW), or *-e/a* (HW):

wuz-əm taw-əy wind-əy
 ‘I saw you’ (Payne 1980: 182).

In Hunza (Passu), it appears that this past tense marker is strongly preferred, perhaps even obligatory with intransitive past tenses

*wuz-əm kand-a (*kand)*
 ‘I laughed’.

3.2.5.3a Anterior to present and conditional perfective

The (simple) past refers to events occurring prior to the time of the speech act, but it also functions as perfective in conditional clauses (see 5.7).

3.2.5.3b Imperfective meanings and conditional irrealis with *-š*

With the particle *-š*, the past expresses various imperfective meanings including progressive, durative, and habitual, as well as unrealized actions, and functions in irrealis conditional clauses (see 5.7.2):

maž-eš yark kert
 ‘I was working’;
yaw-eš sk-i yafč lup mulk hukumat kert
 ‘he ruled over a very great kingdom’;
ar ruz-əš tə ku-š rejd-əy
 ‘he would go to the mountain every day’ (GSK: 625).

3.2.5.4 Distant past

The distant past (“pluperfect”) is derived in two ways:

- (1) from the perfect base by replacing *-k/g* with the past tense form of ‘to be’, HW and lower WW *-tu*, upper WW *-tɔɪw*: pres. *tsar-* ‘do’, past *kərt*, perf. *kər-k*, dist. past *kər-tu*;
- (2) present base + *-ətutɔɪw*: pres. *wɔɪzɔɪm-* ‘bring’, past *wɔzom-d*, perf. *wɔɪz(ɔɪ)m-ət-k*, dist. past *wɔɪzm-ət* (P: 77; GSK: 598–9).

3.2.5.4a Anterior to past event and distant past

The distant past refers to a past event prior to another past event, or to an event in the relatively distant past:

ʒɔɪ vrɔɪt pɑd wəs-tu
 ‘my brother came last year’ (GSK: 626);
wuz də ɣi diwɔɪl bən mɔɪʃ-tu-m ʃat-əɣ, ɣaw mɑʒ-əɣ nə wind-əɣ
 ‘I hid/had hidden myself at the base of a wall, and he didn’t see me’ (GSK: 626).

3.2.5.4b Unrealized actions and irrealis conditional with *-ʃ*

With the particle *-ʃ*, the distant past expresses unrealized action, and appears in irrealis conditional clauses:

tsə dɪŋg-ən-ən ɣɑw-i qəɪɪb ʃi- tɔɪw-əw-ʃək
 ‘they almost killed him by beating him’ (P: 84).

3.2.5.5 Perfect

The perfect consists of the perfect stem plus the pronominal clitics. It functions as follows: the basic indicative function of the perfect is resultative-stative:

δɔɪr kɔɪnd vit-k
 ‘the sickle has become dull/is dull’ (P: 83),

from which develop inferential and mirative senses. Compare the simple past and perfect meanings in the following sentences:

simple past:	<i>salim peʃawar reʃd-e</i> ‘Salim went to Peshawar’ (direct knowledge);
perfect (inference):	<i>salim-i peʃawar reʃ-k</i> ‘apparently Salim went/has gone to Peshawar’;
perfect (mirative):	<i>tu ɣa bilaʃta-ət tɔɪw-ət-k</i> ‘apparently you are very stupid!’ (GSK: 625).

3.2.5.6 Perfect subjunctive

Combined with the perfect, *(h)umɔ*, i.e. non-indicative ‘be’, contributes subjunctive senses to the basic resultative meaning of the perfect. It appears as an optative, in a future perfect sense, and in conditionals:

salim kum-er er gilt mə wez-g humu-t
 ‘I hope Salim hasn’t come down to Gilgit’;
piga batkenit wuz yem yark-e ʃet-k humuy-em
 ‘by tomorrow I will have finished this work’.

3.2.5.7 Distant perfect

The distant perfect is formally a double perfect, consisting of the perfect followed by *tuwɛtk*, the perfect of 'to be'.

3.2.5.7a Unwitnessed/inferred action

This form temporally distances unwitnessed/inferred action:

yukš-ī a-yā xālg dəyt ki yaw tat diy-ətk tuw-ətk
 'the same person killed the mountain goat who (previously had) killed his father' (P: 83).

3.2.5.7b Irrealis wishes in past time

The distant perfect also appears with irrealis wishes in past time:

koški tu mə rəx-k tɔw-ətk
 'if only you had not gone' (P: 83).

4 SYNTAX AND USE OF FORMS

4.1 Coordination

A native overt coordinating element is the clitic *-et* (HW)/*-ət* (WW), which attaches to the final element of the first conjunct, either a NP: *wuz-ət tu* 'you-and I', or a clause (see 4.1). Paired coordinating elements include *nə . . . nə* 'neither . . . nor'; *bɛ . . . bɛ* and *am . . . am* 'both . . . and' (P: 110).

4.2 Noun phrase structure

Wakhi is a head-final language. The unmarked word order in the noun phrase is: determiner, possessive, numeral, adjective, head noun (DET – POSS – NUM – ADJ – N).

ts-əm ži bu lup pətr-(əv)-ən
 from-this my two big son-s-from
 'from these two elder sons of mine' (GSK: 659).

4.2.1 Genitive phrases

Dependent noun phrases precede the head, and are either in OBL₁ (a) or OBL₂ (b). The form with OBL₁ is the native Wakhi structure, and is still the most common. 1s and 2s pronouns have a special genitive form: *ži xun* 'my house'. Less frequent in both Hunza and Wakhan is an ablative construction, approximately 'of NP his X' (c). The Persian *ezāfe* construction (d) may be found as well:

(a) <i>tat xun</i>	'father's house'	<i>vrɔt-əv xun</i>	'the brothers' house'
(b) <i>tat-i xun</i>	'father's house'	<i>dəɾaxt-əv-i palč-išt</i>	'the leaves of the tree'
(c) <i>tat-ən yaw xun</i>	'the father's house'	<i>dəɾaxt-əv-ən yāw palč-išt</i>	'the leaves of the trees'
(d) <i>xun-i tat</i>	'father's house'.		

4.2.2 Definiteness marking

Non-specific indefinite noun phrases are unmarked. Specific, referential indefinites are usually marked by *(y)i(w)* ‘a, one’, which may precede the noun or follow as an enclitic, as in the formulaic folk tale introduction:

tu nə-tu, i potšo tu
 ‘there was or there wasn’t; there was a king’ (GSK: 561);
kum dbızd-i wız-it
 ‘a (some) thief is coming’ (GSK: 561).

Definiteness may be marked by the far demonstrative, which often loses its deictic force and functions as a definite article.

4.3 Clause structure and word order

The unmarked clausal constituent order is SOV. The topic occupies sentence-initial position, while the focus position is pre-verbal:

wuz-em tə xun reyde
 ‘I went home’ (unmarked order);
tə xun-em wuz reyde
 ‘it is I who went home’ (answering the question, ‘who went home?’).

4.3.1 Clitics

The unmarked position for all clitics is second position, attaching to the first major sentence constituent. These include the tense/aspect clitics *-(V)š(V)* (immediate/imperfective) and *-(V)p* (non-immediate/future), the modal clitic *ni* ‘perhaps’, and the pronominal clitics. The position of the pronominal clitics is variable, and has semantic and pragmatic correlates. Compare the following sentences in which the shift of the pronominal clitic to the past form of the verb produces the reading of imminent future action:

wuz-em tə xun reyde
 ‘I went home’;
wuz tə xun reyde-em
 ‘I’m going home (immediately)’.

4.3.2 Questions

Yes-no questions are marked with a sentence-final particle *-o* (upper WW)/*-a* (lower WW; HW). WH-question words are normally immediately pre-verbal.

4.3.3 Nominal sentences

In present tense nominal sentences, the pronominal clitics sometimes perform the copular function. When existence is emphatically affirmed or questioned, the copula *tei* ‘is’ can co-occur with the clitics. In negative sentences, the clitic is obligatory (except for 3s), and attaches either to the negative copula *nast* ‘is not’ or to another constituent:

ži tat pə xun
 'my father is at home' (no clitic) (GSK: 579);
yem čuan -i trok
 'this apricot (tree) is bitter' (clitic *-i*);
miskin-ar roḍow-i baf yark
 'to give to the poor is a good deed' (clitic *-i*);
da pasu-ev dzūŷ-išt tei
 'there are yaks in Passu' (clitic *-ev*; *tei* asserts existence);
dziring xak-i baf nast
 'it is not good to do like this' (clitic; *nast* denies identity).

In the past tense, *tu* 'was' is obligatory, and co-occurs with the pronominal clitics.

4.4 Case marking

4.4.1 Subject marking

Subject marking in Wakhi depends on tense: past or non-past, and sometimes on transitivity. It also involves a choice among at least three ways of indicating the subject:

- (a) personal ending (non-past), or pronominal clitic (past) alone (or zero for 3s);
- (b) noun phrase alone (past);
- (c) noun phrase plus personal ending or clitic; and, with (b) and (c), a choice between DIR or OBL case.

4.4.1.1 Non-past tenses

In both Wakhan and Hunza, subjects of both intransitives and transitives are (usually) DIR in all dialects, and subject agreement is marked on the verb by person-number markers:

*wuzl*maž rəč-əm* 'I am going';
*wuzl*maž taw-i win-əm-əš* 'I see you' (Payne 1980: 180).

4.4.1.2 Past tense and perfect

Case marking of nominal and pronominal subjects varies by dialect:

4.4.1.2a Lower dialect of Wakhan W (Namadgut)

Here all past tense and perfect subjects are DIR, regardless of transitivity, and definite direct objects are OBL₂, yielding a NOM-ACC type pattern in all tenses:

intransitive: *yawišl*yav pə kbik rəŷd-əv* 'they went to the spring';
 transitive: *wuzl*maž taw-i (OBL₂) wind-əm* 'I saw you' (Payne 1980: 180).

4.4.1.2b Upper dialect of Wakhan W

Here transitive subjects are oblique:

maž taw-əy wind-(-əy) 'I saw you' (Payne 1980: 180).

With 1s and 2s pronouns, the only singular forms that distinguish direct and oblique, both transitive and intransitive subjects may be marked either DIR or OBL₁:

maž (OBL₁) *rəʔd(-əy)* ‘I left’ (Payne 1980: 180–1).

4.4.1.2c Hunza W

Here marking of past-tense subjects is similarly complex.

- (a) *Intransitive* subjects are usually DIR in isolated non-contextualized sentences:

wuz-m-eš tə-xun reʔd-e ‘I was going home’ (DIR and clitic).

- (b) *Transitive* nominal and pronominal subjects, however, can be DIR, OBL₁, or OBL₂, depending on grammatical person and number, referential status, and discourse variables.

Case marking thus depends on both discourse and syntactic factors, with OBL noun phrase subjects more likely to appear when one or more of the following conditions obtain:

- (a) the subject/agent is different from that of the previous action;
- (b) the specific identity of the agent is to be stressed;
- (c) the action is perceived as volitional and active rather than “passive”.

4.4.1.3 OBL₁ and OBL₂

Further, there are two OBL cases: the clearly marked OBL₂, and OBL₁ which is overtly identical with the DIR in singular nouns and in most pronouns (see sections 3.1.1, 3.1.3 on nouns and pronouns). In general, OBL₂ has two functions:

- (a) as grammatical OBL case where the DIR/OBL₁ distinction is not marked; and
- (b) to mark discourse prominence.

The appearance of OBL₂ on both subjects and direct objects is thus conditioned partly by grammatical constraints, and partly by discourse semantics.

In non-contextualized sentences elicited in 1986 and 1989 from two Wakhi speakers of Passu, definite nominal and pronominal subjects of transitive past tenses appear consistently with OBL₂:

kaš-e (OBL₂) *maž wind* ‘the boy saw me’;
māž-el̄taw-el̄yow-e (OBL₂) *šapik yit-ik*
 ‘I/you/(s)he has/have eaten food (lit. bread)’.

However, OBL₁ instead of OBL₂ in the following interrogative sentence points to the role of discourse factors:

taw (OBL₁) *šapik yit-k-a?*
 ‘have you eaten (food)?’.

Lorimer noted that OBL₂ marking is optional with OBL₁ pronominal subjects in transitive past tenses, and contrasts the emphatic effect of OBL₂ vis-à-vis the pronominal clitic plus DIR subject-marking option (Lorimer 1958: I, 97–8). A tabulation of subject marking strategies in three of Lorimer’s texts shows that of 54 OBL₂ subjects, 49 of them

occur when there is a change of discourse subject (Bashir 1986). Contemporary texts also show this effect of discourse factors; compare the first two sentences below, both with transitive past tense subjects from the same text. In the second, OBL₂ focuses attention on the identity of the fox and marks a change of subject from the previous sentence:

bu-i naxčir ts-an yit (DIR, clitic *-i*)
 ‘the fox ate two of them’ (Mock 1998: 451);
naxčir-əy ‘*hum hum*’ *xən-etk*, ‘*yi loi bešan-i*’ *xənetk* (OBL₂)
 ‘the fox said, “Hum hum, sing it once more” he said’ (Mock 1998: 452).

Further, in Hunza it appears likely that the behavior of OBL₂ is now being influenced in bilingual speakers by the functions of the Urdu postpositions *ne*, which marks transitive subjects of perfective tenses, and *ko*, which marks specific direct objects and indirect objects.

4.4.1.4 Oblique intransitive subjects

Under certain conditions, even intransitive subjects may be OBL. In upper WW with the 1s and 2s pronouns, past intransitive subjects may be OBL₁:

maž rəyđ(-əy)
 ‘I left’.

