
THE

SINO-TIBETAN

LANGUAGES

There are more native speakers of Sino-Tibetan languages than of any other language family in the world. Our records of these languages are among the oldest for any human language, and the amount of active research on them has multiplied in the last few decades.

Now in its second edition and fully updated to include new research, *The Sino-Tibetan Languages* includes overview articles on individual languages, with an emphasis on the less commonly described languages, as well as descriptions and comments on the subgroups in which they occur.

There are overviews of the whole family on genetic classification and language contact, syntax and morphology, and also on word order typology. There are also more detailed overview articles on the phonology, morphosyntax, and writing system of just the Sinitic side of the family. Supplementing these overviews are articles on Shanghaiese, Cantonese, and Mandarin dialects. Tibeto-Burman is reviewed by genetic or geographical subgroup, with overview articles on some of the major groups and areas, and there are also detailed descriptions of 41 individual Tibeto-Burman languages, written by world experts in the field.

Designed for students and researchers of Asian languages, *The Sino-Tibetan Languages* is a detailed overview of the field. This book is invaluable to language students, experts requiring concise, but thorough, information on related languages, and researchers working in historical, typological, and comparative linguistics.

Graham Thurgood PhD is Professor Emeritus at California State University, Chico, USA.

Randy J. LaPolla PhD FAHA is Professor of Linguistics at Nanyang Technological University, Singapore.

ROUTLEDGE LANGUAGE FAMILY SERIES

Each volume in this series contains an in-depth account of the members of some of the world's most important language families. Written by experts in each language, these accessible accounts provide detailed linguistic analysis and description. The contents are carefully structured to cover the natural system of classification: phonology, morphology, syntax, lexis, semantics, dialectology, and sociolinguistics.

Every volume contains extensive bibliographies for each language, a detailed index and tables, and maps and examples from the languages to demonstrate the linguistic features being described. The consistent format allows comparative study, not only between the languages in each volume, but also across all the volumes in the series.

The Austronesian Languages of
Asia and Madagascar
*Edited by Nikolaus Himmelmann &
Sander Adelaar*

The Indo-European Languages 2nd Edition
Edited by Mate Kapović

The Iranian Languages
Edited by Gernot Windfuhr

The Languages of Japan and Korea
Edited by Nicolas Tranter

The KhoeSan Languages
Edited by Rainer Vossen

The Mongolic Languages
Edited by Juha Janhunen

The Munda Languages
Edited by Gregory D.S. Anderson

The Oceanic Languages
*Edited by John Lynch, Malcolm Ross &
Terry Crowley*

The Romance Languages
*Edited by Martin Harris &
Nigel Vincent*

The Semitic Languages
Edited by Robert Hetzron

The Sino-Tibetan Languages,
2nd Edition
*Edited by Graham Thurgood &
Randy J. LaPolla*

The Slavonic Languages
*Edited by Bernard Comrie &
Greville G. Corbett*

The Tai-Kadai Languages
Edited by Anthony Diller

The Turkic Languages
*Edited by Éva Csató &
Lars Johanson*

The Uralic Languages
Edited by Daniel Abondolo

THE
SINO-TIBETAN
LANGUAGES

Second edition

Edited by
Graham Thurgood 杜冠明
and
Randy J. LaPolla 羅仁地

Second edition published 2017
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
711 Third Avenue, New York, NY 10017

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

© 2017 Graham Thurgood and Randy J. LaPolla for selection and editorial matter; individual chapters, the contributors

The right of Graham Thurgood and Randy J. LaPolla to be identified as authors of the editorial matter, and of the contributors for their individual chapters has been asserted by them in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised 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.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

First edition published by Routledge 2003

British Library Cataloguing-in-Publication Data

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

Library of Congress Cataloguing-in-Publication Data

Names: Thurgood, Graham, editor. | LaPolla, Randy J., editor.

Title: The Sino-Tibetan languages/edited by Graham Thurgood and Randy J. LaPolla.

Description: Second edition. | Milton Park, Abingdon, Oxon; New York, NY: Routledge, [2017] | Series: Routledge language family series | Includes bibliographical references and index.

Identifiers: LCCN 2016030853 | ISBN 9781138783324 (hardback : alk. paper) | ISBN 9781315399508 (ebook)

Subjects: LCSH: Sino-Tibetan languages.

Classification: LCC PL3521.S56 2017 | DDC 495/.4–dc23

LC record available at <https://lccn.loc.gov/2016030853>

ISBN: 978-1-138-78332-4 (hbk)

ISBN: 978-1-315-39950-8 (ebk)

Typeset in Times New Roman
by Sunrise Setting Ltd, Brixham, UK

CONTENTS

<i>List of illustrations</i>	xvii
<i>Preface to the second edition</i>	xxv
<i>List of abbreviations</i>	xxvii

PART 1 OVERVIEW CHAPTERS	1
1 Sino-Tibetan: genetic and areal subgroups <i>Graham Thurgood</i>	3
1 Introduction	3
2 Sino-Tibetan	5
3 Chinese	6
4 Lolo-Burmese branch [Burmese-Lolo]	8
5 Bodish: Tibetan, East Bodish, Tshangla, and probably Tamangic	9
6 Qiangic languages	15
7 The ‘Sal’ languages	20
8 The Northeastern border area	21
9 The ‘Rung’ languages	24
10 Rawang and related languages [Nungish]	25
11 Kiranti	25
12 West Himalayan group	26
13 The Kham, Magar, and Chepang languages	26
14 The Karenic branch	27
15 Tani (Mirish)	27
16 Kamengic [Puroik, less often Puroit, etc.]	28
17 Idu-Digaru (Tawrā)	28
18 Kaman-Meyor [Midzuish]	28
19 Koro-Milang	29
20 Raute, Raji	30
21 Unsubgrouped languages	30
22 Conclusion	31
2 Overview of Sino-Tibetan morphosyntax <i>Randy J. LaPolla</i>	40
1 Sino-Tibetan	40
2 Sinitic	48
3 Tibeto-Burman	48

VI CONTENTS

3 Word order in Sino-Tibetan languages from a typological and geographical perspective <i>Matthew S. Dryer</i>	70
1 Introduction	70
2 Word order in Tibeto-Burman	70
3 Chinese	74
4 Conclusion	81
PART 2 SINITIC	83
4 The Sinitic languages: phonology <i>Zev Handel</i>	85
1 Introduction	85
2 Synchronic phonology: features of modern Chinese varieties	87
3 Diachronic phonology: features of historic Chinese varieties and developmental trends	96
4 Conclusion	110
5 The Sinitic languages: grammar <i>Anne O. Yue</i>	114
1 Introduction	114
2 Grammatical categories	114
3 Sentence structure and word order	129
4 Major sentence types	131
6 The characteristics of Mandarin dialects <i>Dah-an Ho</i>	164
1 Mandarin dialects in historical development	164
2 Some characteristics of Mandarin dialects	166
3 The regional characteristics of Mandarin dialects	167
7 Cantonese <i>Robert S. Bauer and Stephen Matthews</i>	169
1 Introduction	169
2 Phonology	171
3 Cantonese lexicon and script	175
4 Grammar	177
5 Conclusion	181
8 Shanghainese <i>Eric Zee and Liejiong Xu</i>	185
1 Shanghai phonology	185
2 Shanghai morphology and syntax	192
9 Chinese writing <i>Mark Hansell</i>	200
1 Structure and function	200
2 Historical development	204

3	Writing as evidence in historical reconstruction	207	
4	Relation to other languages	208	
PART 3 TIBETO-BURMAN		211	
§3.1 Languages of NE India		213	
10	The Tibeto-Burman languages of Northeast India		
	<i>Mark W. Post and Robbins Burling</i>	213	
1	Introduction	213	
2	Previous research	215	
3	Nomenclature	216	
4	Typology	217	
5	Classification	219	
6	Conclusions	233	
11	Garo	<i>Robbins Burling</i>	243
1	Introduction	243	
2	Phonology	244	
3	Verbs	246	
4	Noun phrases	247	
5	Adverbs and reduplication	253	
6	Complex sentences	254	
7	Language contact and language maintenance	256	
12	Hakha Lai	<i>David A. Peterson</i>	258
1	Introduction	258	
2	Phonology	258	
3	Inflectional morphology	259	
4	Syntax	268	
13	Mongsen Ao	<i>Alexander R. Coupe</i>	277
1	Overview of Ao dialects	277	
2	Genetic affiliation	279	
3	Phonology	280	
4	Suprasegmental features	282	
5	Word formation processes	284	
6	Morphosyntax	286	
14	Karbi	<i>Linda Konnerth</i>	302
1	Introduction	302	
2	Phonology	303	

VIII CONTENTS

3	Word classes	305
4	Morphology	306
5	Syntax	311
15	Tani languages <i>Mark W. Post and Jackson T.-S. Sun</i>	322
1	Introduction	322
2	Genealogy and subgrouping	323
3	Phonology	324
4	Morphology	326
5	Syntax	331
16	Meitei <i>Shobhana L. Chelliah</i>	338
1	Location and classification	338
2	Phonetics and phonology	338
3	Morphology	340
4	Major lexical categories	342
5	Evidentiality and epistemic modality	343
6	Syntax	344
17	Tangsa <i>Stephen D. Morey</i>	350
1	Introduction	350
2	Writing	355
3	Phonology	357
4	Morphology	361
§3.2 Bodish languages		369
18	Classical Tibetan <i>Scott DeLancey</i>	369
1	Old and classical Tibetan	369
2	Phonology	370
3	Nominals and the noun phrase	372
4	Verbs	376
5	Clause and sentence	379
19	Lhasa Tibetan <i>Scott DeLancey</i>	385
1	Tibetic languages	385
2	Phonology	385
3	Nominals and the noun phrase	388
4	The verb	391
5	Word formation	397
6	Syntax	397

20	Kurtöp	<i>Gwendolyn Hyslop</i>	404
1	Introduction		404
2	Background		404
3	Phonology		407
4	Morphosyntax		409
5	Evidentiality and related categories		414
21	Tshangla	<i>Erik Andvik</i>	418
1	Introduction		418
2	Phonology		418
3	Morphophonemics		419
4	Noun phrases		420
5	Syntactic roles		423
6	Case marking		423
7	Tense/aspect		425
8	Negation		426
9	Mirativity		426
10	Copular clauses		427
11	Modality and sentence-types		428
12	Relative clauses		429
13	Complementation		430
14	Adverbial clauses		431
15	Clause chains		432
16	Concatenation		432
17	Grammaticalized non-final verbs		433
22	Dolakha Newar	<i>Carol Genetti</i>	436
1	Introduction		436
2	Phonology		437
3	Morphology		438
4	Verbs		441
5	Syntax		443
6	Narrative text		446
23	Kathmandu Newar (Nepāl Bhāṣā)	<i>David Hargreaves</i>	453
1	Introduction		453
2	Phonology		454
3	Inflectional morphology		455
4	Word formation		459
5	Syntax		461

§3.3 Tamangic	468
24 Tamang <i>Martine Mazaudon</i>	468
1 Introduction	468
2 Phonology	468
3 Typological summary	471
4 Basic sentences	471
5 The noun phrase	475
6 The verb phrase	478
7 Complex sentences	480
8 Comparison	487
9 Questions	487
10 Information structure	489
25 Chantyal <i>Michael Noonan and Kristine A. Hildebrandt</i>	494
1 Introduction	494
2 Phonology	494
3 Morphology	497
4 Word formation	507
5 Syntax	508
26 Manange <i>Kristine A. Hildebrandt and Oliver Bond</i>	516
1 Introduction and typological summary	516
2 Phonology	516
3 Word formation, word classes and nominal morphology	520
4 Verb complex and simple clause structure	526
5 Clause combining	529
27 Nar-Phu <i>Michael Noonan and Kristine A. Hildebrandt</i>	534
1 Introduction	534
2 Phonology	535
3 Morphology	542
4 Word formation	551
5 Syntax	551
§3.4 Rung languages	557
§3.4.1 Rgyalrong languages	557
28 Tshobdun rGyalrong <i>Jackson T.-S. Sun</i>	557
1 Background	557
2 Phonology	557

3	Morphology	558
4	Syntax	565
29	Cogtse rGyalrong <i>Yasuhiko Nagano</i>	572
1	Introduction	572
2	Outline of phonology	573
3	Morphology and morphosyntax	574
30	Stau (Ergong, Horpa) <i>Guillaume Jacques, Lai Yunfan, Anton Antonov and Lobsang Nima</i>	597
1	Introduction	597
2	Phonology	597
3	Verbal morphology	599
4	Noun phrase	604
5	Nominalization	605
6	Complementation	606
7	Evidentiality	608
8	Classification	609
31	Japhug <i>Guillaume Jacques</i>	614
1	Introduction	614
2	Phonology	614
3	Verbal morphology	615
4	Nominal morphology	626
5	The noun phrase	627
6	Simple clauses	628
7	Comparative	628
8	Relativization	629
9	Complementation	631
<i>§3.4.2 Rawang-Dulong-Anong</i>		635
32	Dulong <i>Randy J. LaPolla</i>	635
1	Introduction	635
2	Phonology	635
3	Morphosyntax	636
<i>§3.4.3 Kiranti</i>		646
33	Kiranti languages <i>Boyd Michailovsky</i>	646
1	Introduction	646
2	Phonology	647

XII CONTENTS

3	Non-verbal morphology	651
4	The verb	654
5	Transitivity frames and diatheses	665
6	Other case frames	668
7	Nominalization	671
8	Subordinate clauses	673
9	Information structure and focus	674
34	Hayu <i>Boyd Michailovsky</i>	680
1	Introduction	680
2	Phonology	680
3	Morphophonology	681
4	Verb morphology	682
5	Non-verbal morphology and word classes	685
6	Syntax	687
7	Information structure and discourse particles	693
8	The noun phrase	693
35	Belhare <i>Balthasar Bickel</i>	696
1	Introduction	696
2	Phonology	696
3	Inflectional morphology	698
4	Derivational morphology and compounding	709
5	Syntax	711
36	Camling <i>Karen H. Ebert</i>	722
1	Introduction	722
2	Phonology	722
3	Nominals	723
4	The verb	726
5	Syntax	732
37	Wambule Rai <i>Jean Robert Opgenort</i>	736
1	Introduction	736
2	Phonology	736
3	Morphophonology	739
4	Nominal morphology	740
5	Verbal morphology	743
6	Syntax and clause combinations	753

§3.4.4 <i>Western Himalayan</i>	756
38 Sangla Kinnauri <i>Anju Saxena</i>	756
1 Introduction	756
2 Phonology	756
3 Noun phrase	757
4 Structure of the clause	762
 §3.5 Qiangic	 773
39 Qiang <i>Randy J. LaPolla</i>	773
1 The phonological system	773
2 The noun phrase	775
3 Nominal relational morphology	777
4 The verb complex	780
5 Adverbials	784
6 The clause	784
 40 Prinmi <i>Picus Sizhi Ding</i>	 790
1 Introduction	790
2 Phonology	790
3 Morphology	794
4 Syntax	797
 41 Tangut <i>Gong Hwang-cherng</i>	 805
1 Introduction	805
2 Phonology	805
3 Morphology	809
4 Word formation	813
5 Syntax	815
 42 Lizu (Ersu) <i>Katia Chirkova</i>	 823
1 Introduction	823
2 Phonology	823
3 Morphology	827
4 The noun phrase	827
5 The verb phrase	830
6 Phrase and clause structure	835

§3.6 Naic [Naxi]	840
43 Yongning Na (Mosuo) <i>Liberty A. Lidz</i>	840
1 Introduction	840
2 Phonology	841
3 Word classes	843
4 Syntax	847
5 Conclusion	854
§3.7 Lolo-Burmese	856
44 Burmese <i>Julian K. Wheatley</i>	856
1 Introduction	856
2 Contact	858
3 Periodization	859
4 Phonology	860
5 The writing system	862
6 Word classes	866
7 Syntax	869
45 Zaiwa <i>Yanhua Zhu</i>	877
1 Introduction	877
2 Phonology	877
3 Morphosyntax	878
46 Akha <i>Inga-Lill Hansson</i>	885
1 Introduction	885
2 Phonology	886
3 Word formation	888
4 Syntax	890
5 Sentence particles	897
6 Noun incorporation	898
7 Final particles	900
47 Lisu <i>David Bradley</i>	902
1 Introduction	902
2 Phonology	903
3 Morphosyntax	906

48 Lahu	<i>James A. Matisoff</i>	918
1	Introduction	918
2	Phonology	919
3	Lahu vocabulary and word formation	920
4	Clause structure	921
5	Form classes	921
6	Argument–verb relations	922
7	The noun phrase	924
8	The verb phrase (vp)	926
9	Nominalization and relativization	928
§3.8 Karenic		932
49 Eastern Kayah Li	<i>David Solnit</i>	932
1	Introduction	932
2	Phonology	933
3	Word formation	935
4	Syntax	936
50 Pwo Karen	<i>Atsuhiko Kato</i>	942
1	Introduction	942
2	Sounds	943
3	Parts of speech	944
4	Morphology	948
5	Syntax	950
§3.9 Other languages		959
51 Lepcha	<i>Heleen Plaisier</i>	959
1	Introduction	959
2	Phonology	960
3	Nominal morphology	960
4	Verbal morphology	967
5	Clause-final particles, coordination and subordination	970
6	Text	971
52 Tujia	<i>Shixuan Xu, Meiyian Lu and Hongyan Hu</i>	974
1	Introduction	974
2	Phonology	974
3	Morphosyntax	977

XVI CONTENTS

53 Jinghpaw Keita Kurabe	993
1 Introduction	993
2 Phonology	994
3 Lexicon	996
4 Morphology	996
5 Syntax	998
<i>Index</i>	1011

ILLUSTRATIONS

FIGURES

1.1 Tibeto-Burman versus Chinese: two views	6
1.2 Chinese ‘dialects’ (languages)	7
1.3 Lolo-Burmese	9
1.4 The East Bodish linkage	15
2.1 The subgrouping of Qiangic-Rung	49
4.1 Pitch track of an utterance by a native speaker of Standard Mandarin containing only fourth-tone syllables	93
5.1 Deictic forms of Min, Yue, Hakka, Wu and Gan dialects	122
5.2 Deictic forms of Xiang and Mandarin dialects	123
5.3 Interrelationship of the four constructions with ‘give’ as the main verb	146
7.1 Structure of the Cantonese syllable	173
9.1 Compound characters	203
9.2 Characters which include the ‘horse’ radical	203
9.3 The historical development of Chinese characters	205
9.4 The development of ‘form–sound’ characters	206
9.5 The realization of a phonetic in different dialects	209
13.1 Mongsen Ao tonemes	283
14.1 Semantic map of basic clausal functions of Karbi parts of speech	306
14.2 Postverbal morphology	308
14.3 Karbi noun phrase structure	311
15.1 Provisional family tree of the Tani languages	323
15.2 Classificatory word formation in Galo	329
20.1 Kurtöp coda consonants	408
20.2 Kurtöp complex onsets	408
20.3 Contrasts made in perfective aspect	414
20.4 Kurtöp imperfective aspect suffixes	415
25.1 Relations among the Tamangic languages	495
25.2 The tense-aspect system	504
26.1 F1~F2 plots, Manange vowels	518
26.2 F ₀ traces, sonorant-initial monosyllabic words	519
26.3 F ₀ traces, obstruent-initial monosyllabic words	520
26.4 F ₀ traces, disyllabic words	520
26.5 F ₀ traces, suffixed verbs	521
27.1 F1~F2 plots, Nar-Phu vowels	536
27.2 Nasal-initial pitch traces	539
27.3 Obstruent-initial pitch traces	540
27.4 Waveform and spectrogram of <i>nâŋ</i> ‘obligation’ and <i>nħâŋ</i> ‘inside’	540
27.5 Pitch traces for disyllabic words	541

XVIII ILLUSTRATIONS

29.1	Independent personal pronouns	576
29.2	Morphosyntax of prefixes	578
29.3	Layers of prefixes	579
29.4	Semantic layers of prefixes	579
40.1	The structure of the syllable in Prinmi	791
40.2	A layered analysis of the noun phrase	797
40.3	The structure of simplex sentences	798
44.1	Tonal contrasts (citation)	861
44.2	Initial contrasts	861
44.3	Pronouns	868
44.4	Case and other core postpositions	871
46.1	Relations among the Southern Yiphō languages	886

MAPS

3.1	Order of adjective and noun	73
3.2	Order of adjective and noun in Asia	73
3.3	Worldwide distribution of two orders of relative clause and noun among OV languages	79
4.1	The major Chinese dialect groups	88
10.1	Languages of the Northern Area	221
10.2	Languages of the Central Area	225
10.3	Languages of the Eastern Border Area	229
20.1	Approximate location of the Kurtöp homeland	405
42.1	Distribution of Ersu, Lizu, and Duoxu	824
42.2	Distribution of Lizu dialects	824

TABLES

1.1	Reflexes of the three non-checked PLB tones	9
1.2	Loloish checked tones	10
1.3	Tournadre's eight nuclei	12
1.4	Written Tibetan s+ nasal compared with East Bodish tones	13
1.5	Linking shared innovations in East Bodish	14
1.6	Directional verb prefixes	16
1.7	Directional verb prefixes in core Naish, Shixing, and Namuyi	19
1.8	Reflexes of *ry- in core Naish, Shixing and Namuyi, and in PLB	19
1.9	The 'Rung' languages	24
1.10	West Himalayan pronouns and pronominalization	26
1.11	Additional evidence for TB membership and subgrouping	29
1.12	Reputedly non-Tibeto-Burman cognates	29
1.13	Hruso, Dhimmai, and Levai	30
4.1	The initials of Jiān'ōu dialect	90
4.2	The initials of Chénxī dialect	90
4.3	The finals of Wēnzhōu dialect	91
4.4	The four tones of Standard Mandarin	92
4.5	Some tone sandhi rules in Sūzhōu	94
4.6	Pronunciations of rhotacized syllables	95

4.7	Reconstructing Middle Chinese pronunciations of three characters in the <i>Guāngyún</i> Dōng 東 rhyme	99
4.8	The eight-tone framework for comparative Chinese dialectology	100
4.9	Comparative chart of entering-tone stop-coda developments	102
4.10	Comparative chart of nasal-coda developments	103
4.11	Comparative chart of MC voiced obstruents	105
4.12	Comparative chart of MC nasal initials	105
4.13	Comparative chart of palatalization	105
4.14	Comparative chart of dentalibialization	106
4.15	Comparative chart of affricate series mergers	107
4.16	The tone system of Wēnzhōu	108
4.17	The tone system of Cantonese	108
4.18	The tone system of Xi'ān	108
4.19	The tone system of Xi'ān and Middle Chinese sources	108
5.1	Sample plural forms of personal pronouns	118
5.2	Tripartite distinction in demonstratives	120
5.3	Perfective aspect, affirmative aspect, past tense	124
5.4	Final phase particle	126
5.5	Experiential aspect	127
5.6	Progressive aspect	129
5.7	Durative aspect	129
5.8	Three types of existential sentences	132
5.9	Neutral questions	140
5.10	Four types of double-subject construction—NP1 (Pause) NP2 VP	147
5.11	Comparative constructions	150
6.1	Comparisons of the attributes in Chinese	165
6.2	Comparisons of the attributes of Modern Chinese dialects	165
6.3	Tone sandhi in the Changzhi dialect	167
6.4	Comparisons of Mandarin dialects by region	168
7.1	Cantonese initial consonants	171
7.2	Cantonese vowel phonemes (between //) and allophones (between [])	172
7.3	60 Cantonese rhymes	172
7.4	Lexical items contrasting nine (ten) tone contours on open syllable <i>ji</i>	174
7.5	Register stratification in Cantonese syntax	178
8.1	Consonants	186
8.2	Vowels	187
8.3	Syllable structures	189
8.4	Tone categories	190
8.5	Lexical tone melodies	190
11.1	Syllable initial consonants and consonant clusters	245
11.2	Codas: syllable final consonants and clusters	245
11.3	Pronouns	249
12.1	Segmental inventory	259
12.2	Pronominals	260
12.3	Verbal agreement markers	263
12.4	Jussive agreement and negation	264
12.5	Primary relativization strategies	271
13.1	Phonotactic distribution of Mongsen syllable segments	280

XX ILLUSTRATIONS

13.2 Consonant phonemes of Mongsen Ao	281
13.3 Mongsen Ao vowel phonemes	281
13.4 Mongsen Ao personal pronouns	287
13.5 Mongsen Ao possessive/oblique pronouns	287
13.6 Mongsen Ao case-marking clitics	290
13.7 Predicate structure of Mongsen Ao	292
13.8 Tense- and mood-marking morphemes of Mongsen Ao	294
13.9 Converb suffixes of Mongsen Ao	296
14.1 Syllable-initial consonants	303
14.2 Syllable-final consonants	303
14.3 Vowel phonemes	304
14.4 Personal pronouns and personal possessive prefixes	307
15.1 Proto-Tani consonant inventory	325
15.2 Proto-Tani vowel inventory	325
15.3 Tani first and second person pronouns	328
15.4 Word family based on the initial root <i>bi-</i> ‘stone’	329
15.5 Tani predicate inflections and their lexical sources	331
16.1 Chart of consonant phonemes	339
16.2 Chart of vowel phonemes	339
17.1 The ‘pairings’ of Pangwa sub-tribes according to Lamsham Khilak and Nikam Taikham Joglei	353
17.2 Comparative word list for 11 Pangwa varieties	356
17.3 Comparison of Mueshaungx and Chamchang/Shicyü orthographies	357
17.4 Lakhum Mossang’s orthography	357
17.5 Consonant phonemes in Mueshaungx	358
17.6 Vowel phonemes in Mueshaungx and Chamchang	359
17.7 Approximate pitch and phonation values for the Tangsa tones	359
17.8 Examples of unmarked tone correspondences between eight Tangsa varieties	360
17.9 Examples of marked tone correspondences between eight Tangsa varieties	361
17.10 Free pronouns and possessive prefixes in Mueshaungx and Chamchang	361
17.11 Stem alternation in eight Tangsa varieties	363
17.12 Tangsa agreement markers in various Tangsa varieties	364
18.1 Verbal paradigms in Classical Tibetan	377
19.1 The consonant phonemes of Lhasa Tibetan	386
19.2 The vowel phonemes of Lhasa Tibetan	386
20.1 Kurtöp consonant phonemes, represented in local orthography	407
20.2 Kurtöp vowel phonemes	408
20.3 Contrastive vowel length in Kurtöp	408
20.4 Contrastive (high/low) tone following sonorant consonants of initial syllables and palatal fricatives	409
20.5 Kurtöp case markers	410
20.6 Functions of auxiliary in serial verb constructions	413
20.7 Existential copulas	415
20.8 Equational copulas	415
21.1 Tshangla consonants	419
21.2 Tshangla vowels	419

21.3	Tshangla verb classes	420
21.4	Exceptional nasal/liquid-final verb roots	420
21.5	Exceptional vowel-final verb roots	421
21.6	Personal pronouns	421
21.7	Demonstratives	422
21.8	Affirmative verb paradigm	426
21.9	Negative verb paradigm	427
22.1	Affirmative paradigm of <i>nar-</i> ‘eat’	442
24.1	Initial consonants	469
26.1	The Manange consonants	517
26.2	Minimal sets	517
26.3	Oral and nasal vowels	518
26.4	Tones	519
26.5	Word formation	521
26.6	Honorific nouns and verbs	521
26.7	Personal pronouns	522
27.1	Oral vowels	535
27.2	Vowel minimal/near-minimal sets	535
27.3	Diphthongs and simple vowels	536
27.4	Consonants	537
27.5	Minimal/near-minimal sets	537
27.6	Labio-velar and palatal approximants	538
27.7	Tones	539
27.8	Onset consonant clusters	541
27.9	Consonants in final position	542
27.10	Gender suffixes	544
27.11	Honorifics	544
27.12	Honorific verbs	544
27.13	Cardinal numbers	545
27.14	Ordinal numbers	545
27.15	Personal pronouns	546
27.16	Tense-aspect	548
27.17	Copular verb forms	550
28.1	Tshobdun possessive prefixes	562
28.2	Tshobdun person-number markers on intransitive verbs	562
28.3	Basic Tshobdun orientation prefixes	564
30.1	Consonantal phonemes in Stau	598
30.2	Neutralization of voiced stops in word-initial position	598
30.3	Vowel fusion in Stau	599
30.4	Vowel alternations in open-syllable intransitive verbs in Stau	600
30.5	Irregular intransitive verbs in Stau	600
30.6	The transitive conjugation in Stau	601
30.7	Person indexation in the G.yu.rong dialect	601
30.8	Orientation prefixes in Stau	601
30.9	Denominal verbs in Stau	602
30.10	Anticausative verbs in Stau	602
30.11	Causative verbs in Stau	603
30.12	Vowel fusion in Stau nouns	604
30.13	Nominalization suffixes in Stau and Khroskyabs	609

XXII ILLUSTRATIONS

30.14	Potential lexical innovations	610
30.15	Year ordinals in Rgyalrongic languages	610
30.16	Case markers in Stau, Khroskyabs and Tangut	611
31.1	Stem 2 alternations in Japhug Rgyalrong	615
31.2	Stem 3 alternations in Japhug Rgyalrong	615
31.3	Oriental prefixes in Japhug Rgyalrong	616
31.4	Finite verb categories in Japhug Rgyalrong	616
31.5	Japhug transitive and intransitive paradigms	618
31.6	Examples of anticausative in Japhug	621
31.7	Pronouns and possessive prefixes	624
32.1	The Dulong consonants	636
32.2	The Dulong personal pronouns	637
33.1	Khaling: indicative paradigms of <i>lop</i> vt. (<i>loðmne</i>) ‘catch’ and <i>sop</i> vi. (<i>soðmne</i>) ‘have enough’	656
33.2	Selected forms of Khaling and Bahing	657
34.1	Hayu verb morphology	683
35.1	Belhare phonemes	697
35.2	Pronouns and demonstratives	699
35.3	Case markers and their use	700
35.4	Possessive prefixes	700
35.5	Finite verb inflection	702
35.6	Affirmative (upper form in each cell) and negative (lower form) non-past paradigm	704
36.1	Consonant phonemes	723
36.2	Personal pronouns, possessives and demonstratives	724
36.3	Camling person-number markers NW dialects	728
37.1	Wambule consonants	737
37.2	Wambule vowels	737
37.3	Development of Wambule implosive or preglottalised phonemes	738
37.4	Words contrasting modified and unmodified sonorants /m, n, l/	739
37.5	Possessive pronouns	742
37.6	Demonstratives	742
37.7	First intransitive conjugation (no root alternation)	744
37.8	Middle conjugation (no root alternation)	744
37.9	Third transitive conjugation: distribution of paradigmatically conditioned root forms	746
37.10	Basic morphs of the simplex person and number agreement suffixes	747
37.11	Basic morphs of the imperative person and number agreement suffixes	749
38.1	Consonant phonemes	757
38.2	Vowel phonemes	757
38.3	Personal pronouns (nominative)	761
38.4	Transitivity contrasts of disyllabic verbs with <i>-en</i> vs <i>-ja</i>	764
38.5	Transitivity contrasts of disyllabic verbs with <i>-ja</i> vs <i>-en/-f(i)</i>	764
38.6	Stem initial voicing contrasts related to transitivity	765
38.7	Subject agreement markers	765
38.8	Copula paradigm: past tense	766
38.9	Copula paradigm: present tense	766
38.10	Copula paradigm: future tense	766
38.11	Set 1 and set 2 past tense markers	767

39.1	The Qiang consonants	774
39.2	The Qiang personal pronouns	777
39.3	The Qiang person marking suffixes for intransitive verbs	781
39.4	The Qiang non-actor person marking suffixes	782
40.1	Monophthongs of Niuwozi Prinmi	792
40.2	Consonants of Niuwozi Prinmi	792
40.3	Suprasegmental categories in Niuwozi Prinmi	793
40.4	Attachment patterns of the interrogative and the negative clitics	795
40.5	The directional prefixes	795
40.6	Conjugation of the copula and verbs	796
41.1	Reconstruction of Tangut finals	807
41.2	Inventory of consonants	808
42.1	Lizu initial consonants	824
42.2	Lizu vowels	825
43.1	Initial consonants in Na	841
43.2	Allophones of the Na initials	842
43.3	Na vowels	842
43.4	Structure of the verb phrase	847
43.5	Structure of the noun phrase	848
44.1	Vowels of major syllables	860
44.2	Consonants	863
44.3	Writing open rhymes	864
44.4	Writing closed rhymes	865
44.5	Main clause particles	872
45.1	The initial consonants of Zaiwa	878
45.2	The finals of Zaiwa	878
45.3	The personal pronouns of Zaiwa	879
45.4	The demonstrative pronouns of Zaiwa	880
45.5	The imperative markers in Zaiwa	882
47.1	Consonants of Lisu	903
47.2	Vowels of Lisu	904
48.1	The seven tones of Black Lahu	919
48.2	Consonants and vowels	919
51.1	Consonants	961
51.2	Vowels	962
51.3	Initial consonant clusters	962
51.4	Personal pronouns	966
51.5	Demonstrative and interrogative pronouns	967
52.1	The Tujia onsets	975
52.2	The Tujia rhymes	975
52.3	The tones of Tujia	976
52.4	Tujia personal pronouns	978
52.5	Tujia demonstrative pronouns	978
52.6	Locative postpositions	981
53.1	Jinghpaw phoneme inventory	994
53.2	Tones in Jinghpaw	995
53.3	Jinghpaw word structures	996
53.4	Properties of noun subclasses	999
53.5	Personal pronouns	1000

XXIV ILLUSTRATIONS

53.6	Case markers	1002
53.7	Personal indices	1004
53.8	Conditions for the choice between Set I and Set II	1005
53.9	Paradigms of <i>sa</i> ‘go/come’ and <i>rāy</i> ‘cop’	1005
53.10	Allomorphs of the index <i>-ind-</i> ‘2nd’	1005
53.11	Auxiliaries	1008

PREFACE TO THE SECOND EDITION

In deciding what to put into this thoroughly reworked second edition of *The Sino-Tibetan Languages*, we were still guided by the major consideration underlying the choices behind the first edition in 2003: provide a broad overview, attained by a combination of overview articles and a wide array of articles on individual languages, with an emphasis on less commonly described languages. Many new language descriptions have been added, and the articles retained from the first volume have been updated, some, in light of changes in our knowledge, being almost completely reworked. The new articles fill gaps in our earlier coverage. The coverage of languages in Northeast India, Sichuan, and Yunnan has been expanded greatly due to the opening up of those areas to fieldwork and the large number of people, both local and non-local, now working in those areas. On the other hand, even though the diversity within the Sinitic languages is equally great, little work has been done on different Sinitic varieties from a typological perspective (most descriptions just describe the varieties relative to Mandarin), and so again we were not able to include descriptions of more than a couple of the major varieties.

In Part 1 there are three overviews of Sino-Tibetan as a whole, one on genetic and areal groupings (Graham Thurgood), one on Sino-Tibetan morphology and syntax (Randy J. LaPolla), and one on Sino-Tibetan word order typology (Matthew S. Dryer).

In Part 2, Sinitic, there are overview articles on the phonological (Zev Handel) and grammatical (Anne O. Yue) features of Sinitic languages generally, and one overview of Mandarin dialects (Dah-an Ho). There is also an article on the history and function of the unique writing system (Mark Hansell). Supplementing these overviews are in-depth articles on single varieties of Shanghaiese (Eric Zee and Liejiong Xu) and Cantonese (Robert S. Bauer and Stephen Matthews).

In Part 3, Tibeto-Burman, there are overviews of several geographical and genetic groupings, specifically, Mark W. Post and Robbins Burling discuss Northeast India, Boyd Michailovsky discusses the Kiranti languages, and Mark W. Post and Jackson T.-S. Sun discuss the Tani [Abor-Miri-Dafla] languages. Complementing the overviews are detailed descriptions of more than 41 individual Tibeto-Burman languages, in all cases by one of the world's leading experts. In addition, some 300-plus languages are mentioned in one survey or another. It is possible to quibble about the omission of this language or that, but the chapters in the volume manage to achieve a remarkable depth and considerable breadth.

Almost all the languages mentioned in the individual chapters are found within tentative subgroupings in the Thurgood overview. Alternative names for languages are listed in various ways: where the alternative name is an older designation, it is usually put into square brackets, as in the mention of "Tani [Abor-Miri-Dafla]" above. However Sino-Tibetan subgrouping is still in its infancy. It remains an area rife with controversy, but despite this, we (the editors) were, with quibbling here and there, able to agree to a large extent. While there were, of course, differences of opinion, the degree of consensus was

XXVI PREFACE TO THE SECOND EDITION

striking. Where the subgrouping was clear, we agreed on it; where it was unclear, we agreed that it was unclear.

Without exception, contributors and non-contributors alike have been supportive and helpful, providing their expertise. The individual contributors have put up with us harassing them and have helped out whenever we asked. Non-contributors have often generously given extensive feedback, adding much to the quality of the collection.

The people at Routledge, Isabelle Cheng, Camille Burns, Andrea Hartill, and last but not least Karen Greening, were generous with their help and did much to make this an even better edition than the first.

The Editors
Graham Thurgood
Randy J. LaPolla

ABBREVIATIONS

Publications

<i>BIHP</i>	<i>Bulletin of the Institute of History and Philology</i>
<i>BLS n</i>	<i>Proceedings of the nth Annual Meeting of the Berkeley Linguistic Society</i>
<i>BSOAS</i>	<i>Bulletin of the School for Oriental and African Studies</i>
<i>ICSTLL</i>	<i>International Conference on Sino-Tibetan Languages and Linguistics</i>
<i>LTBA</i>	<i>Linguistics of the Tibeto-Burman Area</i>

1	1st person
2	2nd person
3	3rd person
A	‘actor’ of transitive clause
ABL	ablative
ABS	absolutive
ACC	accusative
ADJ	adjective
ADV	adverb
AFF	affix
AGT	agentive
AMG	location in or among
ANIM	animate
ANT	antipassive
ANTC	anticausative
APPLIC	applicative
APPR	approximative
ART	article
ASP	aspect
AUG	augmentative
AUX	auxiliary
BEN	benefactive
CAUS	causative
CL	numeral classifier
CMPL	completive
COLL	collective
COMIT	comitative
COMP	complementizer
COMPAR	comparative marker
COND	conditional
CONT	continuous
CSM	change of state marker

XXVIII ABBREVIATIONS

CVB	converb
DAT	dative
DECL	declarative
DEF	definite
DEM	demonstrative
DEP	dependent
DETR	detransitivizer
DIM	diminutive
DIR	directional
DS	different subject (switch reference)
DTV	derived transitive verb
du	dual
DUR	durative
DYN	dynamic
E	extension to core
EMPH	emphasis/emphatic
ERG	ergative
ESS	essive (location at)
EVID	evidential
ex	exclusive
EXCL	exclusion particle
F	feminine
FOC	focus
FRUST	frustrative
FUT	future
GEN	genitive
GENL	general
HAB	habitual
HON	honorific
HORT	hortative
HS	hearsay
IMAG	imaginative
IMMED	immediate
IMP	imperative
IMPERF	imperfect
IMPERS	impersonal
IMPFV	imperfective
IN	location in
inc	inclusive
INDEF	indefinite
INDEP	independent
INDIC	indicative
INDTV	indirect directive
INF	infinitive
INFR	inferred
INST	instrumental
INTR	intransitive
L	local gender/derivational suffix
LAT	lative (motion towards)

LGR	Leipzig glossing rules
LINK	linker
LOC	locative
M	masculine/male
MAL	malefactive
MID	middle/middle voice
MIR	mirative (just discovered)
NEG	negation
NF	non first person actor
NGR	nasalizing grade
NOM	nominative
NOMZR	nominalizer
N-PAST	non-past affirmative
NR	near
NRPAST	near past
N-SG	nonsingular
OBJ	object case
OBLQ	oblique (non-subject) case
OPT	optative
P	‘undergoer’ of transitive clause
PART	participle
PASS	passive
PAST	past
PERF	perfect
PFV	perfective
pl	plural
PN	pronoun
POSI	positional
POSS	possessive
POT	potential
PREF	prefix
PREP	preposition
PRES	present
PROG	progressive
PROH	prohibitive
PRSNTV	presentative
PURP	purposive
Q	interrogative/question
R	co-referential
REC	recent
RECIP	reciprocal
REDUP	reduplicated
REFL	reflexive
REL	relative
REM	remote
REQU	request marker
RES	resultative
REV	reverential second person
R/M	reflexive/middle

XXX ABBREVIATIONS

RTV	root transitive verb
S	single direct argument of intransitive clause
Sa	S marked like A
Sd	S marked like dative
sg	singular
Sirr	irregular S
Sp	S marked like P
SS	same subject (switch reference)
STAT	stative
SUB	subordinative
SUBJ	subjunctive
SUPER	superessive (location on a horizontal surface)
TAM	tense-aspect-mood
TMdys	past tense marker, 1 day–1 year ago
TMhrs	past tense marker, within today
TMyrs	past tense marker, years ago
TOP	topic
TR	transitive
UNW	unwitnessed
VERT	vertical
vi	intransitive verb
VIS	visual
VN	verbal noun
vt	transitive verb
WIT	witnessed

Part 1

OVERVIEW CHAPTERS

Review Copy – Not for Redistribution

Randy J. LaPolla - Nanyang Technological University - 28/03/2017

CHAPTER ONE

SINO-TIBETAN: GENETIC AND AREAL SUBGROUPS

Graham Thurgood

1 INTRODUCTION¹

Sino-Tibetan began as a single language but under the gentle push of language internal pressures and the far more intense influence of contact with other languages, it changed, repeatedly splitting and restructuring on the way to becoming the modern Sino-Tibetan language family. Little of this linguistic history is retained in even the earliest written records, but the broad outlines of the ebb and flow in the prehistory of the Sino-Tibetan languages and the peoples who spoke various versions of it are recoverable through the techniques of comparative reconstruction. Much of the relevant work, however, remains to be done. In particular, there are critical gaps in our understanding of Sino-Tibetan subgrouping—how the original language split up over time and who the speakers came in contact with. Accurate subgrouping is needed to distinguish between splits in the phylogenetic subgroups based on shared innovations—typically having as their nonlinguistic counterpart abrupt migrations, the dialect chains—areas in which a language has spread out and subsequently differentiated into separate languages, the linguistic areas—areas with typological similarities brought about by language contact, and straight-out borrowings. Much of the Sino-Tibetan subgrouping is impressionistic or geographic, some of it presenting little or no actual supporting evidence. For the most part, this chapter provides a preliminary sketch of the subgroups for which some compelling evidence has been brought forth, however, because they are often mentioned in the literature, it also discusses some subgroupings which seem to lack any serious supporting evidence.^{2,3}

1.1 Phylogenetic ‘trees’ versus linguistic areas

For phylogenetic subgrouping, the standard assumptions about subgrouping need stating: only linguistic data constitutes evidence for a linguistic subgrouping—not geography, not ethnography, not folklore. Of course, if the resulting subgrouping is at variance with known history, for instance, either or both should be carefully re-examined. Only the shared innovations among the correspondence sets constitute evidence of an earlier period of common development; as a corollary, the value of an innovation for subgrouping varies inversely with the probability that it could have happened more than once independently—the less likely that it could have happened independently, the more valuable it is as a subgrouping tool.

The comparative method often discovers shared innovations in cases of common inheritance, that is, where after language change occurred, one of the languages undergoing the change subsequently went its own way. However, other scenarios exist. One possibility, which Sapir termed ‘drift’, occurs when independent but parallel changes occur in each language: the common starting point provided by a common origin often combines with universal tendencies to provide parallel but historically quite independent

development among genealogically related languages (Thurgood 1985: 378; LaPolla 1994). A second possibility is that the language patterns are not the product of abrupt migrations, but instead reflect a language spreading out in an area before breaking up into dialects and then distinct languages. In neither of the last two cases should one expect to find shared patterns of innovations that break up the languages into non-overlapping sub-groups. Instead of an innovations-based tree structure, the data shows patterns of dialect continua which have broken up into new languages (Ross 1988, 1997). In a dialect continuum, adjacent varieties often share intersecting patterns of shared innovations with their neighbors, sometimes because the innovation was there before the two varieties diverged, sometimes because of contact postdating the diversification.

A theoretical concern is whether relatedness can be established without extensive morphology. Jacques (forthcoming) invokes this with reference to the paucity of morphology in Chinese, citing Meillet (1982 [1914]: 97), who wrote that languages lacking extensive morphology and depending on word order present a problem for determining genetic relatedness because it is difficult to demonstrate that lexical similarities are not simply the result of borrowing. Such paradigmatic richness, even if it may have once, no longer exists in Chinese, but elsewhere one finds complex verbal systems. Within Tibeto-Burman, even some of the areas with reduced morphology and a dependence on word order for case marking may not be a problem. Nichols (1996: 48) notes that the real necessity is not complex morphological paradigms *per se* but “whole systems or subsystems with a good deal of internal paradigmaticity, and involving not only categories but particular shared markers for them.” Directly relevant to Tibeto-Burman, Nichols (1996: 64) notes that complex tone systems fall into this category. Once the relevant reconstructions have been worked out, rich tone systems divide the whole inherited lexicon into what are from a synchronic perspective arbitrary sets of words, making group membership relatively easy to establish. Lolo-Burmese illustrates this clearly.

1.2 Contact with other languages

Language contact is pervasive throughout the Sino-Tibetan region, an area where multilingualism is the norm and language shift is common, but the task of recognizing its influence varies in difficulty depending on the circumstances. It is most transparent when the influence comes from known, unrelated or at most distantly related languages such as Chinese, Tai, Mongolic, or Mon-Khmer. The superstrate Chinese influence on Tibeto-Burman is often obvious (see section 3.1 ‘Contact influences on Chinese’). The substrate influence of Mon-Khmer speakers, who were once prevalent in many of the areas that Tibeto-Burman speakers and some Chinese speakers now occupy, is frequently noticed as is the contemporary Chinese influence on the Tai [Thai] languages in southern China. More difficult are closely related languages such as Burmese and Jingphaw, with borrowing between various Tibetan dialects being even more difficult to separate out. Finally, much, much more difficult to recognize are influences from unknown substrate languages.

1.3 Lingua franca status and restructuring

The use of a variety or a language as a lingua franca, if sufficiently heavy, results in restructuring, often of a simplificatory nature. This is unsurprising as the typically adult speakers who use it as a *langua franca* seldom possess full mastery, nor is full mastery possessed by those who shift to the *lingua franca*. From this perspective, while it is

recognized that Mandarin has been influenced by the Tungus, Mongol, and Manchu languages in the north and the Mon-Khmer and Kra-Tai (Tai-Kadai) languages in the south (Hashimoto 1986), it is not as widely realized that the massive shifts of non-native speakers to Mandarin along with the widespread use of Mandarin as a lingua franca have also played a significant role. Under such influences Mandarin has undergone more restructuring, much of it simplificatory, than have non-Mandarin varieties. Similarly within Tibeto-Burman, variants of languages which served as lingua francas have undergone at least partial simplification under the influence of adult learners shifting, as well as others using it as a lingua franca. If nothing else this makes reliance on just the most widespread and often the most prestigious variant of a language problematic for the reconstruction of language history (see also section 9).

1.4 Language names

No attempt to straighten out all the names has been undertaken here. See the relevant general discussions in Post and Burling, in Post and Sun, and in Michailovsky, as well as more specific languages and subgroups; for still further information, see Matisoff *et al.* (1996). Some groups have multiple names for themselves. Here, we have simply given a common name, sometimes with another possibility added after a slash and sometimes with an older name given in square brackets.

1.5 The genealogical subgroups

Not all Tibeto-Burman languages can be readily subgrouped: Meithei, Mru, Karbi [Mikir], Tujia, and Bai remain partial mysteries, although proposals have been made. The various subgroups differ radically in the strength of their supporting evidence. Some, like Lolo-Burmese, not only have a rich database but are also substantiated by lexical reconstruction. Some like Puroik-Sherdukpen-Sartang have neither, but nonetheless seem to be a valid group. Finally, there are proposed subgroups which lack both a significant database and any extensive reconstruction, although some of these are supported by the presence of marked shared innovations and impressionistic reconstructions. This chapter lists subgroups which have some degree of support, largely ignoring those which lack sufficient data to pass a judgment. Thus, higher-level groupings like ‘Baric’ and languages like Pyu have been omitted, on the one hand because the lower-level foundations are lacking, on the other because the necessary linguistic database is lacking.⁴

2 SINO-TIBETAN

The Sino-Tibetan languages are named after the two most salient members, the Chinese languages and the Tibetan languages, with the term Tibeto-Burman traditionally used to refer to the non-Chinese subset of these languages. The vast majority of specialists agree that the relationship between Chinese and the Tibeto-Burman languages is genealogical, with disagreement over the status of the Chinese component: some view the Chinese component as a sister to Tibeto-Burman; others view the Chinese component as one of the subgroups of Tibeto-Burman languages (see Figure 1.1); and, in light of the lack of much convincing linguistic data, many of us remain agnostic.

In this chapter, the label Sino-Tibetan refers to all these languages; Tibeto-Burman is used as a convenient way to refer to all these languages except Chinese (Sinitic), but without committing to a subgrouping scheme. Both configurations face the same

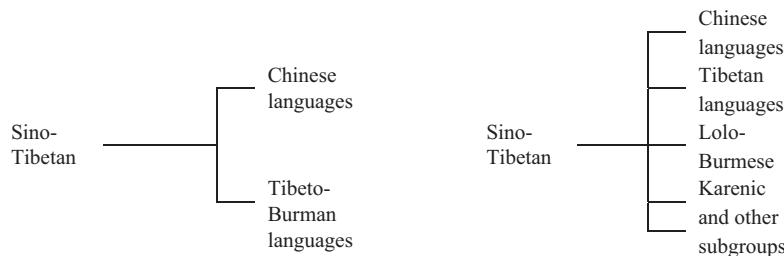


FIGURE 1.1 TIBETO-BURMAN VERSUS CHINESE: TWO VIEWS

questions: how does Chinese subgroup with other Sino-Tibetan languages, and how do the remaining languages subgroup with one another (see Handel 2008; Sün 1988)?

2.1 Wider relationships

Wider relationships linking Sino-Tibetan to other languages have been proposed. Most frequently suggested is a relationship to the Tai languages [[Chinese: Zhuang-Dong]; Kra-Dai [Tai-Kadai]] and to Hmong-Mien [Chinese: Miao-Yao]. Both positions have since been rejected by most Western scholars, but are still widely held among Chinese linguists. Another view is proposed by Sagart (2005a, 2005b), who has argued for Sino-Tibetan-Austronesian, with the Tai-Kadai languages a daughter to the Austronesian family. More far-reaching are Sapir (1920) and Shafer (1957, 1969), who both compared Sino-Tibetan with Na Dene (Athapaskan; Athabaskan, Eyak, Tlingit); more recently Vajda (2010) established a Dene connection, not with Sino-Tibetan, but with Yeniseian. A considerable number of other proposals have been put forth but most such proposals lack convincing data: it is difficult if not impossible when dealing with a very small number of forms to rule out chance relationships and ancient contact, let alone ancient genealogical relationships.

3 CHINESE

Chinese is not a single monolithic language, but rather a family of related languages, each with its own dialects often distinct enough to make even dialects within the same dialect group mutually unintelligible. Rather than linguistically distinct groups, the dialect families seem to reflect dialect networks with different nuclei. The grouping into six dialect families matches Norman's (1988) practice, aside from the placement of Hakka (see Figure 1.2). Alternate groupings are common in the literature (see Handel, this volume).

Each of these dialect groups is further subdivided into subgroups, but the larger groupings are sufficient to make the incredible diversity of Chinese dialects apparent to the thoughtful reader.

The national standard language, Putonghua, is based on the phonology of the dialect of Beijing but the lexicon and grammar are based on the general northern vernacular. Although in this work Chinese may also refer to the collection of Chinese dialects, to any one of the various Chinese languages, or to the national standard, it typically refers to the national language with more specific references used to refer to the dialects or to a specific language.

Northern: Mandarin supergroup: Northeastern dialects Zhongyuan dialects Beijing dialects Lanyin dialects Jilu dialects (Beifang) Southwestern dialects Jiaoliao dialects Jinghuai dialects	Central: Wu dialect family Shanghaiese Xiang dialect family Gan dialect family Hakka dialect group Southern: Yue dialect family Cantonese Min dialect family Hokkien, Taiwanese
---	---

FIGURE 1.2 CHINESE ‘DIALECTS’ (LANGUAGES)

In the 2003 volume there were questions about the status of Tangut [Hsihsia] and Bai. Tangut is now recognized as unquestionably Tibeto-Burman, and, most likely close to Pumi or rGyalrong (Jacques 2012a). The situation concerning Bai is more controversial, more for political than linguistic reasons, but the data on Bai increasingly suggests it is a heavily Sinicized Tibeto-Burman language, otherwise as yet ungrouped. For Bai, a Lolo-Burmese connection is often suggested but the evidence is lacking.

3.1 Contact influences on Chinese

Throughout its history, Chinese has been prone to restructuring under the influence of intense language contact (LaPolla 2001, 2010). In fact, even the basic SVO (Subject–Verb–Object) word order of Modern Chinese seems to be the product of contact (cf. LaPolla 2015). Norman (1988) and Hashimoto (1986) correlated various phonological and syntactic characteristics of the Chinese dialects with the different language contact patterns: the more northerly the Chinese languages are, the more they resemble the non-Sino-Tibetan languages of the north, e.g. Tungus, Mongol, Manchu; the more southerly they are, the more they resemble the non-Sino-Tibetan languages of the south, e.g. Thai of the Tai-Kadai and various Mon-Khmer languages. In his contribution, Dryer documents and extends this analysis of contact, describing clear, typological correlates.

The earliest recoverable Chinese vocabulary already has borrowings from other languages. The words for ‘honey’ and ‘goose,’ found in Chinese but not in Tibeto-Burman, for instance, probably reflect early Indo-European contact of a limited nature. Norman (1988) notes borrowings from Mon-Khmer including ‘tiger,’ ‘ivory,’ ‘crossbow,’ and the word for ‘river,’ found in the Chinese name for the Yangtze (*Jiāng* < *krong), undoubtedly indicating not just contact but also a significant early Austroasiatic (Mon-Khmer) presence throughout that area. Various calendric terms also look to have been borrowed from Austroasiatic, indicating a significant cultural influence on the early Chinese.

Some languages coming into contact with Chinese must have been absorbed with scarcely a trace, but others appear in the historical records. In the north, Chinese came

8 GRAHAM THURGOOD

under intensive, long-term contact with speakers of other languages. For instance, the Sixteen Kingdoms period (roughly 303–439) refers to numerous non-Chinese dynasties that ruled at least parts of northern China at various times. These kingdoms included speakers of what have been termed Altaic languages (Tungusic, Mongolian, and so on), of Tibeto-Burman languages, and of other languages, and began the process of the Sinicization of these languages as these languages were influencing Chinese.

In the north, the influences on Chinese have been Tungusic and Mongol, a presence that stands out in the historical records. The Liao dynasty (916–1125) was a kingdom below the Great Wall that extended from Mongolia into southern Manchuria. Its capital Khitan (Khitai) was the source for the word Cathay, used in medieval Europe to refer to northern China. The Liao established their southern capital in what is now modern-day Beijing. Overlapping with the Liao dynasty was the Jin dynasty (1115–1234), another non-Chinese speaking people who originated in Manchuria and who were the ancestors of the Manchus. Just 30 years after the end of the Jin dynasty came a Mongolian dynasty, the Yuan dynasty (1264–1368) established by Genghis Khan and his successors. Finally, the Manchus, who established the Qing dynasty (1644–1911), the last of the Chinese dynasties, in 1644, spoke Manchu, a Manchu-Tungusic language still spoken by pockets of speakers here and there in parts of northeast and northwest China.

In the south, particularly south of the Yangtze, the influences are not as obvious in the historical record but the linguistic influences may have been as significant. Certainly, there is clear evidence in Chinese of contact with Austroasiatic speakers, Tai-Kadai (Thai) speakers, Hmong-Mien speakers, and Austronesian speakers. Most of the evidence of an Austroasiatic presence is found in lexical borrowings into both Chinese and into Tai-Kadai languages. Speakers of Tai-Kadai languages are still present in significant numbers in southern China; the absorption of many of these speakers is reflected both in borrowings and in structural realignments in southern Chinese dialects. The interaction with both Hmong-Mien and Austronesian is reflected most obviously in borrowings.

Pockets of most of these peoples and their languages can still be found scattered in parts of China, although their linguistic influence has diminished to the point that many of the smaller groups are now in danger of totally disappearing; the absorption of many, many speakers of these languages into various Chinese dialects over a long period of time has had a significant influence on Chinese structures.

4 LOLO-BURMESE BRANCH [BURMESE-LOLO]

Lolo-Burmese (= Burmese-Lolo) constitutes a well-established subgroup with a system of complex, shared innovations involving initials, tones, and rhymes (Burling 1967; Matisoff 1972; Bradley 1979; Thurgood 1974, 1977, 1982, and various other scholars). It is divided into Burmish and Loloish (Figure 1.3). Burmese has several distinct dialects in addition to the Yangon dialect: Arakanese, Tavoyan, Palaw, Merguese, Yaw, Intha, Taungyo, Danu.

To identify a language as Lolo-Burmese it is usually sufficient to establish regular correspondences for Proto-Lolo-Burmese (PLB) tone 3, however, Table 1.1 also presents the reflexes of the other two proto-tones. Note that the Zauzou reflexes for *1 and *2 are obscured by tone sandhi.

Languages which share the complex three-way patterns of tonal reflexes in checked syllables from the interaction of various PLB prefixes and initials in checked syllables are Loloish. Tables 1.1 and 1.2 give my versions of the PLB correspondences, with Table 1.2 dividing all the checked syllables into three classes (for this, see Matisoff

Burmish	Loloish		
Burmese:	Northern	Central	Southern
Yangon	Nusu	Sani-Ahi	Akha
Arakanese	Nasu	Lahu	Hani
Tavoyan	Nosu	Lisu	Haoni
Yaw	Nisu	Lipho	Mpi
Zaiwa (Atsi)	Luquan	Jino, Jinuo	Khatu
Maru		Zauzou	Bisu:
Bola			Phunoi
Achang			Bisu
Lachi			Pyen

FIGURE 1.3 LOLO-BURMESE

TABLE 1.1 REFLEXES OF THE THREE NON-CHECKED PLB TONES

Proto-tones	Initials class	Written Burmese	Zauzou	Lahu	Lisu	Akha	Jino
*1	*voiced	low,	tone sandhi	21	33	55	42
	*voiceless	level,		33			
	*spirantal prefix	normal			33c		
*2	*voiced	fairly high,	tone sandhi	54	33	11	44
	*voiceless	sharp fall,					
	*spirantal	breathy		11			
	*spirantal prefix				55		
*3	*all initials	high, slight fall, creaky	35	33	33c	33	33

1972, Thurgood 1977, 1982 in particular). Checked syllables not following these patterns are borrowings.

Early Mon-Khmer influence on Proto-Lolo-Burmese is evident in the number and the basic character of the Mon-Khmer borrowings (e.g. the word for ‘eat’), mostly from Mon. The earliest Burmese inscriptions were in Mon script, with subsequent inscriptions in the developing Burmese script as well as in Pali (an Indic language), in Mon (a Mon-Khmer language (the largest subgroup of Austroasiatic)), and occasionally even in Pyu (an extinct Tibeto-Burman language). All Lolo-Burmese languages have been subject to Mon-Khmer influence. In Loloish there is also Chinese influence, and Loloish languages now located in Thailand contain numerous recent Thai loanwords.

Wider relationships: It has been claimed that Lolo-Burmese itself is part of a slightly larger subgroup, whose members do not have these initial or tonal reflexes, but, if so, the bases for a wider subgroup remain to be established.

5 BODISH: TIBETAN, EAST BODISH, TSHANGLA, AND PROBABLY TAMANGIC

In this survey Bodish is used as a label for the Tibetan languages; the East Bodish languages: Dakpa [Takpa]-Dzala, Bumthang, and Kurtöp; the Tshangla dialects; and probably but more distantly the Tamangic languages. Scholars consistently suggest that these language groups form a higher-level group.

TABLE 1.2 LOLOISH CHECKED TONES

PLB initials classes	Sani	Ahi	Nasu	Luquan	Jino	Lisu	Lahu	Akha	Sangkok
*s-bak	ph- 55	ph- 55	ph- 55	ph- 55c	ph- 55	p- 55	p- 35	p- 11c	ph- 31c
*s-mak	m- 55	m- 55	m- 55	m- 55c	m̥- 55	m- 55	m- 35	m- 11c	m- 31c
*pak	ph- 44	ph- 44	ph- 32s	ph- 22s	ph- 42	ph- 35c	ph- 54s	p- 33c	ph- 33c
*sak	sz- 44	s- 44	s- 32s	s- 22s	ts- 42	s- 35c	š- 54s	s- 33c	s- 33c
*k-rak	h- 44	h- 44	x- 32s	h- 22s	x- 42	h'- 35c	y- 54s	x- 33c	x- 33c
*s-pak	p- 44	p- 44	p- 32s	p- 22s	p- 42	p- 35c	p- 54s	p- 33c	ph- 33c
*(s)-mak	m- 44	m- 44	m- 32s	m- 22s	m̥- 42	m- 44c	m- 54s	m- 33c	m- 33c
*m-pak	b- 44	b- 44	b- 32s	b- 22s	p- 42	b- 44c	b- 54s	b- 33c	p- 33c
*ryak	h- 22s	x- 44s	x- 34	?h- 55c	x- 55	h- 21?	h- 35	y- 11c	x- 31c
*C-sak	s- 22s	s- 44s	s- 55	s- 55c	s- 55	s- 21?	š- 35	s- 11c	s- 31c
*C-pak	ph- 22s	ph- 44s	ph- 55	ph- 55c	ph- 55	ph- 21?	ph- 21s	p- 11c	ph- 31c
*rak	y- 22s	j- 44s	y- 55	y- 55c	r- 55	y/w- 21?	y- 21s	y- 11c	*r- 31c
*zak	z- 22s	z- 44s	dz- 55	z- 55c	z- 55	z- 21?	y- 21s	y- 11c	*z- 31c
*bak	b- 22s	b- 44s	b- 55	b- 55c	p- 55	b- 21?	p- 21s	b- 11c	p- 31c
*mak	m- 22s	m- 44s	m- 55	m- 55c	m- 55	m- 21?	m- 21s	m- 11c	mb- 31c

Note: b- = voiced stops; p- = unaspirated stops; ph- = aspirated stops; m- = nasals.

Wider affiliations: At one time or another, Western Himalayan, Kham-Magar, and Kiranti have all been suggested for this subgroup, but shared innovations between these groups have not been demonstrated.

5.1 The Tibetan subgroup

Proto-Tibetan is used to refer to the oldest stage, unattested and lacking written records; our knowledge of this stage comes largely from the application of the comparative method. Here the term Tibetan is restricted to languages directly descended from Proto-Tibetan; Tournadre (2014) uses “Tibetic” for these languages. Old Tibetan refers to the earliest written Tibetan; sometimes the term Written Tibetan is restricted in the same way, but this simply invites conflation of all stages of written Tibetan into one. When used carefully, Classical (Literary) Tibetan refers to the language of most documents written after the ninth-century language reforms (DeLancey, this volume), but at times is used as if it were a synonym for any form of written Tibetan. It needs to be emphasized that Written Tibetan, even Old Tibetan, meaning the earliest form of written Tibetan, is not equivalent to Proto-Tibetan. The potential for misanalysis is further compounded by the fact that the two most commonly used Tibetan–English dictionaries, Jäschke (1881) and Das (1902), are panchronic, containing words and spellings from various times and sources, many of which are not found in older records.

Thus, Tibetan languages are distinguished from Tibetan-influenced languages. The regular sound–meaning correspondences between the various Tibetan languages allow a historical linguist to separate out the Tibetan dialects from Tibetan-influenced languages. Proto-Tibetan itself also requires working out the structure of the older forms plus a set of rules connecting Proto-Tibetan to the modern languages. Note that the cumbersome phrase ‘regular sound–meaning correspondences’ is not equivalent to ‘similarities’, but rather the regularity of the correspondence sets. The problem with ‘similarities’ is distinguishing the borrowed from the inherited.

From an external viewpoint, still other evidence marks a language as Tibetan (Tibetic): the presence of a reflex of the innovated Tibetan word for ‘seven’ *bdun*, the presence of the related but obviously borrowed word for ‘seven’ in Japhug *βdumpa* ‘seventh’, notwithstanding (Jacques 2004), and two innovated pronouns. Without exception, all Tibetan languages share the innovation of a second person pronoun **khyot*, **khyet* ‘thou’, the innovation of a third person singular, roughly **kho*.

Internally, the modern Tibetan languages are the residue of widespread dialect networks. The innovations diffuse in intersecting patterns or linkages (Ross 1988: 8, 1997), which suggests the Tibetan speakers entered the area, spread out, and subsequently differentiated in different ways depending upon the patterns of contact. DeLancey (this volume) breaks them up into at least four linguistically established nuclei: Amdo, Khams, Central or Ü-Tsang (*dbus-gtsang*) including Lhasa, and Western. Using both linguistic and non-linguistic evidence, both Nishi (1986) and Tournadre (2014) add other branches to these four. Nishi (1986) has six major branches: Central (or Ü-Tsang), Western Innovative, Western Archaic, Southern, Khams, and Amdo, and is much like the scheme in Bielmeier’s (forthcoming) *Comparative Dictionary of Tibetan Dialects*, although Bielmeier’s splits Khams into two segments. Tournadre’s scheme (2014: 120) presented in Table 1.3 has eight, but is neither a shared-innovations based scheme nor a purely linguistically based scheme (nor is it intended to be); it includes “geographical parameters, migration and language contact factors,” making it, like most Tibetan classifications, a hybrid system. Other recently described Tibetan varieties seem to fall outside of all these schemes (Tournadre 2014; Sun 2014) and still other suggested schemes exist. Tournadre’s scheme (2014: 120–1) is presented with its eight nuclei in Table 1.3:

Wider affiliations: With some frequency various authors have proposed subgrouping the Tibetan languages with East Bodish, Tamangic, and Tshangla. The notion is appealing but convincing evidence has yet to be put forth. The innovated pronouns **khyot*, **khyet* ‘thou’ and **kho* ‘he/she’ which characterize the Tibetan languages are not obviously attested in East Bodish, Tamangic,⁵ nor Tshangla.

5.2 Tamang-Gurung-Thakali-Manange languages (Tamangic)

Tamangic: (according to Noonan 2011)

Tamangic complex:

 Tamang

Gurungic:

 Manange-Nar-Phu complex

 Gurung

Thakali complex:

 Thakali, Chantyal, Seke

Mazaudon’s (1977, 1978, 2005) reconstruction of a three manner–two tone system for Proto-Tamangic [Tamang-Gurung-Thakali-Manange [TGTM; Gurung] subgroup establishes the core Tamangic languages as a subgroup. It is not Tibetan; trivially, it lacks Tibetan *bdun* ‘seven’ and the innovated second and third person pronouns found in Tibetan. Internally, the subgrouping is less clear (Noonan 2011).

Wider relationships: Tamangic membership has been suggested for Ghale (Paudel 2008) and Kaire (Honda 2008) but neither language shares the reconstructed tone

TABLE 1.3 TOURNADRE'S (2014: 120–1) EIGHT NUCLEI

Central	Ü Kad Kongpo	Tsang Lhokha Tö	Lhasa Phenpo Shigatse
Southwestern	Sherpa Kagate Dolpo Langtang Gyalsumdo	Jirel Humla Nubri Kyirong Walung	Lhomu Mugu Tsum Yolmo Tokpe Gola
Northwestern	Balti Zanskari	Ladakhi NW	Purik, Purki
Western	Lahul Jad	Spiti Garzha	Tod Khunu
Southeastern	Khams Yushu Minyak Chaktreng	Hor Nagchu Pembar Dzayul Muli-Dappa	Hor Bachen Rongdrak Derong-Jol Semkyi Nyida
Southern (Sikkim, Bhutan)	Dzongkha Dhromo Lakha Dur Brokkat	Tsamang Drengjong, Lhoke Mera Sakteng Brokpa-ke (Ladkhi)	
Northeastern	Amdo Sikkimese	gSerpa	Khalong
Eastern	Chone Khöpokhok Thewo	Baima Palkyi [Pashi] Zhongu	Drugchu Sharkhok

system. Hence, at most they might constitute a sister to Tamangic, but strong evidence is lacking and alternatives exist: Noonan (2008a, 2008b, 2008c, 2008d), for example, puts Kaire together with Kham, Magar, and Raji.

5.3 East Bodish

The oldest description of an East Bodish language is Hodgson's (1853) description of Dakpa⁶ [Dwags], which Hodgson thought was a Tibetan dialect, a misconception corrected by Shafer (1955). There are two descriptions of the Mama dialect, termed Cuònà Ménbà, one a wordlist in Sün *et al.* (1991) and the other the Lu (1986) dialect; despite the fact that the low tone is 13 in Sün *et al.* (1991) but 35 in Lu, these are probably the same dialect by the same author. Lu's Wenliang dialect (also in Lu 1986) and van Driem's Dakpa (1997) are two other Dakpa dialects. Van Driem (1997) puts Lu's Wenliang dialect together with his own Dzala, terming them both Dzala; the two appear to be dialectal variants of each other, rather sister languages. Two other languages for which we have good descriptions

are Bumthap, a Bumthang language (van Driem 1997, 2015), and Kurtöp (Hyslop 2011), with Hyslop (2011) being a comprehensive grammar with rich diachronic insights.

Preliminary reconstruction establishes the East Bodish languages as a well-substantiated subgroup. For several varieties of Dakpa as well as for Bumthap and Kurtöp, the languages with the richest databases, many of the basic regular sound-meaning correspondences have already been worked out. Much of this was laid out in Michailovsky and Mazaudon (1994); much more is laid out, often quite explicitly, in Hyslop (2009, 2011, 2013).

More evidence is provided by tonogenesis, the earliest stage of which dates back to Proto-East Bodish. An examination of Table 1.4 shows Written Tibetan compared to the East Bodish languages Dakpa, Dzala, Bumthap, and Kurtöp. The earliest stage, still retained in all the richly documented East Bodish languages, evolved from forms with *s-* before a nasal, providing a modern high tone register throughout East Bodish.

Finally, certain intersecting shared correspondences can be seen in the innovated second and third person pronouns and in the shared innovation of **l-* > *y*. The innovated third person singular pronouns *pé*, *bé* and the innovated second person singular *?i* and so on groups all the Dakpa dialects together, while the **l* > *y/j* groups Kheng, Bumthap, and Kurtöp as well as Chalikha together. Despite having limited data, the data in two of the remaining languages shows intersecting shared innovations that show a linking pattern left over from a former dialect network (Table 1.5): Chalika groups with the Dakpa data with reference to the innovated pronouns, but with the Bumthap and Kurtöp cluster with **l* > *y/j*; Phobjip, a ‘Nyenka dialect, clusters with Dakpa except for the second person *khi*. For the other ‘Nyenka dialects possibly including Mangdep, the data is too limited to conclude much.

Following Ross (1997), we term such configurations linkages, further dividing them into dialect chains and networks. In dialect chains the innovations are linked from one language to another in some sort of a line; prototypically a chain is found around the shore of an island. Network chains involve languages in various directions, as one might expect of a group established inland. East Bodish appears to reflect a network (or linkage; see Figure 1.4); following Ross the double line indicates a linkage (were my knowledge more complete this representation would be extended to other parts of this chapter).

TABLE 1.4 WRITTEN TIBETAN S+ NASAL COMPARED WITH EAST BODISH TONES¹

	Written Tibetan	Dakpa in Sün <i>et al.</i> (1991)	Dzala van Driem (1997)	Bumthap van Driem (2015)	Kurtöp Hyslop (2011)
H tone		/H/	/H/	/H/	/H/
green	sñon-po	^h ŋAU ⁵³ po ⁵³	^h ŋäu	—	^h ŋúnti
heart	sñiñ	^h nij ⁵³	^h néñ	nengma -i	^h neng
medicine	sman	^h man ⁵³	—	—	^h mán
nose	sna	^h nA ⁵³	^h nă	^h naphang	^h ná
oil	snum	^h num ⁵³	—	—	—
L tone		/L/	/L/	/L/	/L/
Isg; I	ŋa	¹ ŋe ³⁵	¹ ŋe	¹ nat (A)	¹ ngat (A)
barley	nas	¹ nA? ³⁵	—	[¹ nat]	¹ na?:?
cry; weep	ŋu	¹ ŋu ³⁵	—	—	¹ ŋò
inside; interior	naŋ	¹ nen ³⁵	—	¹ na ¹ naŋ	¹ naŋ
name	miŋ	¹ men ³⁵	¹ men	¹ men	¹ men

1 The Dakpa is from Lu (1986); most of Lu’s pitch labels converted rather readily into disyllabic entities with high register and low register. Other labels have been standardized to facilitate clarity of exposition.

TABLE 1.5 LINKING SHARED INNOVATIONS IN EAST BODISH

	Dakpa Hodgson (1853)	Dakpa Lu (1986)	Dakpa van Driem (1997)	Dzala van Driem (1997)	Phobjip Hyslop (2011)	Chalika Hyslop (2011)	Kheng Hyslop (2011)	Bumthap van Driem (2015)	Kurtöp Hyslop (2011)
1sg	gné, nyé	ŋe ³⁵	ŋe	ŋa	jat	jat; ja	jat (A); jaí (E)	ngat (A); ngai (E)	
2sg	i'	i ³⁵	i	yi	i	we	wet (A); wi (E)	wit (A), wi (E)	
3sg	pé, bé	pe ³⁵	bi	khi	be	gon	khit (A); khi (E); gon (E)	khit (A); kħi (E)	
*l> j	l-	l-	l-	l-	j-	j-	y-	y-	j-

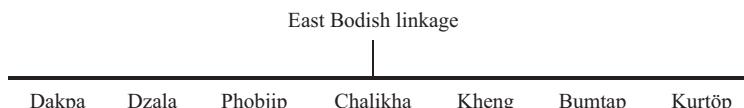


FIGURE 1.4 THE EAST BODISH LINKAGE

Wider affiliations: Although East Bodish is part of a larger group that includes the Tibetan dialects [Tibetic varieties], it is not itself a Tibetan dialect; it has neither the *bdun* ‘seven’ nor the innovated second and third pronoun characteristic of the Tibetan dialects, nor does it share the innovated second person *khyot, *khyet ‘thou’ or the third person *kho ‘he/she.’ Burling (2003) suggests that East Bodish and Tshangla are closer to each other than either is to Tibetan, but Hyslop (2014), for instance, suggests Tshangla is quite distinct.

Other members of this subgroup: It is now clear that two languages once tentatively considered East Bodish are not East Bodish: Sherdukpen (Bradley, personal communication) and Black Mountain Mönpa (Hyslop 2011, and elsewhere; van Driem 2011).

5.4 Tshangla

Tshangla is known as Sharchopka (‘the eastern language’) in Dzongkha, the national language of Bhutan. Andvik describes Bhutanese Tshangla (Andvik 2010, this volume), Cangluo Monpa (= Motuo Monpa, Sün *et al.* 1980; Zhāng 1986; Sün *et al.* 1991), and Central Monpa (Das Gupta 1968) as closely related dialects of the same language, something an examination of comparative wordlists readily confirms.

Not members of Tshangla: Lhokpu and Gongduk are not in the Tshangla branch. An overview of the Gongduk data presented in van Driem (2001: 463–8) and the Lhokpu data in Sharma (2005: 232–8) make it clear that neither language is obviously close to other Tibeto-Burman languages. For example, Gongduk *dayli* ‘water’, *tah* ‘meat’, *dij* ‘wood; firewood’, *rn* ‘tooth’, *um* ‘face’ have no known cognates.

Wider relationships: Hard evidence for Tshangla being Bodish is not available, although the lexicon suggests Tibetan to many analysts. It is certainly not Tibetan; it has neither the *bdun* ‘seven’ nor the innovated *khyot, *khyet ‘thou’ and *kho ‘he/she.’ Nor is Tshangla obviously East Bodish; Hyslop (personal communication) suggests the grammar is quite different.

6 QIANGIC LANGUAGES

On the basis of similarities found among the Qiangic, Pumi, rGyalrongic, and Ersuic as well as certain other languages of the West Sichuan Ethnic Corridor, Sün (1990 [1983] and elsewhere) has labeled these languages a subgroup, which he terms Qiangic. However, Sün’s supporting evidence includes not just similarities due to genealogical inheritance but also similarities due to areal language contact. Chirkova (2012, 2014) argues, “the similarities among certain proposed ‘Qiangic’ languages may be at least as likely to reflect convergence due to language contact as it is due to linguistic genealogy.” Certainly, both influences coexist, but in some cases the genealogical and the areal can be teased apart.

The systems of directional prefixes in some languages (see Shirai 2009; Sün 1981a) appear to be not only genealogically inherited, but to also suggest a tree structure (implicit in Table 1.6).⁸ For the directional prefixes, the broadest group of related languages is Qiang, Pumi, Ersuic, and rGyalrong, which are in turn distinguished from each other by

16 GRAHAM THURGOOD

shared innovations not found in the other subgroups: At the highest level, Qiang, Pumi, Ersuic, Muya, and rGyalrong all share the directional prefix for ‘up,’ an innovation which distinguishes them from non-Qiangic languages. The Qiangic languages are divided into Qiang and Pumi-Muya-Ersuic-rGyalrongic by the latter’s shared innovations of ‘down’ and ‘away.’ This group is divided into rGyalrongic and Pumi-Muya-Ersuic by the latter’s shared innovations of ‘inward, upstream,’ and this group is divided into Ersuic and Pumi-Muya by the latter’s innovation of ‘outwards, downstream.’ Finally Pumi subgroups with Muya rather strikingly.

As noted, each of the lowest-level subgroups has at least one set of innovated prefixes not shared with any of the other four low-level subgroups: Qiang has a unique innovation for ‘down’ and for ‘outwards, downstream,’ Pumi has a unique innovation for ‘toward center,’ Ersuic has a unique innovation for ‘outwards, downstream’ and rGyalrong has a unique innovations for ‘inwards, upstream.’

Notice that, while the evidence for Pumi, Muya, and Proto-Ersuic is intriguing, the evidence for the remaining candidates is weaker. That is, Tangut and Qiang may ultimately be grouped elsewhere, and the placement of Queya and nDraba [Zhaba] is tentative. Finally, other languages within the Qiangic linguistic area have directional verb prefixes with no obvious genetic connection to the genealogically related prefixes in Table 1.6; these will be dealt with in discussions of individual low-level subgroups.

TABLE 1.6 DIRECTIONAL VERB PREFIXES

	‘up’	‘down’	‘away’	‘inward, upstream’	‘outwards, downstream’	‘toward center’
Qiang:						
Mawo	tə-	a-			sə-	
Ronghong (Yadu)	tə-	fia-			sə-	
Longshi	tə-/tə	fiə-			sɪ-	
Taoping	tə ⁵⁵ -				s ³¹ -	
Mianchi	tə -	fiə-			sɛ-	
Pumi:						
Qinghua	tə ⁵⁵ -	nə ¹³ -	thə ¹³ -	kha ¹³ -	xə ¹³	də ¹³ -
Taoba	tə ⁵⁵ -	nə ³⁵ -	thə ³⁵ -	kha ³⁵ -	xə ³⁵ -	də ³⁵ -
Prinmi Niuwozi	tɔ-	nɔ-	thɔ-	(x)ɔ-	gə-/kha-	dʒ-/də-
Muya	tuu ⁵⁵ -	nuu ⁵⁵ -	thuu ^{35/55}	khi ³⁵ -	fiɛ ³⁵ -	
Proto-Ersuic	*de-	*ne-	*tʰe-	*kʰe-	*ŋe-	
Kala Lizu	de-	ne-	the-	khə-		
Mianning Lizu	de-	ne-		khə-		
Kala Lizu	de-	ne-	the-	khe-	ŋe-	
Naiqu Lizu	də-	nə-	thə-	khə-	ŋɛ-	
Qingshui	də-	nɛ-	(the-)	khe-	ŋɛ-	
Zeluo Ersu	də-	nc-		khc-	ŋɛ-	
Tangut [Xixia]		nja ¹ -		kha ¹ -		
Queya, nDraba ?	Λ-	a-			ŋΛ-	
rGyalrong:						
Stau [Ergong]	rə-	nə-				
Tshobdun [Caodeng]	tə-	nə-				
Japhug rGyalrong	tʂ-	pjəu-; nuu-	thə-	lə-		
ICogtse rGyalrong	to-/ta	no/na	tʰw- cʰw-	lʂ-		
				ko/ka		

Constructions with formally similar structures but with unrelated forms are evidence for language contact (Shirai 2009). LaPolla (2003: 30) in a parallel way notes that case markers and existential verbs often have similar patterns but largely unrelated forms (LaPolla 1994, 2003: 30). Various authors are worried about low cognacy rates, but low rates only seem to rule out a genetic relationship with a shallow time depth.

6.1 The Qiang varieties

Subgrouping the Qiang varieties has proven a far more complex task than previously thought. Chang's (1967) work contains the oldest reconstructions. Sūn (1981b) divides the Qiang varieties into Northern Qiang and Southern Qiang, but given its heavy reliance on typological evidence, it is a typological subgrouping. Evans (2001) reconstructs Southern Qiang, often offering data from two Northern Qiang varieties for comparison. Most recently Sims (2016) adds Yonghe and Goudou to the Qiang database, terming the subgroup Southeastern Maoxian. Attempts to subgroup on the bases of shared innovations establishes that neither Yonghe nor Goudou fit with either Northern or Southern Qiang. Sims' work on the phylogenetic grouping suggests more complexity and thus probably more time depth than we originally suspected. As Sims concludes, more work needs to be done (see LaPolla, on Qiang, this volume).

6.2 Pumi and Ersuic

Prinmi [Primi, Pumi]:

Northern Prinmi

Southern Prinmi (Niuwozi Prinmi; see Ding, this volume)

Muya [Minyak; Miyao]

Ersuic:

Ersu

Lizu (Chirkova, this volume)

(but not Duoxu [Tosu])

6.3 The Ersuic languages

Dominic Yu's (2012) reconstruction of what he terms Proto-Ersuic establishes that Ersu and Lizu are closely related. However, the innovated directional verb prefixes and other evidence leave Duoxu outside of not just Proto-Ersuic but outside of Proto-Qiangic. Duoxu seems to have two directional verb prefixes, but they match neither Proto-Qiang nor anything else so far, a pattern that suggests the presence of directional prefixes in Duoxu is areal, not genetic. Sūn's (1990) evidence for the subgrouping of Duoxu with Ersu and Lizu is largely typological, establishing only its inclusion in the Qiangic linguistic area. As for genetic evidence, Chirkova and Handel (2013) present a shared chain of innovations—a change they schematize as $*sN > *NN > *N > \tilde{h} > \tilde{x}$. This constitutes evidence for subgrouping Ersu and Lizu (that is, for Ersuic), but it is quite weak evidence for the inclusion of Duoxu since Duoxu only shares the first part of the chain (the $*s\text{-}N\text{-} > NN\text{-}$), and the devoicing of the nasal onset component of the change is fairly widely distributed in Tibeto-Burman. The picture is further complicated, as Huáng and Yǐn (2012) point out, by Duoxu being analytic like the Loloish languages, rather than agglutinative like most Qiangic languages, suggesting that Duoxu has undergone considerable contact influence.

6.4 rGyalrongic

rGyalrongic (to use Sun’s term (2000a))

rGyalrong (proper) (three main dialects)

Situ (eastern rGyalrong)

lCogtse rGyalrong

Japhug [Chabao; WT Ja-phug] (northeastern rGyalrong)

Tshobdun Sidaba [Caodeng] (northwestern rGyalrong)

Zbu, Showu

Horpa [Ergong: Sün 1990; Daofu: Huang 1991]

Stau [Horpa, Tre-Hor]; Ergong: Sün (1990)

Dgebshes [Chinese Geshizha]

Stodsde [Chinese Shangzhai]

Gyurong Horpa

Khroskyabs [Guanyinqiao > Lavrung: Huang 1991]

Mu’erzong

Sun (2000a, 2000b) establishes that rGyalrongic subgroups together against Qiangic using three striking parallel innovations shared between the rGyalrongic languages: what Sun (2000a: 171) calls “glottality-inversion in past-stem formation, (ii) ablaut, (iii) transitivity marking via vocalic alternation in the orientation prefixes.” Also Jacques (on Stau in this volume) offers compelling evidence that, within the larger rGyalrongic, Horpa-Khroskyabs subgroups together against rGyalrong: (i) the generalization of the inverse forms in the non-local scenario and the complete loss of the inverse in the direct 3→3 forms, (ii) the loss of most traces of the nominalization prefixes, and (iii) a pattern of verbal reduplication unattested elsewhere, and so on.

There is clear evidence of Tibetan influences on rGyalrong resulting in mis-subgrouping it with Tibetan but as Nagano points out (this volume), most Tibetan-looking words are borrowings, while actual cognates between Tibetan and rGyalrongic are few (Jacques 2004: 169–70 reports fewer than 150 examples). The more basic vocabulary is typical of Tibeto-Burman in general, while the Tibetan-like forms are cultural items.

Wider affiliations: More important in terms of a genetic subgrouping are the similarities between the rGyalrong pronominalization system and the somewhat parallel systems in certain Kiranti languages, something also observed by Ebert (1990) and LaPolla (2003, 2006, 2013). More work remains to be done in this area.

6.5 Naish languages (Naxi)

	Core Naish	Outliers
Naxi	Na [Mosuo, Narua]	Shixing [Xumi, Shuhing]
Western		
Dayanzhen	Eastern	Laze [Lare, Shuitian] Namuyi
Lijang Naxi ‘Naxi proper’	Yongning Na	Baidi
Baoshanzhou	Guabie	
	Ninglang [Beiqu]	

Jacques and Michaud’s (2011) preliminary reconstruction with its numerous regular sound-meaning correspondences establish the core Naish languages as a close knit subgroup (see Lidz [LD], Yongning, this volume).

The directional verb prefixes in Table 1.7 resemble the directional verb forms in Qiangic, but the forms are quite distinct; so when Sūn (1990) uses this as evidence for the connection of Naish with Qiangic, it must be an areal connection, not a genealogical connection we are talking about. In addition to Sūn (1990), Sūn (2001) and Bradley (1997) both offer a Qiangic connection, but it must be areal influence since, while the constructions are similar the actual forms are not.

Jacques and Michaud (2011) and Michaud *et al.* (forthcoming) provide ample evidence for a Naish subgroup, with Namuyi and Shixing the closest relatives.

Table 1.8 shows reflexes of *ry- (and, *rw- and *s-wy-) in Core Naish, in Namuyi and Shixing, and in Lisu (as a representative of the Loloish component of PLB (Table 1.2 shows other reflexes of Loloish *ry-)). Dismissing random chance, one could argue the parallelism in the reflexes of *ry- was parallel but independent development—Sapir’s ‘drift’—but it would reconstitute a peculiar-looking group.

Alternately, it might be argued that Naish represents a group split off of Loloish, an analysis consistent with the Jacques and Michaud (2011; see also Chirkova 2012) finding that Naish and Lolo-Burmese—particularly Loloish—share a significant number of common lexical innovations. The remaining question would be is Naish a sister language to PLB or is it connected at a lower level to just Loloish? The preliminary guess is that Naxi (that is, Naish) is, as Matisoff (2003: 40) suggests, outlier Loloish. If the shared lexical innovations were inherited from, say, a hypothetical Burmo-Qiangic branch, the distribution should not be restricted as it is disproportionately to Naish and Loloish.⁹ In any case, without considerably more evidence, it would be premature to jump to any conclusions about Naish and Lolo-Burmese.

Wider relationships: Sūn (2001) and Bradley (1997) offer a Qiangic connection, which is an areal contact phenomenon. Nishida (1973, 1976) argues for a closer relationship with Burmese (hence, with Lolo-Burmese), which if one limits it to Loloish seems true. The evidence for a Burmo-Qiangic node is only suggestive.

TABLE 1.7 DIRECTIONAL VERB PREFIXES IN CORE NAISH, SHIXING, AND NAMUYI

		‘up’	‘down’	‘in’	‘out’
Shixing		dzi ³³ -	mia ³³ -	ji-	wu-
Namuyi		luo ³³ -	mi ³³ -		
Yongning Na (Mosuo)	gy ¹¹		mu ¹¹ ‘downward’		
Guìqióng [Yútōnghuà]	thu-		mi- ‘downward’		

TABLE 1.8 REFLEXES OF *RY- IN CORE NAISH, SHIXING AND NAMUYI, AND IN PLB

Lijiang Naxi	Yongning Na (LD)	Yongning Na	Namuyi	Shixing	PLB	Lisu Fraser
ɛi ³³	ɛi ³³	dʒ ³³ ɛi ⁵⁵	hiɔ ⁵³	dzi ³³ ɛɛ ⁵⁵	*rya ¹	h'ya ⁴ hundred
xw ³¹ -	ha ³³	xa ³¹ kha ³³	xī ³³ qho ³³	—	*ryak	h'ya ⁶ full day
xy ⁵⁵	hi ³¹	xi ³³	hi ¹³	dze ³³ ɛi ⁵⁵	*ryap	h'i ⁶ stand
xo ⁵⁵	ho ¹³	xu ¹³	hi ³³	ɛyi ⁵⁵	*ryat	h'i ⁶ eight
xu ⁵⁵	hū ³¹	xo ¹³ mi ³³	hi ³³ mbi ⁵⁵	—	*s-wyik	hi ⁶ - stomach
xw ³¹	ɛi ³³ -gi ¹³	xi ³¹	hi ⁵⁵ ŋɛ ³⁵	ɸui ⁵⁵ za ⁵⁵	*rwa ¹	-h'a ⁴ rain
x>ɛ-/ i	h-> e-/ i	x-> e-/ i	x->h-/ i	h-> ɛ/ i		

7 THE ‘SAL’ LANGUAGES

Burling (1983, 2012b) posited this upper-level subgroup, consisting of Bodo-Koch, Northern Naga¹⁰ (Konyakian) and Jinghpaw, on the basis of shared lexical innovations, e.g. PTB *b-war ‘fire’ and PTB *tsyar ‘sun’, naming it after the innovation of *sal* ‘sun’ (Burling 1983). As Burling suspected, many of the lexical innovations, instead of being restricted to Burling’s *sal* languages, are more widely distributed in Tibeto-Burman. This finding undermines the relationship between Boro-Koch and Konyak somewhat, and even more so in the case of the Jinghpaw.

Each of the smaller components is established by strong evidence. The first two are reconstructed: Boro-Koch by Burling (1959), and Konyak by French (1983). The third is expanded; Matisoff (2013), based on new descriptions of Sak/Chak (Huziwarra 2008) and of Kadu (Sangdong 2012), has assembled considerable comparative evidence for subgrouping the Asakian languages with Jinghpaw.

7.1 Bodo-Koch [Boro-Garo, Bodo-Garo] (see Burling 2012a)

Boro: Boro [Bodo]
Dimasa
Tiwa [Lalung]
Mech
Kachari
Hill Kachari
Kokborok [Tripuri]

Garo: Garo

Koch: Koch: Tintinkiya Koch, Wa’nang Koch and Pani Koch
Rabha
A’tong
Ruga

Deori [Deuri], Chutia

7.2 Konyakian (Northern Naga)

Tangsa
Yogli [= Jugli]
Lungecang (Lungchang, Longchang)
Nocte [Namsangia]
Wancho [Banpara]
Konyak [Tableng]
Phom (Chingmengnu, Tamlu)
Chang
Khiamniungan
Moshang
Wakching

7.3 Jinghpaw and Asakian

Jinghpaw¹¹ [Jingpho, Singpo (in parts of India)]

Asakian (Luish is apparently pejorative):

Sak/Chak (Huziwarra (2008))

Kadu (Sangdong 2012)

Andro[†]

Sengmai[†]

Chairel[†]

Wider relationships: The once posited subgrouping relationship between Jinghpaw and Rawang [Nungish] on the one hand and between Jinghpaw and Lolo-Burmese has been withdrawn. The frequent connection of Jinghpaw with Lolo-Burmese in part reflects contact-based convergence, not inheritance. Jinghpaw speakers are often bilingual in Zaiwa (Atsi), a Burmish language, in many cases living interspersed with Zaiwa speakers. Other Jinghpaw speakers also know Maru, Lachi, or Bola, and Burmese, all Burmish languages.

8 THE NORTHEASTERN BORDER AREA

The eastern border area of Nagaland, Manipur, and Mizoram, and a neighboring strip of Myanmar is home to Central Naga [Ao], the Angami-Pochuri, the Zeme, the Tangkhul, and the (Kuki-)Chin groups along with the Meitei and Karbi. These languages are often assumed to be a subgroup. However, as Burling (2003) writes, the proof is still lacking, not just for the higher-level subgroup but also for some of the component subgroups. The situation is further complicated by language contact. Until these are better understood, we will be using Post and Burling's geographical designation 'the eastern border area' (see Post and Burling, this volume).

8.1 Central Naga

Proto-Central Naga (PCN)

Proto-Ao

Chungli Ao

Mangmetong Mongsen Ao (Coupe 2007)

Changki (dialect of Mongsen)

Yacham

Tengsa

Sangtam [Tukumi]

Yimchungṛü [Yachumi]

Lotha [Lhota]

Central Naga is thoroughly established by Bruhn (2014), who reconstructs Proto-Ao from Chungli Ao and Mangmetong Ao and then Proto-Central Naga from Proto-Ao, Sangtam, Yimchungü, and Lotha. Correspondences are given for the transition from PTB > PCN > CN (see also Burling 2003; Post and Burling, this volume). The relational morphology, particularly cognate forms of agentive/instrument and locative compounds, give additional support for this grouping.

8.2 The Angami-Pochuri group

Post and Burling (this volume) characterize the Angami-Pochuri group as having two nuclei: an Angami nucleus and a Pochuri nucleus.

Angami [Tenyidie/Tenedyie]	Pochuri [southern Sangtam, eastern Rengma]
Chakhesang [Chokoso, Chokri-Kheza]	Meluri [Anyo]
Mao [Sopvoma, (E)memei, Poumai]	Ntenyi Sumi [Simi, Sema] Rengma proper

8.3 The Zelingrong [Zeme]

Zeme [Empeo, Kachcha•]	Nruanghmei [Rongmei, Kabui]
Mzieme	Puiron
Liangmai [Kwoireng]	Khoirao Marām

Shafer (1955), Garrison (1967), Burling (2003), and Post and Burling (this volume) all agree on this subgrouping, with the latter suggesting these varieties constitute a dialect chain.

8.4 Tangkhul

Proto-Tangkhul	Tusom
Standard Tangkhul	Phadāng (McCulloch 1859 wordlist)
Kachai	Champhung (Brown 1837 wordlist)
Huishu	
Maring	(is at least outside core Proto-Tangkhul, if not even more distant).

Based primarily on Standard Tangkhul, Kachai, and Huishu, Mortensen (2003) and Mortensen and Miller (2013) provide a reconstruction of Proto-Tangkhul with a focus on the rhymes. A wide range of innovations are presented including not just correspondences but also some lexical and morphological items. A tentative internal structure of Tangkhul (Mortensen 2003: 5) is given but not defended.

Having established a convincing set of Proto-Tangkhul correspondences, it is possible to determine that the following languages fall outside of the domain of Proto-Tangkhul: Liangmai and Maram are better placed in Zeliangrong, and Maring (Mortensen 2003) and Sorbung (Mortensen and Keogh 2011) fall outside the group, but it is not clear where they go (see also Burling 2003 and Post and Burling, this volume).

8.5 Chin [Kuki-Chin]

Central Chin ¹²	Maraic	Northern Chin	Southern Chin	Northwestern (Kuki-Chin)	
Mizo [Lushai]	Mara	Tedim [Tiddim]	Daa Chin	Aimol	Langrong
Laizo Lai	[Lakher/	Paite	Hyow	Anal	Monsang
Hakha Lai	Maram]	Sizang [Siyin]	Asho [Shö]	Hallam	Moyon
Laamtuk Thet	Senthang	Gangte	K'Cho Khyang	Bete	Taraو
Bawm	Zothung	Thado	Chinbok	Rangkol	Purum
[Banjogi]	Zophei	Ralte	Khomic	Chothe	Lamkang
Zahao	Lautu	Pawi		Koireng	
Hmar				Kom	

VanBik (2009) and other scholars have assembled strong shared morphological and phonological evidence both for (Kuki)-Chin-(Mizo) as a subgroup of TB and for an internal subgrouping of (Kuki)-Chin.

Externally, (Kuki)-Chin is distinguished from the rest of TB by several distinct shared changes. Phonologically, it has the change PTB *s-, *sy- > PKC *th- (VanBik 2009: 9). Several reflexes at first look like counterexamples: In Tedim, [s-] is an allophone of /t-/ occurring before /_i, as may also be the case in Paité.

Here we use VanBik's classification. Core Kuki-Chin indicates VanBik's KC, which excluded Old Kuki (for want of sufficient data).

Morphologically, it has two striking shared innovations. One is its system of subject agreement proclitics derived "probably as an outcome of denominalization (reanalysis of nominalized verb forms as finite verbs)" (DeLancey 2013a). Long recognized for their subgrouping value (e.g. Thurgood 1985), these possessive forms *kai 'I', *nang 'thou', and *a-mi 'third person' developed into a prefixal proclitic subject–verb agreement system that consists of *ka- 'first', *na- 'second', and *a- 'third.' The second striking shared innovation is the KC system of morphologically determined verb stem alternations (VanBik 2009: 9–17). To greatly oversimplify, Stem I forms are typically associated with main clauses and intransitive predicates, while Stem II forms are typically associated with subordinate clauses and transitive predicates (and often associated with closed syllables). Northwestern (Kuki)-Chin (NW(K)C) forms do not seem to manifest this stem alternation. However, the value of this for subgrouping is greatly reduced by the failure of the stems to correspond.

Finally, perhaps the strongest evidence for being a subgroup is VanBik's (2009) reconstruction of the phonology and lexicon.

Internally, (Kuki)-Chin has three branches: 1. Central, 2. Maraic, and 3. 'Peripheral', the latter the combination of Northern Chin and Chin (Gierson's Northern (Kuki)-Chin) and Southern (Plains) Chin, the latter a suggestion of Peterson. VanBik (2009) offers various shared phonological innovations in support of this configuration. The fourth column contains Northwestern (Kuki)-Chin ['Old Kuki']; the precise nature of its relationship to the rest of KC remains a question. VanBik, in lieu of more data, remains agnostic.

8.6 Mru-Hkongso

Peterson and Wright (2009) note that Mru-Hkongso (probably including Anu) apparently lacks the change PTB *s > th and the stem alternations characteristic of (Kuki)-Chin languages, suggesting in lieu of positive evidence of their inclusion these are not (Kuki)-Chin.

8.7 Karbi [Mikir]

Linguists who have worked with Karbi [Mikir] usually suggest it is closely related to the (Kuki)-Chin and the Naga languages. Extensive contact with Austroasiatic Khasian languages, particularly Pnar, has left its mark on the language (Konnerth 2014; Grüßner 1978). Burling (2003) leaves it unclassified.

8.8 Meithei/Meitei

Like Karbi, the place of Meithei within Tibeto-Burman subgrouping remains to be determined, in part because the picture has been clouded by long-term contact with Kuki and

Tangkhul among others (formerly Manipuri, among various older names; see Chelliah, this volume; Burling 2003; Post and Burling, this volume).

9 THE ‘RUNG’ LANGUAGES

LaPolla (2013 and elsewhere) proposes a subgroup based on, among other things, a cognate hierarchical person marking system (including a first person singular suffix, a second person singular suffix, a dual, and a plural marker), an inverse marker with the same distribution, and a #*-si* reflexive/middle marking verb suffix (not present in rGyalrong) which has largely the same distribution (Table 1.9; LaPolla 2013). The first and second person suffixes are transparently related to PTB **ŋa* ‘I’ and **naŋ* ‘you.’ LaPolla argues that the data provides evidence for an intermediate-level subgrouping termed ‘Rung’ encompassing the rGyalrongic languages, the Dulong-Rawang languages (T’rung, Rawang), the Kiranti languages, and the West Himalayan languages (Kinauri, Almora) as well as perhaps Kham and Chepang (LaPolla 2003, 2013).

LaPolla argues that a crucial part of early Tibeto-Burman migration involved a split into two different dispersal groups, each with its own distinct paths and which developed in different ways. Rung with migrations south down the river valleys represents one line of development and Qiangic (minus rGyalrong) is another. The relevant pattern is not found outside the circle these languages form around the edge of the Tibetan plateau (LaPolla, personal communication; see LaPolla, Chapter 2, this volume).

Two basic interpretations of the data in Table 1.9 exist. LaPolla argues that the languages with this particular configuration represent a complex shared innovation and thus these languages form a subgroup within Tibeto-Burman. In contrast, DeLancey (2010, 2013a, 2014, personal communication) and Jacques (2012b, personal communication) suggest that the verbal indexation system shared by rGyalrongic and Kiranti goes all the way back to Proto-Sino-Tibetan. Paraphrasing Jacques, this proto-system was subsequently completely lost in some languages (Tibetan, Lolo-Burmese, Chinese, Karen, etc.) and considerably restructured in others (Kuki-Chin, Kham, Jinghpao, etc.). In this conceptualization, it is not just the suffixes but the combination of suffixes and prefixes that is crucial. Particular emphasis is placed on the fact that the second person is represented by #*tV-* in both rGyalrongic and Southern Kiranti and on the presence of a common inverse prefix (see LaPolla 1992 for counter arguments).

Certainly reconstructing the system back further than the evidence justifies is not a methodologically sound practice. Notice this is not to say the system does not reconstruct back to Tibeto-Burman; it is only the comment that the crucial evidence is not yet in. If it does go back that far, we should be able to find the evidence.

Independent of the choice of models are the numerous instances in closely related languages where one and the same system is present in one but missing in another.

TABLE 1.9 THE ‘RUNG’ LANGUAGES

	1sg suf	2sg suf	dual	plural	reflexive
P-rGyalrong	#-ŋ	#-n	#-tsh	#-i	—
P-Dulong-Rawang	#-ŋ	#-n	#-si	#-i	#-si
P-Kiranti	#-ŋ	#-n	#-ci	#-i	#-nsi
P-West Himalayan	#-g/-ŋ	#-n	#-si	#-ni	#-si

Note: # indicates a rough reconstruction.

Few of these suggest independent innovation. For most, the system was present at an earlier time but subsequently lost. The most obvious cases are like Jinghpaw (Kurabe, this volume), in which the younger generation is losing the verbal morphology (still present in older records), and in the closely related Singpo, which has lost it entirely (Morey 2012). Sūn and Liu's (2009) Anong similarly shows massive loss of morphology under intense contact with Lisu in a short time period—most of it was still there in the 1950s when Sūn and Liu began recording it. Now only part of the older generation retain it. In Newar, the Kathmandu dialect has lost the system but it is retained in the Dolakhae dialect (Genetti, this volume). All three examples are examples of what is sometimes termed creoloid, that is, creole-resembling. Although the process is sometimes termed creolization, it does not necessarily imply a prior pidgin-to-creole scenario, but instead such creoloids may be simply the product of intense language contact (DeLancey 2013b). Certainly numerous examples also exist outside of Tibeto-Burman: Eastern Chamic, a Chamic language which served as a lingua franca, Mandarin, which has undergone both intense contact and has functions as a lingua franca, and various other prominent languages including English (McWhorter 2002), which have undergone intense language contact (McWhorter 2007). The main relevance of this for subgrouping is that the lack of agreement in language subgroups like Lolo-Burmese and Tibetan is difficult to evaluate.

10 RAWANG AND RELATED LANGUAGES [NUNGISH]

Dulong (an exonym) [other names: Taron, T'rung, Kiu (Qiu), Kiutze (Qiuzi), Kiupa, or Kiao]

Anong

Rawang

Agreement is a Nungish feature found in all dialects, but it is rapidly becoming lost in the Anong described by Sūn and Liu (2009), under intense pressure from Lisu.

Wider relationship: The literature often suggests a genetic connection with Jinghpaw, but strong substantiating evidence is lacking.

11 KIRANTI

Like Werner Winter, Ebert (2003) is tentative about the existence of a Kiranti subgroup; instead Ebert refers to the 30-plus languages as the Kiranti cluster. There is limited evidence for designating them as a subgroup: shared innovations in the pronouns, a related verbal agreement system, and the beginnings of some historical reconstruction (Michailovsky 1994, this volume). The line between similarities inherited from a common proto-language and those due to a long period of mutual contact and interaction is often difficult to distinguish.

Athpare	Hayu (this volume)	Thulung
Bahing	Jero	Wambule (this volume)
Bantawa	Khaling	Yakkha
Belhare (this volume)	Kulung	Yamphu
Camling (this volume)	Limbu	
Chilling	Nachiring	
Dumi	Sunwar	

Contact has restructured all of these languages to some degree, with language shift putting most in danger of disappearing. Nepali, the locally dominant language, has been a major influence. Ebert (2003) notes similarities between the Southeastern Kiranti languages and the Naga and Chin languages in the prefixed person markers and the participle formation with *ka-*, and similarities between the inverse marking of Camling and Bantawa and inverse marking in rGyalrong. Parts of Kiranti show non-Tibeto-Burman influences: parallels with reduplication in North-Dravidian Kurukh, syllable-final *k* in eastern Indo-Aryan and in Munda languages, and a highly agglutinative morphology characteristic of North Munda languages.

12 WEST HIMALAYAN GROUP

Kinauri cluster	Almora cluster
Kinauri, Kanauri	Rangkas
Chamba Lahuli	Byangsi
Kanashi	Chaudangsi
Rangloji, Gondla, Tinan	Darma
Bunan	
Jahri	
Manchad	

Sometimes referred to as West Himalayish and sometimes as Kinauri-Almora, this group is characterized by innovations in the pronouns and in the same shared innovated agreement markers (Thurgood 1985: 390–2) as are found in the other Rung languages (see Table 1.10). Whether or not this represents a dialect chain remains unclear.

Other languages: Raute [Raji] is sometimes added into West Himalayish, but it does not fit well.

Wider relationships: see ‘Rung’, above.

13 THE KHAM, MAGAR, AND CEPHANG LANGUAGES

Kham (see Watters, this volume)

Magar

Cephang

TABLE 1.10 WEST HIMALAYAN PRONOUNS AND PRONOMINALIZATION

	1st sg	1sg suf	2sg	2sg suf	3sg	3sg suf
P-West Himalayan	*gai	*-g/-ŋ	*ga-na	*-n	*do	
Kinauri cluster						
Kinauri	gə	-g	kə	-č/-ñ	do	-ñ
Chamba Lahuli	ge	-ga; -g	ka; ku	-na; -n	du	—
Kanashi	gu	-k	ko	-n	du	—
Manchad	gye	-g	kà?	-ñi	du	-ñi
Almora cluster						
Byangsi	ji	-?	gan	-n	vaii; u	
Rangkas	ji; jin	#-?	ga	*-?	hve; u	
Chaudangsi	ji	-?	gan	-n	vo; u	-ni
Darma	ji	#-?	ge	-n-	?u	—

Kham, Magar, and Chepang have been subgrouped in various ways. Watters (2003, this volume) puts Kham and Magar together, noting that “a careful examination of more innovative vocabulary makes it apparent that Magar is indeed Kham’s closest relative.” DeLancey (1987) likewise puts Kham and Magar together as sisters to Chepang. Kham and Magar have agreement, however, it is worth noting that while eastern Magar dialects have no agreement, western dialects do have it. It is not likely that the western dialects borrowed it from Kham, as Kham has biactantal systems, but the Magar dialects have subject agreement.

Wider relationships: Several authors suggest a relationship with Kiranti.

14 THE KARENIC BRANCH

Northern	Central	Southern
Pa-O	Kayah Li (Karen)	Pwo Sgaw

Brè (= Bwe)
Yintale
Palayachi
Mopwa

Padaung (transitional) between Northern and Central

Externally, Karenic is a well-defined subgroup of Tibeto-Burman. Internally, its division into northern, central, and southern is geographical. Following general linguistic practice, Solnit (this volume) concludes that western Kayah State (Karen) and the adjoining area of Karen State, the area of greatest diversity, is the probable homeland. Solnit points out that the Karenic languages are found at the southeastern edge of Tibeto-Burman territory (along with Lolo-Burmese and Tujia speakers), bringing Karen into close contact with members of two characteristically SVO language families, Tai-Kadai and Mon-Khmer. He notes the Tai influence is strikingly evident in the Proto-Karen consonant system and reports Tai interaction with the evolution of tones. Karen has significant numbers of loanwords from Tai-Kadai and from Mon-Khmer (specifically, the Palaungic and Monic branches). The Mon-Khmer influence is pervasive culturally and linguistically, including considerable bilingualism. Thus, despite the opinions sometimes expressed in older scholarship, modern scholarship sides with Solnit in assuming that Karen’s SVO word order is the result of contact with Subject–Verb–Object [SVO] Tai-Kadai and SVO Mon-Khmer languages (see chapters by Solnit and by Kato).

15 TANI (MIRISH)

Adi [Abor]	Gallo	Nishi [formerly Dafla], Nishing
Apatani/Apa Tani	Milang	Nyisu
Bengni	Miri	Padam Adi
Bokar	Mising (Plains Miri)	Padam-Minyong Adi
Damu	Na Bengni	Padam-Mising

Various scholars have argued that the Tani languages constitute a distinct Tibeto-Burman subgroup, a finding confirmed by Sun’s (1993a, 1993b) reconstruction of Proto-Tani. For a fuller discussion of the internal relationships within Tani, see Post and Sun (this volume). The wider relationships with nearby Tibeto-Burman languages are not clear yet.

16 KAMENGIC [PUROIK, LESS OFTEN PUROIT, ETC.]

The relationship between Puroik [Sulung, reportedly pejorative]; Bugun [Khoa, Khowa]; Chug; Lish; Mey; Rupa; Shergaon; and Sartang [= But Monpa] was apparently first recognized by Sun (1993a: 12 fn. 18), who suggested that Sherdukpen, Bugun, and Lish might form a subgroup, with Puroik being added in somewhat more tentatively. Chug, Lish, and Gompatsse, according to Blench and Post (2011: 3–5), are probably a single language. Sherdukpen labels a cluster formed from the names of the two largest villages (Bradley 1997: 12): Shergaon and Rupa [Tukpen]; despite the paucity of forms, it is clear that Mey fits in Sherdukpen.

The perception of aberrancy is seen in Rutgers' (1999) suggestion that they are isolates and in the Blench and Post (2011) suggestion that the data reflects the influence of a yet-to-be-identified substrata. In any case, Matisoff (2009) recognizes that Puroik (and by extension the group) is certainly Tibeto-Burman. Puroik, the only member for which we have anything other than meager data (Sun *et al.* 1991; Li 2004), yields an abundance of cognates with wider Tibeto-Burman. Van Driem (2001: 479–81) uses the words for ‘water’ *kho* and ‘fire’ *bwa* for Kho-Bwa, his label for the subgroup. Neither word, however, is unique to this cluster of languages: for ‘water; snow’, Hyslop (2011: 40 fn. 14) gives Kurtöp *khwe* ‘water’, Dzongkha *khai* ‘snow’, Bodo *khwa* ‘snow’, and Dakpa *kho* ‘snow’, and could give more. As for the ‘fire’ etymon, *mey ‘fire’ is the most common TB word for ‘fire.’ As Matisoff (2009) notes, it is obvious that this subgroup underwent the change of nasals to voiced stops, e.g. *m- > b-. The limited database kept the change from being obvious; most of these languages only have the relevant data for ‘fire’ and ‘name’ but in the larger Puroik database (Sun *et al.* 1991; Li 2004) there are at least six instances: *mey ‘fire’ > Puroik *bæ*³³; *minj ‘name’ > a³³be³³ŋ³³; *ma ‘not’ > ba³³; *mak ‘son-in-law’ > a³³bu⁵³; *maj ‘dream’ > mə³³bak³³; and, *mi ‘man’ > bi³³.¹³ The forms in Table 1.11 show additional cognate forms, while Table 1.12 (adapted from Blench and Post 2011), despite the suggestion that they are not cognate, show cognacy with wider Tibeto-Burman on the one hand and the frequent uniqueness in their reflexes within the subgroup on the other.

The evidence that Hruso (the first two columns in Table 1.13) and Miji (the next three columns in Table 1.13) has a special relationship to the Miji languages is at best weak, the sporadic *s- > *t- in ‘die,’ ‘three,’ and others (and sometimes further palatalization), the *sw- > *t- in ‘tooth,’ and other teasingly semi-regular correspondence patterns notwithstanding.

17 IDU-DIGARU (TAWRĀ)

Idu [Yidu, Chulikata]

Digaru [Taraon, Tawrā, Darang Deng]

Idu and Digaru are closely related. Kaman, however, is not linguistically close to Idu-Digaru, despite the fact that Kaman is included in the Deng nationality along with Idu (see Post and Burling, this volume).

18 KAMAN-MEYOR [MIDZUISH]

Kaman [Miju, Geman]

Meyor [Zakhring, Zaiwa, Zhá]

Strong evidence for a genealogical relationship between these two is lacking.

TABLE 1.11 ADDITIONAL EVIDENCE FOR TB MEMBERSHIP AND SUBGROUPING

gloss	PTB	JAM	Puroik	Bugan	Chug	Lish	Rupa	Shergaon	Sartang
fire	*mey		bæ ³³	boe	bei	bei	ba	ba	be
name	*minj		a ³³ be ^r ŋ ³³	—	biŋ	biŋ	—	—	—
six	*d-k-ruk		yək ³³	—	syk	ɟ ^b u?	ʃuk	ʃuk	kit
nine	*gəw		dəŋ ³³ ge ^r ə ⁵³	dige	t̪ik ^b u	t̪ik ^b u	d̪ik ^b it	t̪ik ^b i	t̪ek ^b e
eight	*gyat		[la ³³]	[mla]	sarge?	sarge?	sardʒat	sargyat	sardʒe
seven	*ni		[lie ³³]	[milye]	his	jis	sit	sit	si?
star	*kar		—	—	karma	karma	zik	t̪uzuk	t̪ydʒy
leaf	*lap		—	arap	ula?	ulap	alap	alap	arap
four	*b-ləy		vər ³³	—	psi	p ^b əhi	bsi	phsi	pʃi
woman	*mi ²	PLB	a ³³ mu ⁵³	bimi	d̪udma	esma	dʒimi	dʒimi	dʒymy k ^b re
sun	*nəy		—	—	nami	nami	nini	nini	nimi?
leg	*krəy		lae ³³	loe	lai	lei	la	la	le
two	*nis		ŋi ³³	ŋeŋ	nɪʃ	ŋes	jik	jit	nɪʃ
pig	*wak		—	wak	aba?	saba	swok	swag	swa?
tongue	*s-lay		rye ³³	—	loi	loi	lapon	laphð	le
tree	*sij		he ⁿ ³³	hijmua	ſiŋ	hiŋ	sɪntiŋ	hiŋ ^b uŋ	hiŋ
three	*sum		—	im	om	?um	uŋ	uŋ	um
male	*pu		p ^b u	bp ^b ua	pədəŋ	bǔdūn	ſirin	džuhu	džiriŋ

TABLE 1.12 REPUTEDLY NON-TIBETO-BURMAN COGNATES (ADAPTED FROM BLENCH AND POST 2011)

gloss	PTB	Puroik	Bugan	Chug	Lish	Rupa	Shergaon	Sartang
head	—		k ^b ruk	k ^b lo?	k ^b olo?	k ^b ruk	k ^b ruk	k ^b ru?
stomach	*ri:l	θui ³³	lui	hiliŋ	hipiŋ	sliŋ	siriŋ	fəriŋ
mouth	*ka	—	syam	k ^b otfu	hotfok	tʃaw	nitʃaw	tʃonə
dog	*kwəy	—	—	wat ^b i	wat ^b i	b ^b a	p ^b it ^b a	pet ^b e
five	*ŋa	—	kua	k ^b a	k ^b a	k ^b u	k ^b u	k ^b u
bone	*rus	—	—	sukus	sukus	skik	skit	ski?
moon	*s-la	—	—	atnamba	namba	namblu	namblu	namlu?
tooth	*swa	kə ³³ tuan ³³	—	hintunj	ʃintunj	toktʃe	nuthunj	nit ^b inj
water	*twiy?	—	k ^b o	k ^b u	k ^b au	k ^b o	k ^b o	k ^b ow
ten	*tsyay	suat ⁵³ pa ⁵³	sunŋwa	ſan	ſan	sō	sō	sou
hand	*l(y)ak	get ³³	—	hut	hu	ik	ik	ik
eye	*mik	—	—	k ^b um	k ^b umu	kivi	khibi	k ^b a?by
nose	*na	[pɔk ³³]	ep ^b uŋ	heŋp ^b oŋ	hemponj	nəfunj	nup ^b uŋ	ap ^b uŋ

Wider relationships: So far there is no evidence that Miju [Kaman] has a special relationship with any Mishmi language. Suggestions to the contrary probably result from the confusion of ethnicity with language since the Miju speakers are ethnically Mishmi (see Post and Burling, this volume).

19 KORO-MILANG

Only weak evidence for a genealogical relationship for these exists.

TABLE 1.13 HRUSO, DHIMMAI, AND LEVAI

PTB Matisoff 2003	Hruso Shafer 1947	Hruso Simon 1970	Dhimmai Shafer 1947	Dhimmai Simon 1979	Levai Bodt and Lieberherr 2015	
*səy	tsu	dzañe	t'i	ci/_i	tai	die
*g-sum	tzû -i	zi	ge-t'an	githin	kət̪əñ	three
*swa	-t'u	itcu/_u	t'u	thu	mətu:	tooth
*mi	næ-	nina	nə	ñih	—	man
*myak	-nyi	eñi	mre	mih	mejä?	eye
—	p'u	phu	p'oñ	phung	səpi:	hill
*tsa	-sa	sa	zə	zu	mədzu:	child; son
*ba:r	-ba	-ba	—	-boh	-buə:	flower
*s-rij	-śu	-ishshi	—	shin	səñ	alive
*dzya	tsá-	tsa-	sə-	tsuh	teð?	eat
*doŋ-i	t'u	thu	—	thung	tun	drink

20 RAUTE, RAJI

It is unclear what to do with Raute, Raji. Raute [Raji] is sometimes added into West Himalayish, but it does not fit well.

21 UNSUBGROUPED LANGUAGES

The languages immediately below are otherwise ungrouped Tibeto-Burman, and have been particularly resistant to a more precise subgrouping.

Lepcha [Rong] (see Plaisier, this volume)

Newar:

Dolakhae dialect (see Genetti, this volume)

Kathmandu dialect (see Hargreaves, this volume)

It is noteworthy that while the Kathmandu dialect lacks agreement, the Dolakhae dialect has it.

- Pyu. Pyu is an extinct language of a Tibeto-Burman people who once dominated much of what is now northern Burma, while southern Burma was part of a Mon Kingdom. The former influence of the Pyu is reflected in the fact that early Burmese inscriptions were occasionally written in the Pyu script and in the references to them found in Chinese records of the time.
- Tujia. The two mutually unintelligible Tujia varieties are found on the southeastern edge of Tibeto-Burman in northwest Hunan, where Tujia has come under considerable contact pressure. Its subgrouping within Tibeto-Burman remains a mystery (see Xu Shixuan, this volume).
- Bai. In Wiersma's (1990) excellent dissertation on Bai, she notes that some scholars argue that Bai is a Tibeto-Burman language with a heavy layer of Chinese loanwords while others argue that it is an Old Chinese dialect that split off from the rest of Chinese some 3,000 years ago. It has been difficult for scholars to determine whether the similarities between Chinese and Bai reflect the results of long-term contact or reflect inherited features, because Bai has been under the influence of

both Tibeto-Burman languages (for instance, Lisu, Yi, and Naxi) and Chinese. However, the evidence increasingly suggests that Bai is a Tibeto-Burman language influenced by Chinese, rather than the reverse. As Dryer notes (this volume), some of the morphological oddities shared by Chinese and Bai are likely to be contact-induced, just as is some of the shared lexical material. The most compelling evidence is Xu (2015), which demonstrates that where third-century mainstream Chinese and Shāndōng differ dialectally, current mainstream Chinese shows the Shāndōng forms, but Bai shows the third-century mainstream forms. In short, Bai borrowed from the third-century mainstream Chinese rather than inheriting from Shāndōng (see also Hefright 2011).

22 CONCLUSION

The subgrouping of these languages is still in its infancy. For the majority, the evidence is suggestive, and in many even the minimal database for a proper analysis is lacking. However, one is greatly encouraged by the appearance of dissertations that have laid the foundations for the future.

NOTES

- 1 I thank Scott DeLancey, Guillaume Jacques, and Randy LaPolla, who gave generously of their time discussing the hierarchical person agreement systems and the subgrouping of Tibeto-Burman as well as elsewhere. Thanks are also owed to Gwendolyn Hyslop, David Peterson (on Chin), Mark Post, Liberty Lidz, Alexis Michaud, and David Bradley. All caught errors of fact and concept. The remaining errors however are my own and I shall be astonished if the mistakes should be minor and would be grateful for the corrections of the readers.
- 2 It is not a survey of the history of Sino-Tibetan subgrouping; for this, the reader can consult, among the most widely cited works, Shafer (1974), Benedict (1972), Hale (1982), and Matisoff (1991), all partially updated by Bradley (2002: 74–5, 2012). Other chapters in this volume also have a focus on subgrouping: on subgrouping in general see LaPolla, on Northeast India see Post and Burling, on Kiranti see Michailovský, and on Tani see Post and Sun, and most chapters at least mention it.
- 3 In most cases, space limitations as well as time limitations prevent a full discussion of much of the available. Nonetheless this chapter frequently criticizes proposals for not providing sufficient supporting evidence. The irony of this has not been lost on me.
- 4 Genetti (forthcoming) is a valuable overview, giving bottom-up language groupings frequently accompanied by evaluations. More extensive is Bradley (1997) partially updated in Bradley (2002). Given the lack of a clear understanding of higher-level relationships, Bradley organizes his survey in the only way possible: essentially by geographical areas, supplemented by what is known about the genetic subgrouping. Van Driem (2001) is another widely cited source.
- 5 Thurgood's (1985) suggestion that Tamangic had a related velar-initial second person pronoun has not held up (Nishi 1991; Honda 2008).
- 6 Dakpa is sometimes termed Northern Monpa, but Monpa is used so imprecisely and so broadly that it is avoided here.
- 7 Perhaps *sinj ‘tree; wood’ (< Mark Post).
- 8 The vowels of the prefixes often display vowel harmony (ignored in Table 1.6). The semantics have their patterns of variation. The cardinal directions north, south, east, west vary as they were originally the location and orientation of where the speakers lived. The notion of toward the center and upriver may be coded differently or the

- same way; likewise with other prefixes. Matching prefixes across languages, however, is largely straightforward.
- 9 That the Lolo-Burmese tones do not seem to correspond to Naish tones (Lidz, personal communication; also the author, independently) is noteworthy but not a counterexample. One suspects that these formerly Loloish speakers lost their tonal system under contact, and that the current system is a subsequent development. For a parallel case in Jiamao, see Thurgood 1992.
 - 10 French used Northern Naga for this group; in light of the ambiguity of Naga, for clarity of reference we have chosen Konyak.
 - 11 Jinghpaw is the name for the language; Kachin is the name for the ethnic group.
 - 12 Chin and Kuki are essentially synonyms for the same thing, with Chin being used more in India and Kuki more in Burma. However, Kuki has taken on a strongly pejorative flavor so as far as possible Chin will be used in this work.
 - 13 The initials in ‘corpse’ *s-man; ‘ripe’ *s-mij; and *nam ‘smell’, however, remain nasals.

REFERENCES

- Andvik, Erik. 2010. *A Grammar of Tshangla*. Leiden/Boston: Brill.
- Benedict, Paul K. 1972. *Sino-Tibetan: A Conspectus*. Cambridge: Cambridge University Press.
- Bielmeier, Roland *et al.* forthcoming. *Comparative Dictionary of Tibetan Dialects*. Vol. I: Verbs, Vol. II: Nouns. (No further publication information available at present.) (Preprint no.1, 1997.)
- Blench, Roger and Mark W. Post. 2011. (De)classifying Arunachal languages: reconsidering the evidence. Unpublished draft of 1 November 2011.
- Boedt, Timotheus A. and Ismael Lieberherr. 2015. First notes on the phonology and classification of the Bangru language of India. *Linguistics of the Tibeto-Burman Area* 38.1: 66–124.
- Bradley, David. 1979. *Proto-Loloish*. Scandinavian Institute of Asian Studies, Monograph series no. 39. London: Curzon Press.
- Bradley, David. 1997. Tibeto-Burman languages and classification. In: David Bradley (ed.), *Tibeto-Burman Languages of the Himalayas*, 1–72, *Papers in Southeast Asian Linguistics*, no. 14. Pacific Linguistics, Series A, No. 86.
- Bradley, David. 2002. The subgrouping of Tibeto-Burman. In: Christopher Beckwith (ed.), *Medieval Tibeto-Burman Languages*, 73–112. Leiden: Brill.
- Bradley, David. 2012. Tibeto-Burman languages of China. In: Rint Sybesma (ed.), *Encyclopedia of Chinese Language and Linguistics*, 1–72. Leiden: Brill.
- Brown, Nathan. 1837. Comparison of Indo-Chinese languages. *Journal of the Asiatic Society of Bengal* 6.2: 1023–1038.
- Bruhn, Daniel W. 2014. A phonological reconstruction of proto-central Naga. PhD dissertation, University of California.
- Burling, Robbins. 1959. Proto Bodo. *Language* 35.3: 433–453.
- Burling, Robbins. 1967. *Proto-Lolo-Burmese*. *International Journal of American Linguistics* 33.2, Part II. Bloomington: Indiana University.
- Burling, Robbins. 1983. The ‘Sal’ languages. *LTBA* 7.2: 1–32.
- Burling, Robbins. 2003. The Tibeto-Burman languages of Northeastern India. In: Graham Thurgood and Randy J. LaPolla (eds), *The Sino-Tibetan Languages*, 169–192. Routledge Language Family Series. New York: Routledge.
- Burling, Robbins. 2012a. The Stammbaum of Boro-Garo. In: Gwendolyn Hyslop, Stephen Morey and Mark W. Post (eds), *North East Indian Linguistics*, vol. 4, 21–35. New Delhi: Cambridge University Press.

- Burling, Robbins. 2012b. Where did the question ‘Where did my tribe come from’ come from? In: Toni Huber and Stuart Blackburn (eds), *Origins and Migrations in the Extended Eastern Himalayas*, 49–62. Leiden: Brill.
- Chang, Kun. 1967. A comparative study of the Southern Ch’iang dialects. *Monumenta Serica* 26: 422–443.
- Chirkova, Katia. 2012. The Qiangic subgroup from an areal perspective: a case study of the languages of Muli. *Language and Linguistics* 13: 133–170.
- Chirkova, Katia. 2014. The Duoxu language and the Duoxu-Ersu-Lizu relationship. *LTBA* 37.1: 133–170.
- Chirkova, Katia and Zev Handel. 2013. Diachronic developments of voiceless nasals: the case of Ersu, Lizu, and related languages. Updated version of a paper given at the 46th International Conference on Sino-Tibetan Languages and Linguistics, Dartmouth, 2013. Available at: www.academia.edu/7637522/ (accessed December 2014).
- Coupe, Alexander R. 2007. *A Grammar of Mongsen Ao*. Berlin: Mouton de Gruyter.
- Das, Sarat Chandra. 1902. *A Tibetan-English Dictionary with Sanskrit Synonyms*. Alipore, Calcutta: West Bengal Government Press.
- Das Gupta, Kamalesh. 1968. *An Introduction to Central Monpa*. Shillong: Government of NEFA.
- DeLancey, Scott. 1987. Sino-Tibetan languages. In: Bernard Comrie (ed.), *The World’s Major Languages*, 799–810. Oxford: Oxford University Press.
- DeLancey, Scott. 2010. Towards a history of verb agreement in Tibeto-Burman. *Himalayan Linguistics* 9.1: 3–39.
- DeLancey, Scott. 2013a. The history of postverbal agreement in Kuki-Chin. *Journal of the Southeast Asian Linguistics Society* 6: 1–17.
- DeLancey, Scott. 2013b. Creolization in the divergence of the Tibeto-Burman languages. In: Thomas Owen-Smith and Nathan Hill (eds), *Trans-Himalayan Linguistics*, 41–70. Berlin: Mouton de Gruyter.
- DeLancey, Scott. 2014. Second person verb forms in Tibeto-Burman. *LTBA* 37.1: 3–33.
- Ebert, Karen H. 1990. On the evidence for the relationship Kiranti–Rung. *LTBA* 13.1: 57–78.
- Ebert, Karen H. 2003. Kiranti languages: an overview. In: Graham Thurgood and Randy J. LaPolla (eds), *The Sino-Tibetan Languages*, 505–517. London: Routledge.
- Evans, Jonathan P. 2001. *Introduction to Qiang Lexicon and Phonology: Synchrony and Diachrony*. Tokyo: ILCAA, Tokyo University of Foreign Studies.
- French, Walter T. 1983. Northern Naga: a Tibeto-Burman mesolanguage. PhD dissertation, City University of New York.
- Genetti, Carol. forthcoming. The Tibeto-Burman languages of South Asia: the languages, histories, and genetic classification. In: Hans Henrich Hock, Elena Bashir and K. V. Subbarao (eds), *South Asia* (World of Linguistics). Berlin: Mouton.
- Gong, Xu. 2015. How old is the Chinese in Bai? CRLAO, EHESS/INALCO sponsored workshop in Paris.
- Grüßner, Karl-Heinz. 1978. *Arleng Alam, die Sprache der Mikir: Grammatik und Texte*. Beiträge zur Südasiensforschung 39. Wiesbaden: Franz Steiner.
- Hale, [Everett] Austin. 1982. *Research on Tibeto-Burman languages. Trends in linguistics*. New York: Mouton.
- Handel, Zev. 2008. What is Sino-Tibetan? Snapshot of a field and a language family in flux. *Language and Linguistics Compass* 2/3: 422–441.
- Hashimoto, Mantaro J. 1986. The Altaicization of Northern Chinese. In: John McCoy and Timothy Light (eds), *Contributions to Sino-Tibetan Studies*, 76–97. Leiden: E.J. Brill.

- Hefright, Brook Emerson. 2011. Language contact as bilingual contrast among Báí users in Jiānchuān Country, China. PhD dissertation, University of Michigan.
- Hodgson, Brian H. 1853. Sifan and Horsok vocabularies. *Journal of the Asiatic Society of Bengal* 22: 117–151.
- Honda, Isao. 2008. Some observations on the relationship between Kaire and Tamangic. *Nepalese Linguistics: Journal of the Linguistic Society of Nepal* 23: 83–115.
- Huang, Bufan. 1991. Daofuyu [The Daofu language = Stau]. In: Dai Qingxia *et al.* (eds), *Zangmianyu shiwu zhong* [Fifteen Tibeto-Burman Languages], 1–45. Beijing: Yanshan Press.
- Huáng, Bùfán 黃布凡 and Yǐn Wèibīn 尹蔚彬. 2012. Duōxùyǔ gàiikuàng 多须语概况 [An outline of the Duoxu language]. *Hàn-Zàngyǔ Xuébào* 《汉藏语学报》 6: 49–70.
- Huziwarra, Keisuke. 2008. Chakku-go no kizyutut gengogakuteki kenkyuu [A descriptive study of the Cak language]. PhD dissertation, Kyoto University.
- Hyslop, Gwendolyn. 2009. Kurtöp tone: a tonogenetic case study. *Lingua* 119: 827–845.
- Hyslop, Gwendolyn. 2011. A grammar of Kurtöp. PhD dissertation, University of Oregon.
- Hyslop, Gwendolyn. 2013. On the internal phylogeny of East Bodish. *North East Indian Linguistics* 5: 91–109.
- Hyslop, Gwendolyn. 2014. Bodic. *Language and Linguistics Compass* 8.6: 243–270.
- Jacques, Guillaume. 2004. The laterals in Tibetan. Paper given at Himalayan Languages Symposium, Thimphu, Bhutan, 1–3 December 2004.
- Jacques, Guillaume. 2012a. The Tangut kinship system in Qiangic perspective. In: Nathan W. Hill (ed.), *Medieval Tibeto-Burman Languages IV*, 211–257. Leiden: Brill.
- Jacques, Guillaume. 2012b. Agreement morphology: the case of Rgyalrongic and Kiranti. *Language and Linguistics* 13.1: 83–116.
- Jacques, Guillaume. forthcoming. The genetic position of Chinese. In: Rint Sybesma, James Huang, Wolfgang Behr and Zev Handel (eds), *Encyclopedia of Chinese Languages and Linguistics*. Leiden: Brill.
- Jacques, Guillaume, and Alexis Michaud. 2011. Approaching the historical phonology of three highly eroded Sino-Tibetan languages: Naxi, Na and Laze. *Diachronica* 28.4: 468–498.
- Jäschke, Heinrich August. 1881. *A Tibetan–English Dictionary, with Special Reference to the Prevailing Dialects, to which is Added an English–Tibetan Vocabulary*. London: Routledge and Kegan Paul.
- Konnerth, Linda. 2014. A grammar of Karbi. PhD dissertation, University of Oregon.
- LaPolla, Randy J. 1992. On the dating and nature of verb agreement in Tibeto-Burman. *Bulletin of the School of Oriental and African Studies* 55.2: 298–315.
- LaPolla, Randy J. 1994. Parallel grammaticalizations in Tibeto-Burman: evidence of Sapir's 'drift'. *Linguistics of the Tibeto-Burman Area* 17.1: 61–80.
- LaPolla, Randy J. 2001. The role of migration and language contact in the development of the Sino-Tibetan language family. In: R.M.W. Dixon and Alexandra Y. Aikhenvald (eds), *Areal Diffusion and Genetic Inheritance: Case Studies in Language Change*, 225–254. Oxford: Oxford University Press.
- LaPolla, Randy J. 2003. Overview of Sino-Tibetan morphosyntax. In: Graham Thurgood and Randy J. LaPolla (eds), *The Sino-Tibetan Languages*, 22–28. (Routledge Language Family Series.) New York: Routledge.
- LaPolla, Randy J. 2006. Sino-Tibetan languages. In: Keith Brown (ed.), *Encyclopedia of Language and Linguistics* (2nd edition), 393–397. London: Elsevier.
- LaPolla, Randy J. 2010. Language contact and language change in the history of the Sinitic languages. *Procedia: Social and Behavioral Sciences* 2(2010).5: 6858–6868

- (Special issue on *The harmony of civilizations and prosperity for all: diversity in the development of human civilizations*, selected papers from the Beijing Forum, Peking University and Great Hall of the People, Beijing, 2–4 November 2007, ed. by Yan Jun). Available at: www.sciencedirect.com/science/article/pii/S1877042810011705.
- LaPolla, Randy J. 2013. Subgrouping in Tibeto-Burman: can an individual-identifying standard be developed? How do we factor in the history of migrations and language contact? In: Balthasar Bickel, Lenore A. Grenoble, David A. Peterson and Alan Timberlake (eds), *What's Where Why? Language Typology and Historical Contingency*, 463–468. Amsterdam and Philadelphia: John Benjamins.
- LaPolla, Randy J. 2015. Sino-Tibetan syntax. In: William S.-Y. Wang and Chaofen Sun (eds), *The Oxford Handbook of Chinese Linguistics*, 45–57. Oxford: Oxford University Press.
- Li, Daqin. 2004. *Sulong-yu yanjiu [A Study of the Sulung Language]*. Beijing: People's Publishing Co.
- Lu, Shaozun. 1986. *Cuona Menba yu jianzhi [Brief Description of the Cuona Menba Language]*. Beijing: Minzu Chubanshe.
- Marrison, Geoffrey E. 1967. The classification of the Naga languages of North-East India. PhD dissertation, University of London, School of Oriental and African Studies.
- Matisoff, James A. 1972. *The Loloish Tonal Split Revisited*. Berkeley: Center for South and Southeast Asian Studies.
- Matisoff, James A. 1991. Sino-Tibetan linguistics: present state and future prospects. *Annual Review of Anthropology* 20: 469–504.
- Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley: University of California Press.
- Matisoff, James A. 2009. Stable roots in Sino-Tibetan/Tibeto-Burman. In: Yasuhiko Nagano (ed.), *Issues in Tibeto-Burman Historical Linguistics* (Senri Ethnological Studies 75), 291–318.
- Matisoff, James A. 2013. Re-examining the genetic position of Jingpho: putting flesh on the bones of the Jingpho/Luish relationship. *LTBA* 36.2: 15–95.
- Matisoff, James A., Stephen P. Baron and John Lowe. 1996. *Languages and Dialects of Tibeto-Burman*. University of California, Berkeley: Center for Southeast Asia Studies.
- Mazaudon, Martine. 1977. Tibeto-Burman tonogenetics. *LTBA* 3.2: 1–123.
- Mazaudon, Martine. 1978. Consonantal mutation and tonal split in the Tamang sub-family of Tibeto-Burman. *Kailash* 6.3: 157–179.
- Mazaudon, Martine. 2005. On tone in Tamang and neighbouring languages: synchrony and diachrony. In: S. Kaji (ed.), *Crosslinguistic Studies of Tonal Phenomena: Historical Development, Tone Syntax Interface, and Descriptive Studies*, 7996. Tokyo University of Foreign Studies, Tokyo, Japan: Institute for the Study of Languages and Cultures of Asia and Africa. Available at: halshs.archives-ouvertes.fr/halshs-00139489/document.
- McCulloch, W. 1859. *Account of the Valley of Munnipore and of the Hill Tribes: With a Comparative Vocabulary of the Munnipore and Other Languages*. Calcutta: Bengal Printing Co.
- McWhorter, John H. 2002. What happened to English? *Diachronica* 19: 217–272.
- McWhorter, John H. 2007. *Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars*. New York, Oxford: Oxford University Press.
- Meillet, Antoine. 1982 [1914]. *Linguistique Historique et Générale*. Paris: Champion.
- Michailovsky, Boyd. 1994. Manner vs place of articulation in the Kiranti initial stops. In: Hajime Kitamura, Tatsuo Nishida and Yasuhiko Nagano (eds), *Current Issues in Sino-Tibetan Linguistics*, 766–772. Osaka: Organizing Committee, 26th International Conference on Sino-Tibetan Languages and Linguistics.

- Michailovsky, Boyd and Martine Mazaudon. 1994. Preliminary notes on the languages of the Bumthang group. In: Per Kvaerne (ed.), *Tibetan Studies: Proceedings of the 6th Seminar of the International Association for Tibetan Studies, Fagernes 1992*, vol. 2, 545–557. Osaka: Institute for Comparative Research in Human Culture.
- Michaud, Alexis, Limin He and Yaoping Zhong. forthcoming. Naxi (Naish). In: Rint Sybesma, Wolfgang Behre, Zev Handel and C.T. James (eds), *Encyclopedia of Chinese Language and Linguistics*. Leiden: Brill.
- Morey, Steven. 2012. The Singpho agentive—functions and meanings. *LTBA* 35.2: 1–14.
- Mortensen, David R. 2003. Comparative Tangkhul. PhD dissertation, University of California at Berkeley.
- Mortensen, David R. and Jennifer Keogh. 2011. Sorbung, an undocumented language of Manipur: its phonology and place in Tibeto-Burman. *Journal of the Southeast Asian Linguistics Society* 4.1: 64–114.
- Mortensen, David R. and James A. Miller. 2013. A reconstruction of Proto-Tangkhulic rhymes. *LTBA* 36.1: 1–32.
- Nichols, Johanna. 1996. The comparative method as heuristic. In: Mark Durie and Malcolm D. Ross (eds), *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*, 39–71. New York: Oxford University Press.
- Nishi, Yoshio. 1986. A classification of Tibetan dialects [in Japanese]. *National Minority Studies Museum Research Report* 11.4: 837–901.
- Nishi, Yoshio. 1991. The distribution and classification of Himalayan languages (Part II) [In Japanese]. *Bulletin of the National Museum of Ethnology* 16.1: 31–158.
- Nishida, Tatsuo. 1973. *A Study of the Toso-Chinese Vocabulary*. Kyoto: Toso I-yu.
- Nishida, Tatsuo. 1976. His-hsia, Toso, and Lolo-Burmese languages. *Studia Phonologica* 10: 1–15.
- Noonan, Michael. 2008a. Patterns of development, patterns of syncretism of relational morphology in the Bodic languages. In: Jóhanna Barðdal and Shobhana Chelliah (eds), *The Role of Semantics and Pragmatics in the Development of Case*, 879–881. Amsterdam: John Benjamins.
- Noonan, Michael. 2008b. Nominalizations in Bodic languages. In: María José López-Couso and Elena Seoane, in collaboration with Teresa Fanego (eds), *Rethinking Grammaticalization: New Perspective for the Twenty-First Century*, 219–221. Amsterdam: John Benjamins.
- Noonan, Michael. 2008c. Contact-induced change in the Himalayas: the case of the Tamangic languages. In: Peter Siemunds and Noemi Kintana (eds), *Language Contact and Contact Languages*, 81–106. Amsterdam: John Benjamins.
- Noonan, Michael. 2008d. Case compounding in the Bodic languages. In: Greville Corbett and Michael Noonan (eds), *Case and Grammatical Relations*, 219–238. Amsterdam: John Benjamins.
- Noonan, Michael. 2011. Aspects of the historical development of nominalizers in the Tamangic languages. In: Yap Foong, Ha Karen Grunow-Harsta and Janick Wrona (eds), *Nominalization in Asian Languages* (Typological Studies in Language 96), 195–214. Amsterdam: John Benjamins.
- Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge Language Surveys, Cambridge University Press.
- Paudel, Krishna. 2008. Ghale language: a brief introduction. *Nepalese Linguistics* 23: 168–185.
- Peterson, David A. and Jonathan Wright. 2009. Mru-Hkongso: a new Tibeto-Burman grouping. Paper presented at the International Conference on Sino-Tibetan Languages and Linguistics, Chiang Mai.

- Ross, Malcolm D. 1988. *Proto Oceanic and the Austronesian Languages of Western Melanesia*. Canberra: Pacific Linguistics.
- Ross, Malcolm D. 1997. Social networks and kinds of speech-community event. In: Roger Blench and Matthew Spriggs (eds), *Archaeology and Language I: Theoretical and Methodological Orientations*, 209–261. London: Routledge.
- Rutgers, Leopold Roland. 1999. Puroik or Sulung of Arunachal Pradesh. Paper presented at the 5th Himalayan Languages Symposium, Kathmandu, Nepal. (I have not personally seen this.)
- Sagart, Laurent. 2005a. Sino-Tibetan-Austronesian: an updated and improved argument. In: Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds), *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, 161–176. London: Routledge.
- Sagart, Laurent. 2005b. Tai-Kadai as a subgroup of Austronesian. In: Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas (eds), *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, 177–181. London: Routledge.
- Sangdong, David. 2012. A grammar of the Kadu (Asak) language. PhD dissertation, La Trobe University.
- Sapir, Edward. 1920. Comparative Sino-Tibetan and Na-Dene Dictionary. Ms. Ledger. American Philosophical Na 20a.3. Society. [microfilm]
- Shafer, Robert. 1947. Hruso. *Bulletin of the School of Oriental and African Studies* 12: 184–196.
- Shafer, Robert. 1955. Classification of the Sino-Tibetan languages. *Word* 2: 94–111.
- Shafer, Robert. 1957. Note on Athapaskan and Sino-Tibetan. *International Journal of American Linguistics* 23: 116–117.
- Shafer, Robert. 1969. A few more Athapaskan and Sino-Tibetan comparisons. *International Journal of American Linguistics* 35: 67.
- Shafer, Robert. 1974. *Introduction to Sino-Tibetan, Parts 1–2*. Wiesbaden: Harrassowitz.
- Sharma, B. Deben. 2005. *Lhops (Doya) of Bhutan: An Ethnographic Account*. New Delhi: Akansha Publishing House.
- Shirai, Satoko. 2009. Directional prefixes in nDrapa and neighboring languages: an areal feature of Western Sichuan. In: Yasuhiko Nagano (ed.), *Issues in Tibeto-Burman Historical Linguistics*, 7–20. Osaka: National Museum of Ethnology.
- Simon, Ivan Martin. 1970. *Aka Language Guide*. Shillong: North-East Frontier Agency.
- Simon, Ivan Martin. 1979. *Miji Language Guide*. Shillong: North-East Frontier Agency.
- Sims, Nathaniel. 2016. Towards a more comprehensive understanding of Qiang dialectology. *Language and Linguistics* 17.3: 351–381.
- Sūn, Hóngkāi. 1981a. The category of directionality in the Qiang verb. *Minzu Yuwen* 1981.1: 334–342.
- Sūn, Hóngkāi. 1981b. *Qiangyu Jianzhi [A Brief Description of the Qiang Language]*. Beijing: Nationalities Press.
- Sūn, Hóngkāi. 1988. A classification of Tibeto-Burman languages in China [in Chinese]. *Languages and History in East Asia: Festschrift for Tatsuo Nishida on the Occasion of his 60th Birthday*, 61–73. Kyoto: Shokado.
- Sūn, Hóngkāi. 1990 [1983]. The languages of the ‘ethnic corridor’ in western Sichuan. *LTBA* 13.1: 1–31. [In Chinese: 1983. *Chuan Minzu Zoulang Diqu de Yiyuan [Research on the Nationalities of Southwestern China]*, 429–454. Sichuan Nationality Press.] Translated from Chinese by Jackson T.-S. Sun.
- Sūn, Hóngkāi. 2001. On the language of the Qiangic branch in Tibeto-Burman. *Language and Linguistics* 2: 157–181.

- Sūn, Hóngkāi, Huang Bufan et al. 1991. *Zangmianyu yuyin he cihui*. Beijing: Zhongguo Shehui Kexue Chubanshi.
- Sūn, Hóngkāi and Liu Guangkun. 2009. *A Grammar of Anong: Language Death under Intense Contact*. Leiden: Brill.
- Sūn, Hóngkāi, Lù Shàozūn, Zhāng Jichuān and Ōuyáng Juéyà. 1980. *The Languages of Monba, Luoba and Deng* [in Chinese]. Beijing: Nationalities Publishing House.
- Sun, Jackson T.-S. 1993a. *A Historical Comparative Study of the Tani (Mirish) Branch of Tibeto-Burman*, PhD dissertation, University of California, Berkeley.
- Sun, Jackson T.-S. 1993b. The linguistic position of Tani (Mirish) in Tibeto-Burman: a lexical assessment. *LTBA* 16.2: 142–188.
- Sun, Jackson T.-S. 2000a. Parallelisms in the verb morphology of Sidaba rGyalrong and Lavrung in rGyalrongic. *Language and Linguistics* 1.1: 161–190.
- Sun, Jackson T.-S. 2000b. Stem alternations in the verb morphology of Sidaba rGyalrong and Lavrung in Qiangic. *Language and Linguistics* 1.2: 211–232.
- Sun, Jackson T.-S. (ed.). 2014. *Phonological Profiles of Little-Studied Tibetic Varieties* (Language and Linguistics monograph series). Taipei: Institute of Linguistics, Academia Sinica.
- Thurgood, Graham. 1974. Lolo-Burmese rhymes. *LTBA* 1: 98–108.
- Thurgood, Graham. 1977. Lisu and Proto-Lolo-Burmese. *Acta Orientalia* 38: 147–207.
- Thurgood, Graham. 1982. Subgrouping on the basis of shared phonological innovations: a Lolo-Burmese case study. *Berkeley Linguistics Society* 8: 251–260.
- Thurgood, Graham. 1985. Pronouns, verb agreement systems, and the subgrouping of Tibeto-Burman. In: Graham Thurgood, James A. Matisoff and David Bradley (eds), *Linguistics of the Sino-Tibetan Area: The State of the Art. Papers Presented to Paul K. Benedict on the Occasion of his 71st Birthday*, 376–400. Pacific Linguistics Series C. No. 87. Canberra: Australian National University.
- Thurgood, Graham. 1992. The aberrancy of the Jiamao dialect of Hlai: speculation on its origins and history. In: Martha Ratliff and Eric Schiller (eds), *Southeast Asian Linguistics Society I*, 417–433. Tempe: Arizona State University Southeast Asian Studies Publication Program.
- Tournadre, Nicolas. 2014. The Tibetic languages and their classification. In: Thomas Owen-Smith and Nathan Hill (eds), *Trans-Himalayan Linguistics: Historical and Descriptive Linguistics of the Himalayan Area*, 105–130. The Hague: De Gruyter.
- Vajda, Edward. 2010. A Siberian link with Na-Dene languages. In: J. Kari and B. Potter (eds), *The Dene-Yeniseian Connection* (Anthropological Papers of the University of Alaska, new series, vol. 5), 33–99. Fairbanks: University of Alaska Fairbanks, Department of Anthropology.
- VanBik, Kenneth. 2009. *Proto-Kuki-Chin: A Reconstructed Ancestor of the Kiki-Chin Languages* (STEDT Monograph 89). Berkeley: Sino-Tibetan Etymological Dictionary and Thesaurus project, University of California.
- van Driem, George. 1997. Dzala and Dakpa form a coherent subgroup within East Bodish, and some related thoughts. In: Roland Bielmeier and Felix Haller (eds), *Linguistics of the Himalayas and Beyond*, 71–84. The Hague: De Gruyter.
- van Driem, George. 2001. *Languages of the Himalayas: An Ethnographic Handbook of the Greater Himalayan Region: Containing an Introduction to the Symbiotic Theory of Language*. Leiden/New York: Brill.
- van Driem, George. 2011. Tibeto-Burman subgroups and historical grammar. *Himalayan Linguistics* [Special issue in memory of Michael Noonan and David Watters] 10.1: 31–39.

- van Driem, George. 2015. A synoptic grammar of the Bumthang language. A language of the central Bhutan highlands, with contributions by Nathan Hill. *Himalayan Linguistics Archives* 6: 1–77.
- Watters, David E. 2003. Some preliminary observations on the relationship between Kham, Magar, (and Chepang). SIL ms.
- Wiersma, Grace. 1990. A study of the Bai (Minjia) language along historical lines. PhD dissertation, University of California, Berkeley.
- Yu, Dominic. 2012. Proto-Ersuic. PhD dissertation, University of California, Berkeley.
- Zhāng, Jichuān. 1986. *Cāngluò Ménbā yǔ jiānzì [A Sketch of Cāngluò Ménbā]*. Beijing: Minzu Chubanshe.

CHAPTER TWO

OVERVIEW OF SINO-TIBETAN MORPHOSYNTAX

Randy J. LaPolla

1 SINO-TIBETAN

At the earliest reconstructable stage of the development of the Sino-Tibetan (ST) language family, possibly as much as 6,000 years ago (Thurgood 1994),¹ the proto-language was monosyllabic. Matisoff (2014) reconstructs the syllable canon as *(P²) (P¹) Ci (G¹) (G²) V (:) (w/y) (Cf) (s).² It is not clear whether the prefixes in some or all cases entailed a vocalic element. If so, the structure might have been sesquisyllabic (e.g. as in the name *tərùng* ‘T’rung/Dulong’, the vocalic element of the *tə-* prefix is very slight).

There was no relational morphology (LaPolla 1990, 1992a, 1992b, 1994b, 1995a, 1995b, 2004, 2012b), but there was derivational morphology in the form of prefixes, suffixes, and voicing alternations of the initial consonants (Wolfenden 1928, 1929; Benedict 1972; Pulleyblank 1962–3, 1972, 1973a, 1973b, 1977–8, 1991, 2000; Bodman 1980; Mei 1980, 1988, 1989, 2012; LaPolla 1994c; Sagart 1999; Sagart and Baxter 2010, 2012; Jin 1998a, 1998b, 2000, 2004, 2005a, 2005b, 2005c, 2006, 2012; Gong 2000; Matisoff 2003; Handel 2012). In §1.1 are examples of several types of derivational morphology.³ Sections 1.2–1.5 discuss other aspects of morphosyntax common to all of ST. Following that are sections that discuss aspects of the morphosyntax unique to Sinitic or Tibeto-Burman (TB).

1.1 Affixal morphology

1.1.1 *s- prefix

The *s- prefix in most cases had a causativizing, denominative, or ‘intensive’ (change of state) function (Wolfenden 1929; Pulleyblank 1973a, 2000; Bodman 1980; Mei 1989, 2008a, 2008b, 2012; Gong 2000, 2001; Dai 2001; Matisoff 2003; Phua 2004; Sagart and Baxter 2012; Handel 2012). For example, Old Chinese (OC) *mjet (滅) ‘extinguish, destroy’ : *smjet (威) ‘to cause to extinguish, destroy’; OC *mək (墨) ‘ink’ : *smək (黑) ‘black’ : Written Tibetan (WT) smag ‘dark’; OC *C-rjəs (吏) ‘clerk’, ‘minor official’ : *srjə? (使) ‘to cause (someone to be an emissary)’, ‘to send’; *tju? (帚) ‘broom’ : *stu? (掃) ‘to sweep’; *lje? (易) ‘to exchange’ : *sljeks (賜) ‘to give’, ‘gift’; WT grib ‘shade’, ‘shadow’ : sgrīb-pa ‘to shade, to darken’; gril ‘a roll’ : sgril-ba ‘to roll together’, ‘to form into a roll’; mnan ‘smell (intr.)’ : snam ‘smell something (trans.)’. There is broad agreement that there was an *s- prefix with these uses, but there is still much controversy as to which forms are due to this prefix and which might be due to other factors, such as a *voiced prefix or *voicing contrasts (see below).

1.1.2 *Voiced prefix and/or *voicing contrasts

In both OC and TB, we find pairs of cognate lexical items which differ phonetically only in terms of the voicing or aspiration of the initial, and differ semantically in terms of transitivity, where the item with the voiced initial is intransitive, and the item with the voiceless initial is transitive (e.g. see §4.2.3 in the chapter on Classical Tibetan⁴ and §2.5.1 of the Tangut chapter). Benedict (1972: 124) discussed this for TB, but argued that in OC no consistent pattern of morphological alternation could be recognized. Most scholars now would see the OC forms as parallel to the TB forms, and part of a cognate phenomenon. Pulleyblank (1973a, 2000) argues these variant forms should be the result of an intransitivizing prefix *ă- (a non-syllabic pharyngeal glide) which voiced the initial of the original transitive roots. Mei (1989, 2008a, 2008b, 2012) includes this prefix in a paradigm with the *s- directive prefix and the *-s direction of action changing suffix (below).⁵ Both Pulleyblank and Mei base the idea for the prefix mainly on the WT *a-chung* ('small a') prefix (here marked with an apostrophe). Pulleyblank also equates this prefix with the *a-* nominalizing prefix found in Burmese. Baxter (1992) adopts this view in reconstructing OC forms, and uses *h- for the form of the prefix,⁶ e.g. *kens (見) 'see': *hkens (> *gens) (現) 'appear/be visible.' While this analysis is attractive from a systemic point of view, Benedict (1972) points out that the prefixing and the voicing alternation in Tibetan are two different phenomena that interact in the specialization of different forms as 'present,' 'perfect,' 'future' and 'imperative,' such that the present and future forms have the voiced initial and are intransitive or durative, while the perfect and imperative forms have the voiceless initial and are transitive or active. As an example, for the verb 'put off, pull off, take off,' we have present *'bud-pa* and future *dbud*, which derive from an intransitive stem **bud*, and perfect and imperative *phud*, which derives from a transitive stem **pud*. Evidence that it is not the *a-chung* prefix that is involved in the contrast in Tibetan is the fact that in many cases both forms of a pair of contrasting forms have the prefix, e.g. Tibetan *'gril-ba* 'to be twisted or wrapped round': *'khrl̥i-ba* 'wind or coil round, embrace.' Bodman (1980: 54) also mentions that he did not find any Tibetan–Chinese cognates where prefixation or lack of it in Tibetan corresponds with the voicing distinction in OC. We also find the voicing alternation in TB languages independent of prefixation, e.g. **kh(r)jok* (曲) 'bend,' 'bent': **hkh(r)jok* (*g(r)jok) (局) 'compressed,' 'bent', 'curved (body)': *Bahing kuk* 'make bent': *guk* 'to be bent' (TB **kuk* ~ **guk*; Benedict 1972: 125). Pulleyblank's association of the voicing distinction in OC with the *a-* prefix in Burmese is also problematic, as the latter is a nominalizer, not an intransitivizer, and is independent of the voicing distinction, e.g. Burmese *phai* 'break off a small piece from a larger,' 'crumble': *pai* 'to be broken off,' 'chipped'; (cf. also Qiang *he-phe* 'tear (clothes)': *de-pe* 'be torn'; TB **pe* ~ **be*; Benedict 1972: 59) (:: OC **phajs* (破) 'to break': **paj?* (跛) 'lame'). Other examples: OC **prats* (敗) 'to defeat': **hprats* (*brats) (敗) 'to be defeated'; **krujs* (壞) 'to destroy,' 'ruin': **hkrujs* (*grujs) (壞) 'to be ruined'; **trjang?* (長) 'grow tall,' 'increase'; 'elder': **htrjang* (*drjang) (長) 'long'; Bodo *bey* 'to be straight': *phey* 'to make straight' (TB **bley* ~ **pley*; OC **brey* (平) 'level').

In a 2010 debate (published in 2012), Mei (2012) argued for the *s- prefix as the cause of the voicing contrasts (also the view of Dai 2001, Gong 2000, 2001, and Phua 2004) and Sagart and Baxter (2012) argued for a *N- prefix as the cause of many of the voicing contrasts said by the others to be due to *s-. Handel (2012) critically evaluated both positions, and argued that neither is the ultimate answer, and so there should be consideration that there may have been several different productive and non-productive morphological

forms, plus analogical leveling, involved in the contrasts found in different periods before and during the OC period.

There clearly were intransitivizing and nominalizing prefixes in PTB and possibly PST, but these are represented by WT *m-* (e.g. *mkho-ba* ‘desirable,’ ‘to be wished for’: *kho-ba* ‘to wish, to want’; Wolfenden 1929: 27—notice the *a-chung* in the active form, showing that that doesn’t have an intransitivizing function). Wolfenden (1929: 26) described verbs marked with *m-* as “of intransitive nature, or which at most describe an act on the part of the subject which does not entail any change in position on the part of such subject,” and connected this prefix with the *m-* found on many substantives (1928). Matisoff (2003: 117) says “the nasal prefix generally signals *inner-directed states or action* . . . PTB etyma like **m-nwi(y)* ‘laugh’, **m-tu:k* ‘spit’, **m-sow* ‘awaken’ . . . it is sometimes found in paradigmatic opposition to the **s-* prefix, which marks *outer-directed action*, transitivity, causativity: e.g. WT *mnam-pa* ‘smell, stink’ (v.i.) vs. *snam-pa* ‘sniff, take a smell of’ (v.t.)” (italics in original). See Matisoff 2003: 87–156 for extensive discussion of the PST/PTB prefixes with exemplification from many languages.

We also find **b-* and/or **g-*, e.g. T’rung *rut* ‘to tear down (a house)’ : *brut* ‘to collapse (of a house)’; *la* ‘to throw (down)’ : *glà* ‘to fall (down)’⁷ (there is also a separate intransitivizing/stativizing *ə-* prefix in T’rung as well: *tāl* ‘roll (vt)’ > *ətāl* ‘roll (vi)’; LaPolla and Yang 2004; see also LaPolla 2000, 2011 on Rawang). These are independent of the voicing alternation, which is also found in Rawang: *zən-ē* ‘follow (intr.)’ : *sən-ōē* ‘follow (trans.)’. In Ronghong Qiang (LaPolla with Huang 2003; LaPolla 2011) we also find both the voicing distinction (*he-phe* ‘tear (clothes)’ : *de-pe* ‘be torn’) and a reflex of the **s-* prefix (*cteə* ‘feed’ < *tchə* ‘eat’ (with assimilation of the prefix to point of articulation of the initial; see also Sun 1981a:192–3 for more examples of each type from the more conservative Mawo dialect)). So my own view is that all three phenomena exist; while some of the voicing distinctions can be shown to be due to either an **s-* prefix or a *nasal prefix, we need to also recognize the possibility that some of the voicing contrasts can’t be explained by either of these prefixes and so are an independent phenomenon.

1.1.3 **?ay-* (**?ə-*) **?ak-* **?a-* (*ə-*) prefix

Matisoff (2003: 104) reconstructs **?a-* **(?ə-*) **?ə-* **?ay-* **?ak-* as a single prefix for PTB (but see his discussion on p. 108 of that work, which leaves the door open to other possibilities). He says the “essential component of the prefix was its initial glottal stop” (p. 105), and so calls it “the glottal prefix.” He sees all the forms as related, and includes the functions of marking kinship (usually as *a-*), third person possessive and subject, nominalization of verbs, stativity or intransitivizing or causativizing of verbs.⁸ All of the forms can appear in a single language, e.g. Lahu (Matisoff 2003: 108–9). In Rawang (LaPolla 2000) the functions are similar, but distinguishable. The nominalizing prefix *ay-* (which has the allomorph *ak-* before voiceless stops) is the third person pronoun and third person possessive prefix, and is used quite productively to form nominals. Some of these have become lexicalized, such as *aydál* ‘fool (n.)’ (< *dál-ē* ‘to be foolish’), *aywəm* ‘lid’ (< *wəm* ‘to cover’). This prefix is actually more of a general formative prefix, and so can be used on some nouns as well, such as in *aytʰi* ‘liquid’ (< *tʰi* ‘water’), and on classifiers, e.g. *ayfʰayri* ‘the trucks’ (< *fʰay* ‘classifier for round or lump-like objects,’ with the plural marker *-ri*). The prefix *ə-*, when used on verbs, is mainly for intransitivizing (*əya?* **ə?* **ənē* ‘fall over’ < *ya?* **ōē* ‘push over’), which includes forming reciprocals (*ayma?* *əʃə?* **nē* ‘They are arguing/fighting’ < *ʃət-ōē*), but is also involved in some deverbal nominals as well, such as *ayiú* ‘one who cries easily’ (< *yū-ē* ‘to cry’; note the tone change, which can mark

nominalization alone) and *əkʰuú* ‘thief’ (<*kʰuū-ə* / *kʰuú-ədə* ‘to steal’), though it is not very productive. The kinship use of *ə-* marks first person possessive (*əpə* ‘my father’ vs. *nəpə* ‘your father’ and *àypə* ‘his father’). In Sinitic it seems we only find the kinship/vocative use (generalized for vocatives), so this might be a separate morpheme.⁹

1.1.4 *-t suffix¹⁰

The *-t suffix most often has the function of transitivizing an intransitive verb, as in WT *hbye-ba* ‘open,’ ‘separate’ (vi) : *hbyed-pa* ‘open,’ ‘separate’ (vt), Rawang *ŋuū* ‘weep’ : *ŋut* ‘mourn,’ ‘cry for someone (vt,’ but in some cases seems to nominalize intransitive verbs, as in WT *yu-mo* ‘weep’ : *yud-mo* ‘a sob’ (see also Benedict 1991), and in still other cases seems not to have had any effect on the valency, e.g. WT *gči-ba*, *gčid* ‘to urinate’; *bka* ‘word,’ ‘speech,’ *skad* ‘speech’ (for other examples and discussion, see Benedict 1972: 98–102; Dai and Xu 1992; Michailovsky 1985; van Driem 1988; Jin 1998a, 2004; Matisoff 2003: 439ff). In OC we find pairs of related forms that differ only in the final consonant, but no clear derivational pattern can be determined, e.g. **nji* (尼) ‘near,’ ‘close’ : **njit* (昵) ‘intimate,’ ‘familiar’; ‘glue’ (from Pulleyblank 1972: 11; this set is cognate with WT *nye* ‘near,’ *nyen* ‘relative’; see also LaPolla 1994c; Matisoff 2003: 439ff).

1.1.5 *-n suffix

The *-n suffix generally had a nominalizing function, e.g. WT *rku* ‘steal’: *rkun-po* ‘thief’; *nye* ‘near,’ *nyen* ‘relative,’ but in some cases seems to have had a collective sense (Benedict 1972: 99ff; Matisoff 2003: 446ff), e.g. Proto-Tibeto-Burman (PTB) **r-mi* ‘person’ : OC **mjin* (民) ‘the people.’ Pulleyblank (1991, 2000) also suggests that Proto-Sino-Tibetan (PST) had a morphological *-n suffix (as well as a *-t suffix), which could explain the correspondences among pairs such as **nja* (語) ‘speak’ ~ **njan* (言) ‘say’, ‘word’ (see also Jin 1998a for more examples). Following Graham (1983), Pulleyblank argues that the *-n suffix marks a durative or continuative aspect, and *-t marks a punctual or perfective aspect. Norman (1988: 86) argues that the forms **njan* (然) and **w(r)jan* (焉) are fusions of **nja* (如) and **w(r)ja* (于) with an *-n- initial pronoun, possibly **nja?* (爾) or **njak* (若). While a demonstrative may have been the ultimate origin of the *-n suffix, it seems this *-n could have been a more general suffix, and not the result of a chance fusion of isolated lexical items. Especially when we see the patterns of variants, it is hard not to assume there was some systematicity to it, e.g. **nja* (如) ‘like’ : **njan* (然) ‘like this’ : **njak* (若) ‘like’, ‘that.’ There is also **?a* (鳥) : **?an* (安) both ‘interrogative pronoun’ (‘where’), and possibly **?ak* (惡) ‘interrogative pronoun’.¹¹

This is not to say there were no fusions. Some variation within word families may be due to a coalescence of two forms, as suggested for Tibetan by Walter Simon (1941, 1942, 1949, 1977). Simon’s idea was that many of the finals in Tibetan, such as -g, -n, -l, -r, -s were from the coalescence of two syllables, the second of which originally also had lexical content, such as -s < *sa/so* ‘place’.¹² We find synchronic variation in Tibetan that points to this kind of development, such as *da-ra* ~ *dar-ba* ‘type of buttermilk,’ *ža-la* ~ *žal* ‘clay,’ *bu-ga* ~ *bug* ‘hole,’ *lco-ga* ~ *lcog* ‘lark,’ *nya-ga* ~ *nyag* ‘steelyard,’ and *yi-ge* ~ *yig* ‘letter.’ Norman (1988: 85ff.) gives the following as examples of fusion words in OC: **tja* (諸) from **tjə* *?ja* (之於) ‘third person patient pronoun’ + adposition ‘in,’ ‘at,’ ‘to’; **njə?* (耳) from **nja* *ljə?* (而已) ‘linking particle’ + ‘end’; **lja* (歟) from **le hwa* (也呼) ‘sentence final particle’ + ‘question particle’; **gap* (盍) ‘negative question (“why not”) particle’ from **gaj pə* (何不) ‘question word’ + ‘negative particle.’

1.1.6 *-s suffix

The *-s suffix generally had a nominalizing function (Pulleyblank 1973b; Mei 1980, 1989, 2012; Matisoff 2003: 439ff), where the derived noun is the patient of the action represented by the verb, but also had a function that Mei (1980, 1989) and Schuessler (1985) have characterized as ‘change of direction’ or ‘inversion of attention flow’ respectively. Mei (2012) suggests these two functions derive from a marker that was used originally for perfective aspect but was then extended to marking nominalizations of the result of the action. In Modern Chinese this suffix is now reflected in the ‘departing’ tone.¹³ In some cases, the addition of the suffix resulted in the creation of a new Chinese character, but in many cases there are simply two pronunciations for the same character. For example OC *C-rjang (量) ‘measure’: *C-rjangs ‘an amount’ :: WT ‘grang-ba’ ‘to number,’ ‘to count’: grangs ‘a number’; OC *tjək (織) ‘weave’: *tjəks ‘thing woven’ :: WT ‘thag-pa’ ‘to weave’: thags ‘texture,’ ‘web’; OC *nup (納) ‘bring in’: *nups (內) ‘inside’; *mreʔ (買) ‘buy’: *mres (賣) ‘sell’; *djuʔ (受) ‘to receive’: *djud (授) ‘to give.’

1.1.7 *-j suffix

Matisoff (1989, 1995, 2003: 482ff) discusses etyma that show palatal-final and non-palatal-final variants, and posits three different sources for variants with morphological differences: PST *s-waj ✕ *s-jaj ‘go’, ‘motion away,’ for transitive motion/motion away from the deictic centre or emergent quality in stative verbs; PST *ja (✕ *za ✕ *tsa ✕ *dza) ‘child,’ ‘son’ for a diminutive or affective sense; and PST *waj ✕ *raj copula for nominalization, subordination, or other grammatical functions. The clearest examples are in the system of pronouns, where for the first person pronoun we get PTB *ŋa : ŋaj :: OC *ŋa (吾): *ŋaj (我).

1.1.8 *-ʔ suffix

OC seems to have had a glottal stop suffix which developed into the rising tone category (Baxter 1992; Zhengzhang 2013), e.g. *trjang (張) ‘to make long,’ ‘stretch’ : *trjang? (長) ‘grow tall,’ ‘increase’, ‘elder’; *wək (或) ‘someone’ : *wjək? (有) ‘there is’; *kak (各) ‘each’ : *k(r/j)ak? (舉) ‘all.’ In these last two examples I am assuming that the suffix caused the loss of the root final consonant, just as is assumed to have happened with the *-s suffix (Baxter 1992: 323ff; cf. also Bodman 1980: 132), but this assumption is not widely accepted. An alternative possibility, discussed immediately below, is that there was a *-k suffix. Glottalized forms do appear in some TB languages, e.g. T’rung, but it is clear at least in T’rung that these are developments from *-k.

1.1.9 *-k suffix

There may have been a *-k suffix as well, as we find a large number of lexical items in both TB and Sinitic that have open final and *-k final variants, e.g. TB *yu(w) : *yuk ‘descend’ (Benedict 1972: 101); OC *m(r)ja (無) ‘there is not’ : *mak (莫) ‘no one’; *djuʔ (誰) ‘who’ : *djuʔ (孰) ‘which one.’ This possibility was suggested by Pulleyblank (1972: 13, 1973a: 122) as an explanation for some of the pairs given above as examples of the glottal stop suffix: *wjəʔ (有) ‘there is’ : *wək (或) ‘someone’; *k(r/j)ak? (舉) ‘all’; *lift : *kak (各) ‘each.’ Pulleyblank only discusses this in relation to pronominal forms, and says the suffix marks a distributive sense. There is also the set *ŋja (如) ‘like’ : *ŋjak (若)

‘like’, ‘that’ mentioned above. As the largest number of variants involve the difference between an open final and a *-k final (63 out of 99 rhymes in the *Book of Poetry* where the finals differed, as marked in Wang 1980; see LaPolla 1994c for discussion), it may be that there is more than one explanation; some velar stop finals may have dropped due to the influence of the glottal stop suffix, and some may have been the result of a *-k formative suffix (see also Jin 1998b).¹⁴ If PST had a particle similar to Tibetan -ga, which Das (1902: 203) says “is sometimes used as an affixed particle of a word to complete it,” then this would be at least one explanation for the large number of *-Ø ~ *-k variants.

It has long been known that within ST we must deal with word families rather than isolated words (Karlgren 1933, 1956; Wolfenden 1936, 1937, 1939; Wang 1982; Zhang 1999). Given what we now know about these derivational processes, we can see clearly how the word families are created. These forms seem to have formed paradigms (sets of choices), but of derivational possibilities rather than inflectional possibilities. Following are two examples (from Baxter 1992: 317 and 324 respectively; see also Mei 1989; Matisoff 2003: 440): *kat (割) ‘to injure,’ ‘to harm’ (vt) : *hkat (*gat) (害) ‘to suffer harm or injury’ (vi) : hkats (*gats) (害) ‘harm,’ ‘injury’ (n); *trjang (張) ‘to make long,’ ‘stretch’ (vt) : *trjang? (長) ‘grow tall,’ ‘increase’; ‘elder’ (intransitive active verb) : *htrjang (*drjang) (長) ‘long’ (stative verb). To this last set Mei (2012) adds (脹) *trjangs ‘distended, swollen’ (perfective).

Aside from the suffixes mentioned above, Mei Tsu-lin (personal communication, November 1994) has suggested some of the frequent variations found in Chinese between homorganic stop and nasal final might be due to Chinese having had suffixes similar to WT -ma and -pa (which have both gender marking and formative functions). The nasal-initial suffix would cause a final stop to nasalize, while the stop-initial suffix would denasalize a final nasal. We see this sort of development with the diminutive in some dialects of Chinese, where the diminutive suffix reduces to a nasal element (e.g. in Wenzhou, and some areas of Anhui, Zhejiang, Guangxi, and Guangdong), and in some cases nasalizes final stops, e.g. in Xinyi of Guangdong, the nasal suffix -n causes final -p, -t, and -k to become -m, -n, and -ŋ respectively, as in *ap*³³ ‘duck’ > *am*³⁵ ‘duckling.’ Certainly the use of reflexes of PTB *pa (and to a lesser extent *ma) as a gender marker and as a nominalizer (usually producing an agentive noun) is widespread throughout TB, though there is the possibility that many of these were independent parallel developments, such as in the case of the frequent development of diminutives from a word meaning ‘son’ or ‘child’ (Matisoff 1995), and of causatives from a word meaning ‘make,’ ‘cause,’ or ‘send’ (LaPolla 1994b). In Chinese the form *p(r)ja(?) (夫/父/甫) was used as an extra-syllabic suffix for creating agentive nouns, just as in TB (e.g. *din p(r)ja? (田父) ‘farmer’), and this may be the cognate of PTB *-pa.

1.2 Clausal morphology

In terms of clausal morphology, there may have been a clause-final question particle *la, as there is evidence for such a particle in several languages across the family: OC *ʃja (歟), Newar lā, Burmese lá, Meithei la (Matisoff 1995: 73–4). Matisoff (2003: 599) reconstructs PTB *la-j. As mentioned above, though, the Chinese form has been said to be a fusion form, from *le hwa (也呼) (Norman 1988: 95), so may not be cognate with the PTB form.

Unmarked clausal negation in PST took the form of a preverbal particle *ma-j (Matisoff 2003: 488). For PTB we can also reconstruct a prohibitive (negative imperative) particle *ta- ~ *da (Matisoff 2003: 586; see ex. (1) below for a Lahu example), but this is not

found in the Sinitic languages. OC instead had two negative imperative particles **mja* (毋), which was homophonous with the unmarked negator but written with a different character, and **mjat* (or **mjut*) (勿), which is often assumed to be due to fusion of the negative **mja* with another particle (assumed to be the demonstrative pronoun **tjə* (之)).¹⁵

Most languages in the family have not grammaticalized grammatical relations, but many have grammaticalized semantic role marking.¹⁶ For detailed arguments against the existence of syntactic functions in particular ST languages, see Andersen 1987 (Classical Tibetan), Bhat 1991 (Manipuri), and LaPolla 1990, 1993b (Chinese). See also the discussions of Lisu in Hope 1973, 1974 and Mallison and Blake 1981. Benedict (1972: 95ff) also expressed the view that relational morphology was not part of the grammatical system of PTB. A corollary of the fact that very few languages have grammaticalized grammatical relations is that there are few true passive constructions in the family. As the order of NPs is generally determined by pragmatic factors, variations of word order can affect the interpretation of utterances in a way similar to the effect of passives.

1.3 Pronouns

First and second person pronouns are reconstructable to the family as a whole (first person **ŋa* ✕ *yaj*; second person **na* ✕ *nay*), with both variants of each pronoun found on both sides of the family (see LaPolla 1994c; Matisoff 2003), but there is no third person pronoun or plural marker reconstructable to the PST or PTB or even Sinitic level. The deictic pronouns are reconstructable, though (**ndaj* ✕ **ndi*), and they became the source for many of the third person forms. A survey of inclusive and exclusive forms in TB (LaPolla 2005b) showed that the inclusive forms were innovative, and the exclusive form in such systems generally involved the inherited pronoun.

1.4 Classifiers and definite marking

Classifiers were not part of PST, but evolved individually in quite a few of the languages in the family (see, for example, Xu 1987, 1989; Dai 1994, 1997a, 1997b for prosodic reasons for some languages developing classifiers, and some not). Even within some of these groups, such as the Lolo-Burmese languages, while the nouns they derive from may be cognate, the use of the nouns as classifiers is recent (see Bradley 2012 for discussion). There was no definite marking in PST, and only a few languages in TB, such as Qiang (see LaPolla with Huang 2003), have developed something that can be considered as definite marking (in the case of Qiang, the marking seems to have developed from demonstratives). Several languages that have developed classifiers, both in TB and in Sinitic, have developed a use of the classifiers that resembles definite or specific marking. This generally involves use of the classifier without a numeral, e.g. Rawang *lègā tiq bok* [book one CL] ‘one book,’ *lègā bok* ‘the book,’ Cantonese *yat⁵⁵ ga³³ che⁵⁵* [one CL vehicle] ‘one car,’ *ga³³ che⁵⁵* (roughly) ‘the car.’ This feature is an areal feature of part of Southeast Asia (Baron 1973).

1.5 Constituent order

In terms of constituent order (LaPolla 2002, 2015), all ST languages have GENITIVE–HEAD order and MODIFIER–MODIFIED order in N-N structures (the former is actually a subcase of the latter in PST). All ST languages have RELATIVE–NOUN order (Karen also has a less

productive post-nominal relative clause—Solnit 1997: 249ff, this volume). Originally there were no nominalizers or relative markers in relative clauses, but various languages have developed one or the other since that time (see LaPolla 2008a, 2008b for discussion of the process). In cases where the relative clause is nominalized, this construction then is also a subcase of the N-N modifier–modified construction (see LaPolla 2013c for the example of Mandarin and the implications of such a structure). It seems the original position of attributes was after the head, but in many languages (e.g. Burmese), the attribute can be nominalized and appear before the head. This then becomes another subcase of the N-N modifier–modified construction. The overwhelming majority of ST languages have NEGATIVE–VERB order, and where there is a deviation from this, the pattern is either due to reinforcement of the original negative, as in Karen, or due to the grammaticalization of a post-main-verb negative verb out of a negative-auxiliary verb combination. We can therefore assume MODIFIER–MODIFIED order in N-N structures, and GENITIVE–HEAD, HEAD–ATTRIBUTE, NEGATIVE–VERB, and RELATIVE–NOUN word order patterns for PST.

The TB languages generally have verb-final word order with an immediately preverbal unmarked focus position (including in many cases for interrogative words, e.g. in Tangut). At present, the Sinitic languages, the Karen languages, and Bai have an unmarked post-verbal focus position (rather than an immediately preverbal unmarked focus position as in the other languages), and so the patient argument often appears in post-verbal position in the clause.¹⁷ From the fact that we can clearly see changes in the word order of these three languages over time, and cannot see such changes in the TB languages other than Bai and Karen, we assume that it was Bai, Karen, and Chinese that changed rather than all the other TB languages (but see Wheatley 1984, 1985 for possible ongoing changes in the Yi language). As argued in LaPolla 1993a and 2015, these three languages show a remarkable similarity in the particular patterns they developed. In OC, verb-medial order (which implies a post-verbal position for unmarked focus) was the unmarked word order, but there was a marked verb-final word order pattern used for contrastive focus that seems to have been due to an earlier preverbal focus position. In Karen and Bai, we have the same situation as in OC in terms of the major constituents: unmarked verb-medial order, but NP-NP-V as a marked word order possibility. What is significant is that the conditions on the use of the marked word order pattern in Bai are almost exactly the same as those of OC: it is used when the second NP is a contrastive pronoun or when the sentence is negative or a question (Xu and Zhao 1984). Also interesting about the use of the different word order patterns in Bai is the fact that the older people prefer the verb-final order, whereas the younger and more Sinicized people prefer the verb-medial order (*ibid.*). This would seem to point to the change in word order as being relatively recent. Karen (Solnit 1997, this volume) has some similar word order patterns, with genitives and nominal modifiers coming before the noun, and number and classifier following the noun, while adjectival and verbal modifiers (i.e. relative clauses) can follow the head. Karen does not appear to have a preverbal focus position; from the data in Solnit 1997, it seems that focus position is sentence-final as in Modern Chinese. In terms of phrase-internal order, Karen is very similar to OC, differing mainly in terms of having HEAD–ATTRIBUTE order as the unmarked word order, as opposed to Chinese, which has it only as a marked order. Karen and Bai differ from most of the rest of the TB languages mainly in terms of the position of the NP representing the undergoer referent and in terms of having prepositions. Based on the relative frequency of patterns and patterns of change witnessed in some languages, we can assume PST also had the following word order patterns: DEMONSTRATIVE–HEAD, HEAD–NUMBER, NOUN–ADPOSITION, and STANDARD–(MARKER)–ADJECTIVE (see LaPolla 1993a, 1994a/2002, 2015; also Dryer, this volume).

2 SINITIC

In OC, there was a gradual loss of productivity of the derivational morphology sometime around the formation of the characters (roughly 3,500 years ago), and the language became more isolating. A gradual change occurred in the word order and information structure pattern to verb-medial word order and post-verbal focus position (LaPolla 1993a, 2015). There has been no grammaticalization of grammatical relations; the basic structure of the clause is topic–comment rather than subject–predicate (Chao 1955 [1976], 1959 [1976], 1968; Lyu 1979; LaPolla 1990, 1993b, 2009). Information structure is the chief determinant of word order in Chinese (LaPolla 1995c; LaPolla and Poa 2005, 2006). The prepositions now found in the language all derive transparently from verbs. In the past there was an assumption among Chinese linguists that the grammar of all the dialects is roughly the same, and so until recently little serious work was done on the grammar of the dialects. With the work by Yue-Hashimoto (1993, this volume), Huang (1996), Chappell (2001, 2004), and Liu (2008, 2013), serious investigation of the grammar of the dialects has begun, but much more needs to be done to understand the differences between the dialects, particularly in more difficult-to-understand areas such as information structure and its relation to grammatical structure (compare LaPolla 1995c, on Mandarin, and Lee 2002, on Cantonese, in this regard). As Yue (this volume) and Ho (this volume) give us an overview of modern dialect grammar, I will not say more about this, and devote the rest of this chapter to the TB languages.

3 TIBETO-BURMAN

After the split-up of ST into Sinitic and TB due to the divergent migrations (LaPolla 2001), there were a number of developments in the realm of grammar, some of which have areal coverage, some of which are subgroup specific. I will first discuss the different groups and some of their characteristics, and then some more general morphosyntactic phenomena.

3.1 Language groups

Based on morphological paradigms and migration history (LaPolla 2001, 2013a, 2013b), I divide TB into the following groups.¹⁸

The Bodish group: Tibetic¹⁹ and the other languages, such as Tamang, Gurung, Lepcha, Dzongka, and Newar (Newari), derived from the original migrations west into Tibet and then later migrations south down into Nepal, India/Sikkim, and Bhutan; in terms of morphology, this group is characterized by an *-s ablative/ergative suffix on nouns (see LaPolla 1995a). Non-classificational morphological features include development of evidentiality and conjunct–disjunct²⁰ systems in many varieties (see, for example, Hale and Watters 1973: 207ff. on Jirel, Newar, and Sherpa; Hale 1980; Hargreaves, this volume, on Kathmandu Newar; DeLancey, this volume; Tournadre and LaPolla 2014 and references therein on Tibetan), and a lack of a bound pronominal person marking system²¹ or reflexive/middle marking.

The Qiangic group: Qiang, Pumi, Shixing, Ergong, Daofu, Queyu, Guiqiong, Muya, Namuyi, Zhaba, and possibly a few others, the speakers of which migrated only a short distance from the original ST homeland in Northwest China (these languages are now spoken in Sichuan and Yunnan Provinces, China). The genetic rather than area affiliation of this group has been called into question (e.g. Chirkova 2012a; LaPolla 2005a), as the

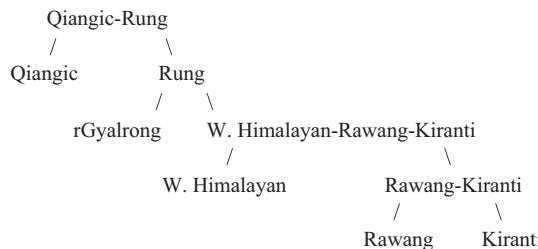


FIGURE 2.1 THE SUBGROUPING OF QIANGIC-RUNG (FROM LAPOLLA 2005A: 394)

similarities seem to be more typological than genetic, and at least two of them, Shixing and Namuyi, have been argued to be closer to Naxi (Chirkova 2012b). Languages of this group characteristically have a set of five to ten directional prefixes on the verb, marking action up, down, up-river, down-river, inward, outward, towards the speaker, away from the speaker, and sometimes towards the mountain, away from the mountain, although the actual forms of the systems in different languages do not all correspond in any clear way (see Sun 1981b; Huang 1991; Chirkova 2012a), and so may be considered an areal phenomenon (Shirai 2009; Chirkova 2012a). They have cognate person marking systems which often have an actor–non-actor contrast (as opposed to a hierarchical system as in many other TB languages—see LaPolla 2010 on the hierarchical system in Rawang). The exception would be Tangut, if this language is in fact to be included in the Qiangic group (Sun 1991, 2001), as the very simple person marking system there is clearly hierarchical (see Ebert 1987; LaPolla 1992a). It may be that the Qiangic system was originally hierarchical and later developed into an actor/non-actor system, as it seems this system may be related at a very deep time depth to the system of the Rung group (see Figure 2.1 and discussion of the Rung group below), which is clearly hierarchical. These languages generally have evidential systems, but it is not clear if there is any cognacy among the systems. The case markers fill similar categories, but generally are not cognate. (See Sun 1982, 1985, 2001; Huang 1991 for more on this group.) Sun (1982, 1985, 2001), who first established the Qiangic group as a group, includes rGyalrong as part of the group, but the relation of rGyalrong to the Rawang and Kiranti groups is much clearer than that to the Qiangic group. The similarities rGyalrong shares with Qiangic are at a deeper level and there is also areal influence.

The Rung group: rGyalrong, T'rung²² (Dulong), Rawang, Kiranti, Kham, and Western Himalayan (Byangsi, Darma, Chaudangsi, Kinnauri), languages that (except for rGyalrong) migrated down along the eastern edge of the Himalayas and then across Burma and into Northern India and Nepal.²³ These languages have clearly cognate complex person marking systems, and all but rGyalrong have a *-si reflexive/middle marking suffix on the verb (adapted from LaPolla 2013b):

	<i>1sg</i>	<i>1pl</i>	<i>2pl</i>	<i>dual</i>	<i>refl/middle</i>
Proto-rGyalrong	*-ŋ	*-i	*-ñ	*-tsh	
Proto-Dulong-Rawang	*-ŋ	*-i	*-n	*-si	*-si
Proto-Kiranti	*-ŋ	*-i	*-ni	*-ci	*-nsi
Proto-W. Himalayan	*-g/ŋ	*-ni	*-ni	*-si	*-si

Within this group then, there is a branching where the languages other than rGyalrong split off from rGyalrong, as it does not manifest the innovation of the reflexive/middle

marking. This accords well with the migrations assumed. A second branching of Western Himalayan off from Rawang and Kiranti is assumed, as the latter two share the innovation of a non-first-person-actor marking prefix.²⁴

The Karenic group: The Karen were one of the earliest groups to migrate down into Burma along the river valleys. As the earliest migrants into Mon and Tai territory, this group has been greatly influenced by the latter two languages. Most striking is the verb-medial word order, prepositions, and post-nominal relative clauses (see Kato, this volume; Solnit, this volume).

The Kuki-Chin group: Now straddling the India–Burma–Bangladesh borders, the speakers of these languages closely followed the Karen down the eastern edge of the Tibetan plateau and into Burma, but went more westerly and so had less contact with Mon and Tai. This group has also innovated person marking, but independent of the system found in the Rung group. In the Kuki-Chin system we find the Proto-Kuki-Chin pronouns **kai* ‘1sg,’ **nay* ‘2sg,’ and **a-ma* ‘3sg’ grammaticalized into the person marking prefixes **ka-*, **na-*, and **a* respectively (Thurgood 1985).

The Lolo-Burmese group: This group came down along the same path as the Karen and Kuki-Chin but at a later time and displaced them in many areas in Burma. They are now stretched from Sichuan and Yunnan Provinces in China (the Yi languages (Nasu, Nosu, etc.), Lisu, Zaiwa, Langsu) down along the migration path to southernmost Burma. Lisu is now found in Northeast India as well. Within Lolo-Burmese, some Loloish languages have been greatly influenced by contact with Tai (e.g. Lahu—Matisoff, this volume), while Burmese has been more influenced by Mon (see Bradley 1980; Wheatley, this volume). Within this group there is no relational morphology that can be reconstructed to the PLB stage. They do not show person marking, and most adpositions and auxiliaries are recent transparent grammaticalizations (see, for example, Matisoff 1991b; see also Bradley 2012 on the phonetic and morphosyntactic characteristics of this group).

The Bai language shared the same origin and territory as the Lolo-Burmese (initially Sichuan and later Yunnan Province in China), but broke off from the main TB group culturally (aligning themselves culturally with the Chinese), forming what was known in Chinese in the eighth century as the *Bái Mán* (White Barbarians), in contrast to the *Wū Mán* (Black Barbarians), the rest of the Lolo-Burmese, who were not as Sinophilic. Because of the cultural orientation of the Bai people, the Bai language came to be heavily influenced by Chinese, and now the lexicon is comprised largely of Chinese loanwords, and the word order is now verb medial (see Lee and Sagart 2008; Matisoff 2008). See Wang (2013) for an alternative view of the affiliation of Bai.

The Tani group: Sun (1993a: ch. 5, 1993b; see also Post and Sun, this volume) argues convincingly that the Tani group (formerly called Mirish or Abor-Miri-Dafla, including the languages of the Adi, Nishi, Bengni, Apatani, and Mishmi peoples) constitutes a separate branch of TB at the highest level. Thurgood (1985: 397) shows there is a high degree of uniformity in the case marking systems of the languages. I have little information about their migration to Southern Tibet and Northeastern India, only anecdotal information about the members of this group now in Arunachal Pradesh having come across northern Burma.²⁵

The Sal or Bodo-Konyak-Jinghpaw group: This group, which includes the Luish/Askaian (Sak/Cak: Bernot 1967, Luce 1985, Huziwara 2008a, 2008b, 2010; Kadu/Kanan: Brown 1920, Sangdong 2012), Bodo-Garo (Burling, this volume, on Garo), Koch, Konyak, and Jinghpaw/Singpho (Kurabe, this volume) languages (Burling 1983; Post and Burling, this volume; see Matisoff 2013 for Askaian-Jinghpaw comparisions), was given central importance by Benedict (1972: 6) partly because of its central geographic

location. There are early Chinese records that seem to point to the Jinghpaw having been in northern Burma in the early part of the current era, but there is nothing definitive on their time of arrival. A number of linguists have grouped Rawang and Dulong (the so-called Nungish languages) with Jinghpaw, but I do not find a pattern of shared innovations that would lead to seeing them as forming a group. While Jinghpaw does have a person marking system, it is not cognate with that of the Rung group. Resemblances between the languages seem to be due to shared retentions rather than innovations, or due to long-term contact.

Aside from these genetic groupings, and a split in prosodic type between a southeastern iambic stress area and a northern trochaic stress area due to contact with Austro-Asiatic languages in the southeast,²⁶ there are two other broader areas of language contact, the Indo-sphere and the Sino-sphere (Matisoff 1990, 1991a). These terms refer to whether the languages are more influenced by Indic languages and culture, or by Sinitic languages and culture. There are certain features that we frequently find in languages in the Indo-sphere that we do not find in the Sino-sphere. In phonology we find, for example, the development of retroflex stop consonants. In syntax we find, for example, post-head relatives or correlatives of the Indic type (relative clauses are generally pre-head and without relative pronouns in ST languages). In Sino-spheric languages we often find the development of tones. Contact with Chinese can also result in monosyllabicity and an isolating structure (the most extreme example of this is Vietnamese).²⁷

3.2 Person marking

Several branches of TB have independently innovated person marking, possibly due to areal influence (LaPolla 1992a, 1994b, 2001). The marking develops from copies of the free pronouns becoming prefixed or suffixed to the verb. Even groups that do not normally have person marking systems, such as Karen, have recently developed such systems in some dialects (see, for example, the Delugong dialect of Sgaw Karen discussed in Dai *et al.* 1991).²⁸ The pattern discussed most often is that of the Rung group, because of its wide geographic distribution along the edge of the Tibetan plateau. This pattern seems to have developed out of the Tangut pattern (1sg *-ŋa², 2sg *-na² (the same forms as for the free pronouns), first and second person plural ni²),²⁹ though the pattern found in the Rung languages expanded greatly after the split of Qiangic and Rung (see Figure 2.1). Attempts to associate the Rung pattern with other patterns in the family and reconstruct it to PTB have been unsuccessful (see LaPolla 1992a, 2012b for discussion).

3.3 Multiple existential verbs

In a number of unrelated languages we find a pattern of multiple existential or locative verbs, with the difference between them being, if there are only two, as in Idu (Sun 1983: 72) a difference between an animate (Idu *i*⁵⁵) and an inanimate (Idu *kha*⁵⁵) referent. A language may have as many as seven different verbs with distinctions between the verbs being of the type animate vs inanimate, abstract vs concrete, location within a container vs location on a plane, and others. For example, Hani has a general existential *dza*³³, an existential for people and animals *dzo*⁵⁵, an existential *bo*³³ for people and their organs, *ds*³¹ for liquids, *de*³¹ for general animates, *kx*³¹ for existence within a group, and one existential verb, *ss*⁵⁵, which is used only in the poetic language (Li and Wang 1986: 54). In Queyu there are seven existential verbs (Wang 1991: 61): *tji*⁵⁵, for animals; *tey*¹³, for location in a vessel or certain area; *bo*³¹, for non-movable objects; *ei*¹³, for movable

objects; *lo^{l3}*, for an object mixed up in another object; *ru^{l3}*, for abstract objects; and *tse^{l3}*, for possession by a person. In Zaiwa (Xu and Xu 1984: 80–1) there are six existential verbs, two of which are specialized for animate beings and can be causativized: *nji^{s1}*, which seems to mark the existence or long-term location of animate beings and has the causative form *nji^{s1}*; *luy^{s5}*, for short-term location of animate beings and has the causative form *luy^{s5}*; *vo^{s5}*, for possession by a person; *tso³¹*, for inanimates; *po^{s1}*, for containment within a vessel; and *toy^{s1}*, for roads and footprints. Other languages that have this feature are Jinghpaw, Apatani, Tamang, Naxi, Nusu, Pumi, rGyalrong, Qiang (this volume) and most of the other Qiangic languages. While some of the categories of existential verbs correspond among the languages, particularly within Lolo-Burmese, such as ‘containment in a vessel or area’ (Hani *tey^{l3}*, Zaiwa *po³¹*), ‘possession by a person’ (Hani *tse^{l3}*, Zaiwa *vo^{s5}*), the forms used in these languages are clearly not cognate.

3.4 Causative constructions

The PST **s*- causative prefix and voicing alternations are no longer productive in most TB languages, and so languages throughout the family (in more than 80 languages and dialects I have counted) have innovated analytical causatives, usually by serializing a verb meaning ‘send on an errand,’ ‘entrust with a commission,’ ‘make,’ ‘do,’ or ‘give’ to create a causative construction (Matisoff 1976, 1991b; LaPolla 1994b). For example, in Lahu the verb *ci* ‘send on an errand’ is used to create causatives, as in *Johnny tha? qay-ci-ve* [OBJ go-CAUSE-PART] ‘Make Johnny run’ (Matisoff 1976: 418). Though occasionally different languages will use cognate verbs to form such causatives (e.g. Lahu and Burmese), the pattern cannot be reconstructed to even some of the lower (e.g. the Proto-Lolo-Burmese) levels; it must have been independently grammaticalized in each of the languages (Matisoff 1976). Even among the very closely related languages and dialects of Northern Burmish we find radically different forms used for causative marking: Longchuan Achang *xu^{s5}*, Xiandao Achang *gəy³¹*, Bola *nɔ̄^{s5}*, and Leqi/Langsu *lɔ̄^{s5}*. In each case we have the independent grammaticalization of a free verb into a post-verbal causative marker.

3.5 Benefactive constructions

Another commonly found development among TB languages is the grammaticalization of a benefactive construction. This most commonly takes the form of an auxiliary verb derived from a verb meaning ‘to give,’ as in Jinghpaw (*-t'a³³*), Tamang (*pin*), Tshangla (*bi*), Camling (*bi*), Belhare (*-per*), and Lahu (*pî*; for third person benefactives; Matisoff 1991b). As can be seen from these examples, the verb used in this construction is often the PST verb **bij* ‘give,’ but the constructions themselves were independently innovated. A fully morphological benefactive such as is found in Rawang, where the suffix *-ā* has an applicative benefactive function (LaPolla 2000; e.g. *rí-ā-ð-ē* (carry-BEN-TR.NPAST-NPAST) ‘(He) is carrying (something) for him’) is rare (see LaPolla and Yang 2007 for the possible origin of this marker).

3.6 Semantic case marking

As mentioned above, there is no relational morphology that we can reconstruct to the PST stage, but there has been grammaticalization of different types of adpositions in every branch of the family (see Hale and Watters 1973; DeLancey 1984; LaPolla 1994b, 2004). These adpositions are also often used for subordinate clause marking (Genetti 1986,

1991; Ebert 1993). There is a regular path for the development of adpositions in the family, where locational markers first develop, then these are extended in use to cover other types of relation, in a predictable way along two different paths: ablative > instrumental > manner adverbial > agentive > anterior or causal clause subordinator (see Hargeaves, this volume for one example of the full set); locative > dative > patient > purposive, temporal, or conditional clausal subordinator (LaPolla 1995b). Large-scale surveys of agentive marking (LaPolla 1994b, 1995a) and ‘object’ marking (LaPolla 1992b, 1994b) were carried out, and the results indicate that although 106 languages (out of 145 surveyed) have an agentive marker, and such a marker can be reconstructed to some of the lower-level groupings within TB, such as Proto-Bodish, there is no form that cuts across the upper-level groupings to the extent that it could be reconstructed to PTB. The conditions on the use of agentive marking in each language were also surveyed. The results point to the existence of at least two major types of ‘ergative’ marking in TB: systemic and non-systemic (or ‘paradigmatic’ and ‘non-paradigmatic’). Non-systemic marking can be seen as a relatively recent development, and has the same function as ‘anti-ergative’/‘anti-agentive’ marking (LaPolla 1992b, 1994b), i.e. disambiguation of two potential agents. It is used only when needed for this purpose and does not pattern paradigmatically, so is unlike what is normally referred to as ‘ergativity.’ Systemic ergativity is much more complex, often involving semantic and pragmatic functions beyond simple disambiguation (see, for example, Genetti 1988; Nagano 1987; Tournadre 1991, 1996; Saxena 1991; Chelliah 2009, forthcoming; Donohue and Donohue 2016). Though discussed as two types for expository purposes, these two types, as they are manifested in TB, are actually points on a continuum of types from completely non-systemic to fully systemic, with movement along the continuum corresponding to degree of grammaticalization.³⁰

From the survey of ‘object’ marking in TB (LaPolla 1992b), it was found that out of 126 languages surveyed, 22 languages had no nominal object marking, 20 languages had nominal morphology consistently marking the patient as object, regardless of whether the clause included another non-agent argument (i.e. was either transitive or ditransitive), and 84 languages, from a broad spectrum of languages in all sub-branches and areas of TB, had a type of marking where the patient in monotransitive clauses is often or always marked with the same postposition as the recipient, beneficiary, or other non-actor argument in ditransitive clauses. For example, in the Lahu examples below (Matisoff 1973: 156–7), *thà?* marks a patient argument in (1a), but a recipient argument in (1b).

- | | | |
|-----|--|---|
| (1) | a. <i>ŋà thà? tâ dɔ?</i> .
1sg OBJ NEG.IMP hit
‘Don’t hit me.’ | b. <i>li? chi ŋà thà? pi?</i> .
book that 1sg OBJ give
‘Give me that book.’ |
|-----|--|---|

I refer to this type of marking as ‘anti-ergative’ or ‘anti-agentive’ marking, as the crucial function of this type of marking is to mark an animate argument that might otherwise be interpreted as an actor being something other than an actor. In this way it is the opposite of the type of ergative marking we find in some of these same languages, which marks an argument as being an actor.³¹ In those languages that have both types of marking, it is often optional whether to use one or the other or both, but the marking is often not systemic, as it is used only to disambiguate two arguments when that becomes necessary due to the semantics of the referents, the actions involved, or the pragmatic viewpoint (see, for example, Matisoff 1973: 155–8 on Lahu *thà?*, Wheatley 1982 on Burmese *kou*). It is especially common for overt marking (either ergative or anti-ergative) to be necessary when the most natural (unmarked) topic, the agent, is not the topic, and instead appears in the preverbal focus position.

Most of the languages have grammaticalized different morphemes to mark anti-ergative arguments, and so while it is possible to reconstruct forms for some low-level groupings such as Tani or Tibetan, in other branches even closely related languages have different anti-agentive markers (e.g. Lahu (*thà?*), Akha (*āŋ*)), or differ in terms of having anti-agentive marking or not (e.g. Akha, which has anti-agentive marking, and Hani, which does not). We can assume that this marking is not of great time depth.

Those languages that have postpositions, but do not have the anti-agentive marking pattern (e.g. Tujia, Hani) generally mark NPs by strictly semantic principles. That is, a locative/goal (when marked) will always be marked the same way, and a patient/theme (when marked) will always be marked the same way, and there are no relation changing (or ‘promotion’) rules (e.g. passive, dative, antidative). We then have two types of role marking in TB. Both are semantically based, but one (agentive and patient marking) is based on what semantic role a referent has, and the other (anti-agentive marking) on what semantic role a referent does not have. The development of both types of marking can be said to be related to the importance of semantic role, pragmatic viewpoint, and animacy to the users of these languages.

3.7 Evidential marking

Evidential marking, the marking of how one came to know the information one is reporting in making a statement (e.g. seen with one’s own eyes, heard from someone else, inferred—see Aikhenvald and LaPolla 2007; Tournadre and LaPolla 2014) has grammaticalized in quite a few languages within TB. The systems may be as simple as having only a contrast between hearsay and non-hearsay (e.g. Rawang, where the hearsay particle *wā* is derived from the verb ‘say’), to more complex systems, as in different varieties of Tibetic (DeLancey 1986, this volume; Woodbury 1986; Sun 1993c; Hongladarom 1993). Some other languages which have evidential marking are Qiang (LaPolla 2003, this volume), Kathmandu Newar (Hargreaves 1983, this volume), and Akha (Hansson this volume; Egerod 1985; Thurgood 1986).

3.8 Reflexive/middle marking

Reflexive marking of different types, using reflexive pronouns or verb suffixes, is found throughout the family, but a small number of languages have independently innovated patterns like that found in French, where marking that was originally used only for true reflexives has been extended to middle voice situations (i.e. situations where there is no clear distinction between the ‘doer’ and the one ‘being done to’; LaPolla 1996, LaPolla and Yang 2004). One pattern found was mentioned above. This is the **-si* suffix found in the Rawang, Kiranti, Kham, and Western Himalayan languages. For example, in Dulong, *āŋ sat-čiú* ‘He is hitting himself’ and *āŋ et-čiú* ‘He is laughing/smiling’ have the same morphological form, but the semantics of the reflexive are less clear in ‘laugh,’ and this verb must take this suffix to mean ‘laugh’ rather than ‘laugh at (someone).’ This suffix has also become extended to use as a detransitivizer in some contexts (see LaPolla and Yang 2004; LaPolla 2000, this volume, on Dulong and Rawang). Several Tani languages, e.g. Padam, Nishi, have a similar suffix **su* (Lorrain 1907; Tayeng 1983; Das Gupta 1969; Sun 1993a), but it is unclear whether this suffix is cognate to the one in Rawang. rGyalrong has a verbal prefix *nə-* which marks indirect reflexives and middles and also functions as an emphaser of intransitiviness (Nagano 1984: 55, this volume; Jin *et al.* 1958: 81). Mizo (Chhangte 1993; Lorrain and Savidge 1898) has a verb prefix *ni-* which marks reflexive, reciprocal, and middle semantics.

Quite a few other frequent patterns could be discussed, but the above should suffice to show that with the loss of the original PST derivational morphology the daughter languages each went their own way in creating new morphology, but due to inherited typological features and areal contact, there were certain regularities in the types of morphology they developed (see LaPolla 1994b for more detailed discussion).

NOTES

- 1 The comparative method is limited in how far back it can go. Arguments trying to link genetic or archeology findings of earlier periods to languages carry no weight in discussions of language history, as there is no necessary connection between these sorts of evidence and language forms.
- 2 P = prefix, C_i = initial consonant, G = glide, : = vowel length, C_f = final consonant, s = suffixal *-s; parentheses mark that the item does not appear in all syllables. This is an expansion of the syllable canon given in Matisoff 1991a: 490 and 2003: 12.
- 3 This list is not exhaustive, and the necessarily brief discussion glosses over many controversies and details. As is always the case in attempting to find ST correspondences, the lack of a single standard for the reconstruction of OC (ideally based mainly on the comparative method) makes comparative work difficult and more conjectural than would otherwise be the case. What constitutes a cognate set using one reconstruction system might not be seen as cognate using another system (see LaPolla 2012a for discussion). I have here used the system of Baxter 1992, as this is the best system I have found to date, though even this system is in flux (see Baxter 1995; Baxter and Sagart 1998, 2014 for discussion of some of the recent changes, and Schuessler 2015 for a critical reaction to the changes).
- 4 See Hill (2014) for a different analysis of the Tibetan data.
- 5 In a slightly earlier paper, Mei (1988) argues for reconstructing a voiced initial rather than a prefix.
- 6 Baxter (1992: 221; following Chang and Chang 1976, 1977) also associates his *N-prefix (posited to account for characters with phonetic elements that appear in syllables with both stop and homorganic nasal initials) with Tibetan *a-chung*. Gong (2000; also following Chang and Chang 1976, 1977) associates Tibetan *a-chung* with a nasal prefix, but uses it to explain the development of Middle Chinese *d-, items that Baxter (1992) reconstructs with *ml- clusters (e.g. Gong: *N-ljək (食), Baxter *mljək (see Matisoff 1995, fn. 1; originally *Ljik in Baxter 1992).
- 7 See also Konnerth's (2009) discussion of the voiced velar nominalizing prefix found in many languages in the family, which she associates with these functions, but those velar prefixes might instead be related to the nominalizing *ay- prefix found in other parts of the family (see §1.1.3).
- 8 Matisoff (2003: 107): states that as far as verbal valence is concerned, “*m- is consistently stativizing/intansitive, *s- is consistently causativizing/transitivizing, while *ʔ- behaves sometimes one way, sometimes the other.”
- 9 Although the intransitivizing prefix Pulleyblank (1973a, 2000) reconstructed (see §1.1.2) was written as *ā-, it was seen as a non-syllabic pharyngeal glide; no one has argued for a prefix such as the one here to explain the voicing and transitivity contrasts found in Chinese, and no one has given arguments why it couldn't have been such a prefix.
- 10 Although not often explicitly mentioned, except by Jin (1998a, 1998b, 2006, 2012), the idea is that some of the finals we find on words are etymological, while others are due to affixation. Here we are only talking about affixation.
- 11 The usual reading of this last character when used as an interrogative pronoun is *?a, but it is written using a character that is in other contexts pronounced *?ak. If it is the same pronunciation as the one otherwise written (鳥), it seems odd to use a character that normally is read with a stop final.

- 12 This may also be the origin of the *-s* nominalizing suffix found in many languages in TB.
- 13 See Wang 1958[1980], Downer 1959 on the derivational process of the departing tone.
- 14 Our answer to this question will affect our understanding of certain word families. For example, Pulleyblank (1991: 30) suggests that **k(r/j)a?* (舉) ‘all’; ‘lift’ has an alloform **kjai* (揭) ‘lift’ (and so the latter would involve a **-t* suffix). This set would stand only if we assume the root did not originally have a **-k* final.
- 15 One might conjecture that the mysterious **-t* final of the OC negative imperative **mjət* is actually the prohibitive **ta*, but we do not find **ma-* and **ta-* occurring together in TB.
- 16 By grammatical relations is meant the grammatical singling out of a particular NP (the ‘pivot’ of a construction) for special grammatical treatment in a construction, such that a restricted neutralization of semantic roles occurs (has conventionalized/grammaticalized) in that position in the construction for the purpose of aiding referent identification. See Van Valin and LaPolla 1997, ch. 6, for the concept of pivot and its relation to grammatical relations. See Dixon 1995, ch. 2, and also Hale and Watters 1973 on semantic marking vs grammatical marking.
- 17 This is due to the fact that crosslinguistically agents are overwhelmingly more likely to be topical and patients are more likely to be focal in discourse. See Du Bois 1987, Mallison and Blake 1981 and Sun and Givón 1985 for statistics.
- 18 These groupings are not definitive, as much more work needs to be done on comparing the morphology (rather than random samples of words) to prove genetic relatedness (see LaPolla 2012a, 2013b for arguments). For earlier hypotheses on the genetic groupings, see Benedict 1972; Burling 1983; Dai *et al.* 1989; DeLancey 1987; Matisoff 1990, 1991a; Shafer 1955; Sun 1988; Thurgood 1984, 1985; Bradley 1997.
- 19 See Tournadre 2014 for this term, used for all of the varieties that can be shown to be derived from Old Tibetan.
- 20 See Tournadre 2008 for arguments against the use of the term “conjunction–disjunction” for the relevant phenomenon.
- 21 Dolakha Newar (Genetti, this volume), a Newar variety surrounded by Kiranti languages, is an exception in not manifesting the sort of conjunction–disjunction or evidential systems found in the other varieties of this language and in having innovated a person marking system on the verb. Classical Newar did not have such a system, and in the only other variety to have innovated such a system, Pahari, the system is not cognate with that in Dolakha, and neither system is cognate with the Rung or other known systems. See Kansakar 1999 for discussion and evidence.
- 22 See LaPolla 1987 on the lack of correspondence in the tone systems. Commonalities shown in that paper between Jinghpaw and T’rung are shared retentions, as both are rather conservative phonologically.
- 23 Ebert (1990) has argued for a Kiranti-rGyalrong-Rawang genetic grouping (see also Thurgood 1985), based largely on the person marking systems; I am including also Western Himalayan in this grouping, based on the person marking and the reflexive/middle marking (LaPolla 2013b). See also Grierson (1909, vol. III), for particular characteristics shared between the Western and Eastern Himalayan pronominalized languages not shared by the Tibetic languages, and Watters (1975: 50) for discussion of the “remarkable similarities” between the pronominals and subject marking systems of the Eastern (now including Kham) and Western Himalayan pronominalized languages. Chang and Chang (1975) also argued for a close connection between rGyalrong and T’rung. Given that the distribution of the group is due to migration along the eastern and southern edges of the Tibetan plateau, we would expect other languages found along that path, particularly in Northeast India, would be members of the group, but so far no language of that area fits the strict criteria established in LaPolla 2013b for membership in the group. The name Rung was coined by Thurgood (1984), but used for a somewhat larger grouping of languages identified using

- different principles and methodology. That original grouping is no longer recognized, and so I have used the name for this grouping, as my original name for the group, GRKW, is less euphonious.
- 24 Thurgood (1984) discusses the fact that rGyalrong, T'rung, and Kham all have a pre-verbal yes–no interrogative particle *ma- (< PTB *ma ‘negative particle’), and argues this is a shared innovation (a reduction of an alternative (A-not A) question) that points to a common parent language. If only these three share this innovation, it would cloud the picture presented above, unless there was an assumption that this form was lost in Kiranti, just as it is now being lost in T'rung.
One interesting commonality is what I have talked about as “transitivity harmony” in Rawang (LaPolla 2008c, 2011) and Kiranti, e.g. Hayu.
- 25 The Rawang people feel that the speakers of the Tani languages are related to the Rawang people, being simply a further extension of the Rawang migration west. They point to the name *Abor* as evidence (*Abur* is a Rawang clan name), and tell stories of Rawangs who have been to India and can speak in Rawang with the people there and be understood. Given the major differences in the languages, this would seem unlikely.
- 26 Cf. Donegan and Stampe's (2004) discussion of the shift in Munda due to the prevailing rhythm patterns in South and Southeast Asia.
- 27 As discussed in LaPolla 2001: 236, there is also the subjective aspect of the training of the scholars documenting the languages, which influences the description of the languages.
- 28 Independent innovation of bound pronominal paradigms in various languages in a family is not unique to TB, but occurred also in Amerindian (Mithun 1991) and Australian (Dixon 1980) languages.
- 29 The correspondence of the latter form with the Western Himalayan first and second person plural marker *ni is interesting in this regard, but the Western Himalayan form may be due to leveling of the original second person plural form to marking both plurals.
- 30 See also DeLancey 2011, Sawada 2012, and the other papers from the two special issues (34.2 and 35.1) of *LTBA* on ‘Optional Case Marking in Tibeto-Burman.’
- 31 The term anti-ergative may be somewhat infelicitous, as, like the term ergative itself, it may lead the reader to credit these particles with more of a paradigmatic nature than they actually have, but this term is already somewhat established in the literature (e.g. Comrie 1975, 1978; LaPolla 1992b), and clearer than Blansitt's (1984) term for this phenomenon, dechtaetaitive. I also do not use Dryer's (1986) term primary object because he defines it as a grammatical function. The use of this type of marking in most of the TB languages that have it is not of the nature of a grammatical function, and in some languages it is also not limited to marking objects and recipients.

REFERENCES

- Aikhenvald, Alexandra Y. and LaPolla, Randy J. (2007) ‘New perspectives on evidentials: a view from Tibeto-Burman’, *Linguistics of the Tibeto-Burman Area* 30.2: 1–16.
- Andersen, Paul Kent (1987) ‘Zero-anaphora and related phenomena in Classical Tibetan’, *Studies in Language* 11.2: 279–312.
- Baron, Stephen P. (1973) ‘The classifier-alone-plus-noun construction: a study in areal diffusion’, ICSTILL 6, University of California, San Diego, October 19–21 1973.
- Baxter, William H. (1992) *A Handbook of Old Chinese Phonology*. Berlin and New York: Mouton de Gruyter.
- Baxter, William H. (1995) ‘Old Chinese Version 1.1 (beta text version)’, Paper presented at the 28th International Conference on Sino-Tibetan Languages and Linguistics, Charlottesville, Virginia.

- Baxter, William H. and Sagart, Laurent (1998) ‘Word formation in Old Chinese: new approaches to Chinese’, in Jerome L. Packard (ed.) *Word Formation: Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*, Berlin and New York: Mouton de Gruyter, 35–76.
- Baxter, William H. and Sagart, Laurent (2014) *Old Chinese: A New Reconstruction*. Oxford: Oxford University Press.
- Benedict, Paul K. (1972) *Sino-Tibetan: A Conspectus (Princeton-Cambridge Studies in Chinese Linguistics II)*, James A. Matisoff, Contributing Editor). Cambridge: Cambridge University Press.
- Benedict, Paul K. (1991) ‘The Proto-Burmese-Yipho nominalizing *-t suffix’, *LTBA* 14.2: 149–153.
- Bernot, Lucien (1967) *Les Cak: Contribution à l'étude ethnographique d'une population de langue loi*. Paris: Centre National de la Recherche Scientifique.
- Bhat, D.N.S. (1991) *Grammatical Relations: The Evidence against their Necessity and Universality*. London: Routledge.
- Blansitt, E.L., Jr. (1984) ‘Dechtaetiative and dative’ in F. Plank (ed.) *Objects: Towards a Theory of Grammatical Relations*, London: Academic Press, 127–50.
- Bodman, Nicholas C. (1980) ‘Proto-Chinese and Sino-Tibetan: data towards establishing the nature of the relationship’, in F. van Coetsem and L.R. Waugh (eds) *Contributions to Historical Linguistics: Issues and Materials*, Leiden: E.J. Brill, 34–199.
- Bradley, David (1980) ‘Phonological convergence between languages in contact: Mon-Khmer structural borrowing in Burmese’, *BLS* 6: 259–267.
- Bradley, David (1997) ‘Tibeto-Burman languages and classification’, in D. Bradley (ed.) *Tibeto-Burman Languages of the Himalayas*, Canberra: Australian National University, 1–71.
- Bradley, David (2012) ‘The characteristics of the Burmic family of Tibeto-Burman’, *Language and Linguistics* 13.1: 171–192.
- Brown, R. G. (1920) ‘The Kadus of Burma’, *Bulletin of the School of Oriental Studies* 1.3: 1–28.
- Burling, Robbins (1983) ‘The Sal languages’, *LTBA* 7.2: 1–32.
- Chang, Betty Shefts and Kun, Chang (1976) ‘The prenasalized stop initials of Miao-Yao, Tibeto-Burman and Chinese: a result of diffusion or evidence of a genetic relationship?’, *BIHP* 47: 467–502.
- Chang, Betty Shefts and Kun, Chang (1977) ‘Tibetan prenasalized initials’, *BIHP* 48: 229–243.
- Chang, Kun and Chang, Betty Shefts (1975) ‘Gyarong historical phonology’, *BIHP* 46: 391–524.
- Chao, Yuen Ren (1955[1976]) ‘Notes on Chinese grammar and logic’, in Anwar S. Dil (ed.) *Aspects of Chinese Sociolinguistics: Essays by Yuen Ren Chao*, Stanford: Stanford University Press, 237–249.
- Chao, Yuen Ren (1959[1976]) ‘How Chinese logic operates’, in Anwar S. Dil (ed.) *Aspects of Chinese Sociolinguistics: Essays by Yuen Ren Chao*, Stanford: Stanford University Press, 250–259.
- Chao, Yuen Ren (1968) *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- Chappell, Hillary (2001) ‘Language contact and areal diffusion in Sinitic languages’, in R.M.W. Dixon and Alexandra Y. Aikhenvald (eds) *Areal Diffusion and Genetic Inheritance: Case Studies in Language Change*, Oxford: Oxford University Press, 328–357.

- Chappell, Hillary (ed.) (2004) *Chinese Grammar: Synchronic and Diachronic Perspectives*. Oxford: Oxford University Press.
- Chelliah, Shobhana L. (2009) ‘Semantic role to new information in Meithi’, in Jóhanna Barðdal and Shobhana Chelliah (eds) *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case*, Amsterdam and Philadelphia: Benjamins, 377–400.
- Chelliah, Shobhana L. (forthcoming) ‘Ergativity and differential marking in Tibeto-Burman’, in Jessica Coon, Diane Massam and Lisa Travis (eds) *Oxford Handbook of Ergativity*, Oxford: Oxford University Press.
- Chhangte, Lalnunthangi (1993) ‘A grammar of Mizo’, unpublished PhD dissertation, University of Oregon.
- Chirkova, Katia (2012a) ‘The Qiangic subgroup from an areal perspective: a case study of languages of Muli’, *Language and Linguistics* 13.1: 133–170.
- Chirkova, Katia (2012b) ‘Extralinguistic factors, language change, and comparative reconstructions: case studies from South-West China’, *Yuyanxue Luncong* 45: 67–111.
- Comrie, Bernard (1975) ‘Antiergative’. Papers from the 11th Regional Meeting of the Chicago Linguistic Society, ed. by R.E. Grossman, L.J. San and T.J. Vance, 112–121.
- Comrie, Bernard (1978) ‘Ergativity’, in W.P. Lehmann (ed.) *Syntactic Typology: Studies in the Phenomenology of Language*, Austin: University of Texas Press, 329–394.
- Dai, Qingxia (1994) ‘Zangmian yu geti liangci yanjiu [a study on numeral classifiers in Tibeto-Burman]’, in Ma Xueliang (ed.) *Zangmian yu xin lun [Recent Contributions to Tibeto-Burman Studies]*, Beijing: Zhongyang Minzu Xueyuan Chubanshe, 166–181.
- Dai, Qingxia (1997a) ‘A study on count-noun classifiers in Tibeto-Burman languages’, in Editorial Committee of the International Yi-Burmese Conference (ed.) *Studies on Yi-Burmese Languages*, Chengdu: Sichuan Nationalities Publishing House, 355–373.
- Dai, Qingxia (1997b) ‘Jingpoyu ci de shuang yinjiehua dui yufa de yingxiang [The influence of bisyllabification of lexical items in Jinghpaw on the grammar]’, *Minzu Yuwen* 1997.5: 25–30.
- Dai, Qingxia (2001) ‘Zangmian yuzu yuyan shidong fanchou de lishi yanbian [Historical changes in the causative category of Tibeto-Burman languages]’, *JCL* 29.1: 1–10.
- Dai, Qingxia, Huang, Bufan, Fu, Ailan, Renzengwangmu and Liu, Juhuang (1991) *Zangmianyu shiwu zhong [Fifteen Tibeto-Burman Languages]*. Beijing: Yanshan Chubanshe.
- Dai, Qingxia, Liu, Juhuang, and Fu, Ailan (1989) ‘Guanyu woguo Zangmian yuzu xishu fenlei wenti [On the problem of genetic subgrouping within the Tibeto-Burman languages of China]’, *Yunnan Minzu Xueyuan Xuebao* 1989.3: 82–92.
- Dai, Qingxia and Xu, Xijian (1992) *Jingpoyu yufa [Jinghpaw Grammar]*. Beijing: Zhongyang Minzu Xueyuan Chubanshe.
- Das, Sarat Chandra (1902) *A Tibetan–English Dictionary*. Calcutta: Bengal Secretariat Book Depot.
- Das Gupta, K. (1969) *Dafla Language Guide*. Shillong: Research Department, North-East Frontier Agency.
- DeLancey, Scott (1984) ‘Etymological notes on Tibeto-Burman case particles’, *LTBA* 8.1: 59–77.
- DeLancey, Scott (1986) ‘Evidentiality and volitionality in Tibetan’, in Wallace Chafe and Johanna Nichols (eds) *Evidentiality: The Linguistic Coding of Epistemology*, Norwood, NJ: Ablex, 203–213.
- DeLancey, Scott (1987) ‘The Sino-Tibetan languages’, in B. Comrie (ed.) *The World’s Major Languages*, New York: Oxford University Press, 799–810.
- DeLancey, Scott (2011) “Optional” “ergativity” in Tibeto-Burman languages’, *LTBA* 34.2: 9–20.

- Dixon, R.M.W. (1980) *The Languages of Australia*. Cambridge, London, and New York: Cambridge University Press.
- Dixon, R.M.W. (1995) *Ergativity*. Cambridge: Cambridge University Press.
- Donegan, Patricia J. and Stampe, David (2004) ‘Rhythm and the synthetic drift of Munda’, in Rajendra Singh (ed.) *The Yearbook of South Asian Languages and Linguistics 2004*, Berlin: Mouton de Gruyter, 3–36.
- Donohue, Mark and Donohue, Cathryn (2016) ‘On ergativity in Bumthang’, *Language* 92.1: 179–188.
- Downer, G.B. (1959) ‘Derivation by tone change in Classical Chinese’, *BSOAS* 22.2: 258–290.
- van Driem, George (1988) ‘Reflexes of the Tibeto-Burman *-t directive suffix in Dumi Rai’, in D. Bradley, E.J.A. Henderson and M. Mazaudon (eds) *Prosodic Analysis and Asian Linguistics: To Honour R.K. Sprigg*, Canberra: Pacific Linguistics C-104, 157–167.
- Dryer, Matthew S. (1986) ‘Primary objects, secondary objects, and antidative’, *Language* 62: 808–845.
- Du Bois, John W. (1987) ‘The discourse basis of ergativity’, *Language* 63.4: 805–855.
- Ebert, Karen H. (1987) ‘Grammatical marking of speech act participants in Tibeto-Burman’, *Journal of Pragmatics* 11.4: 473–482.
- Ebert, Karen H. (1990) ‘On the evidence for the relationship Kiranti-Rung’, *LTBA* 13.1: 57–78.
- Ebert, Karen H. (1993) ‘Kiranti subordination in the South Asian areal context’, in K.H. Ebert (ed.) *Studies in Clause Linkage*, Zürich: Seminar für Allgemeine Sprachwissenschaft, 83–110.
- Egerod, Søren (1985) ‘Typological features in Akha’, in Graham Thurgood, James A. Matisoff and David Bradley (eds) *Linguistics of the Sino-Tibetan Area: The State of the Art. Papers Presented to Paul K. Benedict for his 71st Birthday (Pacific Linguistics C, 87)*, Canberra: Australian National University, 96–104.
- Genetti, Carol (1986) ‘The development of subordinators from postpositions in Bodic languages’, *BLS* 12: 387–400.
- Genetti, Carol (1988) ‘A syntactic correlate of topicality in Newari narrative’, in J. Haiman and S.A. Thompson (eds) *Clause Combining in Grammar and Discourse (Typological Studies in Language 18)*, Amsterdam and Philadelphia: Benjamins, 29–48.
- Genetti, Carol (1991) ‘From postposition to subordinator in Newari’, in E.C. Traugott and B. Heine (eds) *Approaches to Grammaticalization*, vol. II, Amsterdam and Philadelphia: Benjamins Pub. Co., 227–255.
- Gong, Hwang-cherng (2000) ‘Cong Hanzangyu de bijiao kan Shanggu Hanyu de citou wenti [Looking at the problem of prefixes in Old Chinese through Sino-Tibetan comparisons]’, *Language and Linguistics* 1.2: 39–62.
- Gong, Hwang-cherng (2001) ‘Shanggu Hanyu yu yuanshi Han-Zangyu dai r yu l fushengmu de goumu [A reconstruction of consonant clusters with -r- and -l- as the second elements in Old Chinese and Proto-Sino-Tibetan]’, *Humanitas Taiwanica* 54: 1–36.
- Graham, Angus (1983) ““Yun” and “yue” as verbs and particles”, *Acta Orientalia Hungarica* 44: 33–71.
- Grierson, George (ed.) (1909) *Linguistic Survey of India*. Calcutta: Superintendent of Government Printing.
- Hale, Austin (1980) ‘Person markers: finite conjunct and disjunct verb forms in Newari’, *PSEAL* 7: 95–106.

- Hale, Austin and Watters, David E. (1973) ‘A survey of clause patterns’, in Austin Hale and David E. Watters (eds) *Clause, Sentence, and Discourse Patterns in Selected Languages of Nepal*, Part II (*SIL Pub. in linguistics and related fields*, no. 40), Kathmandu: SIL and Tribhuvan University Press, 175–249.
- Handel, Zev (2012) ‘Valence-changing prefixes and voicing alternation in Old Chinese and Proto-Sino-Tibetan: reconstructing *s- and *N- prefixes’, *Language and Linguistics* 13.1: 61–82.
- Hargreaves, David (1983) ‘Evidentiality in Newari’, unpublished MA thesis, University of Oregon.
- Hill, Nathan W. (2014) ‘A note on voicing alternation in the Tibetan verbal system’, *Transactions of the Philological Society* 112.1: 1–4.
- Hongladarom, Krisadawan (1993) *Evidentials in Tibetan: A Dialogic Study of the Interplay between Form and Meaning*. Ann Arbor: University Microfilms International.
- Hope, Edward R. (1973) ‘Non-syntactic constraints on Lisu noun phrase order’, *Foundations of Language* 10: 79–109.
- Hope, Edward R. (1974) *The Deep Syntax of Lisu Sentences: A Transformational Case Grammar* (*Pacific Linguistics B-34*). Canberra: Australian National University.
- Huang, Borong (1996) *Hanyu fangyan yufa leibian [A Categorization of Chinese Dialect Grammar]*. Qingdao: Qingdao Chubanshe.
- Huang, Bufan (1991) ‘Qiangyuzhi (the Qiangic branch)’, in Ma Xueliang (ed.) *Hanzangyu gailun [A General Introduction to Sino-Tibetan Languages]*, Beijing: Peking University Press, 208–370.
- Huziwarra, Keisuke (2008a) ‘Chakku-go no kizyutu engogakuteki kenkyuu [A descriptive linguistic study of the Cak language]’. Doctoral dissertation, Kyoto University, lix + 942 pp.
- Huziwarra, Keisuke (2008b) ‘Cak numerals’, *Dhaka University Journal of Linguistics* 1.2: 1–10.
- Huziwarra, Keisuke (2010) ‘Cak prefixes’, in Dai Zhongming and James A. Matisoff (eds) *Zang-Mian-yu yanjiu shishi nian [Forty Years of Sino-Tibetan Studies]*, Harbin: Heilongjiang University Press, 130–145.
- Jin, Lixin (1998a) ‘Hanzangyu de mingci houzhui *-n [The noun suffix *-n in Sino-Tibetan]’, *Minzu Yuwen* 1998, 1: 43–48.
- Jin, Lixin (1998b) ‘Hanzangyu zhong liang ge xingzhi butong de *-g yunwei [Two different *-g finals with different natures in Sino-Tibetan]’, *Minzu Yuwen* 1998.6: 27–34.
- Jin, Lixin (2000) ‘Lun xingtai zai queding Han-Zang tongyuanci zhong de zhongyao yi yi [The importance of morphology in determining Sino-Tibetan cognates]’, *Minzu Yuwen* 3: 41–49.
- Jin, Lixin (2004) ‘Han-Zangyu de shiyi dongci houzhui *-d [The *-d causative suffix in Sino-Tibetan]’, *Minzu Yuwen* 2: 38–45.
- Jin, Lixin (2005a) ‘Han-Zangyu de wanchengti houzhui *-s [The *-s perfective suffix in Sino-Tibetan]’, *Minzu Yuwen* 2: 1–6.
- Jin, Lixin (2005b) ‘Shanggu Hanyu *m- qianzhui de yi yi [The meaning of the Old Chinese *m- prefix]’, *Yuyan Yanjiu* 1: 73–77.
- Jin, Lixin (2005c) ‘Shanggu Hanyu shengmu qingzhuo jiaoti he dongci de ti [The alternation of voiced and unvoiced initials in Old Chinese and verbal aspect]’, *Yuwen Yanjiu* 4: 27–31.
- Jin, Lixin (2006) *Shanggu Hanyu xingtai yanjiu [Studies of Old Chinese Morphology]*. Hefei: Huangshan Shushe.
- Jin, Lixin (2012) *Shanggu yinlue [On Old Chinese Phonology]*. Hefei: Huangshang Shushe.

- Jin, Peng, Tan, Kerang and Lin, Xiangrong (1958) ‘Jiarongyu Suomohua de yuyin he xingtai (xu) [The phonology and morphology of the Suomo dialect of Jiarong, part 2]’, *Yuyain Yanjiu* 3: 71–108.
- Kansakar, Tej R. (1999) ‘Verb agreement in Classical Newar and modern Newar dialects’, in Yogendra P. Yadava and Warren W. Glover (eds) *Topics in Nepalese Linguistics*, Kathmandu: Royal Nepal Academy, 421–443.
- Karlgren, Bernard (1933) ‘Word families in Chinese’, *BMFEA* 5: 9–120.
- Karlgren, Bernard (1956) ‘Cognate words in the Chinese phonetic series’, *BMFEA* 28: 1–18.
- Konnerth, Linda (2009) ‘The nominalizing prefix *gV- in Tibeto-Burman’, MA thesis, University of Oregon.
- LaPolla, Randy J. (1987) ‘Dulong and Proto-Tibeto-Burman’, *LTBA* 10.1: 1–42.
- LaPolla, Randy J. (1990) ‘Grammatical relations in Chinese: synchronic and diachronic considerations’, unpublished PhD dissertation, University of California, Berkeley.
- LaPolla, Randy J. (1992a) ‘On the dating and nature of verb agreement in Tibeto-Burman’, *BSOAS* 55.2: 298–315.
- LaPolla, Randy J. (1992b) ‘Anti-ergative marking in Tibeto-Burman’, *LTBA* 15.1: 1–9.
- LaPolla, Randy J. (1993a) ‘On the change to verb-medial word order in Proto-Chinese: evidence from Tibeto-Burman’, in H. Kitamura, T. Nishida and Y. Nagano (eds) *Current Issues in Sino-Tibetan Linguistics*, Osaka: National Museum of Ethnology, 98–104.
- LaPolla, Randy J. (1993b) ‘Arguments against “subject” and “direct object” as viable concepts in Chinese’, *BIHP* 63.4: 759–813.
- LaPolla, Randy J. (1994a) ‘Word order patterns in Sino-Tibetan: their significance to theories of explanation in typology’, Paper presented at the Symposium on Language Typology, Tsukuba University, Japan, 19–21 January 1994. Published as LaPolla 2002.
- LaPolla, Randy J. (1994b) ‘Parallel grammaticalizations in Tibeto-Burman: evidence of Sapir’s “drift”’, *LTBA* 17.1: 61–80.
- LaPolla, Randy J. (1994c) ‘Variable finals in Proto-Sino-Tibetan’, *BIHP* 65.1: 131–173.
- LaPolla, Randy J. (1995a) ‘Ergative marking in Tibeto-Burman’, in Yoshio Nishi, James A. Matisoff and Yasuhiko Nagano (eds) *New Horizons in Tibeto-Burman Morpho-Syntax (Senri Ethnological Studies 41)*, Osaka: National Museum of Ethnology, 189–228.
- LaPolla, Randy J. (1995b) ‘On the utility of the concepts of markedness and prototypes in understanding the development of morphological systems’, *BIHP* 66.4: 1149–1185.
- LaPolla, Randy J. (1995c) ‘Pragmatic relations and word order in Chinese’, in P. Downing and M. Noonan (eds) *Word Order in Discourse*, Amsterdam and Philadelphia: Benjamins, 299–331.
- LaPolla, Randy J. (1996) ‘Middle voice marking in Tibeto-Burman languages’, *Pan-Asian Linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics*, Vol. V. Mahidol University, Thailand, 1940–1954.
- LaPolla, Randy J. (2000) ‘Valency-changing derivations in Dulong/Rawang’, in R.M.W. Dixon and A.Y. Aikhenvald (eds) *Changing Valency: Case Studies in Transitivity*, Cambridge: Cambridge University Press, 282–311.
- LaPolla, Randy J. (2001) ‘The role of migration and language contact in the development of the Sino-Tibetan language family’, in R.M.W. Dixon and A.Y. Aikhenvald (eds) *Areal Diffusion and Genetic Inheritance: Case Studies in Language Change*, Oxford: Oxford University Press, 225–254.

- LaPolla, Randy J. (2002) ‘Problems of methodology and explanation in word order universals research’, in Pan Wuyun (ed.) *Dongfang Yuyan yu Wenhua [Languages and Cultures of the East]*, Shanghai: Dongfang Chuban Zhongxin, 204–237.
- LaPolla, Randy J. (2003) ‘Evidentiality in Qiang’, in Alexandra Y. Aikhenvald and R. M. W. Dixon (eds) *Studies in Evidentiality (Typological Studies in Language 54)*. Amsterdam and Philadelphia: John Benjamins, 63–78.
- LaPolla, Randy J. (2004) ‘On nominal relational morphology in Tibeto-Burman’, in Ying-jin Lin, Fang-min Hsu, Chun-chih Lee, Jackson T.-S. Sun, Hsiu-fang Yang and Dah-an Ho (eds) *Studies on Sino-Tibetan Languages: Papers in Honor of Professor Hwang-cherng Gong on his Seventieth Birthday*, Taipei: Institute of Linguistics, Academia Sinica, 43–74.
- LaPolla, Randy J. (2005a) ‘Sino-Tibetan languages’, in Keith Brown (ed.) *Encyclopedia of Language and Linguistics* (2nd edition), London: Elsevier, 393–397.
- LaPolla, Randy J. (2005b) ‘The inclusive–exclusive distinction in Tibeto-Burman languages’, in Elena Filimonova (ed.) *Clusivity: Typology and Case Studies on Inclusive/Exclusive Oppositions*, Amsterdam and Philadelphia: John Benjamins, 289–310.
- LaPolla, Randy J. (2008a) ‘Nominalization in Rawang’, *LTBA* 31.2: 45–66.
- LaPolla, Randy J. (2008b) ‘Relative clause structures in the Rawang language’, *Language and Linguistics* 9.4: 797–812.
- LaPolla, Randy J. (2008c) “‘Transitivity harmony’ in the Rawang language of northern Myanmar”, in *Proceedings of the 2008 Annual Conference of the Australian Linguistic Society*, University of Sydney, 2–4 July 2008.
- LaPolla, Randy J. (2009) ‘Chinese as a topic–comment (not topic–prominent and not SVO) language’, in Janet Xing (ed.) *Studies of Chinese Linguistics: Functional Approaches*, Hong Kong: Hong Kong University Press, 9–22.
- LaPolla, Randy J. (2010) ‘Hierarchical person marking in Rawang’, in Dai Zhaoming (ed.) *Forty Years of Sino-Tibetan Language Studies: Proceedings of ICSTLL-40*, Heilongjiang University Press, 107–113.
- LaPolla, Randy J. (2011) ‘On transitivity in two Tibeto-Burman languages’, *Studies in Language* 35.3: 639–649.
- LaPolla, Randy J. (2012a) ‘Comments on methodology and evidence in Sino-Tibetan comparative linguistics’, *Language and Linguistics* 13.1 (Special issue on Sino-Tibetan Comparative Studies): 117–132.
- LaPolla, Randy J. (2012b) ‘Once again on person-marking in Tibeto-Burman: a reply to DeLancey 2010’, Paper presented at the 45th International Conference on Sino-Tibetan Languages and Linguistics, Nanyang Technological University, 26–28 October. Available at: randylapolla.net/papers/LaPolla_2012_Once_Again_on_Person_Marking_in_Tibeto-Burman_-a_reply_to_DeLancey_2010.pdf.
- LaPolla, Randy J. (2013a) ‘Eastern Asia: Sino-Tibetan linguistic history’, in Immanuel Ness (ed.) *The Encyclopedia of Global Human Migration*, Chichester: John Wiley and, vol. 1, 204–208.
- LaPolla, Randy J. (2013b) ‘Subgrouping in Tibeto-Burman: can an individual-identifying standard be developed? How do we factor in the history of migrations and language contact?’, in Balthasar Bickel, Lenore A. Grenoble, David A. Peterson and Alan Timberlake (eds) *Language Typology and Historical Contingency*. Amsterdam and Philadelphia: John Benjamins, 463–474.
- LaPolla, Randy J. (2013c) ‘Arguments for a construction-based approach to the analysis of Chinese’, in Tseng Chiu-yu (ed.) *Human Language Resources and Linguistic*

- Typology, Papers from the Fourth International Conference on Sinology*, Taiwan: Academia Sinica, 33–57.
- LaPolla, Randy J. (2015) ‘Sino-Tibetan syntax’, in William S.-Y. Wang and Chaofen Sun (eds) *The Oxford Handbook of Chinese Linguistics*, Oxford: Oxford University Press, 45–57.
- LaPolla, Randy J. with Huang, Chenglong (2003) *A Grammar of Qiang, with Annotated Texts and Glossary*. Berlin: Mouton de Gruyter.
- LaPolla, Randy J. and Poa, Dory (2005) ‘Jiaodian jiegou de leixing ji qi dui Hanyu cixu de yingxiang’ [The typology of focus structures and their effect on word order in Chinese], in Xu Liejiong and Haihua Pan (eds) *Jiaodian jiegou he yuyi de yanjiu [Studies on the Structure and Semantics of Focus]*, Beijing: Beijing Foreign Studies University Press, 57–78.
- LaPolla, Randy J. and Poa, Dory (2006) ‘On describing word order’, in Felix Ameka, Alan Dench and Nicholas Evans (eds) *Catching Language: The Standing Challenge of Grammar Writing*, Berlin: Mouton de Gruyter, 269–295.
- LaPolla, Randy J. and Yang, Jiangling (2004) ‘Reflexive and middle marking in Dulong-Rawang’, *Himalayan Linguistics* 2. Available at: randylapolla.net/Papers/LaPolla_and_Yang_2005_Reflexive_and_Middle_Marking_in_Dulong-Rawang.pdf.
- LaPolla, Randy J. (Luo Rendi) and Yang, Jiangling (2007) ‘Cong yuyan, fangyan de bijiao liaojie yuyan de lishi fazhan’ [Understanding the historical development of a language by comparing it with related languages and dialects], *Journal of Sino-Tibetan Linguistics* 1.1: 120–126. [Also published in *Dongfang Yuyanxue* 2 (2007): 11–18].
- Lee, Edith Yuet Ying (2002) ‘The influence of information structure on word order in Cantonese’, MA thesis, City University of Hong Kong.
- Lee Yeon-ju and Sagart, Laurent (2008) ‘No limits to borrowing: the case of Bai and Chinese’, *Diachronica* 25.3: 357–385.
- Li, Yongsui and Wang, Ersong (1986) *Haniyu jianzhi [Brief Description of the Hani Language]*. Beijing: Minzu Chubanshe.
- Liu, Danqing (2008) *Yufa diaocha yanjiu shouce [Handbook for Grammar Fieldwork and Analysis]*. Shanghai: Shanghai Jiaoyu Chubanshe.
- Liu, Danqing, ed. (2013) *Hanyu fangyan yufa yanjiu de xin shijiao—Di wu Hanyu fangyan yufa yanjiu guoji xueshu yantaohui lunwenji [A New Perspective on the Study of the Grammar of Chinese Dialects—Papers from the 5th International Conference for the Study of Chinese Dialect Grammar]*. Shanghai: Shanghai Jiaoyu Chubanshe.
- Lorrain, J.H. (1907) *A Dictionary of the Abor-Miri Language*. Shillong: Eastern Bengal and Assam Secretariat Printing Office.
- Lorrain, J.H. and Savidge, Fred W. (1898) *A Grammar and Dictionary of the Lushai Language*. Shillong: Assam Secretariat Printing Office.
- Luce, Gordon H. (1985) ‘Sak-Kantu’, in Gordon H. Luce (ed.) *Phases of Pre-Pagan Burma: Languages and History*, vol. I, New York/Oxford: Oxford University Press, 36–46.
- Lyu, Shuxiang (1979) *Hanyu yufa fenxi wenti [Questions on Analysing Chinese Grammar]*. Beijing: Commercial Press.
- Mallison, Graham and Blake, Barry J. (1981) *Language Typology: Cross-Linguistic Studies in Syntax* (North-Holland Linguistic Series 46). Amsterdam, New York, Oxford: North Holland.
- Matisoff, James A. (1973) *The Grammar of Lahu* (University of California Publications in Linguistics, 75). Berkeley and Los Angeles: University of California Press.

- Matisoff, James A. (1976) ‘Lahu causative constructions: case hierarchies and the morphology/syntax cycle in a Tibeto-Burman perspective’, in Masayoshi Shibatani (ed.) *The Grammar of Causative Constructions (Syntax and Semantics 6)*, New York: Academic Press, 413–442.
- Matisoff, James A. (1989) ‘Palatal suffixes in Sino-Tibetan’, *ICSTLL* 22, University of Hawaii, 6–8 October 1989.
- Matisoff, James A. (1990) ‘On megalocomparison’, *Language* 66.1: 106–120.
- Matisoff, James A. (1991a) ‘Sino-Tibetan linguistics: present state and future prospects’, *Annual Review of Anthropology* 20: 469–504.
- Matisoff, James A. (1991b) ‘Areal and universal dimensions of grammaticalization in Lahu’, in E. Traugott and B. Heine (eds) *Approaches to Grammaticalization*, vol. II, Amsterdam and Philadelphia: John Benjamins, 383–454.
- Matisoff, James A. (1995) ‘Sino-Tibetan palatal suffixes revisited’, in Yoshio Nishi, James A. Matisoff and Yasuhiko Nagano (eds) *New Horizons in Tibeto-Burman Morphosyntax (Senri Ethnological Studies 41)*, Osaka: National Museum of Ethnology, 35–91.
- Matisoff, James A. (2003) *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley: University of California Press.
- Matisoff, James A. (2008) ‘Zang-Mian yuxi-li de Baiyu xishu wenti’, Chinese translation by Fu Jingqi of ‘On the genetic position of Bai within Tibeto-Burman’, Paper presented at the 34th International Conference on Sino-Tibetan Languages and Linguistics, Kunming, 2001. Published in Bai Cultural Institute of the Dali Bai Autonomous Prefecture’s *Bai-yu Pian [Volume on Bai Language]*, vol. I, Kunming: Yunnan Minzu Chubanshe, 503–544.
- Matisoff, James A. (2013) ‘Re-examining the genetic position of Jingpho: putting flesh on the bones of the Jingpho/Luish relationship’, *LTBA* 36.2: 15–95.
- Matisoff, James A. (2014) ‘On the demise of the Proto-Tibeto-Burman mid vowels’, *Bulletin of the National Museum of Ethnology* 39.3: 375–395.
- Mei, Tsu-lin (1980) ‘Sisheng bieyi zhong de cengci [Strata in the use of tones to distinguish meaning]’, *Zhongguo Yuwen* 1980.6: 427–443.
- Mei, Tsu-lin (1988) ‘Neibu nigou Hanyu sanli [Three examples of internal reconstruction in Chinese]’, *Zhongguo Yuwen* 1988.3: 169–181.
- Mei, Tsu-lin (1989) ‘The causative and denominative functions of the *s prefix in Old Chinese’, *Proceedings of the Second International Conference on Sinology: Section on Linguistics and Paleography*, Taipei: Academia Sinica, 33–52.
- Mei, Tsu-lin (2008a) ‘Jiaguwen li de jige fufuyin shengmu [Some consonant clusters in oracle bone inscriptions]’, *Zhongguo Yuwen* 2008.3: 195–207.
- Mei, Tsu-lin (2008b) ‘Shanggu Hanyu dongci zhuo-qing bieyi de laiyuan [The origin of voicing alternation in Old Chinese verbs]’, *Minzu Yuwen* 2008.3: 3–20.
- Mei, Tsu-lin (2012) ‘The causative *s- and nominalizing *-s in Old Chinese and related matters in Proto-Sino-Tibetan’, *Language and Linguistics* 13.1: 1–28.
- Michailovsky, Boyd (1985) ‘Tibeto-Burman dental suffixes: evidence from Limbu (Nepal)’, in Graham Thurgood, James A. Matisoff and David Bradley (eds) *Linguistics of the Sino-Tibetan Area: The State of the Art. Papers Presented to Paul K. Benedict for his 71st Birthday (Pacific Linguistics C. 87)*, Canberra: Australian National University, 334–343.
- Mithun, Marianne (1991) ‘The development of bound pronominal paradigms’, in W.P. Lehmann and H.J.J. Hewitt (eds) *Language Typology* 1988, Amsterdam and Philadelphia: John Benjamins, 85–104.
- Nagano, Yasuhiko (1984) *A Historical Study of the rGyalrong Verb System*. Tokyo: Seishido.

- Nagano, Yasuhiko (1987) ‘Some ergative phenomena in Tibeto-Burman’, *The Memoirs of the Research Department of the Toyo Bunko* 45: 53–74.
- Norman, Jerry (1988) *Chinese*. Cambridge: Cambridge University Press.
- Phua, Chiew Pheng (2004) ‘Shanggu Hanyu zhong de qing-zhuo bieyi ji qi xiangguan de wenti [Alternations of voiced and voiceless initials in Old Chinese and its problems]’, in Center for Chinese Linguistics, Division of Humanities (ed.) *Ding Bangxin Jiaoshou Rongxiu Jinian Lunwenji [Collected Essays in Honor of Professor Ting Pang-Hsin on his Retirement]*, 1, 67. Hong Kong: Center for Chinese Linguistics, Division of Humanities, Hong Kong University of Science and Technology.
- Pulleyblank, Edwin G. (1962–3) ‘The consonantal system of Old Chinese (2 parts)’, *Asia Major* (new series) 9: 58–144, 206–265.
- Pulleyblank, Edwin G. (1972) ‘Word families in Chinese: a reconsideration’, *Journal of Asian and African Studies* 5: 1–19.
- Pulleyblank, Edwin G. (1973a) ‘Some new hypotheses concerning word families in Chinese’, *JCL* 1: 113–114.
- Pulleyblank, Edwin G. (1973b) ‘Some further evidence regarding Old Chinese *-s and its time of disappearance’, *BSOAS* 36.2: 368–373.
- Pulleyblank, Edwin G. (1977–8) ‘The final consonants of Old Chinese’, *Monumenta Serica* 33: 180–206.
- Pulleyblank, Edwin G. (1991) ‘Some notes on morphology and syntax in Classical Chinese’, in Henry Rosemont, Jr. (ed.) *Chinese Texts and Philosophical Contexts: Essays Dedicated to Angus C. Graham*, LaSalle, IL: Open Court Press, 21–45.
- Pulleyblank, Edwin G. (2000) ‘Morphology in Old Chinese’, *JCL* 28.1: 26–51.
- Sagart, Laurent (1999) *The Roots of Old Chinese*. Amsterdam; Philadelphia: John Benjamins.
- Sagart, Laurent [Shā Jiā’ěr] and Baxter, William H. [Bái Yǐpíng] (2010) ‘Shànggǔ Hànyǔ de N- hé m- qiánzhui [The N- and m- prefixes in Old Chinese]’, *Journal of Sino-Tibetan Linguistics* 4: 62–69.
- Sagart, Laurent and Baxter, William H. (2012) ‘Reconstructing the *s- prefix in Old Chinese’, *Language and Linguistics* 13.1: 29–59.
- Sangdong, David (2012) ‘A grammar of the Kadu (Asak) language’, PhD thesis, La Trobe University.
- Sawada, Hideo (2012) ‘Optional marking of NPs with core case functions P and A in Lhaovo’, *LTBA* 35.1: 15–34.
- Saxena, Anju (1991) ‘Pathways of the development of the ergative in Central Tibetan’, *LTBA* 14.1: 109–116.
- Schuessler, Axel (1985) ‘The function of *qusheng* in Early Zhou Chinese’, in Graham Thurgood, James A. Matisoff and David Bradley (eds) *Linguistics of the Sino-Tibetan Area: The State of the Art. Papers Presented to Paul K. Benedict for his 71st Birthday* (*Pacific Linguistics C, 87*). Canberra: Australian National University, 344–362.
- Schuessler, Axel (2015) ‘New Old Chinese’, *Diachronica* 32.4: 571–598.
- Shafer, Robert (1955) ‘Classification of the Sino-Tibetan languages’, *Word* 11.1: 94–111.
- Shirai, Satoko (2009) ‘Directional prefixes in nDrapa and neighboring languages: an areal feature of Western Sichuan’, in Yasuhiko Nagano (ed.) *Issues in Tibeto-Burman Historical Linguistics*, Osaka: National Museum of Ethnology, 7–20.
- Simon, Walter (1941) ‘Certain Tibetan suffixes and their combinations’, *Harvard Journal of Asiatic Studies* 5: 388–389.
- Simon, Walter (1942) ‘Tibetan *day*, *cij*, *kyin*, *yin*, and *ham*’, *BSOAS* 10: 954–975.

- Simon, Walter (1949) ‘The range of sound alternation in Tibetan word families’, *Asia Major* (new series) 1: 1–15.
- Simon, Walter, (1977) ‘Alternation of final vowel with dental nasal or plosive in Tibetan’, *BSOAS* XL: 51–57.
- Solnit, David (1997) *Eastern Kayah Li: Grammar; Texts, Glossary*. Honolulu: University of Hawai’i Press.
- Sun, Chao-fen and Talmy Givón (1985) ‘On the so-called SOV word order in Mandarin Chinese: a quantified text study and its implications’, *Language* 61: 329–351.
- Sun, Hongkai (1981a) *Qiangyu Jianzhi [A Brief Description of the Qiang Language]*. Beijing: Nationalities Press.
- Sun, Hongkai (1981b) ‘Qiangyu dongci de quxiang fanchou [The category of direction in the Qiang verb]’, *Minzu Yuwen* 1981.1: 34–42.
- Sun, Hongkai (1982) ‘Qiangyu zhishu wenti chutan [On the question of affiliation of the Qiang language]’, in Editorial Board (ed.) *Minzu yuwen yanjiu wenji [Collected Papers on Minority Languages]*, Xining: Qinghai Minzu Chubanshe, 189–224.
- Sun, Hongkai (1983) ‘Yidu Luobayu gaiyao [A brief description of the Idu Lhoba Language]’, *Minzu Yuwen* 1983.6: 63–79.
- Sun, Hongkai (1985) ‘Liujiang liuyu de minzu yuyan ji qi xishu fen lei [The ethnic languages of the Six Rivers area and their genetic affiliations]’, *Minzu Xuebao* 3: 99–274.
- Sun, Hongkai (1988) ‘Shilun Zhongguo jingnei Zang-Mian yu de puxi fenlei [A classification of Tibeto-Burman languages in China]’, in Paul K. Eguchi, Yukio Fujimoto, Nobuyoshi Fukuhara, Masura Hashimoto, Koichi Miya-moto, Atsushi Iwamoto, Tatsuo Kondo, Masaoki Miyamoto, Osamu Sakiyama, Akihiro Sato, David Sell, Norio Shibata, Ken-ichiro Shirai, Mashiro Shogaito, Shiro Yabu and Kazuhiko Yoshida (eds) *Languages and History in East Asia: A Festschrift for Tatsuo Nishida on the Occasion of his 60th Birthday*, vol. I, Kyoto: Shokado, 61–73.
- Sun, Hongkai (1991) ‘Cong cihui bijiao kan Xixiayu yu Zangmian yuzu Qiangyuzhi de guanxi [The relationship between Tangut and the Qiang branch of Tibeto-Burman from the point of view of shared lexical items]’, *Minzu Yuwen* 1991.2: 1–11.
- Sun, Hongkai (2001) ‘Lun Zangmian yuzu zhong de Qiangyuzhi yuyan [On the languages of the Qiangic branch of Tibeto-Burman]’, *Language and Linguistics* 2.1: 157–181.
- Sun, Jackson T.-S. (1993a) ‘A historical-comparative study of the Tani (Mirish) branch’, unpublished PhD dissertation, University of California, Berkeley.
- Sun, Jackson T.-S. (1993b) ‘The linguistic position of Tani (Mirish) in Tibeto-Burman: a lexical assessment’, *LTBA* 16.2: 143–188.
- Sun, Jackson T.-S. (1993c) ‘Evidentials in Amdo Tibetan’, *BIHP* 63.4: 945–1001.
- Tayeng, Aduk (1983) *A Phrase Book in Padam*. Shillong: Director of Information and Public Relations, Arunachal Pradesh.
- Thurgood, Graham (1984) ‘The “Rung” languages: a major new Tibeto-Burman subgroup’, in Claudia Brugman and Monica Macaulay (eds) *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society*, Berkeley: University of California, 338–349.
- Thurgood, Graham (1985) ‘Pronouns, verb agreement systems, and the subgrouping of Tibeto-Burman’, in Graham Thurgood, James A. Matisoff and David Bradley (eds) *Linguistics of the Sino-Tibetan Area: The State of the Art. Papers Presented to Paul K. Benedict for his 71st Birthday (Pacific Linguistics C, 87)*, Canberra: Australian National University, 376–400.

- Thurgood, Graham (1986) ‘The nature and origins of the Akha evidential system’, in Wallace Chafe and Johanna Nichols (eds) *Evidentiality: The Linguistic Coding of Epistemology*, Norwood, NJ: Ablex, 214–222.
- Thurgood, Graham (1994) ‘Tai-Kadai and Austronesian, the nature of the historical relationship’, *Oceanic Linguistics* 33.2: 345–368.
- Tournadre, Nicolas (1991) ‘The rhetorical use of the Tibetan ergative’, *LTBA* 14.1: 93–108.
- Tournadre, Nicolas (1996) *L’ergativité en tibétain: approche morphosyntaxique de la langue parlée*. Paris: Peeters.
- Tournadre, Nicolas (2008) ‘Arguments against the concept of “conjunct”/“disjunct” in Tibetan’, in B. Huber, M. Volkart, P. Widmer and P. Schwieger (eds) *Chomolangma, Demawend und Kasbek: Festschrift für Roland Bielmeier zu seinem 65. Geburtstag, Band 1*, Halle (Saale): International Institute for Tibetan and Buddhist Studies GmbH, 281–308.
- Tournadre, Nicolas (2014) ‘The Tibetic languages and their classification’, in Thomas Owen-Smith and Nathan Hill (eds) *Trans-Himalayan Linguistics: Historical and Descriptive Linguistics of the Himalayan Area*, Berlin: Mouton de Gruyter, 105–130.
- Tournadre, Nicolas and LaPolla, Randy J. (2014) ‘Towards a new approach to evidentiality: issues and directions for research’, *LTBA* 37.2: 240–262.
- Van Valin, Robert D. Jr. and LaPolla, Randy J. (1997) *Syntax: Structure, Meaning, and Function (Cambridge Textbooks in Linguistics Series)*. Cambridge: Cambridge University Press.
- Wang, Feng (2013) *Han-Zang yuyan bijiao de fangfa yu shijian—Han, Bai, Yi bijiao yanjiu [The Method and Practice of Sino-Tibetan Linguistic Comparison—A Comparative Study of Chinese, Bai, and Yi]*. Beijing: Peking University Press.
- Wang, Li (1958[1980]) *Hanyu shigao [Outline History of the Chinese Language]*. (Republished in 1980). Beijing: Zhonghua Shuju.
- Wang, Li (1980) *Shijing yundu [The Rhymes of the Book of Poetry]*. Shanghai: Guji Chubanshe.
- Wang, Li (1982) *Tongyuan cidian [Cognate Dictionary]*. Beijing: Commercial Press.
- Wang, Tianxi (1991) ‘Queyu’, in Dai Qingxia, Huang Bufan, Fu Ailan, Renzengwangmu and Liu Juhuang (eds) *Zangmianyu shiwu zhong [Fifteen Tibeto-Burman Languages]*. Beijing: Yanshan Chubanshe, 46–63.
- Watters, David E. (1975) ‘The evolution of a Tibeto-Burman pronominal verb morphology: a case study from Kham (Nepal)’, *LTBA* 2: 45–80.
- Wheatley, Julian K. (1982) ‘Burmese: a grammatical sketch’, unpublished PhD dissertation, University of California, Berkeley.
- Wheatley, Julian K. (1984) ‘The role of verb serialization in word-order change’, *BLS* 10: 350–359.
- Wheatley, Julian K. (1985) ‘The decline of verb-final syntax in the Yi (Lolo) languages of southeastern China’, in Graham Thurgood, James A. Matisoff and David Bradley (eds) *Linguistics of the Sino-Tibetan Area: The State of the Art. Papers Presented to Paul K. Benedict for his 71st Birthday (Pacific Linguistics C, 87)*, Canberra: Australian National University, 401–420.
- Wolfenden, Stuart N. (1928) ‘The prefix *m-* with certain substantives in Tibetan’, *Language* 4: 277–280.
- Wolfenden, Stuart N. (1929) *Outlines of Tibeto-Burman Linguistic Morphology*. London: Royal Asiatic Society.
- Wolfenden, Stuart N. (1936) ‘On certain alternations between dental finals in Tibetan and Chinese’, *Journal of the Royal Asiatic Society* 1936: 401–416.

- Wolfenden, Stuart N. (1937) ‘Concerning the variation of final consonants in the word families of Tibetan, Kachin and Chinese’, *Journal of the Royal Asiatic Society* 1936: 625–655.
- Wolfenden, Stuart (1939) ‘On the restitution of final consonants in certain word types of Burmese’ *Acta Orientalia* 17: 153–168.
- Woodbury, Anthony C. (1986) ‘Interactions of tense and evidentiality: a study of Sherpa and English’, in Wallace Chafe and Johanna Nichols (eds) *Evidentiality: The Linguistic Coding of Epistemology*, Norwood, NJ: Ablex, 188–202.
- Xu, Lin and Zhao, Yansun (1984) *Baiyu jianzhi [A Brief Description of the Bai Language]*. Beijing: Nationalities Press.
- Xu, Xijian (1987) ‘Classifiers in Jingpo’, *Minzu Yuwen* 1987.5: 27–35.
- Xu, Xijian (1989) ‘On the origin and development of noun classifiers in Jingpo’, translated by Randy J. LaPolla, *Linguistics of the Tibeto-Burman Area* 12.2: 15–23.
- Xu, Xijian and Xu, Guizhen (eds) (1984) *Jingpozu yuyan jianzhi (Zaiwa) [A Brief Description of the Zaiwa Language of the Jinghpaw People]*. Beijing: Nationalities Press.
- Yue-Hashimoto, Anne O.-K. (1993) *Comparative Chinese Dialectal Grammar: Handbook for Investigators*. Paris: Ecole des Hautes Etudes en Sciences Sociales (CRLAO).
- Zhang, Xifeng (1999) *Hanyu cizu congkao [Collection of Chinese Word Families]*. Chengdu: Ba Shu Shushe.
- Zhengzhang, Shangfang (2013) *Shanggu yinxi, di er ban [The Old Chinese Sound System, 2nd edition]*. Shanghai: Shanghai Jiaoyu Chubanshe.

CHAPTER THREE

WORD ORDER IN SINO-TIBETAN LANGUAGES FROM A TYPOLOGICAL AND GEOGRAPHICAL PERSPECTIVE¹

Matthew S. Dryer

1 INTRODUCTION

Word order, both at the clause level and even more at the phrase level, varies among Sino-Tibetan languages. In this chapter, I describe some of this variation and examine it in the light of word order tendencies found among the languages of the world as a whole. In section 2, I briefly summarize some of the variation in word order within Tibeto-Burman (TB) languages, and discuss what features of word order in these languages are typical and atypical. In section 3, I discuss word order in Chinese, identifying some typologically unusual features and discussing possible explanations for them. An overall theme shared by the two sections is that word order in Sino-Tibetan is best understood in an areal context.

2 WORD ORDER IN TIBETO-BURMAN

The discussion in this section summarizes briefly what I discuss in much greater depth in Dryer (2008). It is based on an examination of descriptions of 93 TB languages.

2.1 Order of object and verb and word order features that correlate with it

The distribution of the two orders of object and verb in TB is straightforward: all TB languages are ov, except for Bai and the Karen languages, which are vo (and more specifically svo). Although available data varies in the descriptions, the ov languages within TB generally share a variety of other word order characteristics typical of ov languages, in employing postpositions rather than prepositions, in placing genitive modifiers before the possessed noun, in placing relative clauses (if they are externally headed) before the head noun, in placing postpositional phrases before the verb, in employing clause-final markers for subordinate clauses, in placing markers of polar questions (if they employ them) at the end of sentences, and in placing auxiliary verbs after the main verb. An example of an exceptional feature found in a few TB languages is the placement of manner adverbs. While most TB languages more commonly place manner adverbs before the verb, a few TB languages, all of them Kuki-Chin-Naga languages, commonly if not preferentially,

place manner adverbs after the verb. This is described as the preferred position in Tiddim Chin (Henderson 1965) and Angami (Giridhar 1980) and is illustrated for Tiddim Chin in (1).

- (1) *'Hawi' ci in dawng zel zal a*
 hello say PTCL answer loudly PTCL
 v ADV

'He called out loudly, as if answering someone.' (Henderson 1965: 4, sentence 5)

In some cases, the vo TB languages exhibit the mirror image of the characteristics mentioned above for ov TB languages; however, in many other cases, they do not. Both Bai and the Karen languages employ genitive–noun order. In fact this order is the one word order feature that I have data on that is apparently shared by all Sino-Tibetan languages. In employing this order, Bai and the Karen languages are actually not atypical: as discussed in Dryer (1991), the two orders of genitive and noun are about equally common among svo languages.

The Karen languages exhibit a number of features that are atypical of vo languages. I illustrate this with features from Bwe Karen. While Bwe Karen does employ some clause-initial markers of subordinate clauses, there are also some clause-final subordinators, illustrated by *kha lé* 'if' in (2).

- (2) *nə-də̄e o kha lé, yə-kha o ka*
 2SG-if stay if 1SG-FUT stay then
 'If you stay, I will stay.' (Henderson 1997: 78)

Another feature of Bwe Karen that is atypical for a vo language is the placement of a word meaning 'able' after the main verb, as in (3).

- (3) *kə-pwa phá dō ə-kháchi də̄-ja-nə̄*
 1PL-build granary village POSS'D-near NEG-able-NEG
 'We can't build our granaries close to the village.' (Henderson 1997: 142)

While this is rather unusual among vo languages, it is something found in a number of vo languages in other families in southeast Asia, including Tai-Kadai (e.g. Nung: Saul and Wilson 1980: 47–8, 55), Mon-Khmer (e.g. Chrau: Thomas 1971: 97), and Hmong-Mienic (Hmong Njua: Harriehausen 1990: 179–80).

Bai also exhibits features atypical of vo languages. Foremost among these is the placement of relative clauses before the modified noun, as in (4).

- (4) *[yi⁴² tse²¹tsa⁴² no³³] sr⁵⁵ xə̄⁵⁵ ro⁴²*
 write tidy LINK word read easy
 'Words that are written tidily are easy to read.' (Xu and Zhao 1984: 73)

As discussed in greater detail in section 3 on Chinese, this order is extremely rare among vo languages. There is at least one word meaning 'able' that follows the main verb, as in (5).

- (5) *a⁵⁵na⁴⁴ li⁵⁵ ŋe²¹ ta⁴²*
 where all go able
 'I can go anywhere.' (Xu and Zhao 1984: 22)

And although Bai has some prepositions, it also has some postpositions, like *no³³* marking the indirect object in (6).

- (6) *ŋa⁵⁵ si³¹ nu⁵⁵ no³³ pe²¹xo⁵⁵ ku⁵⁵*
 1PL give 2SG OBJ flower CLSFR
 'We gave you a flower.' (Xu and Zhao 1984: 51)

2.2 Noncorrelating word order characteristics

As discussed in detail in Dryer (1992), there are a number of word order characteristics which, contrary to widespread belief, do not correlate with the order of object and verb. These include the order of adjective, numeral, and demonstrative with respect to a modified noun and the order of degree words with respect to a modified adjective. Among the vast majority of ov languages in Asia that are not TB, these pairs of elements occur in the order modifier–modified, and this has led some linguists to the mistaken belief that these features are to be expected of ov languages. However, as shown in Dryer (1992), it is not the case that these features are typical of ov languages. For example, with respect to the order of adjective and noun, it is actually somewhat more common for these to follow the noun in ov languages outside of Asia. The ov TB languages are in many respects atypical among ov languages in Asia, but normal for ov languages in the world as a whole, in that in most ov TB languages, some of these modifiers normally follow the modified word. The distribution of these word order characteristics among TB languages is also interesting in that there is considerable variation in their distribution, and it is often the case that even within a given subgroup of TB, some languages will employ one order while others employ the opposite order. It is possible to describe this variation only very briefly here; I discuss it in much greater detail in Dryer (2008).

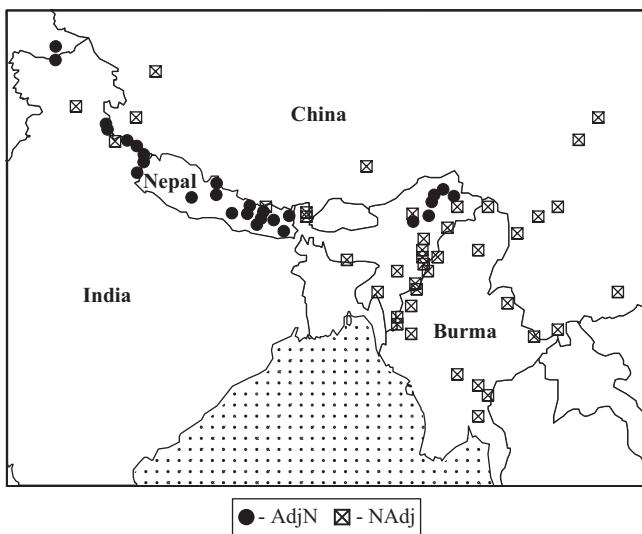
2.2.1 Order of adjective and noun

Both orders of adjective and noun are well-attested as preferred orders among TB languages. Among the 85 TB languages for which I was able to obtain information on this, the preferred order is *ADJN* in 28 languages and *NADJ* in 40 languages, and in 17 languages, both orders occur without any indication in my source that one order is preferred. Furthermore, assuming for the purposes of discussion the classification of TB languages proposed by Bradley (1997), in three of Bradley's six highest-level subgroups of TB (Bodic, North-Eastern India, and Central), there are some languages in which *ADJN* is the preferred order and other languages in which *NADJ* is the preferred order. In the other three subgroups, all of the languages are either *NADJ* or allow both orders, with neither order dominant. Even within a number of groups at the next level down in Bradley's classification, there are four groups containing languages of each of the two types (Bodish, Bodo-Garo, Tani, and Digarish ‘Mishmi’). For example, within Bodo-Garo, Deuri (Brown 1895) is *ADJN*, while Kokborok (Karapurkar 1976) and a few others are *NADJ*.

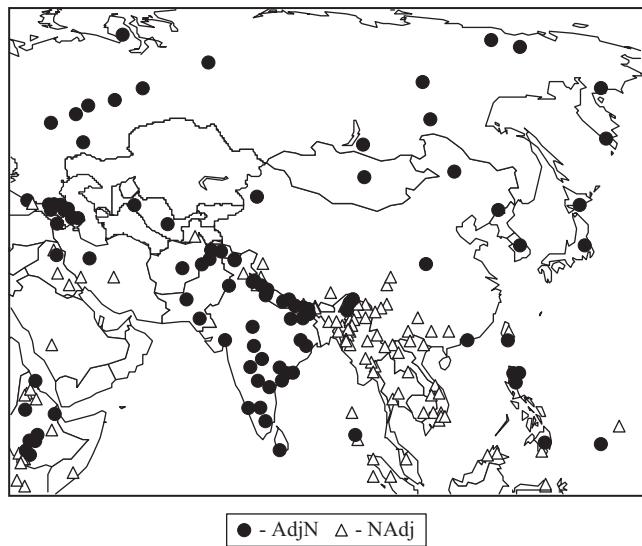
The geographical distribution of the two orders of adjective and noun is shown in Map 3.1. This map makes clear the extent to which *ADJN* order is found in the western part of the area in which TB languages are spoken while *NADJ* order is found more to the east.

This distribution can be understood in terms of the distribution of the two orders in non-TB languages in the surrounding area. The languages to the west and southwest of TB, especially the Indic languages within Indo-European, are consistently *ADJN*, while the languages to the east of TB other than Chinese, namely, Tai-Kadai and Mon-Khmer languages, are consistently *NADJ*. Map 3.2 shows the distribution of the two orders of adjective and noun in a larger area of Asia that surrounds TB languages. The overall impression given by Map 3.2 is two clear areas in South and Southeast Asia, one to the west which is *ADJN* and one to the east which is *NADJ*. But these two areas split TB down the middle.

Note that the clearest exceptions to the tendency for *ADJN* order in the west are a number of *NADJ* TB languages in the extreme western side of TB. These languages are all languages that are closely related to Tibetan, however, and they thus represent either



MAP 3.1 ORDER OF ADJECTIVE AND NOUN



MAP 3.2 ORDER OF ADJECTIVE AND NOUN IN ASIA

languages which have moved into that area relatively recently, and thus have been less subject to influence from Indic languages, or languages whose less accessible location in the Himalayas has also meant that they have been less subject to influence from Indic languages.

2.2.2 *Order of demonstrative and noun*

The overall pattern of the distribution of the two orders of demonstrative and noun among TB languages is somewhat similar to the distribution of the two orders of adjective and noun: DEMN order is more common to the west, while NDEM order is more common to the east. However, DEMN order is more common than ADJN order and there are a number of languages which are DEMN but NADJ. Among the 79 TB languages for which I was able to obtain information on the order of demonstrative and noun, 51 are DEMN, 20 are NDEM, three allow both orders without there being evidence for one order being dominant, and five normally have demonstrative words simultaneously preceding and following the noun, as in the example in (7) from Nishi.