Examining the occurrence of OBL subjects in published WW materials, Bashir (1986) argues that in cases where both DIR and OBL subjects are possible, the case marking of past-tense subjects is conditioned by semantic and discourse factors.

In Hunza, this occurs in both past and non-past sentences. Thus, in the following present tense example, contrastive focus is on *my daughter*, as opposed to the servant girls in the story:

[ž i dəyđ]-ε (OBL₂) *dr-em-εp wεrεš-t*
 ‘**my daughter** will remain here’ (Lorimer 1958: I, Text 4, #74).

In the following past tense sentence, OBL₂ points up the contrast of the grandmother’s reaction with the attitude of the previous speaker, her grandson:

[xəy mum]-əy (OBL₂) *həyron vit-k*
 ‘and (**his**) **grandmother** was/became amazed’ (Mock 1998: 454).

4.4.2 Direct object marking

4.4.2.1 Wakhan Wakhi

In WW, case-marking of direct objects depends on their referential status. Nominal direct objects are zero-marked if indefinite, OBL₂ if definite:

žεč rand ma-rək ‘give me **bread**’ (zero-marked) (GSK: 562);
i loy ayət awqot-əy (OBL₂) *mar-ək wəz-əm*
 ‘please bring me **that food**’ (GSK: 563).

Pronominal direct objects are either OBL₁ or OBL₂:

tu maž (OBL₁) *yund*, *wuz tôqa wəšy-əm*
 ‘take me (with you); alone I am afraid’ (P: 53);
šač taw-əy (OBL₂) *dəyt məy*
 ‘(look out) lest the dog bite you!’ (*məy* = ‘may not’) (P: 51);
mauzman yaw-əy (OBL₂) *pə xingar-ən dəyt-əy*
 ‘Mauzman struck him with (his) sword’ (GSK: 581).

4.4.2.2 Hunza Wakhi

In HW, direct object marking depends mainly on tense. In non-past tenses, pronominal and definite or specific nominal direct objects appear regularly as OBL₂, and non-specific direct objects are zero-marked:

kaš-ep maž-e wind
 ‘the boy will see **me**’;
kaš-ep ya šač-e dič-t
 ‘the boy will beat **the dog**’;
kuy-ep yem-e gox-t
 ‘who will do **this (task)**?’;
sak-eš šapik yaw-ən
 ‘we are eating (**bread**)’.

For past tenses, comparison of elicited sentences with texts reveals a continuum of object-marking strategies. Context-free elicited sentences show a pattern in which a maximally distinguishing morphological marking strategy employing both OBL₁ and OBL₂ emerges. In past tenses, direct objects are OBL₁ (formally identical with the DIR in nouns and 3rd person pronouns), providing contrast with the consistently OBL₂-marked subjects:

kaš-e (OBL₂) *maž* (OBL₁) *wind*
 ‘the boy saw **me**’;
taw-e (OBL₂) *ya* (OBL₁) *žu šič remet*
 ‘**you** sent him to **me**’;
kaš-e (OBL₂) *ya šač* (OBL₁) *dičt*
 ‘the boy beat **the dog**’.

However, OBL₂ exists as an option, depending on referentiality status and discourse factors, including choice of subject marking and degree of contextualization, which may even result in the same direct object marking pattern in present and past. Texts reveal the variation expected from richly connected context and inter-speaker variability.

4.4.3 Causee marking

4.4.3.1 Pronominal causees

Pronominal causees are normally marked as direct objects, and accordingly are in the OBL₂ in non-past, and OBL₁ in past tenses:

yaw-ep taw-e (OBL₂) *reδ-ev-id*
 ‘he will make you run away’;
yaw-e ta(w) (OBL₁) *reδ-ev-ovd*
 ‘he made you run away’.

With causatives of “ingestive” verbs such as ‘eat, drink’, the human causee takes the dative/goal:

wuz tow-ər šapik yuv-um
‘I will feed you’.

4.4.3.2 Causees with analytical causatives

In addition to morphologically derived causatives, causatives of transitives are also formed analytically, with a finite form of the verb *rem-* ‘send’ and the infinitive of the complement verb:

wuz taw-ə [loqpar wəzdɨ-k] rem-im
‘I will get you [to wash the clothes]’;
wuz-ep [ti yark-e xak] rem-im
‘I will have [your work done]’.

If the specific identity of the causee is to be stressed, it may also be indicated by the instrumental circumposition *tsə NP-ən*:

maž-e tsə taw-en [loqpar wəzdɨ-k] remet
‘I got **you** [to wash the clothes]’.

4.5 Semantics and use of verb forms

4.5.1 Possession

Existential possession is expressed by the ablative-comitative form in *-(V)n*, with or without *tse* ‘from’:

tī-n kəž təy-a ‘Do you have a knife’ (GSK: 568),
žū-n-ən dušman yafč tei ‘I have many (*yafč*) enemies’ (Lorimer 1958: I, 106).

Such ablative-comitative forms are also used predicatively:

yəm kitōb tī-n-ən
‘this book is yours’ (P: 55);
yəm xun-i ha-yā halg-ev-en [kumd-ar ki sak-e xi degit detk]
‘this house **belongs to** the people [to whom we gave our daughter]’.

4.5.2 Aspectual-phasal constructions

Most phasal constructions involve infinitival clauses.

4.5.2.1 Prospective action

Prospective action ‘to be about to’, is expressed by ‘to be’ and the dative/goal case of the infinitive:

wuz-əm [ts-am-ən kitâb wuzum-n]-ər
‘I (am) about to [bring the book from him]’;
čiz-em [xan-ak]-ar tu
‘what was (*tu*) I saying?’ (lit. ‘about to say’).

Prospective action may also be expressed by the present participle in *-kbiʒg* from the infinitive:

wuz-əm [rəç-n]-kbiʒg
 'I am going, about to go' (P: 88).

4.5.2.2 Intended action

Pakhalina notes a specialized participial form consisting of infinitive + *-mij*, which indicates *intention to begin* an action: *tuk-n-mij* 'preparing/intending to go/leave' (Pakhalina 1975: 89, 221).

4.5.2.3 Inception

Inception, the beginning of an action, can be expressed in several ways, depending on the dialect:

- (a) *wots-*, past *vit-* 'become' with a locative infinitive construction with *dər* 'into', or *šik* 'to':

[yaw wəzy-ak bərobar] dər yark tsər-ak vit-əy
 [as soon as he came] she began to work' (P: 86)
yaw-i šik [yark xā-k] vit-k
 'he has begun, began to work';

- (b) *wots-*, past *vit-* 'become' with *pərsim-i* + infinitive (*ezāfe* construction):

baytguy-iš [pərsim-i xā-ak] -əv-vit
 'the birds began to sing' (-əv *vit* 'they became') (GSK: 628);

- (c) *wədur-n*, *wədrək* 'to seize, catch' + infinitive (cf. Tajik *gereftan*):

[xōn xā-k]-əm wədort
 'I began to build a house' (Lorimer 1958: I, 189);

- (d) *di-*, *dəyt* 'beat, hit' + plus infinite (cf. Tajik and Persian *zadan*, also Khowar *dik* 'to beat'):

yəm kəmpirzān bə dəyt-i [nəlv-n]
 'this/the old woman also began to cry' (GSK: 630).

4.5.3 Modal constructions

Similar to phasal constructions, most modal constructions involve infinitival clauses.

4.5.3.1 Deontic modality

Deontic modality, i.e. necessity or obligation, may be expressed in two ways:

- (a) *wots-*, past *vit-* 'become' + infinitive:

piga sahār-ep sak [jald giʒ-en] wots-en
 'we will (-ep) have to get up early tomorrow'
 (lit. 'early tomorrow we [are to get up quickly]).

(b) sequence of two finite clauses:

rat-i majbur vit

‘he had to give (it)’ (lit. ‘he gave it, he **was obliged to**’) (Mock 1998: 459).

4.5.3.2 Presumption

The presumption of an action may be expressed by *ni* ‘perhaps’ with the past tense:

maž-e ni xû kitob kuy-er det

‘I must have given my book to someone’ (lit. ‘perhaps I gave’).

4.5.3.3 Ability

4.5.3.3a Wakhan Wakhi

Wakhan Wakhi uses several constructions, including the so-called “potential” construction:

(1) *kə(r)tsər-*, past *kə(r)kaŋn-* ‘be able to’ + infinitive, probably an original Wakhi usage

wuz [nɔw-ak] nə kətsər-əm

‘I am not able [to cry]’ (GSK: 630).

(2) *wots-*, past *vit-* ‘become’ + infinitive

yaw bland, xalg [kɪ-ak] nə-wost

‘it (wall) is high; a man/person can not [come down]’

(lit. ‘it is not possible for a man to come down’) (GSK: 627).

Constructions (1) and (2) correspond to the “potential” construction in Khotanese and Tumshuqese (see Chapter 7) and several other Old, Middle and New Iranian languages. The verbs **kar-* ‘make, do’ and its counterpart **baw-* ‘become’, or their equivalents, are used to express the potentiality of an action, mostly distinguishing transitive and intransitive action.

(3) *bas-wəzɣ-*, past *bas-wəzd-* ‘be able’ + infinitive

wuz [jɔy-ak] nə-bəs-wəzɣ-əm

‘I can not [read]’ (GSK: 629).

4.5.3.3b Hunza Wakhi

Ability can be expressed in HW by *bas wez-* ‘to be able’ + infinitive, optionally marked by the OBL₂:

wuz-ep taw-er xû kitob [redɔw-(e)] bas wez-im

‘I will (-ep) be able [to give] my book to you’.

5 COMPLEX SENTENCES

5.1 Coordinate clauses

In the most common original Wakhi strategy, a sequence of two actions by the same agent is expressed by a sequence of finite verbs in the same tense with no overt coordinating element:

kampir tillah dōžd-ei, tagd-ei
 ‘the old woman took the gold coins, and left’ (Shaw 1876: 173).

The native coordinating element *-ət* (WW)/*-et* (HW), which cliticizes to the first conjunct, appears when the agents of the coordinated clauses differ:

wuz məryov rəč-əm-ət yaw dušanbi rəš-t
 ‘I am going to Murghab, and he is going to Dushanbe’ (Payne 1989: 441).

In sentences like the preceding, either forward or backward deletion of the identical verb is possible.

Upper WW and HW also employ *ǰə* (WW)/*ǰe* (HW), which typically indicates temporal succession in coordinated clauses (P: 111); the subjects of the clauses may be the same:

wuz wez-im ǰe šapik yaw-em
 ‘I will come, and (then) eat’; or different:
wuz-əm wezd-a ǰe taw-ə šapik yit
 ‘I came and (then) you ate’.

A borrowed construction with *wəz* ‘again, and, but’ occurs in both HW and WW:

ži virut-e lahor joy-ītk, wəz ži xuy-e islamabad joy-ītk
 ‘my brother studied in Lahore, but my sister studied in Islamabad’.

yālyo ‘or’ and *lōkin* ‘but’ are borrowed Persian elements.

5.2 Subordination

The principal native subordination marker is *tsə* (pre-verbal)/*tsəy* (clause-final). Wakhi has also borrowed Tajiki clause-initial *ki*, which, like *tsə*, introduces a wide range of subordinate clauses, including relative clauses, temporal, conditional, purpose and reason clauses, and complements of verbs of cognition. Clause-initial *agar* ‘if’ is also borrowed from Persian.

5.3 Relative clauses

5.3.1 Finite relative clauses

Finite relative clauses (RCs) are constructed in several ways.

5.3.1.1 Relative noun in matrix clause

If the relative noun appears in the matrix clause (MC) it can be preceded by a demonstrative element, and the native subordination marker *tsə* immediately precedes the verb of the RC:

a-ya xalg [naql-əʃ tsə tsar-əm]
 ‘that man [(whose) story I am telling]’ (GSK: 637).

5.3.1.2 Relative noun in relative clause

If the relative noun appears in the RC, a resumptive pronoun co-referential with it appears in the MC:

[pārd ya dəraxt dər-əm tsə tu] yaw-ī kot-əv
 ‘[last year (*pārd*) **that tree** that was here], they dug **it** up’,
 i.e. ‘they dug up the tree that was here last year’ (P: 115).

5.3.1.3 Demonstrative + head noun in matrix clause

The head noun is in the MC preceded by a demonstrative element, and the generalized complementizer *ki* introduces the RC:

yem xun-i ha-ya halg-ev-en [kumd-ar ki sak-e xi degit detk]
 ‘this house belongs to **the** people [**to whom** we have given our daughter]’.

5.3.1.4 *ki* and *tsə* in relative clause

ki and *tsə* co-occur in the RC; *ki* introduces the RC, and *tsə* immediately precedes the verb. A demonstrative element precedes the relative noun in the MC, and the borrowed pre-conjunctive restrictive *-i* can sometimes appear:

a-ya nayd-i [ki a-yəm mərda-v gur tsə kart-əv]
 ‘that night [**when** they buried this corpse]’ (GSK: 664).

5.3.1.5 Interrogative/indefinite pronouns

The interrogative/indefinite pronouns *kuy* ‘who’, *kum(d)* ‘which’ + *ki*, with *ki* introducing the RC and a distal demonstrative as correlative element in the MC:

yem xun-i ha-ya halg-ev-en [kumd-ar ki sak-e xi degit detk]
 ‘this house belongs to **the** people [**to whom** we have given our daughter]’.

5.3.1.6 Headless relative clauses

The interrogative/indefinite pronoun + *ki* can also introduce a headless RC:

[kūy ki ʒə dəγd nūng-ī dišt], ʒə dəγd ya-rək
 ‘[**whoever** guesses my daughter’s name], my daughter (goes) **to him**’ (P: 113).

5.3.2 Participial relative clauses

Participial relative clauses are constructed with the present and perfect participles.

5.3.2.1 Present and perfect participles

(1) Present participle (agent)

[*sk-a vadek tuk-kiʒg*] *xalg-i ʒi mormor*
‘the man [walking along the road] is my friend’.

(2) Perfect participle

With the perfect participle, the relative noun can have various thematic roles—subject/agent, theme, or possessor:

- (a) Agent, [*tōi naxʒetk*] *deyd* ‘a girl [who hasn’t married]’ (Lorimer 1958: I, 185);
- (b) Possessor, [*nōng niʒk*] *xalg* ‘a [renowned] person’ (lit. ‘name having come out’) (Lorimer 1958: I, 185).

5.3.2.2 Substantival perfect participle

Substantival use of the perfect participle yields a headless relative interpretation:

[*no-wīn-ətʃk*] *rang qəsa mə ʒan*
‘don’t talk like (*rang*) [one who has seen nothing]’ (agent) (P: 107);
ha-yem xalg tsə [xʒu ʒ-en- etʃk]-ən munkir vit-a (object)
‘this man has repudiated [what he said]’ (theme) (Lorimer 1958: I, 185).

5.4 Temporal clauses

Temporal relations are expressed with both clausal and infinitival structures.

5.4.1 Finite temporal clauses

Finite clauses use the conjunctions *ki*, *tsə(y)*, *tsoyader ki* (< *tsə yow ʒad-er*).

5.4.1.1 Simultaneous actions/events

Simultaneous actions/events can be expressed with *ki* or *tsə(y)*:

[*də ʒbi diyər ki tu*] *ar ruz-əʒ tə ku-ʒ rəʒd-əy*
‘[when he was in his village] he would go to the mountain every day’ (GSK: 625);
[*ʒi pup mərt-i tsəy*] *yaw tsum-sol tu*
‘[when my grandfather died] how old was he?’ (GSK: 654).

5.4.1.2 Sequential actions

Sequential actions are expressed with *tsoyader ki* ‘when’ with the simple past in its complete (perfective) meaning:

[*tsoyader ki jamil wezd-a*] *zā-iʒt-ep xoʒ wots-en*
‘[when Jamil comes] the boys will be happy’.

Such sentences with *ki*, but without a temporal adverb like *tsoyader* ‘when’, may have either realis conditional or temporal interpretation.

5.4.1.3 Punctual event

A punctual event interrupting an ongoing action is introduced by a *ki*-clause following the main clause:

maš-eš šapik yit [ki tu-et wezd-a]
‘I was eating [when you came]’.

5.4.1.4 Time after which

Time after which, ‘since’, may be expressed by finite clauses with *ts-o-γad-an (ki)* ‘since’, lit. ‘from that time that’:

[wuz-əm tsoyən tr-əm drem-əm nieng] tu-et ne wezd-a
‘you haven’t come [since I have been sitting here]’ (Lorimer 1958: I, 223).

5.4.1.5 Time up to which

Time up to which, ‘until’, may be expressed with both clausal and infinitival structures, marked by phrases with *batken* ‘until, up to’. Notice the negative element in this ‘until’-clause, as is found in Urdu (and Tajik):

(1) Finite

wuz ne yaw-əm [tsoyader batken-it ki tu ne wesk]
‘I will not eat [until you come]’ (lit. ‘so long as you have not come’).

(2) Infinitival

[ti reč-an-e batkən] dr-em nezd-əm
‘I shall sit here [until you go]’ (lit. ‘until your going’) (Lorimer 1958: I, 193).

5.4.2 Infinitival temporal constructions

5.4.2.1 Inception of an event

The inception of an event immediately upon completion of a preceding event, ‘as soon as’, may be expressed by the infinitive plus *bərobar* ‘equivalent, same as’:

[potšo-bača γət-ak bərobar də xbi šar] pərst-əy
‘[as soon as the king’s son reached his own city] he asked . . .’ (GSK: 628).

5.4.2.2 Sequence of two events

A sequence of two events can be expressed with the infinitive of the prior action marked as SOURCE, with various supplementary adverbials or prepositions:

γafč peysa jama xa-k-e tsəbas-an]-ep salīm-e tat yi xun gox-t
‘[after collecting a lot of money] Salim’s father will make a house’;
[bad tsə γ ar xša-k-an] kač-an bənyod

[after bringing rocks] the foundation is built' (GSK: 628);
 [žbi rēč-n-ən tər pərūt] . . . 'before my going . . .' (Lorimer 1958: I, 260).

5.5 Reason and purpose clauses

5.5.1 Reason clauses

Reason clauses are expressed:

- (1) with the collocation *ts-a-r-en ki* 'since/because' and a correlative element in the MC:

[tsaren ki mur-eš dixt] **ha-ya dastan sak tə xun hal-en**
 '[since it is raining] **for that reason** we will stay at home';

- (2) with the fused prepositional-demonstrative elements *(a)-sk-əm*, *(a)-sk-ət*, *(a)sk-a* 'so, therefore':

žbi tat mərt-əy [sk-a wuz-əm səyirə vit-əm]
 'my father died [so I became an orphan]' (GSK: 646).

5.5.2 Purpose clauses

Purpose can be expressed in several ways:

5.5.2.1 Infinitival construction

- (1) When the subject of both matrix and subordinate action is the same, the *n*-infinitive of the subordinate verb can be marked as dative/goal:

wuz-əm [yow bar viši-n-ər] wəz-g-i xən-etk
 '“I have come [to sweep his door],” she said' (Mock 1998: 454).

- (2) Another construction employs the *k*-infinitive in a nominal sentence:

wuz safar čaw-əm [kənd yund-ak]
 'I am going on a journey [to bring a wife]' (GSK: 628).

5.5.2.2 Finite-verb clause

The strategy of using two finite verbs in sequence is also employed to indicate purpose. In this construction, the subordinate (purpose) clause contains an unspecified present/future verb carrying an optative sense and follows the main clause:

t-əm mənōč nəzd-en [mōr mə dix-t]
 'let us sit in this rock shelter [so that it does not rain (on us)]' (Lorimer 1958: I, 224).

5.6 Complement structures

Semantic complement structures may be similarly finite or nominalized:

- (1) The finite verb sequence strategy is employed to form a (semantic) complement of 'to allow':

šû tatnan iǰazat ma-r rand-ən [wuz reč-em keđmi-ər . . .]

‘(may) my parents give me permission [that I go to the academy . . .]’ (Reinhold 1998: 345).

In general, clausal *complements of speaking* follow the main clause, with or without *ki*:

maž-ər šan [(*ki*) (wuz) tsoyadər wez-im]

‘tell me [when to come]’;

wuz nə diš-əm [yaw rəš-k tɒw-ət k yā nəy]

‘I don’t know [(whether) he has left or not]’ (P: 83).

- (2) Some *verbs of cognition* like ‘to want’, ‘to know’, take infinitival complements:

sahar-ək, žɒ dur [fayzobod rečək]

‘tomorrow I want (‘my wish is’) [to go to Faizabad]’ (GSK: 627);

wuz-əš [tə xun reč-ən-ə] zok tsar-em

‘I want (‘have desire’) [to go to your house]’;

wuz-eš [angrezi *nivis-en*]-iš diš-em

‘I know [(how) to write English]’.

5.7 Conditional clauses

For purposes of this discussion, conditional clauses are defined as clauses referring to actions which are unrealized or of which the realization status is unknown to the speaker. With realis conditionals the action/event is still potential, while in the case of irrealis conditionals it has not happened/is (assumed to be) not the case.

5.7.1 *Realis*

Realis conditionals pertaining to present to future time can be constructed with present (+ -(*V*)š(*V*)), simple past, present subjunctive, and perfect subjunctive verb forms in the protasis, and the present or imperative in the apodosis. The conjunctions *tse* in clause-final position, or *ki* and *agar* after the clause-initial element, may appear with the present or past in the protasis:

5.7.1.1 Present

yašt ki řem-š wots-ən

‘if they find the time . . .’ (Reinhold 1992b: 5–6).

5.7.1.2 Simple past

mər ki dišt-a, sak to xun hał-en

‘if it rains, we will stay at home’.

5.7.1.3 Present subjunctive

wuz humi-em digār hum-ut, farx nāst

‘be it I, or another, makes no difference’ (Lorimer 1958: I, 187).

5.7.1.4 Perfect subjunctive

žārž tərəšp nə-vit-k um̄t, ta-r wəzm-əm
 ‘if the milk hasn’t become sour, I will bring it to you’ (P: 84).

5.7.2 *Irrealis*

Irrealis conditions appear with the simple past or the distant past plus the imperfective particle $-(V)\xi(V)$ in the protasis, depending on whether the reference is to future or past time, and the past or distant past with $-(V)\xi(V)$ in the apodosis. *Agar* may appear with simple past in the protasis:

5.7.2.1 Simple past

wuz-əš agar fayzobod-əš rəyd-əm, ta-rək iγun čiz-əš wozomd-əm
 ‘if I were going to Faizabad, I would bring you something’ (GSK: 640).

5.7.2.2 Distant past

wuz-əm-šə tuw-ət̄n̄w, ta-r-əm-šə yōrdam kər-ən̄w
 ‘if I had been (there), I would have helped you’ (P: 84).

6 LEXIS AND WORD FORMATION

Compounding in Wakhi shows the typical Iranian patterns, such as the dvandva *tāt-ət nān* ‘father and mother’ = ‘parents’.

There are only a few productive derivational suffixes:

- $-i(\gamma)$ forms abstract nouns from adjectives or designations for residents of specific places: *s̄br* ‘cold’, *s̄br-i(\gamma)* ‘cold(ness)’, *yamg* ‘place name’, *yamg-iγ* ‘person from Yamg’;
- $-ij$ forms terms for secondary non-blood relationships: *pətr* ‘son’, *pətr-ij* ‘stepson/child’;
- $-V/k$ forms a diminutive suffix: *yaš-ək* ‘little horse’, *šüγ-ak* ‘little sister’;
- $-b̄ngl-in$ forms adjectives from nominals indicating time, place or number: *p̄brz* ‘evening’ (noun), *p̄brz-b̄ng* ‘evening’ (adj.);
- $-in$ derives adjectives from nouns indicating the material of which something is made: *šung* ‘wood’, *šung-in* ‘wooden’. This same suffix forms a perfect participle.

7 DIALECTOLOGY AND CONTACT PHENOMENA

7.1 Wakhi dialects

According to Pakhalina (1975: 8–9) and Griunberg and Steblin-Kamensky (1976: 666), the basic dialectal divisions are:

- (1) the Wakhi of (then Soviet) Wakhan, which itself falls into three dialect areas plus the dialects of Sarhad (the villages along the Wakhan River which fall in Afghan territory);
- (2) Hunza Wakhi in Pakistan; and

- (3) Sarikol Wakhi in China. Afghan Wakhi, according to Pakhalina, is close to the central and upper dialects of WW. The Wakhi spoken in Pakistan, with two main variants in Gojal and Ishkoman, is probably close to the central and upper dialects of Wakhan. It is likely that all Wakhi dialects are mutually intelligible to a fairly high degree (Backstrom 1992: 68–9).

The varieties of Wakhan Wakhi constitute a dialect continuum, with influence from Tajiki strongest in the westernmost villages of lower Wakhan. The few differences in the consonantal and syllabic system include: interdental fricative δ in Langar against the tendency of $\delta > d$ elsewhere, and variation in consonantal onsets: *avz* (lower)/*havz* (Langar) 'lake'; *uz* (lower)/*wuz* (upper) 'I'. Noticeable differences in the vowel systems include the following (Lashkarbekov 1975: 73–4):

	<i>i ~ ê</i>	<i>ə ~ ê</i>	<i>u ~ ə</i>	<i>ɪ ~ u</i>
Upper	<i>bix</i> 'root'	<i>rəʃ</i> 'dirt'	<i>wuz</i> 'I'	<i>pɪd</i> 'foot'
Lower	<i>bêx</i> 'root'	<i>rêʃ</i> 'dirt'	<i>wəz</i> 'I'	<i>pud</i> 'foot'

Other differences include the use of the direct and oblique cases discussed above, and the lack of *n*-infinitives in the lower dialect. Lexical differences due to Tajiki influence include even basic vocabulary, such as (Lashkarbekov 1975: 78–9):

	'flower'	'Wakhan'	'bear'	'rain'
Tajiki	<i>gul</i>	<i>vaxon</i>	<i>xirs</i>	<i>boron</i>
Lower	<i>gəl</i>	<i>waxon</i>	<i>xers</i>	<i>boron</i>
Upper	<i>sprəʃ</i>	<i>wuʃ</i>	<i>n(ə)γərdum</i>	<i>wɪr</i>

7.2 Adjacent languages and dialects

WW's present linguistic neighbors are: Shughni and Kirghiz on the north, Ishkashmi and Tajik Persian on the west, Khowar and Burushaski on the south, and Kirghiz on the east. Baroghil Wakhi borders Khowar on the south. HW's neighbors are: Kirghiz on the north, Burushaski on the north and south, Shina on the south and west, and Sarikoli on the east.

7.3 Regional contact languages

In Tajikistan, Wakhi speakers know Tajik Persian and Russian; in Afghanistan, Dari Persian is the lingua franca of the region within which Wakhi is spoken. Shughni is also an important contact language for many Wakhi speakers.

In China, Uighur and Chinese are languages of education.

In Pakistan, HW now interacts intensively with Urdu and English. In the Yarkhun (Chitral) and Ishkoman valleys Khowar is widely known and spoken.