- (7) *sa nyem si*
 here woman this
 ‘this woman’ (Hamilton 1900: 20)

As with the order of adjective and noun, we find both orders represented within the same subgroup. In four of Bradley’s highest-level subgroups there are some languages which are DEMN and others which are NDEM (Bodic, Kuki-Chin, North-Eastern, and South-Eastern). Note that this set of subgroups is very different from the analogous set with the order of adjective and noun: Bodic is the only subgroup in both sets. In other words all six subgroups are inconsistent, either in the order of adjective and noun or in the order of demonstrative and noun. Again we find languages of each type even within lower-level subgroups. For example, among the Burmish languages, Maru (Clerk 1911) is DEMN while Achang (Dai 1985) is NDEM.

2.2.3 *Order of numeral and noun*

There is less variability in the order of numeral and noun among TB languages. The only languages in which NUMN is dominant are Bodic, although both types are about equally common in Bodic. Within Bodic, the distribution is largely predictable from lower-level subgroups in Bradley’s classification: West Himalayish and Kiranti are NUMN while Central Bodish (Tibetan), Western Bodish (Tamangic), and Eastern Bodish (e.g. Monpa) are NNUM. Within what Bradley classifies as Central Himalayan, all are NUMN except for Newari; however, the classification of Newari is notoriously problematic.

2.2.4 *Order of degree word and adjective*

I was able to obtain data on this characteristic for fewer languages, and I did not include affixes expressing degree. Again, there is a split, with 25 languages in which DEG \wedge ADJ is preferred, 11 languages in which ADJ \wedge DEG is preferred, and four in which both orders occur with no evidence that one is preferred. The subgroup with the most ADJ \wedge DEG languages is actually geographically central, namely, Kuki-Chin. And again, we find inconsistencies within lower-level groups. Among the western languages of Central Bodish, Jad (Sharma 1989) is ADJ \wedge DEG while Nyamkad (Sharma 1992) is DEG \wedge ADJ.

3 CHINESE

I use the expression ‘Chinese languages’ to apply to what are traditionally called ‘dialects’ of Chinese, following the use of these terms within linguistics. I will, however, largely

restrict discussion to Mandarin. There are some differences in word order among the different Chinese languages, but I will generally ignore these here.

This section has two primary purposes. The first is to document the extent to which Mandarin is unusual in its word order in a number of ways. The second is to argue that these unusual characteristics are best understood areally and reflect the geographically intermediate position of Mandarin between the languages of Southeast Asia and the languages of Northeast Asia. In section 3.1, I describe the word order characteristics of Mandarin, discussing ways in which it is unusual, and in section 3.2, I discuss possible areal accounts of these unusual characteristics.

3.1 Unusual word order characteristics of Mandarin

The dominant order at the clause level in transitive clauses in Mandarin is *svo*, as in (8).

- (8) a. *wǒ xǐhuān Susan*
1SG like Susan
'I like Susan.'

b. *Susan xǐhuān wǒ*
Susan like 1SG
'Susan likes me.'

sov word order is also possible, as in (9), though this word order is less common (see Sun and Givón 1985 for evidence from narrative and oral text).

- (9) *wō bā shū mǎi le*
 1SG OBJ book buy CURR.REL
 ‘I bought the book.’ (Li and Thompson 1981: 21)

Since the dominant order in Mandarin is VO, we might expect it to have prepositions rather than postpositions. There are in fact a couple of different sorts of words that can be classified as adpositions. There is a set of words that can be described as prepositions, illustrated by the words *ba* ‘object marker’ in (9), *cóng* ‘from’ in (10a), and *dào* ‘to’ in (10b).

- (10) a. *tā cóng Zhōngguó lái le*
 3SG from China come CURR.REL
 ‘He/she has come from China.’ (Li and Thompson 1981: 24)

b. *wǒmen fēi dào Shànghǎi le*
 1PL fly to Shanghai CURR.REL
 ‘We flew to Shanghai.’ (Li and Thompson 1981: 410)

These words are called coverbs by Li and Thompson (1974), and different coverbs seem to be in different stages of grammaticalization from their original status as verbs to their current status as more preposition-like elements. I will assume here that this grammaticalization has proceeded far enough for some of them to justify calling themselves prepositions.

¹ Mandarin also has postpositions, as illustrated by *qī* ‘from’ (< ‘start’) in (11a) and *shìde* ‘like’ in (11b) (these examples kindly provided by Liu Danqing).

- (11) a. *tā* (*cóng*) *míngtiān* *qǐ* *shàngbān*
 3SG (from) tomorrow from go.to.work
 ‘He will go to work from tomorrow on.’

b. *tā* *xiàng* *húli* *shìde* *jiǎohuá*
 3SG like fox like sly
 ‘He is sly like a fox.’

Note how the postposition in (11a) optionally co-occurs with a preposition; however, the postposition in this example and in similar constructions is generally obligatory. The number and frequency of both prepositions and postpositions in Mandarin suggest that it is best not to say that the language is primarily prepositional or primarily postpositional. Nevertheless, the frequency of postpositions is somewhat unexpected of a vo language.

These adpositions combine with noun phrases to form adpositional phrases. But these adpositional phrases more often precede the verb, as in (9), (10a), and (11), though they sometimes follow the verb, as in (10b). Mandarin is thus an instance of a language in which the object normally follows the verb but adpositional phrases more often precede the verb. This makes Mandarin highly unusual from a cross-linguistic perspective: among 199 vo languages in my database that I code for order of adpositional phrase and verb, only three are PP-V, while the other 196 are V-PP. The three that are PP-V are all Chinese languages: Mandarin, Cantonese, and Hakka. The Chinese languages are thus the only known instances of languages of this sort. It is not clear to me, however, how old the PP-V order in Chinese is. The claims in the literature regarding the position of PPs in Classical Chinese are somewhat contradictory. On the one hand, Li and Thompson (1974) claim that classical Chinese was predominantly V-PP and that PP-V is a development since the twelfth century CE. On the other hand, Sun (1987) argues that both orders of PP and V were common in Classical Chinese.

The comparative construction in Mandarin also uses a coverb construction, as in (12).

- (12) *Zhāngsān bǐ tā pàng*
 Zhangsan than 3SG fat
 ‘Zhangsan is fatter than him/her.’ (Li and Thompson 1981: 142)

The construction in (13) is Marker–Standard–Adjective (*bǐ + tā + pàng*). Again, this order is quite rare cross-linguistically; only Mandarin and Hakka in my database have this order, the typical order in vo languages being Adjective–Marker–Standard, as in English *taller than Mary*. In fact, this is the normal order in Cantonese, as in (13).

- (13) *gāmyaht yiht gwo kàhmyaht*
 today hot than yesterday
 ‘Today is hotter than yesterday.’ (Matthews and Yip 1994: 166)

Both orders are possible in both Mandarin and Cantonese, the difference being which is the dominant construction.

Although manner adverbs can either precede or follow the verb in Mandarin, their normal position is preverbal, as in (14).

- (14) *tā kuài-kuài-de zǒu*
 3SG quickly walk
 ‘She/he walked quickly.’ (Li and Thompson 1981: 323)

This is also not the usual order for vo languages, it being more common for manner adverbs to follow the verb, though there are a number of other vo languages outside Chinese which more commonly place manner adverbs before the verb (such as Latvian).

Mandarin is also unusual in being a vo language that places relative clauses before the noun, as in (15).

- (15) *[wǒ gěi nǐ de] shū*
 1SG give 2SG LINK book
 ‘the book [that I gave you]’ (Li and Thompson 1981: 117)

Among 254 vo languages in my database for which I code the order of relative clause and noun, all are NREL, except for the three Chinese languages (Mandarin, Hakka, and Cantonese) and Bai, as illustrated in this section.

There are three characteristics where SVO languages are intermediate between verb-initial and verb-final (Dryer 1991), where some SVO languages exhibit the order associated with verb-initial and others exhibit the order associated with verb-final. For all three of these, Mandarin exhibits the order associated with verb-final languages. One of these is the use of GENN order in genitive constructions, as in (16).

- (16) *tùzi de ērduō*
rabbit LINK ear
'the rabbit's ear' (Li and Thompson 1981: 113)

A second is the use of sentence-final question particles, as in (17).

- (17) *nǐ néng xiě Zhōngguó zì ma?*
2SG can write Chinese character Q
'Can you write Chinese characters?' (Li and Thompson 1981: 547)

A third is the fact that interrogative phrases in content questions do not need to occur at the beginning of the clause, typically occurring *in situ*, in the position that corresponding non-interrogative phrases would occur in, as in (18).

- (18) *nǐmen zuò shénme*
2PL do what
'What are you doing?' (Li and Thompson 1981: 522)

The last two of these characteristics are also ones shared by other vo languages of South-east Asia, in Hmong-Mienic, Tai-Kadai, and Mon-Khmer. The GENN order contrasts, however, with the NG order of Tai-Kadai and Mon-Khmer.

There are other respects in which Mandarin behaves more like a typical vo language. The predicate follows the copula, as in (19).

- (19) *wǒ shì Susan*
1SG be Susan
'I am Susan.'

Words meaning 'able' precede the other verb, as in (17). Verbs meaning 'want' precede the verb denoting what is wanted, as in (20).

- (20) *wǒ yào qù Zhōngguó*
1SG want go China
'I want to go to China.'

To a large extent, the inconsistencies in Mandarin word order can be characterized in terms of the distinction between two types of dependents, what in different grammatical traditions have been called complements or arguments on the one hand vs adjuncts or modifiers on the other. Mandarin typically places the head before a complement, a dependent which is required grammatically and semantically to complete the meaning of the phrase: verb before object, adposition before object, copula before predicate, verbs with meanings like 'want' or 'able' before their verbal complements. On the other hand, Mandarin typically places heads after adjuncts (i.e. modifiers), dependents which are not required grammatically or semantically but which optionally elaborate on the meaning of their phrases; this is reflected in placing nouns after adjectives, relative clauses and other

modifiers of nouns, adjectives after intensifiers or standards of comparison, and verbs after manner adverbs and adpositional phrases. Whether this pattern is more than a coincidence, either synchronically or diachronically, is not clear.

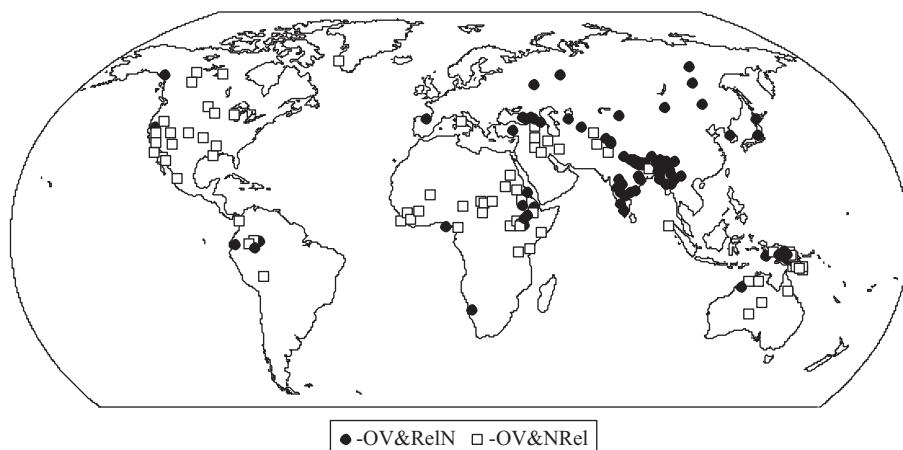
3.2 Chinese word order from a geographical perspective

We have seen in the preceding section that Mandarin (and other Chinese languages) has a number of characteristics that are highly atypical of vo languages. Can we offer anything to explain why Chinese might have these unusual characteristics? I will start with the assumption that Proto-Sino-Tibetan was ov, reLN, and pp-v, and that the reLN and pp-v orders are at least partly a retention of these features from Proto-Sino-Tibetan (LaPolla 1994, Chapter 2, this volume; Liu 1999). Let me focus on two of these characteristics, vo&reLN (vo with prenominal relative clauses) and vo&pp-v (vo with preverbal adpositional phrases). The cross-linguistic rarity of these types implies that there are some causal factors discouraging such languages from arising in the first place and possibly also encouraging such languages to change to some other type if they do arise. Languages elsewhere in the world which were ov&reLN or ov&pp-v and which became vo have apparently also become nrel and v-pp, either simultaneously or shortly after becoming vo. Chinese, however, has apparently retained these characteristics for a long time. reLN order is also the dominant order across TB, suggesting that the reLN goes back to Proto-Sino-Tibetan. Since vo order is also apparently fairly old in Chinese, dating back to Proto-Sinitic or close to that, this means that Chinese has been vo and reLN for a long time. The situation with pp-v order is less clear, as noted above: either it too dates all the way back to Proto-Sino-Tibetan, or Chinese has moved from pp-v towards v-pp and back towards pp-v.

We cannot answer the puzzle by simply saying that reLN and pp-v order are retentions from Proto-Sino-Tibetan, since other instances of ov languages changing to vo have apparently invariably ceased to be reLN and pp-v. What is striking, however, is the extent to which Chinese languages resemble languages to the north, including Japanese, Korean, Mongolian, Tungus, and Turkic. These languages also place relative clauses before the noun and adpositional phrases before the verb. Of course, since these are characteristics associated with ov languages in general, the fact that Chinese resembles languages to the north in these respects is no different from saying that it resembles ov languages elsewhere in the world.

But there are other ways, however, in which Chinese resembles ov languages to the north far more than it resembles ov languages elsewhere in the world. As shown in Dryer (1992), the two orders of relative clause and noun are about equally common in ov languages. The reLN order in Chinese is an ov characteristic only in the sense that reLN languages are usually ov. The opposite is not the case: it is not the case that ov languages are generally reLN. NREL order is in fact as common as reLN order among ov languages. The cross-linguistic distribution of the two orders of relative clause and noun among ov languages partly follows an areal pattern: reLN order is more common in Eurasia, while nrel order is more common elsewhere in the world. Hence the reLN order in Chinese cannot be simply viewed as an ov characteristic; rather it is a characteristic associated with ov languages in Asia, both in TB and those north of Chinese. Map 3.3 shows the cross-linguistic distribution of the two orders of relative clause and noun among ov languages. Map 3.3 shows clearly how reLN order is more common in Eurasia, particularly eastern Asia, in the area surrounding Chinese (except to the south, where the languages are not ov).

The reLN order of Chinese resembles the common reLN order to the north and to the west, in TB. But there are other respects, however, in which Chinese word order



MAP 3.3 WORLDWIDE DISTRIBUTION OF TWO ORDERS OF RELATIVE CLAUSE AND NOUN AMONG OV LANGUAGES

resembles word order in languages to the north more closely than word order in TB languages. Consider, for example, the order of adjective and noun. We saw in section 2.2.1 how both orders of adjective and noun are found in TB, though Nadj order is somewhat more common, particularly in the east, towards Chinese. On the other hand, languages to the north are consistently AdjN. When we look at the distribution of AdjN and Nadj order in Asia in Map 3.2, we see that Chinese is situated between a large set of Nadj languages to the south (Tai-Kadai and Mon-Khmer) and southwest (eastern TB languages) and a large set of AdjN languages to the north. The order of adjective and noun in Mandarin thus more closely resembles the languages to the north than many TB languages, especially those that are situated more closely to Chinese.

The situation regarding the position of demonstratives and numerals relative to the noun and the position of intensifiers relative to the adjective is similar: Chinese languages consistently place the modifiers before the modified element, like almost all languages in northeast Asia and unlike the majority of TB languages. The only TB languages like this are a subset of Bodic languages in Nepal and northwest India and these are the TB languages that are most distant geographically from Chinese. The tendency to consistently place modifying elements before the modified element is a property of ov languages of northern Asia. As discussed above, it is not a property of most ov languages outside Asia. In this way, therefore, Chinese resembles languages of northern Asia far more than it resembles TB languages or other ov languages, suggesting that these characteristics are best understood in terms of areal influence from languages of northeast Asia.

Someone wishing to deny the claim of areal influence from the north could take one of two approaches. One might try to argue that these characteristics reflect word order from an earlier time, perhaps going back to Proto-Sino-Tibetan. Since some of these characteristics (prepositional phrase and manner adverb before verb, standard of comparison before adjective, relative clause before noun) are ones generally found only in ov languages, this hypothesis would have to claim that Proto-Sino-Tibetan was ov and that Chinese has retained these characteristics, despite changing to vo order. Since TB

languages are also largely ov, and share these characteristics, this is not an implausible scenario for these characteristics. But this leaves two things unexplained. First, since these characteristics are so rare in vo languages other than Chinese, why has Chinese maintained them, when languages elsewhere in the world changing from ov to vo order apparently seldom if ever retain these ov characteristics? Second, while this explanation may make sense for the characteristics of Chinese that are generally associated with ov order, it does not explain why Chinese resembles languages to the north in ways that are not associated with ov order, namely, in placing adjectives, demonstratives, and numerals before nouns and intensifiers before adjectives. We would have to say that Proto-Sino-Tibetan also had these characteristics, coincidentally similar to languages to the north. And since the majority of TB languages do not have these characteristics, we would have to say that all these TB languages have lost these characteristics, except in the subset of Tibetic languages that are like Chinese in these ways. However, there is a more obvious explanation for the fact that these Tibetic languages place these modifiers before the modified element: they are also adjacent to languages which consistently place modifiers before the modified element, namely, Indic languages in Indo-European. The Indic languages (and also Dravidian) belong to a large arm of consistently premodifying languages that connects with the area in northern Asia where this is found via Pakistan, Afghanistan, Turkmenistan, Uzbekistan, and Kazakhstan.

In fact, even if the premodifying characteristics of Chinese are retained from Proto-Sino-Tibetan, it is likely that there has still been areal influence from the north in contributing to the Chinese retaining these characteristics. In general, it is probably the case that areal influence more often has an effect in encouraging languages to retain characteristics than in causing changes. In other words, even if Proto-Sino-Tibetan was consistently premodifying, like Modern Chinese, it is unlikely to be a coincidence that the languages that have retained these characteristics are precisely those (Chinese, southern and western Tibetic) that are adjacent to non-Sino-Tibetan languages that have exactly those characteristics, and that the languages which have lost these characteristics are ones that are not adjacent to such languages.

The idea that Chinese word order has been influenced from the north has been suggested by others, especially by Hashimoto (1986). Hashimoto provides a further type of argument for this influence, the fact that syntactic and phonological differences among Chinese languages follow a north–south pattern in that where one finds differences among Chinese languages, the languages to the north tend to be more similar to non-Sino-Tibetan languages (Tungus, Mongolian) to the north of Chinese. However, Hashimoto's discussion assumes (following views shown to be incorrect by Dryer 1992) that the pre-modifying order within noun phrases is an ov characteristic. But the fact that this is not an ov characteristic, the fact that adjectives, demonstratives, and numerals do not tend to precede the noun in ov languages actually provides further support for Hashimoto's position, since one cannot attribute these characteristics to Chinese being ov in the past (or moving towards ov).

An alternative hypothesis is that these characteristics of Chinese reflect internal changes that coincidentally led to characteristics that resemble languages to the north. In most cases, I do not think that this possibility should be ruled out, or even viewed as unlikely. When one examines the geographical distribution of typological characteristics, there are bound to be many instances of adjacent languages being similar by accident. However, the fact that Chinese is so unusual in some of these characteristics lowers the likelihood of coincidental resemblances, since there is a need to explain why Chinese has these characteristics when they are not found elsewhere in the world.

4 CONCLUSION

The most salient overall generalization about word order within Sino-Tibetan is that where one finds differences among languages, the different languages tend to be more similar in word order to adjacent non-Sino-Tibetan languages. In the last section, I have dwelt on the resemblances of Chinese to languages to the north, and have pointed out the resemblances of western and southern Tibetic languages to Indic languages, but I have also pointed out, in section 2, the fact that the more eastern TB languages more closely resemble Tai-Kadai and Mon-Khmer languages to the east. We see this in its strongest form with the Karen languages, which are vo, like languages to the east. We also see it in the overall tendency for postmodifying order for various sorts of modifiers to be more common towards the east of TB, in Lolo, Bai, Qiang, and Pumi. However, even towards the east, we find genn order everywhere, even in Karen, as well as reln order, except in Karen.

On the other hand, the details are much more complex than these overall patterns might suggest. We have seen that for a number of modifiers, such as adjectives modifying nouns, there is considerable diversity, even within subgroups of TB. In addition, Tibetan, and the Bodic languages most closely related to it, do not fit the overall east–west pattern within TB, since they are towards the west, yet they tend to place modifiers after the noun.

Nor, surely, should all the geographical patterns be understood in terms of non-Sino-Tibetan languages influencing Sino-Tibetan languages rather than the other way round. It is precisely because we find such variation within Sino-Tibetan, compared to most adjacent families, that it is possible to see how the variation within Sino-Tibetan can be understood in terms of languages within Sino-Tibetan resembling adjacent groups of languages. In some cases, it may be that the direction of influence may have gone from Sino-Tibetan to non-Sino-Tibetan, but where that might be the case is not clear.

NOTE

- 1 Part of the research for this paper was made possible by Social Sciences and Humanities Research Council of Canada Grants 410-810949, 410-830354, and 410-850540 and by National Science Foundation Research Grant BNS-9011190. I am indebted to Meihan Low for assistance with the sources written in Chinese.

REFERENCES

- Bradley, David (1997) ‘Tibeto-Burman languages and classification’, in David Bradley (ed.) *Tibeto-Burman Languages of the Himalayas*, Papers in Southeast Asian Linguistics, No. 14. Pacific Linguistics, Series A, No. 86, 1–72.
- Brown, William Barclay (1895) *An Outline Grammar of the Deori Chutiya Language Spoken in Upper Assam, with an Introduction, Illustrative Sentences, and Short Vocabulary*. Shillong: Assam Secretariat Printing Office.
- Clerk, F.V. (1911) *A Manual of the Lawngwaw or Maru Language: Containing the Grammatical Principles of the Language, Glossaries of Special Terms, Colloquial Exercises, and Maru–English and English–Maru Vocabularies*. Rangoon: American Baptist Mission Press.
- Dai, Qingxia (1985) *Achangyu Jianzhi* (A brief description of the Achang language), Beijing: Minzu Chubanshe.
- Dryer, Matthew S. (1991) ‘SVO languages and the OV:VO typology’, *Journal of Linguistics* 27: 443–82.

- Dryer, Matthew S. (1992) ‘The Greenbergian word order correlations’, *Language* 68: 81–138.
- Dryer, Matthew S. (2008) ‘Word order in Tibeto-Burman languages’, *LTBA* 31: 1–88.
- Giridhar, Puttushetra Puttuswamy (1980) *Angami Grammar*, Manasagangotri, Mysore: Central Institute of Indian Languages.
- Hamilton, R.C. (1900) *An Outline Grammar of the Dafla Language*, Shillong: Assam Secretariat Printing Office.
- Harriehausen, Bettina (1990) *Hmong Njua*, Tübingen: Max Niemeyer.
- Hashimoto, Mantaro J. (1986) ‘The Altaicization of Northern Chinese’, in John McCoy and Timothy Light (eds) *Contributions to Sino-Tibetan Studies*, Leiden: E.J. Brill, 76–97.
- Henderson, Eugénie J.A. (1965) *Tiddim Chin*, London Oriental Series, Vol. 15. London: Oxford University Press.
- Henderson, Eugénie J.A. (1997) *Bwe Karen Dictionary: With Texts and English–Karen Word List*, London: School of Oriental and African Studies, University of London.
- Karapurkar, Pushpa Pai (1976) *Kokborok Grammar*, Central Institute of Indian Languages, Grammar Series 3. Mysore: Central Institute of Indian Linguistics.
- LaPolla, Randy J. (1994) ‘On the change to verb-medial word order in Proto-Chinese: evidence from Tibeto-Burman’, in Hajime Kitamura, Tatsuo Nishida and Yasuhiko Nagano (eds) *Current Issues in Sino-Tibetan Linguistics*, Osaka: National Museum of Ethnology, 98–104.
- Li, Charles and Thompson, Sandra (1974) ‘Co-verbs in Mandarin Chinese: verbs or prepositions?’ *Journal of Chinese Linguistics* 2.3: 257–78.
- Li, Charles N. and Thompson, Sandra A. (1981) *Mandarin Chinese: A Functional Reference Grammar*, Berkeley: University of California Press.
- Liu, Danqing (1999) ‘Word order in the Chinese clause over time and across dialects’, unpublished manuscript.
- Matthews, Stephen and Yip, Virginia (1994) *Cantonese: A Comprehensive Grammar*, New York: Routledge.
- Saul, Janice E. and Wilson, Nancy Freiberger (1980) *Nung Grammar*, Dallas: Summer Institute of Linguistics and University of Texas at Arlington.
- Sharma, Devidatta (1989) *Tibeto-Himalayan Languages of Uttarakhand*, Vol. 1, New Delhi: Mittal Publications.
- Sharma, Devidatta (1992) *Tribal Languages of Himachal Pradesh, Part II: Studies in Tibeto-Himalayan Languages II*, Delhi: Mittal Publications.
- Sun, Chao-Fen (1987) ‘The adposition *yi* and word order in Classical Chinese’, *Journal of Chinese Linguistics* 19: 202–19.
- Sun, Chao-Fen and Givón, T. (1985) ‘On the so-called SOV word order in Mandarin Chinese’, *Language* 61: 329–51.
- Thomas, David D. (1971) *Chrau Grammar*, Honolulu: University of Hawaii Press.
- Xu, Lin and Yansun, Zhao (1984) *Baiyu jianzhi*, (A brief description of the Bai language), Beijing: Minzu Chubanshe.

Part 2

SINITIC

Review Copy – Not for Redistribution

Randy J. LaPolla - Nanyang Technological University - 28/03/2017

CHAPTER FOUR

THE SINITIC LANGUAGES: PHONOLOGY

Zev Handel

1 INTRODUCTION

The Chinese (also called Sinitic) language family is a diversified group of languages and dialects descended from a common ancestor spoken perhaps no more than about 2,000 years ago. In terms of time depth and degree of linguistic diversity, it has been compared to the Romance language family descended from Latin (Norman 1988: 187). Chinese as a written language is attested as early as 1250 BCE in the Shāng dynasty oracle bone inscriptions. Based on the interpretation of first-millennium-BCE rhyming texts and of the structural elements of Chinese characters a stage of the language spoken in the central plains area of northern China termed Old Chinese (OC) can be reconstructed. (For more on the characteristics of OC, see later in the chapter.) However, many of the features reconstructed for OC are not reflected in later varieties of Chinese. Taking into consideration migration history and the results of comparative-historical analysis of modern Chinese languages, it appears likely that the common ancestor of modern Sinitic was a late OC *koine* that spread rapidly during the great Qín-Hàn imperial expansion toward the end of the first millennium BCE.¹

Chinese is often described, in conformity with the perceptions of Chinese speakers themselves, as a single language with many regional dialects. From a linguistic perspective, however, the degree of diversity and lack of mutual intelligibility among the regional varieties demonstrate that the “dialects” are in fact distinct, if closely related, languages. The Standard Mandarin word for the regional varieties of Chinese, *fāngyán* 方言, is usually translated as “dialect” in English.² However, as many scholars have pointed out (e.g. Mair 1991), this is misleading. The term *fāngyán* literally means “regional speech,” and does not draw any distinction between mutually intelligible and unintelligible varieties. Even when used in a technical sense by Chinese linguists, the term *fāngyán* is not equivalent to the English technical term *dialect*.

It is not known how many distinct Chinese languages there are. Based on the criterion of mutual intelligibility, the number must be in the range of anywhere from several dozen to several hundred (Norman 1988: 187, 2003: 72). But the nature of the dialect continua in China makes it exceedingly difficult to identify language boundaries in a consistent and meaningful way; to my knowledge nobody has ever attempted to do so in a rigorous fashion.

Chinese dialectologists divide the family into between six and twelve major groupings called *dà* 大 (large) *fāngyán*. These major groupings each have millions of speakers and encompass a high degree of linguistic diversity, including mutually unintelligible varieties. They are sometimes called *dialect groups* in English. In the pre-modern era they were conventionally and imprecisely defined by geographic location and impressionistic linguistic features. Within the field of Chinese dialectology in the twentieth century, these older definitions of the dialect groups were refined using historical-phonological criteria in relation to the phonological categories of reconstructed Middle Chinese (MC) (Li 1937

is an influential early example; see Li 1973). In more recent decades the validity of these criteria for classification has been challenged, and some scholars have argued for a more varied set of classificatory criteria based on multiple linguistic sub-systems (e.g. Norman 1988: 181ff.). On a broader level, it is legitimate to ask whether any such classificatory system is worth pursuing; that is, whether the dialect groups that result from the application of classificatory criteria are meaningful in terms of having useful linguistic applications beyond the classification itself. Regardless of how this broader question is answered, the fact remains that the dialect groups remain useful as shorthand categories for bundles of linguistic features, and scholars continue to use them as a framework for discussions of Chinese dialectal diversity and typology.

All of the considerations described above present considerable challenges of nomenclature for scholars who wish to talk about speech varieties within the Chinese language family. The need for technical precision and accuracy must be balanced against the weight of tradition and the misunderstandings that could arise by introducing new terms into an already crowded and complex field.

In this chapter, I will employ terminology in the following way:

- 1 The entire family of languages will be referred to as “Chinese” (equivalent to “Sinitic”).
- 2 The *dà fāngyán* will be referred to as “dialect groups.”
- 3 The term “variety” will be used to refer to various forms of Chinese without prejudging the question of mutual intelligibility.
- 4 The term “dialect” will be reserved for labeling varieties that are limited to a small geographic area such as a city or county.

Note that this approach differs somewhat from that taken by Norman (1988, 2003).

Chinese belongs to a larger language family, Sino-Tibetan (see the introduction to this volume). To continue the analogy invoked earlier, the position of Chinese within a larger and older family is like the position of the Romance languages within the larger Indo-European family. This naturally raises the question of how we can confidently distinguish language varieties that belong to Chinese from those that are related to, but not part of, the Chinese family. Definitionally this is a simple matter. We can say that any speech variety that is descended from the most recent common ancestor of Běijīng 北京 Chinese, Fúzhōu 福州 Chinese, and Guǎngzhōu 廣州 Chinese (dialects within the Mandarin, Mǐn 閩, and Yuè 粵 dialect groups, respectively) belongs to the Chinese language family. That common ancestor, as noted above, was probably a species of OC spoken during the Hán dynasty (206 BCE to 220 CE). Impressionistically, it is also rather simple to identify a speech variety as belonging to Chinese. The family’s typological coherence, large number of obvious cognates within basic vocabulary, and high degree of superficial similarity in syntax, morphology, and phonology all serve to make Chinese languages readily identifiable. Just as one might feel confident identifying a Romance language on the basis of just a few dozen vocabulary words, one can similarly feel confident about identifying a Chinese language based on only a cursory inspection of its linguistic features. In most cases, the line between Chinese and non-Chinese languages is sharp and clear. However, the definition of what is Chinese given earlier does not provide a method for determining whether a given language is Chinese or not, and the impressionistic approach lacks objective rigor.

A rigorous approach to distinguishing Chinese from non-Chinese varieties requires identifying innovations within the early history of Chinese—i.e. at the OC stage—that distinguish it from other members of Sino-Tibetan, and that are part of the common inheritance of all later varieties of Chinese. If a given language demonstrates the presence of

these innovations (assuming that we can exclude the possibility that they originate from contact-induced convergence)—and lacks innovations seen in other branches of Sino-Tibetan—then it can be confidently assigned to the Chinese family. There do in fact appear to be a number of criteria that distinguish Chinese from related non-Chinese languages, including lexical, phonological, and morphosyntactic features (LaPolla 2013: 124). However, the precise formulation of these depends in part on how OC and Proto-Tibeto-Burman are reconstructed, and there is no scholarly consensus on which criteria to use or how to apply them.

Currently, the only language whose “Chinese-ness” is actively debated is Bai. The affiliation of Bai is difficult to determine because of its long history of intimate contact with various Chinese varieties. It is not a simple matter to strip away layers of Chinese borrowing and uncover the core features of Bai in order to determine if they are Chinese or Tibeto-Burman. There are also a number of newly formed hybrid languages developing on the Chinese language frontier (Dede 2016 forthcoming; Ansaldi 2016 forthcoming); their status as Chinese languages is also problematic. We will set aside these marginal cases in the current chapter, and focus on phonological features of those varieties that are uncontroversially part of the Chinese family.

The universally recognized dialect groups are Mandarin (Běifānghuà 北方話), Wú 吳, Xiāng 湘, Gàn 贛, Hakka (Kèjiā 客家), Mǐn 閩 and Yuè 粵. Mandarin speakers number well over 800 million; the other groups all boast tens of millions of speakers. Aside from these seven, many scholars recognize additional groups, based on a combination of geographic, synchronic, and diachronic features. These include a Jin 晉 group concentrated in Shānxī, a Huī 徽 group concentrated in southern Ānhuī, and a Píng(huà) 平(話) group concentrated in Guǎngxī. Because there is no generally accepted set of criteria for dividing dialect groups, the recognition of these groups remains controversial, with some scholars considering them varieties of Mandarin, Wú and Yuè, respectively. As seen in Map 4.1, Mandarin is spoken across a wide swath of China, comprising not only its homeland in the north but also areas of the west and south that have seen large-scale Chinese immigration in the last several hundred years (LaPolla 2001). The remaining six groups are confined to the southeast of the country. There are also significant Chinese-speaking communities outside of China resulting from emigration to other parts of the world.

This chapter provides an overview of the phonological features of Chinese both synchronically and diachronically. Every attempt is made to describe these features in a way that avoids the highly specialized terminology used within the field of Chinese dialectology, terminology that is drawn from the native Chinese tradition of phonological analysis. To some degree, however, the use of such terminology is unavoidable. Some of these terms reflect concepts that are useful for the description and analysis of Chinese, and for which no equivalent generic term exists. Moreover, the use of this terminology is so widespread within the field that to avoid it altogether would prevent the interested reader from pursuing the topic in more detail by consulting other publications. We will therefore introduce and explain some traditional terminology as needed.

2 SYNCHRONIC PHONOLOGY: FEATURES OF MODERN CHINESE VARIETIES

2.1 Common features and generalizations

All varieties of Chinese, with few exceptions, share a number of phonological features. Some of these are areal features seen also in neighboring languages, notably those of southern China and Southeast Asia.



MAP 4.1 THE MAJOR CHINESE DIALECT GROUPS

2.1.1 Syllable structure

The basic level of phonological analysis for Chinese is the syllable. This is not just a matter of tradition. There is a strong tendency in modern Chinese varieties toward what might be broadly called “monosyllabicity.” Most morphemes are one syllable long, and most syllables are identifiable as morphemes. Phonological rules (both synchronic and diachronic) tend to operate within rather than across syllable boundaries, despite the fact that polysyllabic (especially disyllabic) words, most of which are transparently compounded of monosyllabic morphemes, occur frequently in the lexicon. (The one notable exception to this generalization is tone sandhi, discussed later.)

The native Chinese phonological tradition divides the syllable into three component parts: initial (*shēngmǔ* 聲母), final (*yùnmǔ* 韻母), and tone (*shēngdiào* 聲調). The initial (I) is the consonantal onset (or zero). The final (F) contains all of the segmental material that follows the initial. The tone (T) is a phonemically distinctive pitch pattern (incorporating features of pitch height, pitch contour, and sometimes distinctive phonation or length) imposed on the syllable. The initial and final correspond roughly, but not exactly, to the modern concepts of onset and rime.

The syllable schema $\sigma = IF^T$ can be used in the analysis of all modern varieties of Chinese. Most descriptive publications of Chinese varieties do not present phonemes in consonant and vowel charts. Rather, they present separate charts of initials, finals, and tones. We will follow this approach here.

2.1.2 Initials

Across all modern varieties of Chinese, the syllable initial slot can only be filled by a single consonant or zero. Consonant clusters do not occur. (Sequences like [ts] are best analyzed as unitary affricate phonemes /ts/. For consistency and typological convenience we will follow convention and write them without a tie bar.) The zero consonant /Ø/ may be realized phonetically in a variety of ways, such as [?] (Li 1966). Note that not all consonant phonemes found in a Chinese variety can necessarily occupy the initial slot of a syllable. For this reason, one cannot assume that the chart of initials typically found in a published phonological inventory of a Chinese variety reflects all of its consonant phonemes. For example, in Standard Mandarin (Pǔtōnghuà 普通話), the official Chinese of the People's Republic of China, whose phonology is based on the educated speech variety of Bēijīng, /ŋ/ is not a possible syllable initial. It does occur, however, in syllable coda position (i.e. as the last segment of a syllable final), where it contrasts with /n/.

Consonant inventories of Chinese varieties range from about 15 to 30 phonemes. The dialect of Jiàn'ōu 建甌, spoken in northern Fújiàn province and belonging to the Mǐn dialect group, provides an example of a Chinese variety with only 15 initial consonants (counting the zero initial).

The inventory in Table 4.1 is typical of Chinese varieties with fewer consonant phonemes, in that it exhibits a basic three-way place-of-articulation contrast (labial, dental, velar) and a two-way aspiration distinction in (voiceless) stops and affricates. These distinctions are found in all Chinese varieties.

A particularly large consonant inventory is seen in the dialect of Chénxī 辰溪 in Húnán province, located in the border area between the Mandarin and Xiāng 湘 dialect groups. There are a total of 31 initial consonants (counting the zero initial).

The largest consonant inventories found in Chinese varieties, like that in Table 4.2, have a three-way distinction in stops and affricates (adding a voicing distinction to the aspiration distinction) and three sets of sibilants (at dental, retroflex, and palatal places of articulation). Most of these larger inventories are found in varieties belonging to the Wú and Xiāng dialect groups.

Both of the inventories presented in Tables 4.1 and 4.2 are typical of Chinese in that they lack “exotic” sounds (i.e. typologically rare or marked sounds) other than retroflexes.³ As a general rule clicks, implosives, ejectives, uvulars, interdentals and the like seldom occur in Chinese. Below are a few examples of “unusual” consonant sounds in Chinese:

- Twentieth-century descriptions of Xī'ān note that older speakers have a series of labial affricates /pf pfʰ/ corresponding to Standard Mandarin retroflexes before [u] or [w], as in the word for ‘spring’: [pʰfɛ²¹] (cf. Standard Mandarin chūn /ʈʂʰwan⁵⁵/ 春).
- Some varieties have a voiceless lateral fricative [l], for example the Táishān 台山 and Yángjiāng 陽江 dialects of Yuè, e.g. Táishān [ɬam⁵⁵] 三 ‘three’ (cf. Cantonese [sam⁵⁵]).
- A number of southern varieties, notably the Mǐn dialects spoken on Hainán island, have implosive stops. These are believed to result from areal influence. Wénchāng 文昌 has implosives /ɓ d̪/ that contrast phonemically with voiced stops /b d/ and voiceless stop /t/, as in [du²¹] 堵 ‘to block up’ vs. [du²¹] 你 ‘2sg’ vs. [tu²¹] 主 ‘main.’

2.1.3 Finals

Within the field of Chinese dialectology, syllable finals are divided into three components or slots: the obligatory monophthongal main vowel (equivalent to the nucleus), preceded

TABLE 4.1 THE INITIALS OF JIÀN'ÓU DIALECT (ADAPTED FROM BĚIJÍNG DÀXUÉ ZHÖNGGUÓ YÜYÁN WÉNXUÉ XÌ YÜYÁNXUÉ JIÀOYÁNSHÌ 2003: 39)

p	p ^h	m			
t	t ^h	n			l
ts	ts ^h			s	
k	k ^h	ŋ		x	
Ø					

TABLE 4.2 THE INITIALS OF CHÉNXÍ DIALECT (ADAPTED FROM XIÈ 2010: 179)

p	p ^h	b	m	f	v
t	t ^h	d	n		l
ts	ts ^h	dz		s	z
tʂ	tʂ ^h	dʐ		ʂ	ʐ
tɕ	tɕ ^h	dʐ		ɕ	ʐ
k	k ^h	g	ŋ	x	
Ø					

optionally by a “medial” on-glide and followed optionally by an ending which can be either a vocalic off-glide or consonantal coda. Abbreviating these elements as V, M, and E respectively, we can expand our syllable schema to $\sigma = IMVE^T$. The syllable structure of nearly all Chinese varieties can be neatly accommodated by this schema, although some dialects of the Mǐn group are exceptions in having “double endings” consisting of a post-nuclear off-glide and a consonant, as in Zhènqián 鎮前 (a Mǐn dialect of northwest Fújiàn) /kwajŋ²¹/ 汗 ‘sweat.’

Phonological inventories typically treat each contrastive final as a single phonological unit, without attempting a phonemicization of the main vowels. This is a practical approach which however raises interesting challenges of theoretical analysis.

The number of finals found in a Chinese variety ranges from about 35 to about 85. The most crucial factor accounting for this range is the number of endings. The canonical syllable structure of Standard Mandarin permits only the four endings /n ŋ j w/ and has a total of 40 finals. Near the other end of the spectrum, Xiàmén dialect (spoken in Fújiàn province and belonging to the Mǐn dialect group) permits nine endings /m n ŋ p t k ? j w/ and has a total of 80 finals. The number of on-glides ranges from zero (for example, in Standard Cantonese of Hong Kong) to three (as in Standard Mandarin). Stop endings are typically unreleased. The syllables in which they occur are notably shorter than other syllables; they give the impression of being “abrupt” or “clipped” in comparison with English stop-coda syllables.

Table 4.3 lists the finals of the Wēnzhōu 溫州 dialect spoken in Zhèjiāng province and belonging to the Wú dialect group. With a relatively impoverished set of endings, it has only 34 distinct finals.

Chinese dialectologists typically do not make a notational distinction between glides and vowel nuclei, writing /i u y/ where usual linguistic practice would have /j w ɥ/. In most cases this does not cause confusion because, as in Wēnzhōu, /i u y/ must be functioning as short glides when seen in combination with any other vocalic element. Thus /uai/ in Table 4.3 may be interpreted as /waj/ or /uaɪ/. The notation is problematic, however, in those

**TABLE 4.3 THE FINALS OF WĒNZHŌU DIALECT (ADAPTED
 FROM BĚIJīNG DÀXUÉ ZHōNGGUÓ YǔYÁN
 WÉNXUÉ XÌ YǔYÁNXUÉ JIÀOYÁNSHÌ 2003: 20)**

ɿ	i	u	y
a	ia	ua	
ɛ	ie	ue	
e			
ɔ		œ	yo
o			
ai	iai	uai	
ei			
au	iau		
əu	iəu		
øy			
aŋ	iaŋ	uaŋ	
eŋ			
oŋ			yoŋ
ŋ	ɿ		

varieties with /ui/ and /iu/ finals, because the notation is ambiguous between GV and VG interpretations. (The ambiguity is not just a matter of phonemic analysis, but fails to indicate measurable differences in phonetic length of the two vocalic elements.) In most examples given in this chapter, I have substituted glide symbols where the original source has vowel symbols.

Table 4.3 illustrates a number of other typical features, both of data presentation and content. The finals are laid out in columns based on their medial on-glides, with rows grouped together by ending. The first column contains finals with zero medial; the second the final /i/ and all finals beginning with /i/, i.e. with medial /j/; the third the final /u/ and all finals beginning with /u/ = /w/; the fourth the final /y/ and all finals beginning with /y/ = /ɥ/. The lack of phonemicization in the presentation of finals is obvious: /uœ/ and /yo/ have a main vowel which does not occur in isolation; one could entertain the possibility that phonemically the nucleus here is /e/ with a realization of [œ] after a rounded on-glide. The first symbol in the chart, /ɿ/, represents a super-high alveo-apical vowel articulated with some frication, producing an acoustic impression of buzzing. The closest International Phonetic Alphabet (IPA) representation of this sound would be [z]. This vowel sound is quite common in Chinese varieties, and is typically found only after dental sibilant initials. The symbol for it and the corresponding post-alveolar (“retroflex”) vowel /ɻ/ ([z̥]) are among a special set of non-IPA symbols that are typically encountered in Chinese dialect descriptions, usually without explanation (see Handel 2016 forthcoming on the meaning and history of these symbols). Finally, note the occurrence of finals composed of syllabic resonants, which are found in many varieties of Chinese and usually combine only with the zero initial.

2.1.4 Tone

Phonemic tone is found in all modern varieties of Chinese. The number of distinctive tones varies across dialects, and is generally smaller in the northern varieties and larger in

the southern varieties. The Dūnhuáng 敦煌 dialect (of the Mandarin group) has only three tones while Wēnzhōu dialect has eight. Guāngzhōu dialect is frequently described as having the most tones of all the major representative dialects, with nine, but this number is based on the traditional way of counting distinctive tones, which relies on historical categories rather than synchronic phonological patterns. The number can be reduced to six through phonemic analysis.

Generally speaking, the domain of Chinese tone is the syllable, so that each syllable within a word or phrase has a distinct tonal category. (A notable exception is the tone system of Shànghǎi, see Chapter 8 on Shanghaiese in this volume.) The category is manifested by a combination of relative pitch height and contour. While other phonetic factors can be present (such as length and phonation distinctions), these are usually considered predictable secondary features. Tone values can be described using two-word phrases such as “high level,” “low falling,” “mid rising,” “low convex” = “low rising-falling,” and so on. They are also frequently represented using “Chao tone numerals,” devised by Y.R. Chao to indicate pitch and contour in terms of a five-point scale, with 1 indicating the lowest natural pitch of a speaker’s voice and 5 the highest.⁴ A sequence of numbers shows on what pitch the tone starts and ends. Thus “high level” corresponds to 55 (starting high and ending high), while “low convex” corresponds to 121 (starting low, rising slightly, then dipping back down). Tones that are distinctively short can be indicated by using only one number or by underlining, e.g. 5 for a high short (level) tone and 121 for a low short convex tone. Chinese dialectologists append the numbers in superscript form when using IPA notation. This is not standard IPA usage, but is more practical than the IPA tonal notation. The four tones of Standard Mandarin are typically described and notated as in Table 4.4.

In isolation, the third tone is notably longer than the others, and is often accompanied by glottal creak. However, these suprasegmental features are not notated and are not considered to be as fundamental as the pitch contour.

The Chao tone numbers indicate pitches relative to the natural pitch range of a particular speaker’s voice. But even taking this relativity into account, these values should be understood only to apply to “citation tones” uttered in isolation. The actual pitch pattern of a tone varies considerably depending on many factors, including linguistic context. The overall intonational pattern of a phrase and the tones of neighboring syllables have effects on tonal articulation. For example, the Standard Mandarin 51 falling tone can be realized as quite long with a large pitch differential, or as shorter and nearly flat or even slightly rising, as seen in Figure 4.1. As with all purely allophonic differences in language, native speakers perceive these realizations as being the same tone.

A widespread phenomenon in Chinese dialects is the conditioned phonological rules known as “tone sandhi.” Tone sandhi refers to the predictable change in the value of a tone determined by the position of the syllable and/or by the tones of adjacent syllables within the relevant phonological unit. In varieties that exhibit tone sandhi, it occurs

TABLE 4.4 THE FOUR TONES OF STANDARD MANDARIN

Name	Description	Chao tone numerals	Example
First tone	high level	55	tāng /tʰan ⁵⁵ / 湯 ‘soup’
Second tone	high rising	35	táng /tʰan ³⁵ / 糖 ‘sugar’
Third tone	low falling-rising	214	tāng /tʰan ²¹⁴ / 躺 ‘to lie down’
Fourth tone	falling	51	tàng /tʰan ⁵¹ / 燥 ‘scalding hot’

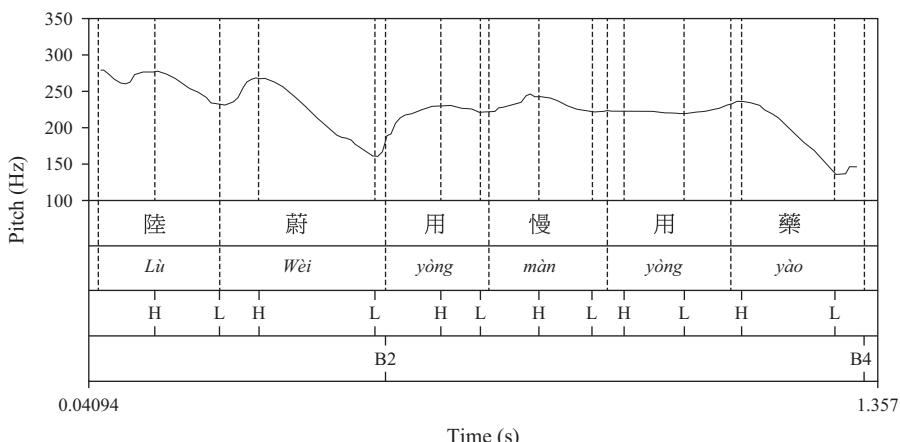


FIGURE 4.1 PITCH TRACK OF AN UTTERANCE BY A NATIVE SPEAKER OF STANDARD MANDARIN CONTAINING ONLY FOURTH-TONE SYLLABLES. THE SENTENCE IS *LÙ WÈI YÒNG MÀN YÒNG YÀO* ‘LÙ WÈI USES SLOW MEDICINE’. ADAPTED FROM YANG (2011: 47)

always in compound words and sometimes in phrases. (In some Chinese varieties, rules that accurately specify phrasal sandhi domains can be difficult to formulate.) The term is usually applied only to categorical tone changes, i.e. to changes resulting in surface forms identical to those of other tonemes, rather than to allophonic changes.

For example, Standard Mandarin is usually described as having a single tone sandhi rule, which is typically formulated as “a third tone changes to a second tone before a third tone” and schematized as /214/ + /214/ ⇒ /35/ + /214/.⁵ As a result of this rule, the words *zhǐdǎo* 指導 ‘to guide’ /tʂɿ²¹⁴/ + /taw²¹⁴/ and *zhídǎo* 執導 /tʂɿ³⁵/ + /taw²¹⁴/ ‘to direct (e.g. a film)’ are pronounced identically, as /tʂɿ³⁵ taw²¹⁴/.⁶ In contrast, the fact that a falling tone /51/ does not descend as far, and so is realized as [53], when preceding another falling tone is not usually considered an instance of tone sandhi, because the [53] value is not identical to that of any other tone category, and can simply be considered a conditioned allotone.

While some varieties of Chinese lack tone sandhi entirely (Cantonese being a good example), some have tone sandhi rules that are notoriously complex. Such tone systems are found in many varieties of the Wú and Mǐn dialect groups. In Sūzhōu of the Wú group, there are seven tones. In disyllabic compounds, all 49 possible tonal combinations induce sandhi changes to one or both syllables. Two of the 49 changes are illustrated in Table 4.5.

2.1.5 Phonotactics

Generally speaking, within a single Chinese variety, all tones can co-occur with all segmental syllables (although for historical reasons, there may be significant and systematic synchronic gaps). There are, however, usually restrictions on permissible combinations of initial and final. Typically these restrictions can be formulated in terms of place-of-articulation features of the initial and features of the beginning sound of the final. Consider Standard Mandarin. It has 22 initials (including zero), 39 finals, and four tones.

TABLE 4.5 SOME TONE SANDHI RULES IN SŪZHŌU

工人	$\text{kon}^{44} + \eta\text{in}^{24} \Rightarrow \text{kon}^{44}\eta\text{in}^{21}$	‘worker’ (Mandarin <i>gōngrén</i>)
奇怪	$\text{dzi}^{24} + \text{kwp}^{412} \Rightarrow \text{dzi}^{22}\text{kwp}^{44}$	‘strange’ (Mandarin <i>qíguài</i>)

If all logical combinations were possible, there would be $22 \times 39 = 858$ segmental syllables $\times 4 = 3,432$ distinct syllables. But in fact the number of possible syllables is reduced by phonotactic restrictions. For example, the retroflex initials *zh ch sh r /tʂ tʂʰ s l/* do not combine with finals beginning with vowels /i/ and /y/ or their corresponding glides /j/ and /ɥ/. Thus we find syllables like /mjen/ and /tejen/ but not */tʂjen/. Because of these restrictions the actual number of possible segmental syllables is slightly over 400. In combination with the four tones, this yields 1,600 possible syllables. (The number of actually occurring syllables is significantly lower, closer to 1,300, because of accidental and historically conditioned gaps. For a syllable chart of Standard Mandarin, see Ramsey 1987: 48.)

2.2 Major dialect groups and particular phonological features

In this section we will briefly lay out some of the distinctive phonological features seen in Chinese languages. We will focus on the largest dialect groups: Mandarin, Wú, Mǐn, and Yuè. In some cases we will note exceptions to the general features described above. Some of the distinctive phonological features will be explained from a historical perspective in the second half of the chapter. No attempt is made to be comprehensive.

2.2.1 Mandarin

Mandarin dialects exhibit a number of common features. They have relatively few tones (four or five is typical, although as few as three and as many as seven are found); two or three phonemically distinct series of sibilant affricates and fricatives (including retroflex if there are three); relatively few consonantal codas (typically just /n/ and /ŋ/); a two-way initial contrast based on aspiration; and a rich inventory of pre-nuclear on-glides. Also typical of Mandarin dialects is a lexical stress pattern in which the second syllable of disyllabic words is unstressed, resulting in shorter syllable length, vowel reduction and centralization, and sometimes allophonic voicing of syllable-initial obstruents. These unstressed syllables lose their lexically contrastive tone contours, and are said to have a “neutral” or “light” tone (notated by the absence of a tone mark). Thus, for example, where Cantonese has /hɔk¹sɔŋ⁵⁵/ 學生 ‘student’, the Standard Mandarin cognate is *xuéshēng* /ɛ̆χε²⁴ʂəŋ/⁵⁵, not */xuéshēng /ɛ̆χε²⁴ʂəŋ⁵⁵/. Outside of the Mandarin group, such “neutral tone” syllables are generally restricted to grammatical particles.

Although it is generally true that Chinese lacks sub-syllabic morphemes like English plural suffix *-s* and past-tense suffix *-ed*, a notable exception is the suffix *-r 兒* found in Standard Mandarin and related dialects. (Cognates and equivalents to this suffix are also found in many non-Mandarin varieties of Chinese, with realizations as varied as nasalization and tonal modification.) This suffix *-r*, which has diminutive and noun-formative functions, should not be confused with the fully syllabic morpheme *ér 儿* ‘child, son’ from which it is derived, and which is written with the same Chinese character. In *pinyin* transcription, derived words employing this suffix are spelled by appending the letter *r* to the spelling of the root, for example *huār 花兒* ‘flower’ (<*huā*>), *jīr 鷄兒* ‘chicken’ (<*jī*>),

gùnr 棍兒 ‘stick’ (<*gùn*>), *kòngr* 空兒 ‘free time’ (<*kòng*>), *jīnr* 今兒 ‘today’ (<*jīn* ‘now’), *shír* 事兒 ‘matter, affair, thing’ (<*shí*>). This spelling obscures, however, complex changes to the final of the root syllable, as shown in the broad phonetic representations in Table 4.6 (with tone omitted). A narrower transcription would explicitly show the derived vowels as rhotacized.

When *-r* is suffixed, root vowels are centralized and rhotacized. The ending *-n* is elided. In the case of ending *-ng*, the derived syllable has a rhotacized and nasalized vowel.

If these derived syllables are considered part of the basic phonological system of Standard Mandarin, then the description given earlier of Standard Mandarin syllable structure and phonotactics must be significantly revised. Such a revised analysis is, however, not typically done in Chinese dialectological studies, which consider rhotacized (*érhuà* 兒化) forms to be a secondary or marginal phenomenon, outside of the canonical phonological system.

2.2.2 *Wú*

Wú varieties are known for having many tones (seven or eight is typical); a three-way distinction of initial stops and affricates including a voiced series (in many northern varieties they are phonetically voiceless with following murmur); and complex patterns of tone sandhi. It is not uncommon to find only two consonantal coda phonemes: glottal constriction (usually realized [?]) and nasalization (realized in a variety of ways, including as [n] or [ŋ], but also as vowel nasalization).

2.2.3 *Mǐn*

The *Mǐn* group is so diverse that it is difficult to generalize about phonological features. Some varieties, like the inland *Fújiān* dialects termed Northern *Mǐn*, have relatively simple syllabic systems with minimal tone sandhi and only two consonantal endings /ʔ/ and /ŋ/. *Fúzhōu*, an Eastern *Mǐn* dialect, has extremely complex patterns of tonal and segmental sandhi associated with lexical compounding. Typically, the tone and vowel of the first morphosyllable change and the onset of the second syllable is weakened or assimilated. For example, the word for ‘storefront’ is [teiŋ⁵⁵naw⁵³], compounded of the morphemes [tajŋ²¹³] 店 ‘store’ and [thaw⁵³] 頭 ‘head.’ Many Southern *Mǐn* varieties such as *Cháozhōu* have a series of prenasalized voiced stops, contrasting with voiceless unaspirated and aspirated stops. As seen in the *Fúzhōu* word for ‘store,’ *Mǐn* varieties are also unusual in that many permit a post-vocalic off-glide followed by a consonant coda.

**TABLE 4.6 PRONUNCIATIONS OF RHOATACIZED SYLLABLES
 (ADAPTED FROM WÁNG AND HÈ 1985)**

Word	Root	Root + <i>-r</i>
<i>huār</i> 花兒 ‘flower’	[xua]	[xuaɪ̯]
<i>jīr</i> 鷄兒 ‘chicken’	[tɕi]	[tɕiaɪ̯]
<i>shír</i> 事兒 ‘matter’	[ʂɿ]	[ʃaɪ̯]
<i>gùnr</i> 棍兒 ‘stick’	[k'uəŋ]	[k'uəɪ̯]
<i>jīnr</i> 今兒 ‘today’	[tɕeɪŋ]	[tɕiaɪ̯]
<i>kòngr</i> 空兒 ‘free time’	[kʰuŋ]	[kʰuɪ̯]

2.2.4 *Yuè*

Yuè dialects generally have a high number of tones with little tone sandhi; few if any pre-vocalic on-glides, a rich inventory of stop and nasal codas, and a relatively impoverished set of syllable onsets. Cantonese, the *Yuè* variety of Guǎngzhōu 廣州 and Hong Kong, has a distinctive series of labiovelar initials /k^w k^{w̥}/, which however are merging with velars for younger speakers. (For more on these varieties, see Chapter 7 on Cantonese in this volume.) Many *Yuè* varieties have a pervasive vowel-length distinction. Because the members of long–short vowel pairs also differ notably in quality, many descriptions don't indicate the length distinction. For example, the low central vowels [ə] and [a:] can be treated as /a/ vs /a:/ but are more typically treated as /ə/ vs. /a/. The length distinction is an areal feature possibly reflecting Tai-Kadai substrate influence, and is most obviously manifested in tonal effects, conditioning a tone split in Cantonese (more on this later).

3 DIACHRONIC PHONOLOGY: FEATURES OF HISTORIC CHINESE VARIETIES AND DEVELOPMENTAL TRENDS

Because of the way that the fields of historical Chinese phonology and dialectology developed in the twentieth century, it is impossible to discuss research in these areas without some understanding of what is usually termed “traditional Chinese phonology.” Traditional Chinese phonology is a set of terms and concepts that emerged from the native medieval tradition of phonological analysis which took shape in the Táng and Sòng dynasties (seventh to thirteenth centuries), was refined by Qīng dynasty (seventeenth to nineteenth century) philologists, and was finally filtered through a modern linguistic sensibility in the twentieth century. Traditional Chinese phonological terms and concepts are used in nearly all scholarly investigations of the history of the Chinese language and synchronic descriptions of modern Chinese varieties, whether in Chinese-language or non-Chinese-language publications. For linguists and other scholars who have not received extensive training in this tradition, published work in these fields can present enormous challenges of understanding. In addition, some Western linguists who work on Chinese would argue that the traditional framework has not just rendered the field opaque to outsiders, but has acted as something of a straitjacket, cutting scholars off from data and methods that are crucial to a complete understanding of the synchronic and diachronic aspects of Chinese dialectology (Norman and Coblin 1995; Akitani and Hán 2012).

A full introduction to the topic of traditional Chinese phonology is well beyond the scope of the current chapter. Jacques (2016 forthcoming) provides an excellent overview for readers interested in a more complete understanding.

3.1 Periodization

The historical stages of Chinese are demarcated based on two criteria: the availability of useful textual sources and broad changes in typological features. In practice, this means that historical phonologists tend to periodize differently from historical syntacticians. Historical phonologists refer to the main stages of Chinese as OC (Shànggǔ Hànyǔ 上古漢語), MC (Zhōnggǔ Hànyǔ 中古漢語), Pre-Modern Chinese (Jindài Hànyǔ 近代漢語) and Modern Chinese (Xiàndài Hànyǔ 現代漢語). Within the general framework on

which all scholars agree, there are numerous different periodizations advocated by different scholars. They vary in the dates demarcating the stages, and on the degree to which the main stages can be sub-divided into smaller periods. These differences are not important for our purposes in this chapter. The periodization presented below is from Handel (2014: 579).

1 Old Chinese (OC):

Early	1250–1100 BCE
Middle	1100–200 BCE
Late	200 BCE–200 CE

2 Middle Chinese (MC):

Early	420–600
Middle	600–900
Late	900–1150

3 Pre-Modern Chinese (in practice synonymous with “Old Mandarin”)

1150–1650

4 Modern Chinese (all dialect groups)

1650–present

3.2 Middle Chinese phonology

MC is associated with two types of textual sources, known as rhyme books (*yùnshū* 韻書) and rhyme tables (*yùntú* 韵圖 or *děngyùntú* 等韻圖) (sometimes spelled “rime books” and “rime tables”). In combination, these two sources are the foundation for establishing the phonological categories of MC. Rhyme books are Chinese character dictionaries originally intended as reference tools for poetic composition. Characters are grouped together in sets of rhyming syllables known as rhyme groups (*yùnbù* 韵部) or simply rhymes (*yùn* 韵), which presumably have the same main vowel and ending, but may differ in initial and medial. By way of illustration, in a similarly-structured dictionary for Modern Standard Mandarin, *guān* /kwan⁵⁵/ 觀 ‘observe’ and *dān* /tan⁵⁵/ 單 ‘single’ would be grouped together within one rhyme group. Characters whose pronunciations are completely homophonous are grouped together within a homophone group. To continue with our analogy, *dān* 單 ‘single’ and *dān* 丹 ‘vermilion’ would be placed together in the same homophone group. The rhyme groups are themselves further organized by tone category (reflecting the fact that identity of tone was a requirement for rhymeability).

If one starts with the assumption that the contents of a rhyme book like *Guǎngyùn* 廣韻, compiled in 1008 as an expanded version of the *Qièyùn* 切韻 of 601, reflect a real phonological system, then its organizational structure immediately reveals significant details of the system: the number of tones, the number of distinct rhymes (= main vowel + ending combinations), the number of distinctly pronounced syllables, and which characters belong to each category. Further phonological information is provided by a sound glossing method known as *fāngqiè* 反切, which expresses the pronunciation of each homophone group in terms of the pronunciation of two common Chinese characters.

For example, in the first rhyme group of *Guāngyùn*, the homophone group containing the character *tóng* /tʰuŋ³⁵/ 同 is glossed *tú hóng* /tʰuŋ³⁵ xuŋ³⁵/ 徒紅. The target pronunciation is achieved by combining the initial of the first character with the final and tone of the second; thus, *tú hóng* yields *tóng*. This example works with Standard Mandarin pronunciation, but many *fāngqiè* glosses do not, because they are based on medieval pronunciation. In the same rhyme group we also find a homophone group containing the character *méng* /məŋ³⁵/ 蒙 glossed with *mò zhōng* /mo⁵¹ tʂun⁵⁵/ 莫中. Based on Mandarin, that gloss would appear to indicate the pronunciation *mōng*, a non-occurring syllable. Through analysis of all the *fāngqiè* “spellings” in the rhyme dictionary, it is possible to determine a more or less consistent set of distinct phonological categories of initial consonants and syllable finals. Phonetic values for these categories could be, in theory at least, determined by looking at pronunciations in a variety of modern Chinese dialects that are hypothesized to be directly descended from MC. In practice, however, additional information from the second main textual source, rhyme tables, is first used to refine the phonological analysis.

Rhyme tables, the earliest complete extant versions of which date to the Sòng (tenth to thirteenth centuries), are elaborated sets of grids on which Chinese characters are positioned according to features of their pronunciation. In very general terms, columns indicate features of the initial consonant and rows indicate features of the final and tone. Each position on a grid therefore corresponds to a unique combination of initial, final, and tone, thus providing the pronunciation of the character. Grid positions that do not correspond to occurring syllables are left blank.

The phonological system implied by the structure of early rhyme tables and that implied by the structure of early rhyme books are not identical, but can be aligned if certain assumptions about the phonological correspondences are made. The categories resulting from this alignment are the basic units of reconstructed “MC phonology” and are then given phonetic value through comparison with modern dialect pronunciations. A very simple example will serve to illustrate the method. Consider the three characters *dōng* 東 ‘east’, *tōng* 通 ‘through’ and *tóng* 同 ‘same,’ all found together in the same rhyme group of the *Guāngyùn*, which is called the *Dōng* 東 Rhyme. Their MC pronunciations can be presumed to rhyme, i.e. to have identical main vowels, endings and tones (thus implying a tone split has occurred in the development of Standard Mandarin). All three are found in different homophone groups with different *fāngqiè* glosses, indicating that their pronunciations differ by initial consonant, medial, or both.

The placement of these characters in the early rhyme table named *Yùnjìng* 韻鏡 gives us still more information. All three characters appear in the same row, indicating identity of their finals (i.e. they all have the same medial or all lack a medial). All three are found in a column designated *shé yīn* 舌音 ‘tongue sound,’ with the initial of *dōng* 東 further categorized as *qīng* 清 ‘clear,’ the initial of *tōng* 通 as *cìqīng* 次清 ‘secondarily clear’ and the initial of *tóng* 同 ‘same’ as *zhuó* 濁 ‘muddy.’

If we combine this structural and nomenclatural information with the pronunciation of these characters in some Chinese varieties belonging to different dialect groups, we can effectively reconstruct the MC pronunciation.⁷ Table 4.7 gives data for the dialects of Běijīng (Mandarin), Sūzhōu (Wú), Shuāngfēng (Xiāng), Nánchāng (Gàn), Méixiān (Hakka), and Guāngzhōu (Yuè). (The standard model of Chinese dialect evolution establishes these six dialect groups as direct descendants of MC, with Mǐn splitting off earlier from OC.)

The reconstruction of the initial of the first two characters is straightforward, given the identity correspondence across all six dialects. The initial of the last character is

TABLE 4.7 RECONSTRUCTING MIDDLE CHINESE PRONUNCIATIONS OF THREE CHARACTERS IN THE GUĀNGYÙN DÔNG 東 RHYME

Mandarin: Běijīng	Wú: Sūzhōu	Xiāng: Shuāngfēng	Gàn: Nánchāng	Hakka: Méixiān	Yuè: Guǎngzhōu	MC Final	MC Initial
東 tuŋ	toŋ	tan	tuŋ	tuŋ	toŋ	*-uŋ	*t-
通 tʰuŋ	tʰoŋ	tʰan	tʰuŋ	tʰuŋ	tʰoŋ	*-uŋ	*tʰ-
同 tʰuŋ	døŋ	dan	tʰuŋ	tʰuŋ	tʰoŋ	*-uŋ	*d-

reconstructed as voiced *d- because: 1. the textual evidence shows it to be phonologically distinct from the initials of the other two characters; 2. the modern reflexes show a distinct pattern of correspondence; 3. some of the modern reflexes are voiced.

Since the rhyme group containing these characters is classified as belonging to the *píng* 平 ‘level’ tone, we can now provide complete MC pronunciations for all three characters:

<i>dōng</i> 東 ‘east’	MC *tun ^{ping} 平
<i>tōng</i> 通 ‘through’	MC *tʰuŋ ^{ping} 平
<i>tóng</i> 同 ‘same’	MC *duŋ ^{ping} 平

We can further conclude that the designation *shé yīn* 舌音 ‘tongue sound’ refers, at least in part, to alveolar stops, and that the terms *qīng* ‘clear,’ *cīqīng* ‘secondarily clear’ and *zhuó* ‘muddy’ correspond to the features voiceless unaspirated, voiceless aspirated, and voiced, respectively. These conclusions are confirmed by additional evidence from the medieval textual sources and the modern Chinese languages.

(I have marked MC reconstructions above with an asterisk. However, in the remainder of the chapter I follow convention by leaving MC reconstructions unmarked.⁸)

In broad terms, the phonology of reconstructed MC most resembles that of southern Chinese varieties. But it differs from all known modern varieties in having a larger and more complex phonological inventory, which may simply reflect its artificiality. A few salient characteristics are:

- 1 four tones, named *píng* 平 ‘level,’ *shǎng* 上 ‘rising,’ *qù* 去 ‘departing,’ and *rù* 入 ‘entering,’ presumably with level, rising, falling, and abrupt contours;
- 2 a three-way contrast in initial stops and affricates (voiceless unaspirated, voiceless aspirated, voiced, e.g. ts- vs. tsʰ- vs. dz-) and a two-way contrast in initial fricatives (e.g. s- vs. z-);
- 3 six consonantal codas (nasals -m, -n, -ŋ, stops -p -t -k) and two semi-vowel off-glides (-j, -w), as well as open syllables; the stop codas co-occur only with *rù* tone and all other syllables co-occur with the other three tones;⁹
- 4 rounded (-w-) and palatal (-j-) medial features, which could co-occur (-wj-);
- 5 three series of sibilants at dental, post-alveolar (retroflex), and palatal places of articulation (e.g. ts- vs. tʂ- vs. tʃ-).

The four tone categories had high and low variants conditioned by the voicing feature of the syllable onset.¹⁰ At some point in the MC period these variants phonologized, resulting in an eight-tone framework in which each of the original four tones has split in two. The eight-tone framework can be schematized as shown in Table 4.8.

TABLE 4.8 THE EIGHT-TONE FRAMEWORK FOR COMPARATIVE CHINESE DIALECTOLOGY

Tone name	Basic tone	Initial type	Other tone designations
<i>yīnpíng</i>	<i>píng</i> ‘level’	voiceless	1 / A1
<i>yángpíng</i>	<i>píng</i> ‘level’	voiced	2 / A2
<i>yīnshàng</i>	<i>shǎng</i> ‘rising’	voiceless	3 / B1
<i>yángshàng</i>	<i>shǎng</i> ‘rising’	voiced	4 / B2
<i>yīnqù</i>	<i>qù</i> ‘departing’	voiceless	5 / C1
<i>yángqù</i>	<i>qù</i> ‘departing’	voiced	6 / C2
<i>yīnrù</i>	<i>rù</i> ‘entering’	voiceless	7 / D1
<i>yángrù</i>	<i>rù</i> ‘entering’	voiced	8 / D2

3.3 Middle Chinese phonology and Chinese dialectology

We are now in a position to understand some of the conventions of modern Chinese dialectology. A Chinese dialectologist might ask concerning a particular Chinese variety whether it has a “*qīng/zhuó* 清濁 ‘clear/muddy’ distinction.” This refers to the rhyme-table designations for voiceless and voiced obstruents, respectively, and whether the MC voiced obstruent initials remain distinct in the modern variety, or have merged with their voiceless counterparts. A Chinese dialectologist might also be interested in whether all three characters of Table 4.7 have the same tone in the modern variety; if so, then she would designate that tone the *píng* ‘level’ tone (even if its pitch contour were, say, 412, which is not level); if not, she would say that the MC tone has split into *qīngpíng* 清平 ‘clear-level’ and *zhuópíng* 濁平 ‘muddy-level’ categories, which could be alternatively designated *yīnpíng* 隅平 and *yángpíng* 陽平. Finally, the dialectologist might inquire about the Dōng-Rhyme reflexes in the variety under examination.

In this way, the synchronic features of a modern dialect’s phonology are discussed not only using technical terminology taken from the native medieval tradition, but as they relate diachronically to MC phonological categories and features. Despite this use of MC, reconstructed phonetic values are not typically presented, and the modern dialect features are not discussed in terms of explicit sound-rule derivations.

As a cursory exploration of a few issues of a modern Chinese dialectology journal like *Fāngyán* 方言 quickly reveals, this is in fact the way that varieties of Chinese are typically described and introduced.

It is worth reiterating here that this kind of approach to dialect description is based on a number of presumptions, all of which are problematic:

- 1 Modern varieties of Chinese are directly descended from, or at least can be treated as if directly descended from, MC;
- 2 MC is a homogeneous entity, corresponding to a single historically existing spoken language;
- 3 Modern dialect phonology is equivalent to the set of pronunciations of written Chinese characters;
- 4 Once the regular correlations between MC categories and phonological features of a Chinese variety have been established, the diachronic and synchronic phonologies of that variety have been explained; no further work is necessary for an adequate description.

The problematic aspects of these assumptions are well summarized in Norman and Coblin (1995).

When it comes to classifying dialects and assigning them to one of the major groups, dialectologists are primarily concerned with a small number of MC features that are considered highly diagnostic. These features and their developments constitute major trends in the history of the Chinese language family.

3.4 Phonological features and their role in classification and dialect description

As noted earlier, recognition of the major Chinese dialect groups predates modern linguistics, and is based on a combination of geography, self-identification, and broad isoglosses. Early twentieth-century attempts to make these designations more rigorous were largely based on correlations with a small number of MC phonological categories, the most frequently referenced of which are: 1. the voiced obstruent (*zhuó* ‘muddy’) initials; and 2. the stop-coda (*rù* tone) syllables. In an influential work on the classification of languages of the Sino-Tibetan family, Li (1937, 1973) defined the major dialect groups mostly in terms of these two categories, as follows:

- Mandarin: devoicing of MC voiced obstruents to aspirated in *píng* tone and unaspirated in the other three tones; complete loss of *rù*-tone stop codas (except in “Eastern” Mandarin, where they survive as -?).
- Wú: preservation of MC voiced obstruents and merger of *rù*-tone stop codas to -?.
- Gàn and Hakka: devoicing of MC voiced obstruents to aspirated regardless of tone; *rù*-tone endings -p -t -k are “more or less preserved.”
- Mǐn: devoicing of MC voiced obstruents to unaspirated in all tones and the preservation of *rù*-tone endings -p -t -k “sometimes in modified and simplified forms.”
- Yuè: preservation of all six MC consonant endings -m -n -ŋ -p -t -k.
- Xiāng: preservation of MC voiced obstruents; loss of *rù*-tone stop codas but preservation of a distinct tonal category for such syllables.

Li (1973: 5) said that, “there are also lexical items more or less peculiar to each of these groups” and acknowledged the existence of Chinese varieties that fell outside of this categorization (as “certain isolated groups”), but did not elaborate.

3.5 Broad trends in Chinese dialectal development

Viewed in terms of developments from MC, a number of broad trends can be seen across the language family. To what degree these changes are the result of internal factors and language-family “drift” and to what degree they have been triggered by language contact remain open matters of debate. Some scholars see a typological divide between northern and southern varieties of Chinese correlating with the “Altaic” typology of the languages historically spoken in the north of China and with the “Southeast Asian” typology of the Tai-Kadai, Austroasiatic, and Hmong-Mien languages historically spoken in the south of China (Yue-Hashimoto 1976; Hashimoto 1976, 1986; LaPolla 2001). According to this view, northern varieties of Chinese tend toward fewer tones, more CV-like syllable structure, and polysyllabicity, while southern varieties tend toward more tones, more CVC-like syllable structure, and monosyllabicity.

Among the most typical developments seen across Chinese varieties are the following (see below).

3.5.1 Weakening of stop codas (*loss of rù tone*)

Nearly all varieties of Chinese have seen some erosion of the MC stop-coda endings -p -t -k of the abrupt *rù* tone. Stop endings appear to first undergo merger, reducing to two or just one stop, then weakening to glottal stop before finally disappearing. In cases where these consonants have weakened to off-glides or zero, as in most varieties of Mandarin, the distinct *rù* tone category is usually lost, the syllables acquiring one of the other phonological tones. In some cases however, as noted for example by Li (1937, 1973) for some Xiāng and Southwestern Mandarin dialects, these syllables retain a distinct tone category even though no longer abrupt. The most faithful preservations of the endings are seen in dialects of the Yuè and Hakka groups, although even here there are reports of coda stop weakening among younger speakers.

The Southern Mǐn dialects are known for having an exceptionally large number of etymological doublets and triplets. The so-called colloquial-layer readings (*báidú* 白讀) are those which have developed as part of the history of the spoken language; the so-called literary-layer readings (*wéndú* 文讀) derive from learned character readings, originally modeled on medieval northern prestige pronunciations that eventually filtered into the spoken language. In general terms, literary-layer pronunciations tend to preserve -p -t -k endings while colloquial-layer pronunciations often reduce them to -? (or even zero in some cases).

Compare the words in Table 4.9 with MC -p, -t, and -k endings. Literary-layer pronunciations in Xiāmén are marked with ^L.

3.5.2 Weakening of nasal codas

A tendency toward merger and loss of nasal codas, paralleling that of stop codas, is also seen throughout Chinese, again with the greatest conservatism found in Yuè and Hakka (Table 4.10). In some varieties of Chinese nasal-coda loss and stop-coda loss are parallel, with vowel nasalization the last stage before total loss of nasals, mirroring the glottal stop as the last stage before total loss of stops. But in many varieties loss of nasals lags behind loss of stops, probably because of the greater phonetic salience of nasal endings as compared to unreleased stops. For example, while MC -p -t -k endings are completely gone in most Mandarin dialects, endings -n and -ŋ are still preserved, with earlier -m having merged to -n. In some Chinese varieties, such as Sūzhōu dialect, the three endings are reduced to a single nasal phoneme, which is realized as either [n] or [ŋ], or is lost altogether, depending on the preceding vowel. Chinese

TABLE 4.9 COMPARATIVE CHART OF ENTERING-TONE STOP-CODA DEVELOPMENTS (MODIFIED FROM BĚIJÍNG DÀXUÉ ZHÖNGGUÓ YÜYÁN WÉNXUÉ XÌ YÜYÁNXUÉ JIÀOYÁNSHÌ 2003). TONES OMITTED

	Mandarin: Běijīng	Wú: Sūzhōu	Gàn: Nánchāng	Hakka: Méixiān	Yuè: Guǎngzhōu	Mǐn: Xiāmén	MC Ending
十 ‘ten’	ʂɿ	ʐɿ?	sɿt	sɿp	sɿp	tsap, ^L sip	-p
葉 ‘leaf’	je	jɪ?	iɛt	iap	jip	i?, ^L iap	-p
日 ‘sun’	rɿ	ɳɿ?	ɳɿt	ɳɿt	jet	lit	-t
月 ‘moon’	ɥɿ	ɳɿ?	ɳɿt̪	ɳjat	jyt	ge?, ^L gwat	-t
客 ‘guest’	kʰɿ	kʰɿ?	kʰak	hak	hak	kʰe?, ^L kʰik	-k
玉 ‘jade’	y	ɳjø?	ɳjuk	ɳjuk	juk	gik, ^L gjok	-k

TABLE 4.10 COMPARATIVE CHART OF NASAL-CODA DEVELOPMENTS (MODIFIED FROM BĚIJÍNG DÀXUÉ ZHÖNGGUÓ YÜYÁN WENXUE XÌ YÜYANXUE JIAOYANSHI 2003), TONES OMITTED

	Mandarin: Běijīng	Wú: Sūzhōu	Xiāng: Shuāngfēng	Gān: Nánchāng	Hakka: Méixiān	Yuè: Guāngzhōu	Mn: Xiāmén	MC Ending
心 ‘heart’	ein	sin	eien	éin	sim	sem	sim	-m
新 ‘be new’	ein	sin	eien	éin	sin	sen	sin	-n
姓 ‘surname’	eiŋ	sin	eien	éjan	sjan	seŋ	síŋ	-ŋ
膽 ‘gallbladder’	tan	te	tâ	tan	tam	tam	tâ	-m
旦 ‘dawn’	tan	te	tâ	tan	tan	tan	tâ	-n
當 ‘should’	tanj	tmŋ	tmŋ	tɔŋ	tɔŋ	tɛŋ	tɔŋ	-ŋ

varieties with a complete loss of nasalization are rare, but the disappearance of nasal segments in favor of vowel nasalization is not uncommon. In Southern Mǐn, many nasal codas in the literary layer correspond to vowel nasalization in the colloquial layer, especially after open vowels. (See also the Cháozhōu examples in the tables in sections 3.5.4 and 3.5.5.)

3.5.3 Initial devoicing

As noted earlier, MC voiced obstruents have devoiced in most dialect groups, with the notable exception of Wú and Xiāng (Table 4.11). The fact that in many varieties the aspiration feature of the devoiced initial is tonally conditioned suggests that there must have been some sort of laryngeal feature, such as breathiness or murmur, associated with some tones of MC (or in their later reflexes), at least when these tones co-occurred with voiced initials. As many studies have noted, the so-called “voiced initials” of northern Wú dialects are actually pronounced as voiceless initials followed by breathy murmur, e.g. Sūzhōu /dɔŋ²⁴/ 同 ‘same’ is actually realized as [t^hɔŋ²⁴] (which could alternatively be transcribed [ton²⁴]).

The Mandarin devoicing pattern, in which voiced obstruent initials of MC *píng*-tone syllables become aspirated and those of non-*píng*-tone syllables become unaspirated, is also the dominant pattern in Mǐn and Yuè, but each of those groups shows notable exceptions to the pattern (above it was noted that Li (1937) pointed out that in Mǐn some *píng*-tone syllables end up with unaspirated initials).

While most Chinese varieties that have undergone devoicing have only an aspiration distinction in initial obstruents, a few have subsequently acquired voiced initials and so have a three-way distinction today. In Southern Mǐn dialects such as Xiàmén in Cháozhōu, nasal initials have denasalized in certain environments, resulting in voiced stops (generally pronounced with prenasalization). Table 4.12 contrasts these two Southern Mǐn dialects with Fúzhōu (Eastern Mǐn) and Jiàn’ōu (Northern Mǐn).

3.5.4 Palatalization

All Chinese varieties show a strong tendency toward palatalization of dental/alveolar affricates and fricatives before high front vowels. In varieties with only one affricate series, as with Cantonese and Xiàmén for example, this palatalization is purely allophonic, and is often not explicitly described in published phonological inventories. Palatalization of velar stops to affricates before high front vowels does not occur in most southern varieties of the Hakka, Mǐn, and Yuè groups, but is common in central and northern China. Where velar palatalization has occurred, mergers with palatalized dentals are common.

Compare the pronunciations of the words in Table 4.13 with MC velar initials and high front vowels, and note where mergers have taken place.

3.5.5 Dentilabialization

MC had the bilabial initial stops p-, p^h-, b-, but no labiodental initials. By the Late Middle Chinese period (Sòng dynasty), these initials, along with m-, had undergone a conditioned split. Although the precise formulation of the conditioning factor is elusive, in general terms we can say that labiodentals developed in syllables containing both rounded and palatal features. The results of this split are seen in all modern Chinese varieties aside

TABLE 4.11 COMPARATIVE CHART OF MC VOICED OBSTRUENTS (MODIFIED FROM BĚIJÍNG DÀXUÉ ZHÖNGGUÓ YÜYÁN WÉNXUÉ XÌ YÜYÁNXUÉ JIÀOYÁNSHÌ 2003). TONES OMITTED

	Mandarin: Bēijīng	Wú: Sūzhōu	Gàn: Nánchāng	Hakka: Méixiān	Yuè: Guǎngzhōu	Mǐn: Cháozhōu	MC Initial and Tone
童 ‘child’	tʰuj	døy	tʰuj	tʰuj	tʰuj	taj, tʰoj	d- <i>píng</i> 平
抱 ‘embrace’	paw	bæ	pʰaw	pʰaw	pʰow	p̥o, pʰaw	b- <i>shǎng</i> 上
舊 ‘be old’	tejow	dziy	teʰju	kʰju	kew	ku	g- <i>qu</i> 去

TABLE 4.12 COMPARATIVE CHART OF MC NASAL INITIALS (MODIFIED FROM BĚIJÍNG DÀXUÉ ZHÖNGGUÓ YÜYÁN WÉNXUÉ XÌ YÜYÁNXUÉ JIÀOYÁNSHÌ 2003). TONES OMITTED

	Mandarin: Bēijīng	Wú: Sūzhōu	Mǐn: Fúzhōu	Mǐn: Jiān’ōu	Mǐn: Xiāmén	Mǐn: Cháozhōu	MC Initial
馬 ‘horse’	ma	mo	ma	ma	be, l̥mã	be, l̥mã	m-
月 ‘moon’	ŋε	ŋɔ?	ŋwɔ?	ŋε	ge?, l̥gwat	gwe?	ŋ-

TABLE 4.13 COMPARATIVE CHART OF PALATALIZATION (MODIFIED FROM BĚIJÍNG DÀXUÉ ZHÖNGGUÓ YÜYÁN WÉNXUÉ XÌ YÜYÁNXUÉ JIÀOYÁNSHÌ 2003). TONES OMITTED

	Mandarin: Bēijīng	Wú: Wēnzhōu	Gàn: Nánchāng	Hakka: Méixiān	Yuè: Guǎngzhōu	Mǐn: Cháozhōu	MC Initial
劍 ‘sword’	tejen	tei	tejen	kjam	kim	kjəm	k-
箭 ‘arrow’	tejen	tei	tejen	tsjen	tsin	tsī	ts-

from the Mǐn group. The overall pattern is that MC p- and pʰ- became /f/, MC b- became /v/ (and then /f/ in varieties with initial devoicing), and MC m- became /w/. Table 4.14 shows contrasting developments.

The lack of labiodentals in Mǐn dialects is one of several features that do not correlate well with Late Middle Chinese, and have led scholars to conclude that Mǐn has a distinct history, splitting off from “mainstream” Chinese dialects soon after the OC period.

3.5.6 Affricate-series mergers

While MC is reconstructed with three affricate series, only a small number of modern Chinese varieties have three series. This rare feature is found in Modern Standard Mandarin, which boasts dental, retroflex, and palatal series as does Chénxī (Table 4.2). However, these three series are not continuations of the distinct MC series. The MC palatals and retroflexes have (mostly) merged to Mandarin retroflexes; the Mandarin palatals result from recent palatalization of velar and dental sibilant initials.

The two most common patterns for affricates in modern Chinese varieties are: 1. A single series, typically dental with palatal allophones, corresponding to all three MC series.

**TABLE 4.14 COMPARATIVE CHART OF DENTILABIALIZATION (MODIFIED FROM
BĚIJÍNG DÀXUÉ ZHÖNGGUÓ YÜYÁN WÉNXUÉ XÌ YÜYÁNXUÉ
JIÀOYÁNSHÌ 2003). TONES OMITTED**

	Mandarin: Běijīng	Wú: Sūzhōu	Gàn: Nánchāng	Hakka: Méixiān	Yuè: Guǎngzhōu	Mǐn: Cháozhōu	MC
板 ‘plank’	pan	pε	pan	pan	pan	pō̄i	pæn
反 ‘reverse’	fan	fε	fan	fan	fan	pō̄i	pjɔn
辦 ‘do, manage’	pan	bε	pʰan	pʰan	pan	pʰō̄i	bɛn
飯 ‘rice, meal’	fan	vε	fan	fan	fan	pʊŋ	bjɔn
慢 ‘slow’	man	mε	man	man	man	maj	mæn
挽 ‘coil up’	wan	wε	wan	van	wan	maj	mjɔn

This is seen, for example, in Cantonese; 2. Two series, one (dental/alveolar) reflecting a merger of all three MC series, and the other (palatal) resulting from recent palatalization of velar and/or sibilant initials. This is seen, for example, in Wēnzhōu. Compare ‘sword’ and ‘arrow’ in Table 4.13 to the data in Table 4.15.

3.5.7 Other segmental mergers

Two of the most common and salient mergers seen in Chinese varieties involve MC n- and l- (distinctly preserved as such in Modern Standard Mandarin) and f- and xw- (also distinct in Modern Standard Mandarin). These mergers manifest in a variety of ways, including free variation, conditioned allophony, and complete loss of one sound in favor of the other. They are common in many varieties of Chinese spoken in the southern part of the country. Free variation of n- and l- is a salient feature of Southwest Mandarin, such as that spoken in Chéngdū 成都, the capital of Sichuān, where the morphemes meaning ‘male’ (Standard Mandarin *nán* 男) and ‘blue’ (Standard Mandarin *lán* 藍) are homophonous. The formal educated register of Hong Kong Cantonese retains the distinction, but in ordinary speech n- is largely giving way to l-, so that one typically hears /lej³⁵/ how³⁵/ instead of /nej¹³/ 你好 ‘hello’ (cf. Standard Mandarin *nǐ hǎo*).

The f-/xw- merger is a similar phenomenon, similarly widespread, but more complex in its details. Stated simply, there is a tendency for historical f- and xw- to show variation or merger (where “w” represents on-glide /w/ or vowel /u/). The result can be free variation; merger to f-; merger to xw-; or partial mergers conditioned by the syllable final. A recent investigation (Xiè 2013) of f- and x- initials in about 30 dialects in and around Chóngqing 重庆 (Southwestern Mandarin varieties) found that 湖 ‘lake’ (Standard Mandarin *hú* [xu³⁵] < Late Middle Chinese velar fricative) and 福 ‘good fortune’ (Standard Mandarin *fú* [fu³⁵] < Late Middle Chinese labiodental) are pronounced identically as [fu] throughout the region. But within the city proper mergers in syllables with finals other than [u] are not common. In the other dialects examined, however, the merger is widespread, so that, for example, 換 ‘to exchange’ (Standard Mandarin *huàn* [xwan⁵¹] < Late Middle Chinese velar fricative) and 飯 ‘rice’ (Standard Mandarin *fàn* [fan⁵¹] < Late Middle Chinese labiodental) are pronounced identically as [fan].

3.5.8 Tone mergers and splits

In Table 4.8, the eight-tone framework of MC was presented. The tone systems of most modern Chinese varieties can be correlated with this framework, although few Chinese

TABLE 4.15 COMPARATIVE CHART OF AFFRICATE SERIES MERGERS
(MODIFIED FROM BĚIJÍNG DÀXUÉ ZHÓNGGUÓ YÚYÁN WÉNXUÉ
XÌ YÚYÁNXUÉ JIÀOYÁNSHÌ 2003). TONES OMITTED

	Mandarin: Běijīng	Wú: Wēnzhōu	Gàn: Nánchāng	Hakka: Méixiān	Yuè: Guāngzhōu	Mǐn: Cháozhōu	MC Initial
祖 ‘ancestor’	tsu	tsøŋ	tsu	tsɻ	tsow	tsou	ts-
珠 ‘pearl’	tʂu	tsɻ	tɛy	tsu	tsy	tsju, ^l tsu	tʃ-
斬 ‘to chop’	tʂan	tsa	tsan	tsam	tsam	tsam	tʂ-

varieties reflect this framework unchanged; mergers and splits are the norm. Most commonly, splits are conditioned by whether the MC initial consonant is voiceless, a voiced obstruent, or a voiced resonant, but other conditioning factors are sometimes seen. The nomenclature used for describing modern tone systems in Chinese publications can be very confusing, since the names of the MC tones are repurposed to name modern tone categories that only partially overlap with the MC categories.

Wēnzhōu is a rare example of a Chinese dialect that retains the canonical eight-tone system (Table 4.16). As a Wú dialect that preserves the MC voicing distinction, the lower-register tones are consistently associated with voiced initials, and have lower pitches than their upper-register counterparts. (It is therefore possible to analyze Wēnzhōu as having a four-tone system, with allophonic pitch variants conditioned by initial voicing, but this is never done in Chinese dialectological publications.)

It is often claimed that Cantonese preserves the MC tone system, but in fact there have been a number of changes. The most obvious is a split in the upper-register entering tone (*yīnrù*) into a high tone (5) and a mid tone (3) according to the syllable's vowel length. The system is therefore traditionally described as having nine tones, as in Table 4.17.

What is not immediately obvious from the chart is that most of the words with the MC *yángshǎng* tone are not in the Cantonese *yángshǎng* tone; rather, they have merged with the *yángqù* tone. An example is [tuŋ²²] 動 ‘to move’ (cf. Wēnzhōu [dɔŋ³⁴], which reflects the original *yángshǎng* tone category). The only syllables to escape this merger are those with resonant initials and a small number of colloquial words (among them Table 4.16’s ‘thick’).

This merger of MC *yángshǎng* with *yángqù* for voiced obstruent initials is in fact one of the most prevalent across Chinese, found in most Xiāng and Gàn varieties, many coastal Mǐn varieties, and in some Wú varieties including Sūzhōu. It is a defining characteristic found throughout the Mandarin dialects, as seen for example in the tone system of Xī’ān 西安 (Table 4.18).

Xī’ān also illustrates another of the most common tone mergers that have occurred in Chinese: the loss of distinct tonal contours when former entering-tone syllables lose their stop codas. In Xī’ān, such syllables have merged with the *yīnpíng* and *yángpíng* tones, depending on their initial type.

The nomenclatural confusion mentioned at the beginning of this section is well illustrated by Xī’ān. Comparing the four tone names with the correlations to MC tones (Table 4.19), we can see that modern tone names like “*shǎng*” and numerical designations like “5” only reflect part of the correspondence pattern. (P = MC voiceless initial; B = MC voiced obstruent initial; M = MC resonant initial.)

It is also worth pointing out that the original register split, with upper register conditioned by voiceless initials and lower register by voiced initials, is not always reflected in the pitch patterns of the modern tone values. As seen in the Wēnzhōu and Cantonese tone

TABLE 4.16 THE TONE SYSTEM OF WĒNZHŌU (ADAPTED FROM BĚIJĪNG DÀXUÉ ZHÖNGGUÓ YÜYÁN WÉNXUÉ XÌ YÜYÁNXUÉ JIÀOYÁNSHÌ 2003: 20)

Tone name	Tone value	Initial type	Example morpheme
1 <i>yīnpíng</i>	44	voiceless	[tʰi ⁵⁵] 天 ‘sky’
2 <i>yángpíng</i>	31	voiced	[ben ³¹] 平 ‘level’
3 <i>yīnshǎng</i>	45	voiceless	[ejaw ⁴⁵] 手 ‘hand’
4 <i>yángshǎng</i>	34	voiced	[gaw ³⁴] 厚 ‘to be thick’
5 <i>yīnqù</i>	42	voiceless	[pʰo ⁴²] 怕 ‘to fear’
6 <i>yángqù</i>	22	voiced	[jup ²²] 用 ‘to use’
7 <i>yīnrù</i>	323	voiceless	[tʰsaj ³²³] 七 ‘seven’
8 <i>yángrù</i>	212	voiced	[ny ²¹²] 月 ‘moon’

TABLE 4.17 THE TONE SYSTEM OF CANTONESE (ADAPTED FROM BĚIJĪNG DÀXUÉ ZHÖNGGUÓ YÜYÁN WÉNXUÉ XÌ YÜYÁNXUÉ JIÀOYÁNSHÌ 2003: 29)

Tone name	Tone value	Example morpheme
1 <i>yīnpíng</i>	55 or 53	[tʰin ⁵⁵] 天 ‘sky’
2 <i>yángpíng</i>	21	[pʰen ²¹] 平 ‘level’
3 <i>yīnshǎng</i>	35	[ʃew ³⁵] 手 ‘hand’
4 <i>yángshǎng</i>	13	[hew ¹³] 厚 ‘to be thick’
5 <i>yīnqù</i>	33	[pʰa ³³] 怕 ‘to fear’
6 <i>yángqù</i>	22	[juŋ ²²] 用 ‘to use’
7a upper <i>yīnrù</i>	5	[tʰvet ⁵] 七 ‘seven’ (short vowel /e/)
7b lower <i>yīnrù</i>	3	[pat ³] 八 ‘eight’ (long vowel /a/)
8 <i>yángrù</i>	2	[jyt ²] 月 ‘moon’

TABLE 4.18 THE TONE SYSTEM OF XĪ'ĀN (ADAPTED FROM BĚIJĪNG DÀXUÉ ZHÖNGGUÓ YÜYÁN WÉNXUÉ XÌ YÜYÁNXUÉ JIÀOYÁNSHÌ 2003: 10)

Tone name	Tone value	Example morpheme
1 <i>yīnpíng</i>	21	[tʰjæ ²¹] 天 ‘sky’
2 <i>yángpíng</i>	24	[pʰiŋ ²⁴] 平 ‘level’
3 <i>shǎng</i>	53	[fei ⁵³] 水 ‘water’
5 <i>qu</i>	55	[twoŋ ⁵⁵] 動 ‘to move’, [qoŋ ⁵⁵] 用 ‘to use’

TABLE 4.19 THE TONE SYSTEM OF XĪ'ĀN AND MIDDLE CHINESE SOURCES

Tone name	Tone value	Middle Chinese source
1 <i>yīnpíng</i>	21	<i>píng</i> -P; <i>rù</i> -PM
2 <i>yángpíng</i>	24	<i>píng</i> -BM; <i>rù</i> -B
3 <i>shǎng</i>	53	<i>shāng</i> -PM
5 <i>qu</i>	55	<i>shāng</i> -B, <i>qu</i> -PBM

systems, Wú and Yuè varieties generally retain the historic pitch distinctions: each *yīn* tone is higher than its corresponding *yáng* tone. But this is notably not the case in Xī'ān, where *yángpíng* has higher pitch (24) than *yīnpíng* (21), and such “register flip-flop” is also typical of many Mǐn varieties.

As mentioned previously, other conditioning factors for tone splits occur, but less commonly. For example, Shàowǔ 邵武 (a Northern Mǐn variety) and Suīchúān 遂川 (a Gàn variety) have tone splits correlated with aspiration of the initial consonant (Norman 1982; Chāng 2012). Many dialects also have morphologically conditioned tone changes, but these are lexical shifts, and strictly speaking do not belong to the study of phonologically conditioned splits and mergers.

Some Chinese varieties have tone systems that do not correlate neatly with the eight-tone framework. Of these the tone systems of the Northern Mǐn group have received the most attention; Jerry Norman (1973) argued that they cannot be derived from the MC phonological system. For more information see Handel (2003).

3.6 Typological shifts from Old Chinese to Middle Chinese

Although not strictly related to the topic of Chinese dialect phonology, a brief description of OC can provide an interesting counterpoint to the overall typological similarity seen within the entire Sinitic family of languages today. OC is reconstructed by reconciling textual evidence (rhyming poetry and phonetic components of Chinese characters) with the phonologies of later stages of Chinese. Although much about OC remains uncertain or controversial, all of the evidence suggests that OC was typologically quite a different species from MC and later varieties of Chinese. Indeed, it appears that a typological shift took place between the OC and MC periods whose magnitude dwarfs even the most dramatic changes seen over the last 1,500 years. This shift involved all sub-systems of the language, syntactic and morphological as well as phonological. The shift appears to have taken place during the time of the great Qín-Hàn imperial expansion at the end of the first millennium BCE. Its effects are seen in all later varieties of Chinese, supporting the hypothesis that the most recent common ancestor of the Sinitic family was a Hán dynasty *koine*. We will here just briefly note some of the ways that OC differed from all modern varieties of Chinese.

3.6.1 Tones

There is now wide (though not universal) consensus among specialists that OC was a language lacking in tone distinctions. The four tones of MC are believed to have developed through a process of tonogenesis, in which non-phonemic pitch differences conditioned by distinct syllable codas were phonologized. The *shǎng* tone from syllable-final *-ʔ (example (2) following) and *qù* tone developed from syllable-final *-s (> *-h) (examples (3)–(6)); those consonants disappeared in the process, along with any preceding stops (examples (5)–(6)). Syllables with syllable-final *-p *-t *-k developed into the distinctively short *rù* tone with retention of the stop codas (example (7)). All other syllables (i.e. those with vocalic or nasal endings) developed the unmarked *píng* tone (example (1)).

3.6.2 Syllable structure, phonotactics, and clusters

While reconstructed OC systems vary considerably in the details, nearly all experts agree that OC syllables tolerated consonant clusters at both the front and back ends of syllables. Many scholars also believe that some clusters were “loose,” with an intervening reduced vowel, so that “syllables” were in some cases actually complexes of minor and major syllables. Thus the familiar IMVE scheme for syllables is not applicable to the OC period.

3.6.3 Consonantal distinctions

Again, details differ, but most scholars agree that OC had significantly more distinctions in initial consonant type than MC. In addition to the familiar three-way voiceless unaspirated / voiceless aspirated / voiced contrast, there were also voiced and voiceless resonants and possibly prenasalized obstruents (both voiced and voiceless).

3.6.4 Type A/B distinction

OC had two distinct syllable types. There is a great deal of disagreement about the phonetic nature of the syllabic feature involved, but the existence of the distinction is not in doubt. For convenience the two syllables types are termed A and B. In the development from OC to MC, Type B syllables show a tendency toward palatalization; Type A syllables do not.

The reconstructed OC syllables (1) through (7), from the system of Baxter and Sagart (2014), illustrate the features mentioned above. In their reconstruction, Type A syllables are marked as pharyngealized (‘). Morphological affixes are marked off by hyphens. (In many cases tonogenetic *-s and *-? can be identified as derivational suffixes.) MC pronunciations (again in the system of Baxter and Sagart, but with modified tonal notation) and Modern Standard Mandarin pronunciations are also provided.

1 雞 ‘chicken’	OC *k ^č e > MC <i>kei</i> ^{ping} 平 > MSM <i>jī</i> /tɕi ⁵⁵ /
2 始 ‘to begin’	OC *ʃə? > MC <i>syi</i> ^{shāng} 上 > MSM <i>shī</i> /ʂɿ ²¹⁴ /
3 放 ‘to place, put’	OC *paj-s > MC <i>pjan</i> ^{qu} 去 > MSM <i>fàng</i> /fan ⁵¹ /
4 謗 ‘to slander’	OC *p ^č an-s > MC <i>paj</i> ^{qu} 去 > MSM <i>bàng</i> /pan ⁵¹ /
5 尿 ‘urine’	OC *kə.n ^č ewk-s > MC <i>new</i> ^{qu} 去 > MSM <i>niào</i> /njaw ⁵¹ /
6 背 ‘to turn one’s back to’	OC *m-p ^č ək-s > MC <i>bwoj</i> ^{qu} 去 > MSM <i>bèi</i> /pej ⁵¹ /
7 客 ‘guest’	OC *k ^č rak > MC <i>khaek</i> ^{rū} 羅 > MSM <i>kè</i> /kʰy ⁵¹ /

4 CONCLUSION

Chinese is a large, highly ramified family comprising many mutually unintelligible varieties spoken over a vast area. Despite its long history and geographic dispersal, most varieties share a number of common phonological features, especially in terms of syllable structure and tone. They also exhibit a high degree of regularity of correspondence, so that it is usually possible to explain phonological divergence in terms of regular historical development.

FURTHER READING

- Norman, Jerry (1988) *Chinese*, Cambridge: Cambridge University Press.
 Norman, Jerry (2003) ‘The Chinese dialects: phonology’, in Graham Thurgood and Randy J. LaPolla (eds) *The Sino-Tibetan Languages*, London and New York: Routledge, 72–83.
 Ramsey, S. Robert (1987) *The Languages of China*, Princeton, NJ: Princeton University Press.

NOTES

1 Some preliminary attempts have been made to reconstruct this ancestral form of Chinese. Schuessler (2009: xi) termed it “Later Han Chinese” and described it as “a hypothetical conservative strain of the language of about the 2nd century AD” and

“the earliest form of Chinese which can be set up without relying heavily on interpretations of [textual evidence],” and which “had all those features which survived as archaisms in later dialects.” Norman (2014: 9) termed it “Early Chinese” and described it as “an early system [of reconstructed Chinese] from which the *Qièyùn* (QY) categories, C[ommon] D[ialectal] C[hinese] and (to a certain extent) Min can be derived.” It is possible that some remnants of pre-Hàn Chinese dialectal diversity and reconstructed OC phonology survive here and there in the modern Sinitic family, but no definitive evidence has yet been discovered.

- 2 Chinese terms are given in the official Hanyu Pinyin romanization scheme, followed at the first use by Chinese characters in the traditional script. Hanyu Pinyin is also used for Standard Mandarin as linguistic data, alongside IPA. Chinese characters are given for data in other varieties when those data are clearly identifiable as cognates of Standard Mandarin or Classical Chinese morphemes that have conventional written forms. Note that many single syllables in Chinese varieties are bound forms; this fact is not noted when providing individual glosses.
- 3 Phonetic studies have shown that for most speakers the so-called “retroflexes” are not true retroflex articulations in which the underside of the tongue makes contact with the palate; they are rather apico-postalveolars.
- 4 One other tonal notation scheme is commonly seen in Chinese-language publications. It uses iconic “tone letters” to visually represent the pitch and contour of a tone, with a vertical bar serving as a standard for the five-point pitch scale. Thus the 55 tone can be represented with |, and the 214 tone with √|. The invention of this notational system is also attributed to Y.R. Chao. The tone letters were officially adopted by the IPA in 1989. While effective in most cases, the five-point scale does not capture all the complexity of Chinese tonal phenomena. For recent work on tonal theory and description, see Zhu (2012).
- 5 Some descriptions also mention a Mandarin tone sandhi rule that affects three-syllable combinations. Chao (1968: 27–8) identifies a rule of “minor importance” in which the middle syllable of a three-syllable group changes from second tone to first tone if the first syllable has the first or second tone, but only at conversational speed. There are 11 possible three-syllable tonal combinations that participate in this change.
- 6 There is actually a long-standing debate about whether the sandhi form of Tone 3 is phonetically identical to Tone 2, e.g. whether *zhidǎo* and *zhídǎo* are truly identical or are subtly distinct. See, for example, Hockett 1947, Martin 1957 and Wang and Li 1967.
- 7 Strictly speaking, the result of this exercise is a reconstruction of the reading pronunciation of a Chinese character, which is not necessarily the same as the reconstruction of words or morphemes of spoken language. The methodology is based on the assumption, almost always unstated, that reading pronunciations undergo regular sound change in the same way that spoken words do. Chinese dialectologists and historical phonologists do not usually carefully distinguish character readings from spoken forms, and therefore often treat written characters as functionally equivalent to cognate sets of spoken morphemes.
- 8 This convention goes back to the pioneering historical linguist Bernhard Karlgren, who believed that the phonological information provided by medieval textual sources was precise enough to be equivalent to written attestation. He therefore did not attach an asterisk to his reconstructions. Few scholars would now agree with this assertion, even though they follow the established convention. The asterisks in the exercise above are intended as a reminder that MC reconstructed values are indeed hypothetical. Nevertheless, the rest of the chapter follows the convention of omitting asterisks to preserve compatibility with other publications in the field.
- 9 In terms of phonemic analysis, it is possible to treat the *rù* tone as a conditioned variant of the *píng* tone. However, the native Chinese literary and linguistic traditions have always considered them to be distinct tone categories. In modern Chinese

varieties that have stop-coda syllables, the shortened tonal contours are perceptually quite distinctive, and this was presumably also the case in the medieval period.

- 10 It is likely that breathy phonation, rather than voicing *per se*, was the proximal factor in tonal conditioning (Thurgood 2002; Zhū 2010).

BIBLIOGRAPHY

- Akitani, Hiroyuki 秋谷裕幸 and Hán, Zhéfū 韩哲夫 (Handel, Zev) (2012) ‘Lishī bǐjiàofǎ hé céngci fēnxífǎ 历史比较法和层次分析法’ (The role of the comparative method and stratum analysis in the reconstruction of Chinese dialect history), *Yǔyánxué Lùncóng* 45: 277–335.
- Ansaldo, Umberto (2016 forthcoming) ‘Pidgins and creoles’, in Rint Sybesma, Wolfgang Behr, Yueguo Gu, Zev Handel, C.-T. James Huang and James Myers (eds) *Encyclopedia of Chinese Language and Linguistics*, Leiden: Brill.
- Baxter, William H. and Sagart, Laurent (2014) *Old Chinese: A New Reconstruction*, Oxford: Oxford University Press.
- Běijīng Dàxué Zhōngguó Yǔyán Wénxué Xì Yǔyánxué Jiàoyánshì 北京大学中国语言文学系语言学教研室 (ed.) (2003) *Hànyǔ Fāngyīn Zihui* 汉语放音字汇 (*Lexicon of Chinese Dialect Character Pronunciations*), revised 2nd edition, Běijīng: Yǔwén.
- Chāng, Méixiāng 昌梅香 (2012) ‘Gānyǔ Suichuān fāngyán de sòngqì fèndiào 赣语遂川方言的送气分调’ (Tone splits conditioned by aspiration in the Sichuan Gan dialect), *Fāngyán* 2012.4: 310–313.
- Chao, Yuen Ren (1968) *A Grammar of Spoken Chinese*, Berkeley and Los Angeles: University of California Press.
- Dede, Keith (2016 forthcoming) ‘Mixed languages’, in Rint Sybesma, Wolfgang Behr, Yueguo Gu, Zev Handel, C.-T. James Huang and James Myers (eds) *Encyclopedia of Chinese Language and Linguistics*, Leiden: Brill.
- Handel, Zev (2003) ‘Northern Min tone values and the reconstruction of “softened initials”’, *Language and Linguistics* 4.1: 47–84.
- Handel, Zev (2014) ‘Historical phonology of Chinese’, in James C.-T. Huang, Y.-H. Audrey Li and Andrew Simpson (eds) *Handbook of Chinese Linguistics*, Chichester, UK: Wiley-Blackwell, 576–598.
- Handel, Zev (2016 forthcoming) ‘Non-IPA symbols in IPA transcriptions’, in Rint Sybesma, Wolfgang Behr, Yueguo Gu, Zev Handel, C.-T. James Huang and James Myers (eds) *Encyclopedia of Chinese Language and Linguistics*, Leiden: Brill.
- Hashimoto, Mantaro J. (1976) ‘Language diffusion on the Asian continent: problems of typological diversity in Sino-Tibetan’, *Computational Analyses of Asian and African Languages* 3: 49–66.
- Hashimoto, Mantaro J. (1986) ‘The Altaicization of northern Chinese’, in John McCoy and Timothy Light (eds) *Contributions to Sino-Tibetan Studies*, Leiden: Brill, 76–97.
- Hockett, Charles F. (1947) ‘Peiping phonology’, *JAOS* 67.4: 253–267.
- Jacques, Guillaume (2016 forthcoming) ‘Traditional Chinese phonology’, in Rint Sybesma, Wolfgang Behr, Yueguo Gu, Zev Handel, C.-T. James Huang and James Myers (eds) *Encyclopedia of Chinese Language and Linguistics*, Leiden: Brill.
- LaPolla, Randy J. (2001) ‘The role of migration and language contact in the development of the Sino-Tibetan language family’, in Dixon, R. M. W. and Aikhenvald, Alexandra Y. (eds) (2001) *Areal Diffusion and Genetic Inheritance: Case Studies in Language Change*, Oxford: Oxford University Press, 225–254.

- LaPolla, Randy J. (2013) ‘Comments on methodology and evidence in Sino-Tibetan comparative linguistics’, *Language and Linguistics* 13.1: 117–132.
- Li, Fang-Kuei (1937) ‘Languages and dialects of China’, in Council of International Affairs (China) *The Chinese Year Book*, Shanghai: Commercial Press, 59–65.
- Li, Fang-Kuei (1966) ‘The zero initial and the zero syllabic’, *Language* 42.2: 300–302.
- Li, Fang-Kuei (1973) ‘Languages and dialects of China’, *JCL* 1.1: 1–13.
- Mair, Victor H. (1991) ‘What is a Chinese “dialect/topolect”? Reflections on some key Sino-English linguistic terms’, *Sino-Platonic Papers* 29: 1–31.
- Martin, Samuel E. (1957) ‘Problems of hierarchy and indeterminacy in Mandarin phonology’, *BIHP* 29: 209–229.
- Norman, Jerry (1973) ‘Tonal development in Min’, *JCL* 1.2: 222–238.
- Norman, Jerry (1982) ‘The classification of the Shaowu dialect’, *BIHP* 53: 543–583.
- Norman, Jerry (1988) *Chinese*, Cambridge: Cambridge University Press.
- Norman, Jerry (2003) ‘The Chinese dialects: phonology’, in Graham Thurgood and Randy J. LaPolla (eds) *The Sino-Tibetan Languages*, London and New York: Routledge, 72–83.
- Norman, Jerry (2014) ‘A model for Chinese dialect evolution’, in Richard VanNess Simmons and Newell Ann Van Auken (eds) *Studies in Chinese and Sino-Tibetan Linguistics: Dialect, Phonology, Transcription and Text*, Taipei: Institute of Linguistics, Academia Sinica, 1–26.
- Norman, Jerry and Coblin, W. South (1995) ‘A new approach to Chinese historical linguistics’, *JAOS* 115.4: 576–584.
- Ramsey, S. Robert (1987) *The Languages of China*, Princeton, NJ: Princeton University Press.
- Schuessler, Axel (2009) *Minimal Old Chinese and Later Han Chinese: A Companion to Grammata Serica Recensa*, Honolulu: University of Hawai'i Press.
- Thurgood, Graham (2002) ‘Vietnamese and tonogenesis: revising the model and the analysis’, *Diachronica* 19.2: 333–363.
- Wáng, Líjiā 王理嘉 and Hè Níngjī 贺宁基 (1985) ‘Běijīnghuà érhuàyùn de tīngbiàn shíyán hé shēngxué fēnxī 北京话儿化韵的听辨实验和声学分析’ (Discrimination test and phonetic analysis of rhotacized finals in Beijing dialect), in Lín Tāo 林焘 and Wáng, Líjiā 王理嘉 (eds) *Běijīng Yǔyīn Shíyànlu* 北京语音实验录 (*Beijing Dialect Experimental Phonetics*), Běijīng: Peking University Press, 27–72.
- Wang, William S-Y. and Li, Kung-Pu (1967) ‘Tone 3 in Pekinese’, *Journal of Speech and Hearing Research* 10: 629–636.
- Xiè, Bóduān 谢伯端 (2010) ‘Húnán Chénxī fāngyán yīnxì 湖南辰溪方言音系’ (The phonological system of Chenxi dialect of Hunan), *Fāngyán* 2010.2: 179–190.
- Xiè, Xiǎoli 谢小丽 (2013) ‘Chóngqìng fāngyán f, h hùnyòng qíngkuàng diàochá yánjū 重庆方言f, h混用情况调查研究’ (Investigative study of f/h alternation in Chongqing dialects), *Journal of Neijiang Normal University* 28.5: 72–75.
- Yang, Chunsheng (2011) ‘The acquisition of Mandarin prosody by American learners of Chinese as a foreign language (CFL)’, PhD dissertation, Ohio State University.
- Yue-Hashimoto, Anne Oi-kan (1976) ‘Southern Chinese dialects: the Tai connection’, *Computational Analyses of Asian and African Languages* 6: 1–9.
- Zhū, Xiǎonóng 朱晓农 (2010) ‘Quánzhōu chíshēng lùn: jiān lùn quán zhuó qīng huà (xiāo chí) dī sòng gāo bù sòng 全浊弛声论: 兼论全浊轻化（消弛）低送高不送’ (Slack voice: the phonetic nature of the ‘murky’ sounds in Middle Chinese), *Yǔyán Yánjiū* 2010.3: 1–19.
- Zhu, Xiaonong (2012) ‘Multiregisters and four levels: a new tonal model’, *JCL* 40.1: 2–17.

CHAPTER FIVE

THE SINITIC LANGUAGES: GRAMMAR

Anne O. Yue

1 INTRODUCTION

The classification of the modern Chinese dialects is not without controversy. We shall follow the scheme of Yuan (1983) which embraces seven major groups—Northern, Wu, Xiang, Gan, Hakka, Yue, and Min—as well as the broader designation of Norman (1988) which groups them into three super groups of Northern, Southern (Hakka, Yue, Min), and Central (Wu, Xiang, Gan), except for Huizhou, which we assign to the Central super group. It will become apparent that major differences in syntax are found between the Northern and the Southern groups, with the Central group as a transitional type that shares features with both the Northern and the Southern dialects. On the other hand, some sub-groups of the Northern dialect such as Southeastern Mandarin or Jianghuai, Southwestern Mandarin (including Hubei), Shanxi (including the Jin group), and Shandong, are sometimes found to share characteristics with the Southern dialects. The differences between the North and the South, moreover, can often be understood in the light of contact with neighboring languages. The outline of grammar presented in this chapter is limited in breadth and depth, since information on the dialects in this capacity is still limited.

2 GRAMMATICAL CATEGORIES

The lack of inflectional morphology in modern Chinese makes it very difficult to define grammatical categories in terms of morphological features. The traditional classification of grammatical categories into *shízì* 實字 ‘full words’ and *xūzì* 虛字 ‘empty words’ is based on the degree of semantic contents associated with them: ‘full words’, such as nouns and verbs, carry lexical meaning while ‘empty words’, such as particles, carry no concrete meaning but only functional meaning. Syntactically the ‘full words’ can be defined most satisfactorily by their co-occurrence patterns with other categories.

2.1 Substantives

2.1.1 Noun and classifiers

The noun (N), which is generally not distinguished by number, can be defined by its co-occurrence ability with a determiner phrase (DET), which contains a measure word or a classifier (CL)¹ with a demonstrative (DEM) or a number (NUM) or both. Countable nouns may take either classifiers which are mostly idiosyncratic to particular nouns or standard measure words while uncountable nouns can only take standard measure words.² For example: 一個橘子 [jí⁵⁵kə tey³⁵ts] ‘an orange’ or 兩斤橘子 [ljan²¹⁴⁺¹¹ tein⁵⁵ tey³⁵ts] ‘two catties of oranges,’ but 三杯水 [san⁵⁵ pei⁵⁵ swej²¹³] ‘three glasses of water.’³

In general, the Southern dialects have a greater number of classifiers than the Northern. The farther north one travels, the smaller the variety of classifiers found. In Dunganese, a Gansu dialect of Northern Chinese spoken in Central Asia, only one classifier, 個 [kə], is used; and this same classifier has almost become the cover classifier for all nouns in Lanzhou of Gansu too. The tendency to use one general classifier for all nouns is also found to a greater or lesser extent in many Shanxi dialects, some Shandong dialects, and even the Shanghai dialect of Wu and Standard Mandarin (SM). The choice of classifiers for individual nouns is particular to each dialect. For example, although the preferred classifier across dialects for ‘human being’ is 個 [kə] and its cognates, 隻 [tsɿ] in its dialect forms is widely used in the Hakka and the Yue dialects of Guangxi and western Guangdong province as well as in the Northern Min dialects and some Xiang dialects in Hunan.

2.1.2 Personal pronouns

While the first and the second person pronouns find cognates across the Chinese dialects, the third person pronoun is most diverse, even within major dialect groups. 他 [t'a⁵⁵] and its cognates are widely used only in the Northern, most Xiang, and a small number of Wu (for example, Yíxíng, Liyáng, Jinhuá, Danyáng, Jingjiang, Chángzhou, Wúxi) dialects;⁴ 佢 [k'-l̩] or 其 [ki] and its cognates are used in the Yue, the Hakka, most Gan, some Wu—especially S(outhern) Wu (Chángshú, Húzhōu Shuanglín, Zhujì, Yúyáo, Ningbo, Huángyán, Wenzhou, Qúzhou, Jinhuá, Yǒngkang), the Huizhou (Jixi, Shèxiān Túnxi, Xiuning, Yixiàn, Qímén, Wùyuán), some W(estern) Min (Yōng'an, Shaxiān, Sanmíng Sanyuán, Jiàn'ou, Nánpíng Xiáyáng, Jiànyáng, Songxi, Zhènghé, Shùnchang Yángdun, Pǔchéng, Jiānglè, Míngxi) and a small number of Xiang dialects (Suining, Chéngbù, Wǔgang, Xùpǔ, Xinhuá, Qíyáng, Máiáng); while 伊 [ji] and its cognates are used in most Min and Wu dialects.⁶

The personal pronouns (PN) are marked by a plural (pl) distinction in three different forms, recalling the different historical stages of the development of such a distinction: in the form of a suffix (highly developed) in the majority of the dialects, in the form of a noun (pre-grammaticalization stage), and in the form of a phrase (where the plural is a descriptive rather than a grammatical concept) (Table 5.1). While the suffix 們 [mən] and its variants are widely used in most Northern dialects, a great variety of other suffixes exist across dialects, including 都 [tou] (Northern: Hánđan and Wǔ'an of Hebei, Huòjia of Henan), 兜 (Northern: Anyáng of Henan, Hakka dialects such as Dongguǎn Qingxi, Cónghuà, Xiānggǎng Shenzhèn, Méngshan Xihé), 岸 [ŋan] (SW Mandarin: Táoyuán of Hunan), [ɣ³¹ ~ i³¹ ~ A³¹ + ts^{i³¹}] (NW Mandarin: Fúfeng of Shaanxi), [le]/[ne] (Gan dialects of Chálíng and Yángxin, Wu dialect of Wenzhou), [t'i]/[ti]/[li] (Xiang: Píngjiang Chángshòu of Hunan, Hakka dialects of Lúchuan and Liánnán, Gan dialects of Lílín, Xinyú, Yífeng, Píngjiang, Xiushuǐ), [na] (SW Mandarin: Línwǔ of Hunan), [no²¹]/[tei⁴⁴]/[tei²¹] (SW Mandarin: Rúchéng of Hunan), [tse] (Gan: Pǔqí of Hubei), [koli] (Gan: Nánchang of Jiangxi), [uo³⁴] (Gan: Tàihé of Jiangxi), [tej²²] (majority of the Yue dialects), and [n] (S Min: Shàntóu of Guangdong). It is not unusual to form the plural by suffixing certain nouns such as 人 [zən] /儂 [noŋ] /伙 [xwɔ] ‘person’ (NW Min, Hakka dialects such as Tónggǔ Sandu, Huizhou dialects such as Shèxiān, Jixi, Túnxi, Xiumíng, Yixiān, Dàtián of Central Min) or 等 [trŋ] (人) ~ 登 [ten] (Hakka dialects of Méixiān, Wengyuán, and Língxiān in Hunan) or 倘 (Danyáng of Wu, Hakka dialects such as Anyì of Jiangxi, Chángting of Fujian) or 多 [twɔ] (Yiyáng of Gan, NW Min such as Jiànníng, Shàowǔ) or 大家 [tateja] ‘everyone’ (Shùnchang of Min, Qímén of Huizhou). After all, plural suffixes are probably derived from earlier nouns.⁷ The plural may also be indicated through

suffixing a phrase such as 多人 ‘several people’ (Hakka dialects such as Nínghuà, Gan dialects such as Yúgan, Nánchéng), 些人 [če zən] / 幾個 [jikə] / 一下 [jieja] ‘several’ (SW Mandarin: Guǎngjì of Hubei for second and third persons) or 幾個人 ‘several people’ (Northern: Sùqian of Jiangsu) or 各 [kʂ] 人 (E Min dialects) or 個人 (Xianju of Wu) or 勒人 [lə zən] (Yǒngkang of Wu).

Sometimes the root and the suffix undergo contraction and produce a composite form (for example, 我 [an] < 我+們⁸ for 1pl is common among Shandong dialects such as Mùpíng, Píngdù, Wéifang, Zibó; also in Guǎngjì of Hubei), found in some Hebei (Northern: Wèixiān has 僮 for 1pl, 您 [nin] for 2pl and [nie⁴³]/[nə⁴³] for 3pl), Shanxi (Hóngtòng and Línfén have [hua] < 我家 for 1pl, [nia] < 你家 for 2pl; Língchuan has [uə²¹³] < 我們 for 1pl, [nə²¹³] < 你們 for 2pl and [t'ə³³] < 他們 for 3pl; Shanyin may have [ua⁵²] for 1pl and [niəu⁵²] for 2pl with 們 optionally suffixed), and Jiangsu dialects ([t'am] for 他們 in Gànýú). S Min dialects such as Shántóu and Cháozhou of Guangdong have contracted forms ending in [ŋ] in the plural ([nan⁵³] for 1inc pl, [un⁵³] for 1ex pl, [nin⁵³] for 2pl and [in³³] for 3pl).

Note that the plural marker may differ for first versus second versus third persons, especially in the Wu dialects: the Suzhou dialect of Jiangsu uses the suffix [to?] for the second and the third persons, but a contraction form [ni] for the first person, Shanghai uses the suffix [ŋi] for the first person, the suffix [la^{53>44}] for the third person, but a contraction form [nʌ²³] for the second person, the Häßián dialect of Zhejiang uses the suffix [la] for the first and the third person but a contraction form [na] for the second person, for example.

Tonal variation as a plural device is found in a number of Yue (Zēngchéng, Yángjiāng, Héshan and the Siyi dialects of Xinhui, Táishan, Kaipíng, Enping, use a glottalized low falling tone), Wu (Shanghai Fèngxián), as well as some Northern dialects of Shanxi (Línyí may use a falling-rising tone besides lengthening the rhyme, Yúnchéng uses a long low rising tone optionally followed by the suffix [ti]) and Shaanxi (Xi'an and Bǎoji use a low falling tone; Shàngxiān).

The feature of visibility figures in the third person of the Xiang dialect of Píngjiāng Chángshòu, which is unheard of in other Chinese dialects: 渠 [k'-] for [+visible] and 他 for [-visible].

The Mùpíng dialect of Shandong distinguishes the feature [proximal] 近 ([+prox]) versus [distal] 遠 ([+dist]) for the plural form of pronouns through incorporating the [+prox] demonstrative [tejə⁵³] 這 ‘this’ in the former and the [+dist] demonstrative [ŋjə⁵³] 乜 ‘that’ in the latter: [tsa⁵³tejəejər⁵¹] 咱這些兒 versus [tsa⁵³ŋjəejər⁵¹] 咱乜些兒 for the first person plural inclusive, [an²¹³tejəejər] 僮這些兒 versus [an²¹³ŋjəejər⁵¹] 僮乜些兒 for the first plural exclusive, [na²¹³tejəejər⁵¹] 嫌這些兒 versus [nd²¹³ŋjəejər⁵¹] 嫌乜些兒 for the second person plural and [t'a⁵¹tejəejər⁵¹] 他這些兒 versus [t'a⁵¹ŋjə⁵³ejər⁵¹] 他也些兒 for the third person plural.

The distinction between inclusive and exclusive first plural is found in many Northern dialects (in, for example, Hebei, Henan, Shandong, Shanxi, Shaanxi), some Xiang (Píngjiāng Chángshòu), some Wu (for example, Wenzhou, Chángzhou, Jiangyin), some Gan (such as Yífeng, Duxiang, Nánchéng), as well as most Min dialects (for example, Fúzhōu, Jiàn'ou, Jiànníng, Shàowǔ, Shùnchang, Pútíán, Xiàmén, Shántóu, Cháozhou, Dàtián).

Case distinction is almost unheard of, but it exists in the Xúnhuà dialect of Qinghai (NW Mandarin) probably as a result of contact with minority languages in the region.⁹

	1	2	3		1	2	3
sg nom	ŋə	ni	t'A	pl nom	ŋəmv	nimv	t'Amv
sg acc	ŋa	nia	tejə	pl acc	ŋama	niamə	tejəmə

In the Hǎiyán dialect of Wu, personal pronouns distinguish between a pre-pause and a post-pause form, often correlating with pre-verbal and post-verbal usage, reminiscent of a case distinction too perhaps. For example, in the Tongyuán subdialect we have for the first and third person pronouns:

- Pre-pause first person singular: fio? no third person singular: ji? nei
Post-pause third person singular: ji? nei third person singular: i

The post-pause form after a verb usually falls within the same breath group with the latter.¹⁰

2.1.3 Pronominalization

Noun phrases (NP) of identical reference that appear in a series of clauses, if understood by context, are seldom pronominalized (see section 4.13 for details). Across sentences, if there is pronominalization, it generally involves only [+human] NPs. However, in the Southern and Central dialects, resumptive pronouns with reference to [-human, -animate] NPs are often found in the object position within the same sentence, whereas in Northern Chinese, not to say resumptive pronouns but pronouns in general are not frequently used for [-human, -animate] NPs except when used in a disposal construction (see section 4.6). For example, suppose in a previous context a radio was bought, and then it was asked how much it cost. A pronoun cannot be used to refer to the radio but the term ‘radio’ has to be repeated, or more often, elision occurs:

- 1a. SM (我wo^{213>35} 買maj^{213>11} 了lə 個kə 收音機sow⁵⁵jin⁵⁵tei⁵⁵)
1sg buy PFV CL radio ‘I bought a radio’
(收音機sow⁵⁵jin⁵⁵tei⁵⁵) 多少two⁵⁵saw 錢te'jen³⁵?
(radio) how-much money ‘how much is (the radio)?’

If a previous context is provided, the ‘radio’ can be pronominalized in a disposal construction:

- 1b. SM 把pa^{213>11} 它t'a⁵⁵ 關上kwan⁵⁵san !
BA 3sg turn-off ‘turn it off!’

In the Wu dialects, the object may be topicalized leaving a resumptive pronoun as its trace in the post-verbal position, for example:

- 2a. Shanghai [die?] 碗湯 倒 脫 伊
this CL soup pour out 3sg ‘throw out this bowl of soup’

In the Southern dialects, however, a resumptive pronoun may be used together with an NP in the object position with or without previous context. In other words, even with previous context provided, a resumptive pronoun may be used within the same NP. For example, in Cantonese:

- 2b. 倒tow³⁵ [tʃɔ³⁵] [ni⁵⁵] 碗wun³⁵ 湯t'ɔŋ⁵⁵ (佢k'ɔy²⁴)
pour PFV this CL soup 3sg ‘throw out this bowl of soup’

TABLE 5.1 SAMPLE PLURAL FORMS OF PERSONAL PRONOUNS¹

Dialect	First person				Second person				Third person			
	-suffix	N	NP	contraction (C)	Tone (T)	-suffix /NP//C//T	-suffix /T	-suff NP/C	t'a/他 k'渠	ki其	i伊	NP/C//T/-suffix
N(orthern)	1 in pl				1 ex pl	2 pl		3 sg		3 pl		
Xiang (Shuangfeng)	tsən-man			tsən咱		wɔ-man	ni-mən	—	t'a			ta-men
Wu (Suzhou / Shanghai)	ŋɔ-li				ŋo-li	n-li						t'o-li
Yue (Siyi/ Guangzhou)	—			ni aʔla(?)	ni/ ^ŋ yu- aʔla(?)	n-to?	—					li-toʔhi-la(?)
Huizhou (Jixi)	(ŋ)ɔ-ŋi				ŋɔj ²¹	/ŋɔj ³³⁻²¹	/ŋjak ²¹	k'ui k'oy				/k'iaŋ ²¹ k'oy-tej
Hakka (Meixian)				ŋaj-ten-ŋin		n+ŋā	nej-tej					ki+ŋā
Gan (Nanchang)				ŋə我kə个li里								ki-ten-ŋin
Min (Fuzhou)				nannja (kouʔ)nneyŋ			/ŋkəli					tejekəli
N (Muping)				tsatɛjɛtɛj远 tsəŋjɛtɛj遠			/ŋwaj					i ~ ~
Min (Xiamen) (Chaozhou)							kouʔnoŋŋy	ny~				
							/antejɛtɛr	natejɛjɛtɛ近(NP)				t'atejɛjɛtɛ近NP
							amŋjɛtɛr	nanŋjɛjɛtɛ遠(NP)				t'anŋjɛjɛtɛ遠NP
							/nat(C)	/nat(C)				
									/lin /mij			/m /ŋ

¹ See Figure 5.1 for the singular forms of various dialects. Tone marks are left out for lack of space. C stands for contracted form.

2.1.4 Demonstratives

The forms of the demonstratives are even more varied and it is almost impossible to establish cognates across dialects. The most common pattern of distinction is two-way: proximal versus distal. The use of a word beginning with an affricate initial for the proximal and one with a dental nasal (a lateral being its equivalent in Jianghuai) initial for the distal demonstrative is prevalent among a vast number of Northern dialects. The majority of the Min dialects also use a form with an affricate initial for the proximal (except Jiān'ou, Jiānyáng, Songxi) but one with a velar fricative or glottal fricative [h] initial for the distal (except Jiān'ou, Jiānyáng, Songxi, Shaxiàn, Yōng'an).

Among other dialect groups, internal diversity is the norm. For example, although the majority of the Yue dialects use a form with a dental nasal/lateral initial consonant for the proximal and a form with a velar plosive [k] initial consonant for the distal deictic, the Siyi dialects, Yángjiāng, and Zhōngshān use forms with exactly the opposite initials for the same distinction, while a few dialects (Nánhǎi, Shùndé, Sanshǔi) use the same initials but variation in vowels for the distinction—higher front/central vowel for the proximal and lower, back vowel for the distal. Most Hakka dialects use a form with a velar stop [k] initial for the distal deictic but varied forms for the proximal deictic, while most of the Gan dialects are just the opposite, using a form with a velar stop [k] initial for the proximal deictic but varied forms for the distal deictic. Although some Xiang dialects (Lúxī, Yuánlíng, Gǔzhàng, Bǎojīng, Xiāngxiāng, Anxiāng, Níngxiāng) have adopted the Northern forms, a sizable number use a form with a velar stop [k] initial for the proximal deictic and one with a dental nasal/lateral initial for the distal deictic, but diverse forms are also used for the latter. The Wu dialects show by far the most variety in the use of the demonstratives, even within a single dialect. It is not uncommon to have two or three variants for each of the demonstratives in the same dialect. There are at least three different ways of expressing the proximal versus distal contrast. Dialects like Suzhou, Lìyáng, Danyáng, Chángzhōu, and Huángyán use forms with front vowels for the proximal and those with non-front vowels for the distal deictic. Dialects like Zhūjī, Qúzhōu, and Jīnhuá use forms with velar initials for the proximal and forms with labial initials for the distal deictic. This second category has a variant in dialects like Shèngxiān Chóngrénn, Shèngxiān Tàipíng, and Níngbo, where a syllable with labial initial is added to the proximal deictic to form the distal deictic. The third type uses cognates of the third personal pronoun 伊 [?i] to indicate the distal deictic, found in dialects such as Bǎoshān Shuāngcǎodūn, Bǎoshān Luódìàn, Nánhuī Zhōupǔ, Shanghai, and Songjiāng.

The feature visibility is reported to exist in the Jiaxiān dialect of Henan, where the distal demonstratives distinguish between [vei²³] with the feature [+visible] and [nai²³] with the feature [-visible].¹¹

A small number of dialects, some Northern dialects in Shandong (Zibó, Shòuguāng, Wéifāng, Mùpíng), Shanxi (Yuánpíng, Shòuyáng, Yángchéng, Shílóu, Zhōngyáng, Línxiān, Liǔlín, Yúshè, Yúxiān, Xiyáng, Jinglè, Línfēn, Xinzhōu, Héshùn, Yángqū, Tàiyuán, Wànróng) and Hubei (Zhōngxiáng, Yǐngshān, Luótíán, Xishūi, Jiayú, Huáng'an, Qíchún, Gōng'an, Hēfēng, Zǐgui, Hóng'an, Máchéng, Huángméi), some Wu (Suzhou, Hái yán Tongyuán, Wúxī, Chángshú, Shèngxiān Chánglè), some Hakka (Lùfēng,¹² Xinfēng) and some Min (Cháo zhōu) dialects have a tripartite distinction of proximal versus medial versus distal.

In addition, the Shèngxiān Chánglè dialect maintains this tripartite distinction together with a distinction between a stative versus an active state. A location is stative while a

TABLE 5.2 TRIPARTITE DISTINCTION IN DEMONSTRATIVES

goal toward which some movement is directed is active. The stative demonstratives are [ku⁴⁴] for proximal (visible and touchable), [løy³⁵] for medial (invisible and untouchable), and [mon³⁵] for distal; while the active demonstratives are [kua⁵³] for proximal, [lia⁵³] for medial, and [mar⁵³]/[mon³⁵ha⁵³] for distal. Ogawa (1981) observed that this distinction in Chinese also exists in various Tai (for example, Siamese) and Austro-Asiatic (Vietnamese, Khasi, Palaung) languages, speculating that it may have been inherited from Proto-Sino-Tibetan and that its loss in most Northern dialects may be due to its non-existence in Mongolian and Manchu.

On the other hand, it is reported that in some Wu dialects such as Kunshan or Huángyán, the younger generation merges the distal versus proximal distinction into one, that is, using the same form for both designations.