Gojal Wakhi until recently has been under the strong influence of Persian, recognizable in the pronunciation of the long vowel [ā] in Persian loans. However, this Persian influence is almost entirely restricted to lexical items and set phrases, and used by male speakers when speaking in public, or by speakers of the older generation, and has not affected the structure of the language (Reinhold 1992b).

7.3.1 *The Pamir linguistic area*

Multilingualism has apparently been widespread in the Pamir region since ancient times. There are indications of one or more substrates (see Steblin-Kamensky 1979; Edelman 1980; Tikkanen 1988, 1995, 1999; and Bashir 1988, 1997). It is also likely that there existed some unknown (pre-) Eastern Iranian language as a lingua franca of the Pamir. Evidence for this is the presence of ancient Iranian loans in the Pamir languages and Khowar (Morgenstierne 1936: 657). Persian has been known in the area at least since the eleventh century (Payne 1989: 422). There are also traces of Kirghiz in the language, reflecting close contact with Kirghiz herdsmen in Afghanistan (see Sharani 1979). The significant number of Wakhi loans referring to items of concrete vocabulary in Khowar argues for a long period of close interaction between Wakhi and Khowar speakers (Morgenstierne 1936, 1938: 441–2; Bashir 2001). Morgenstierne also points to a significant number of Sarikoli borrowings from Wakhi, and a smaller number from Sarikoli into Wakhi (1938: 438). Interestingly, however, Burushaski or Shina loans are lacking in Wakhi (Morgenstierne 1938: 440–1).

Wakhi has several characteristics which argue against the postulation of a unified proto-Pamir group as distinct from the rest of East Iranian (see Chapter 14a Pamir languages). A number of scholars have discussed the areal features of various sprachbund configurations that include the Pamir languages.

The Pamir languages share the following features (here cited from Payne 1989: 422, after Dodykhudoev 1975: 21 n.74):

- (1) vowel alternations to show morphological distinctions—gender, tense, transitivity, and also number to a limited extent;
- (2) three verb stems: present, past, and perfect: Wakhi *al-* (present), *al-d* (past), *al-ət-k-* (perfect) ‘remain’;
- (3) indication of person and number in tenses based on the present stem with personal endings, but the use of mobile agreement particles (pronominal clitics typically enclitic to the first major sentence constituent) in tenses based on the past and perfect stems;
- (4) three-valued demonstrative systems;
- (5) similarity of word order patterns, except for clearly borrowed patterns like the *ezāfe* construction: native ADJ-N pattern, *baf kənd* ‘good wife’ vs. *kənd-i-baf* ‘good wife’ (*ezāfe*).

7.3.2 *Dardic*

The Pamir languages share much with the Dardic area. Vowel alternation to indicate gender, transitivity and number is also characteristic of the Dardic languages. For example, in Dameli, Shumashti, Kalam and Indus Kohistani, Torwali, Kanyawali and Wotapuri vowel fronting or raising marks feminines. Three-valued deictic systems are also characteristic of the Dardic languages, in which the distinctions visible–not visible, or known–not known are grammaticized. Languages with three-term systems are Pashai, Shumashti, Khowar, Kalasha, Kalam Kohistani, Torwali, Indus Kohistani, Shina, and Palula (Bashir 2003). For discussion of other areal configurations, see Edelman (1980), Tikkanen (1988, 1995, 1999), and Bashir (1988 and 1997).

8 SAMPLE TEXT

The following text was written by Ghulamuddin Beg of Gulmit village in Gojal, using the IPA-based script discussed above, and given to Beate Reinhold in 1991 for publication. I thank her for allowing me to include it in this article as an example of recent Wakhi prose composition in Gojal. Reinhold's transcription is maintained, with the exception that <ts> and <dz> have been substituted for Reinhold's <c> and <č>. Morphemic analysis is the author's; the gloss follows Reinhold's translation. In addition to conventional abbreviations, note: GL = goal; SCR = source; IPV = imperfective; PC = person clitic.

wungas tuy.

sparrow wedding.

'The Wedding of the Sparrows'.

wungas mani vit-əy dzaq-ik prinda,

sparrow meaning become-PT small-DIM bird

'Sparrow means a tiny bird.'

woz tuy yow-ər jumat wūzūm-n.

and wedding he-GL wife bring-INF

'and wedding (means) to bring him a wife.'

yəm rəwāj bə də gojal alaqa qadim.

this custom also in Gojal region ancient

'This is also an old custom in Gojal.'

dzirg-š xan-ən ki yūr-k-əš ki peš-n-ər giz-d

like.this-IPV say-PR3p that barley-IPV when ripen-INF-GL come.up-PR3s

'They say that when the barley is about to ripen.'

wungas-išt-əš yūr-k yafč yaw-ən,

sparrow-DIRP-IPV barley much eat-PR3p

'the sparrows (can) eat much barley.'

sk-a-š yəm xudāyi gox-ən

from-that-IPV this sacrifice do-PR3p

'therefore they make a sacrifice (of barley)

woz-əš bar go-n xudāyi,

and-IPV outside do-PR3p sacrifice

'and celebrate the sacrifice in an open area.'

yandi duwā tsar-ən ki

then prayer do-PR3p that

'Then they pray that'

fasl-ər ts-əm wungas-v-ən niyat wos-t

crop-GL from-these sparrow-OBL₁P-SRC blessing become-PR3s

'the harvest be blessed by the sparrows.'

Qadim zamân tə-r bar čiz wūzūm-n fədek ne tu.

ancient time from-down outside anything bring-INF road not was

'In the old times, there were no roads to transport anything from outside (of the valley).'

xalg-v-ər ki fasl-əv-əš nə vit-əy,
 person-OBL₁p-GL if crop-PC3p-IPV not become-PT
 ‘If the people’s harvests did not develop well,

zirič-n-əv-əš mərt-əy. zəmin kam tu.
 weakness-SRC-PC3p-IPV die-PT land little was
 ‘they would die of weakness. Land was scarce.’

sk-a-v-əš wungas-əy yit-ər bə wayt-əy.
 from-that-PC3p-IPV sparrow-OBL₂ eat-INF-GL also lament-PT
 ‘Therefore, they (even) lamented about the sparrows eating (some barley).’

ha-ya waxt yafč muškil waxt tu.
 EMPH-that time very hard time was
 ‘That time was a very difficult time.’

yəm rasm-əš də angrizi mûy jun-ə âxiri hafta
 this custom-IPV in English month June-OBL₂ last week
 ‘This custom in the last week of the English month of June,’

tsə činir-ən das râr jald gox-en.
 from harvest-SRC ten day early do-PR3p
 ‘ten days before the harvest thanksgiving they do (it).’

t-a wungas nung yə yi xudâyi.
 to-that sparrow name this a sacrifice
 ‘This is a sacrificial rite in the name of that sparrow.’

xalg-išt-əš xudâyi-ər duwâ tsar-ən
 people-DIR3p-IPV sacrifice-GL prayer do-PR3p
 ‘People pray during the rite’

ki fasl ts-əm wungas-v-ən bərbâd mə wos-t.
 that crop from-this sparrow-OBL₁p-SRC destroyed not become-PR3s
 ‘that the harvest not be destroyed by these sparrows’.

NOTES

In this summary article, given the uncertainty about the analysis of the Wakhi vowel systems, dialectal variation, and lack of comparable data, the original transcriptions of the authors cited are retained, with the exception that for the voiceless dental affricate <ts> is consistently used. Lorimer’s *u* is rendered as *u*, to avoid confusion with the use of underlining to indicate stress. Examples from my own field materials are in broad phonetic transcription.

ACKNOWLEDGEMENTS

I am grateful to Zafar Iqbal, Manager, NATCO, Gilgit and Ahmad Sakhi Jami, Training Officer, AKRSP, both of Passu, for consulting with me during my 1987 and 1989 field work in Gilgit. This fieldwork was partially supported by a grant from the Smithsonian Institution for the Multi-Disciplinary Project on Pakistani Folk Culture, directed by William Hanaway and Wilma Heston.

BIBLIOGRAPHY AND REFERENCES

- Backstrom, P. C. (1992) 'Wakhi', in Backstrom, P. C. and Radloff, C. F. (eds), *Languages of the Northern Areas*, Islamabad, Pakistan: National Institute of Pakistan Studies and Summer Institute of Linguistics, pp. 57–74.
- Bashir, E. (1986) 'Beyond split-ergativity: Subject marking in Wakhi', in *Papers from the Regional Meetings, Chicago Linguistic Society* 22.1, pp. 14–35.
- (1988) *Topics in Kalasha Syntax: An Areal and Typological Perspective*. Ph.D. dissertation, University of Michigan, Ann Arbor.
- (1997) 'Khowar/Burushaski commonalities', paper presented at the Third International Himalayan Languages Symposium, held July 18–20 at the University of California at Santa Barbara.
- (2001) 'Khowar-Wakhi contact relationships', in Lönne, D. (ed.) *Tohfa-e-Dil: Festschrift Helmut Nespital*, Reinbeck bei Hamburg: Wetzler, pp. 3–17.
- (2003) 'Dardic,' in Cardona, G. and Jain, D. (eds), *The Indo Aryan Languages*, London: Routledge, pp. 818–94.
- Buddruss, G. (1974) 'Neuiranische Wortstudien', *MSS* 32: 9–40.
- (1986) 'Wakhi-Sprichwörter aus Hunza', in Schmitt, R. and Skjærvø, P. O. (eds) *Studia Grammatica Iranica. Festschrift für Helmut Humbach*, München: R. Kitzinger, pp. 27–44.
- (1998) 'Eine einheimische Sammlung von Wakhi-Sprichwörtern aus Hunza: Text, Übersetzung, Glossar', in Bogoliubov, M. N. (ed.) *Tsentral'naiia Azia. Vostochnyi Gindukush* (Griunberg memorial volume), Sankt Peterburg: Peterburgskoi vostokovedenie, pp. 30–45.
- Dodykhudoev, R. Kh. (1975) *Pamirskaia mikrotoponimiia: Issledovanie i materialy* (Pamir microtoponymy: Introduction and materials), Gertsenberg, L. G. (ed.), Dushanbe: Izdatel'stvo Irfon.
- Edelman, D. I. (1980) 'K substratnomu naslediiu tsentral' noaziatskogo iazykovogo soiuza (On the substratal heritage of the Central-Asian Sprachbund)', *Via* 5: 21–32.
- Felmy, S. (1996) *The Voice of the Nightingale: A Personal Account of the Wakhi Culture in Hunza*, Karachi and New York: Oxford University Press.
- Gao, E. (1985) *Tajikeyu Jianzhi* (Concise Grammar of [Sariquli and Wakhi] Tajik). Zhongguo Shaoshu Minzu Yuyan Jianzhi Congshu (Series of Concise Grammars of Chinese Minority Languages), Beijing: Minzu Chubanshe.
- Grierson, G. A. (1921) 'The Ghalchah Languages', in *LSI* X, pp. 455–548.
- GSK = Griunberg, A. L. and Steblin-Kamensky, I. M. (1976) *Vakhanskii iazyk. Teksty, slovar grammaticheskii ocherk* (The Wakhi language. Texts, glossary, grammatical outline). Moskva: Nauka; French translation (1988) *La langue wakhi*, Paris: Editions de la Maison des Sciences de l'Homme.
- Haqiqat, A. (n.d.) *Wakhi Language [xək wōr zik]*, Book 1, Gilgit: Wakhi Culture Association.
- Heston, W. (1986) 'Pashto ambipositions and historical antecedent', in Bashir, E., Deshpande, M. M. and Hook, P. E. (eds) *Select Papers from SALA-7: South Asian Languages Analysis Roundtable Conference, held in Ann Arbor, Michigan, May 17–19, 1985*, Bloomington, Indiana: Indiana University Linguistics Club, pp. 163–81.
- Kieffer, Ch. M. (1978) 'Einführung in die Wakhi-Sprache & Glossar', in de Grancy, R. S. and Kosta, R. (eds) *Grosser Pamir. Österreichisches Forschungsunternehmen 1975 in den Wakhan-Pamir/Afghanistan*. Graz: Akademische Druck-u. Verlagsanstalt, pp. 345–74.
- Klimchitskii, S. I. (1936) 'Vakhanskii teksty' (Wakhi texts), *Trudy tadjikistanskoi bazy AN SSSR. Lingvistika*, tom II, Leningrad: AN SSSR, pp. 75–124.

- Kreutzmann, H. (1996) *Ethnizität im Entwicklungsprozeß – Die Wakhi in Hochasien*. Berlin: Dietrich Reimer Verlag.
- (1998) 'Sprachdifferenzierung im Ost-Hindukusch and Karakorum', in Bogoliubov, M. N. (ed.) *Tsentral'naia Azia. Vostochnyi Gindukush* (Griunberg memorial volume), Sankt Peterburg: Peterburgskoi vostokovedenie, pp. 83–108.
- Lashkarbekov, B. B. (1975) 'O govornykh razlichiiakh vakhanskogo iazyka' (On dialect differences in Wakhi), in *Lingvisticheskie issledovaniia 1975. Issledovaniia po grammatike iazykov narodov SSSR*, Moskva: Nauka, pp. 72–83.
- (1982) 'K istorii vakhanskoi kosvennoi konstruktсии' (On the history of the Wakhi oblique construction), in *AcIr 22 (Monumentum Georg Morgenstierne)*, vol. 2, Leiden: Brill, pp. 1–4.
- (1985) 'Stanovlenie sistemy vakhanskogo glagola na trekh stadiiakh iazykovogo razvitiia' (The emergence of the Wakhi verb system in three stages of language development), *Via* 1: 97–108.
- Lorimer, D. L. R. (1958) *The Wakhi Language*, 2 vols, London: School of Oriental and African Studies, University of London.
- Mock, J. (1998) *The Discursive Construction of Reality in the Wakhi Community of Northern Pakistan*, PhD dissertation, University of California at Berkeley.
- Morgenstierne, G. (1936) 'Iranian elements in Khowar', *BSOS* 8: 657–671.
- (1938) *Iranian Pamir Languages* = Morgenstierne, *IIFL* II.
- (1945) 'Indo-European *k*' in Kafiri', *NTS* 13: 225–238.
- Pakhalina, T. N. (1966) 'Vakhanskii iazyk' (The Wakhi language), in *Iazyki narodov SSSR, I. Indoevropskie iazyki*, Moskva: Nauka, pp. 398–418.
- P = Pakhalina, T. N. (1975) *Vakhanskii iazyk* (The Wakhi language), Moskva: Nauka.
- (1983) *Issledovanie po sravitel'no-istoricheskoi fonetike pamirskikh iazykov* (Introduction to the comparative-historical phonetics of the Pamir languages), Moskva: Nauka.
- (1987) 'Vakhanskii iazyk' (The Wakhi language), in *Osnovy* IV, pp. 408–473.
- (1989) *Sravitel'no-istoricheskaiia morfologiiia Pamirskikh iazykov* (Comparative-historical morphology of the Pamir languages), Moskva: Nauka.
- Pakhalina, T. N. and Lashkarbekov, B. B. (2000) 'Vakhanskii iazyk' (The Wakhi language), in *Iazyki mira. Iranskii iazyki* III, Moskva: Indrik, pp. 174–86.
- Payne J. R. (1979) 'Transitivity and intransitivity in the Iranian languages of the U.S.S.R.', in *Papers from the Conference on Non-Slavic Languages of the U.S.S.R.*, Chicago: Chicago Linguistic Society, pp. 436–447.
- (1980) 'The Decay of Ergativity in Pamir Languages', *Lingua* 51:147–186.
- (1981) 'Iranian Languages', in Comrie, B. (ed.) *Languages of the Soviet Union*, Cambridge: Cambridge University Press, pp. 58–179.
- (1982) 'Relativization in the Iranian languages of the U.S.S.R.', *Folia Slavica* 5.1–3: 344–363.
- (1989) 'Pamir Languages', in *CLI*, pp. 417–444.
- Reinhold, B. (1992a) 'Seven Wakhi poems', *JRAS* 2.2: 203–211.
- (1992a) 'Seven Wakhi poems', *JRAS* 2.2: 203–211.
- (1992b) *Field Report, No. 47*, Culture Area Karakorum Project (CAK), Project: The Wakhi Language of the Northern Areas.
- (1992c) *The Incorporation of Urdu and Persian into the Wakhi of the Karakorams*, ms.
- (1998) 'Women in the rural employment scene in Northern Pakistan: Experiences and expectation', in Bogoliubov, M. N. (ed.) *Tsentral'naia Azia. Vostochnyi Gindukush* (Griunberg memorial volume), Sankt Peterburg: Peterburgskoi vostokovedenie, pp. 341–52.
- (2006) *Neue Entwicklungen in der Wakhi-Sprache von Gojal (Nordpakistan)*. Iranica 10. Wiesbaden: Harrassowitz.

- Shahrani, M. N. M. (1979) *The Kirghiz and Wakhi of Afghanistan: Adaptation to Closed Frontiers*, Seattle: University of Washington Press.
- Shaw, R. B. (1876) 'On the Ghalchah languages (Wakhi and Sarikoli)', *JASB* 45/1: 139–278.
- Sims-Williams, N. (1996) 'Eastern Iranian languages', in *EnIr* 7, pp. 649–652.
- Sokolova, V. S. (1981) 'Kistorii vokalizma vakhanskogo iazyka' (On the history of the vocalism of the Wakhi language), in Bogoliubov, M. N., Griunberg, A. L., Efimov, V. A. and Rastorgueva, V. S. (eds) *Iranskoe iazykoznanie: Ezhegodnik 1980*, Moskva: Nauka, pp. 37–47.
- Steblin-Kamenskii, I. M. (1979) "'Koleni" i "lokti" pamirskogo substrata" ("Knees" and "elbows" of the Pamir substrate), in *Peredneaziatskii sbornik*, III. *Istoriia i filologiiia stran drevnego vostoka*, Moskva: Nauka, pp. 212–214.
- (1981a) 'Istoricheskaia klassifikatsiia vakhanskikh glagolov' (The historical classification of the Wakhi verbs), in Bogoliubov, M. N., Griunberg, A. L., Efimov, V. A. and Rastorgueva, V. S. (eds) *Iranskoe iazykoznanie: Ezhegodnik 1980*, Moskva: Nauka, pp. 57–66.
- (1981b) 'Novye materialy po vakhanskoi leksike' (New materials on the Wakhi lexicon), *ibidem*, pp. 178–185.
- (1982) 'Venok iz Vakhana' (A garland from Wakhan), in *AcIr* 22 (*Monumentum Georg Morgenstierne*), II, Leiden: Brill, pp. 227–250.
- (1999) *Etimologicheskii slovar' vakhanskogo iazyka*, Sankt Peterburg: Sankt Peterburgskii gosudarstvennyi universitet.
- Stein, Sir A. (1921) *Serindia I*, Oxford: Clarendon Press.
- Stilo, D. (1987) 'Ambipositions as an areal response: The case-study of the Iranian zone', in Bashir, E., Deshpande, M. M. and Hook, P. E. (eds) *Select Papers from SALA-7: South Asian Languages Analysis Roundtable Conference, held in Ann Arbor, Michigan May 17–19, 1985*, Bloomington, Indiana: Indiana University Linguistics Club, pp. 308–336.
- Tikkanen, B. (1988) 'On Burushaski and other ancient substrata in Northwestern South Asia', *SO* 64: 303–325.
- (1995) 'Some areal linguistic isoglosses in the transitional zone between Central and South Asia', paper presented at the Third International Hindukush Cultural Conference, September 26–30, 1995, Chitral.
- (1999) 'Archaeological-linguistic correlations in the formation of retroflex typologies and correlating areal features in South Asia', in Blench, R. and Spriggs, M. (eds) *Archaeology and Language IV. Language Change and Cultural Transformation*, London: Routledge, pp. 138–48.

INDEX

- ablative case: Avestan 112–13, 119; Old Persian 112–13, 119
- ablaut: Avestan 70, 72, 75; Pashto 729;
see also vowels
- accusative case: Avestan 105–8, 118–19; Old Persian 105–8, 118–19
- Achaemenid Empire 7, 279
- adjectives: Avestan 79–80; Balochi 655–6; Khotanese 386–7, 399–400; Khwarezmian 344, 358, 371; Kurdish 599, 626–7; Middle Persian 206, 242–6, 259, 262–3; Pamir languages 781; Parachi 697; Pashto 731–2; Persian and Tajik 432–4, 475; Shughni 794, 801–2; Sogdian 290; Tumshuqese 387; Wakhi 829; Zazaki 549–51
- adpositions: Avestan 95; Balochi 657; Khotanese 389; Khwarezmian 344, 346–7, 350; Kurdish 600–2, 605–6, 617–18; Middle Persian 210–11, 224–6; Old Iranian 83; Old Persian 95; Parachi 699; Pashto 735–6; Persian and Tajik 440–4, 513–14; Shughni 796–7; Sogdian 293–4; Wakhi 831–2; Zazaki 552–3
- adverbial clauses: Avestan 160–4; Balochi 681–4; Khwarezmian 366–70; Kurdish 623–6; Middle Persian 257–8; Old Persian 160–4; Parachi 714–15; Pashto 762–6; Persian and Tajik 516–26; Shughni 813–16; Sogdian 319–20; Zazaki 576–81
- adverbials: Persian and Tajik 480–1; Shughni 804
- adverbs: Avestan 80; Balochi 657; Khotanese 389–90, 406–7; Khwarezmian 348; Kurdish 602–3; Middle Persian 206–7; Old Persian 80; Parachi 699; Persian and Tajik 444–5; Shughni 797; Sogdian 294; Wakhi 832; Zazaki 553–4
- affixation: Avestan 167–8; Khotanese 408–9; Khwarezmian 360–1, 370–1; Kurdish 604–5, 606–8, 616; Middle Persian 260–3; Old Persian 167–8; Parachi 716; Persian and Tajik 526–8; Shughni 816–17; Sogdian 322; Zazaki 558, 582
- Afghanistan 766, 787; varieties of Persian 417; Wakhi in 826
- agency, Persian and Tajik 499–500
- agreement: Avestan 102–3; Middle Persian 227; Old Persian 102–3; Pashto 756; Shughni 805; Sogdian 312–13
- Ahadi, S. 496
- Akkadian, timeline 9
- aktionsart constructions: Balochi 675–6; Parachi 706; Persian and Tajik 493–5; Shughni 807–8
- alienability, Shughni 803
- alphabet: Avestan 45–6, 51–2; Balochi 638–40; Khotanese 379–81; Khwarezmian 338; Kurdish 589–90; Middle Persian 197–200; Old Persian 47, 52–3; Pahlavi 202–4; Pashto 722–3; Persian 421–3; Shughni 788; Sogdian 281–4; Tajik 423; Tumshuqese 381; Wakhi 826
- anaphora, Persian and Tajik 503–4
- anaptyxis: Avestan 64–5, 67–8; Old Persian 65
- ‘Ancient Letters’ 325–6
- Andronovo culture 6
- animacy hierarchy, Sogdian 306–7
- Arabic 17; loanwords in Shughni 820; loanwords in Persian and Tajik 529
- Arabization of Persian 419
- Aramaic: script 197–9; timeline 9
- ‘arameograms’, in Parthian and Pahlavi 203
- articles: Khwarezmian 341, 343, 358–9; Shughni 794
- aspect 26; Balochi 661–2; Khwarezmian 352–4; Kurdish 608–11, 619; Parachi 705–6; Pashto 738, 739–42, 743–4; Persian and Tajik 488, 518; Wakhi 838–40, 846–7; Zazaki 559–60
- aspiration, Balochi 644
- assimilation: Avestan 62–3; Indo-Iranian 69–70
- Avesta*, the 47–8
- Avestan 44–6; adjectives 79–80; adverbs 80; cases 103–16; consonant clusters 19; consonants 56–7, 62–6; coordination 147–53; lexis 167–9; metre 68–9; moods 134–8; nouns 71–9; participles 141–5; phonology 18, 53–7; pronouns 80–3; religious terminology 264; script 51–2; stems 75–9; stylistics 169–72, 173–4; subordination 153–66; tenses 126–34; verb stems 87–90; voice 139–41; vowels 54–5, 59–62; word order 94–6; *see also* Old Iranian

- Axenov, S. 652
 Azami, Ch. A. and Windfuhr, G. 337
- Back 202
 Bactria-Margiana Archaeological Complex
see BMAC
 Bactrian, influence on Sogdian 327
 Bailey, H. W. 411
 Bajuwi 819; *see also* Shughni
 Balochi: adjectives 655–6; adpositions 657;
 adverbial clauses 681–4; adverbs 657;
 aktionsart constructions 675–6; alphabet
 638–40; aspect 661–2; case 651–3; clause
 structure 668–9; complement clauses 679;
 conjunctions 678; consonants 19, 642–6,
 648, 650; copula 660–1; definiteness 667;
 dialects 636–8; early location 30; ergativity
 663–4; geography 634–5, 638; literature
 635–6; loanwords 645; modals 677–8;
 mood 662, 671–3, 674–5; nouns 684;
 numbers 658; object marking 669–70;
 orthography 638; phrase structure 667;
 possessives 666–7; promotion 635;
 pronouns 653–5; relative clauses 680;
 stems 659–61; stress 648–9; subject
 marking 669; subordinate clauses 678–9;
 syllable structure 647–8; TAM-forms
 665–6, 671–5; tense 662, 664–5, 673–4;
 transitivity 663–4; verbs 670–8, 684–5;
 voice 662–3; vowels 641–2, 648, 649–50;
 word order 668
 Bartholomae's law, Avestan 62–3
 Barwoz 819; *see also* Shughni
 Benveniste, E. 693
 BMAC 5; culture 6; timeline 9
 borrowed words *see* loanwords
 Bossong, G. 33
 Brāhmī script 380–1
 Brahui 17
 Brunner 218
 Buddhism and Khotanese 378
 Buddruss, G. 652
 Bukhara 420–1
- case 23; Avestan 71, 103–16 [endings 72–5];
 Balochi 651–3; Khotanese 403–4;
 Khwarezmian 342, 343, 346–7, 362; Middle
 Persian 205; Old Persian 103–16; Pamir
 languages 781; Parachi 696–7; Pashto
 730–1; Persian and Tajik 440–4; Sogdian
 305–6; Wakhi 829, 842–6; Zazaki 567–8
 case marking: differential 33–4; ergative
 construction 31–3
 Caspian, consonant clusters 19
 Caucasus, Persian in 417
 causal clauses: Balochi 682; Khwarezmian
 368–9; Kurdish 625; Parachi 715; Pashto
 762–3; Persian and Tajik 521; Shughni 815;
 Zazaki 580–1
 causal-explanatory clauses, Persian and Tajik
 521
 causatives: Balochi 659; Persian and Tajik 448
 causee marking, Wakhi 845–6
 Central dialects, consonant clusters 19
 chiasmus, Avestan 171
 China, Sogdians in 280
 Chinese, influence on Sogdian 328–9
 Christian Sogdian 326; *see also* Syriac
 Cizre–Botan 629–30; *see also* Kurdish
 classifiers, Persian and Tajik 478–9
 clause complementation 35
 clause structure, Balochi 668–9
 clitics: idiomatic 515–16; Persian and Tajik
 471–2
 collocation, Sogdian 323–4
 colloquialisms, Khwarezmian 372
 Common Balochi (ComBal.) *see* Balochi
 comparatives: Avestan and Old Persian 124–5;
 Persian and Tajik 523; Sogdian 290; Zazaki
 581
 complementizers: Balochi 679; Persian and
 Tajik 488, 511–16; Sogdian 314–15; Wakhi
 853–4
 complex sentences, Sogdian 317–21
 compounds: Avestan and Old Persian 168–9;
 Khotanese 409; Khwarezmian 350, 371;
 Kurdish 606, 627; Middle Persian 259, 263;
 Persian and Tajik 476, 528–9; Shughni 798,
 817; Sogdian 323; Wakhi 855; Zazaki
 556–7, 570, 582
 concessive clauses: Balochi 684; Persian and
 Tajik 521; Shughni 814; Zazaki 579–80
 conditional clauses: Avestan 164–6; Balochi
 683–4; Khwarezmian 367–8; Kurdish 623–5;
 Old Persian 159–60, 164–6; Parachi 714;
 Pashto 764–5; Persian and Tajik 519–20;
 Shughni 813–14; Wakhi 854–5; Zazaki
 578–9
 conditional exception clauses, Persian and
 Tajik 523
 conjugation, Khotanese 394–5
 conjunct verbs, Tajik 495–6
 conjunctions: Avestan and Old Persian 149–50;
 Balochi 678; Khotanese 399, 405;
 Khwarezmian 364; Kurdish 612; Middle
 Persian 35, 250–2, 255; Parachi 708, 715;
 Parthian 251–2; Pashto 758–9; Persian and
 Tajik 468–70, 501–2, 534–5; Shughni 800–1;
 Sogdian 317–18; Tumshuqese 405; Wakhi
 840; Zazaki 563
 consecutive clauses: Balochi 683; Pashto 766;
 Persian and Tajik 523; Shughni 815; Zazaki
 581
 consonants: Avestan 56–7, 62–6, 73, 79;
 Balochi 642–6, 648, 650; Indo-Iranian vs.
 Indo-European 48–9; Khotanese 382–3;
 Khwarezmian 340; Kurdish 591–3, 596–8,
 629; Middle Persian 200–2; Old Persian