Table 5.2 summarizes the occurrence of three-way distinctions in the demonstratives across a sample of dialects.¹³

Figures 5.1 and 5.2 provide the basic deictic forms of 18 dialects including demonstratives, personal pronouns, locative and temporal deictics, as well as manner and degree adverbials. The 18 dialects are chosen from the seven major groups, among which only four, Suzhou, Jixi, Tàiyuán and Xi'an, have a tripartite distinction across all deixis. Figure 5.1 includes Southern dialects as well as Gan dialects from the Central group while Figure 5.2 comprises Northern dialects as well as Xiang dialects of the Central group. The two figures show some similarity as well as diversity in how the distinction between the deictics is represented. Very roughly speaking, the basic system of deixis in most of the dialects has a two-way distinction of [+prox] versus [+dist] expressed by means of the place of articulation of the initial consonants. In Figure 5.1 the feature [\pm anterior] marks such distinction in Xiàmen, Fúzhou, Jiānyáng (Min dialect group) and Cantonese/Guāngzhōu (Yue dialect group); and the feature [\pm high] in Nánchāng (Gan) and Wenzhou (S Wu). The vowel feature, [\pm front], is used in Suzhou while the tonal feature of [\pm high] is used in Méixiàn (Hakka). Táishān (Yue) may be marked by the [\pm high] feature in terms of the point of articulation and the feature [\pm nasal] in terms of the manner of articulation. In Figure 5.2 the feature [\pm sonorant], either [\pm nasal] or [\pm lateral], marks the said distinction in Shuangfeng and Chángshā (Xiang), in Yángzhōu (Jiānghuái or SE Mandarin), Wǔhàn (SW Mandarin), Mùpíng and SM (N Mandarin); while double features, [\pm coronal] and [\pm sonorant], are used in Tàiyuán (NW Mandarin), as well as [\pm anterior] and [\pm sonorant] are used in Xi'an (NW Mandarin). The Jixi dialect (Huīzhōu) distinguishes itself by the use of [\pm coronal] for the initial consonant and [\pm back] for the vowel.

2.2 Verbs and aspect

2.2.1 Verbs

The verb (V) can be uniquely defined as negatable. The so-called adjectives are but stative verbs (Vstat), erroneously identified as equivalents to adjectives in familiar Indo-European languages. Vstat are not only negatable but function as predicates just as other kinds of verbs do. The distinction between Vstat and intransitive verbs (Vi) lies in the former being able to be modified by degree adverbs (ADVd) but not the latter. Nor is Vstat the only class of verbs that can be modified by ADVd, the so-called transitive verbs of quality (Vtrq)—for example, 喜歡 ‘to like,’ 愛 ‘to love,’ 憎 ‘to hate’—or optative verbs (Vopt)—for example 敢 ‘to dare,’ 願意 ‘to be willing,’ 會 ‘to be capable of,’ 應該 ‘to ought to’—can be so modified too.

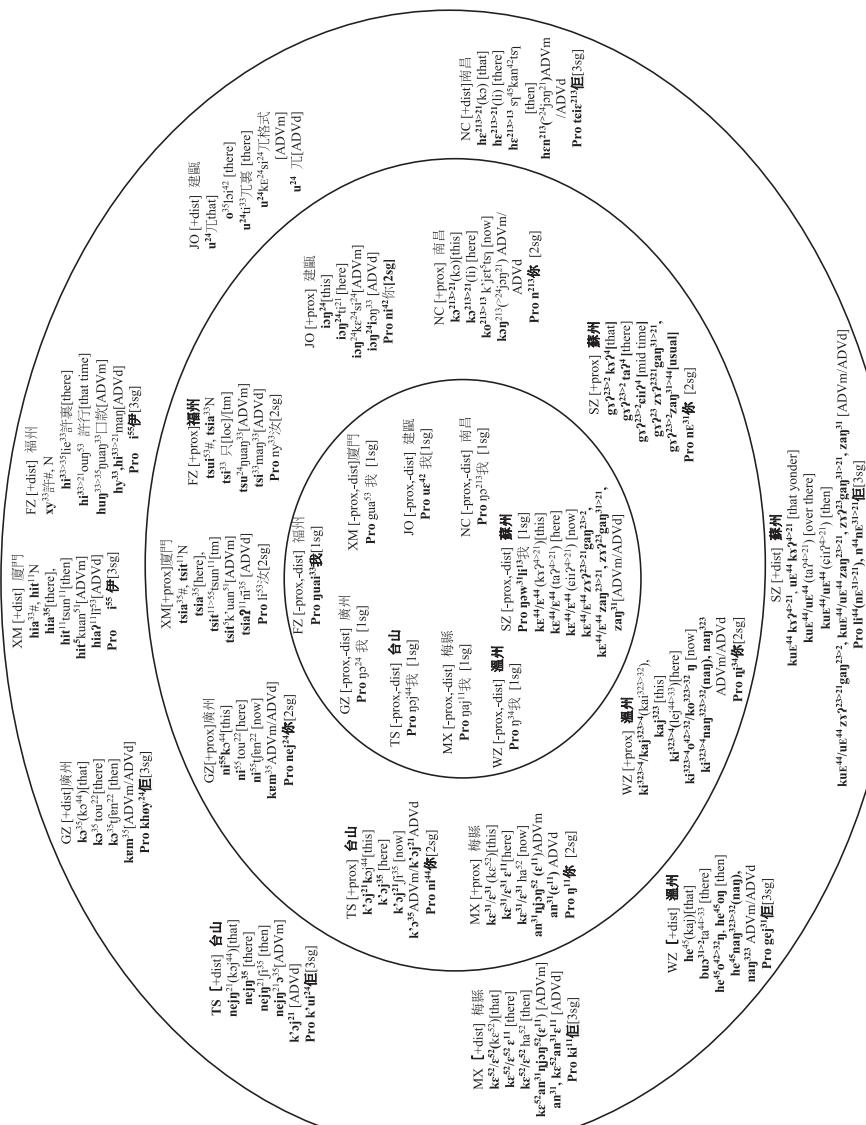


FIGURE 5.1 DEICTIC FORMS OF MIN, YUE, HAKKA, WU AND GAN DIALECTS

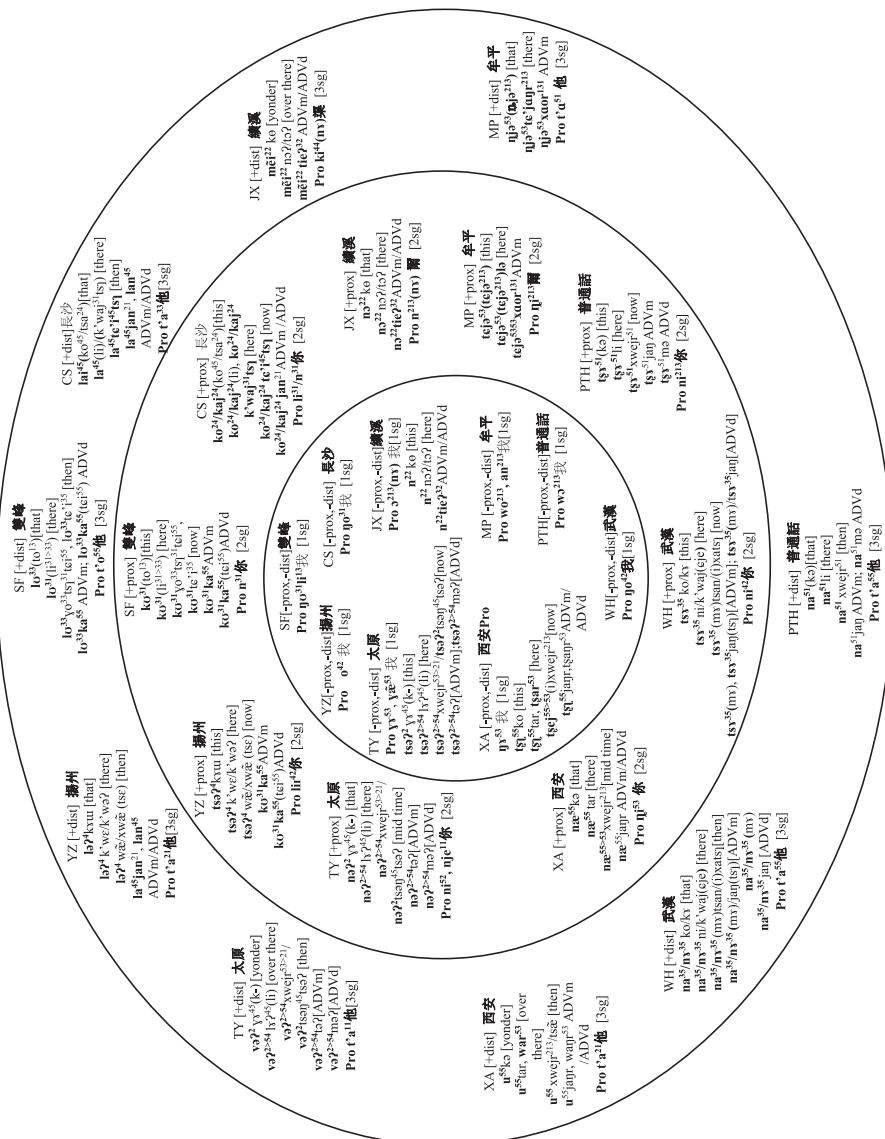


FIGURE 5.2 DEICTIC FORMS OF XIANG AND MANDARIN DIALECTS

2.2.2 Aspect

Aspect markers (ASP) in the modern dialects are mostly derived from verbs. In other words, certain verbs become grammaticalized into function words (Table 5.3). This grammaticalization process took place at different times and at different paces in different dialects. As a result, not to say across dialects, but even within a single dialect, different aspect markers may be at different stages of grammaticalization. In other words, while some dialects such as Min, where grammaticalization has barely started, employ verbs to indicate certain aspectual concepts and for certain aspectual usages, the majority of dialects employ a mixture of suffixes (where grammaticalization has completed) and complements (where grammaticalization is only half achieved). We shall define aspectual suffixes as those that occur in close conjunction with the verbs they modify, allowing no other elements to intervene, and aspectual complements as those that occur in loose conjunction with the verbs they modify, allowing other elements such as potential markers to intervene. Occasionally, tonal or vocalic modification is also used as a device to indicate aspectual categories in some dialects.

The Southern and to some extent the Central dialects have a richer variety of aspect markers than the Northern dialects. Generally speaking, most aspect markers mark specific kinds of actions and therefore occur in sentences or clauses describing specific events, unless they denote a sequence of actions, in which case they usually occur in a complex sentence. Due to the limitation of space, only the most common aspect markers across dialects, the perfective, the change-of-state *cum* new situation, the progressive, the durative, and the experiential, will be presented here.

2.2.2.1 Perfective aspect

The perfective aspect (PFV) signals the realization of certain state or action in the past or in the future, most commonly expressed by a suffix of the verb. In the Northern

TABLE 5.3 PERFECTIVE ASPECT, AFFIRMATIVE ASPECT, PAST TENSE

Dialects	Perfective aspect	Past tense	Affirmative aspect
Majority	(V +) PFV	V1(PFV)V2+ PFV V1V2+ PFV	
Northern	了(lə, li, lou, liou, lio, lə, liau, lie); Huòxiān: V+ou+O+la; NEG+V+lie; Huòjia: Rime change Vj/w>V, VN>v)ŋ, V>Vj/w; Shàngxiān: V lengthening + lowering, T > falling; Hǎiyáng, Mùpíng: V > [+retroflex])	Tàixing: V1PFV V2PFV O	V+了 Huòxiān: VP+li<來
Huizhou	著teo, 到tr, 得te		
Wu	tsə?, də?, de?, te?, tsı, zı, la, da, ta; Yöngkang: VV	Wenzhou, Qingtián: V1PFV+V2	
Xiang	ta/tə, tie, tu, tau; 得te; ka, ga, kə, ku, ko, kua; kata, kota, kətə		
Gan	i, e, ε, ə; te, tə, te, te?; li, le, lo, lə		
Hakka	e, le, li, lei, liau, lə, liə, luə		Vex+V Vex+V
Yue	tʃ, Tone>[+Rise]; Siyi: ə		Vex+V Vex+V
Min	lə, liau; Shántóu: V+好ho ⁵³ (favorable connotation); Quánzhou, Shántóu: V+去k'u (unfavorable connotation)		Vex+V

dialects it is overwhelmingly 了 [lə] derived from a verb with the meaning of ‘to finish,’ and its variants [li]/[lou]/[liou]/[lio]/[lɔ]/[liau]/[lie], although in the Huizhou dialects 著 [teio] or 到 [ts] or 得 [te] is also used; in the Wu dialects it is mostly [tsə?] and its variants [də?]/[de?]/[te?]/[tsi]/[zi]/[la]/[da]/[ta]; in the Xiang dialects it is mostly [ta]/[tɒ] and its variants [tie]/[tu]/[tau] or 得 [te] or [ka]/[ga]/[kɒ]/[ku]/[ko]/[kua] or a combination form [kata]/[kota]/[kɒtɒ]; in the Gan dialects it is [i]/[e]/[ɛ]/[ə] or [te]/[tə]/[te]/[tə?] or [i]/[le]/[lo]/[lə]; in the Hakka dialect it is [e] or [le]/[li]/[lɛi]/[liau]/[lɔ]/[lio]/[luɔ]; and in the Yue dialects it is [tʃɔ] or [ə] (Siyi). In the Min dialects, there is a perfective aspect marker [le] or [liau], grammaticalized to different degrees among different dialects, and such a concept may also be expressed with a complement verb such as 好 [ho⁵³] in Shàntóu (with favorable connotation) or 去 [k'w] in Quánzhou and Shàntóu (with unfavorable connotation).

When a sequence of two verbs occurs in a complement structure such as V1 + V2, the perfective aspect suffix in most dialects occurs only after V2. However, in the Tàixing dialect (Jiangsu, Jianghuai), it occurs after both V1 and V2 (as in example (3) below), while in some Wu dialects such as Wenzhou or Qingtián, it intervenes between V1 and V2, as for example 壓[la]扁 ‘pressed flat’ (Wenzhou).

3. Tàixing 我 說 [ka] 服 [ka] 他 了
1sg talk PFV convince 3sg Fp¹⁴ ‘I (have) convinced him’

The Southern dialects maintain a distinction between the perfective aspect and the past tense (PAST) which have merged in the great majority of the Northern dialects. The so-called past tense is expressed with the existential verb (Vex) occurring before the verb modified. Compare the following (FP = final particle):

- 4a. Cantonese:
 你nei²⁴有jəw²⁴無mow²⁴食sik²飯fan²²啊a⁴⁴? 有jəw²⁴(食sik²)/無mow²⁴(食sik²)
 (Yue) 2sg have not-have eat rice FP have (eat) /not-have (eat)
 ‘did you eat?’ Yes (I) did / No (I) did not’
- 4b. Cantonese:
 你nei²⁴食sik² [tʃɔ³⁵] 飯fan²²未mej²²啊a⁴⁴? 食sik²[tʃɔ³⁵]喇la/([tʃɔŋ²²])未mej²²食sik² 啊a
 2sg eat PFV rice not-yet FP eat PFV FP/ (still) not-yet eat FP
 ‘have you eaten yet?’ (I) have eaten/ (I still) have not eaten’
- 4c. SM: 你ni^{213>11}吃tʂ¹⁵⁵了lə 飯fan⁵¹沒mej³⁵有²¹³? 吃tʂ¹⁵⁵了lə / 沒mej³⁵ 吃tʂ¹⁵⁵
 2sg eat PFV rice not-have eat PFV / not-have eat
 ‘did you eat? / have you eaten? Yes (I) did ~ (I) have/No (I) did not-(I) have not eaten’

It must be pointed out that this Vex preceding another verb also functions as the affirmative aspect (af-asp) in the Southern dialects, affirming the existence of some action or state. For example:

5. Cantonese: 我ŋɔ²⁴屋企ok⁵k'ej³⁵ 有jəw²⁴ 養jœŋ²⁴ 雞kej⁵⁵
 1sg family af-asp raise chicken ‘we raise chickens’

The only known Northern dialect to still maintain a distinction between the perfective and the past by marking it with different suffixes and final particles is Huòxiān (Shanxi):

- 6a. 你 吃 [lou] 飯 [la] 麼? 吃 [la] / (還) 無 吃 [lie]
 2sg eat PFV rice Fp FP eat Fp / (still) not-have eat Fp
 'have you eaten?' (I) have eaten/(I still) have not eaten'
- 6b. 你 吃 飯 [li] 麼? 吃 [li] / 無 吃
 2sg eat rice PAST FP eat PAST/ not-have eat
 'did you eat?' Yes (I) did/ No (I) did not'

The final phase particle [lie] can only and must be used with the negative answer to a question in the perfective aspect. The suffix [li] indicating past tense is probably derived from the verb 賴 ‘to come’ used to denote past experience found in texts since the Tang dynasty, such as the *Taiping Guangji* 《太平廣記}, *Jingde Chuan Deng Lu* 《景德傳燈錄》, *Yuanchao Mishu* 《元朝秘史》, and *Piaotongshi* 《朴通事》.¹⁵

In some dialects, the perfective aspect takes the form of rime change. For example, in the Huòjia dialect (Henan, Northern), it is a simplification of the complex finals, with diphthongs becoming monophthongs, finals with nasal endings reduced to nasalized vowels, and breaking of the simple finals into diphthongs; while in the Shàngxiān dialect (Shaanxi, Northern), it is the lengthening and lowering of the vowel plus a change of each tonal segment into a falling contour. For example:

7. Huòjia 我 去 te³¹ < te³¹ yε³¹ 兩趟
 1sg go two times 'I went/have gone twice'
8. Shàngxiān 病 piȧ⁵⁵¹ < pi̇⁵⁵ 幾 天 啦
 sick several day FP 'been sick for several days'

In other dialects, tonal modification is used, for example, a high rising tone may optionally be used to mark this aspect in Cantonese; or reduplication of the verb, such as in Yōngkang (Southern Wu); or retroflexion of the verb, as in Hǎiyáng and Mùpíng (Shandong).

2.2.2.2 Change-of-state *cum* new situation

Across the dialects, there is a phase particle, Fp, occurring at the end of a phrase, clause, or sentence, which is used to indicate a change of state or a new situation pertaining to the past or to the future but with reference to and effect on the present (Table 5.4). Very often, it occurs together with the perfective aspect in the same sentence. For this final phase particle, in the Northern dialects, a particle with the same phonetic form as the perfective

TABLE 5.4 FINAL PHASE PARTICLE

Dialect	Northern	Wu	Xiang	Gan	Hakka	Yue	Min
Fp	lə	lə?, lɪr?	lo, le, li; ta, kata	得 te, te, tʂ?; 了 li, le, liau, lo	了 le, lei, lo; e	la, lɔ	lɔ, ou

aspect marker but derived from a different historical source is used,¹⁶ for example, see the following sentence:

9. SM: 吃^{tʂ’i⁵⁵} 了^{lə} 飯^{fan⁵¹} 了^{lə}
 eat PFV rice Fp

This sentence has two readings: (1) ‘(it is) ready to eat now’; (2) ‘has already eaten (with the implication that the person does not want to eat any more)’. The first reading has the Fp indicating a new present situation relating to the imminent future (from ‘not yet ready’ to ‘ready’) while the second reading has the Fp indicating a new state of affair relating to the past that bears on the present situation.

Similar to the Northern dialects, the phase particle may share the same form with the perfective aspect marker in some Wu dialects, while in others a different form [lə?]/[li?] is used; in the Xiang dialects, either [lo]/[le]/[li] or the same form as the perfective marker [ta]/[kata] is used; in the Gan dialects, either some form of 得 [te]/[te]/[tx?] or of 了 [li]/[le]/[liau]/[lo] is used; in the Hakka dialects, either some form of 了 [le]/[lei]/[lo], or [e], which may be identical in form with the perfective aspect in some dialects, is used; in the Yue dialects, this final phase particle, [la]/[lo], is distinct from the perfective aspect marker; and in the Min dialects, some weakened form of 了, [lə] or [lou], is used.

2.2.2.3 Experiential aspect

Except for Southern Min, almost all dialects use the suffix 過, [kwɔ̃] and its variants, for the experiential aspect (exp-asp), which signifies the occurrence of a certain action or state at least once in the past. Other suffixes are rare—來 is reported for Wèixiān (Hebei, Northern), 著 [tʂo³¹] is reported for Kunming (Yunnan, SW Mandarin) and [ta] for Jingmén (Hubei, SW Mandarin). In S Min, the verb [bat]/[pak] is used before another verb for the same function, although the suffix 著 [tio?] and 過 [k(w)e] may also be used. In some of these dialects, the latter two may co-occur with [bat]/[pak] to mark this aspect. The use of [k(w)e] is probably due to the influence of either SM or Cantonese (which is the lingua franca in Guangdong province where many S Min dialects are spoken) (Table 5.5).

2.2.2.4 Progressive aspect

Like in many other languages of the world,¹⁷ the majority of the dialects use a form or a phrase derived from an indefinite locative expression with a locative verb (Vloc) and a locative deictic for the progressive aspect (prog-asp), signifying an ongoing action or state. In the Northern dialects, this formation is not very common, but it is reported in Jiaochéng (Shanxi), Xishūi (Hubei) and Chéngdū (Sichuan). In the N Wu dialects, the crystallized locative phrase [lə?he]/[lə?lə]/[lə?to?] and their variants are suffixed to the verb; and in the S Wu dialects such as Wenzhou, it is realized as [zita].¹⁸ In the Huizhou dialects, a similar

TABLE 5.5 EXPERIENTIAL ASPECT

Dialect	Northern	Wu	Xiang	Gan	Hakka	Yue	Min
Ex-asp	過 kwɔ̃; 來 (Wèixiān); 著 tʂo ³¹ (Kunming); ta (Jingmén)			過			S Min: bat+V; V 著 tio?; V 過 k(w)e

crystallized phrase, 是 [nor³]/[kr¹¹][lo¹¹],¹⁹ sometimes with proximal (在 [n²²] [tɔ?]/在 [n²²] [na]/是 [mo⁵³] [le³³]/是 [ka⁵³] [le]) versus distal (是那 [tɔ?]/是 [mēi²²] [tɔ?]/在 [na³¹³] [na]/是 [ka⁵³] [le])/是 [mo⁵³] [le³³]) distinction, forms the progressive suffix.

Some dialects have more than one form for the progressive. In the Yue dialects, while a crystallized locative phrase such as [hej³⁵]道tow²²/[hej³⁵]處sy⁴⁴ (Cantonese) or [ɔ³³] [nein³⁵] (Siyi dialects) may be used,²⁰ there is also a distinct progressive suffix [ken³⁵] (Cantonese)/[kin⁵⁵] (Siyi). While [kin]/[ten] is also used in many Hakka dialects, which is either cognate with the Yue form or a loan from the Yue dialects, [ten³¹] or [ts'əŋ⁵²] [t'i⁵²] is used in the prestige Hakka dialect Méixiān. Other dialects have the progressive share the same form with some other aspect marker. For example, in the Xiang dialects, the perfective aspect marker [ta] also functions as the progressive while the suffixes 起 [te'i] and 倒 [tau]/[tou] function both as the progressive and the durative. In the Northern dialects, the progressive and the durative aspect markers share the same suffix 著 [tʂə] with a weakened form [tə], although for the progressive aspect the final particle (FP) [nə] must co-occur, while the locative verb 在 [tsai] is used as a prefix in a less colloquial style for the progressive aspect. The progressive and the durative share the same suffix 起 or 倒 in SW Mandarin (Guìyáng of Guizhou, northeastern Yunnan) and the same suffix [tɔ?] in many Shanxi dialects. 在 [ts'aɪ] is also used as a prefix in the Gan dialects but a suffix in some Anhui dialects (Héfēi, Jianghuai) for the progressive; it can also be suffixed to other progressive suffixes to indicate the same (Northern—Huòqiū of Anhui, Chénxi of Hunan, Sichuan; Gan—Pǔqi of Hubei). Parallel to the situation in the Wu dialects, in the Min dialects, 在 [te]/[le] is prefixed ([le²¹] in Fúzhōu) to a verb to indicate the progressive aspect and suffixed ([le] in Fúzhōu) to it to indicate the durative aspect. The prefix/suffix 在 may best be regarded as an abbreviated form for the indefinite locative expression mentioned at the beginning of this section. The S Min dialect of Shàntóu renders support to this assumption. In this dialect, among the entire array of some seven variants, [lo]/[to]/[toko] and [na]/[nakə]/[pan]/[panko], prefixed as the progressive and suffixed as the durative, [to] can be identified with the verb ‘to locate,’ [pan] with the verb ‘to place,’ and [ko] with a locative indicator (compare [tsi^{53>24} ko] ‘here’ and [hur⁵³ko] ‘there’ in Cháozhōu).

In some dialects a reduplicated progressive form is used to indicate the progression of some action in the midst of which another action or state is triggered. The progression of action is expressed in a clause always followed by another clause describing the triggered action. In the Northern dialects, the reduplication of the progressive, V[tʂə]V[tʂə], indicates such a progression. In the Xiang dialects, V[ta]V[ta], V[tau]/[tou]V[tau]/[tou], V[ts'an]V[ts'an] (Yíyáng, Xiāngtán), V[te'y]/[dy]V[te'y]/[dy] (Lóudǐ, Xiāngxiāng), are all reduplicated progressive forms. In the Gan dialects, V[a]V is the predominant form, while V住 [k'y]V[k'y]/V到 [tau]V[tau]/V[sā]V[sā] are also used. In the Yue dialects, a suffix 下 [ha³⁵] is used with a reduplicated verb to indicate the same, namely, VV [ha³⁵] or V[ha²⁴]V[ha²⁴]. Some Northern (Huòqiū, Shèxiàn of Anhui) and Wu dialects (Shanghai) simply use a reduplication of the verb, VV, for such a purpose (Table 5.6).

2.2.2.5 Durative aspect

In many Northern and Central dialects, the same marker for the perfective aspect is also used for the durative aspect (DUR) (Table 5.7). However, in many of these dialects, more than one marker is used for the durative. In the Wu dialects, the same form as the progressive aspect is suffixed to the verb to express another form of the durative aspect, namely, V + [ləʔhe]/[ləʔləð]/[ləʔtɔ?]. In the Xiang dialects, while the perfective aspect marker [ta] or [te] and 得 [tʂ] also marks the durative, the progressive suffixes, 起 [te'i] and 倒 [tau]/[tou], are so used too. Similarly, in some SW Mandarin dialects, the progressive suffix 起 or 倒, as mentioned in the previous section, also function as the durative. The suffix 倒

TABLE 5.6 PROGRESSIVE ASPECT

Dialects	prog-asp	
	prog-asp + V	V + prog-asp
Majority	Vloc+locative deictic + V	
Northern	Jiaochéng (Shanxi), Xishūi (Hubei), Chéngdu (Sichuan); 在tsai + V	Other: V+著tʂə/tə ... nə(=F); V+起/倒 (Guīyáng, NE Yunnan); V+tə? (Shanxi); V + 在ts'ai (Héfēi); V+prog-asp+ts'ai (Huòqiu of Anhui, Sichuan, Chénxi of Hunan)
N Wu	在te/lə + V	V+lə?he/lə?lə/lə?to?
S Wu		V + zita (Wenzhou)
Huizhou		V+是nəp ³ /kʂ ¹¹ lo ¹¹ , V+在n ²² tə?/在n ²² na/是mo ⁵³ le ³³ /是ka ⁵³ le; V+是那[tʂ]?/是mēi ²² tə?/在na ³¹³ na/是ka ⁵³ le/是mo ⁵³ le ³³
Yue	hej ³⁵ 道tow ²² /hej ³⁵ 處sy ⁴⁴ + V; ɔ ³³ nein ³⁵ + V (Siyi dialects)	V + _ken ³⁵ ; V + kin ⁵⁵ (Siyi dialects)
Hakka		V + kin/ten; V + ten ³¹ /ts'aj ⁵² t'i ⁵² (Méixiān)
Xiang	V + ta; V + 起te'i /倒tau/tou	V + ta; V + 起te'i /倒tau/tou
Gan	在ts'ai + V	V+prog-asp+ts'ai (Pǔqi of Hubei)
Min	在 te/lə + V; lo/to/toko/na/nako/paŋ/ paŋko+V(Shàntóu)	

TABLE 5.7 DURATIVE ASPECT

Dialects	V + DUR
Northern	V+著tʂə/tə; V+起/倒 (Guīyáng, NE Yunnan); V+tə? (Shanxi)
Wu	V+lə?he/lə?lə/lə?to?
Yue	V + tʂy ²² ; V + kin ⁵⁵ (Siyi dialects)
Hakka	V + tau/tou/tə; V + ten ³¹ /ts'aj ⁵² t'i ⁵² (Méixiān)
Xiang	V + ta/te/得tx; V + 起te'i/倒tau/tou
Gan	V + 倒tau/tou/tʂ
Min	V + le ²¹ (Fúzhōu); V + lo/to/toko/na/nako/paŋ/paŋko (Shàntóu)

[tau]/[təu]/[tə] is also used throughout the Gan dialects as the durative, as well as many Hakka dialects in the form of [tau]/[təu]/[tə]. In Méixiān however, the durative aspect marker is not distinct from the progressive aspect. Most of the Yue dialects use 住 [tʂy²²] as the durative suffix, but the Siyi dialects use the same form as the progressive.

3 SENTENCE STRUCTURE AND WORD ORDER

The structural formula for a sentence and its major components may be summarized as follows, with optional elements in parentheses and obligatory choice of at least one element within braces:

$$S \rightarrow (NP) (TIME (PRT)) VP (Fp, Q, FP)$$

where **TIME** = time words, **PRT** = pause particle

NP → NP' (PRT)

NP' → (DET) N

DET → ({DEM, NUM})* CL

VP → (ADV) (PP (PRT)) V (ASP) (NP) (NP) (COMPL)/(ADV)**

where **VP** = verb phrase, **ADV** = adverb, **PP** = prepositional phrase, **COMPL** = complement

PP → PREP NP

where **PREP** = preposition

The only obligatory category in a sentence is the VP. It is obvious that the NP has a head final structure, the PP has a head initial structure, while the VP has a head middle structure.

The most frequently occurring unmarked word order in a sentence is Subject (S) Verb Object (O). Variation in word order, SOV or OSV, signifies emphasis or contrast. Constituents such as TIME or PP may also be topicalized and occur in initial position of the sentence. However, in the Wu dialects, it is reported that the preferred word order is SOV or OSV—occurring more often than SVO—if O signifies the patient.²¹ On the other hand, in this type of sentence, often a trace of the O is left after the V in the form of a pronoun, which rather suggests that the OV word order is the result of having O topicalized. For example, in (10):

10a. Suzhou, Kunshan: 衣裳 脫 脫 伊

10b. Wuxi, Jiangyin: 衣裳 脫 落 它

10c. Wángjiajīng: 衣裳 脫 其 掉
garment take off 3sg off ‘take off the garment’

Furthermore, in some NW Mandarin dialects, such as those in Qinghai or Gansu, SOV word order does obtain under the influence of neighboring minority languages.

Word order may be associated with meaning. An unmodified, bare noun, or, depending on dialects, a noun modified only by its classifier, occurring in a preverbal position has definite reference while the same occurring in a post-verbal position has indefinite reference. For example, in Cantonese, while 架車 in (11a) refers to any car, the same in (11b) refers to a specific car:

11a. 叫 架 車 [la⁵⁵]
call CL car FP ‘call a car’

11b. 架 車 來 [tʃɔ³⁵] [la]
CL car come PFV Fp ‘the car has come’

There is some structural difference in both the NP and the VP between Northern and Southern Chinese.

The asterisk in the formula for DET refers to the parentheses, which apply to some Southern dialects only. This indicates that while in the Northern dialects the DET consists of a DEM or/and a NUM plus a CL, in the Southern as well as in the Wu (for example, Suzhou, Shanghai, Jinhua, Yongkang, Shuanglin, Changshu, Shengxian, Wenzhou, Yongkang) and the Huizhou (Jixi) dialects the DET may consist of only the CL. As mentioned in the preceding paragraph, the structure CL + N carries definite reference in a preverbal position. For example, in Wenzhou, 枝筆 designates ‘the pen’ in 枝筆不好寫 ‘the pen does not write well.’

The double-asterisked word order for certain adverbs (post-verbal) in the formula for VP is allowed for some Southern dialects only. In other words, all adverbs (except a few

ADVd) in Northern Chinese precede the verb, while some of them may follow the verb in Southern (Min, Yue, Hakka) and Central (Wu, Xiang) as well as certain Huizhou (Jixi, Shèxiān, Túnxi, Xiuning, Yixiān, Qímén, Wùyuán), Jianghuai, SW Mandarin dialects, and even in a few Northern dialects in Hebei (Mǎnchéng), Qinghai, and Inner Mongolia (Huhehote). Post-verbal adverbs are mostly adverbs of quantity (ADVq) such as 多 ‘more,’ 少 ‘less’; adverbs of manner (ADVm) such as 快 ‘fast,’ 慢 ‘slow,’ 好 ‘well,’ 白 ‘in vain’; adverbs of scope (ADVsc) such as 添 ‘in addition,’ ‘merely,’ ‘just,’ ‘all,’ and temporal adverbs (ADVt)²² such as 先 ‘first,’ 快 ‘soon,’ ‘again.’

What should be noted here is that while this V + Adv pattern is prevalent in the dialects mentioned above, its occurrence varies within the same group or even the same dialect. For example, among the Wu dialects, when the Adv is 快, V + Adv is the norm in Yíxíng, Suzhou, Chángshú, Bǎoshan Shuangcǎodun, Nánhuì Zhoupǔ, Songjiang, Húzhōu Shuanglín, Shàoxīng, Yúyáo, and Níngbō; but both V + Adv and Adv + V are used in Jingjiang, Wúxī, Kunshan, Bǎoshan Luódìàn, Shanghai, Wújiāng Lílǐ, Wújiāng Shèngzé, Jiaxīng, Hángzhōu, and Zhují; while Adv + V appears in Liyáng, Jīnhuá, Danyáng, Jiāngyīn, Chángzhōu, Shèngxiān Chónggrén, Shèngxiān Tàipíng, Huángyán, Wenzhou, Qúzhōu, Jīnhuá and Yǒng'an. The V + Adv order probably belongs to the older, native stratum akin with Southern Chinese, the Adv + V order is probably the result of more recent influence from Northern Chinese, while the free use of both types of word order indicate a transitional period where both the old and the new coexist.

4 MAJOR SENTENCE TYPES

The basic word order of a sentence was already given in section 3. There is no difference in word order among declarative, interrogative, or imperative sentences. Occasionally there is variation in the word order of certain constituents such as Measure Complements dependent on whether the sentence is affirmative or negative. This will be discussed in section 4.3 later.

4.1 The copular sentence NP1 + Vcop + NP2

The great majority of the dialects use a form of the verb 是 [ʂɿ⁵¹] as the copula (Vcop), which is derived from the demonstrative of the same form in Old Chinese. Only the Yue and the Hakka dialects employ a different form, 係 ([hɛi²²] in Cantonese).²³

In a simple affirmative copular sentence that involves identity, the copular verb, unmodified by adverbs, is often elided in Northern Chinese, resulting in a structure like NP1 + NP2—for example in SM:

12. 他ta⁵⁵ 北京人peɪ^{214>11}teɪŋ⁵⁵ʐəŋ³⁵
3sg Beijing - person ‘he (is) Pekinese’

But it is far less often elided in the Southern dialects of Yue and Hakka. In a negative copular sentence, however, the copular verb always appears—in other words, *NP1 + NEG + NP2, is ungrammatical.

It is reported that in a few dialects, such as Shùyáng of Jianghuai or Jīnhuá of Wu, the copular verb is rarely used in a copular sentence, affirmative or negative. In the Jīnhuá dialect, the copular verb is used only in emphatic sentences or questions. For example:

13. Shùyáng 那 也 圖書館書
that also library-book ‘that (is) also [a] library book’

14. Jinhuá 這 本 書 弗 我 的
this CL book NEG 1sg NMLZ ‘this book (is) not mine’

Sentences such as these are perfectly grammatical in these dialects but ungrammatical in other dialects, Northern or Southern.

4.2 Existential sentences (TIME) (PLACE) Vex + NP1 / V (PFV/DUR) NP2

Sentences describing existence fall into three major types: those that carry an existential verb such as 有 (Vex, [jōw] in SM) or the copula, those relating to natural phenomena, and those whose verbs are either marked with the durative or the perfective aspect or limited to expressing appearance or disappearance (Table 5.8). They are all characterized by an NP of indefinite reference²⁴ occurring after the verb and signifying the agent or experiencer or natural force. A Time or/and a Place expression may occur before the verb.

The first two types are universally used among the dialects. In the first type, the difference between using Vex and Vcop is that the former denotes simple existence while the latter affirms and identifies existence. This difference is best borne out in the following pair of questions, using SM as example:

15. 裏面li^{213>11}mjen 有jow^{213>11} 什麼səm³⁵mə?
inside exist what
'what is inside?' (not knowing whether there is anything inside)
16. 裏面li^{213>11}mjen 是s̄i⁵¹ 什麼səm³⁵mə?
inside be what
'what is inside?' (knowing something is inside but not knowing what it is)

The second type of existential sentence relating to natural phenomena is also common, although not all sentences describing natural phenomena take the existential form. For example:

- 17a. SM: 下eja⁵¹ 雨uy^{213>11} 了lə
fall rain Fp ‘it is raining now’
- 17b. Cantonese: 落lɔ?² 雨uy²⁴ [la]
fall rain Fp ‘it is raining now’

TABLE 5.8 THREE TYPES OF EXISTENTIAL SENTENCES

Existence	NPloc/Time + Vex + NPindef simple existence with Vex SM: 裏面li ^{213>11} mjen 有jow ^{213>11} 什麼səm ³⁵ mə?	NPloc/Time + Vcop + NPindef existence affirmed with Vcop SM: 裏面li ^{213>11} mjen 是s̄i ⁵¹ 什麼səm ³⁵ mə?
Natural phenomena	NPloc/Time + V + NPindef Cantonese: 落lɔ? ² 雨uy ²⁴ [la]	
Presentative	NPloc/Time + VDUR + NPindef = stative Häiyán: 牆頭浪掛起 (一張)畫[(lo?)ho?]	NPloc/Time + VPfv + NPindef = active Häiyán: 牆頭浪掛了(一張)畫[(lo?)ho?]

Note: indef = indefinite.

The third type, also called the presentative, is not used in the Yue dialects where the same is expressed by using the Vex in combination with a locative expression indicating durative aspect or by using a combination of specific verbs with the Vex. Taking SM and Cantonese as example:

18. SM: 來laj³⁵ 了lə 客人k'ɔŋ⁵¹ zəŋ³⁵
come PFV guest ‘some guest(s) came’
19. Cantonese: 有jɛw²⁴ 人客jen¹¹hak⁴ 來ləi¹¹ [tʃɔŋ³⁵]
Vex guest come PFV ‘some guest(s) came’
20. SM: 牆tɕjan³⁵上ʂan 掛kwa⁵¹ 著tʂə/了lə 一ji⁵⁵幅fu³⁵ 畫兒xwar⁵¹
wall-on hang DUR/ PFV one CL painting
'a picture is hanging / was hung on the wall'
21. Cantonese:
[pon²²]牆tʃøŋ¹¹ 上高søŋ²²kow⁵⁵ 掛kwa⁴⁴[tʃɔŋ³⁵] 有jɛw²⁴ 一jət⁵幅fok⁵ 畫wa³⁵
CL wall above hang PFV Vex one CL painting 'idem'
[pon²²]牆tʃøŋ¹¹ 上高søŋ²²kow⁵⁵ 有jɛw²⁴幅fok⁵ 畫wa³⁵ 掛kwa⁴⁴[hei³⁵]道tow²²
CL wall above Vex CL painting hang Vloc place/prog-asp²⁵
'idem'

In some Wu dialects such as Hǎiyán in Zhejiang, this third type of existential sentence may also take two forms: a stative form which is marked by a locative expression used to indicate the durative aspect and an active form which is marked by the perfective aspect. The former denotes a durative state while the latter indicates the result of some action. For example:

22. 牆頭浪 掛 起 (一張) 畫 [(lo?) ho?] = stative
wall-there hang DUR (one CL) painting (Vloc) place ‘a picture is hanging on the wall’
23. 牆頭浪 掛 了 (一張) 畫 = active
wall-there hang PFV (one CL) painting ‘a picture was hung on the wall’

4.3 Negation

In the use of negative markers, the Northern and the Southern dialects employ different forms, with the Central dialects agreeing either with the Northern or with the Southern group. There are at least three different negative forms: one for simple negation, one or more for aspectual/modal negation, and one for prohibition. Dialects that have more than three negative forms come from the Southern and the Wu group, which register more than one type of modal/aspectual negation.

For simple negation, the Northern and the Central dialects employ forms with a bilabial (labiodental in the Wu dialects) plosive initial [p], the common form being 不 [pu] and its cognates; while the Southern dialects use forms with a bilabial nasal initial [m], the common form being [m] and its cognates.

For prohibition, a complex form which is a combination, and sometimes a contraction, of the simple negative plus an optative verb (Vopt), NEG + Vopt, is generally used. For example, in SM, 別 [pjə³⁵] ‘don’t!’ or 不 [pu^{51>35}] + 要 [jàw⁵¹] /用 [jòŋ⁵³] /准 [tʂwən²¹⁴] /許 [ey²¹⁴] ‘do not!’ / ‘need not!’ / ‘not allowed!’ are used. Among the Central dialects, only Wu follows the Northern pattern, while Xiang and Gan follow the Southern pattern in using prohibitives with the initial [m].

For negating modals/aspect, the great majority of the dialects use a complex form with a bilabial nasal initial which is a combination or contraction of the simple negative and the modal/aspect. In the Northern dialects, 沒 [mei] (< NEG + PFV) and its cognates are used, and in the Central dialects of Xiang and Gan, [mau] (< m + 有) and its cognates are used. In the Southern and the Wu dialects, 無 (a contracted form of NEG + 有) and its variants are used for negating the past tense, a contracted 未 (< NEG + PFV)²⁶ or 未 + 曾 is used for negating the perfective aspect. Examples from the following Southern and Wu dialects illustrate the two distinct forms of negation discussed, which have merged in Northern, Xiang, and Gan:²⁷

24a. Xiàmén (Southern Min)	伊	無	去	:	伊	未	去						
24b. Méixiàn (Hakka)	其	無	去	:	其	未	去						
24c. Cantonese (Yue)	佢	k'øy ²⁴ [mou ²⁴]	去	høy ⁴⁴	:	佢	k'øy ²⁴ 未	mey ²²	去	høy ⁴⁴			
24d. Suzhou (Wu)	伊	[m ^{24>22} pʂɻ ²⁴]	去	:	伊	勿曾	去						
24e. SM (Northern)	他	ta ⁵⁵	沒	mej ³⁵	去	te'y ⁵¹	=	他	ta ⁵⁵	沒	mej ³⁵	去	te'y ⁵¹
24f. Shuangfeng (Xiang)	他	[mʂ ³³]	去	=	他	[mʂ ³³]	去						
24g. Nánchang (Gan)	佢	[mau ²¹]	去	=	佢	[mau ²¹]	去						
	3sg	NEG+Vex	go	:	3sg	NEG+PFV	go						
		‘s/he did not go’		:	‘s/he has not gone yet’								

The negative structure involved appears as NEG + V, but the NEG is actually a complex form of NEG + aspect/past.

For negating other aspects such as the progressive, the durative, or the experiential, the same complex marker is used, but the negative structure is NEG + V + ASP.

Negation in the Min dialects is more complicated. Negative markers are paired with affirmative markers, and negative structures are parallel to and symmetrical with affirmative structures. In Quánzhou (Southern Min), for example, volition is expressed by [be⁴²] and negated with [m³³], possibility/probability is expressed by [ue²¹] and negated with [bue²¹], necessity is expressed by 著 [tio?⁴] and negated with the prohibitive 免 [bian⁴²], the perfective is expressed by 了 [diau⁴²] and negated with 未 [be²²], and existence/affirmation is expressed by 有 [u²²] and negated with 無 [bo²³]. In addition, dependent on verb categories, different negative markers are called for even in the same function. For example, simple negation of Vcop, Vopt and certain cognitive verbs is expressed with [m³³] but simple negation of Vstat is expressed with [bo²³], since [m³³] is indicative of subjective volition and [bo²³] is used for objective, factual negation. In general, the scope of negation includes all elements following the NEG marker. Thus, depending on the location of NEG, different scopes of negation may result. Compare the following in SM:

- | | |
|--|---|
| 25. 你 不 可 以 看 :
ni ^{213>11} pu ⁵¹ k'e ^{213>35} ji ^{213>11} k'an ⁵¹ : | 26. 你 可 以 不 看 :
ni ^{213>11} k'e ^{213>35} ji ^{213>11} pu ^{51>35} k'an ⁵¹ : |
| 2sg NEG maywatch | 2sg may NEGwatch |
| 'you are not allowed to watch' | |
| 27. 你 不 可 以 不 看
ni ^{213>11} pu ⁵¹ k'e ^{213>35} ji ^{213>11} pu ^{51>35} k'an ⁵¹ | : |
| 2sg NEG may | NEG watch |
| 'you cannot but watch' | |

Sometimes negation triggers a difference in word order in the Northern dialects. Measure Complements such as Duration (a time expression indicating a measure or amount of time by which something is completed or to be completed) or Frequency (indicating the number of times an action or motion has taken place or will take place) occur post-verbally in an affirmative sentence but preverbally in a negative sentence. For example in SM:

28. 他t'a⁵⁵病pin⁵¹了lə 三san⁵⁵天tjan⁵⁵了lə
3sg be-sick PFV three day Fp 'he has been sick for three days'
29. 他t'a⁵⁵三san⁵⁵天tjan⁵⁵沒mej³⁵病 pin⁵¹了lə
3sg three day NEG+ PFV be-sick Fp 'he has not been sick for three days'
30. 我wɔ^{214>11}來laj³⁵了lə 三san⁵⁵次ts'i⁵¹了lə
1sg come PFV three time Fp 'I have come three times'
31. 我wɔ^{214>11}三san⁵⁵次ts'i⁵¹沒mej³⁵來laj³⁵了lə
1sg three time NEG+PFV come Fp 'I have not come three times'

In the Southern dialects, except for emphasis, there is no difference in word order for this kind of negative sentence.

The negative potential form will be discussed in section 4.12.2.

4.4 Questions

There are two prominent characteristics of the interrogative forms: the same word order for the declarative sentence obtains and different question particles (Q) occur obligatorily or optionally with different kinds of questions. The major interrogative forms include the Question(Q)-word question, the Yes–No question, the Disjunctive question, and the Neutral question.

4.4.1 Q-word questions

The question particle that may optionally occur with Q-word questions across dialects is [a], which is not a function particle but rather an intonational particle that conveys the tone of voice, mood, and attitude of the speaker, which in the case of [a] carries a softening effect. In some dialects a question particle must occur, for example, in Línxìà (Gansu, Northern), such a kind of question always ends with [za] and in Zhōngníng (Ningxia, Northern) it always ends with 是. The Q-words are as diverse as the deictics. Just take the example of the word for ‘who?’ which is 誰 [sei³⁵] and its cognates in many Northern and even some Yue (such as Siyi) dialects but 哪個, literally ‘which one,’ and its variants in SW Mandarin, Jiānghuái, Xiāng, Gān, and some Yue dialects (such as Cantonese), or 啥人,

literally ‘what person,’ and its variants in the Wu, the Hakka, the Min, and some Yue dialects (such as Yángjiang).

4.4.2 Yes–No questions

The Yes–No question refers to the type that can be answered with ‘yes’ or ‘no.’ It is marked by an obligatory question particle which differs across dialects. In the Northern dialects, it is most commonly [ma] and its variants; in the Yue dialects it is [a²¹] with an extra low falling tone or [me⁵⁵], or [mək²¹] with an extra low falling tone (used in the Siyi dialects). As far as presupposition is concerned, in the Northern dialects, the affirmative Yes–No [ma] question has neutral connotation—the addresser does not express his own evaluation of the situation—while the negative Yes–No question has non-neutral connotation—the addresser expresses his own evaluation of the situation.²⁸ For example, in SM:

32. 你ni^{214>11} 去te'y⁵¹ 嘴ma?
 2sg go Q ‘are you going?’
33. 你ni^{214>11} 不pu 去te'y⁵¹ 嘴ma?
 2sg NEG go Q ‘aren't you going (I thought you're going)?’

However, if the question particle 吧 [pa] (which invites confirmation, such as ‘I suppose’) is used, both the affirmative and the negative Yes–No questions have non-neutral connotation. In the Yue dialects, Yes–No questions are always non-neutral. For example, in Cantonese:

34. 你nei²⁴ 去høy⁴⁴ [me⁵⁵]/[a²¹]?
 2sg go Q ‘are you going (I thought you're not going)?’
35. 你nei²⁴ [m] 去høy⁴⁴ [me⁵⁵]/[a²¹]?
 2sg NEG go Q ‘aren't you going (I thought you're going)?’

A salient feature related to the Yes–No question is that the answer to a negative Yes–No question is in agreement with the truth value of the question itself and not with that of the facts. Thus, the answers to the negative Yes–No question listed above may be the following in SM:

36. 不pù, 我wɔ^{214>11} 去te'y⁵¹ / 是sì⁵¹, 我wɔ^{214>11} 不pu 去te'y⁵¹
 no 1sg go / yes 1sg not go
 ‘No, I am going’ / ‘Yes, I am not going’

4.4.3 Disjunctive question

The Disjunctive question displays a choice of usually two alternatives. Between the two disjuncts, there is usually a conjunction (cnj), while in some dialects a pair of conjunctions may introduce both disjuncts. In the Southern dialects, the preference seems to be for one conjunction between the two disjuncts—in the Wu dialects, the conjunction is 還是 (for example, Suzhou [fiezi]), in the Yue dialects it is 定(係) [tin²²(he)²²] (Cantonese)

or [ma³⁵]_又[jiw³²] (Siyi)—although a pause particle may follow or punctuate the first disjunct. For example, in Suzhou:

37. 倘 是 蘇州 人 呐 還是 溫州 人
?lij zi s̩zutsəi jni̫ nɛ fiezi ?uəntsəi jni̫
3sg be Suzhou- person PRT cnj Wenzhou-person ‘is s/he from Suzhou or from Wenzhou?’

In some dialects, for example, in Zāozhuang (Shandong), no conjunction is used:

38. 你 吃 米飯 吃 煎餅
2sg eat rice eat pancake ‘are you eating rice or pancake?’

4.4.4 Neutral questions

Neutral questions (Table 5.9) refer to those with no presupposition on the part of the addresser. There are three major types: V-not-V, VP-NEG, and ADV-VP. The distribution of these three types has typological significance. The V-not-V type is of Northern origin and is found in many Northern dialects, the VP-NEG type is found in many Southern and Central dialects, while the ADV-VP type is found mainly in a region that now encompasses certain Northern (Anhui, SW Mandarin, Jianghuai) as well as Central (N Wu) and Southern dialects (S Min, southern Jiangxi Hakka). However, the V-not-V type, with the ascendancy of SM as the national language, has become most influential and has spread to every major dialect group. The contemporary scenario displays a complex picture of the V-not-V type co-occurring with other types within one and the same dialect.

The structure of the V-not-V type consists of a complex VP that contains an affirmative VP followed by its negative counterpart. For example, in SM:

39. 你ni^{214>11}去te'y⁵¹不pu去te'y⁵¹
2sg go NEG go ‘are you going?’

When VP contains an object or a complement, there are three possible patterns: VP NEG VP, VP NEG V, and V NEG VP. The full form VP NEG VP is rare and is found only in some Northern dialects in Henan (Huòjia, Wenxiàn), Gansu (Lánzhou), and Shanxi (Yùnchéng), for example:

40. Huòjia 他 應該 來 不 應該 來
3sg should come NEG should come ‘should s/he come?’

The VP-NEG-V pattern is popular in the Northern dialects spoken in Hebei, Shanxi, northern Henan, Shaanxi, Gansu, and Qinghai. For example:

41. Luòyáng (Henan) 你 是 學生 不 是
2sg Vcop student NEG Vcop ‘are you a student?’

This pattern seems to be a reduction from the full form. This observation is based on a historical comparison of Pekinese spoken at the beginning of the twentieth century according to Zhang (1990: 72), and that spoken around the middle of the same century according to Chao (1948). The full form was used in greater frequency than the VP-neg-V

form at the beginning of the century but gradually gave way and was no longer used by the middle of the same century. It should be noted that in all of the dialects that use the full form, the VP-NEG-V pattern is also used.

The V-NEG-VP pattern is by far the most popular. It seems to have originated in the south and is prevalent in SW Mandarin, especially Hubei and Sichuan, but also in Shandong (Jiaodong Peninsula) and Manchuria. For example, in Èzhou (Hubei):

42. 吃冒 吃 飯 啊
eat NEG-PFV eat rice Q ‘have/did (you) eaten/eat?’

This pattern has by now been accepted by the great majority of the Southern dialects of Hakka, Yue, and Min, and even Pekinese and other Northern dialects. It has developed an abbreviated form VV(O) as a result of the contraction of NEG with V or of the ellipsis of NEG, the former is found for example in Fúzhōu (Min), Liánchéng (Hakka) and the latter in some Wu dialects such as Shàoxīng, Zhūjī, Wǔyí, Jīnhuá, Shèngxiān, in Shùndé (Yue) and in Jiānyáng (N Inland Min). For example in Jiānyáng, we have: 今 *kin⁵²* 朝 *tio⁵²* 你 *nɔi⁴²* 去 *kho^{22>55}* 去 *kho²²* ‘are you going today?’ where the change of tone from low to high for the verb ‘to go’ implies the contraction of the simple negative [*oŋ⁴⁴*] and the verb [*kho²²*], yielding [*kho^{22>55}*].

The VP-NEG type has a long recorded history since at least Qin times²⁹ and is native to many Southern and Central dialects as well as the peripheral areas (Shandong, NW Mandarin dialects of Shanxi, Shaanxi, Gansu, Qinghai) of the Northern dialects. The NEG takes different forms in different dialects depending on whether the VP contains ASP. In the Northern dialects, it is often 不 or 沒 (if the VP is marked with ASP). In the Gan dialects it is mostly [mo]/[mə] or [po]/[pɔ] and in the Hakka dialects it is 無 [mɔ/[mɔw]. In the Yue dialects, it is mostly 嗎 [ma] (which is derived from a contraction of the general negative marker [m̩] plus the final particle [a]) or 未 [mej²²] (NEG for the perfective) / [mian¹¹] (a contracted form of 未曾 in, for example, the Siyi dialects), for example:

43. Kaipíng (Yue) 佢 講 得 現 嗎
3sg speak COMPL-mkr clear NEG+Q ‘can s/he speak clearly?’

In the Wu dialects, NEG is mostly 勿 [vɛ?] or 伐 [va] irrespective of whether VP contains ASP (in some dialects, NEG is [vəŋ] (< 勿 + 曾) if VP contains ASP). For example:

44. Jiaxing (Wu) 伊?i 來leE 哩li 伐va
3sg come PFV NEG+Q ‘has s/he come?’

In the S Min dialects, the VP-NEG question sometimes takes the form of VP-a-NEG, namely, with an optional particle occurring between VP and NEG. The NEG may be realized as one of four forms depending on the type of verb and on aspectual/modal types used in the VP: [m̩] (non-aspectual), [bo] (past), [be] (perfective), or [bue] (probability/possibility) in, for example, Quánzhōu. For example:³⁰

45. Quánzhōu 你 明 [na²⁴] [bɔ?̩⁴] 來 (阿 [a²⁴]) [m̩²¹]
2sg tomorrow will come (PRT) NEG ‘are you coming tomorrow?’

46. Xiàmén 有u 芳p'an 無bo
af-asp fragrant NEG ‘is it fragrant?’

The ADV-VP type, prevalent along the southeastern coast of China,³¹ consists of an interrogative marker derived from an adverb. What is special about this marker is that it occurs before the VP and after the subject NP if it appears. This ADV marker has two popular phonetic shapes, marked by either a guttural (including the so-called zero) initial in such forms as [a?], [a], [ɑ?], [ɑ], [ə?], [ɪ?], [hə?] or a velar initial in such forms as [kə?], [k'ə?], [kə], [k'ə?], [ka], [ki?], [ku], [kou], [ku], [kxu], [xa?], [xa], [xə?], [xa], [xæ], [xe], [xə?]. The question may also take an optional question particle at the end. For example:

47. Suzhou (Wu) 俚?lij 阿?a? 曉得 eiætə? [tɕiŋ] 3sg ADV know Q ‘does s/he know?’

In some dialects, the ADV is followed by a negative marker. For example, in Lóngnán (Jiangxi Hakka) the ADV is [æn⁵⁵] followed by the negative marker 不:

- 48a. 你 的 瓜 [æn⁵⁵] 不 甜 2sg ATTR melon ADV NEG sweet ‘is your melon sweet?’

In Mùping, Píngdù, and Róngchéng (Shandong) the ADV is 是 or 可 (only in Mùping) followed by the negative marker 不 or 没 (the latter is possible only with 是), for example:

- 48b. Mùping 可 不 來 ADV NEG come ‘(are you) coming?’

- 48c. Mùping 飯 是 没 熟 rice ADV NEG+PFV ripe ‘is the rice done?’

In Róngchéng, the ADV can be elided, leaving only NEG, for example:

- 48d. 飯兒 (是) 没 涼 rice ADV NEG+PFV cool ‘has the rice gone cold?’

This ADV + NEG combination suggests the attraction of the NEG to the ADV, displaying a hybrid of VP-NEG and ADV-VP, namely, ADV-VP-NEG > ADV-NEG-VP. This can explain the ADV [kam] in Yílán (Taiwan Min), which is probably the result of merging the ADV [ka] with the NEG [m].

Since the modern dialects are generally structured with several strata—native, literary, aboriginal, borrowed—it is not surprising to find all three types of neutral question forms within one and the same dialect. In addition, there are hybrid forms as a result of the merging of usually a native form with a borrowed form indicating a transitional period when a new form is being absorbed before the old form is discarded. Examples from the Shàntóu dialect (Min) illustrate this complex usage:

49. (i) 伊 是 [nia³³] 弟 (啊) [mi³⁵]
(ii) 伊 [k'a?²] 是 [nia³³] 弟
(iii) 伊 [k'a?²] 是 [nia³³] 弟 (啊) [mi³⁵]
3sg ADV be 2gen younger-brother (cnj) NEG ‘is he your younger brother?’

- (iv) 伊 是 [m³⁵] 是 [nia³³] 弟
 3sg be NEG be 2gen younger-brother ‘idem’
- (v) 伊 [k'a?²] 好 來 (啊) [m³⁵] 好
 3sg ADV willing come PRT NEG willing ‘is s/he willing to come?’

While (i) is a VP-NEG question, (ii) is an ADV-VP question, (iii) is a hybrid of these two, (iv) is a V-NEG-VP type borrowed from SM and (v) is a hybrid of (ii) and (iv). Types (i), (ii), and (iii) are equally popular in usage but (iv) and (v) are far less frequently used.

4.5 The double-object construction

Two major word order types of the double-object construction signify a typological difference between the Northern and the Southern dialects. The Northern dialects have the indirect object (O_i) preceding the direct object (O_d)³² while the Southern dialects as well as a number of Jianghuai (Jiangsu dialect of Huáiyīn, Anhui dialects of Tóngchéng, Anqing, Wúhū; Hubei dialects of Huánggāng, Huángpí, Xiāogǎn, Luótán, Yingshan, Máchéng, Yingshan,

TABLE 5.9 NEUTRAL QUESTIONS

Type	Subtype	Distribution	
V-not-V	V-NEG-V	Northern	SM: 你ni ²¹⁴⁻¹¹ 去te'y ⁵¹ 不pu去te'y ⁵¹
	VPNEGVP	N (Henan), NW (Gansu, Shanxi)	Huòjia (Henan): 他應該來不應該來
	VPNEG V	N (Hebei, N Henan, Qinghai), NW (Shanxi), Shaanxi, Gansu)	Luòyáng (Henan): 你是學生不是?
	VNEGVP	SW (Hubei, Sichuan), N (Shandong), NE	Èzhou (Hubei): 吃冒吃飯啊?
	VV(O)	Wu, Hakka, Yue (Shùndé), Min	Jiànyáng (Min): 今kin ⁵² 朝tio ⁵² 你no ⁴² 去khɔ ²²⁻⁵⁵ 去khɔ ²² ?
VP-neg		Southern, Central, NW, N (Shandong)	NEG=不, 沒(N); mo/mɔ, po/pɔ (Gan); 無mɔ/mɔw (Hakka); 嗎ma, 未me ²² (Yue), mian ¹¹ (Siyi); 勿 [ve?]/伐[və] (Wu); [m] (non-aspectual), [bo] (past), [be] (perfective), [bue] (probability/possibility) (Min) Kaiping (Yue-Siyi): 佢講得現嗎 Jiaxing (Wu): 伊zì來leè哩li伐va S Min: Quánzhou: 你明[na ²⁴][bɔ ⁴]來[m ²¹] Xiāmén: 有u芳'ap無bo
ADV-VP	VP-a-NEG	S Min, Central	Quánzhou: 你明[na ²⁴] [bɔ ⁴]來阿[a ²⁴] [m ²¹]
		N-Anhui, Jianghuai, SW Mandarin, N Wu, S Min	ADV=[a?], [a], [a?], [a], [ə?], [i?], [hə?]; [kə?], [k'ə?], [ka], [k'a?], [ka], [ki?], [ko], [kuw], [k'uw], [xa?], [xa], [xa?], [xa], [xæ], [xe], [xə?]. Suzhou (Wu): 倌?lij 阿?a? 曉得ciætə? [teɪd]?
ADV-VP-NEG (hybrid)		Jiangxi Hakka, N (Shandong), S Min	Lóngnán (Hakka): 你的瓜[xən ⁵⁵]不甜? Mùpíng (Shandong): 可不來? 飯是沒熟? Róngchéng (Shandong): 飯兒(是)沒涼? Shàntóu (S Min): 伊[k'a?²]是[nia ³³]弟(啊) [mi ³⁵]?

Note: N = Northern Mandarin, NE = Northeastern Mandarin, NW = Northwestern Mandarin, SW = Southwestern Mandarin.

Suíxiànl, Líshan, Huáng’ān, Anlù, Yìngchéng, Yúnmèng) and SW Mandarin (Hubei dialects of Enshi, Badong, Dangyáng, Jingmén, Jianglíng, Yídu, Hánkǒu, Hányáng, Tianmén, Jingshan), and Central dialects have the direct object preceding the indirect object³³ except when the double-object verb (Vdo) has the inherent feature of [+deprive] (verbs such as ‘to steal,’ ‘to rob,’ ‘to cheat,’ ‘to borrow,’ ‘to buy,’ ‘to win,’ ‘to deduct’). For example, when Vdo has the inherent feature of [+give] (50), and when Vdo has the feature [+deprive] (51):

- | | | | | | | | |
|----------------------|------------------------------|--------------------------|-------------------------|--------------------|--------------------------|--------------------------|------------------|
| 50a. SM | 我 | 給 | 你 | 一 | 本 | 書 | |
| | wɔ ^{213>35} | kej ^{213>35} | ni ^{213>11} | ji ⁵⁵ | pən ^{213>11} | su ⁵⁵ | |
| 50b. Cantonese | ŋɔ ²⁴ | 畀 | pej ³⁵ | | pun ³⁵ | sy ⁵⁵ | |
| 50c. Méixiàn (Hakka) | ŋai | 分 | pun | | su | 分 | |
| 50d. Xinyú (Gan) | ŋɔ | 把 | pa | | pən | su | |
| | 1sg | give | 2sg | one | CL | book | |
| | give 2sg ‘I give you a book’ | | | | | | |
| 51a. SM | 我 | 偷 | 了 | 他 | 一 | 本 | 書 |
| | wɔ ^{213>11} | t'ow ⁵⁵ | lə | t'a ⁵⁵ | ji ⁵⁵ | pən ^{213>11} | su ⁵⁵ |
| 51b. Cantonese | ŋɔ ²⁴ | t'rw ⁵⁵ | tʃɔ ³⁵ | k'øy ²⁴ | jet ⁵ | pun ³⁵ | sy ⁵⁵ |
| | 1sg | steal | PFV | 3sg | one | CL | book |
| | ‘I stole a book from him’ | | | | | | |

As a result of language contact, many Southern and Central dialects allow both types of word order, with the Oi + Od type borrowed from SM. For example:

52. Wǔhàn 他 把 一 本 書 把 得 我 / 他 把 我 一 本 書
(SW Mandarin) 3sg give one CL book give 1sg / 3sg give 1sg one CL book
's/he gave me a book'
53. Xiangxiang (Xiang) 拿 我 錢 / 拿 錢 我
give 1sg money / give money 1sg ‘give me money’

Across the Wu dialects, this is particularly true:

54. Suzhou 我 撥 本 書 傑 / 我 撥 傑 本 書
ŋeu pə? pən sq ne / ŋeu pə? ne pən sq
1sg give CL book 2sg / 1sg give 2sg CL book ‘I gave you a book’

4.6 The disposal form

The so-called disposal form is a construction in which the object NP is fronted before the verb by means of using a disposal marker (BA) under certain conditions. It has the structure of NP1 + BA (把 or [pa²¹³] in SM) + NP2 + VP. Semantically it emphasizes what NP1 has done or will do to NP2. It is widely used in the Northern dialects, but seldom used in the Southern dialects. Conditions for this construction vary across dialects. For SM and most of the Northern dialects, there are two major conditions: NP2 must have definite or generic reference and VP must be specific. While the first two of the following examples fulfill these two conditions and are grammatical in SM, the last two are ungrammatical since the NP2 ('several apples') in (57) is indefinite and the VP of (58) is not specific:

55. 他ta⁵⁵ 把pa^{213>11} 那na⁵¹ 個kə 蘋果p'iŋ³⁵kwɔ 吃tʂ⁷ 了lə
s/he BA that CL apple eat Fp ‘s/he ate that apple’

56. 他^{ta⁵⁵} 把^{pa^{213>11}} 正經事^{tʂəŋ⁵¹tʂŋ⁵⁵ʂi⁵¹} 當作^{tʂəŋ⁵¹tswɔ⁵¹} 兒戲^{ə³⁵ɛi⁵¹}
3sg BA serious matter regard child-play
's/he regards serious matters as trifling matters'
57. *他^{ta⁵⁵} 把^{pa^{213>35}} 幾^{lei^{213>11}}個^{kə} 蘋果^{p'iŋ³⁵kwo} 吃^{tʂ'i⁵⁵}了^{lə}
3sg BA several CL apple eat Fp 's/he ate several apples'
58. *他^{ta⁵⁵} 把^{pa^{213>11}} 那^{na⁵¹} 個^{kə} 蘋果^{p'iŋ³⁵kwo} 吃^{tʂ'i⁵⁵}
3sg BA that CL apple eat 's/he eats that apple'

In some Northern (Cháoxiān of Anhui, Yingshan of Hubei) and Wu (Shanghai) dialects, NP2 can leave a trace in a post-verbal position, namely, the structure can be NP1 + BA + NP2 + V + PN; for example, in Cháoxiān:

59. 把 衣服 洗 乾淨 它
BA garment wash clean PN 'wash the garment clean'

This is why it is reasonable to consider NP2 as being fronted by means of BA to a preverbal position.

Further restrictions on the VP in this construction include non-tolerance of the potential or the negative, both of which must occur before BA. However, these restrictions do not apply to all dialects. For example, the Chángsha dialect (Xiang) allows the potential form in the VP (60), and the NW Mandarin Lánzhōu (Gansu) as well as the Wèinán (Shaanxi) dialect tolerate the negative in the VP (61):

60. 我 把 鎖 打 得 開
1sg BA lock hit can open 'I can open the lock'
61. Wèinán 把 門 不 要 老 閉 著
BA door NEG need always closed DUR 'do not keep the door closed all the time'

Depending on the dialect, not all verbs that take objects can occur in the disposal form. In SM, non-manipulative verbs such as the copula and classificatory verbs, verbs of cognition and of perception, etc. do not occur in this structure. However, this rule does not apply in dialects such as Lánzhōu. For example, the following sentences are ungrammatical in Standard Mandarin but perfectly grammatical in Lánzhōu:

62. 我 把 他們 的 名字 知道
1sg BA 3pl ATTR name know 'I know their names'

The disposal marker differs among dialects. While [pa] and its variants are used in most of the Northern dialects, other markers occur: 紿 (Northern—Luòyáng of Henan; Jiaochéng, Hándān of Shanxi), [pɔ] and its variants (Northern—Línfén of Shanxi, some Xiang and many Wu dialects), [ma] (Northern—Qīngjiàn of Shaanxi, Huángméi of Hubei), 幫 (many SW Mandarin dialects in Yunnan, some Huizhou and Wu and Min dialects), 挨 (many SW Mandarin dialects in Yunnan), [na] and its variants (some Gan, Xiang, Wu, Min dialects), 將 (Yue, Hakka, Min), etc.

4.7 The so-called passive

There is no equivalent to what is generally described as the passive construction in the Indo-European languages, which is distinctively marked with a certain grammatical structure with the patient as the subject. There are several types of construction in Chinese that feature the patient as subject, apart from those with the patient topicalized. There are two major types of this kind of construction that are commonly mislabelled the passive construction in Chinese. One type has the general structure of NP_{Patient} + BEI (+ NP_{Agent}) + VP, characterized by the option of having the agent expressed or unexpressed. This type occurs only in Northern Chinese, where the marker (BEI) is [pej⁵¹] or [kej²¹³] in SM. This marker is often mis-analyzed as an agent marker. The fact that it can occur without an expressed agent argues against this interpretation. For example in SM:

63. 我wɔ^{213>11} 被pej⁵¹ 騙p'ien⁵¹ 了lə
1sg BEI deceive Fp ‘I was deceived’

Therefore [pej⁵¹] or [kej²¹³] is best understood as a marker of infliction. This also accords well with the historical development of the patient-as-subject construction marked with 被 [pej⁵¹] in which the agentless form predates the same with the agent appearing.

The other type of patient-as-subject construction has the general structure of NP_{Patient} + BEI + NP_{Agent} + VP, characterized by the obligatory presence of the agent. This is typical in Southern Chinese, where an unspecified agent cannot be elided but must be overtly expressed with an indefinite noun such as ‘someone’ or ‘people.’ The above example in Cantonese can only be:

- 64a. 我ŋɔ²⁴ 罷pej³⁵ 人jen¹¹ [ŋak⁵] [tʃɔ³⁵] (我ŋɔ²⁴)³⁴
1sg BEI people deceive PFV 1sg ‘I was deceived (by someone)’

Here the BEI marker may be regarded as the agent marker. This second type of patient-as-subject construction also exists in Northern Chinese. In fact within one and the same dialect both types may occur. In SM, the second type is marked with 讓 [zəŋ⁵¹] or 叫 [tejaw⁵¹] or 教 [tejaw⁵⁵] and is popular in colloquial speech. Its equivalent to (63) is:

- 64b. 我wɔ^{213>11} 讓zəŋ⁵¹/叫tejaw⁵¹ 人zəŋ³⁵ 騙p'ien⁵¹ 了lə
1sg BEI people deceive Fp ‘I was deceived (by someone)’

A prominent characteristic of these two types of patient-as-subject construction is its association with undesirable events, especially in colloquial speech. In some dialects, only unfavorable events can be expressed with this type of construction, while favorable events must be expressed in the active. For example in the Yantai dialect (Shandong), while (65a) is grammatical, (65b) is not grammatical and must be rendered as (65c):

- 65a. 他 he 叫 he 婆婆 pəpə 打打了
3sg BEI 3sg wife hit Fp ‘he was hit by his wife’

- 65b. *他 he 叫 he 婆婆 pəpə 誇了
3sg BEI 3sg wife praise Fp ‘he was praised by his wife’

- 65c. 他 he 婆婆 pəpə 誇他了
3sg wife praise 3sg Fp ‘his wife praised him’

When a verb which is neutral in connotation occurs in this kind of sentence, it takes on unfavorable meaning. For example in SM, the verb ‘to hear’ has no undesirable meaning in (66a), but carries unfavorable connotation in (66b):

- 66a. 我wɔ^{213>11} 說swɔ⁵⁵ 的tə 話xwa⁵¹ 他t'a⁵⁵ 聽見t'in⁵⁵təjen 了lə
 1sg say attr mkr word 3sg hear Fp
 ‘what I said was overheard by him/her’
- 66b. 我wɔ^{213>11} 說swɔ⁵⁵ 的tə 話xwa⁵¹ 讓zəŋ⁵¹ 他t'a⁵⁵ 聽見t'in⁵⁵təjen 了lə
 1sg say attr mkr word BEI 3sg hear Fp
 ‘what I said was (unfortunately) overheard by him/her’

On the other hand, there are cases, albeit small in number, where this form is used for favorable meaning, as in the following two examples in SM:

67. 這tʂə⁵¹ 孩子xaj³⁵tsı 常常tʂ'anj³⁵tʂ'anj 被bej⁵¹老師law^{213>11}ʂi⁵⁵ 稱讚tʂ'əŋ⁵⁵tsan⁵¹
 this child often BEI teacher praise
 ‘this child is often praised by the teacher’
68. 他t'a⁵⁵ 被pej⁵¹選cəŋen^{213>11} 為wej³⁵ 總統tsoŋ^{213>35}t'ɔŋ²¹³
 3sg BEI elect be president ‘s/he was elected president’

It is often observed that the broadened usage of the said construction and the generalization of the role of the patient into the recipient are brought about by contact with Western languages where no semantic constraint is imposed on this construction. While the scope of the said construction has certainly widened, it should not be forgotten that even in ancient times certain verbs such as 寵 ‘to favor’ have always been freely used in this construction.³⁵

An interesting aspect that relates the Chinese dialects to their neighboring languages is the identity of the marker of infliction with the verb ‘to give’ and/or with the causative marker. In the Northern dialects, the markers 叫, 讓, and 教 also function as causative markers. For example, in SM, 讓他打了 means either ‘was hit by him/her’ or ‘let him/her hit’ and 叫/教他打了 means either ‘was hit by him/her’ or ‘asked him/her to hit,’ ‘caused him/her to hit.’ Dialects which use 叫 for both functions include many Anhui dialects (Northern—Bàngbù, Língbi, Sùxiàn, Suixi, Dangshan, Fùyáng, Fùnán, Línquán, Jièshōu, Tàihé, Háoxiàn, Woyáng, Fèngtái, Shòuxiàn, Yǐngshàng, Huòqiu, Jinzhài) and Jianghuai dialects (Wúwéi, Lújiang, Shuchéng, Huáinán, Huáiyuǎn, Dìngyuǎn, Jiashan, Lái’ān, Tongchéng). 讓 is used in both functions in Jianghuai (Mǎ’anshan), Xiang (Nánlíng), and Gan (Wàngjiāng, Tàihú, Qiánshan, Yuèxi, Sùsóng) dialects. For these double functions there are other markers such as 得 or [dei] (Línfén in Shanxi), [tʂo] (Jinchéng in Shanxi) or [tʂau] (Dìngxi, Tongwèi, Lóngdié of Gansu, Northern) and in SW Mandarin [tsau] (Chéngdu) or [tʂuo] (Kunmíng) or [tʂok] (Nánníng of Guangxi).

In the Central dialects, the Wu, some Xiang (Línwǔ) and Gan dialects, the said marker is also the verb ‘to give.’ Its form varies across these dialects as we have seen in section 4.5. In some Northern dialects located geographically close to the Southern dialects or in the periphery of the Northern dialects, the same phenomenon can be observed. Such markers with these dual functions include [te] (Xiníng in Qinghai), or in Jianghuai [te] (Rúgao in Jiangsu) or [ha] (Nántong in Jiangsu) and 紿 in Jianghuai (Héféi, Féidong, Féixi, Hánshan, Héxiàn, Lújiang, Huòshan, Huáinán, Dingyuǎn, Chúxiàn, Quánjiào, Tongchéng), Huizhou (Shèxiàn, Jixi, Túnxi, Xiùníng, Qímen, Níngguó), and Gan (Huáiníng, Sùsóng). In some dialects, the different usage may correlate with different pronunciation: for example [kɛ¹¹]

as infilctive marker but [kə?⁴] as the verb ‘to give’ in the Huáiyin dialect (Jianghuai). In a number of Northern dialects, one and the same marker 紿 (Jiaochéng of Shanxi) or 把 (Xishūi of Hubei, Yángzhou of Jiangsu, Lángxi of Anhui, also the Gan dialect of Wàngjiang in Anhui), serves as the verb ‘to give,’ the infilctive marker, and the disposal marker.

Some Southern dialects, such as Yue and Hakka, combine both Northern and Southern features of using the same form to serve triple functions: as the verb ‘to give,’ the infilctive marker, and the causative marker. For example in Cantonese:

- 69a. 佢'øy²⁴ [m¹¹] 界pej³⁵ 我ŋɔ²⁴ 界pej³⁵ 錢tʃin³⁵ 你nej²⁴
3sg NEG let 1sg give money 2sg
's/he does not let me give you money'
- 69b. 我ŋɔ²⁴ [mou²⁴] 界pej³⁵ 佢'øy²⁴ 打ta³⁵ 我ŋɔ²⁴
1sg NEG+PFV BEI/let 3sg hit 1sg
'I was not hit by him/her' or 'I did not let him/her hit me'

There is certainly a semantic relationship among these three functions: ‘to give’ is ‘to cause to receive’ (causative) and ‘to receive’ (BEI) is ‘to cause to give’ (causative). Furthermore, the same marker in the Yue dialects serves as the instrumental marker too. For example in the Siyi dialect of Táishan (Yue), where the verb ‘to give’ is [ʔei⁵⁵].³⁶

70. ʔei⁵⁵ 心[ɿim³³]機[kei³³]讀[uk³²]書[ji³³]
give/use heart study ‘study hard’

In sum, there are four syntactic constructions with the verb [give] and its derivatives sharing the semantic features [causative] and [receive]: the double-object construction with the features [+causative, +receive, +Od] for its main verb, the permissive structure with the features [+causative, +receive, +VP] for its main verb, the BEI construction with the features [-causative, +receive, +VP] for its main verb, and the instrumental structure with the features [+causative, -receive, +VP] for its main verb. Both the permissive and the instrumental are of the VP series structure (see section 4.14). The interrelationship of these four constructions in terms of semantic features and syntactic feature is shown in Figure 5.3.

4.8 The double-subject construction NP1 + NP2 + VP

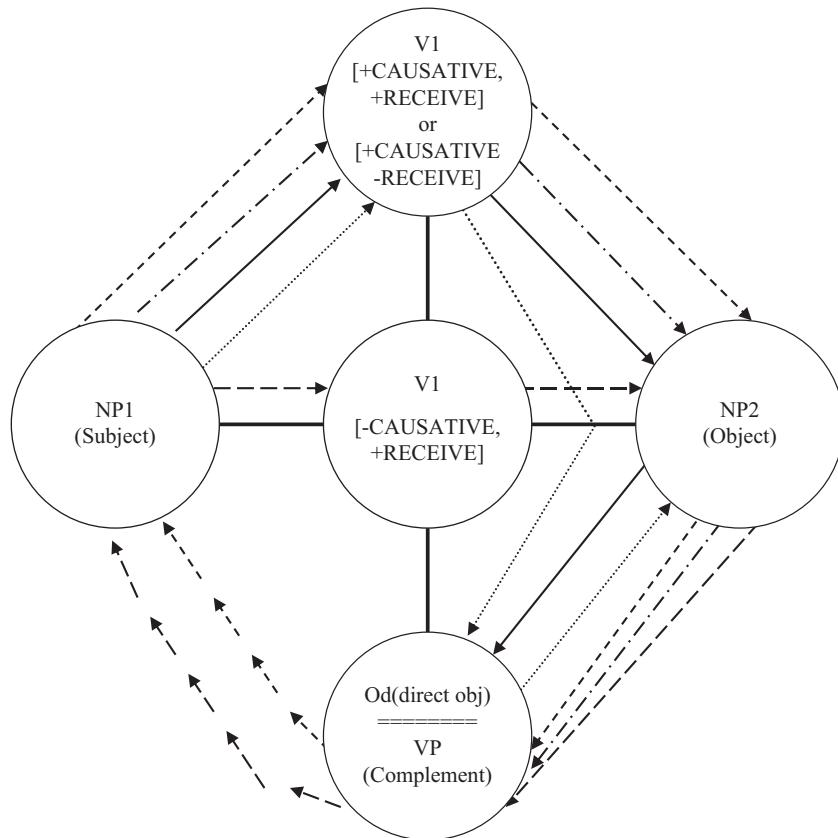
Sentences with two NPs in a sequence initially and before the main verb are described as the double-subject construction if neither of the NPs is topicalized (Table 5.10). The semantic as well as syntactic relationship between these two NPs is diverse. It may fall into roughly four types, relating to body part, the genitive, partitive, or property. For example:

71. SM 象ejaj⁵¹ 鼻子pi³⁵zi 長tʂ'arj³⁵
elephant nose long ‘the elephant has a long nose’
72. SM 那na⁵¹ 個kə 孩子xaj³⁵zi 爸爸pa⁵¹pa 是s'i⁵¹ 大夫taj⁵¹fu
that CL child father be doctor ‘that child has a doctor father’
73. Cantonese
中國tʂɔŋ⁵⁵kwɔk⁴⁴ 人口jen¹¹hew³⁵ 多tʂ⁵⁵
China population numerous ‘China has a big population’

74. SM 三san⁵⁵個kə 孩子xaj³⁵z1 兩ljan^{213>11}個kə 病pin⁵¹ 了lə
 three CL child two CL sick FP
 'two of three children became sick'

Several characteristics can be observed with this kind of construction, in which the relationship between NP1 and NP2 is not one of attribution. Namely, 象ejan⁵¹ 鼻子pi³⁵z1 長tʂ'an³⁵ is not the same in structure as 象ejan⁵¹ 的tə 鼻子pi³⁵z1 長tʂ'an³⁵ 'the trunk of the

- GIVE = V1[+CAUSATIVE, +RECEIVE, +Od] = double-object
- ALLOW = V1[+CAUSATIVE, +RECEIVE, +VP] = permissive
- SUFFER = V1[-CAUSATIVE, +RECEIVE, +VP] = inflictive (> agent marker)
- USE = V1[+CAUSATIVE, -RECEIVE, +VP] = instrumental



- Double-object construction (S V Oi Od) – Mandarin (V=GEI); Wu, Xiang, Gan, Hakka, Min
- Double-object construction (S V Od Oi) – Yue (V=PEI 紣), Taishan (V=[?ei]/[?i])

FIGURE 5.3 INTERRELATIONSHIP OF THE FOUR CONSTRUCTIONS WITH 'GIVE' AS THE MAIN VERB

TABLE 5.10 FOUR TYPES OF DOUBLE-SUBJECT CONSTRUCTION—NP1 (PAUSE) NP2 VP

Body part	Genitive	Partitive	Property
SM: 象ejan ⁵¹ 鼻子 pi ³⁵ zi 長tʂ'anj ³⁵ SM: 那na ⁵¹ 個kə 孩子 xaj ³⁵ zi 呀ja 爸爸pa ⁵¹ pa 是ʂ'i ⁵¹ 大夫taj ⁵¹ fu	SM: 那na ⁵¹ 個kə 孩子 xaj ³⁵ zi 呀ja 爸爸pa ⁵¹ pa 是ʂ'i ⁵¹ 大夫taj ⁵¹ fu	SM: 三san ⁵⁵ 個kə 孩子 xaj ³⁵ zi 兩ljan ^{213>11} 個kə 病piŋ ⁵¹ 了lə 病piŋ ⁵¹ 了lə	Cantonese: 中國tſonj ⁵⁵ kwɔk ⁴⁴ 的確tik ⁵ k'ɔk ⁴⁴ 人口jen ¹¹ hew ³⁵ 多to. ⁵⁵ China certainly population numerous 'China certainly has a big population'

elephant is long.' There are at least three differences. Often an adverb can occur between the NPs, for example:

- 73a. Cantonese 中國tſonj⁵⁵ kwɔk⁴⁴ 的確tik⁵k'ɔk⁴⁴ 人口jen¹¹hew³⁵ 多to.⁵⁵
China certainly population numerous
'China certainly has a big population'

A pause particle may occur between them:

- 72a. SM 那na⁵¹個kə 孩子xaj³⁵zi 呀ja 爸爸pa⁵¹pa 是ʂ'i⁵¹ 大夫taj⁵¹fu
that CL child PRT father be doctor
'that child, has a doctor father'

NP2 + VP can serve as a clause modifying NP1 except for the partitive type exemplified by (74):

- 71a. SM 鼻子pi³⁵zi 長tʂ'anj³⁵ 的tə 象ejan⁵¹
nose long attr mkr elephant 'elephants which have long trunks'

The partitive type, on the other hand, may be understood as expressing the meaning 'among...', that is, (74) may be interpreted as follows:

- 74a. 三san⁵⁵個kə 孩子xaj³⁵zi 裡頭li^{213>11}tow 兩ljan^{213>11}個kə 病piŋ⁵¹ 了lə
three CL child inside two CL sick FP
'among three children two became sick'

4.9 The comparative construction

4.9.1 The comparative degree

The type of comparative structure that displays typological significance among the dialects is the comparative construction of the comparative degree (Table 5.11). The major difference between the Northern and Central (but including N Min and most Hakka) on the one hand and the Southern (but including some Shandong) dialects on the other lies in word order: in the former the compared constituents precede the verb of comparison (Vcomp), while in the latter they flank the Vcomp. The divergence in form of the comparative marker (CPR) is but a lexical matter. The basic Northern structure may be represented as NP1 + CPR + NP2 + (ADV) + Vcomp + (MP) and that of the South as NP1 + Vcomp + CPR + NP2 + (MP) or NP1 + (ADV) + Vcomp + CPR + NP2 + (MP), where MP stands for Measure Complement. The ADV is limited to those with the meaning of 'even.' The CPR is most commonly 比 ([pi²¹³] in SM) in Northern Chinese but varies in Southern Chinese: [k'a?] in

S Min, 過 [kwɔ⁴⁴] in Yue, 起 in Shandong (Mùpíng, Zhuchéng, Píngdù, Wéifang, Zibó). The following examples show the contrast in word order between the North and the South:

75a. SM	他t'a ⁵⁵	比pi ^{213>35}	我wɔ ^{213>11}	高kaw ⁵⁵
75b. Xiàmén (Min)	伊i	k'a?	lo	我gwa
75c. Cantonese	k'oy ²⁴		高kow ⁵⁵	過kwɔ ⁴⁴ 我ŋɔ ²⁴
75d. Mùpíng	他		高	起 我
	3sg CPR	1sg	be-tall CPR	1sg
				‘s/he is taller than me’

The Southern word order reflects an earlier historical pattern. However, this earlier pattern also survives across dialects in a few examples with high frequency Vcomp such as ‘to be tall,’ ‘to be young,’ ‘to be old,’ ‘to be heavy,’ with MP, and without any CPR (or with CPR elided):

76a. SM	他t'a ⁵⁵	高kaw ⁵⁵	我wɔ ^{213>11}	—ji ⁵⁵	個kə	頭t'ow ³⁵
	3sg	be-tall	1sg	one	CL	head

76b. Xiangxiang	我	大	你	三	歲
	1sg	old	2sg	three	year

On the other hand, the Northern pattern is also making its way into various Southern dialects, especially those of big cities, such as Cantonese or Xiàmén, sometimes resulting in hybrid patterns such as NP1 CPR1 NP2 CPR2 Vcomp where CPR1 adopts the Northern marker and CPR2 retains the Southern one.

Apart from Vstat, Vtrq (such as ‘to like,’ ‘to love,’ ‘to hate,’ ‘to hope’) and Vopt (such as ‘should,’ ‘to dare,’ ‘to be willing’) also serve as Vcomp. Furthermore, adverbs derived from Vstat (such as ‘early,’ ‘late,’ ‘first,’ ‘last,’ ‘more,’ ‘less,’ ‘frequent’) may also occur in the comparative construction.

A common characteristic of the Chinese comparative construction, be it in the comparative degree or the positive degree, is the restriction that all compared constituents must be the subject or the topic of the sentence. Comparison of the object, the predicate, or the sentence is expressed by (a) topicalizing these constituents or by (b) rendering the comparison in a two-clause sentence. An expression such as ‘I have more books than you’ is given with alternative (77), while Socrates’ famous saying must be given with alternative (78):

77. SM	我wɔ ^{213>11}	書ʂu ⁵⁵	比pi ^{213>35}	你ni ^{213>11}	多twɔ ⁵⁵
	1sg	book	CPR	2sg	numerous

78. SM	我wɔ ^{213>11}	愛aj ⁵¹	我wɔ ^{213>11}	的tə	老師law ^{213>11}	si ⁵⁵	可k'ə ¹¹	我wɔ ^{213>11}	更kxŋ ⁵¹
	1sg	love	1sg	ATTR	teacher,		but	1sg	even
	愛aj ⁵¹	真理tʂən ⁵⁵	li ²¹³						more

‘I love my teacher but I love truth more’

However, the implication of alternative (b) is not equivalent to the English version, since the former has the positive assumption that ‘I love my teacher’ while the latter is neutral in assumption. English expressions such as ‘he likes to chew bubble gum more than to eat ice cream’ or ‘her lips are redder than the grass is green’ are impossible to render in Chinese.

Since Vstat does not carry any comparative form morphologically, comparison need not be marked in Chinese if appropriate context is given. For example, the answer in (79) is an unmarked comparative sentence with the previous context of a question:

79. 他t'a⁵⁵ 高kaw⁵⁵ (還是xaj³⁵§1) 你ni^{213>11} 高kaw⁵⁵? — 他t'a⁵⁵ 高kaw⁵⁵
 3sg be-tall (or) 2sg be-tall? — 3sg be-tall
 'is s/he taller or you? — he is taller'

4.9.2 The positive degree

There is a distinction between the Equal Degree and the Equalling Degree, the former implies sameness while the latter indicates the extent to which someone or something reaches. There is no difference across dialects for these two subtypes of comparison save the lexical divergence in the use of conjunctions and comparative markers. For the first type, the general structural formula is NP1 + cnj/*resemble* + NP2 + ADV + Vcomp where *resemble* indicates a verb with the meaning of ‘to resemble,’ while ADV here is one that has the meaning of ‘the same.’ For example with cnj (80), and with *resemble* (81):

- | | | | | | |
|----------------------|----------------------------|---------------------|--------------------------|--|--------------------|
| 80a. SM | 我wɔ ^{213>11} | 跟kyn ⁵⁵ | 你ni ^{213>11} | 一樣ji ^{55>35} jan ⁵¹ | 高kaw ⁵⁵ |
| 80b. Shanghai (Wu) | 我 | 搭 | 儂 | 一樣 | 高 |
| 80c. Méixiān (Hakka) | 我 | 同 | 你 | 一般 | 高 |
| 80d. Cantonese (Yue) | 我ŋɔ ²⁴ | 同t'ɔŋ ¹¹ | 你nej ²⁴ | 一樣jot ⁵ jøŋ ²² | 高kow ⁵⁵ |
| 80e. Xiàmén (Min) | 我 | 合 | 汝 | 平(平) | 懸 |
| | 1sg | cnj | 2sg | same | be-tall |
| | ‘I am of your same height’ | | | | |

- | | | | | | |
|----------------------|--------------------------------|---------------------|--------------------|------------------------|---|
| 81a. SM | 你ni ^{213>11} | 像ejan ⁵³ | 他t'a ⁵⁵ | 那樣na ⁵¹ jan | 健壯jen ⁵¹ tʂwan ⁵¹ |
| 81b. Cantonese (Yue) | 你ni ²⁴ | 似ts'í ²⁴ | k'oy ²⁴ | [kem ⁴⁴] | 壯健tʃɔŋ ⁴⁴ kin ²² |
| 81c. Quánzhou (Min) | 汝 | 親像若 | 伊 | | 奐勇 |
| | 2sg | resemble | 3sg | that | strong |
| | ‘you are as strong as s/he is’ | | | | |

For the Equalling Degree type, the formula is NP1 Vex NP2 (ADV) Vstat where ADV is limited to the adverb with the meaning of ‘to that degree.’ The difference between this second type and the first type, namely, between the Equalling Degree and the Equal Degree, is most transparent in the negative form of these two types, since the affirmative forms sometimes convey the same meaning. For example:

- | | | | | | |
|----------------|------------------------------|-----------------------------|--------------------------|-------------------------|--------------------|
| 82a. SM | 我wɔ ^{213>11} | 沒(有)məj ³⁵ (jow) | 你ni ^{213>11} | (那麼na ⁵¹ mə) | 高kaw ⁵⁵ |
| 82b. Cantonese | 我ŋɔ ²⁴ | mow ²⁴ | 你nej ²⁴ | [kem ⁴⁴] | 高kow ⁵⁵ |
| | 1sg | not-have | 2sg | that | be-tall |
| | ‘I do not reach your height’ | | | | |

4.9.3 Superlative degree

To express the superlative degree, all dialects place the superlative adverb ‘the most,’ 最 in its various forms, before the Vcomp.

TABLE 5.11 COMPARATIVE CONSTRUCTIONS

COMPARATIVE DEGREE ADV = 更, 還要, 還, 要	
Northern structure NP1 CPR NP2 (ADV) Vcomp (MP)	CPR = 比 SM: 他t'a ⁵⁵ 比pi ^{213>35} 我wo ^{213>11} 更kəŋ ⁵¹ 愛aj ⁵³ 你ni ^{213>11} ‘he loves you even more than I do’
Southern structure NP1 Vcomp CPR NP2 (MP) or NP1 (ADV) Vcomp CPR NP2 (MP)	CPR = [k'a?] (S Min) Xiāmén: 伊i k'a? lo gwa 過[kwɔ ⁴⁴] (Yue) k'øy ²⁴ kow ⁵⁵ 過kwɔ ⁴⁴ 我ŋɔ ²⁴ 起(Shandong) Mùpíng: 他高起我
with MP and/or ADV	(Xiang) Xiangxiang: 我大你三歲 SM: 他t'a ⁵⁵ 還要haj ³⁵ jaw高kaw ⁵⁵ 我wo ^{213>11} —ji ⁵⁵ 個 kə頭t'ow ³⁵
Vcomp—Vstat, Vtrq (e.g. ‘to like,’ ‘to love,’ ‘to hate,’ ‘to hope’), Vopt (e.g. ‘should,’ ‘to dare,’ ‘to be willing’); adverbs derived from Vstat (e.g. ‘early,’ ‘late,’ ‘first,’ ‘last,’ ‘more,’ ‘less,’ ‘frequent’)	
Restrictions: NP1 can only be Subject NP, Time, NPloc, PP, [VP/Sentence] _{NP} —NOT OBJ Comparison of the object, the predicate, or the sentence is expressed by (a) topicalizing these constituents: SM: 我wo ^{213>11} 書su ⁵⁵ 比pi ^{213>35} 你ni ^{213>11} 多twɔ ⁵⁵ or by (b) rendering the comparison in a two-clause sentence: SM: 我wo ^{213>11} 愛aj ⁵¹ 我wo ^{213>11} 的老師law ^{>1} si ⁵⁵ , 可是k ³ ə ¹¹ si我wo ^{213>11} 更kəŋ ⁵¹ 愛aj ⁵¹ 真理tʂən ⁵⁵ li ²¹³	
POSITIVE DEGREE	
Equal Degree NP1 cnj/resemble NP2 ADV Vcomp	Equalling Degree NP1 Vex NP2 (ADV) Vstat
SM: 我wo ^{213>11} 跟kəŋ ⁵⁵ 你ni ^{213>11} 一樣i ^{55>35} jan ⁵¹ 高kaw ⁵⁵ (Wu) Shanghai: 我搭儂一樣高 (Hakka) Méixiàn: 我同你一般高 (Yue) Cantonese: 我ŋɔ ²⁴ 同t'ɔŋ ¹¹ 你nei ²⁴ 一樣 jət'ɔŋ ²² 高kow ⁵⁵ (Min) Xiāmén (Min) 我合汝平(平)懸 (Min) Quánzhou: 汝親像若伊 兔勇	SM: 我wo ^{213>11} 沒(有)mej ³⁵ (jow) 你ni ^{213>11} (那麼 na ⁵¹ ma) 高kaw ⁵⁵ (Yue) Cantonese: 我ŋɔ ²⁴ mow ²⁴ 你nei ²⁴ [kəm ⁴⁴] 高 kow ⁵⁵
SUPERLATIVE DEGREE	NP + 最 + Vcomp

4.10 Attributive constructions

4.10.1 Modifying clause

The older typological distinction between Northern and Southern Chinese in the modifying clause structure barely exists in the modern dialects, although residues can still be found. The Northern pattern is S/VP/PP/NP + ATTR + NPh where NPh signifies the Head Noun. The ATTR carries a dental initial. The residual Southern pattern is simply S/VP/PP/NP + NPh with zero attributive marker. It survives in a few Jianghuai (Huáiyīn, Shùyáng) and Wu (Chángzhōu, Suzhou) dialects as well as S Min and Yue, where the NPh must contain a DEM in at least the Min and the Yue dialects. For example:

- 83a. Huáiyīn 人家 是 漂漂亮亮 大姑娘
IMPERSONAL BE BEAUTIFUL BIG GIRL ‘she is a beautiful young girl’
- 83b. Kaipíng (Yue) 暖暖nɔn⁴⁴nɔn³⁵ [neir²¹] [kaw³²] 番薯fan⁴⁴si³⁵
SLIGHTLY-WARM THAT CL SWEET-POTATO ‘the slightly warm sweet potato’

The majority of the Southern and Central dialects, however, share exactly the same structure as the Northern dialects except that classifiers function as the attributive marker, in particular, the general classifier 個 or its equivalents often override others. For example:

- 84a. Héngyáng (Xiang) 你手 上 隻 菌子 吃 不得, 有 毒
2sg hand on CL mushroom eat-not-can, have poison
'the mushrooms in your hand are inedible, they are poisonous'
- 84b. Suzhou 阿黃 就 是剛剛走開 隻 小狗
A-Huang exactly be just leave CL puppy 'A-Huang is the puppy that just left'

Two types of modifying clause can be distinguished when the head noun is marked with a determiner that contains a demonstrative. When the modifying clause follows DET, it is descriptive and when it precedes DET, it is restrictive. Semantically the former type simply gives a general description while the latter designates outstanding characteristics. Compare the following pair of examples in SM:

- 85a. 他^ta⁵⁵把pa^{213>11}那na⁵¹枝tʂi⁵⁵新ein⁵⁵買maj^{213>11}的tə 筆pi^{213>11}丟tiw⁵⁵了lə
3sg BA that CL new buy ATTR pen lost Fp
'he lost that newly bought pen'
- 85b. 他^ta⁵⁵把pa^{213>11}新ein⁵⁵買maj^{213>11}的tə 那na⁵¹枝tʂi⁵⁵筆pi^{213>11}丟tiw⁵⁵了lə
3sg BA new buy ATTR that CL pen lost Fp
'he lost that newly bought pen'

While the first example simply narrates the fact with no other implication, the second one implies that he still has some pen(s) left. Syntactically the head noun of the restrictive modifying clause must always have a DEM, indicating that it is always definite in reference. In addition, such a clause cannot occur with Vex. Both of the following examples in SM are ungrammatical:

- 86a. *新ein⁵⁵買maj^{213>11}的tə 一ji⁵⁵枝tʂi⁵⁵筆pi^{213>11}丟tiw⁵⁵了lə
new buy ATTR one CL pen lost FP
- 86b. *有jow^{213>11}新ein⁵⁵買maj^{213>11}的tə 一ji⁵⁵枝tʂi⁵⁵筆pi^{213>11}在tsaj⁵¹這兒tʂər⁵¹
exist new buy ATTR one CL pen locate here

4.10.2 Noun complement clause

When the head noun is a time noun or [+abstract] noun or a “nominalized” NP, the modifiers constitute a noun complement clause rather than a modifying clause. The main difference between the two is that a head noun must always occur with a noun complement but it may be elided in a construction with a modifying clause. Compare the following in SM:

- 87a. 他^ta⁵⁵買maj^{213>11}的tə 原因ψen³⁵jin⁵⁵
3sg buy ATTR reason 'the reason he bought (something)'

87b. *他't'a⁵⁵ 買maj^{213>11} 的tə ‘for which (reason) he bought’

88a. 他't'a⁵⁵ 買maj^{213>11} 的tə 東西tonj⁵⁵ ei
3sg buy ATTR thing ‘the thing he bought’

88b. 他't'a⁵⁵ 買maj^{213>11} 的tə ‘that which he bought’
3sg buy ATTR > NMLZ

In addition, the head noun in a construction with a modifying clause is co-referential with either the subject or the object (including object of a preposition) in the underlying structure of the modifying clause, while the head noun in a construction with a noun complement is not co-referential with such elements but with an adverbial.

To distinguish between the ATTR in (87a) or (88a) which occurs with the head noun and the ATTR in (88b) which does not occur with the head noun, the former should be called a clause marker (ATTR) and the latter a nominalizer. In most dialects, the ATTR and the NMLZ share the same form, either a classifier or some form of 的. 的 assumes the phonetic shape of [tə] or [ni]/[nə]/[ne] (SW Mandarin) and is believed to have derived from 底 of the Tang-Song period and ultimately from 者 of the Qin-Han times. The general classifier 個 ([kɔ⁴⁴] in Cantonese) and its cognates with a velar initial are widely used in the Southern and Central dialects as well as a number of SWS Mandarin dialects distributed mainly in the southern periphery of Hunan (Xintián, Guìyáng, Yízhāng, Línwǔ, Lánshān, Níngyuān, Dàoxiàn, Jiānghuá, Jiāngyōng, Fēnghuáng). It also appears in a reduced form such as [e] in the Min and the Hakka dialects spoken in Taiwan.

4.11 Nominalization

Apart from decapitated modifying clauses exemplified in (88b) in the previous section, which may be considered nominalized and marked with a nominalizer, nominalization is otherwise unmarked in Chinese. A VP or a sentence can be embedded as the subject or the object of another sentence without being nominalized into an NP through marking. The matrix sentence in which an embedded sentence occurs as subject usually has Vcop or Vstat as the matrix verb, and the Vstat serving as such a matrix verb are those which can take [+abstract] subject, such as ‘to be important,’ ‘to be easy,’ ‘to be good,’ ‘to be all right,’ ‘to be common.’ For example:

89. SM 我wɔ^{213>11} 不pu^{51>35} 去te'y⁵¹ 可以k'v^{213>35} ji 嘴ma
1sg NEG go to-be-all-right Q
'is it all right if I do not go?'

Apart from the quotative verb (Vq), ‘to say,’ a great variety of verbs (Vdo, Vtrq, verbs expressing thoughts or perceptions, etc.) can take a VP or a sentence as object. For example in Cantonese:

90. 大家taj²²ka⁵⁵ 都tow⁵⁵ 知道tʃ⁵⁵tow⁴⁴ [mow²⁴] 電話tin²²wa³⁵ [m¹¹] 方便fɔŋ⁵⁵pin²²
everyone all know have-not telephone NEG convenient
'everyone knows that having no telephone is inconvenient'

The above example actually contains a sentence as object of the matrix verb ‘to know’ and within the embedded object sentence, a VP serves as the subject.

4.12 Verbal complements

4.12.1 Manner and extent

Verbal complements follow the main verb. There are two main types that are introduced by markers or complementizers (COMP), the Manner Complement and the Extent Complement. In the Southern dialects, these two types are marked with distinct markers, while in the Northern dialects they are marked with homonymous markers.

The Manner COMPL is also called the Degree or Descriptive COMPL. Constructions with this complement have the structure of NP1 + VP1-COMP + VP2 where VP1 contains a Vt or Vi and VP2 a Vstat. The COMPs are most often verbal suffixes. In the Northern, the Wu, and the Yue dialects, COMP is 得, [tə] (SM) or [ti], [tei], [tiə], [tiæ], [tə?], [tiə?], [lei], [lə?], [nə?] (Shanxi dialects) or [te] (Guiyáng of SW Mandarin) or [tsə?], [tse?], [tsə?], [tə?], [də?], [tə?], [lə?], [te?], [læ], [le], [əi] (Wu) or [tek⁵] (Cantonese); in the Xiang dialects (Chángsha) it is 起 or 得. However, in the Min dialects as well as some Hakka (Miáoli) and Yue (Kaipíng) dialects, COMP is zero. For example:

91a. SM	他t'a ⁵⁵	跑p'aw ^{213>11}	得tə	很xən ^{213>11}	快k'waj ⁵¹
91b. Chángsha	他	跑	起	好	快
91c. Cantonese	佢k'øy ²⁴	走tsaw ³⁵	得tek ⁵	好how ³⁵	快faj ⁴⁴
91d. Kaipíng	佢k'uy ⁴⁴	走taw ⁵⁵		快faj ³⁵	快faj ⁴⁴
91e. Miáoli	其	走		[t'ɔŋ ²⁴]	快faj ⁴⁴
91f. S Min	伊	走		真	緊
	3sg	run	COMP	very	fast 's/he runs very fast'

When an object NP occurs in VP1, it is either topicalized (Northern, Wu) or the verb is repeated (Northern), but in the Xiang and the Yue dialects, it could be placed after V-COMP.³⁷ For example:

92a. SM	(吃tsi ⁵⁵) 藥jaw ⁵¹	吃tsi ⁵⁵	得tə	多twɔ ⁵⁵
92b. Chángsha	吃		得	藥 有
92c. Cantonese	食sik ²		得tek ⁵	藥jøk ²² 多tɔ ⁵⁵
	eat	medicine eat	COMP medicine	much ‘take a lot of medicine’

93. Chángshú (Wu) 菜ts'æ 烹gø 來læ 交關tei²kuoe 好xo
 food cook COMP very good ‘the food is cooked very well’

The Extent COMPL consists of a sentence or a VP. Constructions with this complement have the structure VP1 + COMP + S/VP2, where VP1 contains a Vt, Vi or Vtrq. In the Northern dialects, the COMP for this complement is homonymous with that for the Manner COMPL, but in the Southern dialects, it is a different marker—到[tow⁴⁴] in Yue, [ka?ɔ²²] or [kaw³¹] in S Min. Among Northern Chinese, some Shanxi dialects assign different markers for these two types of complements. Examples of Manner COMPLs are found in (94), while Extent COMPLs are found in (95).

94a. Jíxiàn	你	說	得tei	好
94b. Xiāngyuán	你	說	得tə?ɔ ⁵	真 好
94c. Qinxìan	你	說	勒lə?ɔ ⁵	楞 好
94d. Cantonese	你nej ²⁴	講kɔŋ ³⁵	得te?ɔ ⁵	好how ³⁵
	2sg	speak	COMP	very good ‘you speak very well’

- 95a. Jíxiān 重 [li] 我 都 荷 不 起 啦
 95b. Xiāngyuán 重 得tiə?³ 連 我 也 [xəŋ²¹³] 不 動 了
 95c. Qinxian 扎實 得tiə? 連 我 也 荷 不 動 了
 heavy COMP even 1sg also lift NEG move FP
- 95d. Cantonese 重tʃ'or²⁴ 到tow⁴⁴ 我ŋ²⁴ 都tow⁵⁵ [m¹¹] 抬t'ɔj¹¹ 得tek⁵ [jok⁵]
 heavy COMP 1sg also NEG lift can move FP
 'so heavy that I cannot lift it'

A prominent difference between the Extent COMPL and the Manner COMPL is that the former indicates a causative relationship. The subject NP of VP1 can appear before VP2 if VP1 and VP2 share the same subject NP, but this switch is impossible for constructions with the Manner COMPL; therefore (91a) cannot be turned into (91g):

- 96a. SM 他t'a⁵⁵ 跑p'aw^{213>11}得tə 累lej⁵¹ 極tei³⁵ 了lə
 3sg run COMP tired extreme FP
 'he is extremely tired from running'
- 96b. SM 跑p'aw^{213>11}得tə 他t'a⁵⁵ 累lej⁵¹ 極tei³⁵ 了lə
 run COMP 3sg tired extreme FP
 'the running makes him extremely tired'
- 91f. SM *跑p'aw^{213>11} 得tə 他t'a⁵⁵ 很xən^{213>11} 快k'waj⁵¹
 run COMP 3sg very fast
 *'the running makes him extremely fast'

4.12.2 Resultative and directional

The Resultative COMPL is also called the Causative COMPL. The construction with this complement consists of a sequence of V1 + V2 where V1 and V2 bear a cause and result relationship. V1 is usually Vt or Vi of action or a limited number of modal Vstat and Vtrq and V2 a modal Vstat or Vtrq. An important feature of this construction is its potential form, in which the positive or negative potential marker occurs between V1 and V2. In most dialects, the positive potential marker (POT) is a form of 得 ([tə] in SM and [tek⁵] in Cantonese) while the negative potential marker is the same as the regular negative markers. Whereas non-potential negation has the word order of NEG + V, potential negation in the Northern dialects takes the form of V1 + NEG + V2 and in the Southern dialects the form NEG + V1 + POT + V2, although the Northern form has been widely adopted in the south too.³⁸ While the positive potential in most dialects has the word order V1 + POT + V2, a number of Northern dialects in Shandong (Jinán, Píngyi, Wéifang, Wéixiān, Zibó), Shanxi (Língchuan, Tàigǔ), Shaanxi (Yánchuan), Gansu (Línxia), and Hebei (Changlí, Gúchén, Huòlù) have a different order, V1 + V2 + POT; in addition, POT in such dialects is usually a form of 了 ([liau]/[liə]/[ləu]/[lə], etc.). For example:

- 97a. SM 拿na³⁵ 得tə 動tonj⁵¹, 拿na³⁵ 不pu 動tonj⁵¹
97b. Changlì 拿 動 [liou], 拿 不 動
97c. Cantonese [nɪŋ⁵⁵] 得tek⁵ [jok⁵], [n¹¹] [nɪŋ⁵⁵] 得tek⁵ [jok⁵]
take POT move POT, NEG take NEG POT move
'can move (it), cannot move (it)'

In most dialects, aspect markers occur after V2 only, and this is the reason the combination V1 + V2 has often been treated as a compound. However, in the Jianghuai dialect of Tàixīng, at least the perfective aspect occurs after both V1 and V2, for example:

- 98a. 我 說 [ka] 服 [ka] 他 了
1sg talk PFV convince PFV 3sg Fp 'I convinced him'
98b. 衣裳 曬 [a] 乾 [ŋa] 了
clothes sun PFV dry PFV Fp 'the clothes got dry in the sun'

In the S Min dialects, the past tense marker, in the affirmative or negative form, may occur after V1, for example:

- 99a. 新 冊 買 有 著
new book buy PAST obtain 'the new book was bought'
99b. 衫褲 曝 無 調
clothes sun NEG-PAST dry 'the clothes did not get dry in the sun'

From a comparative point of view, V1 + V2 is best considered a V-COMPL structure and this is exactly its historical origin.

When an object NP occurs together with the Resultative COMPL, the word order is generally V1 + V2 + O; however, in some Wu dialects, the older order V1 + O + V2 obtains—敲伊碎 ‘knock it to pieces,’ 曬伊乾 ‘dry it in the sun,’ etc. are found in, for example, Shanghai. In the potential form, the Northern dialects have the word order of placing the O after V2, namely, V1 + NEG/POT + V2 + O, while some Southern and Central dialects still preserve the older word order of placing O before V2 under certain conditions. In the affirmative potential, V + POT + O + V2 is used in some Xiang (Dòngkǒu, Chángsha), Wu (Danyáng, Chángzhōu, Suzhou, Kunshan, Bǎoshān, Shānghái, Wújiāng, Jiāxìng, Húzhōu, Hángzhōu, Shàoxīng, Zhuījí, Shèngxiàn, Yúyáo, Níngbō, Huángyán, Wēnzhōu, Qúzhōu, Jīnhuá) and Yue dialects,³⁹ for example:

- 100a. Dòngkǒu 果 隻 塘 裡 捉 得 蟻 多 魚 倒
this CL pond in catch POT very many fish obtain
'many fish can be caught in this pond'
100b. Jīnhuá 佢 寫 得 字 來
3sg write POT word obtain 's/he can write'
100c. Cantonese 搞kaw³⁵ 得tek⁵ 佢'øy²⁴ [tim²²]
make POT 3sg straight 'can straighten it'

In the negative potential, there is more than one possible word order: V1 + O + NEG + V2 or NEG + V1 + POT + O + V2 or V1 + NEG + O + V2. The first alternative is the

most popular; the second alternative occurs in Chángsha and the Yue dialects and the last one in Chángsha and some Wu (Shèngxiān Chóngrénn, Níngbo) dialects only. For example:

- 101a. Héngyáng (Xiang) 狗 追 貓 不 到
dog pursue cat NEG reach ‘the dog cannot catch the cat’

- 101b. Shàoxing 我ŋo 打taŋ 伊fi 勿vi? 過ku
‘I am no match in fighting him/her’

- 101c. Cantonese 我ŋɔ²⁴ 打ta³⁵ 佢k'ɔy²⁴ [m¹¹] 過kwo⁴⁴
1sg hit 3sg NEG surpass
‘I am no match in fighting him/her’

- 102a. Chángsha 冒 買 得 肉 到
NEG:PFV buy POT meat obtain ‘did not succeed in buying meat’

- 102b. Cantonese 你nej²⁴ [m¹¹] 打ta³⁵ 得tek⁵ 佢k'ɔy²⁴ 死sej³⁵
2sg NEG hit POT 3sg die
‘you cannot beat him/her to death’

- 103a. Chángsha 他 吃 不 飯 進
3sg eat NEG rice enter ‘s/he cannot eat (send food down)’

- 103b. Shèngxiān Chóngrénn 我ŋɔ 打tã 勿vε? 伊 過kɔ
1sg hit NEG 3sg surpass
‘I am no match in fighting him/her’

The Directional COMPL is actually another type of resultative complement where V2, followed by a directional suffix (DIR), is a directional verb (Vdir) and indicates the direction of the movement of V1. V1 is usually a motion verb (Vm) or action verb. The dir is either 来 ‘toward’ or 去 ‘away from.’ When a locative NP (NPloc) occurs with this complement, it always comes after Vdir. In the Yue dialects, however, if NPloc appears, no dir can occur. For example:

- 104a. SM 跑 p'aw^{213>11} 回 xwej³⁵ 來 laj

- 104b. Cantonese 走 tsaw³⁵ 翻 fan⁵⁵ 來 lej¹¹
run return come ‘run back (toward speaker)’

- 105a. SM 跑 p'aw^{213>11} 回 xwej³⁵ 家 jia⁵⁵ 來 laj

- 105b. Cantonese 走 tsaw³⁵ 翻 fan⁵⁵ 屋企 ok'sk'ej³⁵ 來 laj
run return home come
‘run back home (toward speaker)’

When an object NP co-occurs, it can appear after dir or after V2 (Vdir) or after V1 in Northern Chinese, but it can appear only after V1 in the Yue dialects. For example:

- 106a. SM 拿 na³⁵ 兩 lian^{213>35} 本 pən^{213>11} 書 ʂu⁵⁵ 出 tʂ'u⁵⁵ 來 laj

- 106b. SM 拿 na³⁵ 出 tʂ'u⁵⁵ 兩 lian^{213>35} 本 pən^{213>11} 書 ʂu⁵⁵ 來 laj

- 106c. SM 拿na³⁵ 出tʂ'u⁵⁵ 來laj 兩lian^{213>35} 本pən^{213>11} 書ʂu⁵⁵
 106d. Cantonese [niŋ⁵⁵] 兩lœŋ²⁴ 本pun³⁵ 書sy⁵⁵ 出tʂ'ot⁵ 來lej
 take exit come two CL book exit come
 ‘bring out two books’

4.12.3 Imperative complement

A number of verbs of request or command, which may be called imperative verbs (Vimp), are always followed by a complement in the imperative form, which may be called IMP COMPL. Vimp includes verbs with the meaning of ‘to ask for,’ ‘to persuade,’ ‘to request,’ ‘to implore,’ ‘to beg,’ ‘to urge,’ ‘to hint,’ ‘to force,’ ‘to encourage,’ etc. These have been widely mislabeled “prepivotal” or “telescoping” verbs and misinterpreted as occurring in a pivotal construction. Two facts confirm that we are dealing with an IMP COMPL: when the complement is in the affirmative it can contain emphatic adverbs that occur only in imperative sentences, and when the complement is in the negative the NEG is unmistakably the one that occurs only in imperative sentences. For example, in (107) ‘by all means’ can only be used in an imperative sentence, and in (108) the negative is the NEG-imp and not one that can occur in non-imperative sentences:

- 107a. SM 你ni^{213>11} 勸te'yen⁵¹ 他ta⁵⁵ 千萬te'jen⁵⁵wan 要jaw⁵¹來laj³⁵
 107b. Cantonese 你nej²⁴ 勸hyn⁴⁴ 佢k'oy²⁴ 千祈tʃ'in⁵⁵k'ej¹¹ 要jiw⁴⁴來lej¹¹
 2sg persuade 3sg by-all-means must come
 ‘you persuade him to come by all means’
- 108a. SM 你ni^{213>11}叫tejaw⁵¹我ŋɔ^{213>11}別pje³⁵給kej^{213>11}他ta⁵⁵牛奶niw³⁵naj 吃tʂɻ⁵⁵
 108b. Shanghai 儂 2sg 叫 1sg 別 3sg 勿要 撥 牛奶 伊 eat
 2sg ask 1sg NEG-imp give 3sg milk 3sg eat
 ‘you asked me not to give him milk to drink’

Since an imperative sentence can only have the second person or first inclusive person as the subject of the sentence, the IMP COMPL in question has such a subject NP implied. The 3sg PN in (107) and (108) cannot serve as the subject NP of the IMP COMPL at all, since *他ta⁵⁵ 千萬te'jen⁵⁵wan 要jaw⁵¹來laj³⁵ or *佢k'oy²⁴ 千祈tʃ'in⁵⁵k'ej¹¹ 要jiw⁴⁴來lej¹¹ or *我ŋɔ^{213>11} 別pje³⁵ 紿kej^{213>11} 他ta⁵⁵ 牛奶niw³⁵naj 吃tʂɻ⁵⁵ or *我勿要撥牛奶伊吃 are all ungrammatical.

4.13 Subjoining constructions

The word order of subjoining structure is Dependent Clause preceding the Main Clause. Furthermore, the Dependent Clause may be embedded in the Main Clause, occurring after the subject of the latter, namely, NP + Dependent Clause + VP. Sentences with clauses expressing Time, Cause, Condition, and Concession may be described under this kind of structure. For example:

- 109a. SM 自從tsɻ⁵¹ts'uj⁵¹認識zəŋ⁵¹ʂɻ⁵¹ 你ni³⁵以來ji¹¹laj李四li¹¹s⁵¹沒mej³⁵ 喝xʂ⁵⁵過kwo 酒tejow²¹³
 from know 2sg since Li Si NEG-PFV drink exp-asp liquor

109b. SM

李四li¹¹s₁⁵¹自從ts₁⁵¹ts'uj認識zən⁵¹s₁你ni³⁵以來ji¹¹laj沒mej³⁵喝x₁y⁵⁵過kwo 酒tejow²¹³
 Li Si from know 2sg since NEG-PFV drink exp-asp liquor
 'since knowing you, Li Si has not drunk any liquor'

110a. Cantonese

因為jen⁵⁵wej²²個ko⁴⁴仔tsej³⁵病pej²²所以so³⁵ji²⁴李四lej²⁴sej⁴⁴[mow²⁴] 來lej¹¹
 because CL son sick therefore Li Si NEG-PAST come

110b. Cantonese

李四lej²⁴sej⁴⁴因為jen⁵⁵wej²²個ko⁴⁴仔tsej³⁵病pej²²所以so³⁵ji²⁴ [mow²⁴] 來lej¹¹
 Li Si because CL son sick therefore NEG-PAST come
 'Li Si did not come because his/her son was sick'

111a. SM 要不是jaw⁵¹puʂ₁他t'a⁵⁵來laj³⁵我wo^{213>11}就tejow⁵¹同意t'uj³⁵ji 了le
 if-not 3sg come 1sg then agree Fp

111b. SM 我wo^{213>11}要不是jaw⁵¹puʂ₁他t'a⁵⁵來laj³⁵就tejow⁵¹同意t'uj³⁵ji 了le
 1sg if-not 3sg come then agree Fp
 'if he had not come, I would have agreed'

112a. Cantonese

雖然søy⁵⁵jin¹¹我ŋɔ²⁴畀pej³⁵[tʃɔ³⁵]錢tʃin³⁵佢k'øy²⁴都tow⁵⁵[tʃon²²] [new⁵⁵]
 although 1sg give PFV money 3sg also still angry

112b. Cantonese

佢k'øy²⁴雖然søy⁵⁵jin¹¹我ŋɔ²⁴畀pej³⁵[tʃɔ³⁵]錢tʃin³⁵都tow⁵⁵[tʃon²²] [new⁵⁵]
 3sg although 1sg give PFV money also still angry
 'although I paid, he is still angry'

Notice that when the Dependent Clause and the Main Clause have different subject NPs, as in (111a 要不是他來,我就...) and (112a 雖然我畀[tʃɔ³⁵]錢, 佢都...), the conjunction of the Dependent Clause must precede its subject NP. When both clauses share the same subject NPs, as in (109a 自從(李四)..., 李四沒...) and (110a 因為(李四)個仔..., 李四...), the subject NP appears only in the Main Clause but elided in the Dependent Clause. Where both clauses share the same subject NPs, if the subject NP should appear in the Dependent Clause, as in (109b 李四自從(李四)..., (李四)沒...) and (110b 李四因為(李四個仔)..., 所以(李四) [mow²⁴]...), it must precede the conjunction, just as in (111b 我要不是他..., (我)就...) and (112b 佢雖然我..., (佢)都...), where both clauses have different subject NPs. The structure shared by both types, with and without shared subject NPs, is as illustrated in (b), namely, embedded structure. It may thus be more accurate to describe all these structures as embedding constructions, the Dependent Clause as Embedded Clause and the Main Clause as Matrix Clause.

Another point to note is that no pronominalization is used when both clauses share the same subject NP. When a pronoun appears in the Embedded Clause, it will be understood as having disjoint reference; for example, if a 3sg PN is added in the Embedded Clause

of (109a) or (110a), “他t'a⁵⁵” and “佢k'oy²⁴” will be understood as a person different from Li Si.⁴⁰

- 109c. SM 自從tsr⁵¹ts'ɔŋ 他t'a⁵⁵ 認識zəŋ⁵¹ŋl 你ni³⁵ 以來ji¹¹laj 李四li¹¹sr⁵¹ 沒mej³⁵ . . .
from 3sg know 2sg since Li Si NEG:past

- 110c. Cantonese
因為jen⁵⁵wej²²佢k'oy²⁴個ko⁴⁴仔tſej³⁵病pen²²所以so²⁴ji²⁴李四lej²⁴sej⁴⁴ . . .
because 3sg CL son sick therefore Li Si

Conjunctions either appear in pairs, one in the Embedded and one in the Matrix Clause, or just in either the Matrix or the Embedded Clause, depending on the conjunction. For example, Time (Embedded) clauses usually have the conjunction, unless they designate ‘after,’ in which case if the PFV appears, no conjunction is needed but the Matrix Clause will include some adverb such as ‘then.’ For example:

- 113a. SM 你ni^{213>11} 吃tʂ'ɿ⁵⁵ 了lə 飯fan⁵¹ 再tsaj⁵¹ 去te'y⁵¹
113b. Cantonese 你nej²⁴ 食sik² [tʃɔ³⁵] 飯fan²² [tʃi⁴⁴] 去høy⁴⁴
113c. Suzhou (Wu) [nE] 吃te'iə? [ts] 飯VE 再tse 去te'ij
2sg eat PFV rice then go ‘you go after eating’

With Conditional clauses, sometimes no conjunction appears in either clause at all:

114. SM 你ni^{213>11} 不pu⁵¹ 來laj³⁵ 我wo^{213>11} 不pu^{51>35} 去te'y⁵¹
2sg NEG come 1sg NEG go ‘if you don't come, I won't go’

4.14 Verbal expressions in series (VP series)

Since Chinese need not use markers to distinguish Dependent from Main or Embedded from Matrix clause, VPs can occur in a series. For example:

- 115a. SM 他t'a⁵⁵ 穿tʂ'wan⁵⁵上ʂan 衣服ji⁵⁵fu 跑p'aw^{213>11}出tʂ'u⁵⁵ 大門ta⁵¹mən 來laj
3sg put on clothes run exit front-door come
'he put on his clothes and ran out of the front door (toward the speaker)'

In such a series, word order is important in determining the sequence of events. Given a different order, the above example will have a different reading:

- 115b. SM 他t'a⁵⁵ 跑p'aw^{213>11}出tʂ'u⁵⁵ 大門ta⁵¹mən 來laj 穿tʂ'wan⁵⁵上ʂan 衣服ji⁵⁵fu
3sg run exit front-door come put on clothes
'he ran out of the front door (toward the speaker) and put on his clothes'

The so-called Pivotal construction is a kind of VP series construction in which a post-verbal NP serves as a pivot—object of the preceding verb and subject of the following verb. For example, in (116) “佢k'oy²⁴” is at once the object of ‘to send’ and subject of ‘to go’:

116. Cantonese 你nej²⁴ 派p'aj⁴⁴ 佢k'oy²⁴ 去høy⁴⁴ [pin⁵⁵] 道tow²² 啊a⁴⁴
2sg send 3sg go which way Q 'where do you send him?'

4.15 Conjoining structures

Word order is not as important in conjoining structures, where clauses of the same rank may occur in variant order without affecting the basic meaning. For example:

- 117a. SM 一面ji^{55>35}mien 看k'an⁵¹ 報paw⁵¹ 一面ji^{55>35}mien 吃tʂɿ⁵⁵ 飯fan⁵¹
 117b. Shanghai 一頭 看 報 一頭 吃 飯
 simultaneously read newspaper simultaneously eat rice
 ‘read newspaper and eat at the same time’

- 118a. SM 一面ji^{55>35}mien 吃 飯fan⁵¹ 一面ji^{55>35}mien 看k'an⁵¹ 報paw⁵¹
 118b. Shanghai 一頭 吃 飯 一頭 看 報
 ‘eat and read newspaper at the same time’

Disjunctive structures behave the same way. For example:

- 119a. SM 要嘛jaw⁵¹ma 吃tʂɿ⁵⁵ 飯fan⁵¹ 要嘛jaw⁵¹ma 吃tʂɿ⁵⁵ 麵mjen⁵¹
 119b. Cantonese 一係jet⁵hej²² 食sik² 飯fan²² 一係jet⁵hej²² 食sik² 麵min²²
 either eat rice or eat noodle
 ‘either eat rice or noodles’
- 120a. SM 要嘛jaw⁵¹ma 吃tʂɿ⁵⁵ 麵mjen⁵¹ 要嘛jaw⁵¹ma 吃tʂɿ⁵⁵ 飯fan⁵¹
 120b. Cantonese 一係jet⁵hej²² 食sik² 麵min²² 一係jet⁵hej²² 食sik² 飯fan²²
 either eat noodle or eat rice
 ‘either eat noodles or rice’

FURTHER READING

- Huang, Borong (1996) *Hanyu Fangyan Yufa Leibian* (*Classified Materials on Chinese Dialect Grammar*), Qingdao: Qingdao Press.
- Yue(-Hashimoto), Anne O. (1993) *Comparative Chinese Dialectal Grammar: Handbook for Investigators*, Paris: Ecole des Hautes Etudes en Sciences Sociales Centre de Recherches Linguistiques sur l'Asie Orientale.

NOTES

- 1 “Classifier” refers to specific measure words that co-occur with countable nouns while “measure word” refers to standard measuring units for length, weight, etc. or temporary measuring units such as *a cup of, an occurrence of*, etc.
- 2 Exceptions are rare but do exist. In other words, there are nouns like 人口 ‘population’ that neither take classifiers nor measures.
- 3 All dialect forms are rendered in broad IPA notation with tones designated by superscripted numbers (forms without such numbers are atonal) while sandhi forms are given to the right of an arrow. Such forms within square brackets, when not designated with names of dialects, are Standard Mandarin. Pan-dialect cognates and dialectal forms not given phonetic notation in the original sources are designated with characters.
- 4 For easy identification, all place names below the province level, except for a few well-known cities, will be given tonal designation except for the first tone, to minimize the use of diacritics.

- 5 When [-] is given, it represents unspecified contents which vary among dialects. Here, [k'-] indicates that only the initial [k'] is shared among the dialects using this form while the rest of the word differs.

6 Subgroups of dialects are sometimes given, such as S Wu, W Min.

7 See, for example, Lü (1985): 59–61 for a possible etymological derivation of 門 from 輩.

8 See Lü (1985): 54.

9 Examples taken from Huang (1996): 439.

10 According to Huang (1996): 451–3 many villages and towns within Hǎiyán manifest this distinction and it is not limited to first and third singular but may include both the singular and the plural in all persons.

11 Huang (1996): 469 describes this dialect as Shānxìan 陝縣 in Henan; however, there is no 陝縣, only Jiāxiān 鄭縣 in Henan.

12 The tripartite distinction in this dialect for demonstratives (li-kài/liákài ‘this,’ kaikài/kákài ‘that,’ unkài ‘yonder’) and locatives (li-tse ‘here,’ kai-tse ‘there,’ un-tse ‘yonder’) was first observed in Schaank (1897: 26), which may be the earliest record of such kind of distinction in Chinese.

13 # indicates the occurrence of syntactic boundary such as that for NP. Some forms of the demonstratives do not occur as attributive but as pro-form, in which case they are marked with #.

14 Fp stands for “phase particle.” See section 2.2.2.2.

15 This 來 form is also used to indicate the immediate past in nineteenth century Cantonese texts and even in modern Cantonese in certain contexts: for example, 你去[pin⁵⁵]道來啊? ‘where have you been?’ (literally, ‘where did you go and come back?’), or in Yángjiāng (Yue).

16 The final particle 了 is derived from a weak form of 來 which in turn may be derived from the combination of 了 + 也.

17 See Comrie (1976, section 5.2.1.1).

18 [lə?] and [zɪ] are Vloc while [he]/[lɒ]/[tɔ?] is probably an indefinite locative deictic and [ta] is the distal locative ‘there.’

19 In several Huizhou dialects such as Jixi, Túnxi, Xiuning, Yixiàn, Vcop assumes the function of Vloc too.

20 Both [hei³⁵] and [ɔ³³] are Vloc and [nein³⁵] is the distal locative ‘there.’

21 This is the case in the majority of the Wu dialects. In some, such as Liyáng, Jintán, Danyáng, Jiangyin, Chángzhou, Suzhou, Jiaxing, Hángzhou, and Qúzhou, both SVO and OSV/SOV are used, while in a few, such as Danyáng Tóngjiāqiáo and Jinhúa, only SVO is used. For this kind of mixture of usage, see the explanation below in connection with post-verbal adverbs.

22 ADVt are different from TIME words. The former include adverbs such as ‘already,’ ‘just,’ ‘soon,’ ‘immediately,’ ‘first,’ ‘again,’ ‘finally,’ ‘always,’ etc., while the latter constitute NPs including date expressions and time designations such as ‘January 21,’ ‘today,’ etc.

23 There are some exceptions. The Yue dialect of Liánzhou uses 是 while according to Yang (1974) the Jiahé dialect of Hunan (SW Mandarin) uses both 是 and 系.

24 This NP can carry definite reference only in conditional sentences or for emphasis. For example in SM:

到處 都 有 這 種 人

everywhere emph exist this kind person
‘there is this kind of person everywhere.’ (emph = emphatic marker)

- 25 The indefinite locative expression functions as the progressive aspect. See section 2.2.2.4.

- 26 In some Northern Wu dialects, the neg is 勿 with a [p]/[f] initial rather than 未 with an [m] initial.
- 27 This distinction also prevails in the Huizhou dialects. See Hirata (1998): 274.
- 28 There is a historical explanation for this asymmetry. According to Lü (1942), 嗨 is derived from the combination of a negative plus a final particle.
- 29 不 serving as NEG in VP-NEG appears in the bamboo slips unearthed from Qin tombs in Shuihudi of Yúnmèng in Hubei.
- 30 Example #45 is taken from the field work data on Lichéng of Quánzhou city by Shi Qisheng in 1992 for my Comparative Chinese Dialectal Grammar project.
- 31 Only Yunnan geographically lies outside this southeastern region. However, according to Zhang (1990), Han immigrants into Yunnan are mostly from Anhui province.
- 32 We shall not include those Northern forms where the Oi is marked with 紿 [ge^{j213}] such as: 他送一本書給我 ‘he gave a book to me,’ since 紿 [ge^{j213}] + Oi may be analyzed as a co-verb phrase.
- 33 A few Northern dialects, for example, Luóshan and Xinxian of Henan, are reported to have the Od + Oi word order.
- 34 Notice that the patient may be repeated after the main verb.
- 35 Wang (1989: 285) considers these exceptions as falling into a fixed category of ‘favor from above’ which, just as calamities, cannot be opposed.
- 36 Example taken from Chin (2009): 59.
- 37 In the Yue dialects the Northern option is also used.
- 38 The potential form discussed in this section is limited to that occurring with the Resultive COMPL. With other kinds of VP, the potential is generally expressed with an optative verb, Vopt, such as 能 ([nry] in SM). Furthermore, dialects in Henan, for example, prefer the potential with Vopt even with Resultative COMPL.
- 39 In the modern Yue dialects, this word order obtains only if O is a pronoun; in the Wu dialects, except for Shàoxing, Yúyáo, and Wenzhou, this word order coexists with the Northern word order.
- 40 Some speakers of the younger generation accept pronominalization; moreover, a third person pronoun may have either disjoint reference or co-reference with the subject.

BIBLIOGRAPHY

- Chao, Yuen Ren (1948) *Mandarin Primer*, Cambridge: Harvard University Press.
- Chen, Zhangtai and Li Rulong (1991) *Minyu Yanjiu (A Study of the Min Dialects)*, Beijing: Yuwen Press.
- Chin, Chi On (2009) The Verb GIVE and the Double-Object Construction in Cantonese in Synchronic, Diachronic and Typological Perspectives, University of Washington PhD dissertation.
- Chinese Academy of Social Sciences and Australian Academy of Humanities, eds (1987) *Language Atlas of China*, Hong Kong: Longman Press.
- Comrie, Bernard (1976) *Aspect*, Cambridge: Cambridge University Press.
- Hirata, Shoji (1998) *Huizhou Fangyan Yanjiu (A Study of the Huizhou Dialects)*, Tokyo: Kohbun Press.
- Hou, Jingyi and Wen Duanzheng (1993) *Shanxi Fangyan Diaocha Baogao (Report on a Survey of the Shanxi Dialects)*, Taiyuan: Shanxi Higher Institutions United Press.
- Huang, Borong (1996) *Hanyu Fangyan Yufa Leidian (Classified Materials on Chinese Dialect Grammar)*, Qingdao: Qingdao Press.
- Li, Rulong and Zhang Shuangqing, eds (1992) *Ke Gan Fangyan Diaocha Baogao (Report on a Survey of the Kejia and Gan Dialects)*, Xiamen: Xiamen University Press.