- 58–9, 65–6; Pamir languages 777–80; Parachi 694–5; Pashto 723–5; Persian and Tajik 426–9; Shughni 790–1; Sogdian 285–6; Tumshuqese 383; Wakhi 827–8; West Iranian 18–20; Zazaki 547
- contact languages 8–9, 15–18
- coordinate clauses: Kurdish 620; Parachi 713; Wakhi 849; Zazaki 573–4
- coordination: Avestan 147–53; Middle Persian 221, 250–5; Old Persian 147–53; Shughni 810–11
- copula: Avestan 92, 133–4, 143–4; Balochi 660–1; Khotanese 393–4; Khwarezmian 355; Kurdish 606–7; Middle Persian 216–17, 226–7; Old Persian 133–4; Parachi 702; Parthian 216–17, 227; Pashto 738–9; Persian and Tajik 450–1, 484; Shughni 799; Sogdian 300; Wakhi 835–6; Zazaki 557–8
- counterfactual, Persian and Tajik 458–9
- Cw-clusters 20
- Dardic and Pamir languages 857
- Dari–Persian 767; and Wakhi 856
- Darius 46–7
- dating: Middle Persian 259–60; Parthian 260
- dative case: Avestan 110–11, 119; Old Persian 110–11, 119
- declensions, Khotanese 384–7
- defective verbs, Persian and Tajik 459–60
- definite articles *see* articles
- definiteness: Balochi 667; Parachi 710; Persian and Tajik 483
- demonstrative pronouns: Avestan 119–22; Balochi 655; Khotanese 387; Khwarezmian 344–5; Kurdish 599–600; Middle Persian 209, 223; Old Persian 119–22; Parachi 697–8; Pashto 734; Shughni 794–5; Tumshuqese 388; Wakhi 829–31
- demonstratives: Avestan 82–3; Khwarezmian 358–9; Sogdian 291–3; Zazaki 551
- derived nouns, subordination as, Middle Persian 259
- dialects: Balochi 636–8; Kurdish 628–30; Parachi 717; Pashto 766–7; Shughni 819–20; Sogdian 326–7; Tajik 776; Wakhi 855–6; Zazaki 582–3
- diathesis, Khotanese 398
- differential object marking (DOM) 33–4
- diphthongs: Avestan 55; Sogdian 285
- direct speech: Khotanese 406; Persian and Tajik 513; Sogdian 320
- disjunction: Avestan 150; Middle Persian 253; Old Persian 150
- dissimilation *see* assimilation
- Dodykhudoev, R. Kh. 857
- Dravidian 17–18
- ‘dual’ compounds, Avestan 169
- duality, agreement, Avestan and Old Persian 102–3
- Early New Persian (ENP): loanwords in Khwarezmian 372; verb system 25
- Early Old Iranian, verb system 23–4
- East India Company 418
- Eastern Balochi 637; consonants 646; stress 649; *see also* Balochi
- Elamite, timeline 9
- elative formation, Sogdian 290
- Elfenbein, J. 648
- ellipsis in poetry, Old Avestan 171–2
- elliptic superlatives, Persian 434
- enclitics: Balochi 654–5; Parachi 711; Pashto 756–7; Sogdian 315
- endings, Middle Persian 217–18
- epenthesis: Avestan 65; Old Persian 66
- ergative construction 31–3; Middle Persian 227–9 [word order 247–9]
- ergativity, Balochi 663–4
- evidential forms: Persian 461; Tajik 462–4
- exclamations, Persian and Tajik 482
- existential verb *see* copula
- ezāfe* construction 27–8; Persian and Tajik 470–7; Zazaki 563–7
- Fattah, I. K. 628
- figura etymologica: Avestan 173–4; Middle Persian 246; Old Persian 173–4
- final clauses: Balochi 683; Khwarezmian 369–70; Parachi 715; Shughni 815; Zazaki 580–1
- formality, Persian and Tajik 531
- fricatives, Balochi 643–4
- fused prepositions 293–4
- future participles, Avestan 142
- future tense, Avestan 130
- Geldner, K. F. 45–6
- germination: Avestan 63; Balochi 644; Khwarezmian 341; *see also* consonants
- gender 22; Avestan 71–2, 102; Khotanese 384; Khwarezmian 342–3; Middle Persian 204; Pamir languages 780–1; Parachi 696; Pashto 726–8; Shughni 792–3; Sogdian 288
- genitive case: Avestan 108–10, 119; Old Persian 108–10, 119
- gerund: Khotanese 393, 398; Sogdian 297
- glides, Avestan and Old Persian 66
- Gojal Wakhi 826; *see also* Wakhi
- Gorani: consonant clusters 19; early location 30
- Griunberg, A. L. and Steblin-Kamensky, I. M. 827, 855
- group inflexion, Sogdian 313
- Haig, G. and Paul, L. 628
- head incorporation, Persian and Tajik 508
- hendiadys, Sogdian 323
- Henning, W. B. 343, 360
- hiatus breakers, Persian and Tajik 429

- Hoffmann 46
 holo-kinetic stems, Avestan 76
 homoglossia, Persian and Tajik 535
 homonymy, Persian and Tajik 526
 honorific expressions: Khotanese 410; Sogdian 324–5
 Humbach 125, 204
 Hunza Wakhi 835–6; case 843; object marking 845; plural marking 23; *see also* Wakhi
 Hurrian-Urartian, timeline 9
 hysterokinetic stems, Avestan 76
- i*-epenthesis, Avestan 65
 immigrants, Arab 419
 imperative mood: Avestan 91, 138; Middle Persian 238–9; Old Persian 138; Parthian 238–9
 imperatives: Avestan 91, 99; Old Persian 99
 indefinite articles *see* articles
 indefinite pronouns: Avestan 123–4; Khotanese 388; Middle Persian 209–10; Old Persian 123–4; Pashto 735; Sogdian 293
 indefinite relative clauses, Avestan 159
 indefiniteness, Persian and Tajik 432, 439, 506–7
 Indian: loanwords in Khotanese 409–10; influence on Sogdian 328
 indicative verbal systems 25
 Indo-Aryan 17
 Indo-European, phonology 48–9
 Indo-Iranian (IIr.): consonant clusters 19; phonology 48–50
 infinitival clauses, Persian and Tajik 524–5
 infinitives: Avestan 94, 145–6; Khotanese 392; Middle Persian 243–4; Old Persian 94, 145–6; Parthian 244; Sogdian 296–7, 320–1; Tumulshuqese 392–3
 inflection, Indo-Iranian 69
 initial clusters 21
 injunctive, Avestan 91
 instrumental case: Avestan 113–15; Old Persian 113–15
 interjections: Avestan 153; Sogdian 294
 interrogative pronouns: Avestan 123; Khotanese 388; Middle Persian 209–10; Parachi 698; Pashto 735; Shughni 795
 interrogatives: Avestan 82, 98; Khotanese 402–3; Khwarezmian 344–5, 359; Kurdish 600, 615; Middle Persian 249, 257; Old Persian 82; Parthian 249, 257; Pashto 759; Persian and Tajik 438–40, 481–2; Sogdian 293, 316–17; Zazaki 572
 Iran, varieties of Persian 417
 Iranian: loanwords in Khotanese 409; timeline 9
 irrealis *see* realis
 Ishkashimi *see* Pamir languages
 Islam and Persian 416
- isoglosses, West Iranian 29–31
izāfa construction, Kurdish 612–13
- Johanson, L. 461
 Judeo-Persian 418
- Khānī, A. 587
 Khattak clan 721
 Khotanese: adjectives 386–7, 399–400; adpositions 389; adverbs 389–90, 406–7; affixation 408–9; case 403–4; compounds 409; conjugation 394–5; conjunctions 399, 405; consonants 382–3; copula 393–4; declensions 384–7; diathesis 398; direct speech 406; gender 384; gerunds 393, 398; history 377–8; honorifics 410; infinitives 392; interrogatives 402–3; language death 411; loanwords 409–10; modals 404; moods, 396–8; negation 395; nominalised clauses 407–8; nouns 384–6, 399–400; numbers 390; particles 391–2; person marking 393 4; phonology 18; possession 404; pronouns 387–8, 401–2; relative clauses 406; script 379–81; spelling 381; stems 384–6, 391; stress 383; tense 395–6, 404; vowels 382; word order 401–3
 Khwarezmian: adjectives 344, 358; adpositions 344, 346–7, 350; adverbial clauses 366–70; adverbs 348; affixation 360–1, 370–1; articles 341, 343, 358–9; aspect 352–4; case 342, 343, 346–7, 362; colloquialisms 372; compounds 350, 371; conjunctions 364; consonants 340; copula 355; demonstratives 358–9; DOM 34; ergative construction 32; gender 342–3; history 336–8; interrogatives 359; kinship terms 372; loanwords 372–3; modals 363–4; mood 352–4, 356; nasalization 340; negation 356–7, 360; nominalization 370; nouns 343–4; numbers 342–3, 348; participles 351–2; particles 355–7; person marking 352; pronouns 344–5, 359; relative clauses 365; script 338; stems 349–52; stress 341; subject and object clauses 365–6; subordination 364; tense 351–4; verbs 25–6, 354–5, 362–4; vowels 339–40, 341–2; word order 358–61
- Kieffer, C. 693
 kinship terms: Khwarezmian 372; Pashto 729; Persian and Tajik 532; Shughni 818; Zazaki 583
- Korn, A. 19, 20
 Kurdish 20; adjectives 599, 626–7; adpositions 600–2, 605–6, 617–18; adverbial clauses 623–6; adverbs 602–3; affixation 604–5, 606–8, 616; alphabet 589–90; aspect 608–11, 619; compounds 606, 627; conjunctions 612; consonants 19, 591–3, 596–8, 629; coordinate clauses 620; copula 606–7; dialect 628–30; early written 587; geography