- Lü, Shuxiang (1942) *Zhongguo Wenfa Yaolue* (*Outline of Chinese Grammar*), Shanghai: Commercial Press.
- Lü, Shuxiang (1985) *Jindai Hanyu Zhidaici* (*Deictics in Medieval Chinese*), Shanghai: Xuelin Press.
- Luo, Futeng (1997) *Muping Fangyan Cidian* (*Dictionary of the Muping Dialect*), Nanjing: Jiangsu Education Press.
- Norman, Jerry (1988) *Chinese*, Cambridge: Cambridge University Press.
- Ogawa, Tamaki (1981) ‘Suzhou fangyan de zhishi daici’ (Demonstratives in the Suzhou dialect), *Fangyan* 4: 287–288.
- Peking University Chinese Department (1995) *Hanyu Fangyan Cihui* (*Chinese Dialect Vocabulary*), 2nd edition, Beijing: Yuwen Press.
- Qian, Nairong (1992) *Dang Dai Wuyu Yanjiu* (*A Study of Current Wu Dialects*), Shanghai: Educational Press.
- Schaank, S.H. (1897) *Het Loeh-Foeng-Dialect*, Leiden: E.J. Brill. English translation by Bennett M. Lindauer (1979) *The Lu-Feng Dialect of Hakka [Writing and Language Reference Materials 5]*, BA thesis, the East Asian Studies Department, Princeton University.
- Wang, Li (1989) *Hanyu Yufashi* (*History of Chinese Grammar*), Beijing: Commercial Press.
- Wu, Yunji, ed. (1996) *Hunan Fangyan de Dongtai Zhuci* (*Aspectual Markers in Hunan Dialects*), Hunan: Hunan Normal University Press.
- Xu, Baohua and Tang Zhenzhu, eds (1988) *Shanghai Shiqu Fangyanzhizhi* (*A Survey of the Dialect of Shanghai City*), Shanghai: Shanghai Educational Press.
- Yan, Yiming (1994) *Wuyu Gaishuo* (*Outline of the Wu Dialects*), Shanghai: Huadong Normal University Press.
- Yang, Kai (2009) *Edong Fangyan Cihui Yanjiu* (*A Study of the Vocabulary of the Eastern Hubei Dialects*), Wuhan: Hubei People’s Press.
- Yang, Shifeng (1974) *Hunan Fangyan Diaocha Baogao* (*Report on a Survey of the Dialects of Hunan*), Institute of History & Philology Special Publication No. 66. Taiwan: Academia Sinica.
- Yuan, Jiahua (1983) *Hanyu Fangyan Gaiyao* (*Outline of Chinese Dialects*), Beijing: Wenzi Gaige Press.
- Yue-(Hashimoto), Anne O. (1993) *Comparative Chinese Dialectal Grammar: Handbook for Investigators*, Paris: Ecole des Hautes Etudes en Sciences Sociales, Centre de Recherches Linguistiques sur l’Asie Orientale.
- Zhang, Min (1990) *Hanyu Fangyan Fanfu Wenju de Leixingxue Yanjiu* (A Typological Study of Yes-No Questions in Chinese Dialects: In Diachronic Perspective), Peking University PhD dissertation.
- Zhang, Shuangqing, ed. (1996) *Dongci de Ti* (*Verbal Aspect*), Hong Kong: Chinese University.
- Zhou, Changji (1991) *Minnanhua yu Putonghua* (*Southern Min and Putonghua*), Beijing: Yuwen Press.

CHAPTER SIX

THE CHARACTERISTICS OF MANDARIN DIALECTS

Dah-an Ho

The Mandarin dialects are mainly spoken in the valleys of the Yellow and Huai Rivers, Inner Mongolia, the Northeast, Shaanxi, Gansu, Xinjiang, Sichuan, Yunnan, and Guizhou, and along the Yangtze River, by a population of over one billion. As an official language, Mandarin proper (or Putonghua 普通話 the common language) is in fact spoken throughout all of China. Mandarin dialects are divided into the regions of Northern Mandarin, Southwestern Mandarin, and Jiang-Huai (江淮) Mandarin. Although based on the accent of the educated in Beijing, which is a part of Northern Mandarin, Putonghua is not completely identical to the colloquial spoken in Beijing. For instance, 薄 *po* A2,¹ 學 *çye* A2, 色 *sr* C1, 更 *kay* A1 in Putonghua are *pau* A2, *ciau* A2, *çai* C1, *teiŋ* A1 in the colloquial of Beijing. The Beijing dialect (like many Chinese dialects, including the other Northern dialects listed above) has a distinction between literary and figurative/colloquial strata, and its literary reading is what is taken for Putonghua.

1 MANDARIN DIALECTS IN HISTORICAL DEVELOPMENT

In a study of the historical development of Modern Chinese, Tsu-lin Mei (1997: 82) compared Proto-Chinese and Modern Chinese in terms of the following attributes: 1 morphemes being monosyllabic, 2 possession of tones, 3 absence of consonant clusters, 4 lack of morphological inflection, 5 need for classifiers, 6 modifiers prior to the head, and 7 verb medial word order. The results are given in Table 6.1.

All the attributes of Modern Chinese are newly developed except for 1 morphemes being monosyllabic. From this one is able to see how enormous a change has taken place in the language, and we believe that Mandarin, among the modern dialects, has developed most rapidly. For example, comparing the initials, finals, and the number of tones (see Table 6.2), we can see that Mandarin is the dialect with the lowest number of syllables, which signifies that the sound system of Mandarin is the simplest and so the pace of development is the most rapid.

The reason that Mandarin dialects have evolved so expeditiously is closely related to the fusion of ethnic groups. In Chinese history, the fusion of different Han ethnic groups, and of Han and non-Han ethnic groups in the regions north of the Yangtze River, occurred much more frequently and vigorously than in the south (see LaPolla 2001). Consequently, while the boundaries between different dialects are still securely preserved in the regions south of the Yangtze, in the north, which is greater in area, they have, to a certain degree, gradually converged into a kind of blended dialect.

Convergence has been an important process in the formation of Mandarin dialects. We have obtained sufficient evidence to show that even though Putonghua represents the Mandarin dialect area, this does not mean that the dialects in this area were directly

TABLE 6.1 COMPARISONS OF THE ATTRIBUTES IN CHINESE

	1	2	3	4	5	6	7
Proto-Chinese	+	-	-	±	-	-	-
Modern Chinese	+	+	+	+	+	+	+

TABLE 6.2 COMPARISONS OF THE ATTRIBUTES OF MODERN CHINESE DIALECTS

Dialect	Representative area	Number of initials	Number of finals	Number of tones	Possible number of syllables
Mandarin	Common dialect	16	39	4	2,496
Wu	Suzhou	27	50	7	9,450
Min	Xiamen	15	57	7	5,985
Gan	Nanchang	19	59	6	6,726
Hakka	Meixian	17	69	6	7,038
Yue	Guangzhou	20	53	9	9,540
Xiang	Changsha	23	37	6	5,106

derived from Putonghua, or belong to the same group as Putonghua. Putonghua is only a kind of Mandarin dialect and has been the lingua franca of the country for no more than 200 years.

Zhang Wei (張位) of the Ming dynasty (1368–1644) wrote in *Wen Qi Ji* (《問奇集》) that “in the area north of the Yangtze entering tones are read as level tones; often the words do not have characters, and therefore cannot be completely recorded. South of the Yangtze often the dentals are not articulated clearly. Yet this is also a local variant of Guanhua (官話 ‘official language’) and therefore makes communication between indigenous dialects difficult.” This statement is the first documented appearance of the term Guanhua. This quote tells us that there were different local variations in the Guanhua at that time. According to the understanding today, in the fifteenth to the eighteenth centuries Guanhua was based mainly on the Jiang-Huai (Nanjing) dialect. However, even if there were local variations, as long as the differences were not too great and there were regular correspondences between the dialects, the area of communicability for this Guanhua could have been as great as that of the Mandarin dialect area nowadays. People at that time often called the variant that was spoken in the lower Yellow River area Zhongzhou Yin (中州音, ‘central state language’), Zhongyuan Yin (中原音, ‘central plains language’) or Zhongyuan Yayin (中原雅音, ‘central plains educated language’), all of which were considered to be culturally superior. The Putonghua of today is a branch of this particular variation in the north of Hebei and is the result of the fusion of the Han and the Manchu peoples. It was approximately at the beginning of the nineteenth century that the Beijing dialect replaced the Jiang-Huai dialect as the standard for Guanhua. The standard for Guanhua became the language spoken by the Manchu officials, and the English term for Guanhua, Mandarin (from the Chinese Man Daren (滿大人) Manchu official), reflects the shifting of the standard.

Not only had Zhang Wei mentioned the local variation in Guanhua, he also noticed that this phenomenon made the communicability between indigenous dialects difficult. This shows that Guanhua and the indigenous dialects respectively belonged to two different strata. The local variation in Guanhua was the result of the localization of Guanhua. In some areas, where the standard pronunciation of Guanhua originated, the difference

between the local variation and the indigenous dialects was not great. In the areas where Guanhua and the indigenous dialects did not belong to the same group, the difference between the two was apparent. Most dialects were, and still are, between these two extremes; therefore strata of different degrees arise. For instance, when the original voiced initial stops of Chinese devoiced during the Tang dynasty (618–907), in words with A (level) tone the initials became aspirated voiceless stops, and in words with B, C and D (oblique) tones the initials became unaspirated voiceless stops. This is a change shared between the official language of the Ming and Qing (1644–1911) dynasties and Putonghua. However, documents also indicate that, in the Guanzhong (關中) area of Shaanxi, towards the end of the Tang dynasty, there was another variation: initials that were originally voiced became aspirated voiceless stops in words of all tones. Dialects with this particular variation can still be found in the Guanzhong area, but such a variation only exists colloquially and has been replaced by the pronunciation of Putonghua in the reading pronunciation.

2 SOME CHARACTERISTICS OF MANDARIN DIALECTS

The differences between Mandarin dialects and other Chinese dialects were described in the last section as well as in other parts of this book and therefore will not be repeated here. Some rare but interesting characteristics of certain Mandarin dialects will be introduced in this section.

- 1 Vowel changes—a change in the vowel of a root occurs in the derivation or inflection of the root. Taking the Huojia (獲嘉) dialect of Henan Province as an example, the word *kai* (蓋) ‘to cover, a cover, coverlet’ has four different readings, *kai*, *kər*, *kio* and *ke*, depending on the derivation or inflection. The form *kai* is the unmarked reading (or reference reading); *kər* is the pronunciation which would be equivalent to adding the diminutive suffix *-er* (兒) to the root in Putonghua; *kio* is equivalent to the reading in Putonghua with the formative suffix *-zi* (子) added to the root; *ke* is the pronunciation when it is used as a verb in the perfective aspect (equivalent to a verb followed by the perfective aspect marker *le* (了) in Putonghua). The Huojia dialect demonstrates the marking of the diminutive of nouns and the perfective aspect of verbs by vowel change instead of suffixation, which is quite different from Putonghua and most of the other Chinese dialects. Other dialects that show these vowel changes are found mostly in the north and southeast of Henan and Shanxi provinces.
- 2 Tone sandhi—in most regions, tone sandhi is a phonology-induced phenomenon. In several Mandarin dialect areas, tone sandhi is, however, grammar induced. Taking the Changzhi (長治) dialect of Shanxi as an example, if a word with B tone is the second element of a compound word (a modifier-head compound, coordinative compound, or verb-complement compound), then both the first and second elements generally show tone sandhi (the first syllable changes to a three-five tone contour and the second changes to a five-three tone contour), while in verb-object constructions the first element shows tone sandhi (changing to 35), while the second remains unchanged. If a two-syllable expression has a C1 (upper departing tone) syllable as the second element and an A1 (upper-level tone), A2 (lower-level tone) or B tone syllable as the first, if it is a compound word the second element shows tone sandhi, while the first element stays the same, but if it is a verb-object construction, neither the first nor the second element show any tone sandhi. These patterns are shown in Table 6.3.

TABLE 6.3 TONE SANDHI IN THE CHANGZHI DIALECT

First syllable	Second syllable	B (rising tone) 535	C1 (upper departing tone) 44			
			Compound words	Verb-object construction	Compound words	Verb-object construction
a1 (upper level tone)	213		35 + 53	35 + 535	213 + 53	213 + 44
a2 (lower level tone)	24		35 + 53	35 + 535	24 + 53	
b (rising tone)	535		35 + 53 (535 + 213)	35 + 535	535 + 53	
c1 (upper departing tone)	44		35 + 53	35 + 535	53 + 44	
c2 (lower departing tone)	53	Verb-object construction Compound words	35 + 53	35 + 535	53 + 44 35 + 44	
d (entering tone)	?54		245 + 53	245 + 535	254 + 53	24 + 44

- 3 Infixed word—infixing is a technique of derivation. In the Pingyao (平遙) dialect of Shanxi a form of derivation is to infix *-aʔl-* between the initial (C) and final (V(E)) of a syllable to make the one-syllable CV(E) word into two syllables: *Caʔl V(E)*. For example, adding the infix to *kay* (桿) ‘pole,’ ‘stick’ produces the infixed word *kaʔ lay* (格懶); adding it to *pay* (拌) ‘to stir’ produces *pəʔ lay* (顆浪). Any monosyllabic word can be infixed using this method and the infixed form can substitute for the monosyllabic form. In fact, many words exist in the form of infixed words, with the original monosyllabic words becoming nearly obsolete. Infixing like this developed from a secret language or the so-called Qie Jiao Ci (切腳詞, words with special spelling), which is a common method of derivation in Shanxi, Shaanxi, and Inner Mongolia.

3 THE REGIONAL CHARACTERISTICS OF MANDARIN DIALECTS

Mandarin dialects can be divided into three regions: Northern Mandarin, Southwestern Mandarin, and Jiang-Huai Mandarin. With Putonghua, the Chengdu dialect and the Nanjing dialect as representatives of these three regions, respectively, the phonetic characteristics of these three regions can be compared as in Table 6.4.

The distributions of these characteristics are not completely separable. As far as the development of the original entering tone (words with *-p*, *-t*, or *-k* finals) is concerned, some of the Northern Mandarin dialects and Southwestern Mandarin dialects still preserve a glottal stop final or retain the entering tone category (even though the consonant final was lost). Other characteristics are more or less not in accordance with each other. On the other hand, whether dividing Mandarin dialects into three regions is appropriate and to what extent sub-divisions of each region are to be made are still controversial questions. In our opinion, basic data on several dialects is still insufficient, and therefore no final conclusion can be made on the classification or subgrouping, or even on the criteria to be used in classification. Consequently this chapter can only introduce a few aspects of the Mandarin dialects.

TABLE 6.4 COMPARISONS OF MANDARIN DIALECTS BY REGION

	Common dialect (Northern Mandarin)	Chengdu (Southwestern Mandarin)	Nanjing (Jiang-Huai Mandarin)
The development of entering tone	Final consonant lost; tone category merged with the other three tone categories	Final consonant lost; tone category merges with A2	-? final with entering tone as a separate tone category
Has retroflex initials <i>tʂ</i> · <i>tʂh</i> · <i>ʂ</i>	Yes	No, merged with dentals	Yes
Has opposition of dental nasal <i>n</i> and dental lateral <i>l</i> ?	Yes	No, <i>l</i> merged as <i>n</i>	No, <i>n</i> merged as <i>l</i>
Has opposition of labio- dental fricative <i>f</i> and velar fricative with <i>u</i> vowel?	Yes	No, <i>x</i> before <i>u</i> pronounced as <i>f</i>	Yes
Has opposition of <i>əŋ</i> · <i>iŋ</i> and <i>ən</i> · <i>in</i> ?	Yes	No, merged as <i>ən</i> · <i>in</i>	No, merged as <i>əŋ</i> · <i>iŋ</i>

NOTE

1 The A, B, C, D after each transcribed word stands for the traditional terms level tone, rising tone, departing tone, and entering tone, respectively. 1 and 2 mean upper or lower variety of the tone, respectively.

REFERENCES

- LaPolla, Randy J. (2001) ‘The role of migration and language contact in the development of the Sino-Tibetan language family.’ In R. M. W. Dixon and A. Y. Aikhenvald (eds), *Areal Diffusion and Genetic Inheritance: Case Studies in Language Change*, 225–254. Oxford: Oxford University Press.
- Mei, Tsu-lin (1997) ‘Hanyu qi ge leixing tezheng de laiyuan (The origin of seven of the typological characteristics of Chinese)’. In Tseng Chiu-yu (ed.) *Language and Linguistics of China*, Vol. 4, 81–104. Taipei: Academia Sinica.

FURTHER READING

- He, Wei (1989) *Jinyu Yanjiu (Studies on the Jin Dialect)*, Tokyo: Tokyo University of Foreign Languages.
- Hou, Jingyi (1987) *Hanyu fangyan wengao ji (Collection of Papers on Chinese Dialects)*, Tokyo: Tokyo University of Foreign Languages.
- Jiang, Shaoyu (1994) *Jindai Hanyu yanjiu gaikuang (Overview of Studies on Modern Chinese)*, Beijing: Beijing University Press.
- Yuan, Jiahua (1980) *Hanyu fangyan gaiyao (Overview of the Chinese Dialects)*, Beijing: Wenzi Gaige Chubanshe.

CHAPTER SEVEN

CANTONESE

Robert S. Bauer and Stephen Matthews

1 INTRODUCTION

The term ‘Cantonese’ originally designated the speech of 廣州 *kwɔŋ³-tsəw¹* Guangzhou, or Canton, the name by which the capital of Guangdong province has been known in the West. Since the 1950s, following the establishment of the People’s Republic of China, regional Chinese varieties such as Cantonese in Guangdong began to fall into decline there in the face of the heavy-handed promotion of Putonghua/Mandarin; as a result, the center of Cantonese language and culture shifted away from Guangzhou to Hong Kong (which borders southern Guangdong) where it is now spoken by 90 percent of the ethnic Chinese population of over six and a half million as their usual, daily language (if those who speak it as an additional variety are also included, then the number rises to 96 percent) (Bauer 2015: 31). In addition, Cantonese has historically dominated many overseas Chinese communities in Europe, North America, and Australasia. The Cantonese associated with the Guangzhou district of 西關 *sɛj¹-kwan¹* (literally, ‘west mountain pass’) has been regarded traditionally as the prestige form. Although both the Hong Kong and Guangzhou varieties are relatively similar, some differences in their pronunciation, vocabulary, and grammar are found. In 1997 after 155 years as a British Crown Colony Hong Kong returned to Chinese sovereignty as a Special Administrative Region of China, and it still retains its status as “the Cantonese-speaking capital of the world” (Bolton 2011: 64). Not only do the vast majority of Hong Kong’s seven million residents speak Cantonese as their usual language variety or as another dialect/language, but it is also spoken as the ordinary, regular, default language in official government settings, the law courts, business offices, radio and television broadcasts, and as the medium of instruction in some schools (where it is on the decline, as about 70 percent of primary schools have switched over to using Putonghua as MoI, according to Tam and Cummins 2015: 23). Today, simply on the basis of its intensive usage across a wide range of social domains, Cantonese-speakers can feel fully justified in regarding Cantonese as Hong Kong’s *de facto* official spoken language (Bauer 2000, 2014, 2015). However, over the past two decades of Hong Kong’s reunification with China, the dramatic yet unsurprising increase in the community’s use of Putonghua—it has become the tourism industry’s lingua franca and the medium of instruction in many schools—has left some observers so alarmed about the future of Cantonese that they have called for its legal preservation (Gallagher 2014) and even predicted its death (Tam and Cummins 2015).

1.1 Names for Cantonese

Cantonese speakers themselves call their speech by a variety of names, e.g. 廣州話 *kwɔŋ³-tsəw¹-wa^{6/3}* ‘speech of Guangzhou,’ 廣東話 *kwɔŋ³-tɔŋ¹-wa^{6/3}* ‘speech of

Guangdong,’ 香港話 *haεŋ^l-kɔŋ³-wa^{6/3}* ‘speech of Hong Kong,’ 白話 *pak⁸-wa^{6/3}* ‘plain speech,’ and 唐話 *tʰɔŋ⁴-wa^{6/3}* ‘speech of the Tang’ (i.e. Tang dynasty which has been regarded by many people as the apex of Chinese civilization). In the IPA transcription system used here a Cantonese tone category is represented by a number at the end of the syllable; tone categories are described in detail in the section on tones. In the popular imagination Cantonese enjoys regional dialect status, even among speakers of other varieties of Chinese, but 普通話 *phow³-thoy¹-wa^{6/3}* ‘Putonghua (literally, common speech),’ which is based on 北方官話 *pek^{7a}-fɔŋ¹ kun¹-wa^{6/3}* ‘northern Mandarin,’ is promoted throughout China as the standard form of Chinese speech, and also as the medium of instruction throughout the schools. In the year 2000 the Chinese government adopted its law on the use of languages and dialects across the nation, i.e. 中華人民共和國國家通用語言文字法 ‘law on the nation’s commonly used spoken language and script of the People’s Republic of China.’ This law makes clear that Putonghua and written Chinese hold the predominant position and have superiority over all other language varieties in China and specifies in detail when Putonghua and standard written Chinese must be used, and when Chinese dialects and minority languages can be used (Zhonghua renmin gong-heguo guojia tongyong yuyan 2001).

Within China, Cantonese is classified as a 方言 *fɔŋ¹-jin²* ‘dialect’ (the less negative, more neutral term *topolect* is also sometimes used (DeFrancis 2003: 244)) of 粵 *jy^r⁸* or Yue, one of the seven (or, depending on one’s criteria, ten) major Chinese dialect families that have been recognized by Chinese dialectologists; however, given the mutual unintelligibility among these so-called ‘dialects,’ Western sinologists have usually regarded Cantonese, along with the other major Chinese dialect families of Gan 贛語, Hakka 客家話, Mandarin 官話, Min 閩語, Wu 吳語, Xiang 湘語, along with two transitional varieties of Hui 徽話 and 平話 Pinghua (Bauer 2012; Yan 2006: 222) as *related Chinese languages* (cf. Li 1994a who regarded Yue as an independent language). The Yue dialects are distributed across most of Guangdong and the eastern part of Guangxi. The worldwide population of speakers of Yue dialects has been numbered to be 62.2 million (Lewis *et al.* 2015).

1.2 Origin of Cantonese

It was during the Qin dynasty (221–206 BCE) that Han Chinese soldiers were sent to the South China region which was called 粵 *jy^r⁸* to subdue and occupy it, thus bringing them into contact with its indigenous inhabitants who were called 百越 or 百粵 *pak^{7b}-jyt⁸* ‘hundred Yue.’ It is believed these peoples very likely spoke languages that are now classified as Austro-Asiatic, Tai-Kadai, Miao-Yao, etc. (Yue-Hashimoto 1991b); today ethnolinguistic groups speaking these languages and referred to as 少數民族 *siw³ sow⁵ men² tsok⁸* ‘minority nationalities’ continue to inhabit various areas of southern China, with the largest group being the 壮 *tsɔŋ¹* ‘Zhuang’ who speak varieties of northern Tai in Guangxi (Holm 2013). Early contact between speakers of Qin-Chinese and indigenous languages not only created the need to communicate, but also must have led to intermarriage among the migrants and local inhabitants; such conditions would have set the stage for the formation of pidgins and creoles which, in our view, developed into Norman’s hypothesized Old Southern Chinese which he claimed was the ancestor of the three main Chinese topolects of Yue, Kejia, and Min (Norman 1988: 210–14). That early contact with indigenous languages has influenced the development of Chinese in South China is supported by the identification of Tai-Kadai and Miao-Yao substrata in contemporary Cantonese (Bauer 1987, 1996; Li 1994b; Yuan 1983; Yue-Hashimoto 1991b).

2 PHONOLOGY

2.1 Initial and final consonants

As indicated in Table 7.1, the phonological inventory of standard Hong Kong Cantonese comprises 19 initial consonants (or 20 if we include the so-called zero-initial or non-contrastive glottal stop which occurs as the onset of a syllable beginning with a vowel). The voiceless stops and affricates contrast in aspiration; voiced consonants include the nasals and approximants. Initial approximants *w-* and *j-* are articulated with some friction. In Hong Kong Cantonese, the series of alveolo-palatal consonants (enclosed in square brackets in the table) occur as non-contrastive allophones before certain high and front round vowels (but for some speakers of Guangzhou Cantonese these palatalized consonants are the regular realizations of the sibilant fricative and affricate initials). Our analysis recognizes the series of labialized velars *kw-* and *khw-* (however, some linguists in China have alternatively treated the labial element as a medial in the rhyme which leads to its skewed and unintuitive distribution). The following nine consonants occur in syllable-final position: *-m*, *-n*, *-ŋ*, *-p*, *-t*, *-k*, *-w*, *-j*, *-ɥ* (the rounded semivowel final *-ɥ* occurs only after the mid central round vowel *ə* and so can be considered the assimilated form of *-j*); the voiceless final stops *-p*, *-t*, *-k* are unreleased.

2.2 Vowels, diphthongs, and rhymes

As indicated in Table 7.2, Cantonese includes 11 vowel phonemes and 14 vowel allophones.

Some vowels are long and some short in duration. Most analyses treat only *e* and *a* as contrasting in length; however, one acoustic study by Li (1985) indicates the length difference is the primary cue for distinguishing *e'/ɛ:* and *o''ɔ:* when they occur before velar consonants *-ŋ*, *-k*.

Vowels preceded by the asterisk in Table 7.3 do not occur as independent rhymes in open syllables. The 60 rhymes in Table 7.3 in turn combine with the 20 initial consonants to form 806 syllables out of the possible total of 1,202 (i.e. 1,200 possible rhymes plus the two nasal syllabics) which occur in the pronunciations of standard Chinese characters, colloquial Cantonese words, and English loanwords (Bauer 2013: 34).

The vowels join with the nine final consonants to form a total of 60 rhymes; this number includes those that occur in syllables that are: (1) associated with the standard

TABLE 7.1 CANTONESE INITIAL CONSONANTS

Manners of articulation	Places of articulation				
	Labial	Dental/alveolar	Post-alveolar/palatal	Velar	Glottal
Stop					
unaspirated	/p/	/t/		/k//kw/	(?)
aspirated	/ph/	/th/		/kh//khw/	
Nasal	/m/	/n/		/ŋ/	
Fricative	/f/	/s/	[ç]		/h/
Affricate					
unaspirated		/ts/	[tç]		
aspirated		/tsh/	[tch]		
Approximant	/w/	/l/	/j/		

reading pronunciations of the standard Chinese characters; (2) colloquial syllables; and (3) English loanword syllables; these 60 rhymes are all listed in Table 7.3 (reproduced from Bauer 2013: 33).

2.3 The Cantonese syllable

The Cantonese syllable takes an elegantly simple and symmetrical structure (there is no medial as in Putonghua): every syllable obligatorily comprises a rhyme (or final) as its nucleus which may be optionally preceded and/or followed by a consonant (consonant clusters with the lateral approximant occur in some colloquial and onomatopoeic expressions under special circumstances, but only for some speakers (Bauer and Benedict 1997: 319–23)); the rhyme may comprise either a nuclear vowel which may be followed by one of the final consonants; or alternatively, the rhyme may be a syllabic nasal consonant *ŋ* or *m*. Overriding every syllable is its contour tone. The structure of the Cantonese syllable is presented in Figure 7.1.

2.3.1 Tones

Like the majority of Sino-Tibetan languages, Cantonese is a tone language in which a change in the pitch of the syllable corresponds to a change in its meaning. The number of contrasting tone contours in Cantonese is either six or seven, or nine or ten or even twelve;

**TABLE 7.2 CANTONESE VOWEL PHONEMES (BETWEEN //)
AND ALLOPHONES (BETWEEN ||)**

	Front	Central	Back
High	/i/ [i:] /y/ [y:]		/u/[u:]
High-mid	/e/ [ɛ], [e̚] /ɛ/ [ɛ̚]	/ə/ [ə] /ɐ/ [ɐ̚]	/o/ [o], [oʷ] /ɔ/ [ɔ̚]
Mid			
Low		/a/ [a:]	

TABLE 7.3 60 CANTONESE RHYMES

-j(-ɥ)	-w	-m	-n	-ŋ	-p	-t	-k
[i:]	[i:w]	[i:m]	[i:n]		[i:p]	[i:t]	
[y:]			[y:n]			[y:t]	
*[e̚]				[e̚ŋ]			[e̚k]
*[e]	[ej]						
[ɛ̚]		[ɛ̚:w]	[ɛ̚:m]	[ɛ̚:n]	[ɛ̚:p]	[ɛ̚:t]	[ɛ̚:k]
[ɛ̚̚]			[ɛ̚̚:m]	[ɛ̚̚:n]		[ɛ̚̚:t]	[ɛ̚̚:k]
*[ə̚]	[ə̚]			[ə̚n]		[ə̚t]	
*[ə̚̚]				[ə̚̚n]			
*[ə̚̚̚]	[ə̚̚̚]						
*[ə̚̚̚̚]	[ə̚̚̚̚]						
[a:]	[a:j]	[a:w]	[a:m]	[a:n]	[a:p]	[a:t]	[a:k]
[a:̚]				[a:̚n]		[a:̚t]	
[u:]	[u:j]			[u:n]		[u:t]	
*[oʷ]				[oʷŋ]			[oʷk]
*[o̚]		[ow]					
[ɔ̚̚]	[ɔ̚̚j]		[ɔ̚̚m]	[ɔ̚̚n]	[ɔ̚̚ŋ]	[ɔ̚̚t]	[ɔ̚̚k]

<i>SUPRASEGMENTAL</i> 超音段		
TONE 聲調		
ONSET	RIME 韵母	FINAL F 韵
<i>INITIAL</i> Initial Consonant C_i 聲母	<i>NUCLEUS</i> 核心元音..... Nuclear Vowel V 韵腹, 主要元音 Syllabic Nasal C_{sn} 成音節鼻音	<i>CODA</i> Ending Consonant C_e/G_e 韵尾輔音 韵尾滑音

FIGURE 7.1 STRUCTURE OF THE CANTONESE SYLLABLE

it depends on whether one is talking about Hong Kong or Guangzhou Cantonese, whether one regards the tone contours (which can be level, rising, falling) on syllables with finals *-p*, *-t*, *-k* as separate tones which are distinct from the tones that occur on open syllables and syllables closed by *-m*, *-n*, *-ŋ*, and whether or not one counts the two so-called *changed tones* or 變音 *pin⁵-jem¹* (with tone contours that are identical to two other regular ones) used in word derivation as separate categories. The historical devoicing of originally voiced initial obstruents in Ancient Chinese was the catalyst for doubling the number of tone categories in some dialects; in Modern Cantonese the upper 陰 *jem¹* and lower 陽 *jœy²* tone registers neatly correspond to the historical four-tone category system with its voiced and voiceless initials; that is, the historically voiced initial consonants now have the low register tones of Mid-Low Falling 陽平 *jœy²-pey²*, Mid-Low Rising 陽上 *jœy²-sæy⁴*, Mid-Low Level 陽去 *jœy²-hey⁵*, and Mid-Low Stopped 陽入 *jœy²-jep⁸*.

One interesting (but disappearing) difference between the phonologies of Hong Kong and Guangzhou Cantonese can be found in their tone systems: Some older, educated speakers of Guangzhou Cantonese have undergone the phonemic split of Tone 1 陰平 *jem¹-pey²* by distinguishing between two tone contours: the High Falling contour 上陰平 *sæy⁴-jem¹-pey²* or Tone 1a functions as their usual, default contour which is represented by tone letter and value 152; the High Level contour 下陰平 *ha⁶-jem¹-pey²* or Tone 1b is represented with tone letter and value 155 and functions for these speakers as a so-called *changed tone* or 變音 *pin⁵-jem¹*, i.e. a morphological device to derive certain types of concrete nouns (Zong 1964). Table 7.4 has recognized this tonal split in Guangzhou Cantonese by enclosing the examples 醫 *ji^{1a}* 152 ‘to cure’ and 衣 *ji^{1b}* 155 ‘clothes’ in parentheses. Younger Guangzhou speakers and Hong Kong Cantonese speakers, on the other hand, do not make this tonal distinction, but only have the High Level tone contour which has both lexical and morphological functions as explained in the following section (Bauer and Benedict 1997: 131–2; Zee 1999: 59). Table 7.4 illustrates the ten tone contours with their corresponding Chao tone letters.

2.3.2 *Changed tones*

In addition to the tones listed above, there are two changed tones or 變音 *pin⁵-jem¹* which are used in word derivation; the contours of the two changed tones are identical to those of the regular High Level and High Rising tones, respectively; however, in order to make this morphological change transparent, special symbols are used to here indicate the two changed tones; e.g. the Mid-Low Falling tone on the second syllable of 阿姨 *a-ji²¹*

TABLE 7.4 LEXICAL ITEMS CONTRASTING NINE (TEN) TONE CONTOURS ON OPEN SYLLABLE *JI*

Lexical item	Tone contour	Tone letter	Tone category	English name
1 <i>ji¹</i> 醫 ‘to cure’ (<i>ji^{1a}</i> 醫 ‘to cure’)	High Level (High Falling)	˥˥ (˥˥)	陰平 (上陰平)	Upper Even (High Upper Even)
<i>ji¹</i> 衣 ‘clothes’ (<i>ji^{1b}</i> 衣 ‘clothes’)	High Level (High Level)	˥˥ (˥˥)	陰平 (下陰平)	Upper Even (Low Upper Even)
2 <i>ji²</i> 疑 ‘suspicious’	Mid-Low Falling	˨˩	陽平	Lower Even
3 <i>ji³</i> 椅 ‘chair’	High Rising	˥˨	陰上	Upper Rising
4 <i>ji⁴</i> 耳 ‘ear’	Mid-Low Rising	˧˧	陽上	Lower Rising
5 <i>ji⁵</i> 意 ‘idea’	Mid Level	˧˧	陰去	Upper Going
6 <i>ji⁶</i> 二 ‘two’	Mid-Low Level	˧˨	陽去	Lower Going
7a <i>je^ka</i> 益 ‘benefit’	High Stopped	˥	上陰人	High Upper Entering
7b <i>ja^kb</i> 噥 ‘to eat’	Mid Stopped	˧˧	下陰人	Low Upper Entering
8 <i>je^ks</i> 亦 ‘also’	Mid-Low Stopped	˧˨	陽人	Lower Entering

‘auntie’ typically changes to the High Level *pin⁵-jem¹* which is marked with the small raised circle; the Mid-Low Level tone of the second syllable of 阿二 *a⁵-ji^{6/3}* ‘Number Two servant’ changes to the High Rising *pin⁵-jem¹* and is marked with the asterisk. As noted by Chao (1947: 34) some years ago, both of these changed tones carry “a morphological meaning, namely, that familiar thing (or person, less frequently action) one often speaks of.” The High Level changed tone occurs on some nouns and stative verbs, e.g. 金毛 *kem'mow^{2/1}* ‘person with blond or bleached blond hair,’ 靚仔 *lej⁵-tsəj³* ‘handsome boy’ => 靚仔 *lej^{5/1}-tsəj³* ‘good-for-nothing young guy.’ The High Rising changed tone functions as an especially productive morphological device (Bauer and Cheung 2005), such as changing verbs to nouns, e.g. 掃 *sow⁵* ‘to sweep’ => *sow^{5/3}* ‘broom, brush, duster’; 帶 *ta:j⁵* ‘to guide, lead’ => *ta:j^{5/3}* ‘belt; band; ribbon’; 袋 *tɔ:j⁶* ‘to put into a pocket or bag; to pocket’ => *tɔ:j^{6/3}* ‘bag; pocket.’

2.3.3 Split of Yin Ru

As noted by Norman (1988: 217–18), one distinctive, defininitive characteristic of many (but not all) Yue dialects has been the split of the 陰人 *jem¹* *je^ks⁸* tone category into two subcategories of 上下陰入 *səŋ⁴* *ha⁶* *jem¹* *je^ks⁸*, with this development having been conditioned by vowel length for the standard reading pronunciations of the standard Chinese characters (with only a very few exceptions; in addition, the colloquial pronunciations of some lexical items do not follow this generalization); e.g. the standard reading pronunciations of the standard Chinese characters belonging to 上陰人 *səŋ⁴*

jem¹ jep⁸ carry the high stopped tone which co-occurs with a short vowel; e.g., 态 *hep^{7a}* ‘exactly,’ 乞 *het^{7a}* ‘to beg,’ 黑 *hek^{7a}* ‘black’ (but we note that this latter character’s colloquial pronunciation *hak^{7a}* with long vowel also quite commonly occurs in such colloquial expressions as 黑客 *hak^{7a} mej¹ mej¹* ‘pitch dark,’ 夜晚黑 *je⁶ man^{4/1} hak^{7a}* ‘at night’); in contrast, the standard reading pronunciations of the standard Chinese characters belonging to 下陰人 *ha⁶ jem¹ jep⁸* carry the mid-stopped tone which co-occurs with a long vowel; e.g. 挤 *hap^{7b}* ‘to pinch,’ 犬 *sat^{7b}* ‘to kill,’ 客 *hak^{7b}* ‘guest’ (Bauer and Benedict 1997: 159–62). Given the early historical contact in South China between Proto-Yue and the Tai languages which also contrast long and short vowels in so-called “dead” syllables (i.e. those with final stop consonants *-p*, *-t*, *-k*), this distinctive development in Yue may very well represent the trace of an early Tai substratum (personal communication from the late Danish sinologist Søren Egerod to Bauer in 1994; Bauer 1996).

2.4 Phonetic variation

Like some other southern Chinese dialects Cantonese exhibits the phenomenon of variation between literary and colloquial pronunciations of some lexical items. In Cantonese the literary form typically has a short vowel while the colloquial form has a long vowel; e.g. 生 *sey¹*, *say¹* ‘life,’ 耕 *kay¹*, *kay¹* ‘plough,’ 行 *hey²*, *hay²* ‘walk,’ 爭 *tsøy¹*, *tsay¹* ‘to quarrel,’ 平 *phey²*, *phey²* ‘flat, cheap,’ 瓶 *phey²*, *phey²* ‘vase,’ 頂 *tey³*, *tey³* ‘top, extremity,’ *then¹*, *they¹* ‘to hear, listen’ (Bauer 2013). In addition to differences in vowels, there are some literary vs colloquial readings that show alternation in their initial consonants and tones; for some characters that historically belonged to the 陽去 *jøy² hoy⁵* tone category, the literary reading is now with a syllable that has an unaspirated initial and Low Level tone, while there is also a colloquial reading with an aspirated initial and Lower Rising tone: 伴 *pun⁶* ‘to accompany’ => *phun⁵* ‘companion,’ 斷 *tyun⁶* => *thyun⁵* ‘to cut off, break off,’ 近 *ken⁶* => *khen⁵* ‘to be near to.’

Hong Kong Cantonese is also characterized by sociolinguistic variation that is associated with several ongoing sound changes, some of which have been phonetically conditioned. Over the past 20 years or so, several sociolinguistic studies (reviewed in Bauer and Benedict 1997: 327–42) have investigated the following sound changes: *y* → *m*; *kw-* → *k-* and *khw-* → *kh-* (before back mid round vowel ɔ); *khæy⁵* → *hæy⁵* (only in third person pronoun 佢), *y-* → *O-*; *n-* → *l-*; *-vŋ* → *-vn*, *-vk* → *-vt* (Bauer 1979, 1983, 1986; Law *et al.* 2001; To *et al.* 2015). Rejecting the sociolinguistic explanation of variation and change, some self-appointed language experts have claimed that younger speakers are simply using the so-called 懶音 *lan⁴ jem¹* ‘lazy pronunciation’ which is another way of saying they are simply following the phonetician’s Principle of Least Effort.

3 CANTONESE LEXICON AND SCRIPT

3.1 Colloquial Cantonese vocabulary

Colloquial Cantonese speech includes many vocabulary items which are etymologically unrelated to their semantic equivalents in standard Chinese; some of these words have dialectal characters as their written forms, e.g. 冇 *mow⁴* ‘not have; no’ 馄 *pew⁶* ‘jostle with hips,’ 蹤 *phe⁴* ‘stagger,’ 揽 *tem⁵* ‘droop, hang down,’ 揽 *tep⁸* ‘beat, pound,’ 咩 *them⁴*

‘puddle,’ 掉 *tew*⁶ ‘throw away,’ 挿 *tcŋ*⁵ ‘throw (at target),’ 佢 *khoŋ*⁴ ‘he, she, it,’ 扇 *khe*¹ ‘shit,’ 甲 𠵼 *kat*⁸-*tsat*^{8/3} ‘cockroach,’ 嘛 *khwak*^{7a} ‘loop,’ ‘circle,’ 脍 *nem*² ‘soft, tender,’ 唔 *jam*¹ ‘all right, good,’ 蹤 *jan*⁵ ‘to kick off’ (Cheung and Bauer 2002). In addition, written Cantonese has retained a few ancient characters which are not used in standard Chinese; e.g. 爾 *pej*^{3/2} ‘to give’ dates from the Zhou dynasty 900–700 BCE (Karlgren 1957: 141). Some lexical items are written with standard Chinese characters that have been borrowed solely for their pronunciations to write homophonous colloquial words, e.g. 呢 *ni*¹ ‘this.’ The mouth radical is added to standard characters to create additional characters to semantically unrelated but homophonous colloquial words, e.g. 啟 *tsɔ*² ‘marker of completed action,’ 嚟 *lɛŋ*⁴ ‘to come,’ 嘅 *kɔ*² ‘that.’ At the same time, other morphosyllables may have no written forms at all and so are usually transcribed with the empty box □, e.g. □ *fak*^{7b} ‘to whisk,’ □ *he*⁵ ‘to hang out, idle away one’s time,’ □□聲 *wiw*¹ *wiw*¹ *sey*¹ ‘sound of siren.’ Some lexical items such as these may represent the trace of an ancient non-Han, viz. Austro-Tai and/or Austro-Asiatic, substratum that reflects the historical contact Cantonese and the Yue dialect family have had with languages of these families (Bauer 1987, 1996; Li 1994b; Yuan 1983; Yue-Hashimoto 1991b). In rapid speech some high-frequency two-syllable expressions can contract into monosyllables; e.g. 唔好 *m*² *how*³ ‘don’t => 唔 *mow*², 係唔係 *hej*⁶ *m*² *hej*⁶ ‘is that right?’ => 係咪 *hej*⁶ *mej*⁶.

3.2 English loanwords

No Chinese dialect has been more influenced by a European language than Cantonese; as a result of its long, intimate historical contact with English, which dates back 300 years, a large number of English words have been borrowed into Hong Kong and Guangzhou Cantonese varieties. Many of these loanwords have become so assimilated that they have written forms with Chinese characters (Bauer 2010), and many Cantonese speakers may assume that they are actually Chinese words; e.g. 巴士 *pa*¹-*si*^{6/3} ‘bus,’ 芝士 *tsi*¹-*si*^{6/3} ‘cheese,’ 多士 *tsɔ*¹-*si*^{6/3} ‘toast,’ 唔 *khat*^{7a} ‘card,’ 菲林 *fej*¹-*lem*^{4/3} ‘film,’ 貼士 *thip*^{7a}-*si*^{6/3} ‘tips’ (Bauer 2003, 2006a; Wong *et al.* 2009). In the case of 較 *lip*^{7a} ‘elevator (originally borrowed from British English ‘lift’)’ a new Cantonese character was created to write this loanword (Cheung and Bauer 2002: 122). At the same time, however, there are some loanwords that are pronounced with Cantonese syllables but retain their original English spellings if they appear in written texts, e.g. *DOWNLOAD* => *tan*¹ *low*¹, *PROJECTOR* => *p̥low*⁵ *tsek*^{7a} *t̥a*³, *WARM* => *wɔm*¹. The phonetic adaptation of loanwords through the recombination of existing initial consonants, nuclear vowels, and final consonants has influenced the development of the Cantonese syllabary through the introduction of new syllables that only occur in English loanwords (Bauer 2003; Bauer and Wong 2010). Loanwords have become intimately integrated into the Cantonese grammar (Wong *et al.* 2009), as shown by the following example sentence in which the English loanword *man* ‘manly’ has been used as a stative verb and to which has been attached 啟 *tsɔ*³, the colloquial Cantonese aspect marker of completed action: 你好似 MAN 啟好多! *nej*⁵ *how*³ *tshir*⁵ *men*¹ *tsɔ*³ *how*³ *ts*¹ ‘You seem to have become much more manly!’ (as said by a young woman to her boyfriend).

3.3 Written Cantonese

Hong Kong Cantonese uniquely distinguishes itself from all other regional Chinese varieties through the informal development of its written form in which there is a one-to-one correspondence between words in the colloquial spoken language and their transcription through the combination of standard and non-standard (dialectal) Chinese characters and

letters of the English alphabet (while it is the case that Taiwanese is written on Taiwan, the scale on which this is being done is much smaller in comparison). Writing in Cantonese has become a pervasive phenomenon in Hong Kong newspapers, magazines, comic books, personal correspondence, subway advertisements, government posters, etc.; over time there has been an accumulation of essentially *ad hoc* conventions for writing Cantonese, but these have never been fully standardized, so they are still inconsistent, and it is not unusual to find the same word written with two or more different characters; e.g. colloquial *kem*³ ‘like this, in this way, so’ may be written with both 噉 and 唉 (Bauer 1982, 1984, 1988, 2006b; Cheung and Bauer 2002; Snow 2004). The major reason Cantonese speakers write in Cantonese is simply to transcribe exactly what they say; this has been aptly expressed in the traditional folk saying 我手寫我口 *ŋɔ⁴ səw³ se⁵ ŋɔ⁴ hew³* ‘my hand writes my mouth’ (which was a slogan heard during the May Fourth Movement of 1919 advocating the use of the vernacular in writing in place of 文言文 ‘classical Chinese’). As noted in Cheung and Bauer (2002: 4), “[t]oday writing in Cantonese is perceived by writers and readers as conveying the writer’s message with a greater degree of informality, directness, intimacy, friendliness, casualness, freedom, modernity, and authenticity than writing it in standard Chinese, which is the formal language the Hong Kong Cantonese speaker learns to read and write in school, but its spoken counterpart s/he does not ordinarily use when speaking with coworkers, friends, and family members.” Underpinning the extensive use of written Cantonese in Hong Kong has been the tradition of teaching schoolchildren to read the standard Chinese characters with standard Cantonese pronunciation (in contrast, schoolchildren in Guangzhou and other traditionally Cantonese-speaking areas of the mainland learn to read in Putonghua); unfortunately, this tradition is now threatened as more and more Hong Kong schools switch over to Putonghua as the medium of instruction (giving rise to the 反普教中 *fan³ pʰow³ kaw⁵ tsɔŋ¹* ‘movement opposing Putonghua to teach Chinese-language subjects’). As noted in section 3 ‘Colloquial Cantonese vocabulary,’ many colloquial Cantonese words are not etymologically related to their semantic and functional equivalents in standard Chinese, so these words may not have Chinese characters associated with them as their written forms; as a consequence, Cantonese writers have had to create characters to transcribe such words, as indicated in the following example sentence in which all the characters are dialectal: 佢嚟咗搵乜嘢? *kʰəu⁵ tej⁶ lej² tsɔ² wən³ met^{7a} je⁴* ‘What have they come looking for?’.

4 GRAMMAR

Chinese grammar has often been described as rather uniform across dialects (by Chao 1968, among others); this is so in the sense that, for example, a Cantonese counterpart exists for most Mandarin structures. However, the extent and significance of the differences are also increasingly recognized (Yue-Hashimoto 1993 and this volume; Lucas and Xie 1994; Matthews 1999). Cantonese appears especially distinctive in this respect and studies focusing on the description of Cantonese grammar have borne this out (Cheung 2007; Matthews and Yip 2011; Tang and Cheng 2014). An important characteristic of Cantonese grammar (and that of other Sinitic languages) that has emerged from investigations of these differences is the co-existence of grammatical options which Yue-Hashimoto (1991a, 1993) has termed stratification. In such cases a syntactic structure shared with Mandarin is used in H[igh] register, while L[ow] language uses a more distinctively Cantonese structure as indicated in Table 7.5.

The passive construction with colloquial *pej*³ 倂 superficially resembles Mandarin *bèi* 被, but its origin as a grammaticalized form of the verb *pej*³ 倂 ‘give’ is reflected in

TABLE 7.5 REGISTER STRATIFICATION IN CANTONESE SYNTAX

	High Cantonese	Low Cantonese
Passive voice	<i>pej⁶</i> 被 (NP) V	<i>pej³</i> 倘 NP V
Comparative degree	<i>pej³</i> 比 NP ₁ Adj NP ₂	NP ₁ Adj <i>kwo⁵</i> 過 NP ₂
Excessive degree (too)	<i>thaj⁵</i> 太 Adj	Adj <i>kwo⁵</i> <i>thew²</i> 過頭, Adj <i>tek^{7a}a-tsej⁶</i> 得滯
Possessive marker	NP <i>ke⁵</i> 嘅 N	NP Classifier N
Relative clause	[S . . .] <i>ke⁵</i> 嘱 ^N	[S . . .] <i>kɔ³</i> 嘱 Classifier N

syntactic and semantic differences. In particular, the Cantonese passive with *pej³* 倩 requires an agent phrase. In the following sentence *jen²* 人 ‘person’ is supplied as the generic agent (PERF indicates the perfective aspect marker).

- (1) *khoj⁴ pej³ jen² laj¹ tsɔ³*
佢 倩 人 拉 啟
3sg by person arrest- PERF
'S/he was arrested by someone.'

4.1 Comparative constructions

Among excessive constructions, for example, *thaj⁵ jit⁸* 太熱 ‘too hot’ corresponds to Mandarin *tài rè* and is used in neutral or formal register, while *jit⁸ kwo⁵ thew²* 热過頭 and *jit⁸ tek^{7a} tsej⁶* 热得滯 are more colloquial. There is evidence for change from the indigenous toward the pan-Chinese model: Adj. + *kwo⁵* *thew²* is described in older grammars, such as O'Melia (1941) and is now used mostly by older Hong Kong speakers. Such differences between registers often have typological significance. Having recognized this stratification, we see that in terms of typology the excessive constructions peculiar to Cantonese are head-initial, while the pan-Chinese one with *thaj⁵* is head-final. This applies equally to comparative structures with *kwo⁵* 過 and *pej³* 比 respectively head-initial and head-final:

- (2) *khoj⁴ lek^{7a} kwo⁵ nyɔ⁴*
佢 吠 過 我
3sg smart -er 1sg
'She's smarter than I.'
- (3) *khoj⁴ pej³ nyɔ⁴ tshoy¹ men²*
佢 比 我 聰明
3sg compared 1sg intelligent
'She's more intelligent than I am.'

This is part of the North–South cline within Sinitic, as pointed out by Hashimoto (1983), whereby head-final structures in the north give way to head-initial ones in the south. The proximity of Mandarin and Cantonese to Altaic and Tai languages, respectively, suggests areal diffusion as the source of this difference; for example, the Cantonese comparative marker *kwo⁵* 過 appears to be cognate with Thai *kwàː*, Lao *kua*, etc. In typological terms these are all examples of the ‘exceed’ type of comparative in which the standard NP is treated as the object of a transitive verb meaning ‘to exceed’ or ‘to surpass’ (Ansaldi 2010).

4.2 Noun phrase structure

Despite the above-mentioned tendency towards head-initial typology, the noun phrase is rather strictly head-final. Even relative clauses precede the noun they modify, as in other Sinitic languages. Two types of relative clause may be distinguished (Killingley 1993; Matthews and Yip 2001, 2011): one with the particle *ke⁵* 嘅 corresponding to Mandarin *dè* 的 as in (4), and the other using a demonstrative and classifier as in (5):

- (4) *khəy⁴ tshəy⁵ ke⁵ kɔ¹ how³ sem¹ ke⁵*
 佢 唱 嘅 歌 好 深 嘅
 3sg sing PRT song very deep PRT
 ‘The song(s) she sings is/are rather deep.’
- (5) *khəy⁴ tshəy⁵ kɔ³ səw³ kɔ¹ how³ lan² tshəy⁵ ke⁵*
 佢 唱 嘿 首 歌 好 難 唱 嘅
 3sg sing that CL song very hard sing PRT
 ‘The song she sings is hard to sing.’

The classifier construction has a specific reference and belongs to the colloquial register. Similarly, possessive constructions use either *ke⁵* as in (6) or a classifier (7) to link the possessor and possessed noun:

- (6) *khəy⁴ ke⁵ thej³fat^{7b} thaj⁵ pej¹kun¹ ke⁵*
 佢 嘅 睇法 太 悲觀 嘅
 3sg POSS attitude too pessimistic PRT
 ‘Her attitude is too pessimistic.’
- (7) *khəy⁴ kan¹ ok^{7a} how³ taj⁶ kan¹*
 佢 間 屋 好 大 間
 3sg CL house very big CL
 ‘Her house is very big.’

In addition to the possessive use of the classifier in *khəy⁴ kan¹ ok^{7a}*, we should note here the use of the classifier following the adjective *taj⁶*. In the syntax of classifiers, as in some other respects, Cantonese resembles the neighbouring Tai and Hmong-Mien languages: for example, the classifier possessive construction illustrated above is shared with Hmong. Another such feature is the ‘bare classifier’ construction in which the classifier serves like a determiner to specify that the noun has definite reference:

- (8) *kan¹ ok^{7a} how³ taj⁶*
 間 屋 好 大
 CL house very big
 ‘The house is very big.’

This is not a general Sinitic feature but is characteristic of Yue dialects and others (such as the Min dialect of Chaozhou) which have been in contact with them.

4.3 Verb phrase

Some word order patterns in the verb phrase are distinctively Cantonese. A small number of adverbs, notably *sin¹* 先 ‘first’ and *thim¹* 添 ‘additionally’ (cf. *kwɔ⁵ thew²* and *tek^{7a} tsej⁶* above), follow the verb:

- (9) *ŋɔ⁴-tej⁶ tsəw³ sin¹*

我哋 走 先
1pl go first
'We're leaving now.'

This pattern has been attributed to substrate influence from Tai languages (Lucas and Xie 1994; Matthews 2006). In the double object construction the indirect object of *pej*³ 'give' follows the direct:

- (10) *ηɔ⁴ pej³ tshin^{2/3} lej⁴*
我 傪 錢 你
1sg give money 2sg
'I'm giving you the money.'

This typologically unusual ordering occurs in several southern Chinese dialects. However, the reverse order does occur, especially when the direct object is longer than the indirect or is to be emphasized:

- (11) *ηɔ⁴ pej³ lej⁴ kem⁵ tɔ¹ tshin^{2*}, lej⁴ tow¹ m² jiw⁵?*
我 傪 你 啥 多 錢 你 都 唔 要
1sg give 2sg so much money 2sg all not want
'I give you so much money (and) you don't want it?'

The pretransitive or 'disposal' construction uses *tsəey*¹ 將 (OM, object marker; cognate with Mandarin *jiāng*). Like Mandarin *bā* 把, this serves to place a definite or specific object before the verb, but the construction with *tsəey*¹ is more restricted in function, typically retaining a sense of displacement:

- (12) *ηɔ⁴ tsəey¹ ti¹ sam¹ paj³ hej³ lej⁴ tow⁶*
我 將 D 衫 擺 嘴 你 度
1sg OM PL clothes put LOC 2sg there (PL = plural marker
for following noun)
'I'm putting the clothes in your place/room.'

4.4 Utterance-final particles

The utterance-final particles (PRT) of Cantonese are particularly rich both in sheer number and in their combinatorial possibilities. Some 30 basic particles have been identified (Kwok 1984), and most of them occur in several variant forms and combinations (Bauer and Benedict 1997: 291–5; Leung 2005; Yao 1980). One example involves a set of particles having evidential functions, including *wɔ⁵* denoting surprising or notable information (Luke 1990; Matthews 1998a, 1998b; Leung 2010), as in (13), and *wɔ⁴* which indicates 'hearsay' evidentiality as in (14):

- (13) *khoi⁴-tej⁶ jaw⁶ pun¹ ok^{7a} wɔ⁵*
佢 嘴 又 搬 屋 喎
3pl again move house PRT
'You know, they're moving house again (believe it or not).'
(14) *khoi⁴-tej⁶ jaw⁶ pun¹ ok^{7a} wɔ⁴*
佢 嘴 又 搬 屋 喎
3pl again move house PRT
'They're moving house again (they say).'

5 CONCLUSION

Cantonese is not Putonghua wearing the exotic phonological garb of southern China, although it is sometimes treated as if it were. Forcing Cantonese through the standard Chinese character filter has the unfortunate effect of lopping off many of its most distinctive and interesting features. This brief sketch of Cantonese phonology, vocabulary, and syntax has highlighted some of these features which are shared with other languages of Southeast Asia.

REFERENCES

- Ansaldi, Umberto (2010) ‘Surpass comparatives in Sinitic and beyond: typology and grammaticalization’, *Linguistics* 48.4: 919–950.
- Bauer, Robert S. (1979) ‘Alveolarization in Cantonese: a case of lexical diffusion’, *Journal of Chinese Linguistics* 7.1: 132–141.
- Bauer, Robert S. (1982) ‘D for two in Cantonese’, *Journal of Chinese Linguistics* 10.2: 277–281.
- Bauer, Robert S. (1983) ‘Cantonese sound change across subgroups of the Hong Kong speech community’, *Journal of Chinese Linguistics* 11.2: 303–356.
- Bauer, Robert S. (1984) ‘The Hong Kong Cantonese speech community’, *Cahiers de Linguistique Asie Orientale* 13.1: 57–90.
- Bauer, Robert S. (1986) ‘The microhistory of a sound change in progress in Hong Kong Cantonese’, *Journal of Chinese Linguistics* 14.1: 1–42.
- Bauer, Robert S. (1987) ‘Kadai loanwords in southern Chinese dialects’, *Transactions of the International Conference of Orientalists in Japan* 32: 95–111.
- Bauer, Robert S. (1988) ‘Written Cantonese of Hong Kong’, *Cahiers de Linguistique Asie Orientale* 17.2: 245–293.
- Bauer, Robert S. (1996) ‘Identifying the Tai substratum in Cantonese’, *Proceedings of the Fourth International Symposium on Languages and Linguistics, Pan-Asiatic Linguistics* V. Bangkok: Institute of Language and Culture for Rural Development, Mahidol University, pp. 1806–1844.
- Bauer, Robert S. (2000) ‘Hong Kong Cantonese and the road ahead’, in David C. S. Li, Angel Lin, and Wai King Tsang, eds, *Language and Education in Postcolonial Hong Kong*. Hong Kong: Linguistic Society of Hong Kong, pp. 35–58.
- Bauer, Robert S. (2003) ‘The impact of English loanwords on the Cantonese syllabary’, in David Bradley, Randy LaPolla, Graham Thurgood, and Boyd Michailovsky, eds, *Language Variation: Papers on Variation and Change in the Sinosphere and in the Indo-sphere in Honour of James A. Matisoff*. Canberra: Pacific Linguistics, pp. 275–284.
- Bauer, Robert S. (2006a) ‘The stratification of English loanwords in Cantonese’, *Journal of Chinese Linguistics* 34.2: 172–191.
- Bauer, Robert S. (2006b) ‘The Cantonese script’, in Françoise Bottéro and Redouane Djamouri, eds, *Écriture chinoise: Données, usages et représentations*. Paris: Centre de Recherches Linguistiques sur l’Asie Orientale, École des Hautes Études en Sciences Sociales, pp. 167–183.
- Bauer, Robert S. (2010) ‘The graphemic representation of English loanwords in Cantonese’, in 張洪年 Hung-nin Samuel Cheung and 張雙慶 Song Hing Chang, 主編 eds, 《歷時演變與語言接觸 — 中國東南方言》 [Diachronic Change and Language Contact—Dialects in South East China. *Journal of Chinese Linguistics Monograph Series Number 24*]. Hong Kong: Chinese University Press of Hong Kong, pp. 227–246.

- Bauer, Robert S. (2012) 'Review of *Introduction to Chinese Dialectology* by Margaret Mian Yan,' *Journal of Chinese Linguistics* 40.2: 478–497.
- Bauer, Robert S. (2013) 'Phonetic features of colloquial Cantonese', in Peng Gang 彭剛 and Shi Feng 石鋒, eds, *Eastward Flows the Great River; Festschrift in Honor of Professor William S.-Y. Wang on his 80th Birthday, Essays on Contemporary Linguistics and Language Studies*. Hong Kong: City University of Hong Kong Press, pp. 29–44.
- Bauer, Robert S. (2014) 'Le cantonais de Hong Kong, État des lieux et perspectives', *Critique* 807–808: 755–767. Edited and translated into French by Sebastian Veg.
- Bauer, Robert S. (2015) 'The Hong Kong speech community's Cantonese and other languages', *Global Chinese* 1.1: 27–55.
- Bauer, Robert S. and Benedict, Paul K. (1997) *Modern Cantonese Phonology*, 《摩登廣州話語音學》, Berlin: Mouton de Gruyter.
- Bauer, Robert S. and Cheung, K.H. (2005) 'Database of Cantonese words derived through diminutive bianyin 粵語小稱變音詞語派生的詞彙', in Deng Jingbin 鄧景濱, 主編 ed., 《第九屆國際粵方言研討會論文集》 [Proceedings from the 9th International Yue Dialects Conference (2003)], 澳門: 澳門中國語文學會 [Macau: Macau Chinese Language Association], pp. 174–179.
- Bauer, Robert S. and Wong, Cathy S.P. (2010) 'New loanword rimes and syllables in Hong Kong Cantonese', in Pan Wuyun and 沈鐘偉 Shen Zhongwei, eds, 《研究之樂, 慶祝王士元先生七十五壽辰學術論文集》 [The Joy of Research II, A Festschrift in Honor of Professor William S.-Y. Wang on His Seventy-fifth Birthday], 上海: 上海教育出版社 [Shanghai: Shanghai Education Publishing], pp. 1–24.
- Bolton, Kingsley (2011) 'Language policy and planning in Hong Kong: colonial and post-colonial perspectives', *Applied Linguistics Review* 2: 51–74.
- Chao, Yuen Ren (1947) *Cantonese Primer*. Cambridge, MA: Harvard University Press.
- Chao, Yuen Ren (1968) *A Grammar of Spoken Chinese*, Berkeley: University of California Press.
- Cheung, Kwan-hin and Bauer, Robert S. (2002) *The Representation of Cantonese with Chinese Characters*, *Journal of Chinese Linguistics*, Monograph Series Number 18, Berkeley, CA: University of California Berkeley.
- Cheung, Samuel H.N. 張洪年 (2007) 《香港粵語語法研究》, Hong Kong: Chinese University Press.
- DeFrancis, John (2003) 《ABC 漢英大詞典 ABC Chinese–English Comprehensive Dictionary》, 上海: 世紀出版集團和漢語大詞典出版社.
- Gallagher, Steven (2014) Both city and nation must preserve Cantonese language. *South China Morning Post*. 29 April, C2.
- Hashimoto, Mantaro 橋本太郎 (1983) 《語顏地里類型學》, Beijing: Beijing daxue chubanshe.
- Holm, David (2013) *Mapping the Old Zhuang Character Script: A Vernacular Writing System from Southern China*, Leiden: Brill.
- Karlgren, Bernard (1957) *Grammata Serica Recensa*, Bulletin 29. Stockholm Museum of Far Eastern Antiquities.
- Killingley, Siew-Yue (1993) *Cantonese, Languages of the World* 06, Munich: Lincom Europa.
- Kwok, Helen (1984) *Sentence Particles in Cantonese*, Hong Kong: Center of Asian Studies, University of Hong Kong.
- Law Sam Po, Fung, Roxana S.Y. and Bauer, Robert S. (2001) 'Perception and production of Cantonese consonant endings', *Asia Pacific Journal of Speech, Language and Hearing* 6: 179–195.

- Leung, Chung Sam (2005) *A Study of the Utterance Particles in Cantonese as Spoken in Hong Kong*, Hong Kong: Language Information Sciences Research Centre, City University of Hong Kong.
- Leung, Wai Mun (2010) *On the Synchrony and Diachrony of Sentence-final Particles: The Case of WO in Cantonese*, Saarbrücken: Lambert Academic Publishing.
- Lewis, M. Paul, Simons, Gary F. and Fennig, Charles D. (eds) (2015) *Ethnologue: Languages of the World* (18th edition), Dallas, TX: SIL International. Retrieved from: www.ethnologue.com (accessed 9 January 2016).
- Li, Jingzhong 李敬忠 (1994a) ‘粵語是漢語族群中的獨立語言’ [Yue language is an independent language within the Chinese language family], 《語言演變論》, 廣州: 廣州出版社, pp. 79–115.
- Li, Jingzhong 李敬忠 (1994b) ‘粵語中的百越語成分問題’ [The problem of Bai Yue language elements in Yue language], 《語言演變論》, 廣州: 廣州出版社, pp. 116–134.
- Li, Xingde 李行德 (1985) ‘廣州話元音的音值暨長短對立’ [Phonetic quality and long-short opposition of Cantonese vowels], 《方言》 1: 28–38.
- Lucas, Alain and Xie, Honghua (1994) ‘There is practically one universal Chinese grammar: à propos de xian1 en mandarin et sin1 en cantonais’, *Cahiers de Linguistique Asie Orientale* 23: 189–206.
- Luke, K.K. (1990) *Utterance Particles in Cantonese Conversation*, Philadelphia: John Benjamins.
- Matthews, Stephen (ed.) (1998a) *Studies in Cantonese Linguistics*, Hong Kong: Linguistic Society of Hong Kong.
- Matthews, Stephen (1998b) ‘Evidentiality and mirativity in Cantonese’, *Proceedings of the 6th International Symposium on Chinese Languages and Linguistics*, Taiwan: Academia Sinica.
- Matthews, Stephen (1999) ‘Y.R. Chao and universal Chinese grammar’, in D. Cram et al. (eds), *History of Linguistics 1996*, vol. 1, Amsterdam: John Benjamins, pp. 217–224.
- Matthews, Stephen (2006) ‘Cantonese grammar in areal perspective’, in Alexandra Y. Aikhenvald and R.M.W. Dixon (eds), *Grammars in Contact*. Oxford: Oxford University Press, pp. 220–236.
- Matthews, Stephen and Yip, Virginia (2001) ‘Aspects of contemporary Cantonese grammar: the structure and stratification of relative clauses’, in Hilary Chappell (ed.), *Sinitic Grammar: Synchronic and Diachronic Perspectives*. Oxford: Oxford University Press, pp. 266–281.
- Matthews, Stephen and Yip, Virginia (2011) *Cantonese: A Comprehensive Grammar*, London: Routledge.
- Norman, Jerry (1988) *Chinese*, Cambridge: Cambridge University Press.
- O’Melia, Thomas (1941) *First Year Cantonese* (2nd edition), Hong Kong: Maryknoll House.
- Snow, Don (2004) *Cantonese as Written Language: The Growth of a Written Chinese Vernacular*, Hong Kong: Hong Kong University Press.
- Tam, Arthur and Cummins, Anna (2015) The death of Cantonese? *Timeout Hong Kong*, 1–14 July, 185: 20–26. Retrieved from: www.timeout.com.hk/big-smog/features/73305/the-death-of-cantonese.html (accessed 8 July 2015).
- Tang, Sze-Wing and Cheng, Siu-Pong (2014) ‘Aspects of Cantonese grammar’, in C.-T James Huang, Y.-H. Audrey Li and Andrew Simpson, eds, *The Handbook of Chinese Linguistics*. Malden, MA, and Oxford: John Wiley & Sons, pp. 601–628.
- To, C.K.S., McLeod, S. and Cheung, P.S.P. (2015) ‘Phonetic variations and sound changes in Hong Kong Cantonese: diachronic review, synchronic study and implications for speech sound assessment’, *Clinical Linguistics & Phonetics* 29.5: 333–353.

- Wong, Cathy Sin Ping, Bauer, Robert S. and Lam, Zoe Wai-Man (2009) ‘The integration of English loanwords in Hong Kong Cantonese’, *Journal of the Southeast Asian Linguistics Society* 1: 251–265.
- Yan, Margaret Mian (2006) *Introduction to Chinese Dialectology*, Munich: LINCOM.
- Yao, Shun-chiu (1980) ‘Sentential connotations in Cantonese’, 《方言》1: 35–52.
- Yuan, Jiahua 袁家驥等著 (1983) 《漢語方言概要》[Outline of Chinese dialectology], 第九章, 粵方言 [Chapter 9, Yue dialects], 北京: 文字改革出版社. 第二版.
- Yue-Hashimoto, Anne (1991a) ‘Stratification in comparative dialectal grammar: a case in Southern Min’, *Journal of Chinese Linguistics* 20: 172–200.
- Yue-Hashimoto, Anne (1991b) ‘The Yue dialects’, in William S.-Y. Wang, ed., *Languages and Dialects of China. Monograph Series Number 3, Journal of Chinese Linguistics*. Berkeley, CA: Journal of Chinese Linguistics, pp. 294–324.
- Yue-Hashimoto, Anne (1993) *Comparative Chinese Dialectal Grammar: A Handbook for Investigators*, Paris: EHESS, Centre de Recherches Linguistiques sur l’Asie Orientale.
- Zee, Eric (1999) ‘Chinese (Hong Kong Cantonese)’, in *Handbook of the International Phonetic Association, A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press, pp. 58–60.
- Zhonghua renmin gongheguo guojia tongyong yuyan wenzifa 中華人民共和國國家通用語言文字法. (2001) Retrieved from: big5.gov.cn/gate/big5/www.gov.cn/ziliao/flfg/2005-08/31/content_27920.htm (accessed 7 January 2014).
- Zong, Fubang 宗福邦 (1964) ‘關渝於廣州話陰平調的分化問題’ [On the problem of the split of Cantonese Yinping tone] 《中國語文》5: 376–389.

CHAPTER EIGHT

SHANGHAINESSE

Eric Zee and Liejiong Xu

1 SHANGHAI PHONOLOGY

Eric Zee

The Shanghai dialect, a member of the Wu dialect family, is spoken by approximately 24 million people residing mainly in the city of Shanghai and its vicinities. The city is situated at the southern part of the Yangtze River delta in eastern China. The phonological description of the Shanghai dialect presented here is typical of the educated metropolitan Shanghai speakers in their late fifties.

1.1 Consonants

Shanghai has been described as having maintained: (1) the historical tripartite division of syllable-initial plosives and affricates; and (2) the historical voiced and voiceless distinction of the syllable-initial fricatives (Chao 1928; Sherard 1972; Xu and Tang 1988). However, spectrographic data from the native speakers of Shanghai in their late fifties show that the initials [b d g] in a large majority of the monosyllables in isolation are voiceless. In Chao (1928, 1936), the initials [b d g] were transcribed as [p^h t^h k^h]. The breathiness was described as being realized on the following vowel. This was confirmed by experimental data in Cao and Maddieson (1992) and Ren (1995). Spectrographic data also show that the initials [v z z dz g^w] are more often pronounced as voiceless than voiced, and, as for [h], it is always voiceless. The voiced impression of the historical syllable-initial consonants [b d g g^w dz v z z f] is attributable to the breathiness of the following vowel associated with a low tonal onset. These syllable-initial consonants also differ from [p t k k^w t^h s e h], in that the former become ‘true’ voiced sounds in medial position, whereas the latter remain voiceless. The breathiness of the vowels associated with a low tonal onset has led scholars (Xu and Tang 1988) to transcribe the syllable-initial sonorants as [fim fin fij fiŋ fil fiw] and those followed by a non-breathy vowel as [ʔm ʔn ʔŋ ʔl ʔw]. These transcriptions are misleading, though, as the sonorants are neither preceded by the voiced glottal frication, nor are they glottalized. The consonant system in Shanghai consists of 12 plosives [p p^h b t t^h d k k^h g k^w k^wh g^w], one glottal stop [ʔ], four nasals [m n ŋ ŋ^w], nine fricatives [f v s z e z h h^w], five affricates [ts ts^h tʃ tʃ^h dz], two approximants [w j], and one lateral liquid [l], as shown in Table 8.1.

Phonologically, [b d g g^w] are treated as the allophones of /p t k k^w/ and [v z z dz f] as the allophones of /f s e tʃ h/. The historical [dz] in Shanghai is now [z]. The consonant phonemes in Shanghai include /p p^h b t t^h k k^h k^w k^wh/ [m n ŋ ŋ^w] [ts ts^h tʃ tʃ^h dz] [w j l].

The voiced initial obstruents in Shanghai are associated with the tones having a low tonal onset, such as [¹³] and [¹²⁷] (¹ indicating that the tone is associated with syllables with a glottal ending), whereas the voiceless counterparts are associated with the tones having a non-low tonal onset, such as [¹¹], [¹³⁴], and [¹⁵]. Such an association of consonant

TABLE 8.1 CONSONANTS

	Bilabial	Labio-dental	Alveolar	Alveolo-palatal	Velar	Glottal
Plosive	[p p ^h b]		[t t ^h d]		[k k ^h g] [k ^w k ^{w^h} g ^w]	[?]
Nasal	[m]		[n]	[ŋ]	[ŋ]	
Fricative		[f v]	[s z]	[ç z]		[h f]
Affricate			[ts ts ^h]	[tʃ tʃ ^h dz]		
Approximant	[w]			[j]		
Lateral			[l]			

type with tone does not hold for the syllable-initial sonorants [m n ŋ ɳ w ɿ l], as the sonorants co-occur with any tone. The historical four-way place contrast of the nasals in syllable-initial position is maintained in Shanghai, for example, [me¹³] ‘slow,’ [ne¹³] ‘difficult,’ [nje¹³] ‘twenty,’ and [ne¹³] ‘slow-witted.’ Examples of the monosyllables containing the consonants in syllable-initial position in Shanghai are listed as follows (with [⁹] after the vowel [a] indicating that the velar nasal is weakened, but not dropped, and [V] representing a vowel or diphthong):

- /p/ [pi⁵¹] edge, [pi³⁴] flat, [pri⁵²] pen
[bi¹³] skin, [bir¹²²] other
- /p^h/ [p^hi⁵¹] to criticize, [p^hi³⁴] to lie, [p^hi⁵²] to cleave
- /t/ [ti⁵¹] low, [ti³⁴] shop, [tr⁵²] to fall
[di¹³] sweet, [dr¹²²] enemy
- /t^h/ [t^hi⁵¹] sky, [t^hi³⁴] to substitute, [t^hi⁵²] to kick
- /k/ [ke⁵¹] crafty, [ke³⁴] to select, [ka⁵²] pigeon
[ge¹³] to lean against, [ga⁹]¹²² crowded
- /k^h/ [k^ha⁵¹] to knock, [k^ha³⁴] to rely on, [k^hə⁵²] to cough
- /k^w/ [k^wɛ⁵¹] to close, [k^wɛ³⁴] accustomed to, [k^wə⁵²] bone
[g^wɛ¹³] to throw, [g^wV]¹²² (non-occurring)
- /k^{w^h}/ [k^{w^ha⁵¹] to boast, [k^{w^ha³⁴] fast, [k^{w^hə⁵²] wide}}}
- /ts/ [ts]⁵¹] pig, [ts]³⁴] paper, [tsa⁵²] to tie
- /ts^h/ [ts^hɣ⁵¹] to take out from between, [ts^hɣ³⁴] stench, [ts^ha⁵²] ruler
- /tʃ/ [tʃi⁵¹] pointed, [tʃi³⁴] to mail, [ter⁵²] to catch/receive
[dʒi¹³] to ride (a horse), [dzi¹²²] extreme
- /tʃ^h/ [tʃ^hi⁵¹] to bully, [tʃ^hi³⁴] to go, [tʃ^hi⁵²] seven
- /f/ [fi⁵¹] to fly, [fi³⁴] lung, [fɔ⁵²] fortune
[vi¹³] fat, [vɔ⁹]¹²² to obey

- /s/ [sɔ⁵¹] to cook, [sɔ³⁴] scant, [sɔʔ⁵²] uncle
 [zɔ¹³] to build, [zɔʔ¹²²] cooked
- /e/ [ɛja⁵¹] fragrant, [ɛja³⁴] to think, [ɛiʔ⁵²] snow
 [zɛja¹³] to resemble, [zɛiʔ¹²²] mat
- /h/ [ha⁵¹] laugh (onomatopoeic), [ha³⁴] crab, [haʔ⁵²] blind
 [ɦia¹³] shoe, [ɦiaʔ¹²²] box
- /m/ [mej⁵¹] every, [mej³⁴] good-looking, [mVʔ⁵²] (non-occurring)
 [mej¹³] plum, [məʔ¹²²] ink
- /n/ [ne⁵¹] to hand over, [nV(ŋ)³⁴]/[nVʔ⁵²] (non-occurring)
 [ne¹³] difficult, [nɔʔ¹²²] promise
- /ɲ/ [ɲjx⁵¹] to pinch, [ɲjV(ŋ)³⁴]/[ɲjVʔ⁵²] (non-occurring)
 [ɲjx¹³] cow, [ɲjɔʔ¹²²] meat
- /ŋ/ [ŋju⁵¹] I (literary), [ŋV(ŋ)³⁴]/[ŋVʔ⁵²] (non-occurring)
 [ŋju¹³] I, [ŋɔʔ¹²²] crane
- /l/ [ljx⁵¹] to sneak out, [lV(ŋ)³⁴]/[lVʔ⁵²] (non-occurring)
 [ljx¹³] flowing, [lɪʔ¹²²] to stand
- /w/ [wəŋ⁵¹] lukewarm, [wəŋ³⁴] steady, [waʔ⁵²] to scoop
 [wəŋ¹³] dizzy, [waʔ¹²²] slippery
- /j/ [ja⁵¹] to hide, [ja³⁴] elegant, [jiʔ⁵²] one
 [ja¹³] father (term of reference), [jiʔ¹²²] leaf
- /?/ [ʔəŋ⁵¹] favour, [ʔV(ŋ)³⁴] (non-occurring), [ʔaʔ⁵²] duck,
 [ʔV(ŋ)¹³]/[ʔVʔ¹²²] (non-occurring)

1.2 Vowels

The vowel system in Shanghai consists of [i ɪ y ʏ ε ə ø a ɔ ʊ o ʊ], which are listed in Table 8.2.

TABLE 8.2 VOWELS

Close	i y	u
Near close	ɪ ʏ	ʊ
Close mid		o ɔ
Mid	ə	
Open mid	ɛ ø	ɔ
Near open		
Open	a	

Phonologically, /i y ε ø a ɔ ɤ o u/ are the vowel phonemes and [ɪ ʏ ə ʊ] are allophones of /i y ε u/, respectively. The derivation of [ɪ ʏ ə ʊ] requires the following phonological rules, /i/ → [ɪ]/_ [ŋ ?]; /y/ → [ʏ]/_ [ŋ ?]; /ε/ → [ə]/_ [ŋ ?]; and /u/ → [ʊ]/_ [ŋ]. Phonetically, [ɛ], [ø], and [ɔ] lie between the Cardinal Vowels [e] and [ɛ], Cardinal Vowels [ø] and [œ], and Cardinal Vowels [o] and [ɔ], respectively. [o] lies between the Cardinal close-mid [o] and close [u]. In the speech of speakers in their early fifties and younger, the rhymes [-a⁹] and [-ɔ⁹] have merged to become [-ã⁹], with the vowel being heavily nasalized. Examples of the monosyllabic words containing the vowels are as follows:

- /i/ [i] [ɛi⁵¹] ahead, [tʃ³⁴] shop, [mi¹³] rice
[ɪ] [pi²⁵²] pen, [dr¹²²] enemy
[ɛ] [ɛɪ⁵¹] new, [tnj³⁴] peak, [bjŋ¹³] bottle
- /y/ [y] [ey⁵¹] weak, [tey³⁴] expensive, [ŋy¹³] female
[ʏ] [tey²⁵²] military, [dzy¹²²] to dig
[ɛ] [tevŋ⁵¹] military, [ɛvŋ³⁴] training, [jvŋ¹³] cloud
- /ε/ [ɛ] [ke⁵¹] sly, [ts^hɛ³⁴] vegetable, [mɛ¹³] slow
[ø] [kʷə²⁵²] bone, [fø¹²²] box
[ɛ] [kəŋ⁵¹] root, [təŋ³⁴] to wait, [bəŋ¹³] stupid
- /ø/ [ø] [sø⁵¹] sour, [k^hø³⁴] to see, [mø¹³] full
- /a/ [a] [k^ha⁵¹] to wipe, [ts^ha³⁴] to tear, [da¹³] to wash
[a²⁵²] duck, [ba¹²²] white
[sa^{&51}] uncooked, [sa^{&34}] thrifty, [za^{&13}] long
- /ɔ/ [ɔ] [so⁵¹] to cook, [ho³⁴] good, [lɔ¹³] old
[k^hɔ²⁵²] to cry, [dɔ¹²²] poisonous
[fɔ⁵¹] square, [p^hɔ³⁴] weighty, [gɔ⁹¹³] foolish
- /o/ [o] [po⁵¹] scar, [p^ho³⁴] fear, [bo¹³] to climb
- /ɤ/ [ɤ] [sv⁵¹] to collect, [k^hɤ³⁴] mouth, [lɤ¹³] leaking
- /u/ [u] [tu⁵¹] abundant, [pu³⁴] cloth, [gu¹³] to squat
[ʊ] [sun⁵¹] loose, [t^hʊŋ³⁴] sore, [nʊŋ¹³] you

1.3 Syllabic consonants

The syllabic consonants in Shanghai are [m ɳ] and [ɿ]. While there are only a few monosyllables which contain a syllabic nasal as the vowel nucleus, for example, [m] in the bisyllabic compound [m ma⁵⁻³] ‘mother’ and [ɳ¹³] ‘fish,’ a sizable number of monosyllables contain a syllabic alveolar approximant [ɿ], for example, [ts^hɿ⁵¹] ‘to blow,’ [ts]³⁴ ‘paper,’ and [zɿ¹³] ‘self.’ The approximant [ɿ], called an ‘apical vowel’ in Karlgren

(1915–1926), occurs only after the apico-alveolar sibilants [ts ts^h s z]; it is treated as an allophone of the vowel phoneme /i/.

1.4 Diphthongs

Four diphthongs /ja jo jv ej/ are identified in Shanghai. The first element of the phonetic representations of /ja jo jv/ is shorter than the second, whereas the first element of the phonetic representation of /ej/ is longer than the second. Examples of the monosyllables containing the diphthongs are as follows:

- /ja/ [tja⁵¹] father (term of address), [tçja³⁴] to borrow, [dzja¹³] talented
- [tç^hja⁰⁵¹] gun, [çja⁰³⁴] to think, [lja⁰¹³] bright
- [çja?⁵²] to whittle, [zja?¹²²] to chew
- /jo/ [p^hjo⁵¹] to drift, [pjɔ³⁴] wristwatch, [djo¹³] to exchange
- [tejɔ?⁵²] orange, [njo?¹²²] meat
- /jv/ [ljv⁵¹] to sneak out, [tçjv³⁴] wine, [njv¹³] cow
- /ej/ [hwej⁵¹] dust, [tej³⁴] correct, [wej¹³] stomach

Other diphthongs, such as [je] and [ju], occur in the language. These diphthongs are formed by inserting an on-glide [j] between the vowel [ɛ] or [ʊ] and the preceding alveolo-palatal sound [ç z te^h dz] or [ŋ]. A general on-glide-insertion rule, Ø → [j]/ [+coronal –anterior]_V, where V ≠ [i ɪ y ʏ], is formulated for inserting [j] before any vowel, except [i], [ɪ], [y], or [ʏ], when the vowel is preceded by an alveolo-palatal sound. The ‘true’ diphthongs, such as [ja jo jv ej], may be preceded by a sound other than an alveolo-palatal sound as shown in the above examples.

1.5 Syllable structures

In Shanghai, three types of syllables are distinguished: (c)v, (c)vs, and (c)vn, where the optional ‘c’ is any one of the consonants listed in Table 8.1, ‘v’ a vowel, a diphthong, a syllabic alveolar approximant [ɿ], a syllabic bilabial nasal [m], or a syllabic velar nasal [ŋ], ‘s’ a glottal stop [ʔ], and ‘n’ a velar nasal [ŋ]. The syllabic lamino-alveolar approximant [ɿ] occurs only after an apico-alveolar fricative [s z] or affricate [ts ts^h], and it is not followed by a nasal or glottal ending. The syllabic bilabial nasal [m] and velar nasal [ŋ] occur singly, not preceded nor followed by a consonant. The possible combinations of the syllable-initial consonants with the rhymes in Shanghai are listed in Table 8.3.

TABLE 8.3 SYLLABLE STRUCTURES

V	V + [ŋ]	V + [ʔ]	D	D + [ŋ]	D + [ʔ]
[i]	[iŋ]	[iʔ]			
[p p ^h b m f v t t ^h d l c z te te ^h dz j]	[p p ^h b m t t ^h d l ç z te te ^h dz j]	[p p ^h b m t t ^h d l c z te te ^h dz j]			

Note: V = vowel or syllabic approximant; D = diphthong.

The consonants which occur before a particular rhyme are listed under the rhyme in the same cell.

1.6 Tones

Five of the eight historical tones are preserved in Shanghai (see Table 8.4). The historical yinshang and yinqu have merged. This is also true for yangping, yangshang, and yangqu. The yin and yang tone series in Chinese correspond to the high and low tone registers. The tones in the yin series have a higher tonal onset, whereas the tones in the yang series have a lower tonal onset. Furthermore, the tones in the yin series and those in the yang series are associated with voiceless and voiced syllable-initial obstruents, respectively. Such a correlation between the tonal onset and syllable-initial consonant type is maintained in Shanghai. As shown in Table 8.4 and Table 8.5, the numerical values for the five citation tones in Shanghai are [⁵¹], [³⁴], [¹³], [⁵²], and [¹²²]. The superscript [[?]] indicates that the tone is associated with the short syllables which end with a glottal stop.

1.7 Tone melodies

There are a fixed number of tone melodies on the compounds in Shanghai, as shown in Table 8.5. The tone melodies realized on the bisyllabic and longer compounds are triggered by the lexical tone on the initial monosyllable in the compounds. The tone melodies are elastic, in that their length is adjustable to the size of the compound, as can be seen in

TABLE 8.4 TONE CATEGORIES

Tonal category	Citation form on monosyllables
<i>yin ping</i>	[⁵¹]
<i>yin shang, yin qu</i>	[³⁴]
<i>yang ping, yang shang, yang qu</i>	[¹³]
<i>yin ru</i>	[⁵²]
<i>yang ru</i>	[¹²²]

TABLE 8.5 LEXICAL TONE MELODIES

Type of lexical tone	Citation form on mono-syllables	TM on bisyllabic compounds	TM on trisyllabic compounds	TM on quadrисyllabic compounds	TM on quintesyllabic compounds
I /51/	[⁵¹]	[⁵⁻¹]	[⁵⁻³⁻¹]	[^{5-3↑-3↓-1}]	[⁵⁻⁴⁻³⁻²⁻¹]
II /351/	[³⁴]	[³⁻⁴]	[³⁻⁵⁻¹]	[³⁻⁵⁻³⁻¹]	[^{3-5-3↑-3↓-1}]
III /151/	[¹³]	[¹⁻³]	[¹⁻⁵⁻¹]	[¹⁻⁵⁻³⁻¹]	[^{1-5-3↑-3↓-1}]
IV /451/	[⁵²]	[⁴⁻⁵]	[⁴⁻⁵⁻¹]	[⁴⁻⁵⁻³⁻¹]	[^{4-5-3↑-3↓-1}]
V /12/	[¹²²]	[¹⁻¹²] ([¹⁻³])	[¹⁻¹⁻¹²] ([¹⁻⁵⁻¹])	[¹⁻¹⁻¹⁻¹²] ([¹⁻⁵⁻³⁻¹])	[¹⁻¹⁻¹⁻¹⁻¹²] ([^{1-5-3↑-3↓-1}])

Note: TM = tone melody.

Table 8.5. This is especially obvious with the Types I and V Lexical Tones. The Type I Lexical Tone /⁵¹/ is realized as [⁵⁻¹], [⁵⁻³⁻¹], [^{5-3↑-3↓-1}], and [⁵⁻⁴⁻³⁻²⁻¹] on the bisyllabic, trisyllabic, quadrисyllabic, and quintesyllabic compounds, respectively (the vertical arrows after a numeral indicate that the tone value is either raised or lowered). Similarly, the Type V Lexical Tone /¹²/ is realized as [¹⁻¹⁻¹²], [¹⁻¹⁻¹⁻¹²], [¹⁻¹⁻¹⁻¹⁻¹²], and [¹⁻¹⁻¹⁻¹⁻¹⁻¹²] on the bisyllabic and longer compounds. The Type V Lexical Tone /¹²/ may be optionally realized as [¹⁻³] on some of the bisyllabic compounds, and [¹⁻⁵⁻¹], [¹⁻⁵⁻³⁻¹], and [^{1-5-3↑-3↓-1}] on some of the longer compounds.

In Zee and Maddieson (1980) the tone sandhi in Shanghai is analyzed as an autosegmental process of rightward spreading of the lexical tone on the initial syllable in the compounds, and the lexical tones on the following syllables are deleted. The lexical tones are posited as /HM/, /MH/, /LH/, /H/, and /LH/, which correspond to the Types I, II, III, IV, and V Lexical Tones, respectively, in this study.

A compound in Shanghai may be delineated as a morphological unit which consists of two or more monosyllables and is associated with a single tone melody or domain. It may therefore be regarded as a prosodic word (Selkirk 1984) or phonological word (Nespor and Vogel 1986). A compound in Shanghai rarely consists of more than four monosyllables. Examples of the bisyllabic and longer compounds associated with different tone melodies are presented below. Some of the examples are from Xu and Tang (1988). However, the IPA transcriptions and the numerical tone values are based on the analysis in this study. It should be noted that a syllable-initial or syllable-final glottal stop in the medial position is deleted.

Type I Lexical Tone /⁵¹/:

- [eja⁹ ji]⁵⁻¹ cigarette
- [tʰi vəŋ də]⁵⁻³⁻¹ observatory
- [se ɛŋ lja⁹ ji]^{-3↑-3↓-1} indecisive
- [ke li və(?) ke ka]⁵⁻⁴⁻³⁻²⁻¹ awkward

Type II Lexical Tone /³⁵¹/:

- [hu tsʰo]³⁻⁴ train
- [ta⁹ zl̩ tei]³⁻⁵⁻¹ typewriter
- [tʃɔ r lɔ⁹ tɔ⁹]³⁻⁵⁻³⁻¹ hanging around and doing nothing
- [ejo tʰi dr sŋ vu]^{3-5-3↑-3↓-1} little barber

Type III Lexical Tone /¹⁵¹/:

- [dʒɔ ka⁹]¹⁻³ spoon
- [zəŋ tɛŋ bɪŋ]¹⁻⁵⁻¹ nutty
- [zja tʰi zja di]¹⁻⁵⁻³⁻¹ thank heaven and earth
- [zja⁹ sa(?) jʂ ka zl̩]^{1-5-3↑-3↓-1} pretentious and pompous

Type IV Lexical Tone /⁴⁵¹/:

- [?o(?) tsʰo(?)]⁴⁻⁵ dirty
- [teja(?) da(?) tsʰo]⁴⁻⁵⁻¹ bicycle
- [ha(?) se wo sŋ]⁴⁻⁵⁻³⁻¹ talking nonsense
- [tɔ(?) dŋ mo sŋ kʰɛ]^{4-5-3↑-3↓-1} confident

Type V Lexical Tone /¹²/ (optionally /¹⁵¹/ on some compounds):

- [za(?)ljʂ]¹⁻¹² pomegranate
- [lɔ(?)jŋ tei]¹⁻¹⁻¹² tape recorder
- [vɔ(?) te da fɔ?]¹⁻¹⁻¹⁻¹²⁽¹⁻⁵⁻³⁻¹⁾ Fudan University
- [la(?) li dr ŋi tsŋ]^{1-1-1-1-12(1-5-3↑-3↓-1)} a son with favus on the head (is still a son to be proud of)

In Shanghai, a tone melody may also be associated with a phrasal unit containing words of different grammatical categories, as in the following examples:

[ts^hɔ p^hjɔ]³⁻⁴ [ŋu]¹³ [wε pə(?) ji lə]¹⁻⁵⁻³⁻¹
money 1sg return to 3sg ASP
'I have returned the money to him.'

[ŋu]¹³ [və(?) tə(?) ji kə⁹]¹⁻⁵⁻³⁻¹
1sg (NEG) to 3sg tell
'I will not tell him/her.'

[ŋu]¹³ [tɔ]³⁴ [zə⁹hε te^hi]¹⁻⁵⁻¹
1sg arrive Shanghai go
'I am going to Shanghai.'

[ji]¹³ [te^hi(?) hə lə]⁴⁻⁵⁻¹
3sg eat finished ASP
'He/she has finished eating.'

For more about the phrasal phonology in Shanghai, see Jin (1986) and Selkirk and Shen (1990).

1.8 Concluding remarks on phonology

The foregoing presents a précis of the main aspects of Shanghai phonology based on the speech of the language consultants in their late fifties. It should be pointed out that Shanghai is a fast changing language as far as its sound system is concerned. Much of the change is attributable to the influence of the standard language, Putonghua (Mandarin). The educated young adults in their early twenties have difficulty in pronouncing many monosyllabic words in Shanghai. A phonology of Shanghai based on the speech of the younger speakers is expected to differ to some extent from this presentation.

2 SHANGHAI MORPHOLOGY AND SYNTAX

Liejiong Xu

2.1 Noun phrases

A noun, whether countable or uncountable semantically, may be preceded by a demonstrative plus a classifier or by a numeral plus a classifier, or by both:

gə?[?] pəŋ[¹⁻³] sŋ[⁵¹]
this CL book
'this book'

lia⁹ pej[¹⁻³] sŋ[³⁴]
two CL water
'two cups of water'

gə? gə?¹⁻³] lia⁹ pej[¹⁻³] sŋ[³⁴]
this CL two CL water
'these two cups of water'

As in Mandarin, the numeral *ji?* ‘one’ may be omitted if the noun phrase is in the object position, leaving the classifier alone with the noun:

pəŋ[³]	sɿ[⁵¹]
CL	book
‘book’	

Unlike in Mandarin, omission occurs even with the subject in Shanghainese. Such an incomplete noun phrase can be interpreted as either definite or indefinite, whereas its counterparts in some other Chinese dialects, e.g. Cantonese, must be interpreted as definite.

Noun phrases are strictly head final. A phrasal or clausal modifier can either occur in initial position or immediately precede the head noun:

ji[¹³]	eia gə? ^[3-4]	se pəŋ[⁵⁻¹]	sɿ[⁵¹]
3sg	write MOD.MARKER	three CL	book
‘the three books he wrote’			
se pəŋ[⁵⁻¹]	ji[¹³]	eia gə? ^[3-4]	sɿ[⁵¹]
three CL	3sg	write MOD.MARKER	book
‘the three books he wrote’			

2.2 Verb phrases

The object follows the verb in the canonical form.

də?	tsɿ[¹⁻³]	pəŋ[³]	sɿ[⁵¹]
read	ASP	CL	book
‘have read a book’			

The typical Shanghainese order for the double object construction is v–Direct Object–Indirect Object, when the indirect object is a pronoun (Xu and Tang 1988: 479):

soŋ pəŋ[³⁻⁴]	sɿ noŋ[⁵⁻¹]
give CL	book 2sg
‘give you a book’	

It can be described as derived from the oblique construction by preposition deletion:

soŋ pəŋ[³⁻⁴]	sɿ[⁵⁻¹]	pə?	noŋ[³⁻⁴]
give CL	book	to	2sg
‘give a book to you’			

The preposition cannot be deleted if the indirect object is a full noun phrase instead of a pronoun. An alternative order, v–io–do, is also used by speakers in the city of Shanghai, following the Mandarin pattern:

soŋ noŋ[³⁻⁴]	pəŋ[³]	sɿ[⁵¹]
give 2sg	copy book	
‘give you a book’		

Other expressions that follow the verb are those traditionally known as duration phrases and frequency phrases, both in the form of numerically quantified NP.

də?	tsɿ[¹⁻³]	liaŋ gə? ^[1-3]	tsondʒr[⁵⁻¹]
read	ASP	two CL	hours

'have read for two hours'

dɔ? tsɿ[¹⁻³] se tʰɔŋ[⁵⁻¹]
read asp three times
'have read three times'

Such postverbal adjuncts are analyzed as complements in traditional grammar. In contrast, preverbal adjuncts denoting time, place, manner, etc., are assigned to the category of adverbial.

As in Mandarin, ov order is also found. The object is often preceded by the preposition *ne^{s1}*, equivalent to *bă* in the so-called *bă*-construction in Mandarin.

ne[^{s1}] gə? ɳɛ[¹⁻³] sɿ[⁵¹] gʷɛ tʰə?[¹⁻³]
PREP this CL book throw-away
'Throw away these books.'

Not every vo sentence has an ov variant. There are thematic and prosodic restrictions on the use of the ov order. A resumptive pronoun, identical in form to the third person singular pronoun, can appear in the gap following the verb (Qian 1997: 287). It is always singular, regardless of the number of its antecedent.

ne[^{s1}] gə? ɳɛ[¹⁻³] sɿ[⁵¹] gʷɛ tʰə? ji[¹⁻⁵⁻¹]
PREP this CL book throw away it
'Throw away these books.'

Preverbal objects not introduced by a preposition occur more frequently in Shanghainese than in Mandarin. In Shanghainese the preposition *ne^{s1}* is optional, while omission of its Mandarin counterpart *bă* often results in sentences less acceptable to Mandarin speakers.

The resultative construction consists of a verb plus a resultative complement plus a noun:

ɛia tsʰo tsɿ[³⁻⁵⁻¹] di tsɿ[¹⁻³]
write wrong ASP address
'Write the address wrongly.'

teʰja? wø ji[⁴⁻⁵⁻¹]
eat finish it
'Eat it up.'

The expression following the verb, whether an adjective, an adverb, or a verb in form, functions as a predicate rather than an adjunct. So the first example above means write the address in such a way that it is wrong and the second one means eat it until it is all gone. Thematically, the final noun phrase performs a dual role: it is an argument of the verb and an argument of the resultative element as well. Shanghainese distinguishes itself from Mandarin and many other dialects in that the object, if it is a pronoun, can precede the resultative complement. The inversion is almost obligatory in a negative structure. In the following minimal pair, the former is preferable to the latter.

teʰja? ji və? wø[⁴⁻⁵⁻³⁻¹]
eat it NEG finish
'can't eat it up'

teʰja? və? wø ji [⁴⁻⁵⁻³⁻¹]
eat NEG finish it
'can't eat it up'

2.3 Topic structure

Compared with Mandarin, Shanghainese is even more typically topic-prominent. It has all the properties characteristic of topic-prominent languages. A topic can be a noun phrase, a prepositional/postpositional phrase, a verb phrase, or a clause.

tsɔvə[³⁻⁴] ji[¹⁻³] za^ŋza^ŋ[¹⁻³] və? te^hja?^[1-3]
breakfast 3sg often NEG eat
'Breakfast, he often doesn't eat.'

nøteiŋ lu lɔ^ŋ[¹⁻⁵⁻³⁻¹] ηu[¹³] ciøtə? ji[³⁻⁵⁻¹] zl ts][¹⁻³] za^ŋ jy[¹⁻³]
Nanjing street on 1sg know 3sg live ASP long
'On that street I know he has lived for a long time.'

sɔ[⁵]ve[¹³] ji[¹³] tsə? wej[³⁻⁴] ts^hɔ[³⁴] de[¹³]
cook 3sg only can scramble egg
'As for cooking, he can only scramble eggs.'

The first two sentences above are somewhat similar in structure to their English counterparts. In the first one, the verbal object has been topicalized. In the second one, the postpositional phrase has moved out of the embedded clause, across the boundary of the matrix clause and landed in the sentence initial position. In the third sentence, a verb appears in the topic position. In Shanghainese, a topic need not bind a gap and the expression bound to the topic is not necessarily a pronoun, as in the case of dislocation in English.

tsɔvə[³⁻⁴] ji[¹³] za^ŋza^ŋ[¹⁻³] te^hja?^[5] mipo[¹⁻³]
breakfast 3sg often eat bread
'For breakfast, he often eats bread.'

nøteiŋ lu[¹⁻⁵⁻¹] ηu[¹³] ciøtə? ji[³⁻⁵⁻¹] la[¹²] imi[⁵⁻¹] zl ts][¹⁻³] za^ŋ jy[¹⁻³]
Nanjing street 1sg know 3sg in there live ASP long
'On that street I know he has lived for a long time.'

There is an anaphoric relation between *tsɔvə?* and *mipo*, and between *nøteiŋ lu* and *imi* in the example above. But the anaphors are neither empty nor pronominal. Furthermore, there are topic sentences in which the topic is semantically related to the comment as a whole but not related to any individual expression in it.

bijku[¹⁻³] zə? k^wej[¹⁻¹²] ji? koŋteiŋ[⁴⁻⁵⁻¹]
apples ten dollars one kilo
'As for apples, they cost ten dollars a kilo.'

Topics are often marked by particles. In Mandarin the topic markers are found in other syntactic positions as well, for instance, in sentence-final position. Shanghainese has two particles, *ma* and *zl*, used exclusively to mark topics.

dino mə?[¹⁻⁵⁻¹] ηu zl[¹⁻³] ηa fiɔ^ŋ[¹⁻³]
computer TOP 1sg COPULA layman
'As for computers, I am a layman.'

gə? gə?[¹⁻³] ηu[¹³] zl[¹] mpmə?^[1-3] ηu[¹³] kə[³⁴] za[¹³]
this CL person TOP NEG person dare offend
'This person, no one dares to offend.'

In Mandarin, the typical position for a topic is at the very beginning of the sentence. In Shanghainese, a topic more frequently occurs between the subject and the verb, especially in negative sentences and in yes–no questions.

ji[¹³]	don ^d di[¹⁻³]	mma? ² [¹⁻³]	fu	ku[³⁻⁴]
3sg	money	NEG		pay ASP
'Money, he has not paid.'				
non[¹³]	zo[¹³]	jɔ[³] te ^h ja?	va[⁴⁻⁵]	
2sg	tea	want drink	Q	
'Would you like to drink some tea?'				

2.4 Interrogative structures

Wh-questions have the same word order as declarative sentences with the interrogative words *in situ*:

non[¹³]	ma	tsi[¹⁻³]	samə? ² zj[³⁻⁵⁻¹]
2sg	bought	ASP	what?
'What did you buy?'			
ji[¹³]	sazəŋk ^{wəŋ} [³⁻⁵⁻¹]	le	ku[¹⁻³]
3sg	when		come ASP
'When did he come?'			

Since an interrogative word has an indefinite reading as well, the above sentences are ambiguous. The first one can be interpreted as ‘what did you buy’ or as ‘you bought something.’

A yes–no question can appear in three alternative forms exemplified below:

non[¹³]	te ^h i və? ² [³⁻⁴]
2sg	go Q
'Did you go?'	
non[¹³]	te ^h i və? te ^h i[³⁻⁵⁻¹]
2sg	go NEG go
'Did you go or not?'	
non[¹³]	a? ⁵ te ^h i[³⁴]
2sg	Q go
'Did you go?'	

The last form with an interrogative particle preceding the verb, not available in Mandarin, is typically Shanghainese. Hybrid forms are also permissible (Qian 1992: 1012):

non[¹³]	a? ⁵	te ^h i	və? ² [³⁻⁴]
2sg	Q	go	Q
'Did you go?'			
non[¹³]	a? ⁵	te ^h i	və? te ^h i[³⁻⁵⁻¹]
2sg	Q	go	NEG go
'Did or didn't you go?'			

When the v-not-v form is used and the verb is transitive, Shanghainese is different from Mandarin in word order. While the object can follow either v in Mandarin, in Shanghainese it must follow the second v:

noŋ^[13] tsʰy və? tsʰy^[5-3-1] euaŋi^[5-1]
2sg smoke NEG smoke cigarette?
'Do you smoke or not?'

*noŋ^[13] tsʰy^[5-1] euaŋi^[5-1] və? tsʰy^[1-3]
2sg smoke cigarette NEG smoke?
'Do you smoke or not smoke?'

Occasionally, the object as well as the verb can be reduplicated.

The interrogative particle *a?* cannot immediately precede the negative particle *və?*. To ask a negative question, Shanghainese uses *aʔzj* instead of *a?*. The following sentence means: 'Is it the case that you won't go?'

noŋ^[13] aʔ zj^[4-5] və? teʰi^[1-12]
2sg Q COPULA NEG go
'Did you go or not?'

2.5 Negative structures

Negative sentences are formed by using one of the two negators, *və?* and *mmə?*.

The negative adverb *və?* precedes and modifies an adjective or a verb:

ji^[13] və? tɔŋɛŋj^[1-5-1]
3sg NEG careful
'He is not careful.'

ji^[13] və? lɛ^[1-3]
3sg NEG come
'He won't come.'

As Chinese verbs lack tense inflections, the sentence above can be interpreted as 'he came,' 'he is coming,' or 'he will come,' in different contexts.

The other negator *mmə?* negates a predicate only if the action expressed by it takes place before the reference time (but not necessarily before the speech time or the event time). In the following sentence the reference point is Thursday.

lipa^[1-3]sŋ^[34] ji^[1] kɔŋ^[34] lipa^[1-3]se^[51] mmə?^[1-3] teʰi^[34]
Thursday 3sg speak Wednesday NEG go
'On Thursday he said he had not gone on Wednesday.'

*lipa^[1-3]sŋ^[34] ji^[1] kɔŋ^[34] lipa^[1-3]ŋ^[13] mmə?^[1-3] teʰi^[34]
Thursday 3sg speak Friday NEG go
'On Thursday he said he would not go on Friday.'

The adverb *mmə?* can be used to negate an event (i.e. action, achievement, or accomplishment), but not a state. To negate a state, *və?* is used.

ŋu^[13] və? eiɔtə?^[1-5-1] gə? tsɔŋ^[1-3] zj tʰi^[1-3]
1sg NEG know this CL matter
'I don't know this matter.'

*ŋu^[13] ɿmə?^[1-3] eiɔtə?^[1-3] gə?^[1-3] zɿ t̪i^[1-3]
1sg NEG know this CL matter
'I don't know this matter.'

In the following minimal pair, either of them can appear. But the meaning is different:

ji gə?^[1-3] mi k^honj^[3-4] və?^[1-3] hoŋ^[1-3]
3sg this face NEG red
'His face is not red.'

ji gə?^[1-3] mi k^honj^[3-4] ɿmə?^[1-3] fion^[13]
3sg this face NEG red
'His face has not turned red.'

The latter is acceptable only if it is interpreted as involving a change of state.

The adverb *və?* tends to occur immediately before the expression it modifies. Compare the following pair:

ji^[13] za^hza^h^[1-3] və?^[1-3] kɔ^h^[1-3] fiɛwo^[1-3]
3sg often NEG speak speech
'He often does not speak.'

ji^[13] və?^[12] za^hza^h^[1-3] kɔ^h^[34] fiɛwo^[1-3]
3sg NEG often speak speech
'He does not often speak.'

While the first sentence means that the person is often silent, the second one implies that he does speak, though not frequently. Of the following pair, only the latter is acceptable.

*ji^[13] və?^[1-3] tɛŋ sɔ^h^[3-4]
3sg NEG speak clearly
'He does not speak clearly.'

ji^[13] kɔ^h və?^[3-5-3-1] tɛŋ sɔ^h
3sg speak NEG clearly
'He can't speak clearly.'

It has been suggested by grammarians (Xu and Shao 1998: 148) that cliticized to the verb *kɔ^h*, *və?* in the first sentence cannot extend its scope of negation to cover *tɛŋsɔ^h*. In the second sentence, *və?* directly negates *tɛŋsɔ^h*, which is the focus of the sentence.

If *ɿmə?* is used, the scope of negation may cover the entire verb phrase, not just the expression immediately following it.

ji^[13] ɿmə?^[1-3] kɔ^h tɛŋ sɔ^h^[3-5-1]
3sg NEG speak clearly
'He didn't speak clearly.'

This sentence means that he fails to make himself clear.

The older generation of Shanghai speakers use *və?zəŋ* as an alternative of *ɿmə?*. This expression comes from the dialects spoken in the region between Shanghai and the Yangtze River. Although interchangeable in most contexts, *ɿmə?* and *və?zəŋ* manifest some differences. The former can negate an existential state but the latter cannot.

i gə?^[5-1] zəŋ k^wɔ^h^[1-3] noŋ^[13] ɿmə?^[1-3] lahe^[1-3]
that CL time 2sg NEG in

‘You were not in at that time.’

*i gə? [⁵-¹] zəŋkʷɔŋ [¹-³] noŋ [¹³] və?zəŋ [¹-³] lahe [¹-³]
that CL time 2sg NEG in
‘That time you were not in.’

REFERENCES

- Cao, Jianfen and Maddieson, Ian (1992) ‘An exploration of phonation types in Wu dialects of Chinese’, *Journal of Phonetics* 20: 77–92.
- Chao, Yuen Ren (1928) *Studies in the Modern Wu-Dialects*, Bejing: Tsing Hua College Research Institute, Monograph No. 4.
- Chao, Yuen Ren (1936) ‘Types of plosives in Chinese’, in Daniel Jones and D.B. Fry (eds), *Proceedings of the 2nd International Congress of Phonetic Sciences*, Cambridge: Cambridge University Press, 106–110.
- Jin, Shunde (1986) *Shanghai Morphotonemics*, Bloomington: Indiana University Linguistics Club.
- Karlgren, Bernhard (1915–1926) *Études sur la phonologie chinoise*, Leyde and Stockholm: Brill.
- Nespor, Marina and Vogel, Irene (1986) *Prosodic Phonology*, Dordrecht: Folis.
- Qian, Nairong (1992) *Studies on Contemporary Wu Dialects*, Shanghai: Shanghai Education Press (in Chinese).
- Qian, Nairong (1997) *Grammar of Shanghai Dialect*, Shanghai: Shanghai People’s Press (in Chinese).
- Ren, Nianqi (1995) ‘A fiberoptic and transillumination study of Shanghai stops’, in Eric Zee (ed.), *Studies of Wu Dialects*, Hong Kong: Hong Kong New Asia College, Chinese University of Hong Kong, 261–267.
- Selkirk, Elisabeth O. (1984) *Phonology and Syntax; The Relation between Sound and Structure*, Cambridge, MA: MIT Press.
- Selkirk, Elisabeth O. and Shen, Tong (1990) ‘Prosodic domains in Shanghai Chinese’, in Sharon Inkelas and Draga Zec (eds), *The Phonology–Syntax Connection*, Chicago: Chicago University Press, 313–337.
- Sherard, Michael (1972) ‘Shanghai Phonology’, unpublished PhD dissertation, Cornell University.
- Xu, Baohua and Tang, Zhenzhu (1988) *Shanghai Shiqu Fangyan Zhi (A Study of the Metropolitan Shanghai Dialect)*, Shanghai: Shanghai Education Press (in Chinese).
- Xu, Liejiong and Shao, Jingmin (1998) *Studies on Shanghainese Grammar*, Shanghai: East China Normal University Press (in Chinese).
- Zee, Eric and Maddieson, Ian (1980) ‘Tones and tone sandhi in Shanghai: phonetic evidence and phonological analysis’, *Glossa* 14.1: 45–88.

CHAPTER NINE

CHINESE WRITING

Mark Hansell

The Chinese written language is the vehicle for one of the world's richest and oldest cultural traditions. With its combination of antiquity of origin, continuous vitality, and stability of graphic form and orthography, the Chinese writing system has no rival in the modern world. Like our own Roman alphabet, it spread geographically along with the empire that spoke it, and spread to other languages whose people fell under the cultural influence of the empire. It also inspired neighbouring peoples to invent indigenous writing systems that are structurally and visually similar to Chinese.

Within its cultural realm, the Chinese writing system has always been much more than a means to transmit language across time and space. It has also functioned as religious icon, work of art, symbol of political authority and force for cultural unity.

1 STRUCTURE AND FUNCTION

Writing is the association of a system of permanent visual symbols with units of a spoken language. The system of visual symbols is called a ‘script’, the principles of association between the symbols and the spoken language is called the ‘orthography’, and the resulting body of graphically represented linguistic forms is the ‘written lexicon’.

1.1 Script

Visually, Chinese text is a linear arrangement (either rows or columns) of equally spaced graphs. The graphs, usually called ‘characters’, consist of a varying number of lines or dots (called ‘strokes’). The strokes of each character fill an imaginary rectangle akin to that of Roman block lettering, except in the more exuberant forms of cursive hand.

Chinese writing is built on the principle that Hockett (1963: 9) calls ‘duality of patterning’: a small number of discrete, meaningless elements (strokes) is combined to form a much larger set of meaningful elements (characters). This duality mimics the relationship between phonemes and words in a spoken language, and reflects the same basic principle of writing that English does, with its small set of letters combining to form a huge set of words. For example, here are characters displaying an increasing number of strokes, from one to 30:

一十千木白耳尾朋亮豈梅超亂認鬧墾點壘懷黨蠟竊驗癱攬鑽鑼鑿鬱鸞

Strokes are not only the basic unit of the physical act of writing, but are also the fundamental unit of classification of characters. Indexing and arranging functions that are based on alphabetic order in alphabetically written languages (e.g. arrangement of dictionaries, ordering of names in directories, etc.) are based on the number and configuration of strokes in Chinese writing.

The majority of characters are graphic compounds, composed of two or more components (which may be other, simpler characters, or may be other frequently used multi-stroke elements). This compounding adds a third layer of structure to the script, which is intimately related to the expression of sound and meaning (orthography).

1.2 Orthography

The orthography of Chinese has been the subject of a great deal of controversy and speculation, not all of it well informed. Some of the better-known adjectives used to describe Chinese writing are given below, with explanations of how apt they are:

1.2.1 Ideographic

An ideograph is a graphic symbol that directly represents an idea, without relying on a particular word or sound as an intermediary. For instance, the Arabic numeral <1> has a single meaning but different pronunciations in many different languages (*one, un, uno, eins, ichi, yi*, etc.); it can represent various different English morphemes depending on context (as in <1st> = *fir* + *st*, <10> = *ten*, <11> = *II*); and it can represent different words that are synonyms (<1,000> = *a thousand*/*one thousand*).

Chinese characters superficially seem to be ideographic. For example, <園> has the identical meaning of ‘garden’, but the pronunciations Mandarin *yuan*, Taiwanese *hng*, Cantonese *yuhn*, Japanese *en* or *sono*. In the area of synonyms, however, Chinese characters are utterly unlike Arabic numerals: even for words of the same meaning, Chinese has different characters for each one. For example, each of the following means ‘red’:

紅 *hóng* 赤 *chì* 朱 *zhū* 丹 *dān*

If Chinese characters indeed represented meaning directly, there would be no need for four different characters meaning the colour red, despite the four different pronunciations.

The mistaken classification of Chinese writing as ideographic stems mainly from a confusion of script with orthography. A script may be shared between various languages, but an orthography is language-specific. Since a person reading a Chinese character is also a speaker of a specific language (for example, the Mandarin vs Taiwanese vs Cantonese example above), the fact that a given character is pronounced differently in different vernaculars is irrelevant to how the particular reader uses the character to get at linguistic meaning. The character is always interpreted as a particular word, in that particular dialect, with a particular pronunciation. (By analogy, the fact that the letter <w> is pronounced [w] in English and [v] in German does not mean that the Roman alphabet fails to distinguish semivowels from fricatives – it simply means that English and German, while they share the same script, apply different orthographic rules to that script.)

1.2.2 Pictographic

A pictograph is a highly iconic graph that represents meaning through visual similarity to the referent. Common modern examples include the ‘glyphs’ that direct visitors in international airports to restrooms of the proper gender, restaurants or baggage retrieval carousels. Historically, a number of Chinese characters were indeed pictographs at their earliest stage of existence, and in a very few cases the pictographic origin is still evident: 田 *tián* ‘field’ (a bird’s-eye view of divided fields), 門 *mén* ‘door’ (swinging double doors). Some are much more difficult to interpret, because of modern stylized forms

(standardization of the strokes as straight lines and angles have changed 日 *rì* ‘sun’ and 月 *yuè* ‘moon’ considerably from the original circle and crescent). Most characters of pictographic origin have become so stylized that only the most fertile imagination can impose a horse on 馬 *mǎ* or a tiger on 虎 *hǔ*. Nevertheless, the overwhelming majority of Chinese characters are not of pictographic origin, and are completely uninterpretable pictorially. For example, ‘whale’, ‘cuttlefish’, ‘abalone’, ‘tuna’, and ‘eel’ are all visually distinctive sea creatures, and truly pictographic representations of them should be readily identified even by the uninitiated. As an exercise, here are the characters for those five denizens of the deep; readers should find it impossible to match the graph to the referent by pictographic principles:

鯨 鮑 魷 鰻 魚

1.2.3 Logographic

In ‘logographic’ writing, basic graphs map onto words. Words are represented directly by graphs, rather than one or more graphs representing a sequence of sounds which are then associated with a word through the lexicon of the spoken language. In its lack of phonological mediation, a logograph resembles an ideograph as described above, with one important distinction: synonyms, being different words despite having the same meaning, need to be assigned unique graphs.

Classifying Chinese writing as logographic accounts for many of its characteristics, and is a vast improvement over the ideographic interpretation, but misses two crucial facts: most Chinese words are multimorphemic and are written with more than one character; and most characters do indicate pronunciation. With only a few exceptions, each character represents exactly one morpheme, which corresponds to exactly one syllable. This has led some to focus on the mapping of graph to morpheme and term Chinese a ‘morphographic’ system, and others to focus on the mapping of graph to syllable and term it ‘syllabic’. Perhaps the most useful view is that it is simultaneously both, what DeFrancis (1989) terms a ‘morphosyllabic’ writing system.

1.2.4 A morphosyllabic orthography

A ‘morphosyllabic’ writing system displays characteristics of both a morphographic system (each morpheme has a distinct written symbol) and a syllabic system (the sounds of words are represented at the syllable level). That a single orthography would simultaneously represent meaningful units (morphemes or words) and phonological elements (syllables or phonemes) should come as no surprise to English speakers. Variable spelling of homophones, the bane of the schoolchild, allows graphic differentiation of different words while preserving a certain degree of phonological predictability (e.g. *to/too/two*, *write/rite/right/wright*). Thus exceptions/aberrations in English grapheme-to-phoneme correspondences can function as markers of lexical identity. While Chinese has a somewhat less systematic means of representing sound than English does, its means of graphically differentiating homophones is much more systematic (as should be expected in a language with so many homophonous morphemes).

The majority of Chinese characters are compound characters consisting of two elements, a ‘radical’ (or ‘signific’) that indicates a general semantic category, and a ‘phonetic’ that indicates (more or less) how the syllable is pronounced. Each of the two components of the compound consists of one or more strokes, and is usually a character

鯨 jīng	鮑 bāo	魷 yóu	鰐 mán	鮪 wěi
whale	abalone	cuttlefish	eel	tuna
京 jīng	包 bāo	尤 yóu	曼 mán	有 yǒu
capital city	to wrap	especially	graceful	to have

FIGURE 9.1 COMPOUND CHARACTERS

with its own independent existence in the written language. For example, the five characters for aquatic life-forms mentioned above are all compound characters with <魚> *yú* ‘fish’ as the left-hand element (Figure 9.1, first row). The right-hand elements of these characters, if used alone, would be the characters given in the second row of Figure 9.1. They are used as phonetics in the characters in the first row because of similarity of pronunciation. In the first three pairs the pronunciation is identical, but in the fourth ('eel' and 'graceful') there is a difference in tone, and in the fifth ('tuna' and 'to have') the two differ in significant ways. This discrepancy highlights the variability of the phonetic components, which originally indicated homophones or near-homophones, but in some cases have lost much of their accuracy over time due to sound change.

Certain phonetics are highly regular, while others exhibit extreme variability: fourteen characters with the phonetic <皇> are all pronounced *huáng*, while 22 characters with the phonetic <堯> give 15 different pronunciations (*yāo* *yáo* *yǎo* *jiāo* *jiǎo* *qiāo* *qiǎo* *xiāo* *xiǎo* *shāo* *náo* *nǎo* *nào* *ráo* *rào*) (DeFrancis 1989: 102).

Both radicals and phonetics are derived from independent characters, with the set of possible radicals smaller than the set of possible phonetics. Some characters can perform either function, for example in Figure 9.2, <馬> *mǎ* ‘horse’ is a phonetic in the first row, but a radical in the second row. The ‘horse’ radical demonstrates that semantic variability also exists in radicals, due to diachronic semantic change. While both ‘ride’ and ‘drive’ have direct connections to horses, and ‘camel’ has an obvious similarity as another beast of burden, the others have drifted far from their original meaning.

Over 90 per cent of all Chinese characters are radical–phonetic compounds, but they are not evenly distributed across the lexicon. Low frequency characters are skewed towards radical–phonetic characters, while the highest frequency characters are skewed away. This reflects the general principle that irregularity in language clusters in high frequency forms (because exceptional forms would be forgotten, or go unlearnt, in infrequently used items). Nonetheless, many of the high frequency characters are used as radicals or phonetics in the construction of lower frequency compound characters.

1.3 The written lexicon

The written lexicon of Chinese is vast. It consists of all the lexical items that have a conventional written form in the Chinese orthography. This includes not only the lexicon in

罵 mǎ	碼 má	瑪 mǎ	媽 mǎ	鐫 mǎ
to scold	symbol	agate	mother	masurium
駝 tuó	駁 bó	駐 zhù	騎 qí	駕 jià
camel	to refute	to station	to ride	to drive

FIGURE 9.2 CHARACTERS WHICH INCLUDE THE ‘HORSE’ RADICAL

active or passive use by current writers and readers of standard Chinese, but also lexical items ‘stored’ in literary works, dictionaries and other documents, that are available to be re-admitted to the working lexicon.

Since characters represent morphemes, and most lexical items are multimorphemic compounds, the written lexicon is much larger than the inventory of characters. This is especially evident in the coinage of new terms, which is done by combination and permutation of existing character/morphemes, or by extending the usage of an existing character to novel morphemes. It is often said that to read a newspaper, a reader need know only 2,000 characters (though to understand what is being read, the reader must comprehend many times that number of compounds). Practical dictionaries for daily use contain on the order of 6–8,000 characters, with 50–150,000 compounds. ‘Unabridged’ type dictionaries may contain around 15,000 characters, and the most comprehensive contain nearly 50,000 characters and around 350,000 compounds, most of which are unfamiliar even to a highly educated reader of Chinese.

The nature of the Chinese orthography complicates the relation of the written lexicon to the various vernaculars. An overlapping set of cognate morphemes exists in the various Chinese vernaculars, represented by the same set of characters, therefore novel compounds that are written using combinations of existing characters can spread quite easily from one vernacular to another (or from the Classical literary language into any of the vernaculars, or even from Japanese into any Chinese vernacular). Since the borrowing vernacular already has a standard pronunciation for the characters involved, there is no need to learn a ‘foreign’ pronunciation for the borrowed item, or to ‘respell’ it in accordance with the recipient language’s orthography. For example, in Taiwan, Japanese 中古 *chuuburu* ‘used’, ‘secondhand’ became Southern Min *tiong-kò* and subsequently Mandarin *zhōnggu* without any change in meaning or written form.

2 HISTORICAL DEVELOPMENT

2.1 Development of media

The earliest known examples of Chinese writing are the Shang dynasty Oracle Bone inscriptions of around 1200 BCE. The Oracle Bone inscriptions were incised into the turtle shells and ox scapulae which were used in divination rituals. Since the script is already well developed in the earliest inscriptions, it is likely that the oracle bones represent only the most durable medium of early Chinese writing, not necessarily the earliest and certainly not the most convenient. This is borne out by the appearance of characters written in brush and ink on some oracle bones (Keightley 1978: 46).

The second oldest inscriptions are in another durable medium: Western Zhou bronzes from as early as 1100 BCE. From the Eastern Zhou period (mid-eighth century to mid-third century BCE) comes the first surviving writing done with ink and brush, on silk cloth or wood or bamboo slats. The invention of paper (officially recorded as 105 CE) completed the suite of technology that has served Chinese writing needs for the better part of two millennia; brush and ink on paper is still the aesthetically preferred form of Chinese writing. (Utilitarian writing is done with the familiar pencil, ballpoint, etc.) Indeed, the expressive possibilities offered by brush and ink, and their equal usefulness in writing and painting, are probably responsible for the elevation of calligraphy to high art in Chinese culture.

Relatively cheap paper stimulated the development of printing, first in the form of reproductions of stone inscriptions through rubbing, then as whole-page woodblock prints, and finally as movable type. Despite being at the forefront in the development of

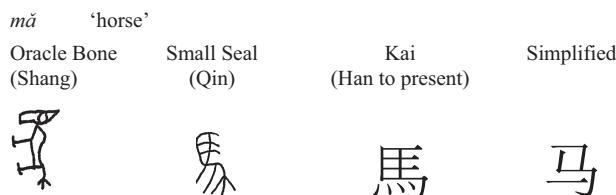


FIGURE 9.3 THE HISTORICAL DEVELOPMENT OF CHINESE CHARACTERS

printing, Chinese fell behind the West technologically in typewriters and early computers, which could not handle the processing demands of thousands of characters. Now, more advanced computers obviate the screen display and printing problems. Use of the alphabetic keyboard originally represented an input bottleneck, but increasingly intelligent predictive text and the development of reliable speech and handwriting recognition systems have removed most impediments.

2.2 Development of the script

In the Oracle Bone and early bronze scripts, some but not all of the originally pictographic characters were already stylized beyond recognition. There was great variation in the writing of individual characters, and in the strokes used to render them. The subsequent development of the script is a process of stylization, standardization and reduction of the process of writing to the repetition of a small number of stereotyped motions (strokes). Curved lines became straight or angled, and pictographic iconicity was completely eliminated.

Following the political unification of China by the first Qin emperor (221 BCE), a standard script was imposed in place of the regional variants that had sprung up. The regularization of the script continued into the Han, by which time the more or less modern script had emerged. Pre-modern forms are still used in some contexts for aesthetic reasons, and various cursive forms have emerged both as convenient shorthands and as calligraphic art forms, but the *Kai* script of the Han dynasty has survived as the model for all subsequent Chinese writing. The most recent change has been the official Mainland Chinese simplifications of the 1950s, which reduced the number of strokes in many characters without fundamentally altering the basic principles of the script (in many cases merely giving official blessing to folk shorthand characters). An example of the historical progression can be seen in Figure 9.3.

2.3 Development of the orthography

Xu Shen's c.100 CE dictionary *Shuowen jiezi*, recognizes two types of characters: *wen* (unitary graphs) and *zì* (compound characters made up of more than one component). They are further classified into six categories (the *liushu*) based on their relation to meaning, to sound and to other characters:

- 1 *zhishi* ('indicate matters'): a diagrammatic representation of a concept (but not a picture of an actual object). For example, the numbers 一 *yī* 'one', 二 *èr* 'two', 三 *sān* 'three' or the position words 上 *shàng* 'above', 下 *xià* 'below';

- 2 *xiāngxíng* ('resembling form'): pictographs, such as 'horse' in Figure 9.3;
- 3 *xíngshēng* ('form–sound'): characters consisting of a radical and a phonetic;
- 4 *huìyì* ('combine meaning'): a character consisting of two components, where both contribute to meaning, and neither is the phonetic. Common examples are 歪 *wāi* 'crooked' (不 *bù* NEG + 正 *zhèng* 'straight'), and 信 *xìn* 'to believe' (人 *rén* 'person' + 言 *yán* 'to speak', 'word');
- 5 *zhuānzhu* ('turned and annotated'): a rare and murky category, apparently involving both etymological and semantic relationships. Xu Shen's example is the connection between 老 *lǎo* 'old' and 考 *kǎo* 'aged', 'old age';
- 6 *jiājìe* ('borrowed'): rebus characters. A character with a particular meaning is used to write a homophonous character of unrelated meaning. For example, 被 *bēi* 'quilt' is used to write the passive marker *bèi*.

The *liushu* classification scheme provides handy categories into which existing characters may be sorted, but should not be viewed as exemplifying the results of six different, unrelated processes of character creation. Chinese writing undoubtedly began with basic characters of types 1 (*zhishi*) and 2 (*xiāngxíng*). Such graphs could function as useful mnemonic devices, but fall far short of full writing because many of the words necessary for full linguistic expression are not conducive to pictorial or diagrammatic representation. The breakthrough that would create a true writing system capable of representing a full range of utterances was the discovery of the rebus principle. Type 6 (*jiājìe*) characters developed when characters of type 1 or 2 began to be used as rebuses to represent homophonous, previously unwritten words. As this principle caught on, the number of words reducible to writing exploded, but so did the number of ambiguous characters – the reader would be unable to tell whether a character represented the original word or a homophonous word represented in rebus fashion. To disambiguate, another character of type 1 or 2 that is semantically related to one of the homophones could be added to differentiate the two words – a radical. The result is a type 3 (*xíngshēng*) character. For example, in Figure 9.4, there is no written form for 'emperor' at Time 1. In Time 2, the use of 'stem' as a rebus allows the creation of a written form for 'emperor', at the same time creating an undesirable ambiguity. The ambiguity is resolved in Time 3 by adding a simplified form of the character 'grass' when the intended meaning is 'stem'.

Thus every use of the rebus principle added to the ambiguity of the system, a problem which could be solved by affixing a radical. Type 3 characters soon outnumbered all other types, and the bipartite radical–phonetic form became the norm. As the number of compound characters increased, they also came to be used as phonetics, resulting in modern characters with more than two components, for example 古 *gǔ* 'ancient' was used as a phonetic (combined with the 'female' radical) to form 姑 *gu* 'father's older sister', which was subsequently used as a phonetic (combined with the 'grass' radical) to form 菇 *gū* 'mushroom'.

Leaving aside the mysterious type 5 (*zhuānzhu*), the above process accounts for all of the categories except type 4 (*huìyì*). The concept of character formation by combination

Time 1	帝	dì	stem, base of fruit	()	dì	emperor
Time 2	帝	dì	stem, base of fruit	帝	dì	emperor
Time 3	蒂	dì	stem, base of fruit	帝	dì	emperor

FIGURE 9.4 THE DEVELOPMENT OF 'FORM–SOUND' CHARACTERS

of semantic primes is intuitively appealing, but such characters are fairly rare, and many purported examples can be revealed to be type 3 characters with obscure phonetics. (Indeed, Boltz (1994), following Boberg (1937), asserts that category 4 characters do not exist, except as isolated exceptions.) The gravest error that casual observers of Chinese writing can make is to see semantic compounding everywhere, treat phonetics as if they were semantic primes, and jump to conclusions about the arcane thought processes of the ancients. For example, Aria (1991: 87) states that 星 *xīng* ‘star’ is:

formed by combining the pictogram ‘sun’ with the character ‘to be born’. Perhaps the ancient Chinese thought that the sparkling points of light they observed in the night sky were like tiny, newborn suns.

Regardless of ancient Chinese astronomical beliefs, the reason ‘sun’ is combined with ‘to be born’ is that ‘sun’ is the radical, and ‘to be born’ *shēng* is the phonetic, being a near homophone with ‘star’ in Old Chinese:

生 *shēng* < sraeng < *srjeng
星 *xīng* < xing < *seng

(reconstructions from Baxter 1992)

The above example raises the question of phonetic latitude in the choice of ‘homophonous’ rebus characters. The syllables are far from homophonous in modern Mandarin, and even in the Old Chinese reconstruction, differences exist. In general, perfect homophony was not a requirement for rebuses; it was sufficient that the nuclear vowel and coda be the same, and that the initials be at least homorganic. Thus the 15 different pronunciations associated with the phonetic <堯> can ultimately be traced back to various kinds of velar obstruents and nasals, occurring with different combinations of glides. For example, Baxter (1992) reconstructs the following with velar nasals:

堯 *yáo* < ngew < *ngew
曉 *xīao* < xew < *hngew
繞 *ráo* < nyew < *ngjew
燒 *shāo* < syew < *hngjew

Two millennia or more of sound change have amplified minor differences to the point that OC near-homophones are sometimes quite distinct, especially in the case of OC initial clusters:

監 *jiān* < kaem < *kram to see, observe
藍 *lán* < lam < *g-ram indigo, blue

(Baxter 1992: 263)

3 WRITING AS EVIDENCE IN HISTORICAL RECONSTRUCTION

The knowledge that a group of characters sharing a common phonetic (called a *xiesheng* ‘harmonizing sound’ series) were near-homophones at the time of character formation is valuable data in the reconstruction of Old Chinese. In alphabetically written languages, the sound values of letters tend to be stable over time and spelling differences between earlier and later forms reflect sound change. But in Chinese, written forms are largely

stable over time, and phonetic information from written forms is purely relational: if words A and B share a certain phonetic, and words C and D share a different one, then A = B and C = D in Old Chinese (with the equal sign interpreted loosely). If modern A = B = C = D, there has been a merger; if A and B are different, there has been a split; if B = C = D but not A, there has been split with merger, etc. To this algebraic skeleton, phonological flesh can be added from other sources.

A more detailed kind of relational information, used in the reconstruction of Middle Chinese, comes from *fanqie* 反切 spellings given in dictionaries and rhyme books, the most important of which is Lu Fayān's 601 CE *Qieyun*. A guide to proper pronunciation of verse, the *Qieyun* gives lists of homophonous characters, and indicates pronunciation by means of two other characters: the first character has the same initial consonant as the target syllable, the second character is identical in all other ways (glide, final and tone.) For example, the character 跨 *kuà* is given the *fanqie* spelling 苦化 *kǔ huà*, the former character representing the [k^h] initial, the latter representing the rest of the syllable. The characters used in *fanqie* spellings create a network of morphemes known to have the same initial consonant or the same rhyme. For example: if A's initial is indicated using B, and B's is indicated by C, and C's by D, then A, B, C and D must all have the same initial consonant.

One further characteristic of Chinese writing is that it is highly etymological. Since the written form is not altered to accommodate sound change, the reflexes of a given morpheme in the various daughter languages are easily identifiable as cognate. Finding cognates between two or more Chinese 'dialects' can be as simple and convenient as asking speakers to 'please pronounce the following characters'. The character represents not just the modern reflex of the ancestral form, but also the etymological root, a tremendous boon to historical inquiry. However, leaning too heavily on this property of Chinese writing can also lead scholars to ignore important colloquial and popular vocabulary that has no standard written form, or to overlook the distorting effect that written forms borrowed from Classical Chinese or other vernaculars can have on the development of the spoken language's lexicon (Norman 1988).

4 RELATION TO OTHER LANGUAGES

For simplicity's sake, the preceding has consistently focused on Chinese writing in relation to a single language, modern Standard Chinese, and its ancestors. Most of what has been said applies to the other languages of the Sinitic branch of Sino-Tibetan (the Chinese 'dialects'), except that many have a weaker tradition of vernacular writing and may have large numbers of morphemes that have no standard written form. The same script is used, and the written lexicons differ depending on the lexicon of the spoken language, but are unified to some extent by cognates inherited from a common ancestor, and by the vast reservoir of lexical items imported from the former common written standard, Classical Chinese. The orthography functions the same way in the various dialects, with the phonetics creating categories that are homophone groups intra-dialectally, cognate sets inter-dialectally. For example, in Figure 9.5 the phonetic <吳> is not intended to associate [u] with [go]: because orthographies are language-specific, this phonetic's value in Mandarin is different from its value in Southern Min. The association between [u] and [go] is a cognate relationship based on common ancestry, and illustrates the history of sound change that this class of syllables has undergone since the creation of these characters (loss of *ŋ-initials in Mandarin, denasalization in Southern Min).

The Chinese script is also used to write other non-Sinitic languages, usually with massive borrowing from the Chinese written lexicon (and, in the case of Japanese, recent

		Mandarin	Southern Min
吳	<i>Wu</i> (<i>a surname</i>)	u ³⁵	gɔ ²⁴
誤	<i>mistake</i>	u ⁵¹	gɔ ²²
娛	<i>amuse</i>	y ³⁵	gɔ ²²

FIGURE 9.5 THE REALIZATION OF A PHONETIC IN DIFFERENT DIALECTS

borrowing in the other direction). Among major literary languages, Japanese, Korean and Vietnamese were all at one time written in Chinese script, though Vietnamese has abandoned it for a variant of Roman, and Korean is phasing out Chinese characters in favour of its homegrown phonetic script. The Chinese orthography's dual morphemic and syllabic nature emerged in the context of monosyllabic morphemes; when the script is applied to a language of significantly different structure, radical orthographic changes are required. Japanese has separated the morphemic from the syllabic principle, with the original script forms (*kanji*) becoming purely morphemic, and simplified script forms (*kana*) becoming purely syllabic. Vietnamese, with a syllabic and morphemic structure closer to Chinese, created new characters for native Vietnamese words in more familiar ways resembling the *liushu*, but with innovative variations (Nguyen 1990). Chinese writing also served as a model for several writing systems that borrowed its outward appearance and some of its orthographic principles, creating Chinese-like scripts that ‘did not borrow already-formed characters, [but] almost completely created them anew’ (Zhou 1989: 45). The best known of these is the script used to write Xixia (or Tangut) (Gong 1985), without which we would have no knowledge of that now-dead Tibeto-Burman language.

FURTHER READING

- Chen, Hsuan-Chih and Tzeng, Ovid J.L. (eds) (1992) *Language Processing in Chinese*, New York: North Holland.
- DeFrancis, John (1984) *The Chinese Language: Fact and Fantasy*, Honolulu: University of Hawaii Press.
- Hannas, William C. (1997) *Asia’s Orthographic Dilemma*, Honolulu: University of Hawaii Press.
- Karlgren, Bernhard (1957) *Grammatica Serica Recensa*, Stockholm: Museum of Far Eastern Antiquities.
- Li, Xueqin (1985) *Eastern Zhou and Qin Civilization* (trans. by K.C. Chang), New Haven, CT: Yale University Press.
- Sampson, Geoffrey (1985) *Writing Systems*, Stanford, CA: Stanford University Press.
- Taylor, Insup and Taylor, M. Martin (1995) *Writing and Literacy in Chinese, Korean and Japanese*, Philadelphia: John Benjamins.
- Tseng, Yuhu (1993) *A History of Chinese Calligraphy*, Hong Kong: Chinese University Press.
- Zhou, Youguang (2003) *The Historical Evolution of Chinese Languages and Scripts*, Columbus, OH: Foreign Language Publications.

REFERENCES

- Aria, Barbara (1991) *The Nature of the Chinese Character*, New York: Simon & Schuster.

- Baxter, William H. (1992) *A Handbook of Old Chinese Phonology*, New York: Mouton de Gruyter.
- Boltz, William G. (1994) *The Origin and Early Development of the Chinese Writing System*, New Haven, CT: American Oriental Society.
- Boodberg, Peter A. (1937) ‘Some proleptical remarks on the evolution of Archaic Chinese’, *Harvard Journal of Asiatic Studies* 2: 329–372.
- DeFrancis, John (1989) *Visible Speech: The Diverse Oneness of Writing Systems*, Honolulu: University of Hawaii Press.
- Gong, Hwang-cherng (1985) ‘Radicals and phonetics in the Tangut script and their generative processes’, *Bulletin of the Institute of History and Philology, Academia Sinica* 56: 719–58.
- Hockett, Charles (1963) ‘The problem of universals in language’, in Joseph Greenberg (ed.) *Universals of Language*, Cambridge, MA: MIT Press, 1–22.
- Keightley, David N. (1978) *Sources of Shang History*, Berkeley: University of California Press.
- Nguyen, Dinh-Hoa (1990) ‘Graphemic borrowing from Chinese: the case of Chu Nom – Vietnam’s demotic script’, *Bulletin of the Institute of History and Philology, Academia Sinica* 61: 383–432.
- Norman, Jerry (1988) *Chinese*, New York: Cambridge University Press.
- Zhou, Youguang (1989) ‘汉字文化圈的文字演变’ (The evolution of writing in the Chinese character cultural sphere), 民族語文 1989. 1: 37–55.

Part 3

TIBETO-BURMAN

Section 3.1: Languages of NE India 213

CHAPTER TEN: The Tibeto-Burman languages of Northeast India	213
CHAPTER ELEVEN: Garo	243
CHAPTER TWELVE: Hakha Lai	258
CHAPTER THIRTEEN: Mongsen Ao	277
CHAPTER FOURTEEN: Karbi	302
CHAPTER FIFTEEN: Tani Languages	322
CHAPTER SIXTEEN: Meithei	338
CHAPTER SEVENTEEN: Tangsa	350

Section 3.2 Bodish Languages 369

CHAPTER EIGHTEEN: Classical Tibetan	369
CHAPTER NINETEEN: Lhasa Tibetan	385
CHAPTER TWENTY: Kurtöp	404
CHAPTER TWENTY-ONE: Tshangla	418
CHAPTER TWENTY-TWO: Dolakha Newar	436
CHAPTER TWENTY-THREE: Kathmandu Newar (Nepāl Bhāṣā)	453

Section 3.3 Tamangic 468

CHAPTER TWENTY-FOUR: Tamang	468
CHAPTER TWENTY-FIVE: Chantyal	494
CHAPTER TWENTY-SIX: Manange	516
CHAPTER TWENTY-SEVEN: Nar-Phu	534

Section 3.4 Rung Languages 557

Section 3.4.1 Rgyalrong Languages 557

CHAPTER TWENTY-EIGHT: Tshobdun Rgyalrong	557
CHAPTER TWENTY-NINE: Cogtse Rgyalrong	572
CHAPTER THIRTY: Stau (Ergong, Horpa)	597
CHAPTER THIRTY-ONE: Japhug	614

Section 3.4.2 Rawang-Dulong-Anong 635

CHAPTER THIRTY-TWO: Dulong	635
----------------------------	-----

Section 3.4.3 Kiranti 646

CHAPTER THIRTY-THREE: Kiranti Languages	646
CHAPTER THIRTY-FOUR: Hayu	680
CHAPTER THIRTY-FIVE: Belhare	696

CHAPTER THIRTY-SIX: Camling 722
CHAPTER THIRTY-SEVEN: Wambule Rai 736

Section 3.4.4 Western Himalayan 756

CHAPTER THIRTY-EIGHT: Sangla Kinnauri 756

Section 3.5 Qiangic 773

CHAPTER THIRTY-NINE: Qiang 773
CHAPTER FORTY: Prinmi 790
CHAPTER FORTY-ONE: Tangut 805
CHAPTER FORTY-TWO: Lizu (Ersu) 823

Section 3.6 Naic [Naxi] 840

CHAPTER FORTY-THREE: Yongning Na (Mosuo) 840

Section 3.7 Lolo-Burmese 856

CHAPTER FORTY-FOUR: Burmese 856
CHAPTER FORTY-FIVE: Zaiwa 877
CHAPTER FORTY-SIX: Akha 885
CHAPTER FORTY-SEVEN: Lisu 902
CHAPTER FORTY-EIGHT: Lahu 918

Section 3.8 Karenic 932

CHAPTER FORTY-NINE: Eastern Kayah Li 932
CHAPTER FIFTY: Pwo Karen 942

Section 3.9 Other Languages 959

CHAPTER FIFTY-ONE: Lepcha 959
CHAPTER FIFTY-TWO: Tujia 974
CHAPTER FIFTY-THREE: Jinghpaw 993

§3.1: Languages of NE India

CHAPTER TEN

THE TIBETO-BURMAN LANGUAGES OF NORTHEAST INDIA¹

Mark W. Post and Robbins Burling

1 INTRODUCTION

‘Northeast India’, or ‘(the) northeast’ as it is usually called in India, primarily refers to the ‘seven sister’ states of Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland, and Tripura.² These seven states are relatively small in area and low in population, but very high in ethnolinguistic diversity. Northeast India is comparatively isolated from the rest of the country, connected as it is only by a narrow (20km wide) strip of land bordered on three sides by Bangladesh, Nepal, and Bhutan.

Northeast India as an ethno-political region is a recent construction. Prior to Indian independence in 1947, administration of most areas within this region was slight, as was post-independence administration, at least until the ‘Chinese aggression’ of 1962 (Guyot-Réchard 2012); the state of Arunachal Pradesh was inaugurated as recently as 1987. Its geopolitical borders are similarly recent, and in some cases (for example, between Arunachal Pradesh and Tibet) remain disputed. Many ethnolinguistic groups of our region also live in neighbouring states; for example, Tshangla speakers are in Arunachal Pradesh and in Bhutan, Garo speakers are in Meghalaya and in Bangladesh, Idu speakers are in Arunachal Pradesh and Tibet, and Tangkhul speakers are in Manipur and Myanmar. Nevertheless, geography, history and geopolitics have conspired to endow our region with a sort of ‘unity in diversity’, such that, with the above caveats regarding its recent origins and porous nature in mind, Northeast India can nevertheless be usefully studied as a whole. And it is in the Northeast Indian region that we find the epicentre of phylogenetic diversity within Tibeto-Burman, at a scale which is perhaps only now beginning to be truly appreciated.

Northeast India has traditionally been host to three major language families other than Tibeto-Burman: Austroasiatic (Khasian, three to six languages³), Tai-Kadai (Southwestern Tai, three to four varieties) and Indo-European (primarily Eastern Indo-Aryan, two to four languages). In recent years, this list has grown to also include Hindi, Nepali, English and Dravidian languages. In addition, we find a handful of Indo-Aryan-based creoles (at least Nagamese and Arunachali Hindi, and there may well be others) and a handful of possible language isolates, about which more later. However, this already substantial phylogenetic diversity is overshadowed by the extraordinary diversity of Tibeto-Burman

in our region, with perhaps 20 independent subgroups, and anywhere from 100 to 300 individual languages (depending on definitions) spoken here.

Geographically, Northeast India centres on the Brahmaputra River, which begins as the Tsangpo in Tibet, descends through the Eastern Himalaya as the Siang, and finally carves out the massive, fertile floodplain of Assam before turning southward and draining into the Bay of Bengal. Surrounding Assam are six ‘hill states’ – a modest way to describe the often towering peaks that rise from the Assam plains to divide India from Myanmar and Tibet at altitudes ranging from 2,000 to 6,000 metres. Geography thus neatly divides our region into ‘hills’ and ‘plains’, a distinction with important, if sometimes imprecise, ethnolinguistic consequences.

The Assam floodplain has long been an ethnolinguistic melting pot. At least 2,000 years ago, Boro-Garo languages appear to have spread and diversified within this region from their likely initial position in the modern-day India/Myanmar northern border region (DeLancey 2012). However, they may have been preceded by speakers of Austroasiatic languages, as suggested by a number of toponyms and areal loanwords (Kakati 1995; Diffloth 2005; Konnerth 2014). The Tai-speaking Ahom invaded and conquered most of modern-day Assam after 1200 CE; they were later followed by a small number of Shan and Jingpho groups, most of whose descendants remained in the river deltas of Upper Assam (Morey 2005). The Eastern Tani language Mising coalesced in later centuries as a result of progressive southward migrations of Siang-area Tani groups following the course of the Brahmaputra River. Other groups have found their way into Assam and its surrounding foothills as a result of first British and later Indian administrative and resettlement policies. In modern times, however, Assam is dominated by several million speakers of Eastern-Indo-Aryan-languages, the outcome of waves of migration from the Bengali-speaking Jamuna/Padma delta region which began in the fourth century CE and continues to the present day (Baruah 1960 [1933]). This has resulted in the emergence of Assamese, a close relative of Bengali which has been considerably reshaped due to its adoption by millions of non-native (mostly Tibeto-Burman) speakers. In modern times, Assamese is spoken by almost everyone in the state of Assam, even if in many cases as a second language; however, significant enclaves of tribal minorities are still found throughout Assam, numbering perhaps more than two million individuals overall.

A similar process has recently, but even more quickly and profoundly, re-shaped the linguistic landscape of Tripura. Two or three generations ago, Tripura State was also predominantly Tibeto-Burman. However, so many Bengalis have now settled there that the indigenous Tibeto-Burman speaking Kok Borok, once the dominant group, have been reduced to a minority in their own homeland.

Meghalaya, a small remnant hill region sandwiched between the plains of lower Assam and Bangladesh, is the only Northeast Indian hill state with a substantial and long established non-Tibeto-Burman population. Khasian groups occupy the eastern three-fifths of this state, forming a bit over half of its population, and speak what may count as four Austroasiatic languages: Khasi, Pnar, War and Lyngngam. Except for the Khasi, the largest group of people in Meghalaya is the Tibeto-Burman Garo, who occupy the western two-fifths of the state together with their Boro-Garo relatives.

Mizoram is dominated by speakers of Kuki-Chin languages, including the namesake Mizo language, but also related languages such as Mara, Lai and Hmar; many of these groups also spill over into the neighbouring countries of Bangladesh and Myanmar, which surround Mizoram on every side except for its much narrower borders with Tripura, Assam and Manipur in the northwest.

Manipur has a more substantial internal diversity. Meitheilon, the language of the ‘non-tribal’ Meitheis, is spoken in and around the state capital Imphal. However, the surrounding hills are dominated by ‘tribals’ who speak languages which mostly fall linguistically within Kuki-Chin, but also a few which are seemingly outside it.

Nagaland has a very high level of ethnolinguistic diversity. While the majority of people in this state are notionally ‘Naga’ in a very loose ethno-politico-linguistic sense, they speak several languages from perhaps four distinct subgroups within Tibeto-Burman (Coupe 2010). They include the Ao, Angami and Sumi, among many others.

The highest level of phylogenetic diversity in our region seems to lie in the state of Arunachal Pradesh. Long a miscellaneous backwater on the fringes of the Ahom and Tibetan empires – though never a formal possession of these or any other regimes – portions of this area were annexed by the British colonial administration in the mid-nineteenth century, and reverted to independent India as the ‘North East Frontier Agency’ in 1947. Thereafter, ‘NEFA’ was administered first from Delhi and later from Shillong and Guwahati, but gained statehood with a capital established at Itanagar-Naharlagun in 1987. Most of modern-day Arunachal Pradesh is still claimed by China, on the basis of the latter’s rejection of the Simla Accord treaty of 1914 between Tibet and British India; China invaded in 1962, and low-intensity border skirmishes still regularly occur. That said, Arunachal Pradesh is internally peaceful, and could not present a more welcoming or more fascinating environment for ethnolinguistic field research. At least 12 independent subgroups of Tibeto-Burman seem to be found inside Arunachal Pradesh, most of them significantly underdocumented. In part due to this high diversity, a Hindi-based creole has been rapidly sweeping the state in recent years. This has led to the sudden and in some cases severe endangerment of several of the state’s indigenous languages (Modi 2006).

2 PREVIOUS RESEARCH

Linguistic research in Northeast India has a complicated history. The first descriptive works were produced by colonial officers and missionaries in the late nineteenth and early twentieth centuries (Robinson 1849; Endle 1884; Lorrain and Savidge 1898; American Baptist Foreign Mission Society 1905; Lorrain 1995 [1907/1910], among others). The sophistication of some of these works is remarkable given their authors’ evident lack of training in and exposure to the linguistics of non-European languages. However, the Latin grammar model typically employed by these authors is inappropriate to most Tibeto-Burman languages of our region; concepts such as ‘Nominative case’ are used to describe the often highly differentiated marking of S/A (Morey 2013b), less-familiar alternations such as person-sensitive aspect marking are generally not identified, and orthographic representations are employed which tend to underdetermine Tibeto-Burman phonologies (for example, tone is usually ignored). Such works therefore remain valuable as historical records, but meet few of the needs of modern scholarship.

This period ended with Indian independence in 1947; for most of the years since then, a strict permissions regime aimed primarily at restricting the activities of foreign religious missionaries – but which has been comparatively much more successful at closing off access to foreign scholars and development non-governmental organizations (NGOs) – together with political unrest has made it nearly impossible for most scholars to conduct research in most areas until quite recently. The resulting gap was to an extent filled by the Central Institute of Indian Languages in Mysore, as well as by the Research Directorates of various states; these agencies have produced a large number of descriptive works, among them Das Gupta (1963), Abraham (1985) and Arokianathan (1987). Unfortunately, these

publications are uneven in quality and reliability, and most suffer from the isolation of their researchers from international scholarship.

In recent years, we have witnessed an increase in descriptive activities by local scholars, many of whom are native speakers of the languages of interest. These include school-books and various other vernacular publications (Karlo *et al.*, 2007a, 2007b), as well as the occasional dictionary (Noram 2008; Choya 2009). Such works are typically privately printed under local sponsorship and sold or otherwise circulated in the immediate area – usually going out of print without the outside world ever hearing about them. Still, the material is available to those who make the effort to seek it out, and even when it falls short of the quality that comparative Tibeto-Burmanists would like, it is the only material we have for many languages, and we must be grateful for it. Similarly, mention should be made of the many descriptive theses and dissertations produced by students at local universities such as Manipur University and Nagaland University, which treat a great number of the languages of their areas. Unfortunately, however, few of them have so far been made available to the wider world.

International scholarship in Northeast Indian languages in the post-independence era proceeded at a trickle, largely due to difficulties of access. Notable exceptions of significant work by international or internationally trained scholars from this period include Burling (1961), Weidert (1987), Chhangte (1993), Sun (1993) and Chelliah (1997). The turn of the century saw this situation begin to change, with the comparatively rapid production of a number of large-scale, modern descriptions (Burling 2004; Jacquesson 2004; Coupe 2007; Joseph 2007; Post 2007; Morey 2010; Konnerth 2014; van Breugel 2014), some high-quality comparative works (Joseph and Burling 2006; VanBik 2007; Button 2011; Bruhn 2014) and the *North East Indian Linguistics* book series (Morey and Post 2008, 2010; Hyslop, Morey and Post 2011, 2012, 2013; Hyslop, Konnerth and Morey 2014). This trend is continuing to gain pace; currently, we know of perhaps 15 large-scale, fieldwork-based descriptive projects led by well-trained researchers which are now in progress. Nonetheless, it must be clearly stated that the majority of languages spoken in Northeast India remain significantly underdescribed, and that overall, most of the descriptive work which has been undertaken over the past hundred years or so has not been completed to the standards of modern-day linguistic scholarship. Since access to all states other than Arunachal Pradesh is now far easier than it once was, since communication between local and international scholars has improved significantly, and since native speakers of many languages have recently obtained degrees in linguistics or have otherwise undergone training in language documentation, we hope and expect that this situation will continue to change.

3 NOMENCLATURE

The task of taking stock of the languages of Northeast India is seriously complicated by their constantly changing nomenclature. With regard to the hill states, the historical record in essence begins when British colonial administrators in the mid-to-late nineteenth and early twentieth centuries first learned about the tribes of the un-administered hinterland from people in areas which were already under their control. This meant that the names that came to be used in English-language publications were often *exonyms* – names used by a group's neighbours rather than by the people themselves. Unsurprisingly, such labels are often pejorative, with meanings along the lines of 'savage', 'wild man' and 'barbarian'. Since independence, more and more groups have insisted upon using their own traditional names, or have even invented new names. People formerly

known as ‘Lushai’ are now called Mizo, ‘Mikir’ have become Karbi, ‘(Plains) Miri’ are now Mising and the ‘Chulikata’, Idu.⁴ Some groups have also realigned their ethnic affiliations, which has complicated the identification and understanding of linguistic relationships. For example, some Rabha and Koch, formerly regarded as separate ‘tribes’, have developed the slogan ‘Rabha are Koch, Koch are Rabha’, hoping to strengthen their political agenda through unity (Karlsson 1997). Such realignments mean that it is not simply the names of groups that change, but also the groups themselves. In this chapter, we will reference ethnolinguistic groups using the terms in current use in Northeast India; however, when older names are commonly found in the literature, they will be included in square brackets; for example, ‘Puroik [Sulung]’.

Terminology is made even more difficult by confusions of ethnic for linguistic affiliation. Fortunately, people themselves often draw ethnic boundaries along linguistic lines, but there are many exceptions. In several cases in Northeast India, speakers of mutually unintelligible languages are understood as belonging to the same ‘tribe’ (for example, Koro and Hruso Aka, whose languages may be as distinct as are English and Hindi). On the other hand, people belonging to different ‘tribes’ can sometimes converse easily (for example, Padam and Mising, which are perhaps no more strongly differentiated than are New York and Midwestern US English). Unfortunately, equations of ethnic and linguistic labels have often led to confusion. A case in point is ‘Naga’: today, the people known as ‘Naga’ typically recognize a common ethnicity, but this may have come only after the British began to use this name to refer to them. This should not fool us into believing that the languages of the ‘Nagas’ must be closely related. In fact, it is clear that the languages spoken by ‘Nagas’ fall into at least two, and possibly as many as seven, distinct subgroups.

Finally, a point which must be made clearly, and prominently: despite the existence of ‘catchall’ labels such as ‘Baric’, ‘North Assam’, ‘Kamarupan’ and now ‘Northeast Indian’, the languages of Northeast India have never been demonstrated to form a genealogical unit within Tibeto-Burman – nor, we believe, can they ever be. Even languages which are close geographical neighbours in our region sometimes differ so radically in their vocabulary and grammar that one must naturally raise the question of whether they should be relatable at all, or whether such resemblances as do occur might not instead be better understood as an outcome of areal sharing (Blench and Post 2012). Although we should welcome any and all data-oriented contributions to the subgrouping of Northeast Indian languages, we must bear in mind that the null hypothesis in historical linguistics is *lack of relationship* unless one can be demonstrated (Thomason and Kaufman 1988: 201–2; Nichols 1996).

4 TYPOLOGY

Few if any typological characteristics set the languages of Northeast India clearly apart from those spoken elsewhere, such that there would be any real basis for considering Northeast India as a whole to count as a ‘linguistic area’ in the traditional sense of Emeneau (1956). That general fact notwithstanding, we nevertheless find dozens of convergence zones scattered throughout our area, in which languages of quite different genealogical statuses do indeed end up sharing certain features, and this is certainly due to widespread and longstanding language contact. For example, Assamese has alveolar consonants rather than the standard Indic dentals and retroflexes, and it is the only Indo-European language with a fully developed classifier system. On a smaller scale, we find evidence of Pnar contact in Karbi (Konnerth 2014), and very similar morphosyntactic profiles among Tibeto-Burman subgroups which seem genealogically distant – for example,

between Boro-Garo, Ao, Tani and Idu-Tawrā. There is no doubt that language contact studies will eventually play a significant role in our understanding of the typologies of and relationships among languages of the Northeast Indian region. However, such studies are really only just beginning to get off the ground.

As in most of the Tibeto-Burman phylum, as well as in the local Indo-Aryan languages (though not in the Austroasiatic and Tai lects), predicates in our area are almost always clause-final. While the widespread occurrence of adpositional relational marking means that other clause constituents can be freely ordered, the most common information structure is topic–focus, and the most common order of main clause constituents is [A O V]; obliques and adjuncts occupy a variety of positions depending on the language. Core argument marking is variable in the region: Tani languages are resolutely accusative (Post and Sun this volume), while Mongsen Ao has ‘optional’ agent-marking (Coupe 2007: 154) and Turung Singpho has semantically oriented non-agentive marking (Morey 2013b: 175). Morphological typology is primarily agglutinating in our region, especially in the predicate; here, one very salient feature in our area is the existence of large systems of ‘predicate derivations’, which expand and modify the predicate with a variety of meanings, including result and manner, lexical aspect/aktsionsart, direction, modality and valence change. Such systems are found in Ao, Boro-Garo, Karbi, Meithei and Tani, among others, and may have anywhere from a few dozen to three or four hundred members. Most languages in the northern, northwestern and southern parts of our region exhibit relatively recent grammatical morphology, almost always suffixes. Some older prefixes are also in evidence; languages to the east/southeast of our area tend to preserve more prefixes, and exhibit a more archaic morphology more generally (DeLancey 2014). A few relatively isolating languages are found in or near to Tai contact areas, for example Turung Singpho, but even they typically exhibit at least some morphology.

As in much of Tibeto-Burman, the syllable is a crucial phonological unit in most Northeast Indian Tibeto-Burman languages, and the inventories of syllable initial consonants are different from and larger than those of finals. Bilabial, alveolar, and velar stops often occur at both the beginning and end of a syllable, but voicing and aspiration are usually only contrasted syllable- or word-initially. Nasals have a freer distribution, and generally include bilabials, alveolars and often palatals and velars, although a few languages, notably Angami, lack final consonants entirely (though it compensates with a comparatively rich array of initials, and a relatively large tone inventory (Blankenship *et al.* 1993)). Affricates and sibilants vary widely, from one or two positions, as in Mising, to the much larger inventories of Idu, Hruso and Miji; bilabial fricatives are notably found among several languages of western Arunachal Pradesh. Liquids typically include *j* and *r* or *ɻ*, and sometimes also *w*. Many stereotypically South Asian features, such as retroflex consonants, are rare among the indigenous languages of our area. A few languages along the eastern border allow such relatively exotic features as pre-consonantal or voiceless nasals (in addition to voiceless *r* and *l*). In addition to the stops and nasals, many languages allow a final glottal stop, often transcribed by *h* or *q* in the now widely used Roman orthographies.

Vowel systems vary from quite simple systems, such as the seven-vowel monophthongal system of Tani languages or the five monophthongs and two diphthongs of Garo, to those with nine or ten simple vowels plus several diphthongs, as in Phom. High central or back unrounded vowels, often represented by *ü*, *ɨ* or *w* in the Roman orthographies, are a characteristic feature of many of these languages, as is *ə*, often represented by *ɛ* or *v*.

Tone remains a vexing issue in the study of Northeast Indian Tibeto-Burman languages. Despite the clear presence of lexical tone in many (perhaps most) area languages,

their analysis is often difficult. This seems to be due on the one hand to systemic factors: many Northeast Indian tone languages have a basically polysyllabic phonological word, making it difficult if not impossible to isolate tone-bearing syllables in elicitation; furthermore, tone often interacts in these languages with other prosodic features such as rhythm and intonation, adding additional layers of complexity to the analysis (Evans 2009; Teo 2014; Post 2015). On the other hand, as Morey (2014) writes, tone often carries a lower ‘functional load’ in Northeast Indian Tibeto-Burman languages than it does in, for example, neighbouring Tai-Kadai languages (also see Konnerth and Teo 2014). This lower functional load can significantly increase variability and complicate analyses. Despite the widespread occurrence of lexical tones in our area, there are several areas in which tones are no longer found. It is not uncommon to find two closely related neighbouring languages, one with tone and the other without; for example, Upper Minyong Adi has tone and the mutually intelligible language Pasi-Padam (Lower Adi) – spoken along the border with Assam – does not. Similarly, while the plains language Boro has tone, its close genealogical neighbours Garo and Atong – spoken nearby to the toneless Khasian lects – do not. Such languages nonetheless are in the minority.

5 CLASSIFICATION

The remainder of this chapter will be devoted to an updated presentation of the classification scheme of Burling (2003). It is ‘updated’ in the sense that it takes stock of most of the research which has taken place in the period since 2003. However, it is not updated in having improved significantly upon Burling’s (2003) method of lexical comparison. Nor, in most cases, have we been able to obtain better data. This is unfortunate, because until we have more, and especially more reliable, lexical *and* grammatical data, it will be impossible to discover regular phonological and morphological correspondences among members of many putative subgroups. Because of this, it will remain impossible in many cases to know whether particular lookalike correspondences are due to common inheritance, to areal diffusion, or to coincidence.

A further problem is presented by the morphology of many Northeast Indian Tibeto-Burman languages, such as most ‘Naga’ languages, Boro-Garo, and all Tibeto-Burman groups west of the Lohit River in Arunachal Pradesh. These languages in general lack the deeply grammaticalized and paradigmatic morphology such as agreement affixes and case alternations that are usually viewed as necessary to the firm establishment of genealogical relationships (Thomason and Kaufman 1988: 201–2; Nichols 1996). Instead, the (albeit copious) grammatical morphology of many such languages seems to be mostly secondary, probably quite recent, and in many cases transparently traceable to lexical roots. In several such cases, we suspect that histories of contact-induced change probably led to the loss of archaic morphology (DeLancey 2010a, 2013a).

Nonetheless, some progress has been made. A handful of solid comparative studies have recently emerged, particularly from the Indo-Myanmar borderlands (e.g. VanBik 2007; Bruhn 2014), some older comparative studies have now been field-checked and either verified or improved upon (e.g. Sun 1993; see Post and Sun, this volume), and we now have adequate data for many languages which had previously been all but unknown (e.g. Morey 2010). Comprehensive, fieldwork-based description of individual languages and low-level subgroups is a slow, laborious and expensive process. However, these studies will eventually yield the data that are needed to form more defensible hypotheses concerning higher-level relationships than may be found in much of the earlier literature.⁵

In what follows, we will take an approximately clockwise, geographic approach to the classification of Northeast Indian Tibeto-Burman languages. We take this approach for two reasons: first, to avoid implying that certain groups are more significant than others in a comparative context, and second, to highlight the relationship between geographical clustering and genealogical grouping that is found in our area. In several cases (e.g. Tani and Milang, or Kaman and Meyor) we do not yet know whether observed similarities are due to common inheritance, to contact or to both.

5.1 Northern Area

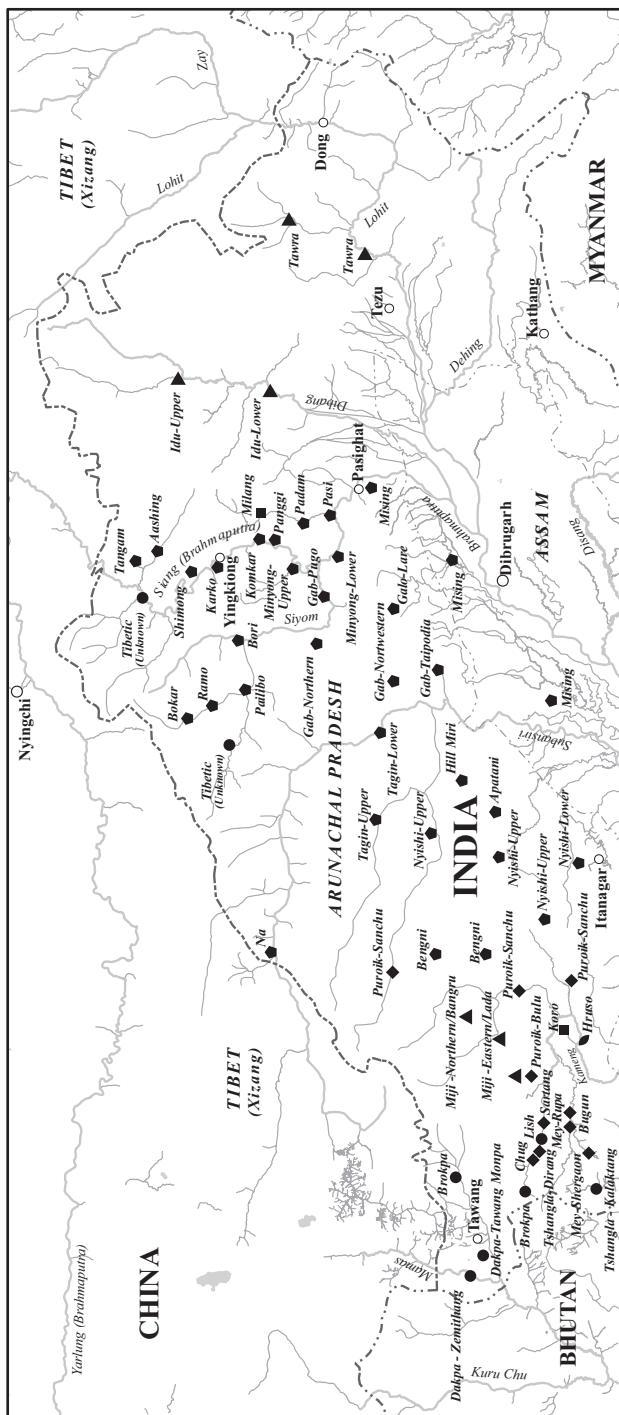
Our ‘Northern Area’ extends from the far northwestern border of Arunachal Pradesh with Bhutan and Tibet to the far northeastern border of Arunachal Pradesh with Tibet and Myanmar (Map 10.1). The far northwestern area is primarily populated by speakers of Tibetic, East Bodish and Tshangla lects, whose main bodies of speakers lie outside Northeast India in present-day Bhutan and Tibet.⁶ These languages are reviewed by Thurgood in Chapter 1 of this volume, and so will not be discussed further here. We will discuss the remaining autochthonous groups and their languages in a basically eastward progression.

5.1.1 Kamengic and Sulung, or Kho-Bwa

A diverse and little-known group of languages is found between the predominantly Buddhist far northwest mentioned above and the more central Tani area to be discussed later. Some of these languages have been suggested to form a subgroup, initially in a tentative observation by Sun (1992) and later more directly by van Driem (2001: 473–6), who credited Roland Rutgers and coined for them the label ‘Kho-Bwa’ (based on Rutgers’ unpublished reconstruction of the putative proto-roots **kho* ‘water’ and **bwa* ‘fire’). These languages are Mey [Sherdukpen], Sartang [But/Boot Monpa], Lish(pa) or Khispi, Chug(pa) or Duhumbi, Bugun or Kho(w)a, and perhaps also Puroik [Sūlung/Sulung].

Mey comprises the two dialects spoken by perhaps 3,000 people in Shergaon and Rupa villages, south of Bomdila in West Kameng district, Arunachal Pradesh. Dondrup (1988) describes the Shergaon variety; no published description of the Rupa variety yet exists. Sartang is spoken in a cluster of four villages by around 2,000 people (Blench and Post 2012); some lexical data from Sartang may be found in Abraham *et al.* (2005). Lish and Chug are two mutually intelligible lects spoken in a small cluster of villages to the northwest of the Mey-Sartang area; Bodt (2014) estimates their population at around 2,000–2,500. Bugun, spoken by around 800 people in a number of villages in West Kameng district, is described by Dondrup (1990); lexical data also appear in Abraham *et al.* (2005). Blench and Post (2012) present a table of lexical data from Mey, Sartang, Chug-Lish and Bugun, which strongly suggest that these languages should form a subgroup. Naming this group ‘Kamengic’, they bifurcate it into Bugun on the one hand, and Chug-Lish, Sartang and Mey on the other.

Puroik [Sulung] is the common name for a cluster of closely related languages spoken by a widely distributed group of traditional foragers primarily inhabiting the Kurung Kumey, East Kameng and Papum Pare districts of western Arunachal Pradesh, as well as adjoining areas in Tibet. Lieberherr (2015) finds three Puroik ‘dialects’ with the autonym *purui?* in the easternmost and largest Sanchu area, and *purun* and *prin(du)* in the smaller, westerly Kojo-Rojo (Lada) and Bulu areas respectively. Sanchu is said to be a dialect chain, however the Kojo-Rojo and Bulu varieties are said to be mutually unintelligible both with Sanchu dialects, and with one another (Lieberherr 2015). The Sanchu variety



MAP 10.1 LANGUAGES OF THE NORTHERN AREA
Legend: ● Tibetan, East Bodish and Tshangla ◆ Kamengic and Puroik ▲ Miji ■ Hruso ◆ Koro-Milang ◆ Idu-Tawrā

is called *Sūlóng* [苏龙] in China, but this is considered derogatory in India. The sources on Puroik include Tayeng (1990), Sun (1991) and Remsangpuia (2008), and Li (2004) is a major descriptive work in Chinese. Lieberherr (2015) maintains Rutgers' and van Driem's (2001) view that Puroik is aligned with Kamengic within 'Kho-Bwa', but this remains to be demonstrated. Furthermore, even if Puroik's alignment with Kamengic languages should prove tenable, it must be acknowledged that Puroik is considerably more different from Kamengic languages than any of them are from each other. Puroik also exhibits both a large number of regional loanwords and a very large stratum of vocabulary which is sparsely distributed, if it may be found at all, in Kamengic and, indeed, in Tibeto-Burman languages more generally. We look forward to further developments, but for now retain the conservative view that Puroik's genealogical affiliation remains undetermined.

5.1.2 *Hruso and Miji [Hrusish]*

Hruso [Aka, Hrusso] and Miji [Dhammai/Dhímmai] are spoken to the east and north of Bugun and Mey in West Kameng, East Kameng and Kurung Kumey districts of Arunachal Pradesh. There may be around 3,000 Hruso speakers, while the more widely distributed Miji-speaking population may be several thousand. These languages were aligned by Shafer (1947), who called Miji 'Hruso A', and Hruso proper 'Hruso B'. In fact, they are not at all dialects of the same language as Shafer implied, as a comparison of Simon's vocabularies of Hruso (1970) and Miji (1979) makes clear.

We know next to nothing about Hruso dialectology. A little more is known about Miji. Blench (2013) divides Miji into three Miji dialect areas, a division with which Bodt and Lieberherr (2015) concur. Blench's Western or Nafra dialect is the variety described by Simon (1979), who reports his speakers' autonym as Dhammai/Dhímmai. Blench's Eastern (Lada) dialect is undescribed, so far as we can see. Blench's Northern dialect is more widely known as Bangru [Levai], which was first reported by Sun (1993) and partially described by Ramya (2011, 2012), as well as by Bodt and Lieberherr (2015).

Do Hruso and Miji form a subgroup? Most scholars have assumed that they do, however it is far from obvious. There are clear resemblances in the lexicon; however, since these languages have long been in contact, a certain amount of sharing ought to be assumed. Until we have comprehensive and reliable descriptions of both languages, the exact relationship between these languages will probably continue to remain mysterious.

5.1.3 *Koro-Milang, or Siangic*

The underdescribed language of perhaps 1,000 Koro people was first reported in Grewal (1997), and later in Abraham *et al.* (2005) and Anderson and Murmu (2010). While Koro has been ethnically aligned with Hruso, these sources demonstrate that Koro is linguistically quite distinct. Post and Blench (2011) found similarities between Koro and Milang, a language spoken by around 2,000 people to the east of the Siang river which had previously been classified as a Tani language (see later); they proposed a tentative grouping 'Siangic' to account for these correspondences. Several possibilities remain open: it may be that Koro and Milang constitute a small subgroup within Tibeto-Burman. Or, it may become possible to align Koro-Milang with Tani (and perhaps also Idu-Tawrā) into a larger unit. A third possibility is that 'Siangic' correspondences between Koro and Milang reflect a shared substrate, found to a lesser extent in Tani languages (as suggested by Post and Blench 2011).

5.1.4 Tani [*Mirish, Misingish, Abor-Miri-Dafla*]

The Tani languages (previously known by the pejorative labels ‘Abor’, ‘Miri’ and ‘Dafla’) form the largest language group in Arunachal Pradesh. Around 600,000 speakers of different Tani languages are found in a basically continuous area from the Kameng to the Siang rivers. In upper Assam, we find around 500,000 speakers of a single Tani language, Mising, settled along the banks of the Brahmaputra River. Much smaller numbers of Tani speakers are found in contiguous areas of Tibet.

Although Sun (1993) divided the Tani languages into two major subgroups, Western Tani and Eastern Tani, he also found that the Tani languages in general constituted a near-perfect dialect chain. Research since 1993 has validated Sun’s conclusions, and found language contact to be a major factor in the evolution of Tani languages, with major consequences for their classification.

Sun’s Western Tani group includes the Galo language (Post 2007), the neighbouring and similar Hill(s) Miri (Simon 1976), several partially described Nyishi, Na/Bangni and Tagin dialects (Das Gupta 1983; Goswami 1994; Tayu 2010) and the more sharply differentiated plateau language Apatani (Post and Tage 2013; Apatani Language Development Committee 2015). To this group we can now add Pailibo (Badu 1994).

Eastern Tani languages comprise the chain of Siang-area dialects known colloquially as ‘Adi’ (literally, ‘hill (people/language)’), including Padam, Panggi, Pasi, Minyong, Komkar, Karko, Shimong and Bori. The plains language Mising, while not ‘Adi’ on ethnic grounds, is nonetheless mutually intelligible with southern Adi lects. In the centre-north of the Tani area we find a number of underdescribed lects which seem basically transitional between Western and Eastern Tani languages; they include Bokar-Ramo (Megu 1990), Aashing (Megu 2003), Damu (Ou-Yang 1985) and Tangam (Badu 2004). Finally, although Milang had been previously classified as Tani, Post and Modi (2011) show that it must descend from a higher position within Tibeto-Burman, and may furthermore contain a substrate of unknown origin (as was also mentioned by Sun 1993).

5.1.5 Idu-Tawrā, or Mishmic [*Digarish*]

Idu [Chulikat(t)a, Bebejia, 义都珞巴 Yìdōluòbā/Lhoba] and Tawrā [Tarað(n), Digaru, 达让儈 Dárlàng Dèng] were traditionally grouped with Kaman in India under the common ethnonym ‘Mishmi’. However, while Idu and Tawrā seem to form a linguistic unit, Kaman falls outside it. This makes comparison with the Chinese sources difficult, as while Tawrā and Kaman are aligned in China as the Dèng nationality, the Idu are lumped together with Tani groups in the Luòbā nationality.

The total number of Idu-Tawrā speakers is unknown, but may be in the neighbourhood of 20,000. Most Idu speakers are scattered in hills surrounding the Dibang River valley in far northeastern Arunachal Pradesh, while most Tawrā speakers are closer to the Digaru River to the southeast of the Dibang. There are several sources for these languages (Talukdar 1962; Philology Section of North East Frontier Agency Research Department 1963; Pulu 1978; Sun *et al.* 1980; Sastry 1984; Pulu 1991; Sastry 1991; Pulu 2002). Unfortunately however, they depart markedly from one another, and there is no easy way to reconcile them. A small number of local publications have also been produced in India (Krisiko 2006; Tawsik 2014), but we have not been able to obtain copies.

It is clear that Idu and, especially, Tawrā share a relatively high percentage of likely cognates both with Tani languages and with Koro-Milang (Sun 1993; Modi 2013). This is striking when one considers that while Idu is spoken in areas which are geographically

contiguous to Tani and Milang speakers, Tawrā is not. While the possibility of a higher-level Idu-Tawrā/Koro-Milang/Tani macro-alignment thus naturally arises, the existing descriptions of most of these languages are not of a quality that will enable us to distinguish genealogical inheritance from the effects of language contact.

Blench and Post (2012) proposed the label ‘Mishmic’ for Idu-Tawrā in favour of the exonymic ‘Digarish’. However, Mishmic may also be problematic given the ethnic alignment of Kaman within ‘Mishmi’. We therefore retain the more conservative label ‘Idu-Tawrā’ in this chapter.

5.2 Central Area

The Lohit-Brahmaputra river valley roughly divides our Northern Area from our Central Area geographically, but also to a degree linguistically. In the Central Area, we begin to find the complex and probably conservative agreement systems which are absent from our Northern Area. They are not found in all Central Area languages – notably, they are not found in the Boro-Garo branch of Sal languages; see later. However, they are found in all major branches of our Central Area, and seem to represent a cleavage among languages in our region (Map 10.2).

5.2.1 Kaman-Meyor, or Lohitic [Midzuish]

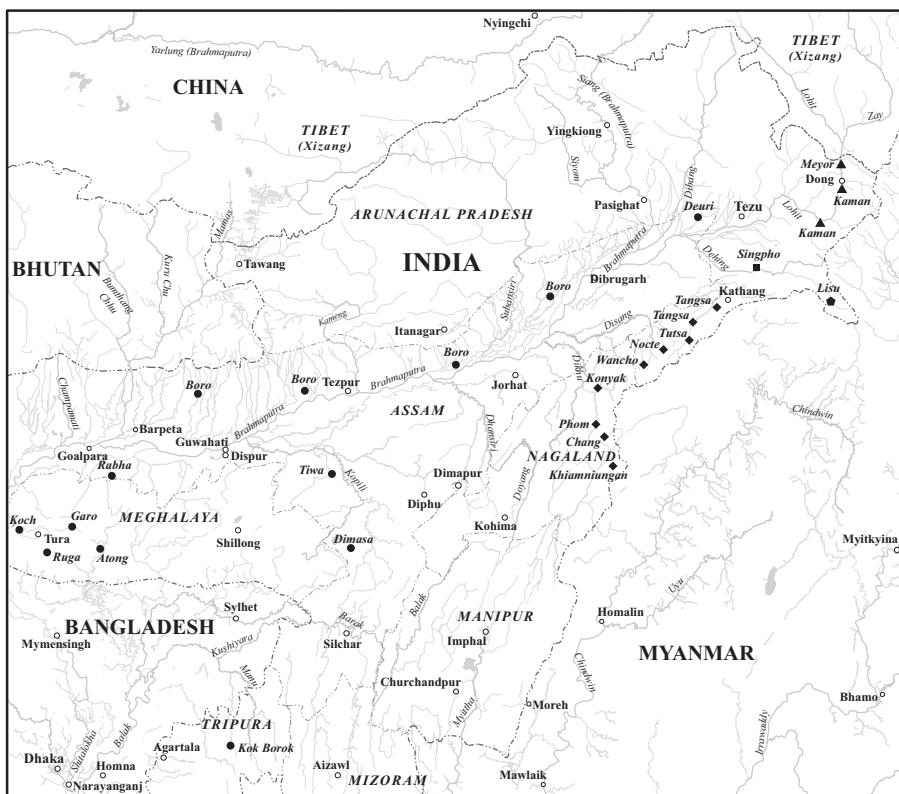
Kaman [Midzu, Miju, 格曼儈 Gémàn Dèng] is spoken by perhaps 2,000 people in the Lohit River area of Arunachal Pradesh, as well as in Tibet (Das Gupta 1977; Boro 1978; Li 2002). Kaman speakers are ethnically aligned with Idu-Tawrā under ‘Mishmi’, however while the Kaman are indeed culturally similar to the Tawrā, their language is very different.

Further up the Lohit River we find the Meyor [Zakhiring, Zaiwa, 扎 Zhá] language, first reported by Grewal (1997) and partly described by Li and Di (2001) in Tibet and Landi (2005) in Arunachal Pradesh. In Arunachal Pradesh, Meyor is reported by Landi (2005) to have been spoken by only 376 people in a handful of villages in Kibithoo and Walong circles of Anjaw district in 2001, while the 2011 *Census of India* reports a total of 989 Meyor tribespeople (Census of India 2011). Presumably, then, there are no more than 1,000 Meyor speakers in India.

Kaman and Meyor share several lexical and grammatical features, including a complex verbal agreement system (not found in Tani, Koro-Milang or Idu-Tawrā), a forward-harmonizing prefixal negator *mV-*, a non-agentive or accusative marker *wi/wik*, and most pronouns: compare Kaman *ki/kin* ‘1SG/1PL’, *nu/nin* ‘2SG/2PL’ and *wi/win* ‘3SG/3PL’ with Meyor *ko/ki*, *no/ni* and *wo/wi*. Before jumping to conclusions, it should be noted that the two languages are areally contiguous, and can be expected to share at least some similarities due to contact. Furthermore, the verbal agreement paradigms, while comparable, seem to exhibit mostly non-cognate morphology. Therefore, while it does seem that genealogical alignment of Kaman and Meyor could eventually be substantiated (as was also suggested by Jacquesson 2001), the matter is not straightforward; thorough and reliable descriptions of both languages are required. For this group, we could retain the conservative label Kaman-Meyor, or adopt a common geographical label Lohitic.

5.2.2 Sal [Boro-Konyak-Jingpho,⁷ Brahmaputran]

The Sal languages,⁸ also known as Boro-Konyak-Jingpho, constitute a major subgroup of Tibeto-Burman, and form the largest recognizable Tibeto-Burman subgroup in Northeast



MAP 10.2 LANGUAGES OF THE CENTRAL AREA

Legend: • Sal: Boro-Garo ◆ Sal: Northern Naga ■ Sal: Jingpho ▲ Lohitic ♠ Lolo-Burmese

India. In Benedict's (1976: 177) view, the Sal languages (represented in his analysis only by Garo and Jingpho) constitute one of the earliest major splits in Tibeto-Burman. Sal has three main branches, which would probably be best labelled Northern Sal, Eastern Sal and Western Sal. However, the labels most commonly used in the literature for these branches are Northern Naga, Boro-Garo and Jingpho-Asakian, respectively. While we hold out hope that the Sal-derived labels might eventually be adopted, to avoid confusion we will maintain the more traditional labels here.

Generally speaking, Northern Naga and Boro-Garo seem slightly closer to one another than either is to Jingpho-Asakian. However, while Jingpho and many Northern Naga languages exhibit complex agreement paradigms, agreement is absent from Boro-Garo and from Asakian languages. It seems likely that lacking agreement among Sal branches and individual languages represents a secondary development (DeLancey 2010b, 2012), however an alternative view is advanced by LaPolla (1992, Chapter 2 in this volume).

Of the three branches of Sal, Boro-Garo [Boro, Boro-Koch, Barish] is the most firmly established, having been recognized at least since the time of the *Linguistic Survey of India* (Grierson 2005 [1903]). Boro-Garo in turn divides into four parts: Boro-Tiwa-Dimasa-Kokborok ('Boro'), Garo, Rabha-Koch-Ruga-Atong ('Koch') and Deori. Data in

Burling (1959) suggest that Garo is closer to Boro than it is to Koch, meaning that '(Deori-)Boro-Koch' would probably have been a better label for this branch. Again however, since the label 'Boro-Garo' has become established in the literature, we will continue to use it.⁹

Deuri [Deori, Chuti(y)a], described by Jacquesson (2004), is spoken by a small number of people in Upper Assam, to the east of the Boro area near Dibrugarh. Deuri is the most strongly differentiated language within Boro-Garo, having lost all syllable-final stops and most nasals, at least some of which are retained in all other languages. It also lacks a large number of cognates which are widespread in the other groups (Burling 2012).

The Boro branch includes several languages spoken both in the valley and in the hilly areas to the south. These have been known by a variety of names, but seem to constitute four languages: Boro [Kachari] in the lower Assam plains, mostly to the north of the Brahmaputra river, Tiwa in the northwestern foothills of Meghalaya, Dimasa in the area spanning southwestern Meghalaya, Dima Hasao and Karbi Anglong districts in Assam and Dimapur in Nagaland, and Kokborok (including Usui and Riyang) to the south, primarily in the state of Tripura. A bifurcation of Tiwa-Boro vs Dimasa-Kokborok is motivated by differences in the proto-syllable-finals they have lost; for example, PBG $*-k \rightarrow \emptyset$ in Tiwa-Boro but $*-t \rightarrow \emptyset$ in Dimasa-Kokborok (Burling 2012). Morān (Gurdon 1904) is an extinct language which retains all proto-finals. It seems to belong in the Boro branch, but it is not clear where.

Garo, described by Burling (2004) is close to Boro, though not as close as any of these languages are to one another. Unlike Boro languages, Garo lacks tones. Though spoken over a wide area in Western Meghalaya and neighbouring areas of Assam and Bangladesh, all dialects of Garo are reasonably mutually intelligible. Finally, Rabha, Koch, Ruga and Atong form a compact group of languages spoken at the western end of Meghalaya and the neighbouring plains. Of these, Rabha and Atong are the most fully described (Joseph 2000, 2007; van Breugel 2014). The Rabha-Koch-Ruga-Atong branch is distinguished by its seeming loss of the nominalizing $*gV-$ prefix surveyed by Konnerth (2012), as well as by some distinctive vocabulary.

The Northern Naga group was originally dubbed 'Eastern Naga' by Konow (1902) and Grierson (Grierson 2005 [1903]), although it has been most persuasively identified and described by French (1983).¹⁰ Burling (2003) labelled this group 'Konyak', however this label properly refers to only one language within the Northern Naga group. Spanning the border regions of far-southeasterly Tirap and Changlang districts of Arunachal Pradesh, northern Nagaland, and adjoining areas of Myanmar, these languages are considerably more heterogeneous than are their neighbours the Boro-Garo languages, and many are currently being identified and investigated by fieldworkers for the first time.

A preliminary division of the Northern Naga group might identify anywhere from six to eight branches, although some of these are much more internally diverse than others.¹¹ In Changlang and Tirap districts of Arunachal Pradesh, these are Tangsa [Tangshang in Myanmar], Nocte [Namsangia], Tutsa, and Wancho [Muthun, Banpara]. In adjacent areas of Nagaland, we find Chang, Khamniungan, Konyak [Tableng], and Phom [Chingmengnu, Tamlu]. A number of seemingly relatable languages are spoken over the modern-day international border with Myanmar, and have recently been added to the *Ethnologue* as Kyan-Karyaw, Lao, Leinong/Lainong, Makyen and Ponyo-Gonwang following a survey conducted by SIL International (Lewis *et al.* 2015). Determination of their exact position in Northern Naga will have to await public disclosure of the survey team's data.

Among these branches, French (1983) offered convincing evidence that the Tangsa languages form a subgroup together with Nocte, and to this group we may add Tutsa. Almost all of these languages as spoken in Northeast India exhibit the clause-final

‘agreement words’ also found in Jingpho (related languages in Myanmar are reported as mostly lacking them; see Morey this volume). Morey (2013a) has recently suggested that the distinction between Tangsa, Nocte and Tutsa may in fact be basically geographical, and that a comprehensive documentation of their respective varieties may reveal a dialect chain. Whatever the case, it seems clear that these lects constitute a defensible genealogical unit. More cautiously and a bit less convincingly, French (1983) has suggested that Wancho and Chang are particularly close to each other, as are Konyak and Phom, and to this group Burling (2003) added Khiamniungan (Kumar *et al.* n.d.). Within this as-yet unsubstantiated group, the position of Wancho in particular seems unclear, and on the basis of vocabulary we might be equally inclined to align it with the Tangsa-Nocte-Tutsa group. Here, and in several other cases, it must be underscored that a very large number of Northern Naga languages are only now being identified and described by fieldworkers for the first time; it might be prudent to wait for their results before wading into the issue of subgrouping any further. Furthermore, it is possible that structural similarities both between Chang, Konyak, Phom, Khiamniungan and Wancho and with their southerly ‘Naga’ neighbours (for example in lacking ‘agreement words’) owe to the effects of language contact rather than to any special genealogical alignment.

The classification of Jingpho [Jinghpaw, Kachin], which is represented by Singpho in Northeast India, has featured especially prominently within Tibeto-Burman studies. While it has been aligned with Lolo-Burmese and with Nungish, these earlier alignments seem to have resulted from confusion due to the large number of Burmese and Jingpho loans in these and other area languages (both Burmese and Jingpho are widely spoken as lingua francas within Myanmar). However, the special similarity of Jingpho to Boro-Garo and the Northern Naga group was noted by Benedict (Benedict 1972, 1976) and Burling (1971, 1983, 2003) and considered most carefully by French (1983). French’s analysis suggests that Jingpho is a bit less similar to Boro-Garo and Northern Naga than the latter two are to each other. However, Jingpho’s closer relationship to the Asakian [Luish]¹² languages Kadu, Andro and Sengmai was noticed long ago by Benedict (1972: 5); this was recently further substantiated by Matisoff (2015), on the basis of new descriptions of the related languages Sak by Huziwara (2008) and Kadu by Sangdong (2012).¹³ A plausible family tree of the Sal languages would thus place Jingpho and Asakian together and coordinate with Northern Naga, with Boro-Garo as a third branch – as was also recently concluded by Matisoff (2015).

Except for Karbi and some settlements of Rengma, both of which are Tibeto-Burman, and Khasi, which is Austro-Asiatic, the gaps between the Boro-Garo languages are mainly filled by Assamese and Bengali. The natural presumption is that the Boro-Garo languages once had a more continuous distribution than they do now, but that they have been separated by the spread of Indic languages. It seems probable that earlier forms of Boro-Garo were, at one time, the predominant languages of the Assam valley and perhaps of some parts of northern Bengal as well. DeLancey (2012) argues that Proto-Boro-Garo may have in fact developed as a lingua franca within the Brahmaputra valley; we find this thesis compelling, as it would explain both the modern-day distribution of Boro-Garo languages and their simplified morphological profile by comparison with their more conservative Northern Naga neighbours.¹⁴

5.2.3 *Lolo-Burmese*

Lolo-Burmese languages are hardly endemic to Northeast India in general, however it appears that around 200 ‘stateless’ Lisu speakers (called ‘Yobin’ in India) are living in far

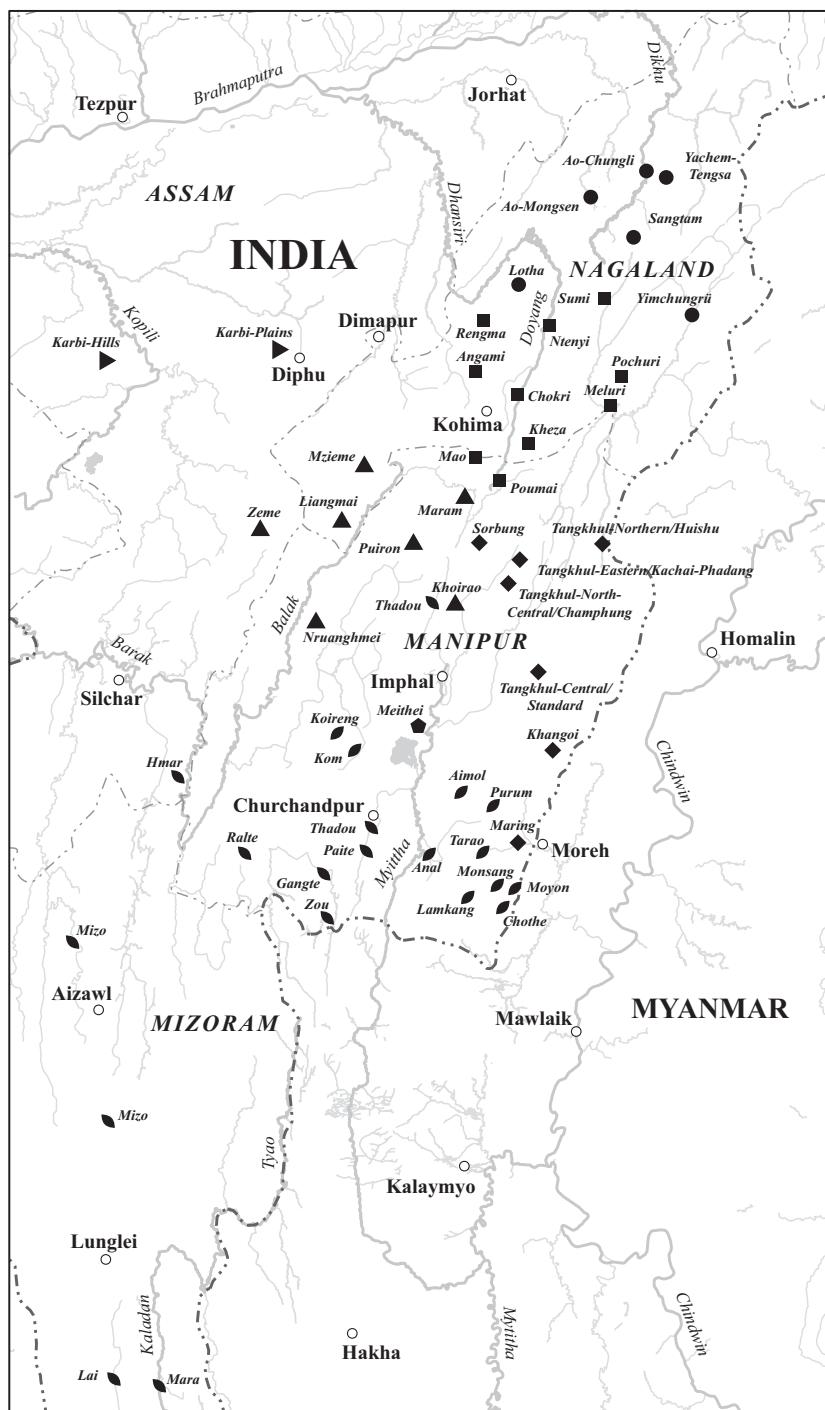
eastern Changlang District's Vijoynagar Circle, in the upper Dihing River valley, where they are surrounded by resettled non-indigenous populations and Indian military installations. An ethnographic description and a language guide corresponding to this group have been produced by Maitra (1988, 1993), however more substantial materials relating to the Lisu have been produced by scholars outside India (see Bradley this volume).

5.3 Eastern Border Area

The states along the Myanmar border exhibit a linguistic heterogeneity that is comparable to our Northern Area (Map 10.3). In spite of valiant efforts by Garrison (1967) and French (1983), these languages are particularly confusing and exhibit a striking level of diversity even among geographical neighbours. This situation is further complicated by the effects of language contact both within and across genealogical groups, migration (often due to conflict), the fact that speakers are often found on both sides of the political border with Myanmar, and the fact that conflict and political considerations have made research in many parts of this area difficult for many decades now. This means that our sources for many languages are either meagre, old or for other reasons lacking in comprehensiveness and reliability.

These considerations have not dissuaded scholars from proposing a range of genealogical groupings in this area, ranging from the small and well-substantiated (e.g. Bruhn's (2014) Central Naga) to the large and as yet unsubstantiated (e.g. Benedict's (1972: 6) Kuki-Naga(-Karbi-Meithe-Mru), which purports to unite all languages of our Eastern Border Area under a single genealogical grouping). As this chapter was nearing completion, DeLancey (2015) presented morphological evidence for aligning the Northern Naga, Jingpho and Kuki-Chin languages, as well as potentially also Kaman/Meyor, Nung and a number of other 'circum-Irawaddy' language groups, under a major high-level branch 'Central Trans-Himalayan'. DeLancey's argument focuses primarily on cognate elements of the 'agreement words' endemic to these groups, features which DeLancey argues (and we would concur) to be conservative. Comparing DeLancey's morphological approach with our present lexical-comparative approach, we find that the morphological evidence produces alignments which we do not identify on lexical grounds (e.g. Northern Naga and Kuki-Chin), but 'divides' branches which we associate on lexical grounds (e.g. the Boro-Garo and Northern Naga branches of Sal). It is important to note that there is no conflict here. On the contrary, DeLancey's argument implies a large language group in our area whose conservative agreement features have been differentially eroded across its branches most likely through the effects of language contact. These same language contact effects would also have introduced complexities in the lexicons of Eastern Border languages. It follows that while the possibility of a large grouping aligning several branches in our Eastern Border Area is real and must be kept in view, it is not currently possible to determine which of the putatively innovative branches (e.g. Central Naga) might align with it and which might not. The best strategy at the present time is therefore to be patient, wait for more comprehensive and reliable data both from India and from Myanmar, and privilege meticulous lower-level reconstructions over more sweeping but less well-substantiated assessments.

As we begin to discuss the languages of the 'Naga' peoples, it is necessary to make clear the fact that 'Naga' is an ethnic label, not a linguistic one. While we have adopted French's label 'Northern Naga', and will below adopt Bruhn's label 'Central Naga', these labels should not be taken as implying a larger 'Naga' group, for which we do not see convincing evidence at the present time. This is in contrast to much of the earlier literature; for example, Shafer (1955) classified all language of the 'Naga' peoples other than



MAP 10.3 LANGUAGES OF THE EASTERN BORDER AREA

Legend: ● Central Naga ◆ Tangkhulic ▲ Western Naga ■ Kuki-Chin ▣ Northwestern Kuki-Chin ◆ Meithei
■ Angami-Pochuri ► Karbi

Northern Naga with those we will call ‘Kuki-Chin’, placing them all in his ‘Kukish’ section. Ever since, Tibeto-Burmanists have tended to presume that ‘Naga’ and Kuki-Chin languages are joined into some sort of coherent branch of Tibeto-Burman. While this may turn out to be correct, Shafer’s discussion cannot be taken as conclusive.

Despite the profound diversity of ‘Naga’ languages, several attempts to classify them have reached similar, though not identical, conclusions. Leaving aside classifications that simply parrot someone else’s opinion, three older sources deserve to be taken seriously: The *Linguistic Survey of India* (Vol. III, Part III, Grierson 2005 [1904]), Shafer (1955) and Garrison (1967). In what follows, they will be supplemented by our conclusions on the basis of a lexical comparison, as well as by some excellent newer materials.

Geoffrey Garrison did Tibeto-Burmanists a monumental service by assembling a vast amount of data on the ‘Naga’ languages. He offers a classification that is more typological than genealogical,¹⁵ and based less on lexical similarities than on phonological, morphological and syntactic comparisons. Even so, he arrives at groupings that have much in common with those offered earlier by the LSI and Shafer, and much in common with our judgements.

Kuki-Chin is listed here as one group among many. This should not be taken to imply that Kuki-Chin has a genealogical status or internal diversity which is necessarily parallel with that of any other group we will review here. It is only to say that we are currently unable to say anything further about Kuki-Chin’s alignment, beyond simply observing DeLancey’s (2015) argument as we have done.

5.3.1 Central Naga [*Ao* group]

The group of ‘Naga’ languages including Ao was first called ‘Central’ in the LSI and re-dubbed ‘Northern’ by Shafer and then the ‘Ao group’ by Burling (2003). Bruhn (2014) has recently placed this group on a firm comparative-historical footing, reviving the LSI’s linguistically neutral label ‘Central Naga’ in the process. We here join Bruhn in adopting this label, albeit at the risk of implying a genealogical alignment with Northern Naga which we do not intend.

Central Naga languages are spoken to the south and west of the Northern Naga languages, primarily in the Mokokchung, Wokha, Tuensang and Kiphire districts of Nagaland, as well as to an extent in Longleng. The major language within the group is Ao, which has at least two major dialects that border on mutual unintelligibility: Chungli and Mongsen; the latter has been described by Coupe (2007). Yacham and Tengsa in the northerly Dikhu river area are similar to Chungli and Mongsen, but even more similar to each other. Garrison (1967) considers Yacham-Tengsa to be dialects of Ao which have been influenced by the neighbouring Phom language. Most scholars have grouped Lotha [Lhota], Sangtam [T(h)ukumi], and Yimchungrü [Yachumi] with Ao, and this practice has been validated by Bruhn’s (2014) reconstruction. Shafer (1955: 106) also included Lepcha (Rong), which is spoken in far-off Sikkim, in the Central Naga branch, an assignment which seems primarily to have been based on the numerals. Lepcha numerals 1–4 and perhaps also 5 do indeed appear close, especially to Tengsa; however, given the comparatively vast dissimilarities in the rest of their lexicons, it is difficult to know what to make of this. More recently, Coupe (2010) has argued that shared patterns of ‘overcounting’¹⁶ in Central Naga and Angami-Pochuri languages imply that these two branches should be aligned. However, given the evidently unstable nature of this feature, the non-cognacy of the parallel forms, and the geographical contiguity of these two groups, one would wish for further supporting evidence before accepting this alignment.

5.3.2 *Angami-Pochuri*

Most languages in this group cluster within the Zunheboto, Phek and Kohima districts of Nagaland. Mysteriously, the LSI called these languages ‘Western’ while Shafer called them ‘Eastern’. This group is less unified than the Ao group, but it has two clear nuclei. The Angami nucleus is found in and around Kohima district, and includes Angami [Tenyidie/Tenedyie], spoken by over 100,000 people, and the closely related Chokri-Kheza [Chokoso, Chakhesang, Eastern Angami] cluster. Mao [Sopvoma, (E)memei, Poumai], spoken by more than 50,000 people in northern Manipur state, stands just a bit further from Angami, but nevertheless within the same nucleus. The second nucleus includes Pochuri [Southern Sangtam, Eastern Rengma, Meluri, Mills’ Rengma C], spoken in the far southeast corner of Nagaland, together with Morrison’s Ntenyi [Northern Rengma, Mills’ Rengma B]. Finally, Rengma proper [Mills’ Rengma A] is spoken to the north of Kohima, and was at least at one time also spoken in a non-contiguous part of Karbi-Anglong district in Assam; no data from Karbi-Anglong Rengma seem to have been published, however, and we do not know whether it is still spoken. Finally, Sumi [Simi, Sema] is spoken in Zunheboto district by around 100,000 people (Teo 2014). Sumi and Rengma stand somewhat apart, both from the Angami and Pochuri nuclei and from one another.

5.3.3 *Western Naga /Zeliangrong/*

The languages of this group are spoken by at least 100,000 people found on both sides of the western border between Nagaland and Manipur, as well as in adjacent parts of Assam. They are even more closely unified than those of the Ao group, and probably in fact constitute a dialect chain. The unity of this group has been recognized at least since Shafer (1955), who called them the ‘Western’ branch of ‘Naga’. The name ‘Zeliangrong’ is of local coinage, and is constructed from the first syllables of Zeme, Liangmai and Rongmei (the latter a variant of Nruanghmei). We are ambivalent about the latter term due to its association with certain local political movements, and our uncertainty regarding whether (or for how long) this label would be accepted by these languages’ speakers. We therefore revert to Shafer’s label, for lack of any more useful ideas.

Among the Western Naga languages, Zeme [Empeo, Kachcha], Mzieme, and Liangmai [K(w)oireng] are particularly close. Nruanghmei [Rongmei, Kabui], Puiron, Khoirao [Thang(g)al] and Maram are a bit more divergent.

The following sets of Eastern Border Area languages, from Tangkhulic through to Kuki-Chin, are found in a contiguous area stretching from northern Manipur state through to Mizoram (as well as across the border in Myanmar). Many of these languages exhibit lexical, phonological and morphological similarities, which naturally raises the possibility of posit-ing a common grouping. Here too, however, it is necessary to factor in the effects of language contact and mutual influence. Despite undeniable progress in recent years, we are not yet at the point where we can clearly distinguish inherited features from areally diffused features in many cases. We therefore maintain Burling’s (2003) practice of treating Tangkhulic, Karbi, Meithei and Kuki-Chin as independent branches, at the same time acknowledging that some higher-order affiliations among some or all of them may eventually be substantiated.

5.3.4 *Tangkhulic*

Around 150,000–200,000 speakers of Tangkhul lects are found in the easterly Ukhrul district of Manipur and adjacent areas of Myanmar, culturally identifying as Nagas.

Although all Tangkhulic lects remain underdescribed, a number of recent essays by Mortensen and colleagues (Mortensen 2003; Mortensen and Keogh 2011; Mortensen and Miller 2013) have clarified their external and internal relationships considerably. Mortensen (2003) identifies a number of phonological and lexical innovations which suffice to establish a Tangkhulic group containing East-Central, Southern, North-Central, and Northern subgroups. The East-Central branch bifurcates into Standard Tangkhul/Ukrul and Kachai-Phadāng, while the Southern branch contains Brown's (1837) 'Central Tangkhul' and 'North Tangkhul' lects, in addition to Khangoi. The North-Central and Northern branches contain only two languages, Champhung and Huishu respectively.

The potential relationship of Tangkhulic languages to Kuki-Chin languages is a vexing question, particularly when data from Maring and the recently discovered language Sorbung are considered. These languages are close to Tangkhulic; however, it is now clear that at least some similarities are due to contact (Mortensen *inter alia*), and that there may be an equal possibility of aligning these languages with Northwest Kuki-Chin. We expect that clarification of the relationship between Tangkhulic and Kuki-Chin languages will become a productive research area in the coming years, as more data become available.

5.3.5 Karbi [Mikir, Arleng]

Karbi is a major language of around 400,000 speakers found primarily in the Karbi Anglong ('Karbi Hills') district of Assam, as well as in smaller parts of surrounding territories. An excellent recent description of Karbi in English by Konnerth (2014) builds on earlier work in German by Grüssner (1978, n.d.). The common exonym 'Mikir' is considered pejorative, while the name 'Arleng' ('person' in Karbi) is not pejorative, but seems to be disfavoured.

The place of Karbi in Tibeto-Burman has never been clear; although it shows resemblances to a number of Eastern Border Area languages, its differences are just as striking. Karbi has been speculatively aligned with Kuki-Chin and Naga languages in most classifications, but there seems to be an at least equal basis for treating it as an isolate within Tibeto-Burman. It is clear that the primary reason for the confusion is language contact, most conspicuously with Meithei (and perhaps other Tibeto-Burman languages of the area) and with Khasian lects such as Pnar [Jaintia]. Karbi numerals provide a good case in point: Konnerth (2014: 11) shows that the subtractive Karbi and Meithei forms for 'eight' and 'nine' (10 – 2 and 10 – 1) are rare in the area, and must represent an instance of calquing. Similarly, Joseph (2010) provides evidence that Karbi 'one' is a Khasian loan.

5.3.6 Meithei(lon) [Manipuri]

The state of Manipur differs from the other 'hill' states of the northeast in having a large central plateau. This permitted wet rice to be grown, supported a denser population, and allowed a more complex political and social system to develop than was possible until the modern era anywhere else among the indigenous populations of the present hill states. The inhabitants of this plateau speak a Tibeto-Burman language, but they are Hindus, and several centuries ago they developed an Indic-inspired script for their language. They are the only Northeast Indian Tibeto-Burman speaking people with a literate tradition that predates the colonial period. Their language, known as Meitei, Meithei or Meitheilon (or, often, as 'Manipuri') shows some lexical resemblances to Kuki-Chin languages and

some to Tangkhulic. These resemblances are not great enough to make the assignment of Meithei to one of the groups obvious; furthermore, they may be a result of borrowing (as in the case of Karbi). None of the earlier classifications include Meithei within one particular group of ‘Naga’ or Kuki languages, though it has often been taken to be a member of the larger Naga-Kuki grouping. As with Karbi, it seems safest to leave Meithei by itself (as is also proposed in Chelliah 1997).

5.3.7 *Kuki-Chin [Kukish]*

The Kuki-Chin languages of Manipur, Mizoram and adjacent areas in Myanmar and Bangladesh, have been the subject of increased study in recent years; nonetheless, the precise structure and external affiliations of this group remain uncertain.

Two morphological features of Kuki-Chin languages especially stand out: the first is the near-universal feature of verb stem alternations. The Stem 1 form is often associated with main clauses or intransitive predicates and usually has an open syllable. The Stem 2 form is often associated with subordinated clauses or transitive predicates, and often has a closed syllable. For example, Hakha Lai *tsōo* ‘buy.1’ and *tsook* ‘buy.2’ (VanBik 2007: 12). The second archetypically Kuki-Chin feature is the isomorphism of nominal possessive prefixes with verbal subject agreement proclitics (DeLancey 2013b).

It is clear that Kuki-Chin has a core group of languages that are both lexically and grammatically close to one another. VanBik (2007: 23) divides these languages into three branches, which he calls ‘Central’, ‘Peripheral’ and ‘Maraic’. The Central branch includes Mizo [Lushai], which is the majority language of Mizoram, and is spoken by well over half a million people. Its close relatives Hmar and Lai are spoken primarily in northern and southern Mizoram within Northeast India; Lai speakers are also found in Bangladesh and in Myanmar, together with speakers of Laamthuk Thet. The Maraic branch was first proposed by Matisoff, and includes Mara [Lakher], spoken by perhaps 40,000 people in the Mara Autonomous District Council area of southern Mizoram, and Zotung and Senthang, both spoken in Myanmar. Finally, several more conservative languages are spoken both to the north and south of Mizo and Mara. These have been called the ‘Peripheral’ branch, following Peterson’s observation that they share the innovation that Proto-Kuki-Chin *r -> g-. The Northern branch of Peripheral Kuki-Chin, sometimes referred to as the ‘Zo(u)’ group, includes well-known languages such as Thado(u), Zo, Tedim, Purum, Ralte and Paite, and also the Sizangic languages Gangte and Vaiphei. The Southern group languages are spoken in Myanmar; they include Cho, Asho and perhaps also Khumi (which, however, lacks stem alternations).

More problematic are the so-called ‘Old Kuki’ languages in southeastern Manipur, which remain little-understood. Although most seem to have subject agreement proclitics, they apparently lack stem alternations. It may therefore be possible to propose a primary split between ‘Old Kuki’, which we prefer to call ‘Northwestern Kuki-Chin’, and core Kuki-Chin as reconstructed by VanBik (2007). Languages in this group include Aimol, Anal, Chothe, Lamkang, Kom, Moyon, Monsang and Tarao.

6 CONCLUSIONS

In the conclusion to Burling’s (2003) version of this chapter, he discussed the cooling of longstanding tensions in several parts of the northeast, the fact that foreigners were once again being welcomed as visitors and researchers, the simultaneous increase in local scholarship, and his hopes and expectations that cooperation between international and

local researchers would continue, and research productivity increase, throughout the region. It is heartening to be able to report that, while there surely remains a great deal to be done, Burling's expectations are beginning to be met. The North East Indian Linguistics Society has conducted a number of conferences in the region, bringing in unprecedented numbers of researchers from several Indian and international institutions and fostering collaborations with local institutions such as Manipur University, Rajiv Gandhi University and especially Gauhati University. A larger number of field projects are currently underway in Northeast India than perhaps at any other time in history, and more (and more reliable) data are now available for more Northeast Indian languages than ever before. Nevertheless, we must again reiterate that the lion's share of the descriptive work which is truly needed in our area has yet to be undertaken. As research access on the Myanmar side gradually becomes a reality, and now that all states other than Arunachal Pradesh have finally lifted their Protected Area Permit regimes, it seems certain that a new age of research on Northeast Indian Tibeto-Burman languages is taking shape. We look forward to the great number of advances that will surely be made in the decade ahead.

NOTES

- 1 The goal of this chapter is to update Burling's (2003) chapter by taking account of recent scholarship, while maintaining as much of its original structure and spirit as was possible. Burling (2003) acknowledged a fellowship from the Fulbright Foundation, which funded seven months' research in Northeast India during 1996–7, and further acknowledged discussions with N. K. Achumi, David Bradley, Thangi Chhangte, Bibhash Dhar, Frederick S. Downs, George van Driem, François Jacquesson, U. V. Joseph, Boyd Michailovsky, Dipankar Moral, Rajesh Sachdeva, L. Mahabir Singh, R. W. Sprigg, Jackson T.-S. Sun and Graham Thurgood (also absolving them from any errors). Post additionally acknowledges the following individuals, who were consulted (in some cases extensively!) in the preparation of this revision: Roger Blench, Timotheus Bodt, Daniel Bruhn, Huziwaru Keisuke, Gwendolyn Hyslop, Linda Konnerth, Ismael Lieberherr, Yankee Modi, Zilpha Modi, Stephen Morey, Amos Teo, Philip Thangliemang and Scott DeLancey. We are aware that not all of their advice has been followed, no doubt to this chapter's ultimate cost, but are nonetheless grateful for their contributions. Finally, a note regarding the terms 'Tibeto-Burman' and 'Sino-Tibetan': acknowledging the controversial nature of these terms and the potentially better fit of the alternative label for this phylum 'Trans-Himalayan', we nonetheless retain 'Tibeto-Burman' in this chapter to avoid a potentially confusing discontinuity with its earlier version, as well as in this volume overall.
- 2 Sikkim and Jalpaiguri Division of West Bengal have 'officially' fallen within Northeast India since the 1990s, however they fall somewhat outside of this region from an ethnolinguistic standpoint, and will not be canvassed in the present chapter.
- 3 Munda languages are also reportedly spoken among tea garden labourers in Assam, although we are not aware of any published descriptions.
- 4 Unfortunately, the *Ethnologue* has not always kept pace with these developments, and retains offensive designations for several groups in their 'three-letter language codes'; for example, *clk* for Idu, from *Chulikata*, *dap* for Nyishi, from *Dafla* and *mrg* from *Miri* for Mising (Lewis *et al.* 2015; Morey *et al.* 2013).
- 5 It is perhaps worth reiterating and emphasizing here that Konow's (1902) seminal identification of the 'North Assam group' was never actually intended to count as a statement of genealogical alignment (in Konow's words, 'the North Assam group is not a well-defined philological group' (Grierson 2005 [1909]: 572)). It is unfortunate, then, that both this non-subgroup and similar constructions such as 'Kamarupan' are nevertheless frequently cited in the linguistics and non-linguistics literature alike as

- though they were (for example, Bielenberg and Nienü 2001; Kang *et al.* 2010, etc.). Evidently, such ‘heuristic’ labels, while in principle methodologically harmless, have nonetheless had a misleading effect.
- 6 Speakers of these languages in Arunachal Pradesh are generally referred to using the ethnonym ‘Monpa’. This label is linguistically imprecise, as it is also used to refer to speakers of several non-Bodic languages of the Kamengic group.
 - 7 Boro is variously spelled *Bodo* and *Boro*; these labels are equivalent. We regularize this spelling to Boro, following Burling (2012).
 - 8 The label ‘Sal’ was developed by Burling (1983) to refer a group of languages in which the word for ‘sun’ is a form of **sal* (*san*, *saing*, *han*, *zhang*, etc.). *Ethnologue* (Lewis *et al.* 2015) for no clear reason that we can discern expands this label to include Kuki-Chin and ‘Naga’ languages as well. We do not support this expansion, and retain the conservative sense of ‘Sal’ here. We acknowledge van Driem’s (2014) label ‘Brahmaputran’, but feel it may be inappropriately restrictive – most languages in this group are nowhere near the Brahmaputra River, and those which are near to it have their own traditional names for this river.
 - 9 ‘Boro-Garo’ may have been encouraged by confusion over Atong and Ruga, whose speakers are ethnically ‘Garo’, although they are linguistically quite different. Benedict formalized the confusion by calling Atong and Ruga ‘Garo A’ while he called Garo proper ‘Garo B’ (1972: 6 fn. 20). Rabha, which is similar to Atong and Ruga, was thus duly included in ‘Garo A’, despite the fact that Rabha are not ethnically ‘Garo’. Rabha, Atong, Ruga and Koch (including the languages or dialects known as Tintinkiya Koch, Wa’nang Koch and Pani Koch, spoken just west of the Garo Hills), should never have been called ‘Garo’. Nor should the utterly mislabelled ‘Garo of Jalpaiguri’ (Grierson 2005 [1903]), who call themselves ‘Koch’ or ‘Rabha’. Locally, this group of languages is known as ‘Koch’, and that is the best term for linguists as well.
 - 10 In geographical terms, Northern Naga languages are in fact mostly spoken to the *northeast* of other ‘Naga’ groups, hence the northern/eastern confusion. Perhaps ‘Northeastern Naga’ would have been a preferable label, and ‘(North)eastern Sal’ might be even better. Nonetheless, we will here maintain consistency with French’s (1983) term, since it seems to be the most widely used.
 - 11 For example, Morey (this volume) identifies around 70 varieties of Tangsa, many of which are mutually intelligible with one another, others of which are not. On the other hand, Tutsa seems to count as only a single language.
 - 12 The label ‘Luish’ is said to derive from *loi*, an exonym which may have a pejorative connotation. We follow Matisoff (2015) in abandoning this label in favour of ‘Asakian’, which is similar to Löffler’s (1964) label ‘Sakisch’.
 - 13 Of these languages, only Andro and Sengmai were once spoken within the modern-day borders of Northeast India (McCulloch 1859). Andro and Sengmai villages still exist, and villagers prize the tradition that they once had their own languages. However, neither language has any living speakers today.
 - 14 Note, however, that DeLancey’s thesis would *not* seem to explain the hypothetical loss of agreement features in Asakian languages, nor for that matter in many other hill languages of the region (e.g. Central Naga).
 - 15 For example, Morrison considers the pre-head or post-head position of negation to be a primary grouping criterion, which it cannot be. If this practice were generalized, then we would find Tani, Boro-Garo, Tibetic and a subset of Morrison’s ‘Naga’ languages to form a single branch based on the criterial innovation of post-head negation. If there is anything to justify discussing these developments in the same context, it may be similarities in areal-typological changes found in different areas of the Tibeto-Burman world (DeLancey 2013a). But this has nothing to do with subgrouping on a genealogical basis.
 - 16 In which lower numerals are expressed in terms of numbers of units ‘on the way to’ a higher numeral; for example, ‘the sixth unit towards twenty’ for ‘16’.

REFERENCES

- Abraham, Binny, Kara Sako, Elina Kinny and Isapdaile Zeliang (2005). *A Sociolinguistic Research among Selected Groups in Western Arunachal Pradesh Highlighting MONPA*. Dallas, TX, Summer Institute of Linguistics.
- Abraham, P. T. (1985). *Apatani Grammar*. Mysore, Central Institute of Indian Languages.
- American Baptist Foreign Mission Society (1905). *English–Garo Dictionary*. Shillong, IN, Assam Secretariat Printing Office.
- Anderson, Gregory D. S. and Ganesh Murmu (2010). ‘Preliminary notes on Koro, a “hidden” language of Arunachal Pradesh.’ *Journal of Indian Linguistics* 71: 1–32.
- Apatani Language Development Committee (2015). ‘Tanw Aguñ: An Audio Dictionary of the Apatani Language.’ Available at: language.apatani.org, accessed 15 August 2016.
- Arokianathan, S. (1987). *Tangkhul Naga Grammar*. Mysore, Central Institute for Indian Languages.
- Badu, Tapoli (1994). *Pailibo Language Guide*. Itanagar, IN, Government of Arunachal Pradesh Directorate of Research.
- (2004). *Tangam Language Guide*. Itanagar, IN, Government of Arunachal Pradesh Directorate of Research.
- Baruah, Kanak Lal (1960 [1933]). *Early History of Kāmarupa, from the Earliest Times to the End of the Sixteenth Century*. Guwahati, IN, Lawyer’s Book Stall.
- Benedict, Paul K. (1972). *Sino-Tibetan: A Conspectus*. [Princeton-Cambridge Series in Chinese Linguistics #2.] Cambridge, UK, Cambridge University Press.
- (1976). ‘Sino-Tibetan: Another look.’ *Journal of the American Oriental Society* 96(2): 167–197.
- Bielenberg, Brian and Zhalie Nieno (2001). ‘Chokri (Phek dialect): Phonetics and phonology.’ *LTBA* 24(2): 85–122.
- Blankenship, Barbara, Peter Ladefoged, Peri Bhaskararo and Nichumeno Chase (1993). ‘Phonetic structures of Khonoma Angami.’ *LTBA* 16(2): 69–88.
- Blench, Roger (2013). *The Miji Language: Distribution, Dialects, Wordlist and Classification*. Cambridge, UK, Kay Williamson Educational Foundation.
- Blench, Roger and Mark W. Post (2012). *Declassifying Eastern Himalayan Languages: Reconsidering the Evidence*. Cambridge, UK and Armidale, AU, Kay Williamson Educational Foundation and University of New England.
- Bodt, Timotheus Adrianus (2014). ‘Ethnolinguistic survey of westernmost Arunachal Pradesh: A fieldworker’s impressions.’ *LTBA* 37(2): 198–239.
- and Ismael Lieberherr (2015). ‘First notes on the phonology and classification of the Bangru language of India.’ *LTBA* 38(1): 66–123.
- Boro, A. (1978). *Miju Dictionary*. Shillong, IN, Government of Arunachal Pradesh Research Department.
- Brown, Nathan (1837). ‘Comparison of Indo-Chinese languages.’ *Journal of the Asiatic Society of Bengal* 6: 1023–1039.
- Bruhn, Daniel Wayne (2014). *A Phonological Reconstruction of Proto-Central Naga*. PhD Dissertation. Berkeley, CA, University of California.
- Burling, Robbins (1959). ‘Proto Bodo.’ *Language* 35(3): 433–453.
- (1961). *A Garo Grammar*. [Deccan College Monograph Series.] Poona, India, Deccan College Postgraduate and Research Institute.
- (1971). ‘The historical place of Jinghpaw in Tibeto-Burman.’ In F. K. Lehman, Ed. *Papers on Tibeto-Burman Historical and Comparative Linguistics from the Second Annual Meeting on Sino-Tibetan Reconstruction, 1969*. Urbana, University of Illinois: 1–54.

- (1983). ‘The Sal languages.’ *LTBA* 7(2): 1–32.
- (2003). ‘The Tibeto-Burman languages of Northeastern India.’ In Graham Thurgood and Randy J. LaPolla, Eds, *The Sino-Tibetan Languages*. London, Routledge: 167–192.
- (2004). *The Language of the Modhipur Mandi (Garo). Volume I: Grammar*. New Delhi, Bibliophile South Asia.
- (2012). ‘The Stammbaum of Boro-Garo.’ In Gwendolyn Hyslop, Stephen Morey and Mark W. Post, Eds, *North East Indian Linguistics Volume 4*. New Delhi, Cambridge University Press: 21–35.
- Button, Christopher T. J. (2011). *Proto Northern Chin*. [STEDT Monograph 10.] Berkeley, CA, Sino-Tibetan Etymological Dictionary and Thesaurus Project.
- Census of India (2011). *A-11: Individual Scheduled Tribe Primary Census Abstract Data and its Appendix: Arunachal Pradesh and Assam*. New Delhi, Government of India. Available at: www.censusindia.gov.in/2011census/PCA/ST.html, accessed 4 April 2016.
- Chelliah, Shobhana L. (1997). *Meithei*. [Mouton Grammar Library 17.] Berlin, Mouton de Gruyter.
- Chhangte, Thangi (1993). *Mizo Syntax*. PhD Dissertation. Eugene, OR, University of Oregon.
- Choya, Tar (2009). *English–Nyishi Dictionary*. Shillong, IN, Living Word Communicators.
- Coupe, Alexander R. (2007). *A Grammar of Mongsen Ao*. Berlin, Mouton de Gruyter.
- (2010). ‘The languages of the Naga and their position in Tibeto-Burman: Historical overview and current perceptions.’ *Proceedings of International Symposium on Sino-Tibetan Comparative Studies in the 21st Century*. Taipei, Academia Sinica: 185–203.
- Das Gupta, Kamlesh (1963). *An Introduction to the Gallong Language*. Shillong, IN, North East Frontier Agency Research Department.
- (1977). *A Phrase Book in Miju*. Shillong, IN, Government of Arunachal Pradesh Research Department.
- (1983). *An Outline on Tagin Language*. Shillong, IN, Government of Arunachal Pradesh Directorate of Research.
- DeLancey, Scott (2010a). ‘Language replacement and the spread of Tibeto-Burman.’ *Journal of the South East Asian Linguistics Society* 3(1): 40–55.
- (2010b). ‘Towards a history of verb agreement in Tibeto-Burman.’ *Himalayan Linguistics* 9(1): 1–39.
- (2012). ‘On the origins of Bodo-Garo.’ In Gwendolyn Hyslop, Stephen Morey and Mark W. Post, Eds, *North East Indian Linguistics Volume 4*. New Delhi, Cambridge University Press India: 3–20.
- (2013a). ‘Creolization in the divergence of Tibeto-Burman.’ In Nathan Hill and Tom Owen-Smith, Eds, *Trans-Himalayan Linguistics: Historical and Descriptive Linguistics of the Himalayan Area*. Berlin, Mouton de Gruyter: 41–70.
- (2013b). ‘The history of postverbal agreement in Kuki-Chin.’ *Journal of the Southeast Asian Linguistics Society* 6: 1–17.
- (2014). The origins of postverbal negation in Kuki-Chin. Paper presented at the 8th International Conference of the North East Indian Linguistics Society. Guwahati, IN, Gauhati University.
- (2015). ‘Morphological evidence for a Central branch of Trans-Himalayan [Sino-Tibetan].’ *Cahiers de Linguistique – Asie Orientale* 44: 122–149.
- Diffloth, Gerard (2005). ‘The contribution of paleoanthropology to the homeland of Austro-Asiatic.’ In Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas, Eds, *The*

- Peopling of East Asia: Putting Together Archaeology, Linguistics, and Genetics.* London, Routledge-Curzon: 77–80.
- Dondrup, Richin (1988). *A Handbook on Sherdukpen Language*. Itanagar, IN, Government of Arunachal Pradesh Directorate of Research.
- (1990). *Bugun Language Guide*. Itanagar, IN, Government of Arunachal Pradesh Directorate of Research.
- Emeneau, Murray (1956). ‘India as a linguistic area.’ *Language* 32(1): 3–16.
- Endle, Sidney (1884). *Outline Grammar of the Kachári (Bárâ) Language as Spoken in District Darrang, Assam, with Illustrative Sentences, Notes, Reading Lessons and a Short Vocabulary*. Shillong, IN, Assam Secretariat Press.
- Evans, Jonathan P. (2009). ‘Is there a Himalayan tone typology?’ *Senri Ethnological Studies* 75: 199–221.
- French, Walter (1983). *Northern Naga: A Tibeto-Burman Mesolanguage*. PhD Thesis. New York, City University.
- Goswami, S. N., Ed. (1994). *Nishing (Bangni) Language Guide*. Itanagar, IN, Government of Arunachal Pradesh Directorate of Research.
- Grewal, Dalvinder Singh (1997). *Tribes of Arunachal Pradesh: Identity, Culture and Languages* [2 Vols]. Delhi, South Asia Books.
- Grierson, George A., Ed. (2005 [1903]). *Linguistic Survey of India, Volume 3: Tibeto-Burman Family, Part 2: Specimens of the Bodo, Nāgā and Kachin Groups*. New Delhi, Low Price Publications.
- , Ed. (2005 [1904]). *Linguistic Survey of India, Volume 3: Tibeto-Burman Family, Part 3: Specimens of the Kuki-Chin and Burma Groups*. New Delhi, Low Price Publications.
- , Ed. (2005 [1909]). *Linguistic Survey of India, Volume 3: Tibeto-Burman Family, Part 1: General Introduction, Specimens of the Tibetan Dialects, Himalayan Dialects, and North Assam Groups*. New Delhi, Low Price Publications.
- Grüssner, Karl-Heinz (1978). *Arleng Alam, die Sprache der Mikir: Grammatik und Texte*. [Beiträge zur Südasiensforschung 39.] Wiesbaden, Franz Steiner.
- (n.d.). *Karbi Dictionary*. Tübingen.
- Gurdon, Philip Richard Thornhagh (1904). ‘The Morans.’ *Journal of the Asiatic Society of Bengal* 73: 36–48.
- Guyot-Récharde, Bérénice (2012). *State-building’s Dress Rehearsal: Re-interpreting the Sino-Indian War of 1962*. Trinity College, Cambridge University.
- Huziwara, Keisuke (2008). *Cakku-go no kizyuu gengogakuteki kenkyuu [A Descriptive Linguistic Study of the Cak Language]*. PhD Dissertation. Kyoto, Kyoto University.
- Hyslop, Gwendolyn, Stephen Morey and Mark W. Post, Eds (2011). *North East Indian Linguistics Volume 3*. New Delhi, Cambridge University Press India.
- (2012). *North East Indian Linguistics Volume 4*. New Delhi, Cambridge University Press India.
- (2013). *North East Indian Linguistics Volume 5*. New Delhi, Cambridge University Press India.
- Hyslop, Gwendolyn, Linda Konnerth and Stephen Morey, Eds (2014). *North East Indian Linguistics Volume 6. [Languages and Peoples of the Eastern Himalayan Region 1.]* Canberra, Asia-Pacific Linguistics.
- Jacquesson, François (2001). ‘Person-marking in TB languages of North-Eastern India.’ *Linguistics of the Tibeto-Burman Area* 24(1): 113–144.
- (2004). *Le deuri: Langue tibéto-birman d’Assam*. [Collection Linguistique publiée par la Société Linguistique de Paris.] Leuven, Peeters.

- Joseph, U. V. (2000). *Rabha–English Dictionary Khúrangnala*. Guwahati, Assam, IN, Don Bosco Publications.
- (2007). *Rabha*. [Brill's Tibetan Studies Library, Languages of the Greater Himalayan Region.] Leiden, Brill.
- (2010). ‘The numeral “one” in Khasi and Karbi.’ In Stephen Morey and Mark W. Post, Eds, *North East Indian Linguistics Volume 2*. New Delhi, Cambridge University Press India: 149–161.
- and Robbins Burling (2006). *The Comparative Phonology of the Bodo-Garo Languages*. Mysore, Central Institute of Indian Languages.
- Kakati, Banikata (1995). *Assamese, its Formation and Development: A Scientific Treatise on the History and Philology of the Assamese Language*. Gauhati, IN, Lawyer’s Book Stall.
- Kang, Longli, Shilin Li, Sameer Gupta, Yingang Zhang, Kai Liu, Jianmin Zhao, Li Jin and Hui Li (2010). ‘Genetic structures of the Tibetans and the Deng people in the Himalayas viewed from autosomal STRs.’ *Journal of Human Genetics* 55: 270–277.
- Karlo, Tado, Bomcak Riba, Ili Riba, Igo Riba and Mark W. Post (2007a). *Galo Agom: Part 1*. Itanagar, IN, Galo Welfare Society.
- (2007b). *Galo Agom: Part 2*. Itanagar, IN, Galo Welfare Society.
- Karlsson, Bengt G. (1997). *Contested Belonging: An Indigenous People’s Struggle for Forest and Identity in Sub-Himalayan Bengal*. Lund, Lund Monographs in Social Anthropology.
- Konnerth, Linda (2012). ‘The nominalizing velar prefix *gV- in Tibeto-Burman languages of Northeast India.’ In Gwendolyn Hyslop, Stephen Morey and Mark W. Post, Eds, *North East Indian Linguistics Volume 4*. New Delhi, Cambridge University Press India: 58–82.
- (2014). *A Grammar of Karbi*. PhD Thesis. Eugene, OR, University of Oregon.
- and Amos Teo (2014). ‘Acoustic-statistical and perceptual investigations of Karbi tones: A peculiar case of incomplete neutralization of F0.’ In Gwendolyn Hyslop, Linda Konnerth and Stephen Morey, Eds, *North East Indian Linguistics Volume 6*. Canberra, Asia-Pacific Linguistics: 13–38.
- Konow, Sten (1902). ‘Note on the languages spoken between the Assam Valley and Tibet.’ *Journal of the Royal Asiatic Society of Great Britain and Ireland* 34(1): 127–137.
- Krisiko, Sutik (2006). *Tawrā: Tako zo bra, käman Püt sai brat*. Tezu, IN, Graphic Systems.
- Kumar, B. B., T. P. Meya and S. Jon (n.d.). *Hindi Khiamngan English Vocabulary*. Kohima, Nagaland Bhasha Parishad.
- Landi, Victor (2005). *The Meyors and Their Language*. Itanagar, Arunachal Pradesh, India, Arunachal Pradesh Directorate of Research.
- LaPolla, Randy J. (1992). ‘On the dating and nature of verb agreement in Tibeto-Burman.’ *Bulletin of the School of Oriental and African Studies* 55(2): 298–315.
- Lewis, M. Paul, Gary F. Simons and Charles D. Fennig, Eds (2015). *Ethnologue: Languages of the World*, 17th Edition. Dallas, Texas, SIL International. Available at: www.ethnologue.com, accessed 4 November 2015.
- Lǐ, Dàqín [李大勤] (2002). 《格曼语研究》 [Geman Yu Yanjiu/A Study of Geman Language]. Beijing, 民族出版社 [Minzu Chubanshe/Nationalities Press].
- (2004). 《苏龙语研究》 [Sulong Yu Yanjiu/A Study of Sulong Language]. Beijing, 民族出版社 [Minzu Chubanshe/Nationalities Press].

- and Dí Jiāng (2001). ‘Zhá huà gǎikuàng.’ *Minzu Yuwen* 2001(6): 61–75.
- Lieberherr, Ismael (2015). ‘A progress report on the historical phonology and affiliation of Puroik.’ In Linda Konnerth, Stephen Morey, Priyankoo Sarmah and Amos Teo, Eds, *North East Indian Linguistics 7*. Canberra, Asia-Pacific Linguistics: 235–286.
- Löffler, Lorenz G. (1964). ‘Chakma und Sak: Ethnolinguistische Beiträge zur Geschichte eines Kulturvolkes.’ *Internationales Archiv für Ethnographie* 50(1): 72–115.
- Lorrain, J. Herbert (1995 [1907/1910]). *A Dictionary of the Abor-Miri Language, with Illustrative Sentences and Notes*. New Delhi, Mittal.
- and Fred W. Savidge (1898). *A Grammar and Dictionary of the Lushai Language (Dulien Dialect)*. Shillong, Assam Secretariat Printing Office.
- McCulloch, William (1859). *Account of the Valley of Munnipore and of the Hill Tribes, with a Comparative Vocabulary of the Munnipore and Other Languages*. Calcutta, Bengal Printing Company.
- Maitra, Asim (1988). *A Guide Book to Lisu Language*. Delhi, Mittal.
- (1993). *Profile of a Little-known Tribe: An Ethnographic Study of Lisus of Arunachal Pradesh*. New Delhi, Mittal.
- Marrison, Geoffrey E. (1967). *The Classification of the Naga Languages of North-East India*, 2 Vols. PhD Dissertation. London, University of London SOAS.
- Matisoff, James A. (2015). ‘Re-examining the genetic position of Jingpho: Putting flesh on the bones of the Jingpho/Luish relationship.’ In N. J. Enfield and Bernard Comrie, Eds, *Languages of Mainland Southeast Asia: The State of the Art*. [Pacific Linguistics 649.] Berlin, de Gruyter: 111–151.
- Megu, Arak (1990). *Bokar Language Guide*. Itanagar, IN, Government of Arunachal Pradesh Directorate of Research.
- (2003). *Aassing Language Guide*. Itanagar, IN, Government of Arunachal Pradesh Directorate of Research.
- Modi, Yankee (2006). *The Complexity and Emergence of Hindi as Lingua Franca in Arunachal Pradesh*. Vital Voices: Endangered Language and Multilingualism CIIL, Mysore, India, Foundation for Endangered Languages.
- (2013). The nearest relatives of the Tani group. Paper presented at the 19th Himalayan Languages Symposium. Canberra, Australian National University.
- Morey, Stephen (2005). *The Tai Languages of Assam: A Grammar and Texts*. Canberra, Pacific Linguistics.
- (2010). *Turung: A Variety of Singpho Language Spoken in Assam*. Canberra, Pacific Linguistics.
- (2013a). Grammatical change in North East India: The case of Tangsa. Paper presented at the 23rd Conference of the Southeast Asian Linguistics Society. Bangkok, Chulalongkorn University.
- (2013b). ‘The marking of noun phrases: Some observations on the languages of North East India.’ In Gwendolyn Hyslop, Stephen Morey and Mark W. Post, Eds, *North East Indian Linguistics Volume 5*. New Delhi, Cambridge University Press: 171–191.
- (2014). ‘Studying tones in North East India: Tai, Singpho and Tangsa.’ *Language Documentation and Conservation* 8: 637–671.
- and Mark W. Post, Eds (2008). *North East Indian Linguistics*. New Delhi, Cambridge University Press India.
- and Mark W. Post, Eds (2010). *North East Indian Linguistics Volume 2*. New Delhi, Cambridge University Press India.

- Mark W. Post and Victor A. Friedman (2013). The language codes of ISO 639: A premature, ultimately unobtainable, and possibly damaging standardization. Paper presented at the conference on Research, Records and Responsibility (RRR): Ten Years of the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). Melbourne, University of Melbourne, 2–3 December.
- Mortensen, David (2003). *Comparative Tangkhul*. Berkeley, CA, University of California at Berkeley.
- and Jennifer Keogh (2011). ‘Sorbung, an undocumented language of Manipur: Its phonology and place in Tibeto-Burman.’ *Journal of the Southeast Asian Linguistics Society* 4(1): 64–114.
- and James A. Miller (2013). ‘A reconstruction of Proto-Tangkhulic rhymes.’ *LTBA* 36(1): 1–32.
- Nichols, Johanna (1996). ‘The comparative method as heuristic.’ In Mark Durie and Malcolm Ross, Eds, *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*. Oxford, Oxford University Press: 39–71.
- Noram, Tameng (2008). *Noram’s Gompi Makeng* (Dictionary). Pasighat, IN, privately published.
- Ou-Yang, Jue-Ya (1985). 《珞巴族语言简志珞巴族言简志（崩尼-博嘎尔语）》 [Luobazu Yuyan Jianzhi (Bengni Boga’eryu)/Brief Description of the Luoba Nationality Language (The Bengni-Bokar Language)]. Beijing, 民族出版社 [Minzu Chubanshe/Nationalities Press].
- Philology Section of North East Frontier Agency Research Department (1963). *A Dictionary of the Taraon Language for the Use of Officers in NEFA*. Shillong, North East Frontier Agency Research Department.
- Post, Mark W. (2007). *A Grammar of Galo*. PhD Thesis. Melbourne, La Trobe University.
- (2015). ‘Tones in Tani languages: A fieldworker’s guide.’ In Mark W. Post, Stephen Morey and Scott DeLancey, Eds, *Language and Culture in Northeast India and Beyond: In Honor of Robbins Burling*. Canberra, Asia Pacific Linguistics: 182–210.
- and Roger Blench (2011). Siangic: A new language phylum in North East India. Paper presented at the 6th International Conference of the North East Indian Linguistics Society. Tezpur, IN, Tezpur University.
- and Yankee Modi (2011). ‘Language contact and the genetic position of Milang (Eastern Himalaya).’ *Anthropological Linguistics* 53(3): 215–258.
- and Kanno Tage (2013). ‘Apatani phonology and lexicon, with a special focus on tone.’ *Himalayan Linguistics* 12(1): 17–75.
- Pulu, Jatan (1978). *Idu Phrase-Book*. Shillong, IN, Government of Arunachal Pradesh Directorate of Information and Public Relations.
- (1991). *A Phrase Book on Taraon Language*. Itanagar, IN, Government of Arunachal Pradesh Directorate of Research.
- Pulu, Jimi (2002). *A Handbook on Idu Mishmi Language*. Itanagar, IN, Government of Arunachal Pradesh Directorate of Research.
- Ramya, Tame (2011). *An Ethnographic Study of Bangrus of Kurung Kumey District of Arunachal Pradesh*. MPhil Dissertation. Itanagar, IN, Rajiv Gandhi University.
- (2012). ‘Bangrus of Arunachal Pradesh: An ethnographic profile.’ *International Journal of Social Science Tomorrow* 1(3): 1–12.
- Remsangpuia (2008). *Puroik Phonology*. Shillong, IN, Don Bosco.
- Robinson, William (1849). ‘Notes on the languages spoken by the various tribes inhabiting the valley of Assam and its mountain confines.’ *Journal of the Royal Asiatic Society of Bengal* 18(1): 311–318, 342–349.

- Sangdong, David (2012). *A Grammar of the Kadu (Asak) Language*. PhD Thesis. Melbourne, La Trobe University.
- Sastry, G. Devi Prasada (1984). *Mishmi Grammar*. [CIIL Grammar Series.] Mysore, Central Institute of Indian Languages.
- (1991). *Mishmi–English–Hindi Dictionary*. Mysore, Central Institute for Indian Languages.
- Shafer, Robert (1947). ‘Hruso.’ *Bulletin of the School of Oriental and African Studies* 12(1): 184–196.
- (1955). ‘Classification of the Sino-Tibetan languages.’ *Word* 11: 94–111.
- Simon, Ivan Martin (1970). *Aka Language Guide*. Shillong, IN, North East Frontier Agency Directorate of Information and Public Relations.
- (1976). *Hill Miri Language Guide*. Shillong, IN, Government of Arunachal Pradesh Directorate of Research.
- (1979). *Miji Language Guide*. Shillong, IN, Government of Arunachal Pradesh Directorate of Research.
- Sun, Hong-Kai, Ed. (1991). 《藏缅语语音和词汇》 [Zàng-Miǎn-yǔ yǔyīn hé cíhuì/ *Tibeto-Burman Phonology and Lexicon*]. Beijing, 中国社会科学出版社 [Zhongguo Shehui Kexue Chubanshe/Chinese Social Sciences Press].
- Sun, Hong-Kai, Ji-Chuan Zhang, Shai-Zun Liu and Jue-Ya Ou-Yang (1980). 《门巴, 罗巴, 登人的语言》 [*Menba, Luoba, Dengren de Yuyan/The Language of the Menba, Luoba and Dengren*]. Beijing, 中国社会科学出版社 [Zhongguo Shehui Kexue Chubanshe/ Chinese Social Sciences Press].
- Sun, Tianshin Jackson (1992). ‘Review of Zangmianyu Yuyin He Cihui [*Tibeto-Burman Phonology and Lexicon*.]’ *LTBA* 15(2): 73–113.
- (1993). *A Historical-Comparative Study of the Tani (Mirish) Branch of Tibeto-Burman*. PhD Dissertation. Berkeley, CA, University of California.
- Talukdar, G. C. (1962). *A Phrase Book in Idu*. Shillong, IN, North East Frontier Agency Research Department.
- Tawsik, Somodang (2014). *Kaman-Tawrah*. Tinsukia, IN, Chitra Printers.
- Tayeng, Aduk (1990). *Sulung Language Guide*. Itanagar, IN, Government of Arunachal Pradesh Directorate of Research.
- Tayu, Toku Stephen (2010). *Nyishi Grammatical Sketch and Dictionary*. Itanagar, IN, All Purpose Editions.
- Teo, Amos (2014). *A Phonological and Phonetic Description of Sumi, a Tibeto-Burman Language of Nagaland*. Canberra, Asia-Pacific Linguistics.
- Thomason, Sarah Grey and Terrence Kaufman (1988). *Language Contact, Creolization and Genetic Linguistics*. Berkeley, CA, University of California Press.
- VanBik, Kenneth (2007). *Proto-Kuki-Chin: A Reconstructed Ancestor of the Kuki-Chin Languages*. [STEDT Monograph 8.] Berkeley, CA, Sino-Tibetan Etymological Dictionary and Thesaurus Project.
- van Breugel, Seino (2014). *A Grammar of Atong*. Leiden, Brill.
- van Driem, George (2001). *Languages of the Himalayas: An Ethnolinguistic Handbook of the Greater Himalayan Region, Containing an Introduction to the Symbiotic Theory of Language*. Leiden, Brill.
- (2014). ‘Trans-Himalayan.’ In Thomas Owen-Smith and Nathan W. Hill, Eds, *Trans-Himalayan Linguistics: Historical and Descriptive Linguistics of the Himalayan Area*. Berlin, de Gruyter: 11–40.
- Weidert, Alfons (1987). *Tibeto-Burman Tonology*. Amsterdam, John Benjamins.

CHAPTER ELEVEN

GARO¹

Robbins Burling

1 INTRODUCTION

The language that is known to everyone except its own speakers as ‘Garo’ is spoken by about 700,000 people in Northeast India and in Bangladesh. Most of these Garos live in a hilly district in the western part of the Indian state of Meghalaya, but about 100,000 live across the border in Bangladesh, most of them just south of the Garo Hills. Smaller settlements are found in several locations in Assam, in the Khasi Hills of Meghalaya, in Tripura state, and near Modhupur in Bangladesh. Most of these people prefer to call themselves ‘A’chik’ or ‘Mande’ but neither of these terms has gained general acceptance, and until one does, I have little choice but to call them ‘Garo.’ Two small enclaves, one with people known as ‘A’tong,’ the other with ‘Ruga,’ are found within the Garo Hills. Each of these groups has its own language but they consider themselves to be Garos and are accepted as such by all other Garos. (See Chapter 10 on Northeast Indian languages, this volume). The Garos are generally regarded as a “hill tribe” and most of them have lived in hilly, even mountainous, country. Around the edges of their territory, however, particularly in Bangladesh, many live in very flat low lands, where, like their Bangladeshi and Assamese neighbors, they practice wet rice agriculture. Except for A’tong and Ruga all dialects spoken by Garos are mutually intelligible, although speakers who are unfamiliar with a dialect that is spoken far from their own home may need some patience and an occasional explanation in order to understand.

A written form of Garo was developed by American Baptist missionaries during the last decades of the nineteenth century. The missionaries based the orthography upon the dialect of the northeastern corner of the Garo Hills. This was the area with the first substantial number of educated and literate Garos, and their dialect has influenced the speech of educated Garos everywhere. The northeastern dialect on which the written language is based is sometimes called ‘A’we.’ The dialect that covers the western part of the Garo Hills and that is spoken in Bangladesh is known as ‘A’beng’ or ‘Am’beng.’ Matchi, Chisak, and Dual are found in smaller areas in the central and southern part of the district, but all of these dialects grade into one another without sharp breaks. Garos compare their dialects with curiosity and amusement, but they do not correspond to important social divisions within the larger Garo community. The examples in this chapter are from the dialect that has become the *de facto* standard, originally that of the northeastern corner of the Garo hills.

Brief wordlists were collected about 1800 by British officials (Eliot 1794; Hamilton 1940 [1820]) but more extensive descriptions of the language had to wait almost a century before American Baptist Missionaries produced the first grammars and dictionaries. The romanization introduced by the missionaries is now well established, and the language has been used as a medium of elementary education in the Garo Hills for many decades. A few collections of Garo stories have been published, and a few thin weekly newspapers appear, but apart from school books, the most important publications are religious, and

since most Garos are now Christians these include the Bible, which has long been available in Garo translation, hymn books, and various other Christian texts. The written language is used for private correspondence and a few signs are posted in the language, and a fluent Garo reader would certainly require no more than a few months to read the entire corpus of Garo printed literature.

By the standards of Northeast India, the Garo community is well served by Garo to English dictionaries. The most widely available is the modestly titled but reliable *The School Dictionary, Garo to English* (Nengminza 1946 and later). It can be supplemented with two others: Marak 1975, and Holbrook 1998 [1940]. An early English to Garo dictionary (Mason 1905) is reprinted periodically, and dictionaries based closely upon it but under the names of other authors have been published as well. The English to Garo dictionaries all consist of Garo definitions of English words rather than Garo equivalents for English words. This is useful for Garos, for whom the dictionaries were written, but it is awkward for someone who does not know Garo. Garo grammars are less satisfactory than the dictionaries. Keith (1874) and Phillips (1904) were early sketches of the grammar written by missionaries. After a period of ethnographic fieldwork, I wrote a somewhat amateurish grammar of the Garo Hills dialect of Garo (Burling 1961). A considerably fuller grammar that focused on the form of Garo spoken in Bangladesh was published later (Burling 2004).

2 PHONOLOGY

The Roman orthography, designed by American missionaries and used by the Garos, is very good, and in order to ease comparison with other publications I will stay close to the conventional spelling. Thus I write *ch* and *ng* where *c̄* and *ŋ̄* would be more conventional among linguists, and I write *p*, *t*, *k*, where *ph*, *th*, and *kh* would be more accurate phonetically. I even use the apostrophe rather than *ʔ* for the glottal stop because Garos themselves use either an apostrophe or raised dot when they write it. For a linguist, the major defects of the conventional spelling are: first, a failure to mark syllable boundaries, which results in a fair number of ambiguities; second, the omission of some /i/s in situations where the vowel is, admittedly, very short, and third, the tendency of many writers to omit the symbol for the glottal stop, probably because it is has no part in what Garos regard as the ‘English’ alphabet, and so it does not seem like a real letter. I indicate syllable boundaries with a dot, and by writing a few /i/s where Garo writers do not.

Garo does not have tones, but its syllable structure is very much like that of an East or Southeast Asian tone language. Syllable boundaries are phonologically sharp, and except in borrowed words (of which there are a great many), 90 percent of syllable boundaries probably correspond to morpheme boundaries. A single syllable virtually never represents more than a single morpheme, but some two-syllable morphemes are found, even among native words.

Syllable initial consonants and consonant clusters are shown in Table 11.1, and finals in Table 11.2. Initial /p, t/ and /k/ are aspirated and /b, d/ and /g/ are voiced, very much as in English. Syllable final stops are unvoiced but also unaspirated and unreleased. The nasals are all very much as in English, even to the extent that /ng/ does not occur initially. /s, ch/ and /j/ are all more palatalized than English /s/ but less so than English /ch/ or /j/. /s, ch/ and /j/ are homorganic. /r/ is a flap. Except in borrowed words, /l/ does not occur as a syllable initial (hence the parentheses in the chart) and /r/ does not occur as a syllable final, so they are in complementary distribution and could be transcribed with the same symbol.

TABLE 11.1 SYLLABLE INITIAL CONSONANTS AND CONSONANT CLUSTERS

p	t	k	pr	tr	kr
b	d	g	sp	st	sk
			spr		skr
m	n		mr		
s	ch	j	sr	chr	jr
w	t	(l)	h		

TABLE 11.2 CODAS: SYLLABLE FINAL CONSONANTS AND CLUSTERS

p	t	k	'
m	n	ng	
m'	n'	ng'	
l	l'	(r)	(s)

Even ignoring the problem of borrowed words, however, a system of writing that does not mark syllable boundaries is made clearer by writing them differently. *Mol.a* ‘tobacco mixture’ and *mo.ra* ‘round basketry stool’ are pronounced very differently. /s/ occurs as syllable final only in borrowed words.

Glottal stops can occur syllable finally, either alone or in combination with a nasal or an /l/. Minimal pairs for the presence and absence of a glottal stop are plentiful: *cha-a* ‘grow,’ *cha'-a* ‘eat’; *ring-a* ‘drink,’ *ring'-a* ‘sing.’ When used with a nasal or lateral, the glottal stop is pronounced right in the middle of the other phone, but Garos conventionally write it second. This has the advantage of marking the end of the syllable and so avoiding a few ambiguities, but it has the disadvantage of making a rather simple morphophonemic process seem more complex than it really is. A glottal stop never occurs in the final syllable of a Garo word, and whenever a glottal threatens to appear as word final, an echo vowel is inserted that protects it. For example, the combining form *do'-* ‘bird’ (*do'-ni* ‘bird’s; *do'-tip* ‘nest’) becomes *do'.o* when no other bound morpheme follows. Syllables that end with both a glottal and another consonant undergo a similar change: *gol'-* ‘stick’ (*gol'-chok* ‘pointed stick,’ ‘stake’; *gol'-ko* ‘stick’ accusative) becomes *go'.ol* when used without a suffix. The rule that inserts the echo vowel suggests that the glottal stop is rather insecurely joined to the other consonant of its ‘cluster,’ and I have even suggested that the glottal stop is a rather tone-like constituent of the syllable (Burling 1992). For a contrary opinion, see Duanmu (1994).

Orthographic Garo has five simple vowels. This is phonologically appropriate and the only serious complication is that /i/ embraces both high front and high back unrounded vowels. Since, in native Garo words, high front vowels are found only in open syllables and high back unrounded vowels only in closed syllables, they are in perfect complementary distribution. Writing them with the same vowel would be entirely appropriate if syllable boundaries were consistently marked. Since they are not, ambiguities occasionally arise. In fact, similar though less salient ambiguities arise with all the vowels, since they are all shorter in closed than in open syllables, but even native speakers find the phonetic difference between open and closed syllable /i/ to be highly salient while the rather modest length variation shown by other vowels is hardly noticed.

The glottal stop never occurs in the second syllable of a word. The loss of the glottal stop is apparent in many pairs such as *pil-a* ‘return’ and *kat-pil-a* ‘run back’ (*kat-* ‘run’). The glottal stop reappears in third syllables, as in *kat-ba-pil-a* ‘run back here’ (*-ba-* ‘in this direction’). It is difficult to construct fully convincing examples where a syllable with an underlying glottal stop appears as the fourth syllable in a word, but it seems to disappear in that position just as it does in second syllables.

Any two vowels can be adjacent if they occur in successive syllables without intervening consonants. It is difficult to find clear criteria by which to consider some vowel sequences but not others to be diphthongs, but /ai/ and /ao/ occur regularly enough with no morpheme break between them to suggest that they should be counted as diphthongs and thus to constitute a single syllable: *ai.ao* ‘wow,’ ‘my gosh!’.

3 VERBS

At its simplest, a Garo sentence requires nothing except a verb base and a tense suffix. Optionally, one or more nouns, noun phrases, pronouns and adverbs can precede the verb, and with the help of additional affixes, the verb itself can be made very complex. Here is a very ordinary Garo verb that can act as a complete sentence.

<i>A.gan</i>	<i>-chak</i>	<i>-tai</i>	<i>-ja</i>	<i>-wa</i>	<i>-kon.</i>
speak	-answer	-again	-NEG	-FUT	-probably
‘[He] will probably not answer again.’					

The only obligatory parts of this verb are the verb base *a.gan-* ‘speak’ (a two-syllable morpheme) and *-wa*, a tense marker for ‘future.’ The three morphemes that occur between the verb base and the tense marker are examples of an extensive class that I will call ‘adverbial affixes.’ A much smaller number of suffixes can follow the tense marker. I will call these ‘post-tense suffixes.’

The future is shown by *-wa* only when it follows the negative *-ja-*. In positive sentences the future marker is *-gen*. *-gen* and *-wa* form the only fully suppletive pair in the language and their alternation is one of the very few genuine morphological irregularities. Usually, morphemes follow each other with almost no phonological modifications. In addition to *-gen/-wa*, the tense markers include *-a* ‘present,’ ‘neutral,’ *-a-ha* ‘past,’ *-gin-ok* or *-na-jok* ‘immediate or intentional future,’ *-bo* ‘imperative,’ *-na-be* ‘negative imperative,’ and *-jok* a suffix that indicates a change of state. *-jok-* can often be translated by a perfect tense: *cha'-jok* ‘has eaten,’ but a more literal translation would be ‘has changed from the state of not having eaten to a state of having eaten.’ The literal meaning becomes clear in the negative: *cha'-ja-jok* ‘not eat any more,’ or, more literally, ‘has changed from a state of eating to a state of not eating.’

Post-tense suffixes include *-chim* a kind of perfect or irrealis marker. It shows that the proposition is untrue now but that it was true once or might be true at some other time: *re'ang-gen chim* ‘would go’ (*-gen* ‘future’); *re'-ang-a-ha-chim* ‘had gone’ (*-a-ha* ‘past’); *re'-ang-gin-ok-chim* ‘would like to go’ (*-gin-ok* ‘intentional future’).

Other post-tense suffixes include *-kon* ‘probably,’ *-ma* ‘question marker for yes-no questions’; *-mo* ‘question marker used when expecting agreement’ (i.e. tag questions), *-na* ‘it is said’ (quotative), *-ne* ‘please’ (used to soften imperatives).

In addition to the suffixes that can follow a tense marker, Garo has scores of affixes that can be placed between the verb base and the tense marker. These include non-productive affixes that can only be added to a limited number of verb bases and that sometimes confer quite idiosyncratic meanings. For example, *-chak-* generally indicates some action

directed towards another person: *a.gan-chak-a* ‘answer’ (*a.gan-a* ‘speak’), *dak-chak-a* ‘help’ (*dak-a* ‘do,’ ‘make’), *ra'-chak-a* ‘borrow’ (*ra'-a* ‘take,’ ‘bring’), *ka-sa-chak-a* ‘feel pity’ (*ka-sa-a* ‘love’). Several affixes that show direction of motion can be used freely with verbs that describe motion, but not with others: *-ba-* ‘in this direction’; *-ang-* ‘away’; *-on-* ‘downward.’ Still others are fully productive: *-rong-* ‘habitually,’ *-be-* ‘very,’ *-tai-* ‘again,’ *-grik-* ‘each other, reciprocal,’ *-at-* ‘causative,’ *-tok-* ‘all,’ *-ku-* ‘still,’ ‘yet,’ *-ja-* ‘negative,’ *-eng-* ‘progressive,’ and many others. Several of these affixes can be used together with the same verb and their order is almost completely fixed. The least productive affixes are always closest to the verb base while the increasingly productive ones come later. The following examples illustrate a few of the most productive of these affixes:

a-song -tai -ku -ja -eng -a
sit -again-yet-NEG-PROG-PRES
‘is not yet sitting again’

dak-chak -grik -at -a
help -RECIPROCAL -CAUSE -PRES
‘make [them] help each other’

bil-on -rong -a-ha
fly-down -habitually -PAST
‘regularly flew down’

The pieces that are glued together to form verbs can be chosen with great freedom. Except for the least productive of these affixes, their meanings are transparent and consistent. They might almost be regarded as separate words rather than bound morphemes, but there is an intonational unity to the entire set of morphemes that form a verb, and even when they are widely separated, the two obligatory parts, the verb base and the tense marker, pull the whole set together. Some, though by no means all, of the affixes that go between the verb base and the tense marker are transparently derived from independent verbs: *pil'-a* ‘return’; *re'-ba-pil'-gen* ‘will come back.’

4 NOUN PHRASES

4.1 Word order

Garo sentences require nothing except a verb. Even though Garo has no hint of verb agreement, neither agent, patient, nor any other participant needs to be explicitly mentioned so long as the larger verbal and non-verbal context makes the intended meaning clear. Nouns, noun phrases, pronouns, and adverbs are, of course, frequently used to flesh out the meaning of a sentence, and noun phrases can have great internal complexity. Nevertheless, they are not essential.

Like most of its Tibeto-Burman cousins, Garo is a verb final language. When asked to provide linguistic examples, speakers almost always place the verb last. In running speech, however, it is not uncommon for a pronoun, or more rarely a noun, to be moved to the postverbal position. Occasionally, they are even copied to the postverbal position.

Cha'-ku-ja, ang-a-de
eat-yet-NEG 1sg -EMPH
‘[I] haven’t eaten yet, not me.’

Bi-a gip-in song -o-na kat -ang -a-ha, bi-a.
 3sg different village -LOC-DAT run -away -PAST 3sg
 'He/she ran away to a different village.'

Postverbal pronouns such as these are clearly set apart from the rest of the sentence by their intonation, and most pronouns and noun phrases come before the verb. The order of preverbal pronouns and noun phrases relative to each other is quite free. The subject more often precedes than follows the object, but the role of each pronoun and noun phrase is so clearly shown by its case marker that the order can easily be changed.

4.2 Pronouns

Among the simplest noun phrases are the personal pronouns, but unlike nouns, several pronouns have a free (or nominative) form that differs from the form to which other case markers are attached.

Notice that four of these pronouns end with *-a*. When any (other) case marker is used, it replaces the *-a*, and the *-a* can be regarded as a nominative suffix that is used only with monosyllabic pronouns. The *-a* forms of the pronouns are also used as the free or citation forms. Polysyllabic pronouns, like nouns, lack any overt mark for the nominative. *na'-a* 'you singular' has an irregular combining form, *nang'*, another of the handful of morphological irregularities in Garo. All the case markers can be added to pronouns as easily as to nouns: *ang-o* 'with/'by me' (locative), *nang'-ni* 'your' (genitive), *ching-ko* 'us' (accusative), *na'-sim-ang-na* 'to you all' (dative).

Garo pronouns do not have phonologically reduced forms. If a pronoun is pronounced at all, it is fully stressed, and where English might use a reduced form, Garo simply leaves out the pronoun entirely, relying on the context to provide the sense.

4.3 Complex noun phrases

Noun phrases that are more complex than pronouns can have some or all of the following constituents although none is obligatory, not even the noun: 1 demonstrative; 2 genitive; 3 classifier phrase; 4 modifier (deverbal adjective, relative clause); 5 noun; 6 case marker and 7 postposition. Demonstratives and genitives always come first, while case markers and postpositions always come last, so, when present, these two constituents frame the noun phrase. Case markers are suffixed to the final constituent of the noun phrase (except for any following postposition), whatever that may be, so they are more accurately called 'clitics' than 'suffixes.' Classifier phrases and simple modifiers more often follow the noun than precede it, but they can, and often do, precede instead. Relative clauses always precede. When both a modifier and a classifier phrase are used in the same noun phrase, there is some tendency to put one, often the modifier, before the noun and the other after it, although any order is possible. If both the modifier and classifier phrase are placed on the same side of the noun, the modifier is always closer to the noun. A noun is not a required constituent of a noun phrase. A demonstrative, a classifier phrase, or a modifier can be used with no noun at all, but they will still take a case marker like any other noun phrase. See Table 11.3.

4.3.1 Demonstratives

The most important demonstratives are *i-a* 'this' and *u-a* 'that.' They can be used as either adjectives or pronouns. In a language without obligatory articles *u-a*, and less often *i-a*, is

TABLE 11.3 PRONOUNS

	Free/Nominative	Combining
I	<i>ang-a</i>	<i>ang-</i>
you, sg.	<i>na'-a</i>	<i>nang'</i>
he, she	<i>bi-a</i>	<i>bi-</i>
we, exclusive	<i>ching-a</i>	<i>ching-</i>
we, inclusive	<i>an'-ching</i>	<i>an'-ching-</i>
you, pl.	<i>na'-sim-ang</i>	<i>na'-sim-ang-</i>
they, human	<i>bi-sim-ang</i>	<i>bi-sim-ang-</i>

often used when a definite meaning is essential: *u-a man-de* ‘that person,’ ‘the person.’ As pronouns, *i-a* and *u-a* are the nearest equivalents to English ‘it’ and, like other pronouns, they can take case markers. Like other monosyllabic pronouns, they drop their final *-a* when another case marker is added.

Ang-a u -ko nik -a-ha.
 1sg DEM -ACC see -PAST
 ‘I saw it, I saw that.’

4.3.2 Genitives

Genitives are formed by suffixing the case marker *-ni* to a noun or pronoun. A genitive always precedes the name of the thing possessed: *ang-ni jak* ‘my hand’; *nang'-ni ma'-gip-a* ‘your mother’; *bi-ni nok* ‘his/her house’; *nok-ni bol-gru* ‘the ridge pole of the house.’ A genitive can be used without mentioning the thing possessed and can even be followed by another case marker:

Ang-ni-ko ni -bo.
 1sg-GEN-ACC look-IMP
 ‘Look at mine.’

4.3.3 Classifier phrases

Garo has a rich set of numeral classifiers that are used with numbers and chosen according to the nature of the thing being counted: people, animals, roundish things, thin flat things, long thin things, poles, posts, slices, portions, parts, teams, groups, kinds, number of times, abstract things such as stories or ideas, and many others. Even unsophisticated speakers are sufficiently aware of the classifier system to advise learners to use *-ge* when a more specific classifier is not known. *-ge* covers a residual category of mostly physical objects.

In addition to these core classifiers, three other sets of words are used in very much the same way as classifiers, though each has its own special characteristics.

- 1 Containers: the name of any container can be used to count units of the amount it can hold. Borrowed words pose no problem. *gil-es-gin-i* ‘two glasses of,’ *nok-git-tam* ‘three houses of’ (i.e. ‘three families’).
- 2 Time words: units of time can be used with numbers just as classifiers can, but unlike ordinary classifiers the resulting phrase cannot be used with a noun: *sal-sa* ‘one day’; *wal-gin-i* ‘two nights’; *ja-git-tam* ‘three months’.

- 3 Measures: all units of weight and size can be used with numbers to indicate the amount of some material that is being counted. Again, borrowed words pose no problem: *ba'-ra mik-sa* ‘one cubit’ (the length from finger tip to elbow) ‘of cloth,’ *mail-bri* ‘four miles,’ *gong-bri* ‘four rupees.’

Classifiers are never used without a number, and numbers only rarely without a classifier. Classifiers are sometimes, though not always, omitted when counting. It is quite possible to count *sak-sa*, *sak-gin-i*, *sak-git-tam* . . . ‘one person, two people, three people . . .,’ but *sa*, *gin-i*, *git-tam* . . . ‘one, two, three . . .’ will also do. Classifiers with numbers often modify nouns. Typically they follow the noun but occasionally they precede it: *me'-chik sak-git-tam*, *sak-git-tam me'-chik* ‘three women’ (*sak-* ‘classifier for people’); *meng-go mang-bong-a* ‘five cats’ (*mang-* ‘classifier for animals’). Classifiers with numbers are often used with no noun at all. Their semantic specificity lets them convey considerable information about what is being counted, and often this is all the information that is needed in the context. It would be entirely normal to report having seen three people without using any noun, as long as it is not necessary to say what sort of people they are:

Sak -git-tam-ko nik-a-ha.
people-three -ACC see-PAST
'[I] saw three people.'

It is often possible to choose among several alternative classifiers for the same noun: *te'-rik rong-sa* ‘two bananas’; *te'.rik pang-sa* ‘two banana trees’; *te'-rik gal-sa* ‘one small bunch (hand) of bananas’; *te'-rik ol-sa* ‘one large bunch (arm) of bananas.’ This, and the ease with which classifiers can be used with no noun at all, makes it impossible to regard them as constituting some elaborate system of gender in which the choice of classifier is governed by the noun.

A few other morphemes than just the numerals can be used with classifiers: *prak* ‘each,’ *sak-prak* ‘each person,’ *mang-prak* ‘each animal’; *gim-ik* ‘whole,’ ‘entire,’ *mang-gim-ik* ‘whole animal,’ ‘whole body (of an animal),’ *sal-gim-ik* ‘all day’; *gip-in* ‘other,’ ‘another,’ *mang-gip-in* ‘another animal’; *-san* ‘alone,’ ‘only’ (cf. *sa* ‘one’), *sak-san* ‘alone (of a person).’ Garo lacks obligatory articles, but a classifier with *-sa* ‘one’ is often used where an indefinite article would be used in English: *mat-chu mang-sa* ‘one tiger,’ ‘a tiger.’

4.3.4 Modifiers

Most meanings that, in English, are conveyed by adjectives are, in Garo, conveyed by words that act syntactically like intransitive verbs. Only a handful of core adjectives have distinct syntactic characteristics and these are quite idiosyncratic. Both intransitive and transitive verbs can be used to modify a noun.

A verb is put into a form that can modify a noun by means of a nominalizing suffix that is placed in the position that would otherwise be occupied by a tense marker: *-a* is either the same as the present-neutral tense marker or homophonous to it; *-gip-a* is used in essentially the same circumstances as *-a*, but its fuller form makes it more explicit and it is more likely to be used in long and complex constructions. *-gip-a* is also more likely to be used when a modifier precedes the noun while *-a* is often used when it follows. The longer form is appropriate to the less common (or ‘marked’) position of the modifier, but either suffix is possible in either position: *mat-chu dal'-a*, *dal'-gip-a mat-chu* ‘big cow.’

Ang-a mat-chu dal'-a-ko nik -a-ha.
1sg cow big -NOM-ACC see-PAST
'I saw the big cow.'

Transitive verbs can modify nouns just as intransitives can, but they are often most naturally translated into English as relative clauses.

Cha'-eng-gip-a man-de ok-a-gen.
eat-PROG-NOM person satisfied-FUT
'[The] person(s) who is/are eating will be satisfied (full).'

Satellites of the verb can be drawn into a Garo modifier just as they can be drawn into an English relative clause. In the following example, most of the first sentence is turned into a relative clause in the second sentence, where it modifies *man-de* 'person.'

Ang-a u-a man-de -ko me'-ja -o nik -a-ha.
1sg that person -ACC yesterday -LOC see -past
'I saw that person yesterday.'

Ang-ni me'-ja -o nik-gip-a man-de da'-al -o re'-ang-a-ha.
1sg-GEN yesterday -LOC see-NOM person today -LOC move-away-PAST
'The person I saw yesterday went today.'

As the examples suggest, the line between modification by an adjective and modification by a relative clause is less sharp in Garo than in English. However, as soon as satellites are pulled into a modifier the resulting clause must precede the noun rather than follow it. Single-word modifiers more often follow. As befits its relative complexity and its position before the noun, a modifier with many constituents is also more likely to be marked by *-gip-a* rather than *-a*. Like demonstratives and classifier phrases, modifiers can be used without a noun, or, perhaps, such words should themselves be regarded as nouns: *Dal'-a-ko nik-a-ha-ma?* 'Did [you] see the big [one]?' When suffixed to a verb base, *-gip-a* means, approximately, 'the one who' and the resulting word can be used either alone or to modify a noun: *cha'-eng-gip-a* 'the one who is eating,' 'the eater,' *cha'-eng-gip-a me'-chik* 'the woman who is eating'; *dal'-gip-a* 'the big one,' *dal'-gip-a man-de* 'the big person.'

4.3.5 Nouns

Demonstratives, classifier phrases, and modifiers can all be used without a noun, but when a noun is present it forms the center of its phrase. Together with verbs, nouns form one of the two largest Garo word classes, but unlike verbs, nouns are frequently used with no suffix at all. Many nouns are both monomorphemic and monosyllabic, but Garo probably has more bi- and tri-syllabic nouns (many of them also bi- and tri-morphemic) than some other Tibeto-Burman languages. The largest number of polysyllabic nouns are compounds, at least one part of which has a transparent meaning, even if that part never occurs alone. *Mik-* the Garo reflex of the widespread Tibeto-Burman word for 'eye,' never occurs alone in Garo, but it does occur in many compounds where it clearly means 'eye': *mik-ron* 'eye'; *mik-chi* 'tear' (*chi* 'water,' 'liquid'); *mik-gil* 'eyelid' (cf. *bi-gil* 'skin'); *mik-sim-ang* 'eyebrow' (cf. *pak-sim-ang* 'underarm hair,' *re-sim-ang* 'male pubic hair') etc. Of the morphemes appearing in these compounds, only *chi* 'water' can be used alone. All the rest occur only in compounds.

Many compounds begin with a classificatory morpheme, almost always a single syllable. For example, dozens of names for birds have *do'-* as their first syllable: *do'-til-eng* ‘wood-pecker,’ *do'-po* ‘owl,’ etc. The second part of such words gives the word its specific meaning but many of these second parts are never used otherwise. As far as I am aware, *-til-eng* and *-po* never occur except in the words for ‘woodpecker’ and ‘owl.’ Many names for fish start with *na-* and many names for varieties of trees start with *bol-*. Many body part terms are constructed in the same way. *Jak* ‘hand,’ ‘arm’ and *ja'-a* ‘leg,’ ‘foot’ can be used alone, unlike *mik-* ‘eye,’ but they also enter into many compounds: *jak-pa* ‘palm,’ *ja'-pa* ‘sole of the foot’; *jak-sku* ‘elbow,’ *ja'-sku* ‘knee’; *jak-si* ‘finger,’ *ja'-si* ‘toe,’ and many others. Many plant parts have *bi-* as their first syllable: *bi-gil* ‘bark,’ ‘skin,’ ‘peel,’ *bi-bal* ‘flower,’ *bi-gron* ‘pit,’ ‘large seed of a fruit.’ Unlike the more generalized ‘prefixes’ that occur in some Tibeto-Burman languages, these classifying first syllables of Garo nouns have readily identifiable meanings.