- 30–1, 587–8; interrogatives 600, 615; *izāfa* construction 612–13; marking 615–16; modals 619–20; modifiers 614; mood 608–11; negation 627; nouns 598–9, 626–7, 629–30; numbers 603–4; particles 616; phrase structure 612–13; pronouns 599–600, 615; relative clauses 621–2; stems 604–6; stress 594; subject and object clauses 622–3; subordinate clauses 620–1; tense 608–11; terms of address 628; transitivity 608; verbs 616–17, 618–20, 630; vowels 591, 595–6, 628; word order 613–14; and Zazaki 546
- Kushan, influence on Sogdian 328
- language contact, Sogdian 327–9
- language death: Khotanese and Tumshuqese 411; Sogdian 329–30
- Lazard, G. 508
- left-branching vs. right-branching typology 27–9
- length of vowels, Avestan 61–2, 67–8
- light verb constructions (LVC), Persian and Tajik 496–8
- literature, Balochi 635–6
- loanwords: in Balochi 645; in Khotanese 409–10; in Khwarezmian 372–3; in Middle Persian 263; in Pamir languages 784; in Parachi 716–17; in Persian and Tajik 529–30; in Sogdian 327–9; Tajik in Shughni 820; in Tumshuqese 410; in Wakhi 856–7
- locational clauses, Balochi 681–2
- locative case: Avestan 115–16; Old Persian 115–16
- locative clauses, Persian and Tajik 519
- lowering: Avestan 97–8; Old Persian 97–8
- MacKenzie, D. N. 30, 34, 204, 341, 344, 361, 362, 618, 628
- Maggi, M. 383
- Manichaeism 280–1
- Manichean script 196–7, 199, 282–4
- manner clauses, Balochi 682
- manuscripts, Avestan 45–6
- marginal phonemes, Sogdian 286
- marking: Kurdish 615–16; Persian and Tajik 483–6
- Median 43; consonants 57; dialects 15; Empire 7
- Mem ū Zīn* (Khānī) 587
- metre: Avestan 68–9, 96–8; Old Avestan 169–70; Old Persian and word order 96–8
- middle endings, Sogdian 299
- Middle Persian 196–7; conjunctions 35; consonant clusters 19; consonants 200–2; coordination 250–5; loan words 263; moods 234–9; nominal morphology 204–12; religious terminology 263–4; script 45, 202; speech 264–5; subordinate clauses 255–60; tenses 229–32; verb morphology 213–21; verb system 24–5; vowels 200; word order 246–9
- middle voice: Avestan 139; Old Persian 139
- Minorsky, V. 6
- modals: Balochi 677–8; Khotanese 404; Khwarezmian 363–4; Kurdish 619–20; Parachi 706–7, 712; Pashto 751–3; Persian and Tajik 489–93, 534; Shughni 808–10; Sogdian 310–12; Wakhi 847–8; Zazaki 558, 571
- Modern Literary Tajik (MLT) *see* Tajik
- Modern Persian, consonant clusters 19
- Modern Standard Kurdish (MSK) 587–9
- modifiers, Kurdish 614
- mood: Avestan 134–8; Balochi 662, 671–3, 674–5; Khotanese 396–8; Khwarezmian 352–4, 356; Kurdish 608–11; Middle Persian 234–9; Old Persian 134–8; Parachi 702–4, 705–6; Persian and Tajik 451, 453, 456–9, 488, 505–6, 517–18; Sogdian 298–9, 305, 310, 311; Tajik 465–8; Wakhi 837–40; Zazaki 559 62, 572–3
- Morgenstierne, G. 693, 776
- morphophonology: Avestan 59–70; Old Persian 59–70; Persian and Tajik 430
- Mountainous Badakhshan Autonomous Region 773, 787
- MPers. *see* Middle Persian
- Munji *see* Pamir languages
- names, Shughni 817
- narration, Shughni 816
- narrative past tense: Avestan 128–9; Old Persian 129
- nasalisation: Avestan 55; Khwarezmian 340
- negation: Avestan 129; Khotanese and Tumshuqese 395; Khwarezmian 356–7, 360; Kurdish 607–8, 627; Middle Persian 221 [correlative 253]; Old Persian 129; Parachi 702; Pashto 743; Persian and Tajik 452; Shughni 799; Sogdian 304; Wakhi 836
- neologisms, Persian and Tajik 530–1
- nominal clauses: Khotanese 407–8; Middle Persian 256–7; Parachi 715–16; Persian and Tajik 514–15, 523–6; Sogdian 320; Zazaki 576
- nominal morphology: Balochi 651–8; Khotanese and Tumshuqese 384–91; Khwarezmian 342–8; Kurdish 598–604; Middle Persian 204–12; Pamir languages 780–1; Parachi 696–700; Pashto 726–37; Persian and Tajik 430–46; Shughni 792–7; Sogdian 287–95; Wakhi 828–33; Zazaki 548–55
- nominal verbs, Middle Persian 215
- nominalization: Khwarezmian 370; Persian and Tajik 509–11
- nominative case: Avestan 103–4, 107–8 [and pronouns 118]; Old Persian 103–4

- North Pamir group 776; *see also* Pamir languages
- noun phrases, Pashto 754–5
- nouns: Avestan 71–9; Balochi 684; Khotanese 384–6; Khotanese and Tumshuqese 399–400; Khwarezmian 343–4, 370; Kurdish 598–9, 626–7, 629–30; Middle Persian 204–6, 221–3, 242–6, 259, 261–3; Old Persian 71–9; Parthian 221–3; Persian and Tajik 431–2; Shughni 802; Zazaki 548–9
- number: Avestan 71, 84–6, 101–2, 124; Balochi 658; Khotanese 390; Khwarezmian 342–3, 348; Kurdish 603–4; Middle Persian 205, 211–12; Old Persian 84–6, 101–2, 124; Pamir languages 781, 782; Parachi 696, 699–700; Pashto 736–7; Persian and Tajik 431, 445–6; Shughni 793–4, 797; Sogdian 295; Tumshuqese 391; Wakhi 832–3; Zazaki 554–5
- Nyberg 204
- object clauses: Khwarezmian 365–6; Kurdish 622–3; Parachi 714; Pashto 760–1; Shughni 812
- object marking: Balochi 669–70; differential (DOM) 33–4; Middle Persian 233–4; Parachi 710–11; Parthian 233; Shughni 804; Wakhi 844; Zazaki 570
- Öhl, P. and Korn, A. 516
- Old Avestan (OAv.) 44, 46, 48; adverbial clauses 160; conditional clauses 164–5; consonants 56–7; ellipsis in poetry 171–2; metre 68; relative clauses 155–8; vowels 54–5; word order, poetic 169–70; *see also* Avestan
- Old Iranian: dialects 51; *see also* Avestan; Old Persian
- Old Persian 46–8; adverbial clauses 161; adverbs 80; cases 103–16; conditional clauses 166; consonant clusters 19; consonants 65–6; coordination 147–53; lexis 167–9; moods 134–8; nouns 71–9; participles 141–5; phonology 18, 58–9; pronouns 80–3; relative clauses 159; script 52–3; stylistics 172–4; subordination 153–66; tenses 126–34; verb stems 87–90; voice 139–41; word order 94–6; *see also* Old Iranian
- optative mood: Avestan 90, 92, 136–7; Middle Persian 236–8; Old Persian 90, 136–7; Parthian 237–8
- Ormuri, consonant clusters 21
- orthography *see* spelling
- Ossetic: ergative construction 32; phonology 18
- Pahlavi 196, 419; past perfect tense 232; religious terminology 264; script 45, 197–9, 202–4; subjunctive 236
- Pakhalina 827, 855
- Pakistan 766; Wakhi in 826
- Pamir languages 781; case 781; consonants 777–80; gender 780–1; loanwords 784; number 781, 782; pronouns 781–2; semantic sets 783; tense 782–3; vowels 777–8
- ‘parabolic’ optative: Middle Persian 237–8; Parthian 238
- Parachi: adjectives 697–8; adpositions 699; adverbial clauses 714–15; adverbs 699; affixation 716; aktionsart constructions 706; aspect 705–6; case 696–7; conjunctions 708, 715; consonants 21, 694–5; coordinate clauses 713; definiteness 710; dialects 717; enclitics 711; gender 696; geography 693; loanwords 716–17; modals 706–7, 712; mood 702–4, 705–6; negation 702; nominalized clauses 715–16; number 696, 699–700; object clauses 714; object marking 710–11; person marking 701–2; phrase structure 708–9; possessives 712; plural marking 22–3; relative clauses 713–14; stems 701; stress 695; subject marking 710; subordinate clauses 713; tenses 702–3, 705–6; voice 707–8; vowels 694; word order 709
- parallelism: Avestan 170–1; Old Persian 170–1
- parataxis: Avestan 147; Middle Persian 254; Old Persian 147
- parenthetical clauses, Avestan 99
- Parthian 196; conjunctions 251–2; consonant clusters 19; dating 260; imperative mood 238–9; imperfect tense 230; infinitives 244; modification of nouns 221–3; object marking 233; optative mood 237–8; ‘parabolic’ optative 238; participles 245; questions 249, 257; reported speech 257; scripts 198; subjunctive mood 235–6; subordinate clauses 235–6; tenses 218–19, 231; ‘to be’ 216–17, 227; *see also* Middle Persian
- participle of necessity, Middle Persian 242
- participles: Avestan 93–4, 141–5, 150–3; Khwarezmian 351–2; Middle Persian 205–6, 239–41, 244–6, 254–5; Old Persian 93–4, 141–5, 150–3; Parthian 245; Pashto 742; Persian and Tajik 449; Shughni 798–9; Sogdian 297, 302–3; Wakhi 834–5
- particles: Khotanese 391–2; Khwarezmian 355–7; Kurdish 616; Pashto 757; Sogdian 300, 310–12; Zazaki 558
- Pashto: ablaut 729; adjectives 731–2; adpositions 735–6; adverbial clauses 762–6; agreement 756; alphabet 722–3; aspect 738, 739–42, 743–4; case 730–1; conjunctions 758–9; consonants 723–5; copula 738–9; dialects 766–7; early written 721–2; enclitics 756–7; gender 726–8; interrogatives 759; kinship terms 729; modals 751–3; negation

- 743; numbers 736–7; object clauses 760–1; participles 742; particles 757; phrase structure 754–5; pronouns 733–5; relative clauses 759–60; stems 739–42; stress 21, 725–6; subject clauses 760; tense 739–42, 746–51; transitivity 745; voice 753–4; vowels 723; word order 755
- Paul, L. 533
- perfective aorist *see* preterite
- periphrastic adnominal constructions, Middle Persian 223–4
- periphrastic forms, Sogdian 302–3
- permansive particle, Khwarezmian 355–6
- Persian: alphabet 421–3; adjectives 432–4, 475; adpositions 440–4, 513–14; adverbial clauses 516–26; adverbs 444–5, 480–1; affixation 526–8; agency 499–500; aktionsart constructions 493–5; anaphora 503–4; aspect 488, 518; case 440–4; causatives 448; classifiers 478–9; clitics 471–2; complement clauses 488, 511–16; compounds 476, 528–9; conjunctions 468–70, 501–2, 534; consonants 426–9; copula 450–1, 484; defective verbs 459–60; definiteness 483; direct speech 513; DOM 33; early location 30–1; elliptic superlatives 434; evidential forms 461; exclamations 482; *ezāfe* construction 470–7; formality 531; head incorporation 508; hiatus breakers 429; history 533; homoglossia 535; homonymy 526; indefiniteness 432, 439, 506–7; infinitival clauses 524–5; interrogatives 438–40, 481–2; kinship terms 532; light verb constructions 496–8; loanwords 529–30; marking 483–6; modals 389–93, 534; mood 451, 453, 456–9, 488, 505–6, 517–18; morphophonology 430; negation 452; neologisms 530–1; nominal clauses 514–15, 523–6; nominalization 509–11; nouns 431–2; numbers 431, 445–6; participles 449; phonemes 423–4; plural marking 22–3; possessives 435, 474; preterite 454, 518; preverbs 448; progressive 461–2, 534; pronominal clitics 486–8; pronouns 434–40; realis 500; relative clauses 503–11, 524; specificity 483, 485–6; stems 446–7; stress 429–30; subordinate clauses 502–3; syllable structure 427–8; tenses 452–6; terms of address 532–3; varieties 417–19; verbal nouns 475–6; voice 498–9; vowels 425–6, 533–4; word order 479–82
- person marking: Avestan 91–3; Khotanese 393–4; Khwarezmian 352; Parachi 701–2; Zazaki 557
- personal pronouns: Avestan 81–2, 116–19; Balochi 653–4; Cizre–Botan 629–30; Khotanese 387; Khwarezmian 344–5; Kurdish 599; Middle Persian 207–8; Old Persian 116–19; Parachi 697; Pashto 733–4; Persian and Tajik 434–5; Shughni 794; Sogdian 290–1; Tumshuqese 388; Wakhi 829–31; Zazaki 551
- phonemes, Persian and Tajik 423–4
- phrase structure: Balochi 667; Kurdish 612–13; Parachi 708–9; Pashto 754–5; Sogdian 314
- pluperfect, Avestan 92
- plural marking 22–3
- possession: Balochi 666–7; Khotanese 404; Parachi 712; Persian 435; Persian and Tajik 474; Shughni 804–5; Wakhi 846; Zazaki 570
- possessive pronouns: Avestan 82, 119; Kurdish 600; Middle Persian 208, 223; Old Persian 119; Parthian 223; Pashto 735; Shughni 795; Sogdian 307; Wakhi 831; Zazaki 551–2
- potential constructions, Old Persian 145
- prefixation *see* affixation
- prepositions, Sogdian 293–4; *see also* adpositions
- present marking 24, 26
- present participles: Avestan 141–2; Old Persian 141–2
- preterite: Middle Persian 230–1; Persian and Tajik 454, 518; Sogdian 301, 309
- preverbs: Avestan 95–6; Middle Persian 214; Old Persian 95–6; Persian and Tajik 448
- priests and the Avesta 44–5
- principle of anticipation, Khwarezmian 360
- progressive: Persian 461–2, 534; Tajik 465, 534
- pronominal clitics, Persian and Tajik 486–8
- pronouns: Avestan 80–3, 84, 116–24; Balochi 653–5; Khotanese 387–8, 401–2; Khwarezmian 344–5, 359; Kurdish 599–600, 615; Middle Persian 207–10, 223; Old Persian 80–3, 84, 116–24; Pamir languages 781–2; Parachi 697–8; Pashto 733–5; Persian and Tajik 434–40; Shughni 794–5; Sogdian 290–3; Tumshuqese 388; Wakhi 829–31; Zazaki 551–2, 575
- proterokinetic stems, Avestan 75–6
- Proto-Indo-Iranian, development 5
- Proto-Iranian, phonology 50
- Psalter manuscripts 198
- purpose clauses: Kurdish 625–6; Persian and Tajik 522–3; Wakhi 853
- questions *see* interrogatives
- raising: Avestan 60–1, 97; Old Persian 97
- realis, Persian and Tajik 500
- reason clauses, Wakhi 853
- reciprocal pronouns: Tumshuqese 388; Zazaki 552
- reciprocity: Avestan 123; Old Persian 123; Persian and Tajik 437
- reflexive pronouns: Avestan 82–3, 122; Kurdish 600; Middle Persian 208–9, 229; Old Persian 82–3, 122; Parachi 698; Pashto 735; Persian