Nouns can also be formed from verbs by means of the nominalizing suffix, *-a-ni*. This yields a noun with an abstract meaning something like ‘that which is . . .’: *cha'-a-ni* ‘that which is eaten,’ ‘food’ (*cha'-a* ‘eat’), *chan-chi-a-ni* ‘that which is thought,’ ‘thoughts’ (*chan-chi-a* ‘think’).

A number of suffixes can be added to nouns, but none is obligatory. *-rang* ‘plural,’ precedes any case marker that may be suffixed to the noun. The absence of *-rang* does not necessarily imply ‘singular,’ and to make singularity explicit, a classifier phrase with *-sa* ‘one’ must be used. Several other noun suffixes can follow case markers: *-de* ‘emphatic,’ *-sa* ‘only,’ *-ba* ‘also.’ *ang-ko-sa* ‘only me’ (accusative); *te-bil-o-ba* ‘on the table also’ (locative).

4.3.6 Case markers and postpositions

The final constituent of a noun phrase is the case marker, sometimes with a following postposition. It is difficult to draw an absolutely clear line between case markers and postpositions, but the following certainly count as case markers.

- Zero/-*a*, nominative: only monosyllabic pronouns and demonstratives have *-a* in the nominative. The *-wa* of *sa-wa* ‘who?’ and *ja-wa* ‘someone else’ can be regarded as an irregular nominative suffix that, like the *-a* of monosyllabic pronouns, is lost when some other case marker is added. Nouns and longer pronouns lack an overt marker for the nominative.
- *-ko*, accusative: Garo is a straightforward nominative–accusative language and most objects are marked with *-ko*, but the *-ko* can be omitted when its noun immediately precedes the verb and when the meaning is not definite. Adding *-ko* gives the noun phrase a definite sense: *bol den'-a* ‘chop wood,’ *bol-ko den'-a* ‘chop the wood’; *mi-ko cha'-a* ‘eat the rice’; *mi cha'-a* ‘eat a meal.’
- *-na*, dative, ‘for, to’: *ang-a nang' -na ki-tap-ko on' -gen* ‘I will give you the book.’ A number of postpositions regularly follow datives: *ang-na skang* ‘before me.’
- *-ni*, genitive: *-ni* is a straightforward genitive case marker that can be used to show possession of body parts, kinsmen, and physical objects. Many postpositions follow *-ni*. Some of them are transparently derived from nouns: *ang-ni jang-gil-o* ‘behind me,’ ‘at my back,’ (*jang-gil* ‘back of the body,’ *-o* ‘locative’). Other postpositions have no obvious etymology, although they are parallel in form: *bi'-sa-ni gim-in* ‘because of [the] children,’ (*bi'-sa* ‘child/children,’ *gim-in* ‘because of’).
- *-o*, locative, either temporal or spatial, ‘in,’ ‘at,’ ‘on’: *nok-o* ‘at the house,’ ‘at home,’ *kin-al-o* ‘tomorrow,’ *wal-o* ‘at night.’

- *-o-na* ‘towards,’ ‘in the direction of’: *-o-ni*, ‘from,’ ‘in the direction away from.’ *Tu-ra-o-ni Reng-sang-gri-o-na* ‘From Tura to Rengsanggri’; *pring-o-ni wal-o-na* ‘from morning to night.’
- *-chi*, locative, spatial only: *-chi* sometimes indicates movement with respect to a place rather than simply a position at a location: *Tu-ra-chi* ‘in Tura,’ ‘to Tura,’ ‘from Tura.’
- *-chi*, instrumental, ‘with,’ ‘by means of’: this is homophonous with the spatial locative but the meanings seem quite distinct: *ru-a-chi* ‘with an axe.’
- *-ming*, ‘along with,’ ‘accompanying’: *ang-ming* ‘with me.’

Large numbers of postpositions can follow one or another of the case markers, the genitive *-ni* taking the most. It is not always clear whether an ending should count as a case marker or a postposition. A postposition that follows the nominative case (which is generally not marked) would usually be indistinguishable from a case marker, except when following a monosyllabic pronoun where the nominative is distinguished by a final *-a*. Unfortunately for those who like their grammar unambiguous there is some vacillation. Either *ang-a gri* or *ang-gri* can be used to mean ‘without me.’ The first *gri* seems like a postposition, since it follows a distinctive nominative form of the pronoun. The second is like a case marker, since it is attached directly to the pronoun’s combining form. Even the *-na* of *-o-na* ‘towards’ and the *-ni* of *-o-ni* ‘from’ might conceivably be considered postpositions that follow the locative *-o*.

Other words that consistently follow one of the case markers are less ambiguously postpositions: *gim-in* ‘because of,’ *a-chak-ni gim-in* ‘because of the dog’; *king-king* ‘until,’ *kin-al-o-na king-king* ‘until tomorrow’; *pal* ‘instead of,’ *bi-ni pal* ‘instead of him’; *skang* ‘before,’ *ang-na skang* ‘before me,’ and many others.

A simple comparative can be formed by inserting the affix *-bat-* ‘more’ into a verb: *chon-bat-a* ‘smaller,’ *neng'-bat-gen* ‘will be more tired.’ Depending on the context *-bat-* can also imply the superlative: *Sa-wa dal'-bat-a?* ‘Who is the biggest?’ When the object with which something is compared is explicitly mentioned, it is, like all nouns, placed before the verb. It takes the dative case marker *-na*, which is followed by *bat-e* ‘more.’ The *-na-bat-a* in turn is followed by a later echoing *-bat-* of the verb: *chang-ro-bat-a* ‘taller.’

U-a me'-chik bi-ni se-gip-a-na bat-e chang-ro-bat-a
that woman 3sg-GEN husband-DAT than tall-more-PRES
'That woman is taller than her husband.' (literally: 'That woman, more than her
husband, is taller.')

5 ADVERBS AND REDUPLICATION

In addition to its nouns, pronouns, verb bases and classifiers, all of which join with other morphemes to form words, Garo has a large class of adverbs, many of them reduplicative or partially reduplicative, that take no affixes at all: *pang-nan* ‘always’; *bak-bak* ‘quickly’; *sruk-sruk* ‘quietly,’ ‘secretly’; *jol-jol* ‘directly,’ ‘systematically’ *pil-ap-pil-ap* ‘in a flapping manner,’ *pil-eng-pil-eng* ‘rocking back and forth,’ *gu-rung-ga-rang* ‘aimlessly’ (of wandering about). Many reduplicative, or partially reduplicative, adverbs are transparently derived from verbs: *ring-reng-ga-reng* ‘in a back and forth swinging manner,’ from *ring-reng-a* ‘to swing back and forth’; *rip-ong-rip-ong* ‘flying around’ from *rip-ong-a* ‘to fly around’; *srot-srot* ‘in a sliding manner,’ from *srot-a* ‘slip,’ ‘slide.’ Such adverbs, however, cannot be productively created from any verb at all. Adverbs are often

placed directly before the verb and thus after any noun phrases that the sentence may have, but they can come earlier in a sentence as well.

Some adverbial affixes that are used within verbs also have reduplicated forms, many of them conveying a sense of repetition or continuous action: *chot-tip-tip-a* ‘break (string) into bits,’ from *chot-a* ‘break’; *ru-kring-krang-a* ‘pour all around,’ from *ru-a* ‘pour’; *sel-gol-gol-a* ‘leak a lot,’ ‘rapidly,’ from *sel-a* ‘leak.’ A few verb bases, again often conveying repetitive actions, are reduplicated in form: *jok-jok-a* ‘bounce (as when riding in a bus),’ *deng-deng-a* ‘squirm,’ ‘wiggle.’

Finally, reduplication of numbers to convey a distributive sense is fully productive: *gong-gin-i-gin-i* ‘two rupees each’ (*gong-* ‘classifier for rupees,’ *gin-i* ‘two’); *le-ka king-git-tam-git-tam-ko on'-bo* ‘give three sheets of paper to each’ (*le-ka* ‘paper,’ *king* ‘classifier for thin flat things,’ *git-tam* ‘three’).

6 COMPLEX SENTENCES

The simple Garo sentences that I have considered so far can include various noun phrases and adverbs and a verb. Complex sentences are built from two or more simple sentences of this kind. Only a few of the most common types of complex sentences can be illustrated here.

Very often, one sentence is turned into a subordinate clause that precedes the main clause. The subordination is shown by a verb suffix that fills the position that would be occupied by the tense suffix in a sentence-final verb. The most common of these subordinating suffixes are *-e* and *-e-ming*:

Ang-a nok-chi sok-ang-e, cha'-a-ha.
1sg house-LOC arrive-away-SUB eat -PAST
'Having arrived at the house, I ate.'

Ang-ni a-bi-tang nam-en neng'-be -e-ming, nok-chi nap-ang-a-ha
1sg-GEN sister very tired-very -sub house-LOC enter-away-PAST
'My sister, being very tired, went into the house.'

Subordinating constructions of this sort are exceedingly common in Garo and they are used to tie together what would otherwise be separate sentences. This gives unity to the discourse by creating what amount to long run-on sentences. They give an impression like an English monologue in which the sentences are linked with phrases such as ‘and then’ or ‘so.’

A new sentence can also begin with a subordinate form of the same verb that completed the previous sentence. This, too, ties the successive sentences together. In the following example, the verb *sok-* completes one sentence and introduces the next. The second sentence has a short introductory *-e* clause followed by a somewhat longer main clause:

Dos ba-ji -o ang-a song -chi sok -ang -a-ha. Sok-ang-e ang-a mi cha'-e tu-si -a-ha.
ten o'clock-LOC 1sg village-LOC arrive -there -past. Arriving -there -SUB 1sg rice eat -SUB sleep -PAST

‘At ten o’clock I arrived at the village. Having arrived, I ate rice and slept.’

A set of subordinating suffixes with more specific meaning than *-e* includes *-o-de* ‘if’; *-o-ba* ‘although’; *-o-sa* ‘only if.’

Bi-a re'-ba-ja-o-de, ang-a ka-o-nang -be -gen
3sg come-here-NEG-if 1sg angry -very -FUT
'If he doesn't come I will be very angry.'

The infinitive suffix, *-na*, also allows two verbs to be used together.

An'-ching re'-ang-na nang-a.
1plNCL go-away-INF need-PRES
'We need to go.'

Notice that we have now considered three types of verb suffixes, all of which occur in the same position within the verb:

- 1 Tense suffixes, one of which completes each sentence.
- 2 Nominalizing suffixes such as *-gip-a* that put a verb into a form that can act as a noun or modify a noun.
- 3 Subordinating suffixes such as *-e* and *-e-ming*.

Every verb requires a suffix in this position, and the type of suffix is determined by the syntactic role of the verb or clause: main, nominalized, or subordinate.

Verbs can also be subordinated by means of a conjunction such as *gim-in* 'because of' or *ja'-man-o* 'after' that follows a verb that has been nominalized by *-a-ni*:

Sa -a-ni gim-in, ang-a re' -ang -na man' -ja.
sick -NOM-GEN because 1sg go -away -INF can -NEG
'Because of sickness, I cannot go / Because I am sick, I cannot go.'

Mi song -a-ni ja'-man-o, ching-a cha' -a-ni -ko on' -gen
rice cook -NOM-GEN after -LOC 1plexCL eat -NOM-GEN -ACC give -FUT
'After cooking rice, we will serve the food.'

In addition to questions made with question words and simple yes–no questions formed with the verb suffix *-ma*, Garo speakers frequently use balanced questions in which a yes–no question is immediately followed by a corresponding negative clause which may, but need not, have a question marker as well:

Cha'-gen -ma, cha'-ja -wa -ma?
eat -FUT -QUEST eat -NEG -FUT -QUEST
'Will you eat or not?'

Mong-ma-ko nik-jok -ma, nik-ku-ja?
elephant-ACC see-CSM-QUEST see-yet-NEG
'Have you seen elephants or not yet?'

Finally, Garo has a considerable number of conjunctions that can be used to tie successive clauses together. Some of these have specific meanings: *u-ni-gim-in* 'therefore,' literally 'because of that'; *un-bak-sa-ba* 'in addition to that'; *in-di-ba* 'but,' 'however'; *ong'-ja-o-de* 'or' literally 'if [it] is not.' Others mean little more than 'and then,' 'so': *u-ni-ko*, *u-non*, *in-di-de* and others. Speakers of other languages, however, may find the language curiously lacking in a simple equivalent for 'and' that can be used to conjoin not only entire clauses but simple nouns or simple verbs. Instead of conjoining two verbs with a word that means 'and,' the first verb is more often subordinated to the second. Two nouns may be used beside each other with no overt

conjunction at all, or both may be suffixed with *-ba* ‘also’: *na'-a-ba ang-a-ba* ‘you and I,’ ‘both you and I.’

7 LANGUAGE CONTACT AND LANGUAGE MAINTENANCE

On their northern, western, and southern borders most of the neighbours of the Garos speak Bengali or a closely related dialect of Assamese. Garos always call this language ‘Bengali’ and they have probably been borrowing from it for many centuries. The impact of Bengali is particularly strong among the Garos of Bangladesh, where all primary education is conducted in Bengali, and where it is needed for everyday dealings with peddlars, shopkeepers, government officials, and church leaders. All adult Garos in Bangladesh are able to use Bengali for practical purposes and many are fluent. Bengali presses less insistently upon Garos living in India, for it is not the language of education or government, and even Bengali traders learn enough Garo to deal with their customers in their own language. Nevertheless, even the dialects of the most remote areas of the Garo Hills have absorbed large numbers of borrowed words. The influence of English is more recent, but it now competes with Bengali as a source of borrowings. The influence of English is stronger among the Indian Garos than among Bangladeshi Garos in direct proportion to the relative weakness of Bengali.

The most obvious impact of Bengali and English comes with borrowed words. Many Bengali words are thoroughly assimilated into Garo, but educated and bilingual Garos borrow freely and on the fly. It would be impossible to speak about education, politics, Christianity or modern technology without calling on borrowed words, and Garos feel free to use any word from Bengali or English that they believe their listeners will understand. These words bring some innovations to the phonology, though mostly by placing familiar sounds in new positions rather than by introducing entirely new sounds.

The impact of borrowed words is great enough, particularly in Bangladesh, to worry some Garos. A few despair at the flood of Bengali words that they feel are corrupting their language, but they feel powerless to avoid them. Garos in Bangladesh receive all their education in Bengali and even in the Garo Hills, Garo medium education stops after elementary school. All high school subjects come with their foreign vocabulary. Too little has been printed in Garo to sustain a richly literate community, and well-educated Garos must rely upon English or Bengali for many literate purposes. Nevertheless, with 700,000 speakers, Garo is not yet on the list of endangered languages. Even in outlying areas like Bangladesh, most children of Garo parents still learn Garo as their first language. Whether they will still be doing so a century from now is by no means certain.

NOTE

- 1 My association with the Garos and with their language began in the 1950s when I spent two years conducting anthropological fieldwork in the Garo Hills. Starting in the 1980s, when it was impossible for a foreigner to work in Northeast India, I made several trips to work among the Garos living in Bangladesh. Only in 1996–7 was I able to return to Northeast India and to visit, once again, the people whom I had known 40 years earlier. My trips have been made possible by the splendid help of the Ford Foundation, the National Science Foundation and the Fulbright Foundation, and I am deeply indebted to all of them. Too many individuals have helped me over the years to let me list them all, but as always, I am most of all indebted to the hundreds of Garos who with great good cheer have helped me to learn about, and even to speak, their language.

REFERENCES

- Burling, Robbins. 1961. *A Garo Grammar*. Deccan College Monograph Series 25, Poona: Deccan College Postgraduate and Research Institute.
- Burling, Robbins. 1992. Garo as a minimal tone language. *Linguistics of the Tibeto-Burman Area* 15.5: 33–51.
- Burling, Robbins. 2004. *The Language of the Modhupur Mandi*, Vol. I: *Grammar*. New Delhi and Morganville, NJ: Bibliophile South Asia, in association with Promilla. Vol. II: *The Lexicon*, and Vol. III: *Glossary*: name.umdl.umich.edu/bbv9808 or www-personal.umich.edu~rburling/.
- Duanmu, San. 1994. The phonology of the glottal stop in Garo. *Linguistics of the Tibeto-Burman Area* 17.2: 69–82.
- Eliot, John. 1794. Observations on the inhabitants of the Garrow Hills, made during a publick deputation in the years 1788 and 1789. *Asiatic Researches*, Vol. 3, edited by Sir William Jones. Reprinted at New Delhi: Cosmo Publications, 1979: 17–37.
- Hamilton, Francis. 1940. *An Account of Assam: With Some Notices concerning the Neighbouring Territories*, edited by Suryya Kumar Bhuyan. Gauhati, Assam: Government of Assam (first published in 1820).
- Holbrook, L.M. 1998. *Ku'rongdik: A'chikku into English*. Tura, Meghalaya: Garo Literature Society (first publication from a 1940 manuscript).
- Keith, T.J. 1874. *Outline Grammar of the Garo Language*. Sibsagar, Assam: Baptist Mission Press.
- Marak, H.W. 1975. *Ku'bidik: A Garo–English–Assamese Dictionary*. Gauhati: Assam Academy for Cultural Relations.
- Mason, Marcus C. 1905. *English–Garo Dictionary*. Shillong: Assam Secretariat Printing Office.
- Nengminza, D.S. 1946 and later. *The School Dictionary: Garo into English*. Tura: Miranda Library.
- Phillips, Elnathan Gooding. 1904. *Outline Grammar of the Garo Language*. Shillong: Assam Secretariat Press.

CHAPTER TWELVE

HAKHA LAI¹

David A. Peterson

1 INTRODUCTION

A Kuki-Chin language spoken primarily in and around the city of Hakha in Chin State, Burma, and in adjacent areas of India and Bangladesh by over 100,000 people, Lai is also used extensively as a second language by speakers of other Chin languages in the Chin Hills. In recent years, a considerable diaspora population has developed.

Lai has an orthography developed by missionaries during the early part of the twentieth century which is used extensively, although it does not represent vowel length or tone, two essential features of the language's phonology. Certain aspects of the orthography, such as how to represent an alveolar/retroflex distinction in stops and where to mark word boundaries, are subject to ongoing debate.

Lai is a Central Chin language, closely related to Laizo or Zahao (Osborne 1975; King 2010; spoken in the Falam area), Bawm (Reichle 1981; spoken mostly in Bangladesh), and Mizo (Chhangte 1993; spoken chiefly in Mizoram State). Central Chin languages constitute a clear subgroup of the family, with numerous phonological and grammatical innovations distinguishing them from more northerly and southerly Chin languages (VanBik 2009).

2 PHONOLOGY

2.1 Segmental phonology

Table 12.1 gives the segmental phoneme inventory. Where the orthography used does not reflect more or less standard phonetic values, a conventional phonetic transcription is included.

There is little allophonic variation. Stops are unreleased finally. The distinction between short and long vowels, which corresponds to a measurable length difference, also manifests itself in terms of quality: the short vowels are in most cases phonetically more central than their long vowel counterparts. Lastly, there is an allophone [ʃ] of /s/ before [i].

2.2 Suprasegmental phonology

Tone has only a small functional load in distinguishing lexical minimal pairs, so it has often gone unnoticed in previous descriptions. Nonetheless, tonal distinctions are important (Hyman and VanBik 2002a, 2002b, 2004). In isolation, monosyllabic words distinguish two tones: falling and high level. When these tones occur in various morphological contexts, however, three tonal contrasts emerge, with two tones (falling and rising) corresponding to the isolation falling tone. It is difficult to find a minimal triplet, but the three-way tonal contrast may be clearly heard in the last syllables of the following

TABLE 12.1 SEGMENTAL INVENTORY

Consonants					Vowels	
Stops					Simplex vowels	
p	t	t̪	k	?	i [i], ii [i:] (y)	u [u], uu [u:] (w)
ph [p ^h]	th [t ^h]	th̪ [t̪ ^h]	kh [k ^h]		e [ɛ], ee [e:]	o [o], oo [ɔ:]
Affricates					Diphthongs	
	c/ts				a [ə], aa [a:]	
	ch/tsh [ts ^h]				Diphthongs	
	tl [tl]				uy, oy, ooy	iw, ew, eew,
	thl [tl̪]				ay [ey], aay	aw [ow], aaw
Fricatives and laterals					Triphthongs	
f	s [ʃ]			h	ia, iaa, ua, uaa	
v	z				uay	iaw
	r	l				
	hr [r̪]	hl [l̪]				
Nasals						
m	n		ŋ			
hm [m̪]	hn [n̪]		hŋ [ŋ̪]			

forms: falling *?a-diy* ‘he drinks’ vs level *?a-diy* ‘he is honest/it is straight’; level *?an-?ii-law* ‘they’re similar’ vs rising *?an-láw* ‘they disappear’; and falling *?a-mán* ‘it sells’ vs rising *?a-mán* ‘its price’.

Glottalization is a suprasegmental feature involved in marking distinctions in particular verbal ablaut classes and in a transitivizing derivation (see later). Where it is relevant, with open and stop final syllables it is realized as a final glottal stop; with sonorant finals it creates the phonetic impression of a glottalized sonorant.

2.3 Syllable structure and morphophonemics

Syllables have the form cv: or cv(:)c. Short vowels do not occur in open syllables. Long diphthongs and triphthongs do not occur in closed syllables. Any consonant of the consonant inventory may occur in syllable-initial position, but laryngeal distinctions (voicing and aspiration) are neutralized in syllable-final position.

Lai is monosyllabic in that there is an almost perfect one-to-one correspondence between the syllable and the morpheme, especially for function morphemes, but this is not to say that words are monosyllabic. Distributional and semantic considerations suggest highly complex word structure. Lai is almost purely agglutinative in that virtually no morphophonemic processes (other than tonal sandhi) occur at formative boundaries. The only consistent segmental morphophonemic process is shortening of long vowels in open syllables in the first member of compounds.

3 INFLECTIONAL MORPHOLOGY

Form classes include nominals (pronouns, nouns, and relational nouns) and verbals (verbal and adjectival). Demonstratives, discourse deictics (markers of information status),

quantifiers, and classifier-numeral compounds are minor class components of nominal phrases, and independent adverbs can further add to the structure of verbal phrases.

3.1 Nominal inflection

For the structure of nominal phrases, see section 4.1.

3.1.1 *Pronominals*

Independent pronouns are listed on the left-hand side of Table 12.2.

Independent (non-bound) pronouns distinguish three persons and two numbers and consist of the generic demonstrative element *ma?* combined with a pronominal element proper. The use of *-ni?* with plural pronouns as opposed to *-ma?* (i.e. *kan-ni?*, *nan-ni?*, etc.) contrastively focuses the pronoun (see Lehman and VanBik 1997). *-taa* added to either of these forms yields pronominals used in headless possessive phrases (translatable as, e.g. ‘mine’). Independent reflexive/reciprocal pronouns have the structure *pronoun=lee* pronoun (*kayma?=lee kayma?* ‘myself’, *kanma?=lee kanma?* ‘ourselves/each other’).

3.1.2 *Demonstratives*

Nominals may be modified by demonstrative elements. The generic demonstrative element is *ma?*. More specific demonstrative elements, and an admittedly oversimplified depiction of their semantics, include *khaa* (near addressee), *tsuu* (not visible), *hii* (near speaker), and *khii* (distal). See in particular Barnes (1998) and Bedell (2001a) for extensive discussion of the complex syntax and semantics of Lai demonstratives and other deictic elements.

3.1.3 *Possession*

There is no formal indication of possession other than juxtaposition of two nominals (and accompanying predictable tone sandhi). The first of two nominals in sequence will be interpreted as possessor of the second, as in *paalaw nuu* ‘Paalaw’s mother’. If there is no overt nominal possessor, and optionally even if there is one, possessed nouns bear one of a set of prefixes, also given in Table 12.2. These prefixes express person and number of the possessor and are identical in form to the verbal subject participant markers (see section 3.2.2). There are no differences in possessor marking based on semantic characteristics of the possessed entity (e.g. there is no alienable/inalienable distinction).

TABLE 12.2 PRONOMINALS

	Independent		Possessive prefixes	
	sg	pl	sg	pl
1	key- <i>ma?</i>	kan- <i>ma?</i>	ka-	ka- <i>n-</i>
2	naŋ- <i>ma?</i>	nan- <i>ma?</i>	na-	na- <i>n-</i>
3	?a- <i>ma?</i>	?an- <i>ma?</i>	?a-	?a- <i>n-</i>

3.1.4 Case and spatio-temporal relationships

With the exception of absolutives, which are unmarked, nominal phrases bear one of a set of clitic case particles. Agents of most notionally transitive verbs bear the ergative case clitic =*ni?*. The locative-allative case clitic =*?a?* marks static locations, locations towards which a figure moves, or temporal location. The instrumental-ablative case clitic =*?in* marks the instrument with which the action is performed, the location from which the action originates, or the area/medium through which the action occurs. Comitative nominals bear the case clitic =*hee*, standards of comparison bear the (etymologically complex) clitic =*naak-?in*, and standards in equative constructions are marked by =*thluk-?in*, ('be.equal-instrumental/ablative'). In subordinate clauses, the ergative, locative, and instrumental-ablative case particles have an optional allomorph, =*?ii*.

Most spatial notions are encoded through the use of abstract nouns which refer to locations, e.g. *tshuy* 'area inside', *tsuj* 'area on top', *tay* 'area underneath'. These typically occur as the possessed entity in a possessor–possessed relationship with the entity they relate to and are marked obliquely by either the locative or the instrumental-ablative case clitic (e.g. *?in tshuy=?a?* 'inside the house', *?in tay=?a?* 'underneath the house', etc.)

A few elements are categorially intermediate between true relational nouns like these and case particles in that they may occur with or without oblique case markers. The element *tiaj* occurs with nouns which express the extent (spatial or temporal) to which an action occurs. *tshuy*, which is primarily a relational noun, also has a marginal existence as a case particle attached to time phrases to indicate the duration of an activity. *koy* is used to indicate 'about, concerning'.

3.1.5 Number

Marking of plurality is not typically required in noun phrases; the verb bears obligatory marking for plurality of the subject and object. =*lee* (which also conjoins nominals and sometimes phrases), =*hnaa* (which also marks non-first person object plurality in the verbal complex), =*tee*, and =*pool*, are sometimes used to mark collectives. All of these particles have such a low text-frequency, that a reliable assessment of their function is impossible to make at this point.

3.1.6 Information status

Aside from an extensive battery of valence-affecting constructions which mark deviations from unmarked information status for given argument types, Lai has a set of postposed elements which play an essential role in indicating the discourse status of the nominal phrases they are associated with. The morphology of these discourse deictics is virtually identical to that of prenominal demonstratives, *hii* and *khii* have largely spatial connotations, even in the position of a discourse deictic; but *tsuu* occurs in many instances as a topicalizer in a strict, discourse-internal sense, while *khaa* tends to mark preceding entities as being within the shared knowledge of both speaker and listener more generally. See also section 3.2.4.4. on the utterance-final usage of related elements.

3.1.7 Numeral classifiers

Lai has a reasonably large set of numeral classifiers, which are compounded with numeral roots to form nominal attributes (e.g. *mii pa-khat* 'person classifier-one = one person').

Some classifiers referring to special semantic fields include *muu-* ‘granular substances’, *thluan-* ‘elongated items’, *tlaap-* ‘flat items’, *pum-* ‘round or oblong items’, *dor-* ‘drops of liquid’, *zuun-* ‘clothing’, *fan-* ‘units of money’, container classifiers (e.g. *kheen-* ‘plate’, *dur-* ‘small container’), and group classifiers (e.g. *buu-* ‘animal group’, *tua?*- ‘paired items’). Otherwise, classifiers may simply be a copy of the head noun itself. There is furthermore a default classifier *pa-*, which also occurs as part of the citation form of cardinal numbers.

3.2 Verbal inflection

The verbal complex consists of a (possibly derived) verb stem preceded by up to three prefixal or proclitic elements and often followed by several postverbal elements.

3.2.1 Ablaut

Sentence level morphosyntax is dominated first and foremost by a system of verbal ablaut and concomitant alternations in the case marking of nominals. Most verbal roots have two allomorphs, one phonologically largely predictable from the other (the various alternations involve different tonal properties, presence/absence or manner of articulation for a final consonant, vowel length/quality, and presence/absence of glottalization). In affirmative, indicative, main clauses, form 1 occurs if the verb is intransitive (1), and form 2 and an ergative case-marking pattern occur if the verb is transitive (2). However, there are also notionally transitive clauses in which form 1 of the verb is used (3), with accompanying caseless NPs, and a less referential/identifiable P.

- (1) *paalaw ?a-thii*
Paalaw 3sA/S-die,
‘Paalaw died.’
- (2) *paalaw=ni? thil (khaa) ?a-ba?*
Paalaw=ERG clothes DEIC 3sA/S-hang.up,
‘Paalaw hung up the clothes.’
- (3) *paalaw (khaa) thil ?a-bat*
Paalaw DEIC clothes 3sA/S-hang.up,
‘Paalaw hangs up/hung up clothes.’

The difference between the ergative construction and the alternative construction seen in (3) is a subtle one, and, like the comparable alternation in closely related Laizo, it is intimately connected to information structure. It has been pointed out by Kathol and VanBik (2001) that the construction in (3) bears considerable resemblance to an antipassive construction in terms of its information structure properties. Moreover, it turns out that while discourse deictics, which are themselves closely tied to information structure, may readily occur with the absolute argument in (2), they may not be associated with the object argument in (3). However, it should be clear that there are many respects in which (3) could not be considered to be a prototypical antipassive construction. In particular, the object argument is not ommissible, and while in some sense it may be syntactically more inert than the object of a monotransitive, it is not overtly marked as an oblique.

Other morphosyntactic contexts ‘override’ (Kathol and VanBik 2001) this basic system, and require either the form 1 or the form 2 ablaut grade. For instance, regardless of case marking, the polar interrogative marker (4) and the negative marker (5) require form 1 of the verb.

- (4) *paalaw(=ni?) thil ?a-bat=moo*
 Paalaw(=ERG) clothes 3sA/S-hang.up₁=INTERR
 ‘Did Paalaw hang up the clothes?’
- (5) *paalaw(=ni?) thil ?a-bat-law*
 Paalaw(=ERG) clothes 3sA/S-hang.up₁-NEG
 ‘Paalaw did not hang up the clothes.’

On the other hand, subordinate clauses usually require the form 2 grade, even if the subordinate clause is negated (6):

- (6) *ka-paa=ni? tsaʔuk ?a-haʔw-law tik=?a? ka-nuu=ni? ?a-zuar*
 1sA/S-father=ERG book 3sA/S-need₂-NEG time=LOC 1sA/S-mother=ERG 3sA/S-sell₂
 ‘When my father did not need the book, my mother sold it.’

3.2.2 Agreement

Finite verbs are accompanied by a sequence of one or two agreement prefixes, s/A preceding P, and sometimes one suffix (indexing P), which exhibit a nominative–accusative alignment. Table 12.3 shows these elements. A refers to the agent argument associated with the prototypical transitive verb, s refers to the single argument associated with intransitive verbs, and P refers to the patient argument associated with prototypical transitive verbs.

The A/s markers for singulars are straightforward, as are the A/s markers for plurals, since the latter are simply a combination of the former and a plural element *-n-*. The markers for P in the first person are the same as the markers for A and s. Third person P is zero-marked, but in the third plural, a postverbal element *-hnaa* indicates plurality of the object. Second person P agreement exhibits allomorphy between *?in-*, which occurs after a consonant-final (i.e. plural) A/s marker, and *ń-*, which occurs following a vowel-final (i.e. singular) A/s marker. The latter allomorph involves a high tone realized on the nasal portion of the A-P combination. Again, as in the third plural, plurality of the object in the case of second person plural objects is indicated by the postverbal element *-hnaa*. If P is coreferential with A, there are special object prefixes, which can be given either a reflexive or a reciprocal interpretation (seen at right in the table). There is no distinction for person in these forms.

There are special subject agreement forms in the jussive mood (cohortative – ‘let first person v’, imperative, and exhortative – ‘let third person v’), as seen in Table 12.4.

TABLE 12.3 VERBAL AGREEMENT MARKERS

	A/s	P	Reflexive object A _i P _i
1s	ka-	-ka-	-a-
2s	na-	-ń~?in-	-a-
3s	?a-	-ø-	-a-
1p	ka-n-	-ka-n-	-?ii-
2p	na-n-	-ń... -hnaa~?in... -hnaa	-?ii-
3p	?a-n-	-ø... -hnaa	-?ii-

TABLE 12.4 JUSSIVE AGREEMENT AND NEGATION

	Singular	Dual	Plural
cohortative			
1 affirmative	-nij	-?u-si?	-hnaa-?u-si?
negative	-hla?-nij	-hla?-?u-si?	-hnaa-hla?-?u-si?
imperative			
2 affirmative	-ø		-?u-ø
negative	-hla?-ø		-hla?-?u?-ø
exhortative			
3 affirmative	-se?		-hnaa-se?
negative	-hla?-se?		-hnaa-hla?-se?

3.2.3 *Directionals*

Verbal complexes may contain one of a class of directional markers occurring between the subject and object agreement markers, though the semantics of a number of these reportedly has become opaque for younger speakers. In form, the directional *rak-* resembles the verb *raa~rat* ‘come’, and in semantic terms, it marks a venitive (motion towards a deictic centre). *rak-* has also grammaticalized as an indicator of past tense. The directional *va-* is an andative (motion away from a deictic centre):

- (7) *tsakay-pool ka-va-kaap-hnaa-laay tia? ?a-tii*
 tiger-COLL 1sA/S-andat-shoot₁-PL.P-IRR QUOT 3sA/S-say_i
 “I’m going to go and shoot tigers!” he said.”

Another directional particle, *hay-*, is an andative like *va-*, but involves motion directed over a shorter distance. *von-* indicates that the action of the verb is performed suddenly and in the immediate vicinity. *ruj-* marks motion from a point above the speaker towards the speaker. *huj-* indicates motion upwards towards the speaker and *vuj-* motion downwards away from the speaker, respectively. At this point, however, the latter two particles are not used consistently.

3.2.4 *Other inflection in the verb complex*

Verb roots may be followed by sequences of up to several bound particles which distinguish modal, aspectual/aktionsart, tense, mood, and various adverbial categories.

3.2.4.1 Modality

There are several modal elements, which cut across categories in terms of their morphosyntax, ranging from more to less bound elements. First, there are elements which are bound but show the ablaut alternation characteristic of independent verb stems, such as the potential marker *khaw~kho?*. There are also modal elements, such as the desiderative *du?*, the potential *thiam*, the permissive causative *sian~siay*, and obligative *haaw*, which are probably best regarded as independent verbs taking bare verb complements. These modal elements themselves bear agreement morphology, usually to the exclusion of the complement verb. In no case is it actually impossible for these to occur as non-agreeing, bound postverbal elements, however, though some speakers show a preference for one or the other construction type.

3.2.4.2 Aspect/aktionsart

Lai makes a large number of subtle aspectual/aktionsart distinctions. Some of the more basic, high-frequency aspectual markers include *-liaw* (progressive), *-tsay* (perfect), and *-laay* (irrealis, which in its most basic use marks future tense, but which combines with other tense and aspect particles to provide a variety of epistemic modal and subtle aspectual senses). A number of aspectual distinctions have to do mainly with future events: immediate prospective (*-hnik*), neutral prospective (*-deey(maay)*); and a number focus on the event's imperfectivity: continuous activity involving effort (*-leen*), continuous but ineffective activity (*-seek*), neutral continuous activity (*-pey*), and negative neutral continuous/superfluous activity (*-hley*). Temporary activities are marked by *-taa* and *-tshuy*. Other categories include habitual (*-toon*), perseverative (*-ri?*), experiential perfect (*-bal*), instantaneous (*-tso?l* – the action occurs instantaneously), instantaneous unexpected (*-duak*), iterative (*-leymay*), permanent (*-be?*), exhaustive (*-di?*, and for older speakers, also *-thluu* – all of an absolute entity is effected by the action), repetitive (*-thaan* – the action is performed again; *-hoy* – the action is unfortunately performed again), additional (*-vee* – another subject performs the action), associative (*-tii* – the action is performed jointly by a plural subject), accidental (*-sual*), and unpremeditated (*-tshom*). Note that, as Bedell (2012) suggests, some of these elements, and others discussed below, may in fact be independent adverbial particles rather than suffixes or clitics, pending more explicit argumentation to support their morphologically bound status.

3.2.4.3 Tense

There is a basic tense distinction between future (marked by the irrealis marker *-laay*) and non-future events (unmarked). In addition, the directional prefix *rak-* has developed a past tense sense which may be used to explicitly mark past tense, especially in conjunction with various aspectual markers.

3.2.4.4 Marking of epistemic modality, evidentiality, and subjective evaluation

There are a few markers which indicate the speaker's evaluation of the accuracy of the proposition, or an emotional response of the speaker towards the content of the proposition. *-kaw* indicates the speaker's certainty, or at least assumed certainty, in the accuracy of the proposition. *-ruaa*, on the contrary, indicates that the speaker has no direct knowledge of the accuracy of the proposition. The use of *-tuj* in the verbal complex implies that the content of the proposition is counter to the expectations of the speaker. Additionally, there is a set of postverbal elements clearly related to demonstratives/discourse deictics (*hi?*, *khi?*, *tsu?*, *kha?*), which also convey quite intricate spatial and evidential information. *?aay* indicates regret on the part of speaker or subject and *?ee* generally indicates excitement on the part of the speaker *vis-à-vis* the content of the proposition.

3.2.4.5 Ideophonic elements

There is a virtually open class of postverbal particles which are comparable in function to what are variously dubbed ideophones, mimetics, or expressives. These conform to a couple of different prosodic templates, and reportedly add to the vividness of the picture which a clause describes in highly specific ways. For instance, in (8):

- (8) *Puytsaw ?a-baw-duapmap/?a-baw-diapmap*
dog 3sA/S-bark₁-IDEO/3sS-bark₁-IDEO
'The dog barked (big, bellowing dog)/(small, yapping dog).'

The first ideophonic element creates the impression that the dog is large, with an appropriate bark, whereas the second ideophonic element conveys the picture of a little dog and its corresponding bark. The syntactic distribution of ideophonic elements is more complex than the usual postverbal particle. Though they usually occur sandwiched between the verb root and a number of the tense and aspectual particles, in some cases they may occur as nominal modifiers.

3.2.4.6 Comparison

Comparative and superlative constructions require the use of the particles *-deʔw* and *-biik*, respectively, as seen in (9) and (10).

- (9) *paalaw ɻa-nuu=naak?in ɻa-saaŋ-deʔw*
Paalaw 3s.POSS-mother=STAND 3sA/S-tall₁-COMP
'Paalaw is taller than his mother.'
- (10) *paalaw ɻa-saaŋ-biik*
Paalaw 3sA/S-tall₁-SUPERL
'Paalaw is tallest.'

3.2.4.7 Negation

The negative marker in indicative clauses is *-law*. In the jussive mood (and actually in non-finite clauses generally), as we saw in Table 12.4, negation is marked instead by *-hla?*.

3.2.4.8 Mood

Indicative mood is morphologically unmarked. We saw, in section 3.2.2, that the primary exponents of the various jussive categories are a separate set of subject person/number markers, in addition to a distinct negative marker; other markings remain the same in jussives. Finally, there is a marker *-hyaa*, which marks the apodosis of (past) counterfactual conditional clauses, in addition to occurring in a disparate range of epistemic modal contexts (a typical example is in (11)).

- (11) *ka-nuu sin=ɻa? khan ka-rak-thlun-vee-kaw-hyaa-law*
1s.POSS-mother vicinity=LOC DEIC 1sA/S-PAST-return₂-ADD-AFF-SUBJ-NEG
'Oh, why did I not return with my mother?'

3.3 Derivational morphology

3.3.1 Compounding

Compounding is highly productive: e.g. *may-thal* 'fire-bow = gun', *si-vaay* 'medicine-wander = poison', *thut-dan* 'sit-separate/filter = seat'. Often the resulting compound is non-compositional.

3.3.2 Nominal derivation

-naak, which has numerous etymologically related elements throughout the morphology, functions as a deverbal nominalizer which productively creates locative, instrumental,

and action nominals: *sam-meʔ-naak* ‘hair-cut-nomlzs = barbershop’, *thil-tsook-naak* ‘thing-buy-nomlzs = shop’, *?ay-din-naak* ‘eat-drink-nomlzs = restaurant’, *hmuʔ-naak* ‘see-nomlzs = seeing’, *peek-naak* ‘give-nomlzs = giving’. Besides *-naak*, there are two elements which derive nouns from noun bases: *-pii* (augmentative) and *-tee* (diminutive).

3.3.3 Ordinal numbers

Ordinal numbers are derived from numeral roots by means of the suffix *-naak*: *pa-hni?* ‘two’, *hniʔ-naak* ‘second’, *pa-hli* ‘four’, *hli-naak* ‘fourth’, etc.

3.3.4 Valence-affecting morphology

Besides the effect that ablaut may have on verbal valence, the verbal complex has additional resources for affecting valence. This morphology is always adjacent to or is lexicalized within the verbal root.

3.3.5 Middle

The reflexive/reciprocal prefix (*?ii-* and its allomorphs) also has semi-productive derivational properties, producing what is essentially a middle voice.

3.3.6 Causatives and applicatives

Most valence-affecting morphology involves transitivization. There are two levels of derivation, one older and restricted in productivity, and the other of more recent origin and highly productive. The older system produces direct causatives. First, there are a few items showing a causative in *-sak* (e.g. *hmuʔ-sak* ‘to show’). Next, a number of causatives involve non-causative/causative pairs (e.g. *tlaak* ‘to fall’, *thlaak* ‘to fell’) marking the causative member by aspiration (in the case of stop/affricate-initials) or voicelessness (in the case of sonorant-initials). These perhaps reflect the widely recognized Tibeto-Burman **s-* causative prefix. These **s*-causatives are restricted to occurrence with non-stative intransitive roots. Causativization of some stative intransitive roots, on the other hand, involves a glottal feature which is realized either as glottalization of a final sonorant (and in the case of some roots, a change in place of the final consonant from *ŋ* to *n*), debuccalization of a stop consonant (neutralization to *?*), or by addition of a final glottal stop to vowel-final roots. When the latter element occurs in conjunction with non-stative or transitive bases, the result is a dative/goal or benefactive applicative verb stem. These features apparently reflect the PTB **-t* transitivizing suffix, as suggested for Kuki-Chin already by Benedict (1972: 101–2).

The newer system of causativization, on the other hand, effected by the addition of an element *-ter*, is quite productive; the forms it produces may be interpreted as involving indirect, as well as direct, causation. The system of applicatives, which is likewise highly productive (even with intransitive roots), involves the addition of one of seven postverbal elements to the verbal complex, depending on the semantics of the applicative object. Here, (12) to (18) give examples of the applicative morphology.

- (12) benefactive/malefactive applicative
law ?a-ka-thloʔ-piak
 field 3sA/S-1sP-hoe₂-BEN.APP
 ‘He hoed the field for me/in my place.’

- (13) additional benefactive applicative
law *?a-ka-thloʔ-tseʔm*
 field 3sA/S-1sP-hoe₂-ADD.BEN.APP
 ‘He hoed the field for my benefit (in addition to his own benefit).’
- (14) comitative applicative
law *?a-ka-thloʔ-pii*
 field 3sA/S-1sP-hoe₂-COM.APP
 ‘He hoed the field along with me.’
- (15) allative/malefactive applicative
law *?a-ka-thloʔ-hno?*
 field 3sA/S-1sP-hoe₂-MAL.APP
 ‘He hoed the field to my deriment.’
- (16) priorititive applicative
law *?a-ka-thloʔ-kaʔn*
 field 3sA/S-1sP-hoe₂-PRIOR.APP
 ‘He hoed the field ahead of/before me.’
- (17) relinquitive applicative
law *?a-ka-thloʔ-taak*
 field 3sA/S-1sP-hoe₂-RELINQ.APP
 ‘He left me and hoed the field.’
- (18) instrumental applicative
tuuhmuy *law* *?a-thloʔ-naak*
 hoe field 3sA/S-hoe₂-INST APP
 ‘He hoed the field with a hoe.’

Only in the case of *-hno?*, which is generally malefactive, but which may have an allative meaning with some roots expressing motion (e.g. *kal* ‘go’ and *kaay* ‘climb’), is there much deviation from the semantics indicated by these examples. In each of these constructions, the applicative object has more ready access to a number of object properties than the object of the base verb (*law* ‘field’) does (e.g. accessibility to topicalization and object agreement, potential to be associated with discourse deictics, ability to control zero anaphora in certain types of clause chaining), and the narrative text-based study of applicative discourse-function in Peterson (2007) shows that for many applicative constructions in Lai, using a variety of metrics, the applicative object is more topical than a co-occurring base object. In cases where there is an oblique alternative instantiation of an applicative object (e.g. with instruments, which can otherwise be expressed with a case clitic), the non-applicative, obliquely expressed object does not have these properties.

4 SYNTAX

4.1 NP syntax

Elements in nominal phrases exhibit the following basic order:

[demonstrative] [relative] [possessor] <head> [classifier-numeral] [quantifier] [case]
 [discourse deictic]

Although headless noun phrases may occur (e.g. headless relatives), the central component is typically a head noun, which may either be possessed by a preceding noun, or bear a possessive prefix. Quantifiers follow the head noun, and numeral quantification requires a classifier element compounded with a numeral root. Next, all non-absolutive noun phrases bear a clitic case particle. The final element in a nominal phrase is often a discourse deictic (marker of information status); purely spatial deixis always involves a demonstrative element at the beginning of the noun phrase. Relative clauses, including expressions corresponding to adjectives in other languages, generally precede their heads, though the head may also occur internal to the relative clause, giving the impression in some cases of a postnominal relative; adjectival roots may also occur in a distinct, non-finite construction following the nominal they modify.

Discourse deictics show limited agreement in case: if the case of the phrase is anything other than absolute, the deictic bears the agreement marker *-n*. This behaviour might seem somewhat anomalous for this type of particle, but presumably it simply derives from these particles' earlier status as demonstratives. If a noun marked by the instrumental-ablative case clitic is followed by a discourse deictic, often the case clitic is omitted, and simply understood from the oblique marking of the accompanying discourse deictic.

4.2 Clausal syntax

4.2.1 Word order

Arguments are often not instantiated by full NPs, their instantiation being assumed by verbal participant marking. Word order is usually verb-final, with variable placement of any full NPs depending on pragmatic factors. There is a left dislocation position, which attracts heavy constituents and is also used extensively for topicalization. Constituents also may appear in a postverbal afterthought position.

4.2.2 Grammatical relations

Marking of the basic S, A and P functions has already been discussed in passing in the sections on case marking particles (3.1.4), verbal ablaut (3.2.1) and verb agreement (3.2.2). In short, the basic system of grammatical relations as marked on nominals has an ergative/absolute alignment, and verbal ablaut allows for the effect of an antipassive construction, in which both arguments of a transitive verb are absolute, but in which the object does not have access to all of the properties generally available to transitive objects (in particular, the object is restricted in its ability to co-occur with discourse deictics). On the other hand, verbal agreement has a largely nominative/accusative alignment.

Multiple objects exhibit a primary object alignment; with a ditransitive verb, the recipient is marked on the verb rather than the patient. This tendency apparently stems largely from the tendency to mark animates to the exclusion of inanimates.

Finally, it has been noted by Bickel (2000) that Lai agreement, as elsewhere in Sino-Tibetan, is not always of the canonical 'identificational' type, but may instead be partitional, appositional or relational, the latter seen particularly in psycho-collocational constructions like the one in (19), in which agreement is not with the third singular 'my heart', but rather with the entity to which the expressed emotion relates.

- (19) *ka-lunj kan-rook*
1s.POSS-heart 1pA/S-break.down,
'I am disappointed at us.'

4.2.3 Coordination

Coordination of NPs involves the clitic =*lee*, which also sometimes marks clausal conjunction. The most frequent clausal coordination construction makes use of an encliticized particle =*?ii*, which may otherwise appear clause-initially as an independent conjunction; prosodic inclusion in the first clause distinguishes the enclitic use from its clause-initial use. A full treatment of Lai coordination may be found in Peterson and VanBik (2004).

4.2.4 Subordination

Subordination includes three basic types of construction: adverbial subordination, relativization and complementation.

4.2.5 Adverbial subordination

Finite indicative clauses may, without any formal modification, occur as a complement to the oblique case particles =*?a?* and =*?in* to yield adverbial subordinate clauses expressing the circumstance under which an event described in the main clause occurs, as in (20).

- (20) *ka-tiin=?a?* (*khan*) *?akal*
 1sA/S-come.home₂=LOC DEIC 3sA/S-go₁
 ‘When I came back home, he left.’

Related to this subordination pattern is the marking of the protases of conditional clauses, seen in (21), which involves addition of the locative case particle and the discourse deictic *tsu-n*, which tends to mark topicalization, to the corresponding indicative clause:

- (21) *ka-tiin=?a?* *tsun ?a-kal-laay*
 1sA/S-come.home₂=LOC DEIC 3sA/S-go₁-IRR
 ‘If I come home, he will leave.’

In addition, there is a class of constructions consisting of relational nouns with more or less concrete semantics which are marked obliquely (by the locative or instrumental-ablative case clitic) and inserted after zero-nominalized clauses to indicate a variety of adverbial relations. In some cases the relational elements involved are used independently as relational nouns (e.g. *v hnuu=?a?* ‘back (=after) v’, *v hlaan=?a?* ‘front (=before) v’, *v tshuy=?a?* ‘inside (=while) v’, *v laay=?a?* ‘towards (=when about to) v’, *v tsaa=?a?* ‘sake (=because) v’). In other cases, the relational elements only occur in these constructions (e.g. *v tik=?a?* ‘when v’).

Besides these relatively compositional subordination strategies, there are some less compositional subordinators, though parts of them are identifiable. Concessive clauses are marked by *-naa=?in* and purposive clauses are marked by *-diy=?a?*. Besides the *v tsaa=?a?* construction mentioned above, reason clauses may also be indicated by *-koo*; *-koo* reason clauses, as opposed to most other adverbial subordinate clause forms, are usually non-finite: they take no subject agreement markers and they have a negative marker like that found in the jussive mood. *-buu=?in* and *-pa=?in* mark a subordinate action which is simultaneous with that of the main clause, but unlike the other subordinate clause types, their subject must also be identical to that of the main clause. Lastly,

a more complex purposive clause is marked by a combination of a number of otherwise straightforward elements: *-naak-tsaa-diy=ʔa?*

4.2.6 Relativization

Relative clauses may be externally or internally headed, though both strategies are not available for all target types. Externally headed relativization is illustrated in (22) to (27). Externally headed relatives are preposed finite clauses with a gap corresponding to the target of relativization.

(22) s target:

<i>in=ʔii</i>	<i>?a-it-mii</i>	<i>lawthlawpaa</i>	<i>ka-hmu?</i>
house=LOC	3sA/S-sleep ₁ -REL	farmer	1sA/S-see ₂
'I saw the farmer who slept in the house.'			

(23) A target with *-tuu*:

<i>thil</i>	<i>?a-bat-tuu</i>	<i>lawthlawpaa</i>	<i>ka-hmu?</i>
thing	3sA/S-hang ₁ -REL	farmer	1sA/S-see ₂
'I saw the farmer who hung up the clothes (lit. the things).'			

(24) A target with *-mii*:

<i>thil</i>	<i>?a-bat-mii</i>	<i>lawthlawpaa</i>	<i>ka-hmu?</i>
thing	3sA/S-hang ₁ -REL	farmer	1sA/S-see ₂
'I saw the farmer who hung up the clothes.'			

(25) P target:

<i>lawthlawpaa=ni?</i>	<i>?a-baʔ-mii</i>	<i>thil</i>	<i>ka-hmu?</i>
farmer=ERG	3sA/S-hang ₂ -REL	thing	1sA/S-see ₂
'I saw the clothes the farmer hung up.'			

(26) Locative target:

<i>lawthlawpaa=ni?</i>	<i>thil</i>	<i>?a-baʔ-naak</i>	<i>thiijkuy</i>	<i>ka-hmu?</i>
farmer=ERG	thing	3sA/S-hang ₂ -REL	tree	1sA/S-see ₂
'I saw the tree the farmer hung the clothes up on.'				

(27) Instrumental target:

<i>lawthlawpaa=ni?</i>	<i>ŋa</i>	<i>?a-tan-naak</i>	<i>naam</i>	<i>ka-hmu?</i>
farmer=ERG	fish	3sA/S-cut ₂ -REL	knife	1sA/S-see ₂
'I saw the knife the farmer cut the fish with.'				

Relativization on given targets requires a particular ablaut grade and an invariant relative clause particle. Table 12.5 summarizes the morphological devices involved for different target types.

There is no clear difference between the *-tuu* and *-mii* relativizers except in terms of their potential relativization targets, but *-tuu* has a much lower text frequency.

For certain targets, relative clause heads may occur internally as well as externally. This is shown in (28)–(30) for s, A, and P targets.

TABLE 12.5 PRIMARY RELATIVIZATION STRATEGIES

Role of target	S	A	P	Locative, Instrumental
Ablaut grade	form 1	form 1	form 2	form 2
Relativizer	-mii	-mii/-tuu	-mii	-naak

- (28) *?in=?ii lawthlawpaa ?a-it-mii ka-hmu?*
 house=LOC farmer 3sA/S-sleep₁-REL 1sA/S-see₂
 'I saw the farmer who slept in the house.'
- (29) *nikum=?ii lawthlawpaa thil ?a-bat-mii ka-hmu?*
 last.year=LOC farmer thing 3sA/S-hang₁-REL 1sA/S-see₂
 'I saw the farmer who hung up the clothes last year.'
- (30) *nikum=?ii lawthlawpaa=ni? thil ?a-baʔ-mii ka-hmu?*
 last.year=LOC farmer=ERG thing 3sA/S-hang₂-REL 1sA/S-see₂
 'I saw the clothes the farmer hung up last year.'

It is important to note that the relative clause of all of these examples includes an adverbial, which is marked with the oblique case clitic =?ii rather than =?a?. As discussed earlier in the section on case marking, =?ii, occurs only in subordinate clauses, so the presence of this adverbial in each sentence unambiguously indicates that the target of relativization is syntactically internal to the (subordinate) relative clause in question. Internally headed relative clauses are not possible with the -tuu relativizer. For instruments, the internal relative clause head is not marked obliquely; if such marking occurs, the relativization is interpreted as targeting a locative. Internally headed relatives targeting locatives may have an obliquely marked head, but such sentences are ambiguous between a locative target and an instrument target.

4.2.7 Complementation

As mentioned in the discussion of modal elements, it appears that the best analysis for some modal elements is as modal auxiliaries which take bare verb stem complements. However, there are also a number of finite complement types marked by specialized complementizers. The most general complementizer is identical to one of the relativizers, -mii. -mii complements appear with a number of verbs of cognition, as in example (31).

- (31) *paalaw =ni? ?a-tsoo ?a-zuar-mii khaa ?a-thaʔy*
 Paalaw = ERG 3s.POSS-cow 3sA/S-sell₂-COMP DEIC 3sA/S-know₂
 'He knows that Paalaw sold his cow.'

There is no distinction between direct and indirect speech. Verbs of speaking and other verbs of cognition require a (quotative) complementizer probably arising from the proverb -tii 'do/say', *tia?*, as in (32).

- (32) *paalaw=ni? ?a-tsoo ?a-zuar tia? ?a-tii/?a-zuʔm*
 Paalaw=ERG 3s.POSS-cow 3sA/S-sell₂ QUOT 3sA/S-say₁/3sA/S-believe₂
 'He said/believes Paalaw sold his cow.'

Finally, *dij=?a?*, which acts as a purposive clause marker in other contexts, functions as a complementizer:

- (33) *lawthlawpaa=ni? ?a-tsoo ?a-zuar dij=?a? paalaw=ni?*
 farmer=ERG 3s.POSS-cow 3sA/S-sell₂ comp=LOC Paalaw=ERG
?a-leem/?a-nool
3sA/S-persuade₂/3sA/S-request₂
 'Paalaw persuaded the farmer to sell his cow/requested that the farmer sell his cow.'

4.3 Major sentence types

4.3.1 Indicative

Indicative sentences are not coded by any special marking.

4.3.2 Copular

Equational copular sentences are formed with a predicate nominal and an appropriately conjugated form of the copular verb *sii*. Existential copular sentences use a separate copular predicate, *?um*.

4.3.3 Jussive

As discussed in conjunction with verbal inflection, jussive sentences involve special subject person and number, as well as negative markers. Otherwise, these sentences do not differ materially from indicative sentences in their syntax. Another common way to form imperatives is simply to postpose the particle *=tua?* to a form 1 verb stem. *=loo* marks imperatives and cohortatives as more polite.

4.3.4 Interrogative

Polar interrogatives are indicated by the sentence-final particle *=moo* (for some speakers the particle *=maa* is used), as in (34).

- (34) *na-min thooytseew ?a-sii=maa*
2s-name Thawng Ceu 3sA/S-be₁=INTERR
'Is your name Thawng Ceu?'

Content questions are characterized by (optionally utterance-initial) dedicated question words (*zay* 'what', *?ahaw* 'who', *khoj/khoykaa=?a?* 'where', *zay tii=?in* 'how', *zay ruay=?a?* 'for what reason', etc.). In most cases, the particle *=da?* is also added to the question word, as in (35).

- (35) *faalaam=?a? ?ahaw=da? na-tha?y*
Falam=LOC who=QUES 2sA/S-know₂
'Who do you know in Falam?'

FURTHER READING

As an important language of wider communication in the Chin Hills, Lai was the subject of a pair of grammatical descriptions made at the end of the nineteenth century by British military personnel (Macnabb 1891; Newland 1897); a somewhat later description is Haye-Neave (1948).

While a large number of Tibeto-Burman languages have been given more or less satisfactory grammatical descriptions in the last few decades, few languages have been subject to the kind of work which has been carried out on Lai. In the last decade, no less than three groups of researchers have conducted investigations on Lai grammar: Bedell in Japan, Lehman in Illinois, and several graduate students at the University of California, Berkeley under the guidance of Matisoff have all produced independent studies on a wide variety of grammatical topics (see Bibliography). In terms of lexical resources, Kenneth

and David VanBik had been preparing a Lai–English dictionary to complement David’s English–Lai dictionary; since David’s untimely death in 2000, Ken has continued work on this project and hopes to make it available online in the near future.

NOTE

- 1 Many thanks to Ken VanBik for commenting on the original version of this sketch and for providing and discussing Lai data on demand for it, and for this slight revision. Almost all of the insights in this chapter are a result of Ken’s extensive work and collaborations with others, including the present author. Thank you also to Bernard Comrie and to the editors for helpful criticisms and suggestions on the original and present versions.

BIBLIOGRAPHY

- Barnes, Jonathan. 1998. Tsuu khaa tii hla?: deixis, demonstratives and discourse particles in Lai Chin. *LTBA* 21.53–86.
- Bedell, George. 1995. Agreement in Lai. *Papers from the Fifth Annual Meeting of the Southeast Asian Linguistics Society*, Tempe, Arizona: Program for Southeast Asian Studies, Arizona State University, 21–32.
- . 1996. Clitic climbing in Lai. *Pan-Asiatic Linguistics, Proceedings of the Fourth International Symposium on Languages and Linguistics*, Vol. 2, Nakhon Pathom: Mahidol University, 405–415.
- . 1997a. Causatives and clause union in Lai (Chin). *Mon-Khmer Studies* 27.219–232.
- . 1997b. Benefactives and clause union in Lai. *Papers from the Seventh Annual Meeting of the Southeast Asian Linguistics Society*. Tempe, Arizona: Program for Southeast Asian Studies, Arizona State University, 21–34.
- . 1999. Word combination in Lai. *Papers from the Ninth Annual Meeting of the Southeast Asian Linguistics Society*. Tempe, Arizona: Program for Southeast Asian Studies, Arizona State University, 201–219.
- . 2001a. The syntax of deixis in Lai. *LTBA* 24.157–171.
- . 2001b. Passives and clefts in Lai. *Papers from the Sixth Annual Meeting of the Southeast Asian Linguistics Society*. Tempe, Arizona: Program for Southeast Asian Studies, Arizona State University, 15–31.
- . 2007a. Nominal auxiliaries in Lai. *Papers from the Eighth Annual Meeting of the Southeast Asian Linguistics Society*. Canberra: Research School of Pacific and Asian Studies, Australian National University, 11–24.
- . 2007b. Genitive constructions in Lai. *Papers from the Twelfth Annual Meeting of the Southeast Asian Linguistics Society*, 2002. Canberra: Pacific Linguistics, 9–21.
- . 2007c. Negation in Lai. *Papers from the Thirteenth Annual Meeting of the Southeast Asian Linguistics Society*, 2003. Canberra: Pacific Linguistics, 9–21.
- . 2010. Postpositions and relational nouns in Lai. *JSEALS* 3.1–21.
- . 2012. Tense and aspect in Lai. *Language in India* 12.290–303.
- Bedell, George and Kenneth VanBik. 2002. Lexical and syntactic causatives in Lai. *Papers from the Tenth Annual Meeting of the Southeast Asian Linguistics Society*. Tempe: Arizona: Program for Southeast Asian Studies, Arizona State University, 75–86.
- Benedict, Paul. 1972. *Sino-Tibetan: A Conspectus*. Cambridge: Cambridge University Press.
- Bickel, Balthasar. 2000. On the syntax of agreement in Tibeto-Burman. *Studies in Language* 24.583–609.
- Chhangte, Lalnunthangi. 1993. Mizo syntax. University of Oregon PhD dissertation.

- Chinzah, Lalremzami and George Bedell. 2012. Adjectives in Lai. Paper presented at the Thirty-fourth All India Conference of Linguists, Shillong.
- Haye-Neave, D.R. 1948. *Lai Chin Grammar and Dictionary*. Rangoon: Superintendent of Government Printing and Stationery, Burma.
- Hyman, Larry M. and Kenneth VanBik. 2002a. Tone and syllable structure in Hakha Lai. *BLS 28, Special Session on Tibeto-Burman and Southeast Asian Linguistics*: 15–28.
- 2002b. Tone and stem2-formation in Hakha Lai. *LTBA* 25.113–120.
- 2004. Directional rule application and output problems in Hakha Lai tone. *Language and Linguistics* 5.821–861.
- Kathol, Andreas. 2001. The morphosyntax of Lai relative clauses. In Ronnie Cann, Claire Grover and Phillip Miller (eds) *Grammatical Interfaces in HPSG*, Stanford, CA: CSLI Publications, 137–152.
- Kathol, Andreas and Kenneth VanBik. 1999. Morphology – syntax interface in Lai relative clauses. In Pius Tamanji, Masako Hirotani and Nancy Hall (eds) *Proceedings of the 29th Annual Meeting of the North Eastern Linguistic Society, University of Massachusetts, Amherst*: GLSA, 427–441.
- Kathol, Andreas and Kenneth VanBik. 2001. The syntax of verbal alternations in Lai. Paper presented at the Linguistic Society of America Annual Meeting, Washington DC.
- Kavitskaya, Darya. 1997. Tense and aspect in Lai Chin. *LTBA* 20.173–213.
- King, Deborah. 2010. Voice and valence-altering operations in Falam Chin: a role and reference grammar approach. UT Arlington PhD dissertation.
- Lehman, F.K. 1963. *The Structure of Chin Society: A Tribal People of Burma Adapted to a Non-Western Civilization*, Illinois Studies in Anthropology, No. 3. Urbana: University of Illinois Press.
- 1996. Relative clauses in Lai Chin, with special reference to verb stem alternation and the extension of control theory. *LTBA* 19.43–58.
- Lehman, F.K. and Kenneth VanBik. 1997. Notes on Lai Chin personal pronouns and overt case marking. *Studies in the Linguistic Sciences* 27.81–86.
- Macnabb, D.J.C. 1891. *Hand-Book of the Haka or Baungshe Dialect of the Chin Language*. Burma: Superintendent, Government Printing.
- Melnik, Nurit. 1997a. The sound system of Lai. *LTBA* 20.9–19.
- 1997b. Verbal alternations in Lai. *LTBA* 20.163–172.
- Newland, A.G.E. 1897. *A Practical Hand-Book of the Language of the Lais*. Burma: Superintendent, Government Printing.
- Osborne, Andrea Gail. 1975. A transformational analysis of tone in the verb system of Zahao (Laizo) Chin. Cornell University PhD dissertation.
- Patent, Jason D. 1998. A willy-nilly look at Lai ideophones. *LTBA* 21.155–200.
- Peterson, David A. 1998. The morphosyntax of transitivization in Lai (Haka Chin). *LTBA* 21.87–153.
- 2006. High and low applicatives: evidence from Lai. *BLS* 30.353–364.
- 2007. *Applicative Constructions*. Oxford: Oxford University Press.
- Peterson, David A. and Kenneth VanBik. 2004. Coordination in Hakha Lai. In Martin Haspelmath (ed.) *Coordinating Constructions*, Amsterdam: John Benjamins, 333–356.
- Reichle, Verena. 1981. *Bawm Language and Lore: Tibeto-Burman Area*. Bern: Peter Lang.
- Roengpitya, Rungpat. 1997. Glottal stop and glottalization in Lai (connected speech). *LTBA* 20.21–56.

- VanBik, David. 1996. *English–Chin (Hakha) Dictionary*. Hakha: Chin Association of Christian Churches.
- VanBik, Kenneth. 1998. Lai psycho-collocation. *LTBA* 21.201–233.
- . 2002. Three types of causative constructions in Hakha Lai. *LTBA* 25.99–122.
- . 2009. *Proto-Kuki-Chin: A Reconstructed Ancestor of the Kuki-Chin Languages*. Berkeley: STEDT.
- Yamashita Smith, Tomoko. 1998. The middle voice in Lai. *LTBA* 21.1–52.

CHAPTER THIRTEEN

MONGSEN AO¹

Alexander R. Coupe

1 OVERVIEW OF AO DIALECTS

Ao [ISO 639-3 code: njo] is a Tibeto-Burman language spoken in the Mokokchung District of central Nagaland, north-east India, by a population of approximately 261,000.² The language's major dialects – Mongsen [mʊŋ¹¹sən³³] and Chungli [tʂʊŋ¹¹li¹¹] – are sufficiently divergent to border on mutual unintelligibility without extensive prior exposure. Approximately 90 Ao villages are found on the six mountain ranges of Mokokchung District. Of this number, 30 speak the Mongsen dialect as their first language, 45 speak Chungli as their mother tongue, and another seven villages are mixed, with both Mongsen- and Chungli-speaking populations typically living in separate wards organized according to clan membership.

In addition to the major dialects of Mongsen and Chungli, a third Ao variety known as Changki [tʂəŋ³³ki³³] is spoken in eight villages on three western ranges of Mokokchung District. Changki could be considered a sub-dialect of Mongsen, as it shares paradigmatic features such as the pronominal and case-marking systems, and it is close to being mutually intelligible with Mongsen. According to Mills (1926: 333), a number of sub-dialects of Chungli are additionally spoken in another five villages located on the eastern flank of the Ao territory. Language contact with neighbouring languages of Tuensang District (in particular, the Konyak languages Chang and Phom) is almost certainly responsible for these understudied varieties diverging from Chungli, because some villages contain wards speaking entirely different languages. With one notable exception,³ mixed populations in multilingual villages have probably resulted historically from annexations or large-scale kidnapping of women by more powerful villages encroaching on the traditional Ao territory from the east (see Coupe [2011b: 23–4] for further discussion, also Hutton [1987(1929): iii–iv] for observations relating to the annexation of neighbouring villages and land by the Chang during the colonial period).

Many Mongsen speakers are familiar with the prestige dialect Chungli, as it is traditionally the language of Christian proselytization and is now taught up to the level of an undergraduate major. Chungli is also currently the only dialect with an orthography and regularly published newspapers, which further promotes its standing as the prestige dialect and helps to spread its influence in central Nagaland. Its higher status derives from the fact that an American Baptist missionary first worked on the language of a Chungli-speaking village (Melong Yimchen) in the late nineteenth century, with the result that this village's variety became the standard written variety. Chungli-speaking evangelists then spread Christianity to other tribes in neighbouring Tuensang District via the Chungli Bible and other materials written in Chungli during the twentieth century, thereby familiarizing other speech communities of central Nagaland with this Ao dialect.⁴ For predominantly tribe-internal political reasons, Mongsen and Changki remain unwritten dialects.

Typically situated on the spines of ridges at altitudes of up to 1,570 metres, Mongsen-speaking villages are mostly found in the south and west of Mokokchung District on mountain ranges known locally as the Ongpangkong and Changkikong respectively (*kūŋ* = ‘range’), Changki-speaking villages are located on the Changkikong, Japukong and Tsürangkong ranges at the western boundary of Mokokchung District, Chungli-speaking villages predominate on the eastern range known as the Langpangkong (a few more are additionally located at the northern end of the Changkikong), and mixed Mongsen and Chungli villages are mostly found on an irregularly shaped range known as the Asetkong, situated between the Changkikong and the Langpangkong and partly bounded by streams (*āsāt* = ‘island’).

This present-day distribution suggests a historical westerly infiltration of the Ao across the ranges of the Indo-Burmese Arc, with Changki- and Mongsen-speaking groups probably forming the vanguard of this migration. Historical population movements are likely to have resulted from the intermittent pressure of waves of hostile tribes advancing from the south and the east, with each new pulse pushing older waves of settlers off their colonized lands.⁵ Stories of people known as Isangyongr, Molungr and Nokrangr being ousted by the Ao crop up in oral narratives of the tribe’s history and are briefly discussed by Mills (1926: 9–11), who proposes that they were early Konyak colonists. In the same passage, Mills reports that the power of the Ao was waning at the end of the nineteenth century when the tribe’s land was annexed by the British, and enduring evidence of this decline persists in the observation that speakers of other Tibeto-Burman languages of Nagaland (e.g. Chang, Sangtam, Lotha, Phom) now occupy villages with Ao names. Historical annexations logically account for the existence of some multilingual villages at the peripheries of the traditional Ao territory.

Mongsen and Chungli diverge slightly in their phonology, substantially in their morphology, and minimally if at all in their syntactic structure. Mongsen maintains a two-way contrast in the voice onset time of stops and affricates, while Chungli has no voice onset time contrasts whatsoever (Temsunungsang 2009: 12). Both dialects have three tones,⁶ and both superimpose intonation melodies over the lexical tones of individual syllables. Intonation serves to delineate phrasal and clausal boundaries, or to signal the continuity of a topic across a series of finite clauses.

The greatest cleavage between the dialects is to be found in the grammatical morphology, with non-cognate forms of high-frequency case-marking clitics and verbal affixes creating considerable difficulties in intelligibility, particularly for Chungli speakers with limited exposure to Mongsen. Language contact and convergence between Chungli and other Tibeto-Burman languages is primarily responsible for the differences attested between the two major dialects (Coupe 2011b). Mongsen appears to be the more conservative dialect when its relational morphology is compared with other Ao group languages of central Nagaland (see §2), and also on the basis of phonological reconstructions – the synchronic phonology of Mangmetong Village Mongsen, for example, diverges minimally from the reconstructed phonological inventory of Proto-Ao (see Bruhn 2014: 45–7), and various village varieties of Mongsen retain a voiceless sonorant series that has completely merged with the voiced series in the Chungli dialect.

Bi-dialectalism is an interesting feature of some Ao villages, in which one ward of a village speaks Mongsen and the other ward speaks Chungli (Coupe 2003: 45–50). The Mongsen varieties spoken in these villages tend to lose their marked features, such as the voiceless sonorant series that is found in some monodialectal Mongsen villages, and fluctuating degrees of Mongsen~Chungli lexical mixing are also observed. This is demonstrated by variant pronunciations of the same word, e.g. *məzəm* ~ *mījim* ‘poison’, *āzā* ~ *ājā*

‘grass’. Some divergent pairs of lexical items may additionally result from language contact with neighbouring Tibeto-Burman languages, e.g. *tʃəm* ~ *jūŋ* ‘drink’ (*jūŋ* is most likely borrowed by Mongsen from Chang *juy*¹¹ ‘drink’).

2 GENETIC AFFILIATION

Mongsen, Chungli and Changki are members of a cluster of languages known as Ao. Since the mid-twentieth century it has generally been assumed that the Ao dialects form a lower-level genetic grouping with the Western Sangtam, Lotha and Yimchungrü languages spoken in central Nagaland, the basis for this grouping being lexical correspondences and geographical proximity (e.g. Shafer 1950, 1955). Additional support for this grouping is provided by the relational morphology, which is characterized by cognate forms of agentive/instrumental case-marking clitics and ablatives formed from agentive/instrumental + locative case compounds (Coupe 2011b: 26–7).

The only exception to this is found in Chungli, which has borrowed some forms of relational case markers from its Konyak neighbour Chang; for example, the agentive/instrumental/allative form of Chungli is *i*, whereas all other Ao dialects have a syncretic form with a dental nasal and a central vowel, viz. *nə* or *nā*.⁷ Despite its aberrant form, it is significant that Chungli nevertheless demonstrates the same extremely rare agentive/instrumental/allative syncretism that characterizes the Ao dialects, and it also has an ablative that originates from a compound involving the locative and the aforementioned syncretic form, suspiciously in common with other languages of the Ao group.

The members of this cluster additionally share evidence of an obsolescent overcounting numeral system (Coupe 2012). Jejara (Barkman 2014) and Makuri (Vong 2009), both spoken in the Saramati Range dividing southern Nagaland from the Sagaing Division of Myanmar, may also belong to this grouping of central Nagaland languages. Jejara in particular is notable for retaining the extremely rare overcounting cardinal numeral system that was historically shared by members of the Ao group and other languages of southern Nagaland (see §6.1).

Arguments for proposing an intermediate proto-language for the languages of central Nagaland are gaining a more solid footing with recent work on phonological reconstruction (e.g. Bruhn 2014), but relating this intermediate node to a specific branch of Tibeto-Burman is still problematic. Previous attempts at establishing a branch affiliation for the languages of central Nagaland have been tenuously based on little more than conjecture and geographical proximity. It remains the case that the field cannot yet determine if the Ao group forms a separate higher-level branch of Tibeto-Burman, or if it merges with a presupposed existing branch. Coupe (2012) presents evidence of obsolete overcounting systems that appear to be exclusive to many languages of the Ao and Angami-Pochuri groups and tentatively suggests that this may constitute evidence of a shared innovation, possibly pointing to an intermediate-level node linking the Ao group of central Nagaland to the Angami-Pochuri languages of southern Nagaland. Proposed affiliations beyond this level presently remain unsubstantiated by credible evidence.

Until recently, the historical term ‘Naga’ had been infelicitously applied to all the languages spoken in the mountainous area bounded by the Mishmi Hills in the north, the Manipur Valley in the south, the Chindwin River in the east, and the westernmost ranges at the edge of the Plain of Assam.⁸ Yet as some papers have recently pointed out (e.g. Burling 2003; Coupe 2012), the term ‘Naga’ is highly misleading, as it implies that these geographically contiguous languages somehow constitute a well-attested branch of Tibeto-Burman. Of all the languages spoken in the Indo-Burmese Arc south of the Mishmi hills,

it is only the Konyak group that has hitherto been proven to form an independent branch of Tibeto-Burman. The Konyak languages share a set of lexical innovations that links them to Bodo-Garo and Jinghpaw in a grouping originally named the ‘Sal’ languages (Burling 1983), based on the innovative word for ‘sun’ that many of them evince, along with a number of other unique lexical innovations.

To distinguish the non-Konyak languages of central and southern Nagaland from the Sal languages of central-east and northern Nagaland, it is proposed that the neutral and geographically inspired term ‘Indo-Burmic’ henceforth be used in place of the confused and confusing ‘Naga’ label, at least until definitive evidence can be proposed for establishing branch affiliations in these languages. Such an appellative is an improvement over Kuki-Naga (Benedict 1972), Southern Naga (Bradley 1997) or indeed Angami-Ao (Coupe 2012), as it specifically refers to the languages of central and southern Nagaland but does not favour any one linguistic group or perpetuate a historical phylogenetic fallacy.

3 PHONOLOGY

Mongsen Ao syllables minimally consist of a vowel that functions as the nucleus and carries an associated tone.⁹ Onsets are obligatorily simple and can be filled by any consonant with the exception of the glottal stop, which behaves in the manner of a word-level prosody and is consistently deleted before a morpheme boundary in word formation processes (see §5). Because of its peculiar behaviour and unique phonotactic distribution, the glottal stop cannot be accorded the status of a segmental phoneme in a synchronic description.

In common with many tonal languages of Southeast Asia, the coda position is highly restricted and only permits unaspirated plosives, voiced nasals and the voiced retroflex approximant. In most village varieties of Mongsen, only one phonetic diphthong can co-occur with a velar nasal or velar plosive coda in monosyllabic roots, e.g. *t/hāwk* ‘abandoned’, *āwy* ‘jungle’.

The syllable canon is represented as follows (Table 13.1):

$$(1) \quad \sigma = (C_1) V (G) (C_2) T$$

Two varieties of Mongsen Ao can be recognized on the basis of their divergent phonological inventories, the fundamental difference being whether a voiceless series of sonorants has been retained or historically merged with the voiced series. This potentially adds

TABLE 13.1 PHONOTACTIC DISTRIBUTION OF MONGSEN SYLLABLE SEGMENTS (MANGETONG VILLAGE VARIETY)

(C ₁)	V	(G)	(C ₂)	T
p p ^h t t ^h k k ^h				
ts ts ^h tʃ tʃ ^h	i u	w j	p t k	H(igh)
m n ñ	ə		m n ñ	M(id)
m̥ n̥ ñ̥	a ə		ɹ	L(ow)
z s h				
w l r j				
m̥ l̥ r̥ j̥				

TABLE 13.2 CONSONANT PHONEMES OF MONGSEN AO (PHONEMES IN PARENTHESES ARE NOT ATTESTED IN ALL VARIETIES)

	Bilabial	Dental	Post-alveolar	Palatal/pal-alv.	Velar	Glottal
Plosive						
unaspirated	p	t			k	
aspirated	p ^h	t ^h			k ^h	
Affricate						
unaspirated		ts		tʃ		
aspirated		ts ^h		tʃ ^h		
Fricative						
voiced		z				
voiceless		s				h
Nasal						
voiced	m	n			ŋ	
voiceless	(m̥)	(n̥)			(ŋ̥)	
Lateral						
voiced		l				
voiceless		(l̥)				
Approximant						
voiced	w		i	j		
voiceless	(w̥)		(j̥)	(j̥)		

TABLE 13.3 MONGSEN AO VOWEL PHONEMES

	Front	Central	Back
High	i		u
Mid		ə	
Low		a ə	

seven more phonemes to a basic consonant phoneme inventory of 20 segments shared by the vast majority of Mongsen Ao varieties (Table 13.2).

The dental affricates /ts/ and /tsh/ have a highly restricted distribution and only occur in the environment before schwa, yet still respectively contrast with /tʃ/ and /tʃh/ in this position, e.g. *mātsā* ‘spit’ ~ *tjhāmpāŋ* *mātʃāmpūŋ* ‘Yellow-backed Sunbird’, and *mā-tshək* ‘pinched’ ~ *mātʃhəp* ‘kissed’. This is indicative of an evolving phonemic merger levelling out the contrast between affricates at the dental and palato-alveolar places of articulation.

Most Mongsen Ao varieties have a small vowel phoneme inventory of four segments (Table 13.3): this forms a triangular system of three peripheral vowels and one central vowel contrasting three degrees of height, three degrees of backness, and two degrees of rounding. The high back rounded vowel phoneme /u/ occupies an expansive articulatory space that encompasses both [u] and [o], and these two allophones occur in free variation, a characteristic of all Ao varieties. The low central vowel /a/ additionally demonstrates a creaky voice-modal voice phonation contrast, although the creaky vowel is only attested in a handful of verbs and nouns, and only in the environment after a labial velar approximant onset, e.g. *wāpāt* ‘slope’, *mā-tsā* (bamboo-water) ‘bamboo sap’, *wā-pā?* ‘slice-NMLZ’, *tánəm wāpiūŋ* ‘Great Pied Hornbill cock’ and *tánəm wātsā* ‘Great Pied Hornbill hen’. The creaky voice vowel occurs in the absence of a glottal stop, on words that are

clearly not onomatopoeic, and on syllables carrying different tones. Given that it also occurs in contrastive distribution with its modal voice counterpart, e.g. *wā-pā?* ‘go-NMLZ’, and that it is not found to be a manifestation of a particular tone, it must therefore be recognized as an independent vowel phoneme, albeit one with a very light functional load (see Coupe 2003: 43–5, 81–6 for further discussion).

Some varieties of Mongsen spoken in villages on the Changkikong range (Waromung, Khar and possibly others) have a sixth high central rounded vowel phoneme /u/ occurring predominantly in grammatical morphemes and corresponding with /i/ in other Mongsen varieties. This appears to be a developing sound change mostly affecting high-frequency morphemes (see (8) later for an example of this in the Waromung Village variety of Mongsen).

4 SUPRASEGMENTAL FEATURES

4.1 Lexical tone

Mongsen has three lexically contrastive tones, in common with other languages of the Ao group, e.g. *ní* QPTCL, *nī* ‘day’ and *nì* 1SG. All three tones occur on both open and closed syllables and are essentially register-like, in that they do not change significantly in fundamental frequency values over the duration of their rhymes. This renders pitch height at rhyme onset the primary perceptual cue for their recognition. Figure 13.1 plots fundamental frequency as a function of absolute mean duration for one male speaker of Waromung Village Mongsen on open syllables. The findings of the instrumental analysis correlate with an auditory analysis and the results of a perception test that identified three contrastive level tones (see Coupe 2003: 95–9, 2014).

4.2 Tone spreading, grammatical tones and floating tones

Tone spreading frequently occurs across morpheme boundaries, particularly when segment deletions are triggered by word formation processes (see §5 for further discussion). To illustrate, a number of Mongsen varieties of the Ongpangkong Range optionally delete the segmental representation of the agentive case-marking clitic *nā* when it case-marks *nì*, the first person singular pronoun; the dissociated mid tone of the case marker then coalesces with the low tone of the first person pronoun and the resulting low rising tone [ni¹³] in the output encodes 1SG:AGT. This is the only attested example of tonal coalescence producing a contour tone in Mongsen Ao.

Superlative adjective derivations also involve a floating tone that serves a grammatical purpose. To illustrate, adding the nominalizing prefix *tā-* to the verb root *nīj* ‘be ripe’ first derives the deverbal adjective *tā-nīj*, and the comparative is then formed by attaching the nominalizing suffix *-pā?*, which is always realized with a mid tone in comparative adjective stems. This results in the word form *tā-nīj-pā?* (Nzp-be.ripe-NMLZ) ‘riper’, rendering such comparative derivations formally identical to nominalized non-stative monovalent verbs functioning as relative clauses (Coupe 2007: 61; also see §6.2). To derive the superlative degree, the intensifier *-thī* is suffixed immediately after the verb root. Note that this suffix has two tones associated with its syllable: a mid tone associated with the vowel *i*, as well as a floating high tone (represented here by the preceding acute diacritic) that is intrinsic to its morphological representation. The floating high tone productively displaces the adjacent tone of the verb root in the derivation of superlative adjectives, e.g. *tā-nīj-thī-pā?* (Nzp-be.ripe-INTENS-NMLZ) ‘ripest’. Alternatively, a superlative can be

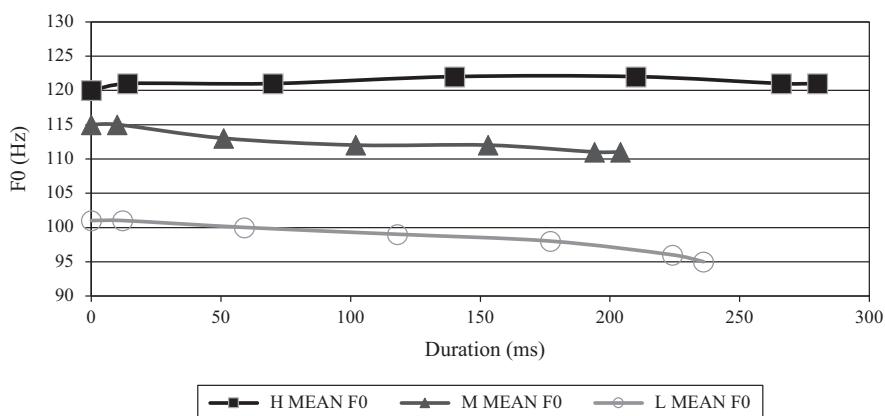


FIGURE 13.1 MONGSEN AO TONEMES (BASED ON FIVE TO SIX TOKENS FOR EACH TONE CATEGORY)

formed by eliding the segmental representation of the intensifier suffix and its associated mid tone completely, leaving just the floating high tone to encode the superlative meaning. This results in the word form *tā-nīŋ-pá?* (Nzp-be.ripe:INTENS-NMLZ), which identically expresses ‘ripest’. In addition to the tone of the verb root being replaced by the high floating tone of the intensifier suffix, the tone of the nominalizing suffix now also becomes high, ostensibly as a result of perseverative tone spreading and the iconicity of emphatic meaning being associated with higher pitch. Tone can thus serve as the only morphophonological feature distinguishing the comparative and superlative adjective derivations.

4.3 Intonation

In addition to the lexical tone system, Mongsen prosody is characterized by the use of intonation; this extends over the tones of individual syllables in various phrasal and clausal constructions, often significantly distorting their realizations.¹⁰ Rising intonation is employed as a boundary signal to delineate phrasal boundaries and convey non-finality, and it is used to mark topic continuity across a series of matrix clauses (Coupe [2007: 73–5, 2014: 467–8]). To illustrate, intonation applying to different phrasal and clausal domains is indicated by the notation ↗ in the following (interrupted) introduction to a narrated text.

- (2) *nì tānīŋ↗ sāsīmūnglā. nì tālījān↗ tūngtānmàn. māpāng à kū↗ ... [ākām kà? sāŋ átè] ... nì ākām↗ .ñūk.iā. thānī.*
- nì tā-nīŋ↗ sāsīmūnglā. nì tā-lī-tʃān↗ tūngtān-màn.*
 1SG RL-name pers.name 1SG NZP-stay-LNMLZ1 log.drum-LNMLZ2
māpāng à kū↗ ... [ā-kām kà? sā-āŋ átè] ...
 time one LOC NRL-year also say-IMPER PTCL
nì ā-kām↗ .ñūk.iā. thānī.
 1sg NRL-year sixty-SEQ seven
- ‘My name is Sashimongla. I live in the LogDrum Ward. Once upon a time... [interrupter: ‘Also say your age.’] ... my age is sixty-seven.’

Rising intonation has additionally developed the morphosyntactic function of distinguishing a non-finite verb stem inflected by the sequential converb suffix from a finite matrix verb inflected for present tense. These two verb forms (both of which are demonstrated in [3] below) share a syncretic termination in *-ə̄.i* and occur clause-finally.¹¹ Whereas the present tense marker *-ə̄.i* carries a low tone, the sequential converb's underlying tone cannot be determined, as it remains obscured by the rising intonation that typically occurs with this type of converb marker. Both morphemes derive historically from an older nominal form that has undergone multiple reanalyses of function, including agentive nominalization, genitivization, relativization, sequential converb marking and present tense marking (Coupe 2013).

- (3) *tè-ə̄.i> thākūj pā wàzà? tòpètī à nē tāsū? à táj kēmə̄.i> lì.i*
 tè-ə̄.i> thākū ī pā wàzà? tò-pètī à nē
 thus-SEQ now+LOC PROX 3SG bird NZP-be.big one AGT
 tā-āsū? à táj kēm-ə̄.i> lì-ə̄.i
 NZP-be.small one just become-SEQ be-PRES
 ‘And now she, a big bird, has become just a small one.’ (Coupe 2007: 485)

A rising intonation is also used to distinguish negated sequential converb clauses (which typically denote a privative meaning) from negated realis matrix clauses. For example, *mò-wā-lā* (NEG-go-NEG:SEQ) expresses the non-finite conversal meaning of ‘without going’ when accompanied by a rising intonation, whereas *mò-wā-lā* (NEG-go-NEG:REAL) without the overarching intonation pattern can only express a finite meaning of ‘didn’t go/hasn’t gone’.

Intonation thus disambiguates the non-final clause-linking function of the sequential converb from identically inflected matrix clause verb forms, and this distinguishing function must have evolved after the agentive nominalizer underwent reanalysis as a present tense marker (Coupe 2013; also see §6.4 later). Once intonation became established as a marker of non-finality, this usage then generalized and spread to other types of non-final clausal and phrasal structures. It is now a pervasive prosodic feature of all Mongen Ao discourse.

5 WORD FORMATION PROCESSES

Word formation in Mongsen Ao is moderately synthetic and agglutinative, with nominal and verbal stems consisting of up to six morphemes representing the upper limit of morphological complexity.

5.1 Diphthongization, dissimilation, assimilation and deletion

Vowel length is not contrastive, and tautosyllabic long vowels are prohibited by the phonotactics. If two identical vowels become juxtaposed as a consequence of word formation processes, then one of the vowels is commonly deleted, e.g. *sānā* ‘speak’ + *āŋ* IMPER → *sānāŋ* ‘speak!’. When crucial morphological information might be lost as a result of vowel deletion involving suffixation, the affected vowel instead undergoes a dissimilation to preserve its morpheme’s grammatical meaning. This only applies to high front vowel juxtapositions with grammatical suffixes consisting of empty onsets, such as when the unrealis marker *-i* and/or the causative marker *-iʔ* are concatenated with a verb stem terminating in an identical high front vowel, e.g. *hàmsi* ‘use’ + *-i* IRR → *hàmsə̄j* ‘will/would use’. In addition to dissimilation of the initial high front vowel of the juxtaposition to schwa, a phonetic diphthong forms across the morpheme boundary, and the second vowel of the resulting V₁V₂ sequence is reduced to a non-syllabic offglide.

Concatenated vowels of differing qualities may form phonetic diphthongs under similar circumstances if V₂ is a high vowel, e.g. *tʃhà* ‘make/do’ + *-i̥kù?* ANT → *tʃhāwkwù*, and *sə* ‘die’ + *-i* IRR → *səj*. In this case, the tone associated with V₂ is lost. Less frequently, the word may be resyllabified so that each vowel forms the nucleus of an independent syllable carrying its own tone.

Glottal stops are consistently deleted in the environment of a morpheme boundary, thus limiting their realization to the word final position, e.g. *á-ŋá?* ‘NRL-fish’ + *tə-kəp* ‘RL-skin’ → *ŋā-kəp* (fish-skin) ‘fish scale’. This offers evidence that the glottal stop functions as a prosody at the level of the word, rather than as a full-fledged segmental phoneme. Although it has a divergent behaviour in word formation processes, nevertheless it still encodes a contrastive function phonologically, e.g. *ā-mī?* ‘NRL-person’, *ā-mī* ‘NRL-spear’.

5.2 Compounding and grammaticalization

Compounding is a common feature of word formation in Mongsen Ao and is arguably the historical source of much of the nominal and verbal morphology, particularly the grammaticalized lexical suffixes (see §6.4). A widely attested example of an erstwhile noun that now functions as a suffix in many Tibeto-Burman languages is the diminutive *-zā*, which corresponds with Proto-Tibeto-Burman (PTB) *za *tsa ‘child’ (Matisoff 2003: 31), and is still preserved in the Mongsen elaborate expression *tə-zā tə-nū* ‘children’. This morpheme is now used to express diminutives of both animate and inanimate nouns, e.g. *ā-hən-zā* (NRL-domestic.fowl-DIM) ‘domestic fowl chick’, *ā-jūŋ-zā* (NRL-RIVER-DIM) ‘creek, stream’, demonstrating that it is fully grammaticalized as a functional morpheme.

PTB *lak ‘hand, arm’ has undergone a similar grammaticalization process in N-N compounds and now productively expresses a semantically related meaning of ‘terminal part’ when suffixed to noun stems, e.g. *təmījūŋ-lāk* (finger-TERM) ‘fingertip’ and *tə-mī-lāk* (RL-tail-TERM) ‘tail-tip’. In common with the diminutive suffix, its status as a semantically bleached functional morpheme is demonstrated by the possibility of it occurring on the stems of inanimate nouns, e.g. *səŋtūŋ-lāk* ‘apex of tree’ and *mī-lāk* (fire-TERM) ‘flame’. In keeping with the head-final typology of Mongsen Ao, it is likely that the morphemes *zā* and *lāk* were originally the heads of compound nouns before losing their status as independent words in this type of construction.

The original PTB form *lak no longer expresses ‘hand’ in Mongsen, having since been replaced by *tə-khát* ‘RL-hand’, but it still occurs with a relational meaning, e.g. *āliŋ tə-lāk phīnð* (lower.range RL-END ABL) ‘from the end of (a) lower range’. In a further surprising development, a morpheme with an identical phonological shape and metaphorically related semantics is used on verb stems to express the telic boundary of an activity, and this almost certainly has the same historical source as the terminative suffix of nouns. Both morphemes share the semantics of denoting an end point: one in space, and the other in time.

- (4) *tə ākī tsū tʃhà ... tʃhālākəi*, (Coupe 2007: 465)

tə *ā-kī* *tsū* *tʃhà...* *tʃhā-lāk-əi*
thus NRL-house DIST make make-TERM-SEQ

‘Thus, having finished [narrator stutters] building [his] house ...’

Compounding is also a likely historical source of the case-marking clitics (§6.3), the majority of which have grammaticalized from nouns expressing relational and orientative meanings.

6 MORPHOSYNTAX

Mongsen Ao is a predominantly suffixing, predicate-final language in which pragmatics determines the order of its pre-verbal core NP arguments. Topical arguments tend to occur first in the clause. Dependent clauses precede their matrix clauses, in keeping with the head-final typological orientation.

6.1 Word classes

Mongsen Ao has open classes of nouns and verbs, and closed classes of pronouns (personal, possessive, demonstrative, interrogative, indefinite), nominal modifiers (nominal deictics, quantifiers, case-marking clitics), time words, adverbs, particles, interjections, onomatopoeia and discourse connectives. Discourse connectives are a grammaticalized closed class that has historically developed out of collocations involving the quotative particle and various converb suffixes. They have a clausal conjunctive function, particularly in tail-head linkage (see §6.6).

Three types of noun can be distinguished on the basis of their morphological structure. Relational nouns are bound morphemes that obligatorily take a relational prefix *tə-*. These denote kinship terms (e.g. *tə-pā?* ‘father’, *tə-tī* ‘older sibling’), body part terms (e.g. *tə-kūlūk* ‘brain’ and *tə-məlī* ‘tongue’), or express a part–whole relationship to another living entity (e.g. *tə-tūŋ* ‘trunk, stem’, *tə-mətshə?* ‘bud’).

Non-relational nouns are bound morphemes obligatorily taking the non-relational prefix *ā-*. These denote entities belonging to a diverse range of semantic classes, including many cultural artefacts (e.g. *ā-thī* ‘granary’, *ā-tūŋ* ‘shield’), some animal names (e.g. *ā-hī?* ‘rat’, *ā-hān* ‘domestic fowl’), features of the biosphere (e.g. *ā-nīŋ* ‘sky’, *ā-tsā* ‘water’), edible and inedible plants (e.g. *ā-mī* ‘taro’, *ā-žā* ‘cane’), terms for types of people (e.g. *ā-zəntī* ‘old person’, *ā-nū* ‘child’), and a small number of semantically unrelated entities (e.g. *ā-jū* ‘word’, *ā-t̥ən* ‘money’, and *ā-kām* ‘year’).

The third type of noun is a free morpheme that encompasses a miscellaneous collection of semantic classes, including names of entities belonging to the natural world (*ūpūŋ* ‘dust’, *lātā* ‘moon’), names of wild and domestic animals (e.g. *tsəjī?* ‘dhole’, *wājū* ‘crow’, *māsə?* ‘ox’), insects, arachnids and invertebrates (e.g. *ləp̥jū* ‘cockroach’, *sūdāk* ‘spider’, *mānət* ‘water leech’), plants (e.g. *máphú?* ‘pumpkin’, *pūlūŋ* ‘jackfruit’, *səyā* ‘tea’) as well as cultural artefacts and various other creations of human beings (e.g. *khāŋ* ‘cage’, *thījā?* ‘contract wages’, *tsəŋjəžm* ‘god’, *tūŋpāŋ* ‘war’).

In common with most other Tibeto-Burman languages, Mongsen Ao lacks an undrived class of adjective. Property concepts instead find expression via deverbal adjectives derived from stative verbs by means of a nominalizing prefix *tə-*, e.g. *məžəžm* ‘be.red’ → *tə-məžəžm* ‘NPZ-be.red’; a derived adjective can then be used to modify an NP’s head noun, e.g. *ā-sə tə-məžəžm* (NRL-shawl NPZ-be.red) ‘red shawl’.

The personal pronouns are listed in Table 13.4. There are three persons and three numbers, and an inclusive~exclusive distinction characterizes the first person dual and plural forms, in common with many other Tibeto-Burman languages. The dual forms have transparently grammaticalized from collocations involving the root of the numeral *ā-nāt* ‘NRL-two’. A good deal of allomorphy in free variation is apparent in the third dual category and in the entire plural series, which is mostly due to phonological reduction.

The possessive/oblique forms of pronouns are listed in Table 13.5. With the exception of the first person singular *kə* and third person plural *páŋ*, their forms are very

TABLE 13.4 MONGSEN AO PERSONAL PRONOUNS

		1st	2nd	3rd
SG		<i>nì</i>	<i>nàŋ</i>	<i>pā</i>
DU	INC	<i>īnāt</i>	<i>nāŋjāt</i>	<i>pānāt ~ tūŋnāt ~ tūŋjāt</i>
	EXC	<i>kānāt</i>		
PL	INC	<i>isā ~ isāŋjālā</i>	<i>nòŋlā ~ nòŋkhòlā</i>	<i>túŋlā ~ tūŋjālā ~ tūŋkhòlā</i>
	EXC	<i>ilā ~ ikhòlā</i>		

TABLE 13.5 MONGSEN AO POSSESSIVE/OBLIQUE PRONOUNS (FORMS IN PARENTHESES ARE IDENTICAL TO PERSONAL PRONOUNS)

		1st	2nd	3rd
SG		<i>kā</i>	<i>nā</i>	<i>(pā)</i>
DU	INC	<i>(īnāt)</i>	<i>(nāŋjāt)</i>	<i>(pānāt ~ tūŋnāt ~ tūŋjāt)</i>
	EXC	<i>(kānāt)</i>		
PL		<i>ī ~ āj</i>	<i>nīŋ</i>	<i>pá</i>

similar if not identical to the corresponding personal pronouns previously presented in Table 13.4.

It is generally the case that independent pronouns and possessive/oblique pronouns are used in free variation for marking possession. For some speakers, the possessive pronoun co-occurs with the relational prefixes of bound nouns, e.g. *kā tā-nīŋ* ‘my name’, while other speakers replace the relational prefix with a possessive pronoun, e.g. *kā-nīŋ*, ‘my name’. This optionality demonstrates the intermediate status of possessive pronouns as words that are in the process of becoming bound morphemes. A third option demonstrated in texts is for a personal pronoun to be used with a pre-fixed bound noun to encode possession, e.g. *nì tā-nīŋ* ‘my name’. Possessive/oblique forms of pronouns are obligatorily used with the oblique case markers (e.g. see [13] for an example).

Nouns may be modified by the proximate and distal demonstratives *ī-pī~īpī* and *tū* to encode spatial deixis. An anaphoric demonstrative *sā*, the third member of this closed word class, is used to encode the antecedent reference of an entity that is not visible. In everyday conversation the anaphoric demonstrative is reserved for determining NPs whose referents are uniquely identifiable to interlocutors. The distal demonstrative *tū* is used as a default determiner, in addition to determining the NPs of referents that are distant from the deictic centre. In extended texts it is commonly used to determine NPs initially until their reference is well established, after which the anaphoric demonstrative *sā* gradually takes over this role.

Quantifiers include terms that express amounts, e.g. *tāmāy* ‘all’, *itjā?* ‘few’, and cardinal numerals. Numerals ONE to THREE and SIX have prefixes that correspond with reconstructed PTB prefixal forms, e.g. *ā-khātā* ‘NRL-one’, *ā-sām* ‘NRL-three’, *tā-āuk* ‘RL-SIX’. A phonologically reduced form of the numeral ONE is additionally used as an indefinite determiner, e.g. *khān māpāŋ* à *kū* (once time one LOC) ‘once upon a time’, in common with many of the world’s languages.

The cardinal numeral system of Mongsen Ao is now uniformly decimal, but until the beginning of the twentieth century all but one Ao dialect had a typologically

unusual overcounting system (Coupe 2007, 2012). This presented as a decimal system up to the numeral FIFTEEN, then from SIXTEEN to NINETEEN the overcounting system operated to express units of progress towards the next whole decade. To illustrate, SIXTEEN was expressed by the formula *mākī mā-pōn tā-.ūk*, (twenty NEG-complete RL-six), literally ‘the twenty not complete, the six’, or the sixth unit towards 20. Overcounting used to operate in every decade from the sixth to the ninth unit from the teens onward in Ao, and overcounting was also documented in many other Indo-Burmic languages of central and southern Nagaland before becoming obsolete in the early twentieth century.

Verbs can be divided into monovalent and bivalent classes, according to how many core arguments each may take. There is no justification for identifying a trivalent class of verb, as prototypically three-place predicates such as ‘give’ cannot be distinguished on syntactic grounds from ordinary bivalent verbs with an additional oblique argument. The recognition of traditional transitivity classes is of limited value, since any bivalent verb can occur with just one argument, especially when a resultant state needs to be expressed. This is simply achieved by omitting the actor argument. Whereas bivalent verbs comfortably occur with either one or two core arguments, a distinguishing feature of monovalent verbs is that they only take one core argument. The only exception to this is when the valency of a monovalent verb stem is increased via causativization.

A copula *l̥* is used to express existence and location; a grammaticalized form of the transitive verb *tjhà* ‘do, make’ is used to encode existence or location in the past (see (21)),¹² and *tjhà* additionally serves as a vehicle for conveying other finite temporal and modal distinctions of copula clauses via suffixal morphology (see Coupe [2007: 371–4] for further discussion and examples).

6.2 The noun phrase and nominal morphology

The template of (5) presents the basic structure of the NP. In addition to these six constituents, a number of discourse particles can be used with a relatively free distribution within the NP to express restrictive, additive or emphatic semantics of elements in the NP. The head is filled by a noun, personal pronoun or demonstrative pronoun and is the only obligatory constituent.

(5) (A₁) B (A₂) (C) (D) (E)



Non-relativized attributes are represented by deverbal adjectives, e.g. *ānītì tā-sān tʃū* (woman NZP-be.new DIST) ‘that new woman’, and the attribute is necessarily restricted to the A₂ post-head position. In contrast, relativized attributes may precede or follow the head noun in the mutually exclusive A₁ or A₂ position, sometimes with significantly different semantics in each case. Pre-head relativized attributes generally express a restrictive meaning, whereas post-head relatives have elaborating appositional semantics and typically express a non-restrictive meaning. This contrast is demonstrated by the examples of (6a–b). The Ao consider *tsəŋjì* ‘rain’ to be unique and universal, i.e. there cannot be different subsets

of ‘rain’ in the world. However, the restrictive semantics of the pre-head relative clause in the elicited example of (6b) (based on the actual textual example of (6a)), does indeed imply that a specific subset of rain is being singled out here, so this semantic constraint renders the pre-head position of the relative clause logically incongruous in this context.

- (6) a. *tə̄. tsəŋjī. lāpà?* tāj tʃū,
- tə̄-əI [tsəŋjī] [lā-pà?]
- thus-SEQ rain come-NMLZ only DIST
- ‘And, only the rain that comes,’
- b. ?? tə̄-əI [[lā-pà?]] tsəŋjī tāj tʃū]_{NP} (Coupe 2007: 220)

In contrast, (7) demonstrates the obligatory pre-head position of the relative clause of *māpāng* ‘time’, which restricts the temporal reference of the relativized event to a specific time of the year.

- (7) *tʃūŋlijimtī līmā kū, álú jímpà?* māpāng kū, ...
- tʃūŋlijimtī līmā kū [[álú jímpà?]] māpāng kū]_{NP}
- village name country LOC NRL-field cultivate-NMLZ time LOC
- ‘In Chungliyimti country, at the time of cultivating the fields, ...’

Internally headed relative clauses are a possible but rarely occurring type of nominal modifier. Access to internally headed relativization appears to be restricted to a notional core argument of a bivalent verb, as in the following Waromung Village Mongsen example demonstrating relativization of an O argument.

- (8) *nì kà?* *ātī nā āsā à ənərə. lāpà?* pū mūk.
- nì kà? [[ā-tī nā ā-sā à ən-ərə
- 1SG also VOC-elder.sibling AGT NRL-shawl one take-SEQ
- lā-pà?]] pū]_{NP} mūk-Ø
- come-NMLZ PROX wrap-PST
- ‘I also wore this shawl that Elder Sister brought.’¹³

Nominalizations functioning as relative clauses can also be headless. In the following textual example, the nominalized verb stem functions as a headless relative clause, and the elided head is subsequently restored by the speaker as an afterthought to clarify the reference of the headless relativization. Both the nominalized verb stem and the following noun *tākī?* ‘bamboo vessel’ are determined by nominal demonstratives, so this is clearly representative of an appositive structure that adds a qualifying elaboration.

- (9) *pā nō tʃhijākpà?* sō, *tākī?* sō hēnərə, khī.
- pā nō [tʃhijākpà? sō]_{NP} tākī? sō hēn-ərə,
- 3SG AGT keep-NMLZ ANAPH bamboo.vessel ANAPH carry-SEQ
- khī?-əI give-PRES
- ‘After bringing that which he had kept, the aforementioned bamboo vessels, he gives them.’

6.3 Case-marking morphology

Forms of the case markers are listed in Table 13.6. These have a clitic-like distribution and take as their host the last constituent of an NP. There is only one core case marker – the agentive *nō* – which is used in pragmatically determined situations to mark the actor

TABLE 13.6 MONGSEN AO CASE-MARKING CLITICS

		Core	Oblique	
<i>nā</i>	Agentive	Non-local	Instrumental	Primary local
		<i>nā</i>		<i>nā</i>
		<i>lī</i>	Dative	<i>phīnā</i>
		<i>ātāmākā</i>	Benefactive	<i>kū</i>
		<i>thān</i>	Comitative	

arguments of both monovalent and bivalent predicates.¹⁴ Seven case-marking clitics are used for encoding oblique cases: four of these denote non-local relations (instrumental, dative, benefactive, comitative), and the remaining three are used to encode spatial relations (allative, ablative, locative).

With the exception of the benefactive marker, which almost certainly has a converbal source (Coupe 2007: 181), the relational morphology has grammaticalized from N_1-N_2 compounds or appositions in which N_2 was originally a noun denoting a spatial or relational meaning. The process continues to be repeated, with the emergence of nascent postpositions resulting in diachronic layers of newer case markers. The common targets of such grammaticalizations are nominal roots of body part terms and relational nouns in appositional constructions, e.g. ‘back’, ‘mouth’, ‘face’, ‘stomach’, ‘side’.

- (10) *tshèlūylā tʃū ākhūlā sīn kū mənər*, (Coupe 2007: 188)

tshèlūy-lā tʃū ākhū-lā sīn kū mən-ər

fox-F DIST tiger-F BACK LOC sit-SEQ

‘Fox, having sat down behind Tiger, ...’ (lit. ‘at Tiger’s back’)

In the following recorded conversation, *tā-mā* ‘RL-face’ is used by a speaker with a super-cessive meaning in her response to a question about her age:

- (11) *nāj ākām kūjā?? nī ālāktfūkā, ūwá tā, hāw?*, *ālā nātā, tāmāw?*

nāj ā-kām kūjā?

2SG NRL-year how many?

‘How old are you?’

nī ālāk-tfūk-ā ūwātā hāw? *ālā nātā, tā-mā-ù?*

1SG forget-PFV-PRES EMPHAT1 yes and.then seventy RL-face-DECL

‘I forgot. Oh yes, above seventy.’ (lit. ‘seventy’s face’)

The NP of a human goal of movement is always marked by the postposition *tāy* together with one of the primary local cases, as is the NP of a human referent to whom speech is directed. The lexical source of this postposition is the noun *tāyī* ‘side’. The following example captures the intermediate stage of the grammaticalization of *tāy* in its categorial transition from noun to postposition, as it is still determined by the distal demonstrative.

- (12) *tāyā, tʃū nā tāpā? khā tājā nātā tāy tʃū nā wā,* ...

tāyā, tʃū nā tā-pā? khā tā-jā nātā tāy tʃū nā

other DIST AGT RL-father CONJ RL-mother two SIDE DIST ALL

wā-ər

go-SEQ

‘Others went to the mother and father,...’ (lit. ‘the mother and father’s side’)

The distribution of the agentive marker is remarkable for the fact that its use is not dictated by syntax or the valency status of the clause. It occurs on the actor NPs of both monovalent and bivalent verbs, but generally only under pragmatically determined circumstances, and many bivalent clauses lack case marking on their core arguments entirely if the context of use permits an unambiguous assignment of semantic roles.

Like a number of other TB languages with pragmatically motivated marking (e.g. see Chelliah and Hyslop 2011, 2012 and papers therein), core case marking in Mongsen Ao does not accord with any alignment pattern hitherto proposed in the typological literature. When agentive case marking *is* used, it is to disambiguate semantic roles, especially when the referents are both animate (and especially human), or to encode that a participant is performing an activity with increased agency. It also appears when the referent of an NP is taking responsibility for the occurrence of an event (such as breaking something), or is making a personal choice to perform an activity.

An inkling of this nuance is evident in the main clause of the following textual example, in which the speaker threatens to leave her husband if her demand is not met. In pragmatically neutral contexts, the copula verb *lī* ‘be at, stay, exist’ does not normally require an agentive case-marked actor (see Coupe 2007: ch. 5, 2011a, 2011b for further examples and discussion).

- (13) *nàj nō ānū ī mètāpsàtʃūkpàlā, nì nō nōthān mèljìù?*, *tò sā, ɻāpītsə́i tʃū nō.*
- | | | | | | | |
|------------------|-------------------|-------------|--------------|--------------------------|------------|-----------|
| <i>nàj</i> | <i>nō</i> | <i>ā-nū</i> | <i>ī</i> | <i>mè-tāpsàt-ʃūkpàlā</i> | <i>nì</i> | <i>nō</i> |
| 2SG | AGT | NRL-child | PROX | NEG-kill-PFV-COND | 1SG | AGT |
| <i>nō-thān</i> | <i>mè-lī-i-ù?</i> | <i>tò</i> | <i>sā-</i> Ø | <i>ɻāpītsə́i</i> | <i>tʃū</i> | <i>nō</i> |
| 2SG.POSS/OBL-COM | NEG-stay-IRR-DECL | thus | say-PST | stepmother | DIST | AGT |
- ‘If you do not kill this child, I will not stay with you, said the stepmother.’

There are just a few conditions under which the agentive case marker must obligatorily be used on an argument of a verbal clause. First, the actors of verbs of vocalization generally have agentive case marking on their NPs, as demonstrated by the clause-final NP of the preceding example. Second, causer arguments of causativized clauses obligatorily take agentive marking on their NPs, because causatives typically involve two human participants, and this creates the possibility of ambiguity arising as to which NP argument is the causer and which is the causee. Agentive marking presumably resolves potential ambiguity in this situation. Third, an obligatory marking pattern is observed in verbal clauses expressing habitual activities. The marking of a habitual actor with an agentive case marker is also reported in Meithei by Chelliah (2009: 391–2). This may be a manifestation of the agency or personal choice associated with habituated activities performed by animate referents.

Given the ubiquity of a syncretic agentive/instrumental form *nā-nō* in the Ao group, it is plausible that this is reconstructible to the level of Proto-Ao (PAo) as **na* (Coupe 2011b: 32), which in turn may be a reflex of PTB **ʔ-nam* ‘side/rib’, as reconstructed by Matisoff (2003: 100, 604). The noun ‘side’ is a known lexical source for locatives in Chinese and some French-based pidgins and creoles (Heine and Kuteva 2002: 272), and (12) above demonstrates that a noun with the meaning of ‘side’ is again undergoing exactly the same kind of shift to a locative postpositional use in a new cycle of grammaticalization. In a comparable investigation, Beames (2012 [1875]: 257) concluded that the locative declension of Sanskrit *kákṣa-* ‘armpit’ was the most likely lexical source of the Hindi dative marker *ko*. Reinöhl (2016: 58–9), citing Beames and other literature, proposes that the semantic extension of the meaning of *kákṣa-* from ‘armpit’ to ‘side (of the body), flank’ provides a logical explanation for its grammaticalization as a marker of a

goal, and subsequently to the marking of recipients and patients. This lexical source and its trajectory of grammaticalization correlates with the goal-marking function of Mongsen *n̄s*, but it is atypical that this relational morpheme has undergone a metaphorical extension to marking actors, rather than recipients or patients.

The Proto-Ao form **na* would have originally grammaticalized as a non-specific type of local case marker from the relational head of a compound noun. This is generally consistent with LaPolla's (1995) observations concerning the diachronic origins of Tibeto-Burman case-marking systems, except for the interesting fact that allative/agentive/instrumental isomorphism is an unattested syncretism; an allative is reported to be more commonly paired with a locative- or a dative-marking function that is then extended to marking a core argument in a non-agentive semantic role, as demonstrated by Hindi *ko*.

The established pathway is for constructions with concrete meanings to develop more abstract meanings via metaphorical extension (e.g. Heine *et al.* 1991: 123ff.), so the core marking function must have been innovated after the development of the local marking function and was probably motivated by a need to disambiguate the semantic roles of animate core arguments, as suggested earlier in regard to causativized clauses. The lack of syntactic systematicity in core case marking precludes identifying this as an ergative-absolutive marking pattern at the present point in time, but this certainly appears to be a pathway by which such an alignment might eventually develop.

6.4 Verbs and verbal morphology

All Mongsen Ao verbs are negated by the negative prefix *m̄d-* and can be inflected by tense-, aspect- and mood-marking suffixes. The verbal predicate potentially has 11 slots consisting of one prefix position and nine suffix positions in addition to the root; stems typically have two to four morphemes. Table 13.7 summarizes the grammatical categories represented in the verb stem. Slots 3 and 7 constitute ‘zones’ (signalled in the table by an asterisk) that permit multiple instantiations of the same category.

The form of the prohibitive mood prefix is *t̄d-* (< PTB **ta* ✕ *da*, Matisoff [2003: 162]). The nominalizing prefix derives types of deverbal adjectives from monovalent verb stems (see §4.2 and §6.1 for examples and discussion). The fourth verbal prefix is the admonitive mood marker *āsá-*, which is used for chiding and warning.

The lexical suffix zone consists of approximately 17 morphemes that are mostly traceable to lexical verb roots, such as the terminative aspect marker *-lāk*, discussed in

TABLE 13.7 PREDICATE STRUCTURE OF MONGSEN AO

-
1. Negative prefix, prohibitive mood prefix, nominalizing prefix, admonitive mood prefix
 2. Root
 3. *Lexical suffix zone
 4. Reciprocal/collective suffix
 5. Directional suffix
 6. Aspectual suffixes
 7. *Modality suffix zone
 8. Resultant state suffix
 9. Perfective aspect suffix
 10. Causative suffix
 11. Negative suffix, positive imperative suffix, converb suffixes, tense/mood suffixes, general nominalizer suffix
-

section 5.2. These have developed out of compounds and express an extensive range of meanings related to their lexical sources. As noted earlier, their position of occurrence in the predicate is best analysed as a ‘zone’, as more than one lexical suffix can occur in tandem in this slot. The same applies to the modality suffix zone, which also permits more than one modality suffix to occur in concatenation. Following are some examples demonstrating the use of lexical suffixes, their likely lexical sources, and a selection of other verbal suffixes.

- (14) *-thān* TOGETHER < $\sqrt{āthānsi}$ ‘gather’
ākhūlā nā wā.ə, sāŋjā tōlik tāŋ kū āj? tōnthān-jāŋ.
 ākhū-lā nā wā-ə AGT sāŋjā tōlik tāŋ kū
 tiger-F go-SEQ monkey group SIDE LOC
 āj? tōn-thān-T?-āŋ
 cane pluck-TOGETHER-CAUS-IMPER
 ‘Tiger went and said to a group of monkeys “Pluck and gather cane for me.”’

- (15) *-thāŋ* SEVER < $\sqrt{thāŋ}$ ‘block’, or possibly ‘sever, divide’
tōmēlī wāt̪hāŋə, pūŋilā sā sāt̪fūk.
 tō-mēlī wā?-thāŋ-ə AGT pūŋi-lā sā sāt̪fūk-Ø
 RL-tongue slice-SEVER-SEQ wild.pig-F ANAPH die-PFV-PST
 ‘Having sliced off its tongue, the aforementioned wild pig died.’

- (16) *-thām* END < $\sqrt{thām}$ ‘finish, end’
ājilā nā áwklā t̪āŋsā t̪ū tōmáŋ t̪ēmthāmt̪fūk āj ...
 ājí-lā nā áwk-lā t̪āŋsā t̪ū tōmáŋ t̪ēm-thām-t̪fūk āj
 dog-F AGT pig-F footprint DIST all trample-END-PFV BECAUSE
 ‘Since Dog had trampled all over Pig’s footprints, ...’

A morphological causative suffix *-i?* (-ū? in the Khar and Waromung Village varieties of Mongsen) adds a core argument in the role of a causer, and this obligatorily takes agentive case marking, as previously discussed in §6.3. In the Khensa and Mekuli Village varieties of Mongsen, the form of the causative is *-pī?* This makes it likely that the lexical source of the causative suffix is PTB **bəy* ‘give’ (Matisoff 2003: 132), as ‘give’ is known to be a common diachronic source of morphological permissives and causatives in South-east Asia (Heine and Kuteva 2002: 152). See examples (14) and (17).

- (17) *.ātū nīŋ-ə, nā jāsà nā.ā, āmāŋ hēnō.ā.ə, māŋmātūŋ kū mātūŋt̪ākī?*
.ātū nīŋ-ə AGT nā jāsà nā.ā ā-māŋ hēnō.ā.ə
 village.name name-ANMLZ AGT person’s name NRL-body bring-SEQ
māŋmātūŋ kū mātūŋ-t̪āk-ī?-Ø
 village.name LOC be.erect-RS-CAUS-PST
 ‘The people of Ritu Village brought the [smoked and dried] corpse of Yasa Naro and left it propped up in Mangmetong village.’

Forms of tense- and mood-marking suffixes are listed in Table 13.8. The tense/mood-marking system appears to have originally consisted of just two mood categories: an unmarked realis, and an overtly marked unrealis form *-i*. The language subsequently reanalysed its agentive nominalizer *-ə* as a present tense marker and its purposive nominalizer *-ā* as an immediate future marker (both now distinguished by low tones in these reanalysed functions), and the old zero-marked realis was reinterpreted as encoding an unmarked past in a new paradigm of tense marking. The unrealis continues to be used for encoding non-actualized events that are unspecified for a temporal location.

TABLE 13.8 TENSE- AND MOOD-MARKING MORPHEMES OF MONGSEN AO

-Ø	Past	-i	Irrealis
-ð.i	Present	-ūkù	Anterior
-ñù	Immediate Future		

This explanation logically accounts for the syncretic forms of the tense markers and nominalizing suffixes, and also for the peculiarity of the past tense being the unmarked form. As noted in Coupe (2013), nominalizing morphology is cross-linguistically a common source of tense-marking morphology. Purposive nominalizers in particular are ideal targets for the grammaticalization of futures, because their semantics allude to non-actualized events. The development of a present tense marker from an old agentive nominalizer is less easily accounted for, but most likely arose through the loss of a copula and reanalysis of the now clause-final nominalizer on the verb stem as being in paradigmatic opposition to other clause-final tense/mood markers.

There are a number of constraints on the use of the immediate future and anterior suffixes. Verbs inflected by these suffixes cannot be negated, and the immediate future can only be used on verbs whose actor referents have the epistemic authority to make assertions about predicted activities or states. This generally limits its use to the first person, or to the reported speech of first persons. See Coupe (2007: 344–5, 351–4) and Coupe (2013: 1121ff.) for additional discussion and examples.

6.5 Basic sentence types

Mongsen Ao has verbless clauses (18)–(20), copula clauses (21), monovalent verbal clauses (23), and bivalent verbal clauses (24).

Verbless clauses consist of a verbless clause topic and a verbless clause complement and express timeless equational meanings denoting the identity or class membership of a referent. The usual presentation in narrative texts is for the topic NP to precede the complement, but this constituent order can be manipulated for pragmatic effect, as in (19). Neither constituent of verbless clauses receives any special marking, although topic NPs can be formally identified by a topic particle (e.g. see (20)).

To facilitate their recognition, verbless clause topics and their complements are respectively bracketed and notated with the subscript labels _{VCT} and _{VCC} in the following three examples.

- (18) *ātñū kúták thùñipà? mīw?*
 [ā-ñū]_{VCT} [kúták thùñ-i-pà?] _{VCC}
 NRL-DIST heaven reach-IRR-NMLZ person-DECL
 ‘Those are the people who will reach heaven.’ (Coupe 2007: 469)
- (19) *nì nō sājākə̄ līpā? sə̄ pīw? tə̄*
 [nì nō sā-jā-ākə̄ lī-pà?] _{VCC} [sə̄] _{VCT} [pī-ū?] _{VCT} tə̄
 1SG AGT say-CONT-SIM be-NMLZ ANAPH PROX-DECL PTCL
 ‘This is what I kept saying would happen.’ (Coupe 2007: 360)

Verbless clauses are additionally used for ascribing properties to referents. The property is expressed via a nominalized verb stem that functions as a deverbal adjective in the complement clause.

- (20) *nàj wàzà?*, *nàj nā ātʃhū lā ā:ñkā tōn:ñ?* *tòpákukā?*, *nàj ásá?* *tfū lā tásón:ñ?*.
nàj wàzà? nàj nā ā-tʃhū lā ā:ñkā tōn:ñ?
2SG bird 2SG AGT NRL-SONG TOP be.good-SIM sing-PRES-DECL
tò-pákukā? [nàj ā-sá?] tfū lā]vCT [tā-són:ñ?]vCC
thus-CONCESS 2SG NRL-meat DIST TOP NZP-be.sour-DECL
'You, Bird – you sing songs very sweetly. Even though that is the case, your flesh
is sour.' (Coupe 2007: 366)

Copula clauses encode existence and possession. Copulas also serve as carriers for tense/mood marking when the identity or ascribed property of a referent requires a non-present temporal or modal specification (e.g. [22]).

- (21) *ātshā tòpàtī sà táj phāñjā ātsā tfū kū hūylī līkā tfhā.*
ā-tshā tè-pètī sà tāj phāñjā ā-tsā tfū kū
NRL-mithun NZP-be.big EMPHAT2 just five NRL-water DIST LOC
hūylī-ər lī-kā tfhā
stand-SEQ exist-SIM COP:PST
'There were five huge mithuns (*Bos frontalis*) standing in the water.'
- (22) *‘i, sítāk, mìpūñmāj sā? tfhājù? .xā.’* (Coupe 2007: 473)
i sítāk mà-pūñ-mā? -i sā? tfhā-i-ù? .xā
PROX be.correct NEG-be.good-CMPL-IRR PTCL COP-IRR-DECL PTCL
'This is right. Something not good at all will happen.'

The following two examples demonstrate the structure of monovalent and bivalent verbal clauses respectively. Example (23) additionally demonstrates the use of *-ñù* as a purposive nominalizer alongside its reanalysed function as a finite immediate future tense marker *-ñù*, as discussed in §6.4.

- (23) *tòpà? xāy! apā? ñnāt lā ãm à tsāk.ñù wā.ñù.*
tòpà? xāy! apā? ñnāt lā ãm
dear.one come-IMPER voc-father 1DU.INC TOP NRL-Aspidistra sp.
tsāk.-ñù wā.-ñù
pluck-PNMLZ go-IMM
'Come, dear one! We two will go to pluck Aspidistra (leaves).'
- (24) *ila mùñsān kin nā lùñsà lī phīñā tāmñtñj tfā? mñtñj.*
ilā mùñsān kin nā lùñsà lī phīñā
1PL.EXC Mongsen clan AGT Longsa land ABL
tāmñtñj tfā? mñtñj
NZP-NEG-consume-IRR nothing NEG-consume-PRES
'We Mongsen clans from Longsa country eat nothing that is not to be eaten' (i.e.
food that is taboo).

6.6 Converbs and clause linkage

Mongsen discourse is characterized by the absence of a native conjunction word class, this clause linkage function instead being accomplished by means of non-finite conversbs. The diachronic sources of conversbs in Mongsen mirror those of other languages of Eurasia, in that the vast majority have developed out of older nominalizing and case-marking morphology (Coupe 2006). Table 13.9 lists the ten forms of converb suffixes and their associated semantics.

TABLE 13.9 CONVERB SUFFIXES OF MONGSEN AO

Form	Gloss	Semantics
- <i>ə</i>	(SEQ)	sequential activity, anterior event, sometimes temporal qualification or reason
- <i>ākā</i> ; - <i>lī?</i> ; - <i>ā</i>	(SIM)	simultaneous activity, attendant circumstance, sometimes sequential activity or reason
- <i>lī</i>	(ALT.IT)	iterative alternation of two activities
- <i>pālā</i>	(COND)	irrealis conditional
- <i>kūlā</i>	(CIRCM)	circumstantial/hypothetical and realis conditional
- <i>pākūkā?</i>	(CONCESS)	concessive conditional
- <i>likā?</i>	(CONTEMP)	actualized contemporaneous event
- <i>tʃān</i>	(DUR)	durative activity
- <i>pānā</i>	(CAUSAL)	consequential circumstance
- <i>ku</i>	(LOC.CV)	temporal sequence, simultaneous activity/state

A distinction can be drawn between the clause chaining converb forms (principally the sequential converb, and to a much lesser extent the simultaneous converb), and the specialized converb forms that are used to express specific meanings generally correlating with adverbial presuppositions. Whereas the sequential converb is used to form multiple chains of non-final clauses, specialized converbs expressing presuppositional meanings such as reason, temporal overlap, condition, concession etc. are much more likely to occur in binary structures consisting of a non-finite converb clause and its matrix clause. Conditional and causal converb clauses exemplify this bi-clausal type of syntactic structure.

- (25) *āmī?* tsā tʃālū-pālā, ākhū tsā kā? tʃālū-tʃūk-ā!¹⁵
ā-mi? tsā **tʃālū-pālā** ā-khū tsā kā? tʃālū-tʃūk-ā!
 NRL-person DIST be.sick-COND NRL-tiger DIST also be.sick-PFV-PRES
 (Speaking of lycanthropy) 'If the person (i.e. the familiar) is sick, the tiger also becomes sick.'
- (26) *kijūŋā* à sā-pānā, thānī lā ilā áhlú nā mūwā-à?
kijūŋā à **sā-pānā** thānī lā ilā á-hlú nā
 neighbour one die-CAUSAL today TOP 1PL.EXC NRL-field all
 mā-wā-à-ù?
 NEG-GO-PRES-DECL
 'Because a neighbour died, today we are not going to the fields.'

A nominal demonstrative sometimes intervenes between the nominalized verb stem and the instrumental case marker of causal converbs. This betrays the nominal origin of many converb suffixes, which have developed a predicative function out of an older referential function. Most have grammaticalized from collocations involving the nominalizing suffix of a post-head relative clause and a case-marking clitic or other functional morphemes. The conditional converb is similar in originating from the general nominalizer -*pā?* and the topic marker *lā*, the converbal form of which has been reanalysed as -*pālā*, as demonstrated in examples (13) and (25).

The sequential converb overwhelmingly carries the heaviest functional load of all the converb suffixes and is used like a conjunction in better-known European languages. The following extended example (taken from Coupe 2006: 146) demonstrates the clause-chaining function of this verb form.

- (27) a. *tə̄.rə̄, tə̄nì khō̄ tə̄zātə̄nū tʃū, tū̄jā̄.l kī nā̄ n̄.rə̄ wā̄.rə̄ n̄.lūktʃā̄kī.ə̄,*
t̄-ə̄.r̄ **t̄-n̄l** **khō̄** **t̄-zā̄** **t̄-nū** **tʃū** **tū̄jā̄.l** **kī** **n̄.r̄**
thus-SEQ **RL-WIFE** **CONJ** **RL-CHILD** **RL-CHILD** **DIST** **other** **house** **ALL**
n̄.l-ə̄.r̄ **wā̄-ə̄.r̄** **n̄.lūktʃā̄kī-ə̄.r̄**
lead-SEQ **go-SEQ** **lead-INTO-RS-CAUS-SEQ**
- b. *īmt̄sā̄ŋpā̄? tʃū, ā̄jī tʃū tā̄tsə̄ŋj à̄ ă̄tʃā̄tłā̄kə̄,*
īmt̄sā̄ŋ-p̄? **tʃū** **ā̄-jī** **tʃū** **tā̄tsə̄ŋj** **à̄**
pers.name-M **DIST** **NRL-rice.beer** **DIST** **bamboo.container** **one**
ă̄tʃā̄tłā̄k-ə̄.r̄
squeeze-DESCEND-SEQ
- c. *tʃ̄m̄ə̄k̄ ă̄tʃūŋ kā̄?, kūsū kū ā̄m̄i m̄ə̄tūŋtʃā̄kə̄,*
t̄ʃ̄m̄-ə̄k̄ **ă̄-tʃūŋ** **kā̄?** **kūsū** **kū**
drink-SIM **NRL-shield** **also** **fireplace.shelf** **LOC**
ă̄-m̄i **m̄ə̄tūŋ-tʃā̄k-ə̄.r̄**
NRL-spear **be.ERECT-RS-SEQ**
- d. *ă̄tʃūŋ kā̄?, ā̄nūk tʃū kā̄? jū̄ŋsñ̄.l̄,*
ă̄-tʃūŋ **kā̄?** **ă̄-nūk** **tʃū** **kā̄?** **jū̄ŋsñ̄.l̄**
NRL-shield **also** **NRL-machete** **DIST** **also** **hone-RPET-SEQ**
- e. *h̄ə̄m̄ə̄tə̄ ă̄jī tʃū tʃ̄m̄ə̄k̄ pā̄ m̄ə̄nə̄.l̄?*
h̄ə̄m̄ə̄t-ə̄.r̄ **ă̄-jī** **tʃū** **tʃ̄m̄-ə̄k̄**
hold-SEQ **NRL-rice.beer** **DIST** **drink-SIM**
pā̄ **m̄ə̄n-ə̄.r̄** **l̄?**
3SG **SIT-SEQ** **wait-PST**

‘And after leading his wife and children into a neighbour’s house and leaving them there, Imtisangba squeezed some rice beer into a cup, drank, propped the shield and the spear against the fireplace shelf, the shield also, honed the machete and held [it] drinking the rice beer as he sat and waited.’

Example (27a) is additionally significant for demonstrating how a collocation involving the quotative particle *tə̄* and the sequential converb suffix *-ə̄.r̄* is used as a discourse connective *tə̄.r̄* to create linkage in extended narration.¹⁶ This is likely to have developed out of a tail-head linkage pattern, in which the matrix verb of a preceding clause is repeated in a recapitulation to establish a cohesive narrative structure. Example (28) illustrates this syntactic pattern.

- (28) a. *m̄āpā̄ŋ à̄ kū tə̄jā̄ à̄ ... ā̄m̄i? à̄ n̄s ā̄nū ā̄pā̄ŋtʃā̄ŋjā̄.zā̄ à̄ sù̄.*
m̄āpā̄ŋ **à̄** **kū** **t̄-jā̄** **à̄ ...** **ā̄-m̄i?** **à̄** **n̄s**
time **one** **LOC** **RL-mother** **one** **NRL-person** **one** **AGT**
ā̄-nū **ā̄pā̄ŋtʃā̄ŋjā̄.zā̄** **à̄** **sù̄-Ø**
NRL-child **male-ANMLZ-DIM** **one** **bear-PST**
‘once upon a time, a mother ... a person gave birth to a little boy.’
- b. *tə̄tʃā̄.l ā̄nū tʃū sù̄.l, sīn tʃū tə̄jā̄ tʃū sə̄tʃūk.*
t̄-t̄ʃā̄-ə̄.r̄ **ā̄-nū** **tʃū** **sù̄-ə̄.r̄** **sīn** **tʃū**
thus-do-SEQ **NRL-child** **DIST** **bear-SEQ** **BACK** **LOC**
t̄-jā̄ **tʃū** **sə̄-tʃūk-Ø**
RL-mother **DIST** **die-PFV-PST**
‘And then, the child having been born, after that the mother died.’

The collocation *tò-t/hà-ɔ̡.l*, consisting of the quotative particle, a generic verb meaning ‘do, make’ plus the sequential converb suffix functions very much like a clausal conjunction equivalent to ‘and’, and is similarly used for linking finite clauses. The discourse connective *ta̡.l* seen in the preceding example (27a) is a phonologically reduced form of this collocation.

ABBREVIATIONS

ALL	allative case
ANAPH	anaphoric demonstrative
ANMLZ	agentive nominalizer
ANT	anterior tense
BACK	‘back’ nascent postposition
BECAUSE	causal subordinating particle
CAUSAL	causal converb suffix
CMPL	completive aspect
COM	comitative case
CONCESS	concessive converb
COND	conditional converb suffix
CONJ	phrasal conjunction
COP	copula
DECL	declarative mood clitic
DESCEND	‘descend’ lexical suffix
DIM	diminutive suffix
DIST	distal demonstrative
DU	dual number
EMPHAT1	emphatic particle 1
EMPHAT2	emphatic particle 2
END	‘end’ lexical suffix
EXC	exclusive
F	feminine natural gender
IMM	immediate future
IMPER	imperative mood
IMPFV	imperfective aspect
INC	inclusive
INST	instrumental case
INTNS	intensifier suffix
INTO	‘into’ lexical suffix
IRR	irrealis mood
LNMLZ1	locative nominalizer 1
LNMLZ2	locative nominalizer 2
NMLZ	general nominalizer
NRL	non-relational noun prefix
NZP	nominalizing prefix
PFV	perfective aspect
PL	plural number
PNMLZ	purposive nominalizer
POSS/OBL	possessive/oblique pronoun
PROX	proximate demonstrative
PST	past tense

PTCL	illocutionary force particle
QPTCL	interrogative particle
REAL	realis mood
RL	relational noun prefix
RPET	repetitive aspect suffix
RS	resultant state marker
SEQ	sequential converb suffix
SEVER	‘sever’ lexical suffix
SG	singular number
SIDE	‘side’ nascent postposition
SIM	simultaneous converb suffix
TERM	terminative
TOGETHER	‘together’ lexical suffix
TOP	topic particle
VOC	vocative prefix

NOTES

- 1 Thanks are due to Guillaume Jacques, Randy J. LaPolla, Alexis Michaud, Nathan Straub, Amos Teo, T. Temsunungsang and Graham Thurgood for their comments and suggestions on an earlier draft. I alone am responsible for any remaining errors of analysis or fact. Research was made possible by a Singapore Government Ministry of Education Tier 2 grant (MOE2012-T2-1-100), and the paper was written during a visit at the Centre of Excellence for the Dynamics of Language (CoEDL) and the School of Literature, Languages and Linguistics at the Australian National University in May 2015. I am grateful to both of these institutions for their support.
- 2 See Lewis *et al.* 2016.
- 3 Chakpa (a.k.a. Jakpa) village has had separate Chungli- and Chang-speaking wards since it was founded in 1891. According to my consultants, this situation came about by a mutual arrangement between Chang and Chungli speakers to form a new village during the head-hunting era, the motivation for the amalgamation being that the respective kinship ties of the two wards would protect all the inhabitants from raids by other Ao and Chang villages.
- 4 Easton and Stebbins (2015) discuss a parallel example of a Christian mission’s language (Wedau) being promoted to the status of a lingua franca in Milne Bay Province, Papua New Guinea, by the act of reducing it to writing. As in the case of Chungli, the creation of an orthography by missionaries had a significant impact on extending the use of a particular language well beyond its traditional boundaries.
- 5 This is suggested by the observations of Hutton ([1987] 1929: iii–iv), who reported that the Chang village of Tuensang had only been established for approximately 11 generations in 1917. Founded by members of the Yimchungrü tribe from the south and Konyaks from the northeast, Tuensang village was also the source of offshoots that conquered and annexed their Phom, Sangtam and Ao neighbours.
- 6 The /High/ toneme is marked in phonemic transcriptions by an acute accent, the /Mid/ toneme is marked by a macron, and /Low/ toneme is signalled by a grave accent. Pitch is indicated by superscript tone numbers 1–5 in phonetic transcriptions (Chao 1930), with σ^{11} signalling the lowest and σ^{55} the highest pitch level. Tone is discussed in greater detail in §4.
- 7 The form of the Chang ergative/instrumental marker is $e\bar{j}^{55}$.
- 8 This area arguably extends along Barail Range of southern Assam to the eastern flank of the Jowai Plateau, as the Zeme language of Dima Hasao District (formerly the North Cachar Hills) has also traditionally been included in the so-called ‘Naga’ grouping.