- and Tajik 436–7; Shughni 795; Sogdian 291; Tumshuqese 388; Wakhi 831; Zazaki 551–2
- relative clauses (RCs): Avestan 154–60; Balochi 680; Khotanese 406; Khwarezmian 365; Kurdish 621–2; Middle Persian 255–6; Old Persian 154–60; Parachi 713–14; Pashto 759–60; Persian and Tajik 503–11, 524; Shughni 811–12; Sogdian 318–19; Wakhi 849–51; Zazaki 574–5
- relative phrases: Avestan 100–1; Old Persian 100–1
- relative pronouns: Avestan 82; Khwarezmian 344–5; Old Persian 82; Sogdian 293
- religious terminology, Middle Persian 263–4
- repetition as coordination: Avestan 147–9; Middle Persian 253; Old Persian 147–9
- reported speech: Avestan 99–100; Old Persian 99–100; Parthian 257
- request clauses, Shughni 816
- restriction, Middle Persian 254
- result clauses, Persian and Tajik 522
- retroflex consonants, Balochi 643
- rhotacisation, Sogdian 285
- 'rhythmic law' 286–7
- right-branching vs. left-branching typology 27–9
- rounding, Avestan 60–1
- Rumi 418
- Russian, loanwords in Persian and Tajik 530
- Saka groups 8
- sandhi: Avestan 67; Old Persian 67
- Sangesari 15, 337–8; consonant clusters 21
- Sanglichi *see* Pamir languages
- Sarikoli *see* Pamir languages
- Scheucher, B. 32
- Schindler 46
- script *see* alphabet
- Scythia 8
- semantic sets, Pamir languages 783
- settlement of Arabs into Iran 419
- Shahdara 819; *see also* Shugh(an)ni
- shibboleths 530
- Shughn(an)i 775; adpositions 796–7; adjectives 794, 801–2; adverbial clauses 813–16; adverbs 797, 804; affixation 816–17; agreement 805; aktionsart constructions 807–8; alienability 803; alphabet 788; articles 794; compounds 798, 817; conjunctions 800–1; consonants 790–1; coordination 810–11; copula 799; dialect 819–20; gender 792–3; kinship terms 818; loanwords 820; modals 808–10; names 817; negation 799; nouns 802; number 793–4, 797; object clauses 812; object marking 804; participles 798–9; possession 804–5; pronouns 794–5; relative clauses 811–12; stems 797–8; stress 791–2; subject marking 804; subordinate clauses 811; tense 800, 806–7; terms of address 818; verbs 802; voice 800; vowels 789; word order 801, 803
- Shughn(an)i–Rushani group 776, 787–8
- sibilants, Avestan 57
- Siever's law 69
- Sims-Williams, N. 21, 286, 310
- sociolinguistics, Sogdian 324–30
- Sogdian: adjectives 290; adpositions 293–4; adverbial clauses 319–20; adverbs 294; affixes 322; agreement 312–13; case-marking 305–6; collocation 323–4; complex sentences 317–21; compounds 323; conjunctions 317–18; consonants 285–6; copula 300; death of 329–30; demonstratives 291–3; dialects 326–7; direct speech 320; early stages 325–6; ergative construction 32; honorific expressions 324–5; infinitives 296–7, 320–1; interjections 294; interrogatives 316–17; language contact 327–9; modals 310–12; moods 298–9, 305, 310, 311; negation 304; nominal clauses 320; nominal stems 288–9; nouns 288, 290; numbers 295; participles 297, 302–3; particles 300, 310–12; periphrastic forms 302–3; plural marking 23; possessives 307; preterite, 301, 309; pronouns 290–3; relative clauses 318–19; script 281–4; stems 305–6, 312–13; stress 286–7; tenses 305, 307–9; verbal stems 295–6, 298–9, 301–2; voice 303–4; vowels 284–5; word order 314–16
- Sogdiana 279–80
- Southeast Iranian, consonant clusters 21
- Southern Balochi 637; consonants 645; stress 649; *see also* Balochi
- specificity, Persian and Tajik 483, 485–6
- speech, Middle Persian 264–5
- spelling: Balochi 638; Khotanese 381; Sogdian 287
- spirantisation, Avestan 63–4
- Steblin-Kamenskii, I. M. 21
- stems: Avestan 70–2, 75–9, 87–90; Balochi 659–61; Khotanese 384–6, 391; Khwarezmian 349–52; Kurdish 604–6; Middle Persian 213–14; Old Persian 71–2, 87–90; Parachi 701; Pashto 739–42; Persian and Tajik 446–7; Shughni 797–8; Sogdian 288–9, 295–6, 298–9, 301–2, 305–6, 312–13; Wakhi 833–4; Zazaki 555–7
- Stilo, D. 27
- stress 21–2; Avestan 67–8; Balochi 648–9; Khotanese 383; Khwarezmian 341; Kurdish 594; Parachi 695; Pashto 725–6; Persian and Tajik 429–30; Shughni 791–2; Sogdian 286–7; Wakhi 828; Zazaki 547
- stylistics: Avestan 169–72, 173–4; Old Persian 172–4
- subject clauses: Khwarezmian 365–6; Kurdish 622–3; Pashto 760

- subject marking: Balochi 669; Parachi 710; Shughni 804; Wakhi 842–4; Zazaki 569–70
- subjunctive mood: Avestan 90, 92, 135–6; Middle Persian 234–6; Old Persian 90, 135–6; Pahlavi 236; Parthian 235–6
- subordinate clauses 35; Balochi 678–9; Kurdish 620–1; Middle Persian 235–6, 255–60; Parachi 713; Parthian 235–6; Persian and Tajik 502–3; Shughni 811; Zazaki 574
- subordination: Avestan 153–66; Khwarezmian 364; Old Persian 153–66; Wakhi 849
- substantival *that*-clauses: Avestan 153–4; Old Persian 153–4
- suffixation *see* affixation
- Sulaimani Kurdish *see* Kurdish
- superlatives, Avestan 125
- syllable structure: Balochi 647–8; Persian 427–8
- Syriac: script 282–4; spelling 287
- Tajik 420–1: adjectives 432–4, 475; adpositions 440–4, 513–14; adverbial clauses 516–26; adverbs 444–5, 480–1; affixation 526–8; agency 499–500; aktionsart constructions 493–5; alphabet 423; anaphora 503–4; aspect 488, 518; case 440–4; causatives 448; classifiers 478–9; clitics 471–2; complement clauses 488, 511–16; compounds 476, 528–9; conjunct verbs 495–6; conjunctions 468–70, 501–2, 534–5; consonants 426–9; copula 450–1, 484; defective verbs 459–60; definiteness 483; direct speech 513; dialects 776; evidential forms 462–4; exclamations 482; *ezāfe* construction 470–7; formality 531; head incorporation 508; hiatus breakers 429; homoglossia 535; homonymy 526; indefiniteness 432, 439, 506–7; infinitival clauses 524–5; interrogatives 438–40, 481–2; kinship terms 532; light verb constructions 496–8; loanwords 529–30 [in Pamir languages 784; in Shughni 820]; marking 483–6; modals 389–93, 534; mood 451, 453, 456–9, 465–8, 488, 505–6, 517–18; morphophonology 430; negation 452; neologisms 530–1; nominal clauses 514–15, 523–6; nominalization 509–11; nouns 431–2; numbers 431, 445–6; participles 449; phonemes 423–4; possessives 474; preterite 454, 518; preverbs 448; progressive 465, 534; pronominal clitics 486–8; pronouns 434–40; realis 500; relative clauses 503–11, 524; specificity 483, 485–6; stems 446–7; stress 429–30; subordinate clauses 502–3; tenses 452–6; terms of address 532–3; verbal nouns 475–6; voice 498–9; vowels 426, 533–4; word order 479–82
- Tajikistan 773–5, 787; Shughni in 788; Wakhi in 826
- Talyshi, consonant clusters 19
- TAM-forms, Balochi 665–6, 671–5
- Tati group: aspect marking 26; consonant clusters 19
- temporal clauses: Balochi 681; Kurdish 623; Parachi 714; Pashto 762–3; Persian and Tajik 519, 522; Shughni 813; Wakhi 851–3; Zazaki 576–8
- tense: Avestan 126–34; Balochi 662, 664–5, 673–4; Khotanese 395–6, 404; Khwarezmian 351–4; Kurdish 608–11; Middle Persian 218–19, 229–32; Old Persian 126–34; Pamir languages 782–3; Parachi 702–3, 705–6; Parthian 218–19; Pashto 739–42, 746–51; Persian and Tajik 452–6; Shughni 800, 806–7; Sogdian 305, 307–9; Wakhi 837–40; Zazaki 559–62
- tense-split ergatives 31–3
- terms of address: Kurdish 628; Persian and Tajik 532–3; Shughni 818; Zazaki 583
- 'to be' *see* copula
- transitivity: Balochi 663–4; Kurdish 608; Middle Persian 219–21; Pashto 745
- v*-clusters 21
- Tumshuqese: adjectives 387; conjunctions 405; consonants 383; history 377–8, 379; infinitives 392–3; language death 411; loanwords 410; negation 395; nouns 399–400; numbers 391; pronouns 388; script 381; vowels 383
- Turkic 17; loanwords in Persian and Tajik 530
- Turkification of Persian 419–20
- v*-epenthesis, Avestan 65
- Uighur: influence from Sogdian 329–30; orthography 285
- verb morphology: Balochi 658–67; Khotanese 391–8; Khwarezmian 349–57; Kurdish 604–11; Pamir languages 782–3; Parachi 701–8; Pashto 738–54; Persian and Tajik 446–68, 469; Shughni 797–800; Sogdian 295–305; Wakhi 833–40; Zazaki 555–62
- verb stems: Avestan 87–90; Old Persian 87–90
- verbal nouns, Persian and Tajik 475–6
- verbal quincunx systems 23–6
- verbs: Balochi 670–8, 684–5; Khwarezmian 354–5, 362–4; Kurdish 616–17, 618–20, 630; Middle Persian 213–21, 241–4; Parthian 244; Shughni 802; Zazaki 570–1
- vocative case: Avestan 104–5; Old Persian 104–5
- voice: Avestan 90–1, 139–41; Balochi 662–3; Old Persian 90–1, 139–41; Parachi 707–8; Pashto 753–4; Persian and Tajik 498–9; Shughni 800; Sogdian 303–4
- voicing, Avestan 62

- vowels: Avestan 54–5, 59–62, 74; Balochi 641–2, 648, 649–50; Indo-Iranian vs. Indo-European 49; Khotanese 382; Khwarezmian 339–40, 341–2; Kurdish 591, 595–6, 628; Middle Persian 200, 202; Old Avestan, initial 60; Old Persian 58; Pamir languages 777–8; Parachi 694; Pashto 723; Persian 425–6; Persian and Tajik 533–4; Shughni 789; Sogdian 284–5; Tajik 426; Tumshuqese 383; Wakhi 826–7; Zazaki 547
- Wackernagel's law 315
- Wakhi 15; adjectives 829; adpositions 831–2; adverbs 832; alphabet 826; aspect 838–40, 846–7; case 829, 842–6; causee marking 845–6; complement structures 853–4; compounds 855; conditional clauses 854–5; conjunctions 840; consonants 827–8; coordinate clauses 849; copula 835–6; dialects 855–6; geography 825–6; loanwords 856–7 [in Shughni 820]; modals 847–8; mood 837–40; negation 836; number 832–3; object marking 844; participles 834–5; possession 846; pronouns 829–31; purpose clauses 853; reason clauses 853; relative clauses 849–51; stems 833–4; stress 828; subject marking 842–4; subordination 849; temporal clauses 851–3; tense 837–40; vowels 826–7; word order 840–2
- West Iranian: consonant clusters 18–21; isoglosses 29–31
- Western Balochi 637; consonants 645; stress 648–9; *see also* Balochi
- Western Middle Iranian, influence on Sogdian 328
- Windfuhr, G. 20, 30, 452
- wish clauses, Shughni 815
- word order 27–9; Avestan 94–6; Balochi 668; Khotanese 401–3; Khwarezmian 358–61; Kurdish 613–14; Middle Persian 246–9; Old Avestan, poetic, 169–70; Old Persian 94–6, 172; Parachi 709; Pashto 755; Persian and Tajik 479–82; Shughni 801, 803; Sogdian 314–16; Wakhi 840–2; West Iranian 29–30; Zazaki 567–8, 571–3
- Xuan-zang 411
- Yaghnobi 15, 327; *see also* Sogdian
- Yaz culture 6
- Yazghulami *see* Pamir languages
- Young Avestan (YAv.) 44, 46; adverbial clauses 160–1; conditional clauses 165–6; consonants 56–7; metre 68; relative clauses 158; vowels 54–5; *see also* Avestan
- Zagros group 28
- Zazaki: adjectives 549–51; adpositions 552–3; adverbial clauses 576–81; adverbs 553–4; affixation 558, 582; aspect 559–60; cases 567–8; compounds 556–7, 570, 582; conjunctions 563; consonants 19, 547; coordinate clauses 573–4; copula 557–8; dialect 582–3; *ezāfe* constructions 563–7; history 30, 545–6; interrogatives 572; kinship terms 583; modals 558, 571; mood 559–62, 572–3; nominal clauses 576; nouns 548–9; numbers 554–5; object marking 570; particles 558; person marking 557; possessives 570; pronouns 551–2, 575; relative clauses 574–5; stems 555–7; stress 547; subject marking 569–70; subordinate clauses 574; tense 559–62; terms of address 583; verbs 570–1; vowels 547; word order 567–8, 571–3
- Zoroastrian texts 196