- 9 The only exception to this appears to be the relational and non-relational prefixes occurring on bound noun roots. In many cases, these appear to copy their tone from the initial syllable of the root; in others, they appear to have lexicalized a contrastive tone. See Coupe (2003: 23–4) for further discussion of their tonal status.
- 10 The relationship of tone to intonation in Ao correlates closely with Chao's (1968: 39) 'small ripples riding on large waves' metaphor.
- 11 This morpheme also serves as a linking device in compounded numerical expressions, as demonstrated in (2).
- 12 Such a usage is not reported for 'do, make' in Heine and Kuteva (2002: 117–20).
- 13 This Waromung Village Mongsen example demonstrates the *i* > *u* sound change affecting some high-frequency morphemes in that variety (see §3).
- 14 The only marginal exception to this applies to some bivalent verbs of surface contact, e.g. *ják* 'beat'; verbs belonging to this semantic class require their O arguments to take locative case marking.
- 15 This example is taken from a text narrated in the Waromung Village variety, which has the form *tsə* for the distal demonstrative.
- 16 Mangmetong Village Mongsen speakers often re-syllabify the collocation as two syllables, thus *ta.jə* in this example.

REFERENCES

- Barkman, Tiffany. 2014. A descriptive grammar of Jejara (Para Naga). Unpublished MA thesis, Payap University, Chiang Mai.
- Beames, John. 2012 [1875]. *Comparative Grammar of the Modern Aryan Languages of India, to Wit, Hindi, Panjabi, Sindhi, Gujarati, Marathi, Oriya, and Bangali*. Volume 2: *The Noun and the Pronoun*. Cambridge: Cambridge University Press. First published 1875.
- Benedict, Paul K. 1972. *Sino-Tibetan: A Conspectus* (Contributing ed. James A. Matisoff). Cambridge: Cambridge University Press.
- Bradley, David. 1997. Tibeto-Burman languages and classification. In: Bradley, David (ed.), *Tibeto-Burman Languages of the Himalayas, Papers in South East Asian Linguistics* 14. Canberra: Pacific Linguistics, 1–71.
- Bruhn, Daniel W. 2014. A phonological reconstruction of Proto-Central Naga. Unpublished PhD dissertation, University of California, Berkeley.
- Burling, Robbins. 1983. The Sal languages. *Linguistics of the Tibeto-Burman Area* 7.2:1–32.
- Burling, Robbins. 2003. The Tibeto-Burman languages of northeastern India. In: Thurgood, Graham and Randy J. LaPolla (eds) *The Sino-Tibetan Languages* (1st edn.). London/ New York: Routledge, 169–191.
- Chao, Yuen Ren. 1930. A system of 'tone-letters'. *Le Maître Phonétique* 30: 24–27.
- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Chelliah, Shobana L. 2009. Semantic role to new information in Meithei. In: Barðdal, Jóhanna and Shobhana L. Chelliah (eds) *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case*. Amsterdam/Philadelphia: John Benjamins, 377–400.
- Chelliah, Shobana L. and Gwendolyn Hyslop. (eds) 2011. Special issue on optional agentive marking in Tibeto-Burman, Part I. *Linguistics of the Tibeto-Burman Area* 34.2.
- Chelliah, Shobana L. and Gwendolyn Hyslop. (eds) 2012. Special issue on optional agentive marking in Tibeto-Burman, Part II. *Linguistics of the Tibeto-Burman Area* 35.1.

- Coupe, Alexander R. 2003. *A Phonetic and Phonological Description of Ao: A Tibeto-Burman Language of Nagaland, North-East India*. Canberra: Pacific Linguistics.
- Coupe, Alexander R. 2006. Converbs. In: Brown, Keith (ed.) *Encyclopedia of Languages and Linguistics* (2nd edn.) Oxford: Elsevier, 145–152.
- Coupe, Alexander R. 2007. *A Grammar of Mongsen Ao*. Mouton Grammar Library 39. Berlin and New York: Mouton de Gruyter.
- Coupe, Alexander R. 2011a. Pragmatic foundations of transitivity in Ao. *Studies in Language* 35.3: 492–522.
- Coupe, Alexander R. 2011b. On core case marking patterns in two Tibeto-Burman languages of Nagaland. Special issue on optional agentive marking in Tibeto-Burman, edited by Shobana L. Chelliah and Gwendolyn Hyslop, *Linguistics of the Tibeto-Burman Area* 34.2: 21–47.
- Coupe, Alexander R. 2012. Overcounting numeral systems and their relevance to subgrouping in the Tibeto-Burman languages of Nagaland. Special issue on Sino-Tibetan comparative studies, *Language and Linguistics* 13.1: 193–220.
- Coupe, Alexander R. 2013. Tense, but in the mood: diachronic perspectives on the representation of time in Ao. *Language and Linguistics* 16.6: 1105–1113.
- Coupe, Alexander R. 2014. Strategies for analyzing tone languages. *Language Documentation & Conservation* 8: 462–489.
- Easton, Catherine L. and Tonya N. Stebbins. 2015. On becoming an object of study: legitimization in the discipline of linguistics. In: De Busser, Rik and Randy J. LaPolla (eds) *Language Structure and Environment: Social, Cultural, and Natural Factors*. Amsterdam/Philadelphia: John Benjamins, 319–351.
- Heine, Bernd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd, Ulrike Claudi and Friederike Hünnemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.
- Hutton, John H. 1987 [1929]. *Chang Language: Grammar and Vocabulary of the Language of the Chang Naga Tribe*. Delhi: Gian Publishing House. First published 1929 (revised and edited by Satkari Mukhopadhyay).
- LaPolla, Randy J. 1995. On the utility of the concepts of markedness and prototypes in understanding the development of morphological systems. *Bulletin of the Institute of History and Philology* 66.4: 1149–1185.
- Lewis, M. Paul, Gary F. Simons and Charles D. Fennig (eds). 2016. *Ethnologue: Languages of the World* (19th edn.). Dallas, TX: SIL International. Available at: www.ethnologue.com, accessed 18 August 2016.
- Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley: University of California Press.
- Mills, James P. 1926. *The Ao Nagas*. London: Macmillan.
- Reinöhl, Uta. 2016. *Grammaticalization and the Rise of Configurationality in Indo-Aryan*. Oxford: Oxford University Press.
- Shafer, Robert. 1950. The Naga branches of Kukish: vocalism. *Rocznik Orientalistyczny* 16: 467–530.
- Shafer, Robert. 1955. Classification of the Sino-Tibetan languages. *Word* 11: 94–111.
- Temsunungsang, T. 2009. Aspects of the prosodic phonology of Ao: an interdialectal study. Unpublished PhD dissertation, School of Language Sciences, English and Foreign Languages University, Hyderabad.
- Vong, Tsuh Shi. 2009. Discourse studies of Makuri Naga narratives. MA thesis, Payap University, Chiang Mai.

CHAPTER FOURTEEN

KARBI

Linda Konnerth

1 INTRODUCTION

Karbi is spoken by around half a million people¹ in the Karbi Anglong District, Assam, Northeast India, as well as adjacent areas. The geographical spread of the Karbi-speaking area is very large in the regional context, extending westwards into present-day Bangladesh and in the north into Arunachal Pradesh. To the south and east, it appears to be mostly restricted within Assam. In historical times, the Karbi-speaking area also covered parts of the Garo Hills and the Khasi Hills in what is now Meghalaya, and to the east extended into present-day Burma (Walker 1925; Grierson 1904).

Karbi society is organized into five major clans, each of which is further divided into various subclans. Clan membership is passed down from the father to the children. While a large portion of the population now lives in towns, village life is still centered on agriculture and mostly wet rice cultivation. Some remote villages still practice slash and burn agriculture.

There are two major dialect groups, which can be roughly distinguished as the varieties spoken in the hills and the ones spoken in the plains. The hills varieties are spoken by a much higher percentage of the population across both the eastern and western parts of the Karbi Anglong District. The plains varieties are spoken particularly in the Assam plains as well as in the Assam–Meghalaya border area. The plains varieties are also sometimes referred to by the cover terms Dumra, Kamrup, or, highly controversially, ‘Amri Karbi.’²

Linguistic resources on Karbi (mostly the hill varieties) include a comprehensive description of the phonology and morphology of the language by Grüßner (1978), a sketch grammar by Jeyapaul (1987) from the Central Institute for Indian Languages, and a full descriptive grammar by the present author (Konnerth 2014a). Dictionaries include an unpublished manuscript by Grüßner (not dated) focused on the hills varieties and with tones marked, as well as a dictionary focused on the plains varieties, which also gives hills variety forms, by Teron and Tumung (2007), with no tones marked.

Regarding the phylogenetic affiliation of Karbi within Tibeto-Burman, it is inspectionally obvious that Karbi is more closely related to Kuki-Chin and the “Naga” groups than to any other branches in Northeast India. There are both apparent cognates as well as morphological evidence (see §§3, 4) that currently point in this direction. However, the exact position of Karbi vis-à-vis those closely related languages remains obscure, and this has likely to do with prolonged historical language contact with the neighboring Austroasiatic Khasian languages and especially Pnar. Preliminary evidence for this contact scenario can be found in Grüßner (1978) and Joseph (2009) and is backed by historical accounts of the impact of the historically dominant Pnar kingdom on the Karbi population. One possible hypothesis is to link certain aspects of Karbi grammar to creolization (in the sense of DeLancey 2012, 2013, 2014) as the result of this contact scenario. Specifically, what we find is a relatively higher prominence of pragmatic flexibility compared to syntactic rigidity in the grammatical domains of word classes (§3.1), role-marking of

NPs (§5.1.2), and relativization of any clausal participant (§5.2), as well as the simple universal nominalization construction with *ke-* for all grammatical purposes (§5.2), and the universal character of head noun marking with *a-* for any kind of preposed modification (§4.1).

2 PHONOLOGY

2.1 Consonants

Consonants in onset position are shown in Table 14.1. The system includes stops at four places of articulation, fricatives at three places of articulation, two nasals, /l/ and /r/. Notably not included are the voiced velar stop or the velar nasal (the latter does occur in coda position, however). <Angle brackets> indicate orthographic representation.

Palatal /j~dʒ/ constitutes free variation between a stop and a glide production. Also, allophonic alternations typical for the area include /p^h~ɸ/ (within the same speaker) and /r~r̥~ɹ̥/ (intergenerational and interdialectal).

Consonants in coda position include unreleased stops and nasals at three places of articulation as well as the rhotic (but not /l/; Table 14.2).

2.2 Vowels

Vowel phonemes in the standard Rongkhang and Chinthong Hills Karbi varieties include the five cardinal vowels as well as three diphthongs, /ai/, /oi/, and /ui/ (Table 14.3). (Note that the diphthongs can alternatively be analyzed as monophthongs with a glide coda.) In the Amri Hills Karbi dialect (not to be confused with the plains variety sometimes referred to as ‘Amri’), there are two additional vowel phonemes, the monophthong /ɪ/ (e.g. *rī* ‘base of tree’ vs *rī* ‘rope’) and the diphthong /ei/ (e.g. *ingvèi* ‘to fly around’ vs *ingvài* ‘to choose’). These are merged with /i/ and /ai/ respectively in the standard varieties.

TABLE 14.1 SYLLABLE-INITIAL CONSONANTS

	Bilabial		Alveolar			Palatal		Velar		Glottal	
Stops	b	p	p ^h ~ɸ	<ph>	d	t	t ^h ~θ	<th>	j~j	c	<ch>
Fricatives	β~w	<v>			s						h
Nasals	m				n						
Lateral					l						
Rhotic					r~r̥						
Glide							j~j				
							<j>				

TABLE 14.2 SYLLABLE-FINAL CONSONANTS

	Bilabial	Alveolar	Palatal	Velar
Stops	p~(<p>)	t~(<t>)		k~(<k>)
Nasals	m	n		ŋ~(<ng>)
Rhotic		r~r̥~ɹ̥		

TABLE 14.3 VOWEL PHONEMES

Monophthongs	/i/	(/ɪ/)	/e/ (/ɛi/)	/a/ (/ai/)	/o/ (/oi/)	/u/ (/ui/)
Diphthongs						

2.3 Syllable structure

Karbi syllables may be open and follow the structure (C)(C)V(V), or they may be closed and follow the structure (C)(C)VC. Possible onset consonant cluster combinations are as follows: /pl,pr,pʰl,pʰr,tʰr,kl,kr,kʰr/. Note that /tʰr/ and /kʰr/ are marginal, occurring only in two words and one word, respectively.

Closed syllables may either have a sonorant coda, i.e. a nasal /m,n,ŋ/ or rhotic /r/, or may be checked and have an unreleased stop coda /p,t,k/.

2.4 Tone system with a low functional load

Karbi has a considerable number of tone minimal pairs (and some minimal triplets), which fall into three categories: ‘low’ (L), ‘mid’ (M), and ‘high’ (H). While L and H tones are indeed distinguished by pitch height, the pitch range of the M tone is variable. However, the M tone is associated with glottalization, which distinguishes it from the L and H tones. All three tones occur on syllables of any type (open, sonorant-final, checked).

Minimal triplets for non-checked monosyllabic roots include *phi* (L) ‘grandmother,’ *phī* (M) ‘to roast,’ *phí* (H) ‘to give birth,’ and *rōng* ‘village,’ *rōng* ‘plant,’ *róng* ‘to borrow.’ For checked syllables, we find *hūt* ‘to dig (a small hole),’ *ahūt* ‘during’ (<-*hūt* ‘time’), *hút* ‘to question/examine a wrongdoer.’ A minimal triplet for disyllabic roots is *ingthī* ‘to kill a louse,’ *ingthī* ‘to wash/rinse (an object),’ *ingthí* ‘to wash/clean (head, hair).’

Despite the existence of such minimal triplets and many more minimal pairs, the status of the tone system is problematic in a number of respects due to its low functional load. One is that even after extensive training, native speaker consultants cannot consistently compare and identify tones, not even on all monosyllabic roots. Furthermore, there is an intriguing mismatch between tone production and perception: an acoustic study showed that there is variability in tone production between speakers such that pitch height differences between the M and H categories are statistically significant in one speaker but not another speaker. A follow-up perception study showed even greater variability in listeners, and surprisingly demonstrated that those stimuli in which M and H categories were consistently distinguished by pitch height did not help listeners in tone category identification. Large error rates in tone category identification suggest a weak status of the M/H tone distinction, with interesting implications for the phonological representation of the three ‘tones’ in Karbi (Konnerth and Teo 2014). Another characteristic of the Karbi tone system is that prosody can layer over lexical tone such that the tone categories are obscured.

These are some of the characteristics that suggest that the tone system does not carry a high functional load. Among the factors that may be enabling the tone system to have a low functional load is the agglutinative morphological profile of Karbi and specifically the suffixal class of predicate derivations (§4.2.2.1). Since the verb stem often consists of a root plus a predicate derivation in conventionalized collocations, the meaning of the root becomes uniquely identifiable independent of its tone.

Despite the low functional load of the tone system, there are some morphophonemic alternations based on tone. For example, the reflexive/reciprocal *che-* and the

auto-benefactive/malefactive *cho-* trigger tone changes in the root such that a low tone turns into a mid tone (e.g. *dām* ‘go’ > *che-dām* ‘RR-go’) and a mid tone into a high tone (e.g. *kūp* ‘cover’ > *che-kúp* ‘RR-cover’).

2.5 Stress

The stress pattern in Karbi is generally iambic as is typical for southeastern Tibeto-Burman languages (LaPolla, Chapter 2, this volume). There is a tendency in Karbi for unstressed syllables to be toneless, but this is not a systematic feature: prefixes are always toneless, whereas suffixes do carry tone, even the highly grammatical and always unstressed ones such as *-lò* ‘realis,’ *-pò* ‘irrealis1,’ and *-ji* ‘irrealis2.’ Derivational suffixes, which are semantically rich, are the stressed syllable in the new verb stem they create, due to the iambic stress pattern. Compare *lè-lè* ‘reach-NEG’ (where the stress is on the suffix and indicated by underlining) or *dām-bōm* ‘go-CONT’ with *nè=tā* ‘1EXCL=ADD,’ where the first person exclusive pronoun is stressed and not the additive enclitic.

Karbi has sesqui-syllabic features in the form of coda-less and toneless prefixes with vowel allomorphy (see §4.2.1).

3 WORD CLASSES

3.1 Major word classes

Lexical roots can be divided into the two categories of nouns and verbs. Noun roots function as noun stems without any further marking, and as noun stems can take the possessive prefix *a-*, e.g. *a-hēm* ‘POSS-house.’ Verb roots function as verb stems without any further marking, and as verb stems can take the nominalizing prefix *ke-*, e.g. *ke-dām* ‘NMLZ-go.’ Property-concept terms (i.e. the semantic class of elements corresponding to adjectives in languages that have them) can also take the *ke-* nominalizer and are best analyzed as a subgroup of verbs, e.g. *ke-thè* ‘NMLZ-be.big.’ Major subclasses of nouns include classifiers and relator nouns. Classifiers occur in classifier-numeral words which can then be used to modify nouns (§5.1.1). In this construction, the numeral follows the classifier (with the exception of ‘one,’ which precedes it). Classifiers are a subclass of nouns as they may also occur with *a-* ‘possessive.’ Relator nouns are bound morphemes that typically occur with *a-* ‘possessive’ but in some instances also occur with personal possessive prefixes.

While the majority of lexical roots are monosyllabic, Karbi has a large number of disyllables that start with either *ing-* or *ar-*, which appear to be fossilized reflexes of Proto-Tibeto-Burman (PTB) **m-* and **r-*, respectively (Wolfenden 1929; Benedict 1972: 109–10; Matisoff 2003: 127–9).

What is important to note is that while predicates are most often formed from verbs, it is quite striking that apparently any element can function as a predicate, e.g. the interrogative pronoun *pí* ‘what’ in (2) later (§4.2). While this is a common feature of Sino-Tibetan languages further to the east, within Northeast India it is restricted to the larger languages (both in terms of speakers and geographic spread) that are mostly found in the valley, e.g. Bodo. Figure 14.1 illustrates the flexibility within the major word classes of nouns, verbs, and property-concept terms (PCTs) with respect to the general clausal functions of reference, modification, and predication, following Croft’s (2001: 99) model. What this shows is that any of the three major word (sub-)classes can function as predicates without requiring any derivational marking. Furthermore, nouns occur in reference, modification,

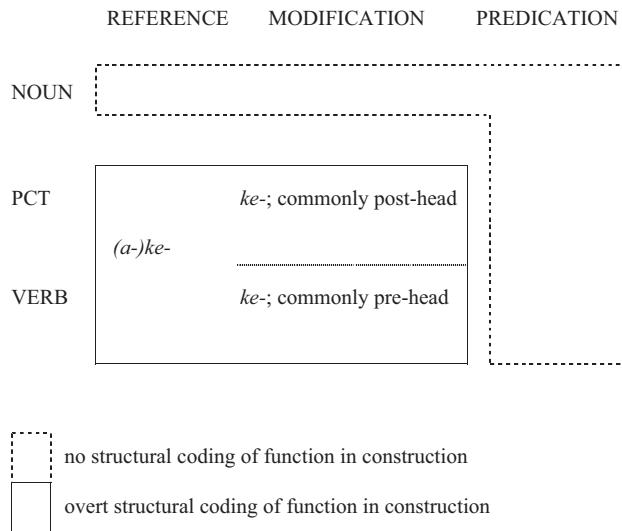


FIGURE 14.1 SEMANTIC MAP OF BASIC CLAUSAL FUNCTIONS OF KARBI PARTS OF SPEECH (FOLLOWING CROFT'S (2001: 99) MODEL)

and predication constructions without requiring any structural coding. Hence word classes are handled in more pragmatically based terms rather than being syntactically strictly structured.

3.2 Closed word classes

3.2.1 Pronouns

Karbi distinguishes between first person exclusive and inclusive forms; the exclusive form also acts as the singular form of the pronoun (Table 14.4). Plural forms of the pronouns feature the plural noun *-tūm* (see also §4.1 and §5.1.1). Honorific forms carry *-li*. Possessive prefixes are of the same form as personal pronouns, with the exception of the 3rd person prefix, where both the pronoun form *alang(li)-* and the possessive *a-* (§4.1) can be used.

3.2.2 Copulas

Copulas include the positive existential *dō*, the negative existential *avē*, and the negative equational copula *kalī*. The positive equational construction is simply the juxtaposition of the two noun phrases; no copula is used.

4 MORPHOLOGY

4.1 Nominal morphology

There is little nominal morphology in Karbi, and some constructions are even typical of an isolating morphological profile despite the highly agglutinative nature of the verb

TABLE 14.4 PERSONAL PRONOUNS AND PERSONAL POSSESSIVE PREFIXES

Personal pronoun (honorific)	Gloss	Possessive prefixes (honorific)
<i>nè(li)</i>	'1EXCL'	<i>ne(li)-</i>
<i>e(li)-tūm</i> (only plural)	'1INCL'	<i>e(li)-</i>
<i>nāng(li)</i>	'2'	<i>nāng(li)-</i>
<i>alāng(li)</i>	'3'	{ <i>alāng(li)-</i> ; <i>a-</i> }

complex. For example, productive plural marking occurs via a periphrastic plural noun construction rather than affixation: *Karbi a-tūm* ‘Karbi POSS-PL > Karbis.’

The periphrastic plural is based on the possessive construction that involves *a-* ‘possessive’ attaching to the possessed noun, schematically [N_{POSSE}] [**a**-N_{POSD}]. In addition to plural marking, a number of syntactic categories and grammatical functions originate in this possessive construction, including relator nouns and relator noun-derived subordinators and adverbs, as well as the emphatic *-māt*; the noun phrase delimiter *abāng* (§5.1.3); role markers *-phān* and *-lōng* (§5.1.2); and restrictive focus *-nāt*. Example (1) shows various instances of this construction within a single sentence.

- (1) [...] *Naka akhei atum aphan adunghétpen*
 [[[Naká a-khéi a-tūm] a-phān] a-dūng-hét=pen
 PN POSS-community POSS-PL POSS-NSUBJ POSS-near-firmly=from
kethedamlong apot ning ingsam'o, neli
 ke-thèk-dām-lōng] a-pót] nīng ingsām-ò nè-lì
 NMLZ-see-GO-GET POSS-because mind be.cold-much 1EXCL-HON
 ‘[...] because I could see the Naga tribes from very near, I was very happy’ [SiT,
 HF 058]

In *Naka akhei atum*, the *a-* on *khéi* represents the synchronic possessive construction; the second occurrence of *a-* represents the plural construction with *-tūm*. The third *a-* occurs on the non-subject marker *-phān*. In the adverbial *adunghetpen* ‘from very near,’ *a-* has become fossilized on the relator noun *-dūng* ‘near.’ Finally, the subordinator *apot* ‘because,’ must also be derived from a relator noun, to explain the etymological *a-* prefix here, which is generally found in subordinators (see also §5.2).

4.2 Predicate morphology

Predicate morphology can reach extensive agglutinating structures as illustrated by words such as *pinepinanedetjima* in (2).

- (2) *te mo pinepinanedetjima ko jirpo pu*
 te mò pí-nē~pinā-Cē-dēt-jí=ma ko jīrpō pu
 therefore future what-INDEF~DISTR.PL-NEG-PFV-IRR2=Q buddy:voc friend QUOT
 ‘and there won’t be any (difficulties/problems/dangers), my friend?’ [HK, TR 140]

What we see here is the interrogative pronoun *pí* ‘what’ as the root, followed by the indefinite marker *-nē*. Together they form a new stem which subsequently is quasi-reduplicated with a typical vowel change resulting in ~*pinā* to indicate distributive plurality. This new stem *pinepina* is followed by the negative suffix *-Cē*, which reduplicates

the onset of the last syllable of the stem. Then there are two inflectional-type suffixes *-dèt* ‘perfective’ and *-jí* ‘irrealis2,’ followed by the question marker *=ma*.

Another morphologically complex example is (3). Here, two preverbal slots are filled with *nang*= ‘cislative’ and *che-* ‘reflexive/reciprocal’ (see §4.2.1). Postverbally, there are two predicate derivation suffixes (*-mék* ‘in advance’ and *-èt* ‘all:s/o’), the negative suffix *-Cē*, perfective *-dèt*, and realis *-lò* (see §4.2.2).

- (3) [...] *adappen hadakpen nangchesikmek 'et' edetlo*
 a-dàp=pen hádák=pen nang=che-sík-mék-èt-Cē-dèt-lò
 poss-morning=from there=from CIS=RR-prepare-in.advance-all:s/o-NEG-PFV-RL
 '[...] from the morning from there we hadn't prepared it (well)' [SH, CSM 062]

4.2.1 Preverbal morphology

Karbi preverbal morphology is limited to one proclitic (with two variants) and four prefixes spread across four position classes altogether. The proclitic is *nang*=, which indexes non-subject speech-act participants (optionally alternating with *ne*= and *e*= for specific indexation of non-subject first person exclusive and inclusive participants, respectively), as well as marking the *cislative* (see Konnerth 2015). As a person marker, *nang*= (*/ne=e*) most commonly indexes first/second person primary objects, but its use is not entirely obligatory, and besides primary objects, participants in other non-subject roles may get indexed. There is no other person marking in Karbi.

The four prefixes include the nominalizer *ke-* (~ *ki-* ~ *ka-*),³ which has further grammaticalized to mark the imperfective in main clauses (§5.2); the causative *pe~pa*⁴ (e.g. in (23),(26));⁵ the reflexive/reciprocal *che-* (e.g. in (3), (10), (23)–(26); see §4.3); and the auto-benefactive/malefactive *cho-* (e.g. in (6)). These prefixes go in three slots: *che-* and *cho-* cannot co-occur. With respect to the ordering of the slots, *ke-* is always first, but between *pe~pa* and *che-* or *cho-*, either one can go first with resulting differences in scope.

4.2.2 Postverbal morphology

There are five morphological slots following the root (or compound stem), while multiple elements from the highly diverse first slot (‘predicate derivations’) after the root can co-occur (Figure 14.2).

4.2.2.1 Predicate derivations

The class of predicate derivations⁶ contains a very large number of suffixes (more than 200). Their productivity in terms of which stems each suffix may occur with varies. Based on their function, most suffixes can be sorted into the following categories: manner derivations; result derivations; direction, (associated) motion, and path derivations; derivations

ROOT (/stem)	DERIV* (incl. modals)	RDPL	NEG	ASP	mood, aspect, subordinate, non-declar. speech act markers
-----------------	--------------------------	------	-----	-----	--

FIGURE 14.2 POSTVERBAL MORPHOLOGY

that modify or highlight arguments and/or argument structure; and aspect/aktionsart and time derivations.

In (4), the result derivation *-ràk* ‘RES:little.wound’ is illustrated.

- (4) [...] “*o bang voarbipi akam kechomathale neno*
[o bàng vōarbípī a-kám ke-chomathā=le
VOC CLF:HUM:PL bird.sp POSS-work NMLZ-think.with.bad.intentions=FOC:IRR
nanglutchock nangarkerakrakdetkema?”
ne-nò nang=lüt-chòk nang=arkè-ràk~ràk-dèt=ke=ma]
my-ear CIS=enter-disappearing CIS=scratch-RES:little.wound~DISTR.PL-PFV=TOP=Q
“O Voarbipi, what were you thinking, coming into my ears and scratching **and**
wounding me!?” [RBT, ChM 034]

Examples (5) and (6) illustrate the subclass of suffixes that function to modify or classify clausal participants, i.e. bearing some resemblance to what Peterson refers to as verbal classifiers (Peterson 2008). In (5), *-vàng* ‘plural:S/A’ indexes subject plurality. Interestingly, *-vàng* is a negative polarity item, which only occurs in combination with negative *-Cē*. It is homophonous with the lexical verb *vàng* ‘come.’

- (5) *pisi kithurvangvedetma {mm}*
pīsi ke-thùr-vàng-Cē-dèt=ma mm
why NMLZ-get.up-PL:S/A-NEG-PFV=Q AFF
‘why didn’t you (pl) get up?’ [HK, TR 154]

In (6), *-bòr* indicates the smallish appearance of a clausal participant, which for *-bòr* is always the S argument.

- (6) *hongpharla si kachodangsi*
hòng-pharlá=si sá kV-cho-dáng-si
outside-outside.part.Karbi.house=FOC tea(<Ind) NMLZ-MID-put.on.stove-NF:RL
totborlo apenan abangke
tót-bòr-lò a-penàn abàng=ke
squat-appearing.small:S-RL POSS-husband NPDL=TOP
‘while she was getting ready, he didn’t have anything to answer and quietly on the porch
he prepared himself tea and sat there very smallish, the husband’ [SeT, MTN 040]

As suggested by the large size of this class of suffixes, new members are apparently recruited fairly easily. One such pathway is verb serialization as illustrated with *-dàm~dām* ‘go’ in (7), which has grammaticalized from *dàm* ‘go.’ In this example, *-dàm~dām* ‘go’ marks the translocative, as the narrative viewpoint is a tree branch where the two protagonists are sitting and from where one of them is falling down as described here.

- (7) *latum achitimsi klodamduplo [...]*
là-tūm a-chitim=si kló-dàm-dùp-lò
this-PL POSS-half=FOC:RL fall-go-falling.sound.from.high.solid.obj-RL
‘he fell down right in the middle of them [...]’ [HK, TR 189]

4.2.2.2 Reduplication in the verbal complex

Reduplication of the verb stem occurs either as full reduplication of the immediately preceding syllable, or involves a vowel change following a phonological pattern.

Functionally, verbal reduplication may indicate a habitual reading or an iterative reading, plurality of an argument, or intensification. In (8), the reduplication of the last syllable of the stem *ke-chō-dūn* ‘IMPF-eat-JOIN’ indicates habituality.

- (8) *netumta nangtum nangpi po longle thak*

[ne-tūm=tā	nang-tūm	nang=pī-pō	longlē	athāk
1:EXCL-PL=ADD	2-PL	1/2:NSUBJ=give-IRR1	earth	on.top

nangbokchom titisi, neta

nang=bók-chōm	titī-si]	nè=tā
1/2:NSUBJ=serve.small.items-a.little	habitually-NF:RL	1:EXCL=ADD

nangkechodundun

nang=ke-chō-dūn~dūn

CIS=NMLZ-eat-JOIN~HAB

‘to us you would also give us, on the ground always you would serve us (food), I also **used to** eat like that’ [KK, BMS 060]

4.2.2.3 Negation

The negative suffix *-Cē* repeats the full onset of the last syllable of the verb stem. In negation, *lē* ‘reach’ becomes *lē-lē*; *krōi* ‘agree’ becomes *krōi-krē*; *ingtōn* ‘conclude’ becomes *ingtōn-tē*. Vowel-initial syllables phonetically begin with a glottal stop, which is accordingly repeated in the negative suffix as well as in the following complex stem that involves a root/stem and a suffix: *chō-ò* [*chō-?ò*] ‘eat-much’ becomes *chō-ò-ē* [*chō-?ò-?ē*].

4.2.2.4 Prefinal aspect position class

Two aspectual suffixes share one slot: the frequently occurring exhaustive perfective *-dēt* as well as the quite rarely used experiential *-pín*. Perfective *-dēt* allows negative *-Cē* to co-occur with the irrealis suffixes *-pō* and *-jí* (see §4.2.2.5), which is otherwise not possible. There is an exhaustive semantic component to it such that *chō-dēt* ‘eat-PFV’ regularly implies that no food is left afterwards.

4.2.2.5 Final position class

The final position class contains elements that belong to a variety of functional categories. Included are a set of three mood markers: *-lō* ‘realis,’ *-pō* ‘irrealis1,’ *-jí* ‘irrealis2’; subordination markers such as *-si* ‘non-final:realis,’ *-ra* ‘non-final:irrealis,’ and the general non-finite *-pen* (§5.4), as well as conditional *-te* ‘if’; an aspectual marker *-lāng* ‘still’; and finally, non-declarative speech act suffixes including imperative and hortative markers (§5.3).

4.2.3 Realis vs irrealis status of the main predicate

There are three sets of markers that indicate, or are sensitive to, the realis or irrealis status of the predicate. The first set of *-lō* ‘realis,’ *-pō* ‘irrealis1,’ and *-jí* ‘irrealis2’ directly marks the (ir)realis status of the main predicate. The other two sets of markers do not directly mark (ir)realis but instead occur on elements other than the main predicate and are sensitive to the (ir)realis status of the main predicate. They include on the one hand non-final

markers *-si* ‘non-final:realis’ and *-ra* ‘non-final:irrealis,’ which occur on non-final verbs in the clause-chaining construction (§5.4), and on the other hand the focus enclitics *=si* ‘focus:realis’ and *=le* ‘focus:irrealis’ as they occur on noun phrases (§5.1.4). What (ir)realis-sensitivity means for these two categories of non-final and focus markers is that certain types of irrealis status (including, respectively, some of the typical ones such as negative, subordinate, or non-declarative) require the irrealis version of the non-final or focus markers, whereas the realis version occurs otherwise in what can be considered the unmarked scenario.

4.3 Reflexive/reciprocal marking

Reflexives and reciprocals occur in the same construction. It consists of the reflexive/reciprocal prefix *che-* on the verb, and optionally the reflexive/reciprocal pronoun (i.e. adding personal possessive prefixes to the reflexive/reciprocal nouns *-metháng* or, less commonly, *-mená* ‘self’). Reflexives are not limited to situation of identity between A and O arguments in Karbi, but also hold in the case of identity between A and the possessor of O (or another core argument, such as a goal in a motion event, e.g. *a-hém che-dám* ‘poss-house RR-go’ > ‘go (to one’s own) home’). In addition, *che-* marking occurs in some middle voice contexts.

5 SYNTAX

5.1 Noun phrase

5.1.1 Elements inside the NP

The categories of elements inside the Karbi NP are given in Figure 14.3. There are two types of elements that always occur before the head noun: demonstratives and possessor NPs. Three types of elements can go on either side of the head noun: enumeration constructions, relative clauses (RCs), and PCT modifiers (see §3.1). Finally, the plural marker (see §4.1) always occurs last.⁷

Figure 14.3 shows that preceding the head noun, there are four slots; the dotted line between the categories of enumeration constructions and RCs or PCT modifiers indicates that the relative ordering between the two is not fixed. While RCs and PCT modifiers can occur on either side of the head noun, RCs almost exclusively occur preposed, while PCT modifiers typically occur postposed. In those marked cases where PCT modifiers occur before their head noun, there is a greater focus on the PCT modifier.

Enumeration constructions include a number of different structural strategies of indicating number, which typically involve classifiers. In addition, there are constructions

(DEM)	(NUM)	(RC) (PCT modifier)	([NP]POSR)	HEAD NOUN	(PCT modifier) (RC) (NUM)	(PL)
-------	-------	------------------------	------------	----------------------	---------------------------------	------

FIGURE 14.3 KARBI NOUN PHRASE STRUCTURE

related to but going beyond enumeration, particularly constructions based on the numeral ‘one’ (such as indefiniteness marking or indicating ‘all’ or ‘same’).

The presence of any preposed modifier requires the head noun to carry *a-* ‘possessive’ (§4.1). Thus, even just a noun phrase consisting of a demonstrative and a head noun (HN) involves *a-* marking, e.g. *lasō a-monít* ‘this poss-person.’

Examples (9) and (10) illustrate noun phrases with the structures [NUM]-[HN]-[PCT modifier] and [HN]-[PCT modifier]-[PL], respectively. In (9), the head noun *khabor* ‘news’ is marked by *a-* because of the preceding classifier-numeral word, whereas in (10), there is no preceding modifier and hence the head noun *ovè* ‘generation’ is not marked by *a-*.

- (9) “[...] *pinike ne eson akhobor mesen arjulong*”
 [pini=ke nè [e-sòn a-khobór mē-sén] arjū-löng]
 today=TOP 1EXCL one-CLF:thing POSS-news(<Ind) be.good-INTENS hear-GET
 “[...] today I got good news” [HK, TR 132]

- (10) [...] *setame ove kimi atum chethanlong Bokolapo*
 setämē ovè ke-mī a-tūm che-thán-löng Bokolā-pō
 nevertheless generation NMLZ-be.new POSS-PL RR-tell-GET NAME-male

abihā kedam pu
 a-bihá ke-dàm pu
 POSS-trade NMLZ-go say

‘[...] the new generations get to tell each other a saying, “Bokolapo is going to the market” (meaning instead of going to your destination, you’re going the other way)’ [HI, BPh 020]

5.1.2 Role-marking of NPs

Role-marking is based on a combination of syntactic, pragmatic, and semantic factors, and the pragmatic factors are at times more prominent than the syntactic ones. We can formally distinguish between three such grammatical roles: unmarked NPs; those marked with *-phān* ‘non-subject’,⁸ and those marked with *-lōng* ‘locative’.

Any kind of NP may remain unmarked if it is clear from context what kind of syntactic/semantic role it plays in the clause. Thus, clearly core arguments such as S, A, and O may be unmarked (S arguments in intransitive clauses are, in fact, always unmarked, A arguments in transitive clauses are typically unmarked, and O arguments may or may not be unmarked as there is differential object marking in Karbi⁹). However, in addition, oblique participants may remain unmarked if their clausal role is clear from context. For example, in (11), the location of the eating event is marked by the relator noun *-ngsóng* ‘high up’ in the first clause (‘the tree house’), but is unmarked in the second clause (‘the field hut’).

- (11) [...] *nangpole hemtap angsong chote, nangtumke*
 [[nang-pō=le [hēmtāp a-ngsóng] chō-tē] nang-tūm=ke
 2-father=FOC:IRR tree.house POSS-high.up eat-if 2-PL=TOP
- mandule cho*
 [mandú=le] chō
 field.hut=FOC:IRR eat
- ‘[...] if your father takes his meal in the tree house, you eat in the field hut’ [CST, RO 017]

Moving on to the next type of role-marking, participants marked with *-phān* ‘non-subject’ are typically core arguments in syntactic O or R roles. For example, in (12), the O argument *chonghō* ‘frog’ of *sáp* ‘beat with a flexible object’ is marked with *-phān*.

- (12) *chongho aphan jamir abupen sapphratphratdet [...]*
chonghō a-phān *jamír* *a-bú=pen* *sáp-phrát~phrát-dét*
frog *POSS-NSUBJ* *grain.sp* *poss-bundle=with* *beat.w/flexible-IDEO~DISTR.PL-PFV*
 ‘and with a bundle of jamir they beat **the frog** [...]’ [RBT, ChM 079]

However, *-phān* can also mark oblique participants, as with the topical first person participant in (13) where the predicate is the property-concept term *mē* ‘be good.’

- (13) *neliphanke me'ongchotlo*
ne-li-phān=ke *mē-óng-chót-lò*
1EXCL-HON-NSUBJ=TOP *be.good-be.much-very-RL*
 ‘this is very good (i.e. a very good opportunity) for me’ [SiT, HF 003]

Finally, *-lòng* ‘locative’ (cf. *-lòng* ‘CLF:place’ in (25)) generically marks oblique locational expressions of any semantic type, i.e. it can replace relator nouns that express more specific locational relations such as ‘in,’ ‘at,’ ‘around,’ ‘near,’ and so on. In addition, *-lòng* marks human R-like locational arguments, as in (14), and human O-like locational argument, as in (15).

- (14) *[...] aphi along thondamkoklo*
[a-phì **alòng]** *thòn-dām-kòk-lò*
poss-grandmother *LOC* *drop-go-in.a.fixed.place-RL*
 ‘[...] and she left (the child) **with the grandmother**’ [KK, CC 032]
- (15) *apiso along chidunkri*
[a-pisò **alòng]** *chV-dūn-kri]*
poss-wife *LOC* *RR-join-follow.closely*
- laso abamonpi along dunkrilo*
[[lasō **abamón-pī** **alòng]** *dùn-kri-lò]*
this *poss-wise.person(<Ind>-female)* *LOC* *join-follow.closely-RL*
 ‘he followed his wife closely, he followed **this bamonpi** closely’ [KK, BMS 082]

As mentioned above, oblique participants are typically marked by relator nouns in Karbi. However, in addition to the relator nouns, there is one enclitic *=pen*, which marks the instrumental, as in (12) and (22); the comitative, as in (18) and (25); as well as the ablative, as in (1) and (3). In addition, the diachronically same element *-pen* is a clause final marker (23).

5.1.3 Noun phrase delimiter *abàng* ‘NPDL’

The noun phrase delimiter *abàng* (‘NPDL’) (cf. *-bàng* ‘CLF:HUM:PL’) marks noun phrases and relator noun phrases in any clausal role (§5.1.2) and with any information status (§5.1.4). Nevertheless, it typically occurs on noun phrases unmarked for clausal role and either unmarked for information structure status or marked with topic *=ke*; it thus appears to correlate with subject and topic to a large degree. The position of the noun phrase delimiter is the end of an NP, either preceding or following any role markers, and always before any information structure clitics. Examples are shown in (16).

- (16) *an Bey Ke'et pu atum abangke*

án [Bēy ke-èt pu a-tūm abàng=ke]
and.then CLAN NMLZ-be.yellow QUOT POSS-PL NPDL=top

asomar abang etpik tangho

[a-oso-màr abàng] èt-pík tànghò
POSS-child-PL NPDL be.yellow-very REP

‘and then, with respect to the so-called Bey the Fair and his family, his children were very fair’ [WR, BCS 010]

5.1.4 Information status clitics

There are four information status clitics: =*ke* ‘topic,’ =*tā* ‘additive,’ =*si* ‘focus:realis,’ and =*le* ‘focus:irrealis.’

Topic =*ke* is used frequently and has the following functional distribution. It almost always occurs on elements at the beginning of the clause, second only to discourse connectors and additional elements marked with =*ke* (although generally no more than two =*ke* marked elements occur in a row). Topic =*ke* most frequently occurs: (a) on S arguments in equational clauses; (b) on framing elements including spatial or temporal adverbs as well as, for example, affected possessors, as in (17); or, (c) in contrastive topic constructions such as (18).

- (17) [...] “*nangke nangdin dolang*”

[nàng=ke] [nang-dín] dō-làng]
2=TOP 2POSS-day(<Asm) exist-still

‘[...] “you still have your life to live (lit. days)” (he said)’ [KK, BMS 084]

- (18) *e nanghem nangritlo, nangke nangcheleroklo*

[e nang-hém nang-rít-lò] [nàng=ke] nang=che-lē-ròk-lò]
DSM 2POSS-house 2POSS-field-RL you=TOP CIS=RR-reach-completed-RL

netumke damthekthelo, nangphipen

[ne-tūm=ke dám-thék-Cē-lò
1EXCL-PL=TOP go-know.how-NEG-RL

nangphuké

nang-phi=pen nang-phù=ke]
2POSS-grandmother=with 2POSS-grandfather=TOP

‘it’s your house and property, **you**’ve already reached, (but) **we** don’t know how to go, your grandmother and your grandfather’ [KK, BMS 096]

While =*ke* ‘topic’ typically marks elements inside the clause, it may also appear on entire clauses. In that case, the clause marked by =*ke* is rarely a subordinate clause but more typically represents a main clause that is pragmatically linked with the context by providing the background to the topic at hand.

Additive =*tā* covers a broad range of functions (Konnerth 2014b). Besides marking simple additivity (‘also’) and, related to this function, bisyndetic coordination, it also marks scalar additivity (‘even’), and universal quantification (‘all (of a particular set)'). Its information structuring functions are related to topic-switch or perspective-switch marking. An example is (19), where the frog jumping on the squirrel’s ladder and breaking it leads to the squirrel in turn getting angry. This direct cause-and-result relationship between the two events is marked by =*tā* on *karlēsibōngpō* in this example.

- (19) *karlesibongo adon chonrai*
[*karlēsibōngpō* a-dón] *chōn-rài*
squirrel.sp POSS-bridge jump-RES:solid.obj.breaking
amat karlesibongpota aning thilo
amāt karlēsibōngpō=tā [*a-nīng* *thī-lò*]
and.then squirrel.sp=ADD POSS-mind be.short-RL
'(the frog) jumped on the ladder of the squirrel so it broke, and then the squirrel
in turn got angry' [RBT, ChM 018-9]

The two focus markers *=si* and *=le* represent (ir)realis counterparts such that *=le* ‘focus:irrealis’ occurs on NPs if the predicate falls into one of the relevant irrealis categories such as non-declarative speech acts (e.g. imperatives or hortatives) or negation (see §4.2.3). Both *=si* and *=le* occur in corrective focus constructions but also in other contexts that apparently aim at directing the interlocutor’s attention to the focus-marked constituent.

5.2 Nominalization

While nominalization may function as a simple derivational process in Karbi (deriving nouns from verbs, as in *a-ki-kú* ‘POSS-NMLZ-crow’ > ‘his (the rooster’s) crowing’), the great majority of occurrences represent predicate or clausal nominalization constructions. Karbi only has one nominalizer, which is *ke-* ~ *ki-* ~ *ka-*, a reflex of the pan-Tibeto-Burman nominalizing velar prefix (Konnerth 2009). Relative clauses and modifiers derived from the verbal subclass of PCTs are typically nominalized, and so are a number of complement clause constructions and adverbial clause types.¹⁰ In addition, there are nominalized main clause constructions. The range of functions carried out by nominalization is typical for languages in the area (Noonan 1997).

Example (20) illustrates relativization of an O argument. The relative clause verb *pòn* ‘take away’ is nominalized with *ke-* and the head noun *theseré* ‘fruit’ is marked with *a-* (§4.1).

- (20) *alang kepon athesere do’anta klolaplo*
[[*alàng* *ke-pòn*]RC *a-theseré*]HN *dō-án=tā* *kló-láp-lò*
3 NMLZ-take.away POSS-fruit exist-all=ADD fall-completely-RL
'all of the fruit that he was taking away fell out' [SiT, PS 030]

The exact same relativization construction shown in (20) is also used to relativize on locatives, as in (21), or possessors, as in (22).

- (21) *lasi la thap ketok alengpumta otdunno, [...]*
lasì [[*lá*] [*thàp* *ke-tòk*]RC *a-lengpūm=tā*]HN *ót-dùn-nō*
therefore this rice.beer.cake NMLZ-pound POSS-pebble=ADD touch-JOIN-be.bad
'the pestle with which the rice beer cake is ground shouldn't be touched, [...]'
[WR, BCS 037]

- (22) [...] *lapenke arlong achetpen sarnung kidip*
lapèn=ke [*arlōng* *a-chèt=pen* *sarnūng* *ke-díp*]RC
and=TOP stone POSS-small.piece=with roof NMLZ-cover
ahemta nelitum thekdamlong
a-hém=tā]HN ne-li-tūm thèk-dām-lōng
POSS-house=ADD 1EXCL-HON-PL see-go-GET
'[...] and then, we also went to see the houses, whose roofs (they) cover with slabs
of stone' [SiT, HF 048]

In (23), *dàm* ‘go’ is nominalized to function as a complement clause of *pangchèng* ‘start.’

- (23) *chepaklangdampen... latum kedamthu pangchengló*
 che-pe-klàng dām-pen... là-tūm ke-dàm-thū pangchèng-lò
 RR-CAUS-appear go-NF this-PL NMLZ-go-again start-RL
 ‘after going to show them, they again start walking’ [SiT, PS 041]

Example (24) begins with an adverbial clause based on subordinator *aphí*, which has grammaticalized from a relator noun and still carries *a-* ‘possessive.’ The whole adverbial clause construction thus transparently relates to the relative clause construction shown above.

- (24) *laso hem nangkachiri aphi, apenan abang sunjoi*
 [[lasō][hēm] nang=ke-che-rī aphí] a-penàn abàng sūn-jòi
 this house CIS=NMLZ-RR-search after POSS-husband NPDL descend-quietly
 ‘[...] after she went back to search for it in the house, the husband quietly came down [...]’ [SeT, MTN 042]

In addition to subordination, there are nominalization constructions involving main clauses. One such instance consists in the nominalization of the main verb followed by the existential copula *dō*, as in (25).

- (25) *Naka akhei amonit so'arlopen pusetame*
 Náká a-khéi a-monít so'ärlo=pen pùsetämē
 TRIBE POSS-community POSS-man women:COLL=with likewise
sopinsopen pusetame elong longni lason
 sopinsō=pen pùsetämē e-lòng lóng-ní lasón
 boy:COLL=with likewise one-CLF:place CLF:place-two that.way
kechetong doke dohe
ke-chetòng dō=ke dō=he
NMLZ-meet exist=top exist=AFTERTHOUGHT
 ‘[...] from both the women and men, (only some of them) I got to meet in a place or two like that’ [SiT, HF 004]

In addition, ‘stand-alone’ nominalization without a copula is also found in Karbi in two constructions. One is the modern imperfective construction, whereby the nominalizer *ke-* has been reanalyzed as an imperfective marker. The other is a focus construction which typically involves focus marker *=si* (§5.1.4), and/or interrogative pronouns or adverbs.

5.3 Non-declarative main clause types

In polar interrogatives (e.g. ‘did you eat?’) and disjunctive interrogatives common in Southeast Asia (schematically, ‘eat not-eat?’), the question particle *=ma* typically occurs at the end of the clause. In content questions, *=ma* may occur but only in specific pragmatically non-neutral contexts. Content questions include interrogative pronouns or adverbs that typically start with the bound interrogative morpheme *ko-*, such as *komāt* ‘who,’ (*ko*)*pí* ‘what,’ *komantú* ‘when.’

For imperatives, the bare stem may be used, but there are also four suffixes that indicate various types of contextuality of a command, i.e. whether or not particular circumstances and (potential) consequences are relevant and if so how. The four suffixes are

-nōi, *-nōn*, *-thā*, *-tū*. The prohibitive is *-rī*, although simple declarative negation can be used as well (analogous to the bare stem imperative). In a pattern also found, for example, in Rongmei Naga (Deb and Singha 2014), the prohibitive *-rī* can be combined with any of the four imperative suffixes to convey the type of contextuality afforded by each suffix.

The general (co-)hortative (directed at first and second person) is *-nāng*, while extended forms *-lonāng* and *-sināng* also exist. The jussive construction (directed at third person) includes the causative prefix *pe~pa-* in combination with hortative *-nāng*.

5.4 Clause combining

Most subordination is structurally carried out via nominalization (§5.2). The only non-nominalized dependent clause construction is clause chaining via *-si* ‘non-final:realis,’ *-ra* ‘non-final:irrealis,’ or *-pen* ‘non-final:with’ (which derives from the NP clitic *=pen* ‘with (instrumental/comitative)'). The two suffixes *-si* and *-ra* form (ir)realis counterparts as outlined in §4.2.3. Clause chaining typically marks events in temporal sequence as in (26), although further constructions built on the basic clause-chaining construction and carrying out other functions exist as well.

- (26) *e anke apaita pharla dam,*
e [ánke a-pāi=tā pharlá dàm]
DSM and.then POSS-mother=ADD:DM outside.part.Karbi.house go
theng akhangra okóksi, hem damsi,
[thēng a-khangrá ó-kòk-si] [hēm dàm-si]
firewood Poss-basket.for.firewood leave-in.a.fixed.place-NF:RL house go-NF:RL
hongkup ingnilúnsi, mok chepachusi,
[hongkūp ingni-lùn-si] [mōk che-pa-chū-si]
entrance.area.Karbi.house sit-big:AO-NF:RL breast RR-CAUS-suck-NF:RL
“*dojoinoi, po!*”
[dō-jōi-nōi pō]
stay-quietly-INF.COND.IMP father
‘and then, the mother went and unloaded the firewood in the Pharla (Veranda), then went inside the house, sat down in the Hongkup, gave the child the milk, (and said) “be quiet, Daddy”’ [KK, CC 015]

5.5 Discourse constructions

A number of areally wide spread discourse constructions also are common in Karbi. This includes the use of final particles for pragmatic functions, as well as tail-head linkage, whereby a portion (typically the verb) at the end of one intonation unit is repeated at the beginning of the next intonation unit. Common are also elaborate expressions and parallelism (Solnit 1995). For example, in (27), the elaborate expression *vōtēk vōsō* collectively refers to ‘wild birds’ and is embedded into a parallelism construction that consists of the repetition of the verb inside this headless locative relative clause.

- (27) [...] *hi’ipi abangke etum aphan ha votek ingrengre*
[hī’ipī abàng=ke] [e-tūm aphan] há [[vōtēk ingrèng-Cē]
witch NPDL=TOP 1PL.INCL-PL NSUBJ over.there wild.bird call(small.animals)-NEG

<i>voso ingrengre along ekethondamti</i>			
[vōsō ingrēng-Cē]	along]	e=ke-thòn-dām-tí	
EE:vōtèk call(small.animals)-NEG	LOC	1PL.INCL=NMLZ-drop-go-get.rid.off	
‘[...] that witch, she went and dropped us over there where the birds don’t chirp’			[CST, HM 062]

ABBREVIATIONS

ADD	additive focus (=tā)
CAUS	causative (<i>pe-</i> ~ <i>pa-</i>)
CIS	cislocative (<i>nang</i> =)
DISTR	distributive
DSM	discourse section marker
EE	elaborate expression
EXH	exhaustive
GO	grammaticalized ‘go’
HON	honorific
IRR1	irrealis1 (- <i>pò</i>)
IRR2	irrealis2 (- <i>jí</i>)
JOIN	grammaticalized ‘join’
NF	non-final
NPDL	noun phrase delimiter
NSUBJ	non-subject
PCT	property-concept term
PRF	perfect (- <i>é</i> t)
QUOT	quotative (<i>pu</i>)
RC	relative clause
RL	realis (- <i>lò</i>)
RR	reflexive/reciprocal

NOTES

- 1 This figure as an approximate comes from the Karbi Literary Association called Karbi Lammet Amei, and includes speakers of all dialects. The *Ethnologue* lists 420,000 speakers of Hills Karbi from the 2001 census, and 125,000 speakers of what they refer to as ‘Amri Karbi’ (see fn. 2) (Lewis *et al.* 2015).
- 2 The name ‘Amri Karbi’ for the plains dialect of Karbi is a political issue. The designation is rejected by the Karbi Lammet Amei (which represents all Karbi people and all variants of the language, although its headquarters are in the Hills District). I was told by the Karbi Lammet Amei that among those speaking one of the plains varieties, some are in favor of using this term and others are not.
- 3 The allomorphs of the nominalizer occur as follows. The form *ki-* precedes monosyllabic roots with a +high vowel, i.e. /i/ or /u/, whereas *ke-* precedes all other monosyllabic roots. Curiously, the form *ka-* can generally be described as occurring with roots that consist of more than one syllable. There is, however, variation across (and sometimes even within) speakers as to where *ke-* and where *ka-* are used, both in the case of multisyllabic roots as well as in the case of monosyllabic roots in the presence of other prefixes (i.e. the causative *pe-*, the reflexive/reciprocal *che-*, the auto-benefactive/malefactive *cho-*).
- 4 Some speakers use *pa-* exclusively, whereas others use *pe-* before monosyllabic roots (with no other prefixes present) and *pa-* elsewhere.

-
- 5 Causative prefixes with a similar form (i.e. a /p/-initial open syllable) are also found in other TB languages of Northeast India (e.g. Maraa and Khumi (both Kuki-Chin), Dimasa (Bodo-Garo), Angami (Angami-Pochuri); see also Matisoff (2003 :132–3)). Interestingly, these kinds of causative prefixes are also widespread in Austroasiatic, reconstructing to Proto-Austroasiatic *pa-/ap- (Sidwell 2008: 253). Diffloth (2008: 98), citing data from Shafer (1952, 1965), suggests that the Khasi causative prefix *pn-* was borrowed into Tibeto-Burman.
- 6 This terminology follows Post (2007, 2009) in his work on Galo (Tani). Similar suffix classes are found in various other Tibeto-Burman languages of Northeast India besides Karbi and Galo, such as Garo (Burling 2004) or Meithei (Chelliah 1997), and may also have some degree of overlap with the extended classes of verbal classifier suffixes in Kuki-Chin languages discussed by Peterson (2008).
- 7 The plural marker may or may not be included in this diagram as it is structurally best analyzed as the head noun of the possessive construction; see §4.1.
- 8 This type of marker, which mostly occurs differentially on primary objects, is also referred to as the ‘anti-ergative’ or ‘anti-agentive’ in Tibeto-Burman linguistics (LaPolla 1992, 1994).
- 9 Differential object marking in Karbi depends on an interplay of semantic and pragmatic factors. While +human arguments are more likely to receive marking, there are also occurrences in the corpus of object-marked –human arguments.
- 10 Curiously, the Karbi corpus contains multiple examples, in which overt nominalization via prefixation of *ke-* is lacking in constructions that need to be analyzed as requiring synchronic nominalization (i.e. relative clauses, complement clauses, adverbial clauses, and PCT modifiers). There does not appear to be a particular pattern underlying the instances where the prefix is absent.

REFERENCES

- Benedict, Paul K. 1972. *Sino-Tibetan: A Conspectus*. New York: Cambridge University Press.
- Burling, Robbins. 2004. *The Language of the Modhupur Mandi (Garo)*. Vol. 1. New Delhi: Bibliophile South Asia.
- Chelliah, Shobhana Lakshmi. 1997. *A Grammar of Meithei*. Berlin: Mouton de Gruyter.
- Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford; New York: Oxford University Press.
- Deb, Debjit and Kh. Dhiren Singha. 2014. “Negation in Rongmei Naga.” In *North East Indian Linguistics*, edited by Gwendolyn Hyslop, Linda Konnerth and Stephen Morey, 6:95–104. Canberra, Australian National University: Asia-Pacific Linguistics Open Access.
- DeLancey, Scott. 2012. “On the Origins of Bodo-Garo.” In *North East Indian Linguistics*, edited by Gwendolyn Hyslop, Stephen Morey and Mark W. Post, 4:3–20. New Delhi: Foundation / Cambridge University Press India.
- . 2013. “Creolization in the Divergence of Tibeto-Burman.” In *Trans-Himalayan Linguistics: Historical and Descriptive Linguistics of the Himalayan Area*, edited by Nathan Hill and Thomas Owen-Smith, 41–70. Berlin: Mouton de Gruyter.
- . 2014. “Sociolinguistic Typology in North East India: A Tale of Two Branches.” *Journal of South Asian Languages and Linguistics* 1 (1): 59–82.
- Diffloth, Gérard. 2008. “Shafer’s ‘Parallels’ between Khasi and Sino-Tibetan.” In *North East Indian Linguistics*, edited by Stephen Morey and Mark W. Post, 1: 93–104. New Delhi: Foundation/Cambridge University Press India.

- Grierson, George Abraham. 1904. *Linguistic Survey of India*. Vol. 3, parts 1–3, Tibeto-Burman family. Calcutta: Superintendent of Government Printing, India. Available at: www.joao-roiz.jp/LSI/search (accessed 6 December 2014).
- Grüßner, Karl-Heinz. not dated. *Karbi Dictionary*. Tübingen: University of Tübingen.
- . 1978. *Arleng Alam, die Sprache der Mikir: Grammatik und Texte*. Beiträge zur Südasiensforschung 39. Wiesbaden: Franz Steiner.
- Jeyapaul, Vethamuthu Yesunesiah. 1987. *Karbi Grammar*. CIIL Grammar Series 15. Mysore, India: Central Institute of Indian Languages.
- Joseph, U.V. 2009. “The Numeral ‘One’ in Khasi and Karbi.” In *North East Indian Linguistics*, edited by Stephen Morey and Mark W. Post, 2:149–161. New Delhi: Foundation / Cambridge University Press India.
- Konnerth, Linda. 2009. “The Nominalizing Prefix *gV- in Tibeto-Burman.” MA thesis, Eugene, OR: University of Oregon.
- . 2014a. “Additive Focus and Additional Functions of Karbi (Tibeto-Burman) =tā.” In *Proceedings of the 38th Annual Meeting of the Berkeley Linguistics Society*, edited by Kayla Carpenter, Oana David, Florian Lionnet, Christine Sheila, Tammy Stark and Vivian Wauters, 206–222. Berkeley, CA: eLanguage (Linguistic Society of America).
- . 2014b. “A Grammar of Karbi.” PhD dissertation, Eugene, OR: University of Oregon.
- . 2015. “A New Type of Convergence at the Deictic Center: Second Person and Cislocative in Karbi (Tibeto-Burman).” *Studies in Language* 39 (1): 24–45.
- Konnerth, Linda and Amos Teo. 2014. “Acoustic-Statistical and Perceptual Investigations of Karbi Tones: A Peculiar Case of Incomplete Neutralisation of F0.” In *North East Indian Linguistics*, edited by Gwendolyn Hyslop, Linda Konnerth, Stephen Morey and Priyankoo Sarmah, 6:13–37. Canberra, Australian National University: Asia-Pacific Linguistics Open Access.
- LaPolla, Randy J. 1992. “Anti-Ergative Marking in Tibeto-Burman.” *Linguistics of the Tibeto-Burman Area* 15 (1): 1–9.
- . 1994. “Parallel Grammaticalizations in Tibeto-Burman Languages: Evidence of Sapir’s ‘drift.’” *Linguistics of the Tibeto-Burman Area* 17 (1): 61–80.
- Lewis, Paul M., Gary F. Simons and Charles D. Fennig, eds. 2015. *Ethnologue: Languages of the World* (18th edition). Dallas, TX: SIL International. Available at: www.ethnologue.com (accessed 6 December 2014).
- Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley, CA: University of California Press.
- Noonan, Michael. 1997. “Versatile Nominalizations.” In *Essays on Language Function and Language Type. In Honor of T. Givón*, edited by Joan Bybee, John Haiman and Sandra A. Thompson, 373–394. Amsterdam/Philadelphia: John Benjamins.
- Peterson, David A. 2008. “Bangladesh Khumi Verbal Classifiers and Kuki-Chin ‘Chiming’.” *Linguistics of the Tibeto-Burman Area* 31 (1): 109–138.
- Post, Mark W. 2007. “A Grammar of Galo.” PhD dissertation, Melbourne: La Trobe University, Research Centre for Linguistic Typology.
- . 2009. “Predicate Derivations in the Tani Languages: Root, Suffix, Both or Neither?” In *North East Indian Linguistics*, edited by Stephen Morey and Mark W. Post, 2:175–197. New Delhi: Foundation / Cambridge University Press India.
- Shafer, Robert. 1952. “Études sur l'austroasién.” *Bulletin de la Société de Linguistique de Paris* 48 (1): 111–158.
- . 1965. *Studies in Austroasian II*. Vol. 30: 5. Helsinki: Studia Orientalia edidit Societas Orientalis Fennica.

- Sidwell, Paul. 2008. "Issues in the Morphological Reconstruction of Proto-Mon-Khmer." In *Morphology and Language History: In Honour of Harold Koch*, edited by Claire Bowern, Bethwyn Evans and Luisa Miceli, 251–298. Amsterdam/Philadelphia: John Benjamins.
- Solnit, David B. 1995. "Parallelism in Kayah Li Discourse: Elaborate Expressions and Beyond." In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society: Special Session on Discourse in Southeast Asian Languages*, 127–140. Vol. 21. Berkeley, CA: Berkeley Linguistics Society.
- Teron, Manik and Rajendra Tumung, eds. 2007. *Karbi Lamarjong*. Diphu, Karbi Anglong, Assam, India: Kamrup Karbi Lammet Amei (Longki Phangcho).
- Walker, George David. 1925. *A Dictionary of the Mikir Language*. Shillong: Assam Government Press.
- Wolfenden, Stuart Norris. 1929. *Outlines of Tibeto-Burman Linguistic Morphology*. London: Royal Asiatic Society.

CHAPTER FIFTEEN

TANI LANGUAGES

Mark W. Post and Jackson T.-S. Sun

1 INTRODUCTION

Tani refers to a compact cluster of Tibeto-Burman languages situated at the eastern end of the Himalayas, in a primarily mountainous area skirted on four sides by Bhutan, Tibet, Burma and the Brahmaputra River in Assam. The main concentration of Tani languages covers the central part of the Indian state of Arunachal Pradesh, comprising the bulk of the following districts: East Kameng, Upper Subansiri, Lower Subansiri, Upper Siang, West Siang, East Siang, Lower Siang and Dibang Valley. Tani languages are spoken by populations who either are currently or at one time were ‘tribally’ identified as Apatani, Nyishi, Na, Bangni, Tagin, Galo, Hill Miri or Adi (various subtribes), in addition to numerous smaller, usually clan-based or village-based self-identifications (such as Komkar). The single largest population of Tani language speakers, however, is the Mising tribe of upper Assam. The Tani languages are spoken in a basically continuous area, bordered in the west by Miji, Puroik, Koro and Hruso speakers, in the east by Idu speakers, in the south by speakers of Boro and Assamese, and in the north by speakers of Tibetic languages. Lingua francas spoken in the Tani area include Assamese in the south and Arunachali Hindi in the central area. Tibetic languages were used as lingua francas in earlier decades, although their use has recently waned.

The 2011 *Census of India* reports a total of 1,380,878 individuals in Arunachal Pradesh and Assam who self-identified as belonging to a tribe whose mother tongue is a Tani language (Census of India 2011): these included 631,042 Tani tribespeople in Arunachal Pradesh, and 680,424 Mising in Assam (plus an additional 7,412 Mising in Arunachal Pradesh). These figures must be treated with caution, as it is not the case that all people who align tribally with a primarily Tani language-speaking population speak this language themselves; given the rapid spread of Hindi throughout Arunachal Pradesh (Modi 2005), and of Assamese among the Mising in Assam (Pegu 2011: 157), the number of fluent Tani language speakers as a percentage of this total must be considerably lower.

Scattered Tani communities spill over the Sino-Indian border into adjacent areas in Motuo (Miguba and Misinba tribes), Milin (Bokar and Tagin tribes), and Longzi (Bengni, Na, Bayi, Dazu and Mara tribes) counties of Tibet (Ou-Yang 1985: 76), where they are lumped with certain linguistically non-Tani peoples (e.g. the Idu, Sulung and Bangru) to form the Luoba nationality. Very little current information regarding these groups is available (although see Huber 2012).

Tani languages vary greatly in terms of number of speakers, adequacy of description, and degree of endangerment. Mising has perhaps half a million speakers, but lacks a modern, comprehensive dictionary and grammar, and is currently undergoing considerable retraction under the influence of Assamese. On the other hand, Tangam has only 150 speakers; however, it has a forthcoming modern (if relatively slight) comprehensive description, and all Tangam children seem – for the time being, at least – to be learning Tangam as a first language and speaking it fluently. Relatively urgent descriptive priorities include

Nyishi, a large and important language about which very little current information of certain reliability is available, Apatani, a relatively divergent language with a robust and intricate tone system which lacks a comprehensive and reliable description,¹ as well as all Tani varieties still spoken in Tibet.

2 GENEALOGY AND SUBGROUPING

Tani languages appear to constitute a distinct subgroup within Tibeto-Burman, as argued in Sun (1993). All Tani languages shown on the family tree in Figure 15.1 share a large amount of common core vocabulary, exhibit a very large number of regular sound correspondences, and in general support firm reconstruction of the phonology and vocabulary of their common proto-language ‘Proto-Tani’. Koro and Milang, spoken on the western and eastern fringes of the Tani area, show many similarities to Tani languages, and Milang in particular has previously been included within the Tani subgroup. Post and Modi (2011) argued against the inclusion of Milang in Tani proper, showing that the strongest resemblances to Tani languages were undoubtedly due to contact with Padam Adi. Around the same time, Post and Blench (2011) placed Koro and Milang in a distinct ‘Siangic’ subgroup. In addition to Siangic languages, Tani languages show many lexical and grammatical resemblances to the easterly Idu-Tawrā languages, and to Tawrā in particular. Unfortunately, however, the current state of description does not enable us to determine whether these resemblances are due to common inheritance or to earlier language contact.

Sun (1993) divided Tani languages into two major branches, on phonological and lexical grounds: Western Tani (WT), clustered around the Subansiri and Kameng Rivers, and Eastern Tani (ET), clustered around the Siang/Brahmaputra. The WT/ET division almost

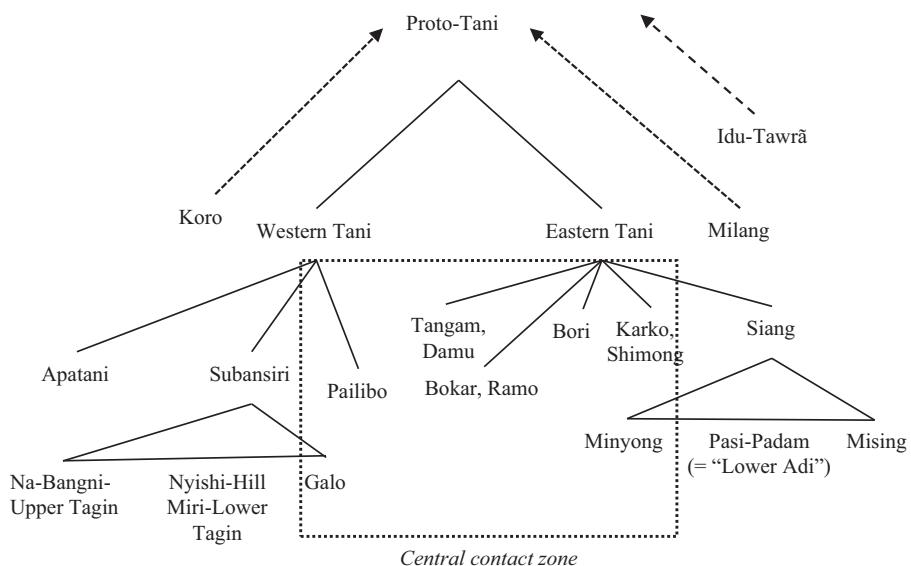


FIGURE 15.1 PROVISIONAL FAMILY TREE OF THE TANI LANGUAGES (REVISION OF SUN 1993: 297)

certainly has at least some sort of genealogical basis as Sun proposed; however, language contact has played an at least equally powerful role in the evolution of Tani languages, and has led to convergence and shared innovation both within and across genealogical branches. Post (2015) reviewed 15 Tani phonological innovations exhibited by 15 Tani languages, and found that no two high-level innovations identified the same set of genealogical branches. In general then, it is best to view the Tani languages as a dialect continuum, with a reconstructible shared ancestor, but with a subsequent evolution marked at least as much by areal sharing of innovations as by sharp genealogical branching. Figure 15.1 provides a heuristic Tani family tree on the basis of what might be called ‘aggregate sharing of innovations’. Depending on which of these innovations are attributed to genealogical inheritance, and which are attributed to contact, a large number of alternative trees could be drawn (particularly with respect to geographically central Tani languages).

3 PHONOLOGY

Syllable and word are both important organizational units in Tani phonology. The Proto-Tani (PT) syllable canon is (C_i)(G)V(X), in which C_i is an optional initial consonant, V is an obligatory nucleus, G is an optional approximant *r* or *j*, and X is an optional coda, either a nucleus-identical V or a final consonant C_f (generally drawn from a relatively restricted set). In at least one language – Apatani – the X constituent may also be represented as nasalization over the nuclear vowel. Tani syllables with an X constituent are bimoraic, or heavy; syllables without an X constituent are monomoraic, or light. Many phonological processes in Tani languages depend on this weight distinction, including syncope and prominence assignment in Galo (Post 2007: §4.2.3), and the realization of tones in Apatani (Post and Tage 2013). The PT syllable canon is well-preserved by the modern languages, albeit in different ways; generally speaking, onsets (including clusters) are better preserved in WT languages, while ET languages are more conservative with rhymes.

Tani segment inventories are relatively simple. Consonants are found at labial, alveolar, (alveolo-)palatal, velar and glottal places of articulation (Table 15.1). Dental and retroflex consonants are not found in Tani languages – nor indeed are they found in either the Arunachali variety of Hindi or in the neighbouring Eastern Indo-Aryan language Assamese; occasional transcriptions of dentals and retroflexes in the sources would accordingly seem to be errors.

All Tani languages contrast voiceless and voiced plosives. Affricates pattern with plosives in languages which retain them; in Siang-area languages such as Pugo Galo and Lower Adi, *tč* historically merges to *c* and thence *s* or *h*. PT contrasted voiced and voiceless fricatives, but this distinction is maintained by very few modern languages other than Na/Bangni; in most modern languages, fricatives have merged to a two-way distinction between a glottal *h*, and a medial fricative, either the conservative *č* (mostly in northerly lects) or the innovative, Indo-Aryan-influenced *s* (towards the south). A small number of languages, including Apatani, have also innovated a velar fricative *x*. A breathy/glottal onset distinction among vowel-initial words is found in some languages, with the breathy onset ^h- reconstructible to a subset of PT fricatives, and the glottal onset ^ʔ- reflecting vowel-initial morphemes at the PT stage; for example, compare Northwestern Galo ^h*āa* ‘come; set (sun)’ (< PT **vay*, cf. Bangni *vaa*-) with ^ʔ*ap*- ‘shoot’ (< PT *^(?)*ap*, cf. Bangni *ap*-). Neither onset is found word-medially: compare the cognate first and second formatives in Northwestern Galo ^h*ag-jāa* ‘fermented soya bean’ and *pej-ák* ‘soya bean’. Other Tani languages, such as lower dialects of Lare Galo and Lower Adi, have merged these categories completely, retaining only a single fricative *s* or *h*.

TABLE 15.1 PROTO-TANI CONSONANT INVENTORY, BASED ON SUN (1993: 56)

	Labial	Alveolar	(Alveolo-)Palatal	Velar	Glottal
Plosive [-VD]	*p	*t	*t̪	*k	(*ʔ)
Plosive [+VD]	*b	*d	*d̪z	*g	
Nasal	*m	*n	*ɲ	*ŋ	
Fricative [-VD]	*f		*ç		*h
Fricative [+VD]	*v		*z		*ɦ
Lateral		*l			
Rhotic		*r			
Semivowel			*j		

**TABLE 15.2 PROTO-TANI VOWEL INVENTORY,
BASED ON SUN (1993: 67)**

	Front	Central	Back
High	*i	*ɪ	*u
Mid	*e	*ə	*o
Low		*a	

Most Tani languages preserve the PT seven vowel inventory (Table 15.2); although Apatani split and merged PT *ə to the polar vowels (e.g. Apatani *já-mù* ‘fire’ < PT *mə ‘fire’), the same vowel ə has re-emerged due to a vowel centralization rule in the environment of rhotics (e.g. Apatani *tà-dár* ‘worm’ < PT *dor ‘worm’). An unusually large ten vowel system has been reported by Ou-Yang (1985: 77) for the Bodic-influenced and possibly mixed language Damu: *a, i, u, e, o, ə, i, y, ə, ɿ* (= syllabic z); this report remains unconfirmed.

Although most instances of contrastive vowel length in underlying forms of morphemes may be secondary, vowel length is a robust phonological feature found in nearly all Tani languages:² compare Galo *banám* ‘to vomit’ and *baanám* ‘to move the head’. While contrastive word-medial vowel length is usually easy to detect, underlying word-final vowel length is more difficult to detect in context-free utterances, and is absent from the majority of descriptions. However, word-final vowel length can typically be discerned in phrase-medial contexts: for example, Galo *aló gó* ‘some salt’ vs *alóo gó* ‘a day’. True diphthongs are largely absent from the underlying forms of PT morphemes; diphthongs in the modern languages are typically either recent loans or contracted compounds.

Tani languages vary greatly in their inventories of syllable codas. The language that preserves the largest number of the original proto-codas is (Adi) Padam, with *-p*, *-t*, *-k*, *-m*, *-n*, *-ŋ*, *-r*, and *-l*. Milang is the only other language documented that exhibits an *-l* coda, albeit usually in Padam loans. The velar nasal *-ŋ and the dental/alveolar stop *-t codas are most prone to loss, especially in WT. The tendency towards coda attrition is epitomized in Apatani, where only two PT codas remain: *-?* (from the original stop codas) and *-r*. New final consonants, including *-s*, *-l*, and *-m*, have been imported via recent Indo-European loans, such as *pas* ‘five’ (< Indo-Aryan). In Nyishi, Galo and Hill Miri, non-original codas are created via the clipping of short final vowels from original disyllables, leading to an innovative coda voicing contrast: e.g. Nishi *ab* ‘father’, *tap* ‘pumpkin’ (Chhangte 1992: 4); cf. Bengni *a-buu* ‘father’, *ta-pii* ‘pumpkin’.

Phonological words are minimally bimoraic, and tend strongly to be disyllabic in Tani; among languages in our Tani database for which more than 1,000 lemma have been

recorded, no language has more than 0.5 per cent morphologically simplex, monosyllabic lexical words. When such words occur, they seem to be idiosyncratic retentions from a basically monosyllabic proto-language; for example, Galo *níi* ‘person’ but Minyong *?a-mí* ‘person’ and Apatani *mí-yù* ‘person’. Rhythm tends strongly to be trochaic in Tani, and has a number of phonological outcomes, including coda-reductions and syncope processes, post-head clitic incorporation, and perhaps even the overall disyllabification of the basically monosyllabic proto-lexicon (Post 2006, 2011b).

Phonemic tone is found in most relatively well-described Tani languages, including Apatani (all dialects), Galo (all dialects), Na/Bangni, Upper Minyong and Tangam, as well as in Milang. Generally speaking, tones are more robust and have a higher functional load in WT languages than in ET languages; however, it would appear that both tone and non-tone languages may be found in each branch. Analysis of tone in Tani languages requires reference to both the surface phonological word and its underlying morpheme constituents. In the case of underlying morphemes, no more than two contrastive tones have been attested for any Tani language: an unmarked H and a marked L. These ‘underlying tone-bearing units (TBUs)’ then interact with syllable and word structure to produce a surface word tone contour, rules for the derivation of which differ from language to language. In Galo, a ‘pitch peak’ over an L word is attracted to heavy syllables, thus interacting with rhythmic processes; for example, *ri-diú* ‘do-IPFV’ > [‘rí.dù] but *ri-diúu-kú* ‘do-IPFV-COMPL’ > [rí.‘dùu.kú] (here and elsewhere in this chapter, á is a (mid-)high tone, à is a low tone, and á is a (rising-)high-falling tone). In Apatani, two underlying H and L tones interact with the structure of word-final syllables to produce three disyllabic word tone contours: HH or HL for words with final light syllables (depending on the underlying specifications of constituent morphemes), and LH for words with final heavy syllables (independent of morpheme constituency). Underlying tones of words with final heavy syllables in Apatani can be discerned via tone spreading to following articles or suffixes. For example: *?ámi* (*ké*) ‘(an) elder sister’, *?ámi* (*ké*) ‘(a) tail’, *?ápi* (*ké*) ‘(some) cooked rice’, *?ámi* (*ké*) ‘(a) name’. At least some Tani languages or dialects thereof seemingly lack lexical tones, including Lower Adi, Mising and Nyishi. Unfortunately, due to an ongoing lack of adequate description (particularly of ET languages), reconstruction of PT tone categories is problematic.

Phonological and morphophonological alternations are extensive in Tani, particularly syllable-structure adjustments or assimilation. The rich allomorphy of the nominal prefix **a-* in Bokar provides a typical example: *a-ŋaa* ‘child’; *i-kii* ‘dog’; *u-puk* ‘arrow’; *e-tce* ‘clothes’; *o-ŋoo* ‘fish’; *ə-jək* ‘pig’; *i-liŋ* ‘stone’. Morpheme identification can be tricky without knowing the sandhi processes at work. For example, given Bokar *luyin* ‘fingernail’, four morphophonemic rules have to be undone to retrieve the original shapes of the component morphemes *lok-* ‘hand’ and *jin-* ‘nail’: namely, initial *j-* simplification (→ *lok-in*), obstruent voicing assimilation (→ *log-in*), spirantization (→ *loy-in*), and vowel assimilation (→ *luy-in*), compare Bokar *a-lok* ‘hand’, *la-jin* ‘bird’s claw’ (compare also Bengni *lak-sin* ‘fingernail’, from PT **lak-zin*). In general, morphophonological processes are more intricate among WT languages than among ET languages, perhaps reaching an apex in Tangam.

4 MORPHOLOGY

4.1 Diachrony

Tani morphology is best understood against the backdrop of its evolution. Most Tani morphemes are reconstructible as simplex monosyllables, and presumably reflect simplex,

monosyllabic lexical words at some early stage (whether PT or earlier). Similarly, most Tani grammatical morphemes are reconstructible to earlier lexical roots; for example, most tense-aspect morphemes reconstruct to an earlier set of clause-final existential or auxiliary verbs (e.g. Galo *-diu* ‘IPFV’ < PT **duj* ‘sit; be at (sitting)’, or Apatani *-dà?* ‘cos’ < PT **dak* ‘stand; be at (standing)’), and most predicate derivations originated as constituents of earlier serial verb constructions (e.g. Minyong *laa-bi* ‘take-BEN’ < PT **lay-bi* ‘take-give’). The modern languages are synthetic and agglutinating to varying degrees; most nouns and adjectives are disyllabic compounds (or further compounds thereof), while earlier serial verb constructions have amalgamated into a single, morphologically complex and expansive predicate word (1).

(1) Galo

<i>tolo</i>	<i>teàa-lèn-lìi-máa-diùu</i>
	< * <i>tə</i> * <i>lo</i> * <i>teay</i> * <i>len</i> * <i>lij</i> * <i>maj</i> * <i>duj</i>
DST.UP.LOC	ascend-OUT-DESD-NEG-IPFV
	DST.UP LOC ascend exit want not.(have) sit/stay

‘(I) don’t want to go up there.’

4.2 Word classes

PT appears to have had nouns and verbs, together with classifiers, postpositions, demonstratives, articles, conjunctions and particles. The modern languages have nouns, most often disyllabic prefixations or compounds of earlier nouns, adjectives, which are noun-like in structure and often derive from deverbal nominalizations, as well as classifiers, postpositions, demonstratives, articles, ideophones and particles. Adverbs are not well-represented in either the early or the modern languages; manner adverbials in the modern languages are typically derived from adjectives via a postposition or suffix **pə* (prob. < PTB nominalizer **pa*), as in Lower Adi *kampo pə* ‘beautiful AVZR’ ‘nicely; beautifully’, while temporal expressions occur as oblique noun phrases, such as Lower Adi *juma lo* ‘in (the) evening’. A few particles and predicate derivations also have adverb-like meanings and/or functions, such as the Adi ‘versatile’ particle *ruj* ‘certainly; definitely’ or the Adi predicate derivation *-man* ‘playfully’. Verbs in the modern languages are best discussed in terms of predicate structure (section 4.3.2).

Nouns are most easily distinguished by their ability to take articles, as in Adi *ami ko*, Apatani *míyù kè* (person IND) ‘a person’. Verb roots must be nominalized to be used in referential phrases, as in Adi *do-nam ko* (eat-NZR IND) ‘some food’. Conversely, nouns cannot usually take predicate morphology; Galo **nii-diu* ‘person-IPFV’ is unacceptable. In many languages, adjectives can be distinguished from verb roots and nouns by their ability to occur in both an article-derived copula construction, like nouns, and in an inflected predicate, like verb roots (2).

(2) Lower Adi

<i>apej=ə!</i>	<i>yo</i>	<i>apee-duj.</i>
	tired=COP	1.SG

tired-IPFV

‘How tired (I am)!; ‘I’m tired.’

Singular, dual and plural pronouns occur in three persons in the majority of Tani languages (Milang has an additional clusivity distinction, seemingly absent from Tani proper). However, only the first and second person forms appear to be reconstructible; third person forms often reflect earlier **ba-hi*, being a composition of **hi* ‘self’ with an

TABLE 15.3 TANI FIRST AND SECOND PERSON PRONOUNS

	SG	DL	PL
1	* <i>yo</i> 'I'	* <i>yo ji</i> 'I + two'	* <i>yo luy</i> 'I + group'
2	* <i>no</i> 'you'	* <i>no ji</i> 'you + two'	* <i>no luy</i> 'you + group'

unknown morpheme (as in Pailibo *mai* '3.sg'), or a demonstrative **hə* (as in Tagin *ᵑθə* '3.sg') or else *ko* as in Bokar (possibly <**ko* 'one'). Tangam has an especially unusual third person singular form *nodi* – the second formative in fact cognate with **hi* 'self', but the first formative oddly homophonous with the second person singular *no*, and different from the 3_{DL} and 3_{PL} forms *dani* and *datəŋ* (Table 15.3).

Classifiers seem to be a PT innovation, not found in Milang (Post and Modi 2011). PT classifiers are monosyllabic and simplex: e.g. **tiŋ* ‘CLF: GROUP (OF ANIMALS)’, **soŋ* ‘CLF: LONG, SLENDER’, **buŋ* ‘CLF: LONG, HOLLOW’ and **tak* ‘CLF: FRAGMENT OF SPHERE’. Modern Tani languages tend to have dozens of classifiers; in some languages, classifier roots are prefixed in their citation form, as in Galo *'a-biu'* ‘CLF: LONG, HOLLOW’, or else compounded to numeral roots or a small number of adjectival roots, as in Galo *biuu-ní* ‘CLF: LONG, HOLLOW-two’ or *biuu-tɔ* ‘CLF: LONG, HOLLOW-big’. In other languages, such as Apatani, classifier roots only occur in compounded forms. In Bokar, classifiers are reported to not occur with numerals higher than one (Sun *et al.* 1980: 128). Notably, all known Tani classifier systems lack a generic classifier (comparable to Mandarin *ge*), as well as a classifier for human beings. Instead, humans and unclassifiables are enumerated without the help of a classifier.

Particles are a broad category, expressing a host of modal, evidential and illocutionary-force meanings. In general, inventories are much larger in WT languages than in ET, sometimes by a factor of six or seven. Some of the best attested Tani particles are **ju(kə)* 'REP', **di* 'WONDER', **pə* 'INFL' and **la* 'CQ'; typically, particles follow the focus of a clause, occurring clause-finally in the majority of cases, or else after a focused NP in a cleft/focus construction (section 5.1). Postpositions and articles are reconstructible, and will be discussed in section 5.2. At least two topographical-deictic distal demonstratives **ta* 'DST.UP(RIVER)' and **ba* 'DST.DOWN(RIVER)' are reconstructible, with a third category 'DST.SLEV' also reconstructible, but of uncertain phonetic value (possibly **ə*). A proximate **ci* is also reconstructible; languages vary in the remainder of their demonstrative inventories, but generally include either a medial-distal or an addressee-proximate, both of which also track anaphora. Ideophones are found widely in Tani, but remain poorly described. Many have non-prototypical phonology, and typically occur in construction with the verb 'say' (or a cognate suffix).

- (3) Galo
cubúk! óm-làa ó-lik-káa
 sound.of.silent.entry.into.water say-NF fall.from.height-INSERT-PRF
 ‘He slipped “splooosh!” into the water.’

4.3 Word structures

4.3.1 Nouns and adjectives

Tani nouns and adjectives generally have one of the structures [ROOT] or, more often, [PFX-ROOT] or [ROOT-ROOT], as in Galo *nii* ‘person’, *a-dir* ‘tired’ and *lə-cək* ‘numb, of feet

(<leg-numb>) or Apatani *dzii* ‘black’, ?*a-mí?* ‘eye’ and *mi?*-láa ‘tear (< eye-juice)’. The most commonly attested prefixes are **a*-, **ɛ(j)a*-, **ja*-, **ta*-, and **pa*;-; **ɛ(j)a*- is attached to roots denoting higher animals; **ta*- is used for lower animals, insects, trees, male proper names and small objects in general; **ja*- is used with colours, diseases, other negative- connotation items and female proper names; **pa*- is confined to bird names and related nouns, e.g. Bokar *pi-tir* ‘chicken carrying basket’. As for **a*-, it is found on many basic nouns, including kinship terms, body parts and basic features of the natural world, as well as on many basic adjectives; presumably, **a*- derives from the early TB nominalizer. The prefixes **ɛ(j)a*- and **pa*- appear to derive from earlier initial formatives of compounds, reflecting perhaps PTB **sja* ‘flesh/meat/animal’ and **bja* ‘bird’ (Benedict 1972: 46, 177).

Compounds of the form [ROOT-ROOT] are common among nouns and adjectives, often following a classificatory template in which the first formative names a generic class or ‘type’ and the second a specific ‘exemplar’ (Figure 15.2). The exemplar formative may in turn stand as a new type; the resulting word networks and families can be large (Table 15.4).

Semi-reduplicative expressive compounds of the form A-B A-C are also common among Tani nouns, adjectives, and verbs alike, as illustrated by these Bengni examples: *a-tuj a-juj* ‘garbage’, *a-bak a-jak* ‘thick (liquid)’, *dzi-sit dzi-mit* ‘to pass something around’, and *jip-mii jip-maa* ‘sleepy’. The familiar Tibeto-Burman voicing alternation in

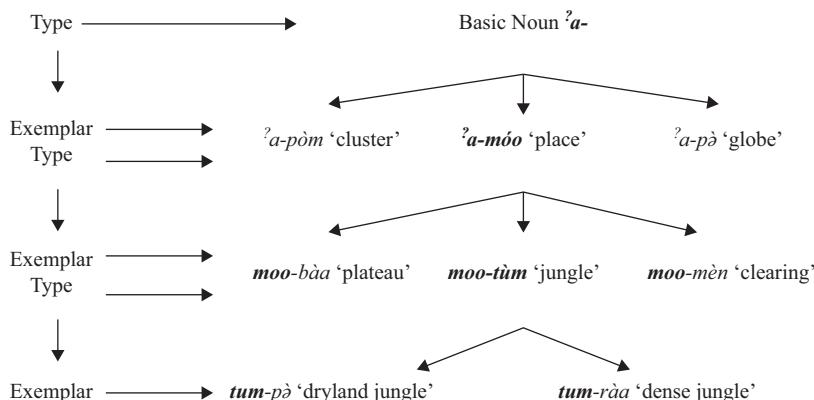


FIGURE 15.2 CLASSIFICATORY WORD FORMATION IN GALO (FROM POST 2012)

TABLE 15.4 WORD FAMILY BASED ON THE INITIAL ROOT *lii-* ‘STONE’ (FROM POST 2006)

<i>liiteák</i>	‘pebble’	<i>liinò</i>	‘boulder’
<i>liimik</i>	‘gravel’	<i>liipùm</i>	‘stone pile’
<i>liimik</i>	‘algae’	<i>liitòr</i>	‘hard stone’
<i>litàk</i>	‘giant boulder’	<i>liijää</i>	‘soft stone’
<i>likàa</i>	‘igneous rock’	<i>liikòr</i>	‘green bead var.’
<i>lipiùu</i>	‘marble’	<i>liitùm</i>	‘green bead var.’
<i>litecik</i>	‘cooking tripod’	<i>liipò</i>	‘sharpening stone’

the formation of simplex-causative pairs is preserved in some measure, albeit unproductively: for example, Galo *dír-* ‘break (of a long thing)’ vs *tír-* ‘break (a long thing)’, or *bíuk-* ‘burst’ vs *pík-* ‘pop (knuckles)’.

4.3.2 Predicates

Most Tani predicates have the basic structure [v.ROOT(-DERIVATIONS)-INFLECTIONS], an example of which can be found in (3); adjectives may also head a predicate (e.g. (2)), but take a smaller set of the available derivations and inflections. Many verb roots are cognate with noun roots, and are preferentially used together with a preceding cognate nominal (usually non-referential and unmarked for case), as in the following Bengni examples:

- | | | | |
|-----|---------------|----------------------|------------------------------|
| (4) | to lay egg | <i>pi-pi pi-</i> | (cf. <i>pi-pi</i> egg) |
| | to spit | <i>ta-teir teir-</i> | (cf. <i>ta-teir</i> spittle) |
| | to dream | <i>jip-maa maa-</i> | (cf. <i>jip-maa</i> dream) |
| | to wear shoes | <i>li-kjam kjam-</i> | (cf. <i>li-kjam</i> shoe) |

Derivation in the Tani predicate is prolific. Many predicate derivations are semantically rich, homophonous with verb roots, and in simple stems may be interpretable as compounded lexical roots: for example, Galo *in-mèn* ‘go-AS.PLAY’ ‘stroll’. However, predicate derivations tend to be fully productive, and as many as four or five can be ‘stacked’ onto a single stem, as in the following Minyong example (5). Predicate derivations encode meanings related to manner, result, direction, aktionsart, modality, argument structure change (e.g. applicatives) and word class change (Post 2010).

- | | | | |
|-----|------------|--------------|---------------------------------------|
| (5) | <i>ami</i> | <i>ə=kom</i> | <i>gok-ta-ki-ram-hi-kaa-to=i.</i> |
| | person | IND=ADD | call-INCP-TENT-FRUS-REFL-EXP-PFV=QTAG |
- ‘The guy also tried in vain to have a go at calling, eh.’

Nominalization is highly functional in Tani languages, forming relative clause constructions in addition to action and participant nouns (section 5.3). The most widely attested nominalizers are realis *-na ~ -nə* ‘subject/actor NMLZ’, *-nam* ‘action/undergoer NMLZ’ *-nana* ‘instrumental NMLZ’ and *-ko* ‘locative NMLZ’; some languages also have unrealis/purposive nominalizers, such as Galo *-há*.

Tani is similar to most other eastern Himalayan Tibeto-Burman subgroups, but different from several groups to the east and west, in lacking grammaticalized person-indexing on the predicate. The relevant inflectional categories of the predicate are aspect and mode, much as in Tibetan. Some amount of variation is found among Tani inflectional systems; the most widely attested set of predicate inflections is given in Table 15.5, together with their lexical source values (when known); a few languages exhibit additional inflections with uncertain provenance, such as Apatani *-ci* ‘intentional’ and *-yé* ‘imperative’, and Milang *-kač* ‘unrealis’.

In most Tani languages, predicate formation is sensitive to a relationship between subject person and speaker knowledge, comparable in many respects to the ‘egophoric’ systems found in Bodic languages and their neighbours. In general, declarative statements with a non-first person subject in perfective aspects are unacceptable unless qualified, as in these Bengni examples:

- | | | |
|-----|----|-------------------------|
| (6) | a. | <i>ŋuu kanoo-pa</i> |
| | | 1SG hungry-PFV |
| | | ‘I have become hungry.’ |

TABLE 15.5 TANI PREDICATE INFLECTIONS AND THEIR LEXICAL SOURCES

Domain	PT	Value	Source
Negation	* <i>maj</i>	NEG	‘not (have)’
	* <i>dunj</i>	IPFV	‘sit; be (at)’
	* <i>do</i>	STAT	‘lie down; be (at)’
	* <i>dak</i>	COS	‘stand; be (at)’
Perfectives	* <i>to</i>	PFV	‘have/be associated to’
	* <i>ka</i>	PRF	‘have/exist’
	* <i>ku</i>	COMPL	Pre-PT
Irrealis	* <i>rje</i>	IRR	‘live/exist’
Anterior	* <i>ai</i>	ANTR	PTB * <i>waj</i> ‘cop’ (?)

- b. *tatik. kanoo-pi-dii* (**kanoo-pa*)

Tatik. hungry-PFV-WONDER
 ‘Tatik has become hungry.’

In some paradigms, as in the Galo ‘direct’ perfective, dedicated suffixes index the ‘egophoric’ (a.k.a. ‘conjunction’, or ‘self-person’) category of first person in declaratives/second person in interrogatives, and the ‘alterphoric’ (a.k.a. ‘disjunction’ or ‘other-person’) category of second/third in declaratives, first/third in interrogatives (7a–b). These are semantic rather than morphosyntactic indices, however, a fact which can be demonstrated by the acceptability of other persons when evidential particles which alter the construal are employed (Post 2013) (7b). Finally, note that unlike in Tibeto-Burman languages to the north and west, egophoricity in Galo is not sensitive to agentivity or volitionality.

(7) Galo

- a. *ηó* (**nó*, **bìi*) ?*atcín* *dó-tó-bá*
 1.SG (2.SG, 3.SG) cooked.rice eat-PFV:EGO-PFV:DIR
 ‘I’ve just had my meal (I experienced it).’
- b. *nó/bìi* (**ηó*) ?*atcín* *dó-ggée-bá*
 2.SG/3.SG (1.SG) cooked.rice eat-PFV:ALTER-PFV:DIR
 ‘You/he/she just had your/his/her meal (I witnessed it).’
- c. *bìi* ?*atcín* *dó-tó-bá* *jukò*
 3.SG cooked.rice eat-PFV:EGO-PFV:DIR REP
 ‘He, it is said, “had my meal.”’
- d. *ηó* (**nó*, **bìi*) *ò-lòo-tó-bá* (**o-loo-gee-ba*)
 1.SG (2.SG, 3.SG) fall-DOWN-PFV:EGO-PFV:DIR fall-DOWN-PFV:ALTER-PFV:DIR
 ‘I fell down (I experienced it).’

5 SYNTAX

5.1 Clause and phrase types and structures

Tani information structure has the basic order [TOPIC FOCUS]. Syntactically, clauses can be divided into predicative and appositive subtypes. A predicative clause is headed by an

inflecting predicate (4.3.2), which usually occurs last, in the FOCUS position (7); more topical noun phrases occur furthest from the predicate, with the most frequent resulting order among core arguments of multi-argument clauses being agent > recipient > patient. However, different orders are straightforwardly possible (8):

- (8) Galo (popular song by Marto Kamdak)
- | | | |
|-----------------|--------------|-------------------|
| <i>nó-m̥=ne</i> | <i>ŋó...</i> | <i>káa-tó-láa</i> |
| 2.SG-ACC=NAGT | 1.SG | look-PFV-NF |
| 'I saw you.' | | |

A basic appositive clause consists of two noun phrases arrayed in the order [TOPIC FOCUS]. Appositive clauses generally handle equative and attributive functions. In many if not all Tani languages, the demonstrative **ha* ‘DIST/ANAP’ has developed via a definite article into a copula in this construction, and can predicate an adjectival as well as nominal attributive focus.

- (9) Tangam
- a. *hi apoy e* (*e* < **ha*)
PROX rice.beer COP
'This is rice beer.'
 - b. *yo kenoy e*
1.SG hungry COP
'I'm hungry.'

Posture verbs **døy* ‘sit’, **dak* ‘stand’ and **do(y)* ‘lie’ also developed locative existential copula functions in the same construction; however, languages vary in the extent to which this construction is retained. Galo appears to have merged this construction into the predicative clause type – choosing to inflect its locational posture verbs – and some Adi varieties have generalized **døy* in all locative-existential functions. However, the full paradigm is retained in Apatani, in which *dà?* predicates entities with legs, or in a standing position, *dó* predicates other inanimate entities, and *dìu* predicates other animate entities (10).

- (10) Apatani
- a. *ni-kà pàrò? so dà? kèe*
2.SG-GEN fowl PROX.LOC be.at.LOC(STANDING) POL
'Your chicken is (standing) here.'
 - b. *mòo-kà mjii dìu*
3.SG-GEN wife be.at.LOC(ANIMATE)
'His wife is here.'
 - c. *yí-kà póosá dó*
1.SG-GEN money be.at.LOC(INANIMATE)
'I have money (with me).'

Negation in appositive clauses is generally handled by post-modifying the focus with a negator *ma(y)* or *mo(y)* (the latter mainly found in north-central-area languages). Only Apatani is known to have a suppletive negative copula *nímá*. Illocutionary force is generally signalled via a wide range of particles. In most cases, particles occur clause-finally; however, in a marked focus construction, they mark the focus. When a marked focus construction is based on a predicative (not an appositive)

clause, and when the predicate is outside the focus, the predicate must either be nominalized, or else is marked by a nominalizer-derived suffix, as in the following Lower Adi example:

- (11) Lower Adi

no	kolō	la	gi-ye-n?
2SG	where	CQ	go-IRR-Q(< NMLZ:SBJ)

‘Where are you going?’

Within the noun phrase, compound nominals are typically MODIFIER–HEAD, as in Lower Adi *kalam diitə* ‘Mt. Kalam’. Coordinative compounding is also common at this level, with each noun signifying an exemplar and the whole signifying a superordinate type, as in Galo *donám-tinám* ‘eat.NMLZ imbibe.NMLZ’ ‘provisions’. Possessors are pre-head (10), while classifiers and numerals follow the head in that order, as in Apatani *sii dór-ni* ‘cattle CLF-two’, and are followed by relator nouns, articles and/or postpositions, as in Galo *?yńi kookŃi lo* ‘after a year (year back LOC)’. A striking characteristic of demonstratives is they can occur on both flanks of the head (e.g. Bengni *sii əki sii* ‘this here dog’). Nominalization-based relative clauses also occur on both sides of the head, as in Lower Adi *dəm-nam ami ~ ami dəm-nam* ‘person who was hit’ (section 5.3). Adjectival modifiers, if formed as relative clauses, in the same way may occur on both sides; in languages which allow direct adjectival modification (without a nominalizer), the adjective can only post-modify, as in Lower Adi *mimum kampo* ‘woman beautiful’.

5.2 Referential and relational marking

Tani noun phrases are marked differently depending on whether they represent generic/non-referential, indefinite (new, usually focal), or definite (given, usually topical) information, and on whether they are in subject, non-subject core or oblique functions. Indefinite NPs are typically marked by an article *(*a*)*ko* (< ‘one’), while definite NPs are marked either by an appropriate demonstrative, or by a demonstrative-derived article, usually **hə*. Generic/non-referential NPs are typically unmarked.

Relational marking follows a basically nominative–accusative pattern: subjects tend to be definite, but are unmarked for case, whether functioning as the single intransitive argument or as the actor of a transitive verb. It is common to encounter identifications of a ‘nominative case’ marker *ə* in early descriptions of Tani languages; in fact, this is a reflex of the definite/topic marker **hə* (12).

- (12) Lower Adi

a.	<i>bi</i>	<i>si-kai</i>
	3SG	die-PRF
‘He died.’		

b.	<i>bi-k</i>	<i>əsu</i>	<i>ə</i>	<i>si-kai</i>
	3SG-GEN	mithun	TOP	die-PRF
‘His mithun died.’				

Non-subject core arguments are marked by a form of the accusative case when definite. Objects are distinguished via an unmarked accusative *-*m* (for pronouns, see (8)) ~ **həm* (for full NPs; see (13b)); note that the latter incorporates an earlier demonstrative **hə*,

which explains its restriction to definite noun phrases. Non-referential undergoers are generally predicate-adjacent, and unmarked for case (as in (7)). Indefinite arguments are never marked for case (13a).

(13) Bokar

- a. *ηoo pətay̩ ako(*-m) koypoŋ-pa*
1SG bird INDEF(*-OBJ) see-PFV
'I saw a bird.'
- b. *ηoo pətay̩-hə-m koypoŋ-pa*
1SG bird-DEF-OBJ see-PFV
'I saw the bird.'

In most languages, there is an additional ‘non-agentive’ object marker usually descended from **mi* (possibly itself the source of our pronominal accusative), which functions to highlight the object status of a highly topical/animate participant; an example of this is in (8) (using the innovative Galo form *nē*). Recipients are marked in the same manner as undergoers, with animacy and context apparently sufficing to disambiguate. Relational marking of agents has not been attested, except for demoted agents which are marked by an instrumental postposition in the passive constructions of some languages (14):

(14) Minyong

- bi himjo kiŋ do-ko-hi-to*
3.SG tiger INST eat-PASS-REFL-PFV
'He was eaten by a tiger.'

A genitive case marker **ka* is widespread and marks adnominal noun subordination (10), as well as subjects of both intransitive and transitive nominalized clauses (section 5.3). The allative cases are often identical to, or built on, the locative; the ablative and comitative are likewise usually derived from the locative, e.g. *lokə* ‘from’ < LOC=GEN. An instrumental case marker **kiŋ* is found in some though not all languages (see (12)); in languages which lack it, the instrumental is typically built on the ablative and/or genitive (as in Galo *lokkə*).

5.3 Complex constructions

Tani languages are particularly rich in the variety of complex constructions they exhibit. Nominalization-based constructions are especially prominent; complex constructions not based on synchronic nominalizations include clause-chaining, postpositional and/or adverbial subordination, and complementation. Serial verb constructions are absent from the modern languages.

Nominalization-based constructions include action nominalizations, several types of participant nominalization, internally headed, externally headed and headless relative clause constructions, cleft/focus constructions ((11) earlier), and clausal nominalizations (see Post 2011a). In general, a set of three to five ‘primary’ nominalizers (4.3.2) mark the participant/action and often also the modality status of the nominalized clause. In (15a), the Galo nominalizer *-há* produces an unrealis action nominalization, while in (15b) the same nominalizer produces a gapped externally headed non-subject relative clause construction (note here the genitive subject marking).

- (15) Galo (Post 2011a: 267–9)

- a. *bulù-kò* *?agóm=əm* *porì-há* *káa-máa*
 3.PL-GEN speech=ACC study-NMLZ:IRR.EVENT have/exist-NEG
 ‘There would be no (need/cause) to study their language.’
- b. *higi* *ŋó-kə* *dó-há* *jaráa=ə*
 SPRX.IND 1.SG-GEN Ø_i eat-NMLZ:IRR.NONSBJ goods=COP
 ‘This is the thing that I’ll eat (lit. my to-be-eaten thing).’

Various types of adverbial subordination, including manner, purposive and temporal clauses are marked by postpositions, typically **pə* (< DAT), **lo* (< LOC) and *(*fiə*)*m* (< ACC), and/or by dedicated markers built at least partly on one of them. Similar to nominalized clauses, many postpositionally subordinated clause types also exhibit a genitive subject (16a); however, postpositionally subordinated clause heads are often inflected (unlike most nominalized clause heads), typically lose genitive subject marking when more advanced in their development, and more generally tend to more closely resemble finite clauses than do nominalized clauses (16b).

- (16) a. Galo (Post 2011a: 282)

- tanií=gə* *hobá=əm* *pá-dák=lo* *ŋó* *?in-níi-tó*
 3.SG=GEN mithun=ACC strike.to.cut-COS=LOC 1.SG GO-ABANDON-PFV
 ‘I left as Tani killed the mithun (lit. at Tani’s killing of the mithun).’

- b. Galo (Post 2009: 84)

- nó* *hagií-booló* *hagií-rə*
 2.SG sigh-COND sigh-IRR
 ‘If you sigh, (it too) will sigh.’

Complex and chained events in a thematically united sequence are typically represented using clause chains, most often marked via a non-final suffix (< additive coordinator **lay*). Depending on a variety of factors, non-final predicates may or may not be inflected, and often exhibit additional suffixes (such as completive or attainment suffixes) which serve to ‘fine-tune’ the temporal-episodic sequencing of events, as in example (17) from Upper Minyong:

- (17) Upper Minyong

- dəl* *ammo hidla...* *də i?ala,* *yolu* *məhaalaŋkula...*
dəlo *ammo hit-la* *də i-pa-la* *yolu* *mə-hay-lay-ku-la*
 ANAP.LOC paddy dibble-NF ANAP DO-ATTN-NF 1.PL put-UPWARD-NF-COMPL-NF
 ‘And then we planted paddy...and after doing that, we would leave it to grow...’

Clause chaining and action nominalization are also frequently interwoven to produce episodic breaks amid thematic continuity, as is widely the case in Tibeto-Burman (18).

- (18) Upper Minyong

- də* *bipi* *jə-bo-hi-duj.* *jə-bo-hi-la* *i-nam=ə...*
 ANAP 3.DL live-RECP-REFL-IPFV live-RECP-REFL-NF do-NMLZ:RLS.EVENT TOP
 ‘So, the two of them lived together. Because of that (lit. Living together and that being so)...’

ADDITIONAL ABBREVIATIONS

ADD	Additive	ALTER	Alterphoric	ANAP	Anaphoric
ANTR	Anterior	ATTN	Attainment	AVZR	Adverbializer

COS	Change of state	CQ	Content interrogative	DIR	Direct evidential
EGO	Egophoric	EXP	Experiential	FRUS	Frustrative
INCP	Incipient	IND	Individuator	INFR	Inferential
NAGT	Non-agentive	NF	Non-finite/non-final	POL	Polite
QTAG	Question tag	REP	Reportative	RLS	Realis
SLEV	Same (topographic) level	SPRX	Speaker-proximate		
STAT	Stative	TENT	Tentative		

NOTES

- 1 Simon (1972) and Abraham (1985, 1987) have been field-checked and found to be incomplete and/or unreliable in very many respects. Apatani Language Development Committee (2015) is reliable, but far from comprehensive.
- 2 Kepor Mara, a trained linguist and native Tagin speaker, has indicated that vowel length may be non-contrastive in Tagin, however this remains unconfirmed.

REFERENCES

- Abraham, P. T. (1985). *Apatani Grammar*. Mysore: Central Institute of Indian Languages.
- (1987). *Apatani–English–Hindi Dictionary*. Mysore: Central Institute of Indian Languages.
- Apatani Language Development Committee (2015). ‘Tanw Aguñ: An Audio Dictionary of the Apatani Language.’ Available at: language.apatani.org (accessed 15 August 2016).
- Benedict, Paul K. (James A. Matisoff, contributing editor). (1972). *Sino-Tibetan: A Conspectus*. New York: Cambridge University Press.
- Census of India. (2011). *A-II: Individual Scheduled Tribe Primary Census Abstract Data and its Appendix: Arunachal Pradesh and Assam*. New Delhi: Government of India. Available at: www.censusindia.gov.in/2011census/PCA/ST.html (accessed 4 April 2016).
- Chhangte, Thangi. (1992). Phonology of some Nishi (Dafla) dialects. Paper presented at the 25th International Conference on Sino-Tibetan Languages and Linguistics, Berkeley, University of California.
- Huber, Toni. (2012). Micro-migrations of hill peoples in Northern Arunachal Pradesh: Rethinking methodologies and claims of origin in Tibet. In: *Origins and Migrations in the Extended Eastern Himalaya*, ed. by Toni Huber and Stuart Blackburn. Leiden: Brill, 83–106.
- Modi, Yankee. (2005). Emergence and Impact of Hindi as a Lingua Franca in Arunachal Pradesh. New Delhi: Jawaharlal Nehru University M. Phil Thesis.
- Ou-Yang, Jue-Ya. (1985). 罗巴族语言简志（崩尼-博嘎尔语）[Luobazu Yuyan Jianzhi (*Bengni Boga'eryu/Brief Description of the Luoba Nationality Language (The Bengni-Bokar Language)*)]. Beijing: 民族出版社 [Minzu Chubanshe/Nationalities Press].
- Pegu, Jugendra. (2011). Morpho-syntactic variation in the pagro and sa;jan dialects of the Mising community. In: *North East Indian Linguistics* Vol. 3, ed. by G. Hyslop, S. Morey and M. W. Post. New Delhi: Cambridge University Press India, 155–170.
- Post, Mark W. (2006). Compounding and the structure of the Tani lexicon. *Linguistics of the Tibeto-Burman Area* 29.41–60.
- (2007). A Grammar of Galo. Melbourne: La Trobe University PhD Thesis.
- (2009). The semantics of clause linking in Galo. In: *The Semantics of Clause Linking: A Cross-Linguistic Typology*, ed. by R. M. W. Dixon and A. Y. Aikhenvald. Oxford: Oxford University Press, 75–91.

- (2010). Predicate derivations in the Tani languages: Root, suffix, both or neither? In: *North East Indian Linguistics* Vol. 2, ed. by S. Morey and M. W. Post. New Delhi: Cambridge University Press India, 175–197.
- (2011a). Nominalization and nominalization-based constructions in Galo. In: *Nominalization in Asian Languages: Diachronic and Typological Perspectives*, ed. by F. H. Yap, K. Grunow-Härsta and J. Wrona. Amsterdam: John Benjamins, 255–287.
- (2011b). Prosody and typological drift in Austroasiatic and Tibeto-Burman: In: *Against ‘Sinosphere’ and ‘Indosphere’*. *Austroasiatic Studies: Papers from ICAAL4*. Mon-Khmer Studies Special Issue No. 3, ed. by S. Srichampa, P. Sidwell and K. J. Gregerson. Canberra: Pacific Linguistics, 198–221.
- (2012). The language, culture, environment and origins of Proto-Tani speakers: What is knowable, and what is not (yet). In: *Origins and Migrations in the Extended Eastern Himalaya*, ed. by T. Huber and S. Blackburn. Leiden: Brill, 153–186.
- (2013). Person-sensitive TAME marking in Galo: Historical origins and functional motivation. In: *Functional-Historical Approaches to Explanation: In Honor of Scott DeLancey*, ed. by T. Thornes, J. Jansen, G. Hyslop and E. Andvik. Amsterdam: John Benjamins, 107–130.
- (2015). Sharing innovations across branches: A case study of the Tani subgroup of Trans-Himalayan [Tibeto-Burman]. Paper presented at the International Conference on Historical Linguistics 22. Naples, Italy.
- Post, Mark W. and Roger Blench. (2011). Siangic: A new language phylum in North East India. Paper presented at the 6th International Conference of the North East Indian Linguistics Society, Tezpur University, Tezpur, Assam, India.
- Post, Mark W. and Yankee Modi. (2011). Language contact and the genetic position of Milang (Eastern Himalaya). *Anthropological Linguistics* 53.215–258.
- Post, Mark W. and Kanno Tage. (2013). Apatani phonology and lexicon, with a special focus on tone. *Himalayan Linguistics* 12.17–75.
- Simon, Ivan Martin. (1972). *An Introduction to Apatani*. Gangtok, Sikkim: Government of India Press.
- Sun, Hong-Kai, Ji-Chuan Zhang, Shao-Zun Lu and Jue-Ya Ou-Yang. (1980). 门巴、珞巴、僈人的语言 [*Menba, Luoba, Dengren de Yuyan/The Language of the Menba, Luoba and Dengren*] Beijing: 中国社会科学出版社 [Zhongguo Shehui Kexue Chubanshe/Chinese Social Sciences Press].
- Sun, Tian-Shin Jackson. (1993). A Historical-Comparative Study of the Tani (Mirish) Branch of Tibeto-Burman. Berkeley: University of California PhD Dissertation.

CHAPTER SIXTEEN

MEITEI

Shobhana L. Chelliah

1 LOCATION AND CLASSIFICATION

Meitei, also known as *Manipuri*, *Meithei*, and *Meiteirón*, is a Tibeto-Burman language spoken in the Indian state of Manipur, which is bordered by Myanmar to the east, Mizoram to the south, Nagaland to the north, and Assam to the west and northwest. The state is 22,356 square kilometres, 1,813 square kilometres of which are level country approximately 750 metres above sea level. This level area is populated mainly by the Meitei and 120,000 Muslims who are the progeny of the intermarriage of Muslim traders and labourers with Meitei women. The 20,543 square kilometres of hill territory are populated by about 657,000 people belonging to Tibeto-Burman speaking Naga and Kuki-Chin tribes. There are 1,180,000 native speakers of Meitei, although the number of actual speakers is higher since Meitei is used as a lingua franca in the state (Lodrick 2016). Meitei is the state language.

In Matisoff's (1991) heuristic model, consisting of seven geographical groups, Meitei belongs to *Kamarupan* (from the Sanskrit word *Kāmarūpa* for Assam). Traditionally, the genetic subgroups postulated for this area are Kuki-Chin-Naga, Abor-Miri-Dafla and Bodo-Garo. Earlier classifications put Meitei in a Kuki-Chin (Grierson 1919) or Kuki-Chin-Naga sub-group (Voegelin and Voegelin 1965: 17). However, it has generally been recognized that the Mikir, Mru and Meitei languages do not fit readily into this or other sub-groups of the area. DeLancey (1987: 800) postulates a distinct Mikir-Meitei sub-branch and, Burling (this volume) classifies Meitei as an independent branch.

The description in this entry is of the Imphal dialect of Meitei, which is considered to be the standard. Other documented dialects include Sekmai (H.S. Singh 1988) and Pheyeng (Devi 1988), both spoken close to Imphal, and the Kwatha village dialect spoken on the Indo-Burmese border near Moreh (W.R. Singh 1989). Thoudam (1980) also lists the following dialect names: Kakching, Thanga, Nongmaikhong, Ngaikhong, Moirang, Langthel, Palel and Tokcing. I am not aware of any study which describes these dialects, so it is difficult to say if these are simply geographical terms or truly distinct dialects.

2 PHONETICS AND PHONOLOGY

An inventory of the consonant phonemes in Meitei is given in Table 16.1.

The aspirated and unaspirated voiced stop and affricate series can be attributed to large scale borrowing of Indo-Aryan words. The aspirated affricate /čʰ/ is phonetically realized as [s], [s^h], [ʃ] or [š^h] in native words. Additionally, /č/ is realized as [ts] before /i/, as in [tsin] 'hill'. /l/ has two allophones: [l] and a flapped [r] which occurs in intervocalic position. In native words, voiceless stops contrast with the voiced stops in word medial position only.

An inventory of the vowel phonemes in Meitei is given in Table 16.2.

TABLE 16.1 CHART OF CONSONANT PHONEMES

	Labial	Alveolar	Palatal	Velar	Laryngeal
Stops	p p ^h b b ^h	t t ^h d d ^h		k k ^h g g ^h	
Affricates			č č ^h j j ^h		
Fricatives		s			h
Nasals	m	n		ŋ	
Lateral/Flap		l			
Trill		r			
Semivowels	w		y		

TABLE 16.2 CHART OF VOWEL PHONEMES

	Front	Central	Back
high	i		u
mid	e	ə	o
low		a	

Voiced stops and affricates appear word initially only in Bengali, Hindi, or English borrowings and in Meitei ideophones. There are no indigenous words beginning with /a/. The vowel /e/ occurs in initial position in a few words.

2.1 Syllable structure

The Meitei syllable consists of an onset and a nucleus and may include a coda. The nucleus consists of a vowel. Onsets may be simple or complex. For native words, in word initial position, onsets may be /p, p^h, t, th, č, č^h, k, k^h, m, n, ŋ, w, y, h, l/. Voiced stop onsets occur word medially, and word initially only in ideophones. Vowel initial syllables of prefixes and roots are always preceded by a glottal stop: *əibə* [?ə-?ibə] ‘writer’ from *ə-* ‘attributive’ and *ibə* ‘to write’. Refer to the section on major phonological processes for the treatment of vowel initial suffixes and enclitics. Complex onsets may consist of a voiced un aspirated stop, fricative or voiceless aspirated stop and /l/, /w/ or /y/. Onsets of borrowed words may consist of voiced un aspirated or aspirated stops, affricates and fricatives in both word initial or medial position. Complex onsets are also limited to consonant-liquid or consonant-glide sequences. The coda in native words may consist of /p, t, k, m, n, ŋ, l/. There are no complex codas. This restriction on codas is upheld for borrowed words also.

2.2 Tone

Meitei exhibits a two-way contrast on roots between low and high tone. Suffixes and prefixes have no tone associated with them; instead, the pitch values observed for these are derived through the spreading of lexically specified root tones through the rules of

downstepping and upstepping. Low tone roots trigger upstep, which results in an augmentation of pitch through the word. High tone roots trigger downstep, which results in a downscaling of pitch through the word. High tone roots are marked with an acute accent; low tone roots are unmarked.

2.3 Major phonological processes

Syllable-initial voiceless unaspirated stops are voiced between voiced segments. Voice assimilation fails to apply with prefixation and applies in compounds only when contiguous syllables have sonorants in coda position. To date, voicing assimilation has affected a limited number of compounds but appears to be spreading to all compounds. The changing status of the rule is supported by the presence of doublets (Chelliah 1997:49–59).

Diphthongization and gemination rules work to ensure that syllables have an onset. When a root is concatenated with a suffix that begins with a vowel, a diphthong is formed, e.g. *u-* ‘see’ + *-i* ‘declarative’ result in *uy* ‘sees’. In those cases where diphthongization does not apply, the vowel sequence may be broken up through: (i) the insertion of a glide when the first vowel is front and high, e.g. *pí-* ‘give’ + *-u* ‘imperative’ results in *píyu* ‘Give!?’; (ii) the insertion of a glottal stop when the first vowel is back, e.g. *pu-* ‘carry’ + *-o* ‘solicitive’ results in *pu?o* ‘won’t you carry?’.

Syllables without onsets also arise when stems ending in consonants are concatenated with vowel initial suffixes. In such cases the final consonant is copied and provides the required onset, e.g. *tham-* ‘keep’ + *-u* ‘imperative’ results in *thammu* ‘Keep!?’.

In sequences of identical oral stops the second stop may be dissimilated. Thus in *čót-* ‘go’ + *-u* ‘imperative’, gemination results in *čáttu* which is followed by the dissimilation of the second consonant to l, resulting in *čátlu* ‘Go!?’.

3 MORPHOLOGY

Gender and number are not obligatory categories. Gender is indicated through compounding with a noun which may be optionally affixed by derivational suffixes indicating gender and number. Gender marking occurs in a limited set of quasi-kinship terms where *-pi* ‘female’ and *-pa* a ‘male’ can indicate gender, e.g. *nupi* ‘woman’, ‘wife’ and *nupáa* ‘man’, ‘husband’ where the stem *nu* means ‘human’. The feminine and masculine suffixes roots *pi* and *pa* also appear in adjectives and traditional proper names which are based on adjectives, e.g. the eldest male child is named *Tomba* and the eldest female child is named *Tombi*, from based on *tómn* ‘singletop’. Gender distinctions can also be seen in occupational titles, e.g. *usúba* ‘carpenter’ composed of *u* ‘wood’ *čʰu* ‘make’. Metonymic extension of *pá* and *pi* can be found in Chelliah (2004). Plurality is indicated with *-siŋ* (e.g. *əyásisŋ* ‘children’), but this suffix cannot occur with pronouns.

The sole nominal inflectional category is semantic role marking through cliticization. An example is the inflectional paradigm of *nupá* ‘boy’ which is as follows: *nupána* ‘by the boy’, where *-nə* is agentive/instrumental; *nupábu* ‘the boy’ (patient), where *-pu* is accusative; *nupáda* ‘at/to the boy’, where *-ta* is locative; *nupádəgi* ‘from the boy’, where *-təgi* is ablative; *nupágı* ‘of the boy’, where *-ki* is genitive; and *nupágə* ‘with the boy’, where *-kə* is associative.

Non-category changing derivational prefixes are the *i-* first person possessive, *nə-* second person possessive and *mə-* third person possessive prefixes, respectively. These may be affixed to kinship terms or inalienably possessed nouns: for example, *məpá* ‘his grandfather’ where *pá* is ‘grandfather’, *imít* ‘my eye’ where *mít* is ‘eye’ and *nəkhóy* ‘your foot’ where *khóy* is ‘foot’. The scope of inalienable possession does not include objects or dwellings that

are momentarily possessed, but can be metaphorically extended to include the ancestral home, e.g. *məyum* ‘his house’. The use of the second person possessive prefix with *má* ‘mother’ is taboo, so that ‘your mother’ is expressed as *nəŋgi imá*, literally ‘your my mother’.

Verb roots are bound forms. They must be minimally suffixed by an inflectional suffix or may be nominalized. Verb inflection consists of eight illocutionary mood markers: the unmarked declarative *-i* which contrasts with the assertive declarative *-e* and with other non-declarative suffixes, i.e. the optative *-ke*; the imperative *-u*; the prohibitive *-nu*; the solicitive *-o*; the supplicative *-si*; and the permissive *-sənu*. Only one inflectional morpheme may appear with a given verb root. Verb inflection may be preceded optionally by three derivational categories. Suffixes in these categories are considered derivational because they have restricted productivity and occur closer to the root than prototypical inflectional morphology. An example is given in (1).

(1) *phúgayrəmle*

<i>phú</i>	<i>-khay</i>	<i>-ləm</i>	<i>-lə</i>	<i>-e</i>
beat	-totally affect	-indirect evidence	-perfect	-assertive
Root	1st level derivation	2nd level derivation	3rd level derivation	Inflection

‘destroyed’

First level derivation consists of eight suffixes that describe the extent to which an agent desires or intends to affect some object or the direction (up, down, in or out) and manner (see later for examples) in which an action is performed. Second level derivation consists of suffixes that occur in free order with respect to each other. They include deictic markers, which indicate where an action is carried out with respect to the speaker, e.g. *-lə* indicates that an action has occurred at the site of the speaker, *-lək* marks the emergence of an entity towards the speaker after the completion or successful instigation of an action, and *-lu* indicates an act has been performed away from the speaker. The deictics also cover aspectual oppositions, i.e. *-lə* indicates prospective aspect, *-lək* indicates perfect aspect, and *-lu* indicates inchoative aspect. The first and second level derivation suffixes are most likely grammaticalized roots. Examples of the first level derivational suffixes are *-khay* ‘totally affect’ from *kháy-* ‘cut with a knife’; *-that* ‘partially affect’ from *thát-* ‘break by pulling’; *-hət* ‘affect with undue psychological or physical influence’ from *hát-* ‘kill’. Examples of second level derivational suffixes are *-min* ‘comitative’ from *min-* ‘be together’; *-pi* ‘v to or for someone other than self’ from *pi-* ‘give’; *-čə* ‘v for sake of self’ from *sá* ‘body’; *-mən* ‘v in excess’ from *man-* ‘be greedy’; *-həw* ‘inceptive’ from *həw-* ‘begin’, ‘grow’; *-ləm* ‘indirect evidence’ from *ləm-* ‘path’; and *-lək* ‘distal’ from *lak-* ‘come’.

Third level derivation consists of suffixes that signal modality and aspect, such as ‘potential’, ‘non-potential’, ‘necessity’, ‘obligation’, ‘probability’, ‘intention’, ‘progressive’ and ‘perfect’. The number, person, and gender of arguments is not indicated on the verb.

There are three category-changing derivational verb affixes. The prefixes *mə-* and *khut-* derive nouns from verbs, e.g. *məčá* ‘small one’ from *čá-* ‘be small’ and *khutká* ‘manner of climbing’ from *ká* ‘climb’. The suffix *-pə* changes verbs to nouns, which form the base for adjectives (see discussion in section 4) and relative clauses (see (14)).

Nominal compounding is highly productive, resulting in both right- and left-headed endocentric or exocentric compounds, e.g. *phigá* ‘undergarment’ from *phi* ‘cloth’ and *kha-* ‘be under’, *khóylám* ‘footpath’ from *khón* ‘foot’ and *lám-* ‘path’ and *mítña* ‘organs’ from *mít* ‘eye’ and *na* ‘nose’. There are no right- or left-headed verbal compounds; the common second stems in left-headed compounds have grammaticalized to form the first

and second level derivational suffixes as described above. The few verbal compounds that exist are exocentric, e.g. *čáthák* ‘dine’ from *čá-* ‘eat’ and *-thák* ‘drink’.

Lexical collocations are formed through a variety of re-duplicative processes and echo word formation. A constituent or part of a constituent can be either partially or fully duplicated or a constituent or part of a constituent can be paired with a rhyming word, e.g. *čótphám lakphám* ‘place of much activity’, literally ‘go -place come -place’.

4 MAJOR LEXICAL CATEGORIES

Nouns and verbs can be distinguished on formal grounds: nouns are free roots whereas verbs are bound roots; moreover, and inflectional and derivational morphology for these two categories come from mutually exclusive sets.

Verb roots may also be used to form verbal nouns, adjectives and adverbs. Verbal nouns are formed through the suffixation of a nominalizer to a verb root, e.g. *čát-* ‘go’ with the nominalizer *-pə* becomes *čátpə* ‘to go’, ‘going’. An adjective is derived through the affixation of the attributive prefix *ə-* to a verbal noun, e.g. the adjective *əčawba* ‘big’ is derived from the stative verb *čaw-* ‘be big’, the nominalizer *-pə* and the attributivizer *ə-*. Manner adverbs are formed through the suffixation of *-nə* ‘adverbial’ to a verb root: for example, *loynə* ‘completely’, ‘all’ from *loy* ‘complete’, ‘finish’. A manner adverb can be negated with the suffixation of the negative *-tə* before it is adverbialized with *-nə*: thus *wánə* ‘sadly’ becomes *wádənə* ‘not sadly’. Locative adverbs are derived through the prefixation of *mə-* ‘nominalizer’ to noun or verb roots, as in *məkha* ‘below’, ‘underneath’ from *kha* ‘south’ and *mətuy* ‘behind’ from *túy-* ‘be in back’. Temporal adverbs are frozen compounds: *yəsi* ‘today’, *yərəŋ* ‘yesterday’ and *həyəŋ* ‘tomorrow’ are most likely bimorphemic although the meaning of the individual roots is unknown.

In nominalized verb–verb constructions, as illustrated by (16), the nominalized verb carries the primary semantic meaning, while the second verb is semantically bleached and functions like an auxiliary. Verbs that function as auxiliaries include *hów-* ‘start’, which indicates the initiation of an action, *loy-* ‘finish’, which indicates the completion of an action, *ka-* ‘roast to a burn’, which indicates excessive carrying out of an action, and *ya* ‘possible’, which indicates possibility.

The singular personal pronouns are *áy* ‘I’, *náy* ‘you’ and *má* ‘he/she’. Possessive pronouns are formed through the suffixation of *-ki* ‘possessive’ to these personal pronouns: *žygi/náygi/mági yum* ‘my/your/his or her house’. The plural is indicated by *khoy* ‘this and others like this’. For the second and third person, the plural forms are not built on the regular personal pronouns, but with possessive prefixes, e.g. *žykoy* means ‘we’, *nəkhoy* means ‘you all’ and *məkhoy* means ‘they’. The dual is indicated by *-bani*. Thus *ibani* means ‘we two’, *nəbani* means ‘you two’ and *məbani* means ‘them two’. Indefinite pronouns are composed of question words and *-no* ‘inquisitive’, *-su* ‘also’, or *-kum* ‘like’, e.g. *kəna* ‘who’ in *kənano* ‘someone’, *kənasu* ‘nobody’ and *kənagum* ‘someone’. There is no relative pronoun to mark relative clauses; instead, the relativized noun occurs directly after a nominalized clause. Reflexive pronouns are formed on noun *sá* ‘body’ and the possessive prefix, e.g. *isá* ‘myself’, *nəsá* ‘yourself’ and *məsá* ‘herself/himself/itself’. To form an emphatic reflexive, the sequence *-məknə* consisting of the quantifier *-mək* ‘only’ and the agentive/instrumental *-nə* is suffixed to the reflexive pronoun, e.g. *nəsáməknə* ‘you yourself’.

The determiners *-si* ‘proximate’ and *-tu* ‘distal’ are stems that function as enclitics. *-si* indicates that the object or person being spoken of is near or currently seen or known to

be near, even if not viewable by the speaker, or is currently the topic of conversation; *-tu* indicates something or someone not present at the time of speech or newly introduced in the conversation. There are two pronouns based on these stems: *ədu* ‘it (there)’ and *əsi* ‘it (here)’ where *ə-* is the attributive prefix. Two other demonstrative pronouns based on *si* and *tu* are *məsi* ‘this’ and *mədu* ‘that’.

There are six basic question word forms all of which begin with *kə-*, from the Proto-Sino-Tibetan interrogative **ka* (Benedict 1984). The question words are morphologically frozen nominals and can be inflected like nouns. Thus *kəday* ‘where’ (approximately) can occur as *kədaydəgi* ‘from where’ and *kədaywaydə* ‘from around where’; *kədomdə* ‘which way’ can occur as *kədomdəgi* ‘from which side’; *kəna* ‘who’ can occur as *kənagə* ‘with whom’, *kənagi* ‘whose’ and *kənadə* ‘to whom’, and *kənadəgi* ‘from whom’; *kəya* ‘how many’ (for count nouns) can occur as *kəyadə* ‘for how much’ (price), *kəyanə* ‘how much needed’, *kəyarək* ‘how many times’ and *kəyasubə* ‘of what number’ (ordinal); *kərəm/kəm* ‘how’, ‘in what way’ can occur as *kərəmkandə* ‘at what time’, *kərəmdəwnə/kəmdəwnə* ‘how’, ‘by what means’ and *kərəmbə* ‘which’; *kəri* ‘what’ can occur as *kərigə* ‘with what’, *kərigi* ‘of what’, *kəridə* ‘on what’, ‘for what reason’, *kəridəgi* ‘from what’, and *kərinə* ‘by what means’, ‘with what instrument’. *Kəyám* ‘how much’ is used for mass nouns.

Most quantifiers in Meitei are lexicalized forms consisting of the unproductive prefix *khV-* (where the vowel can be *ə*, *i* or *u*). These are *khərə* ‘some’ which indicates an indeterminate amount; *khitáy* ‘ever so little’, ‘a particle’ (composed of *khit* ‘a little’ and *tāy* ‘exclusive’) of some tangible material; and *khájikta* which indicates a short amount of time.

Enclitics in Meitei fall into six categories: determiners (*-si* proximate determiner, *-tu* distal determiner); semantic role markers (listed in the morphology section); the copula *-ni*; mood markers (listed in the morphology section); inclusive/exclusive markers (*-ti* ‘delimitative’, *-mək* ‘only’, *-tā* ‘exclusive’) and pragmatic peak markers (*-nə* ‘contrastive’, *-pu* ‘adversative’); and attitude markers (*-ne* ‘shared information’, *-tə* ‘contrary to expectation’, *-ye* ‘confirmative’ *-hé* ‘exasperative’, *-ko* ‘invariant tag’).

5 EVIDENTIALITY AND EPISTEMIC MODALITY

Indirect evidence and degrees of certainty are indicated through derivational verb morphology. For example, as illustrated in (2), Bhat and Ningomba (1997: 270) note that there is an implication with *-hə w* ‘inceptive’ which implies that the speaker is a witness to the initiation of an action.

- (2) *tūren pahəwwí*
túlen pa-həw-í
river overflow-start-declarative
'The river began overflowing.'

As illustrated in (3), the shared information marker *-ne* – which means ‘as you know ...’ – suggests that a proposition contains shared information known to be true by both the speaker and hearer.

- (3) *Kakčijdənīne*
Kakčij-tə-ni-ne
Kakching-locative-copula-shared information
'(As I'm sure you know), it is to Kakching (that I am going).'

Finally, evidential values are indicated in the complementation system through choice of nominalizer or complementizing quotative. The quotative *háyna* is used to indicate that only the speaker's belief or hearsay evidence can support the truth of the subordinated proposition; *háybədu* is used when there is an eyewitness to the subordinated proposition. As illustrated in the nominalized clause sentence in (4), the lexical nominalizer *-jat*, which literally means 'type', indicates knowledge that an action or state has occurred or come into being on the basis of some indirect evidence; the speaker sees that some object is battered and from this draws the most probable conclusion that it was beaten.

- (4) *məsi* *phúrbəbəjatni*
 mə-si phú-lóbə-jat-ni
 noun marker-determiner beat-having-nominalizer-copula
 this is a type of having been beaten
 'It looks like it might have been beaten.'

6 SYNTAX

The Meitei clause minimally consists of a verb and the arguments (i.e. noun phrases) the verb subcategorizes for – there is no evidence in Meitei for a verb phrase constituent. As indicated by the asterisk which follows the noun phrase in (5), the verb may occur with any number of noun phrases.

- (5) sentence → noun phrase* verb

The most common word order is agent/actor verb. The patient may occur initially or medially. Both agents and patients can be omitted from a clause. Compared to other semantic roles, however, agents are more likely to be zero anaphors (i.e. NPs that fulfill the argument structure of a predicate but are not phonetically realized in the same clause as that predicate). As discussed in Chelliah (2013), the fact that agents are animate and tend to be referenced often in natural speech lend to their cognitive accessibility and contribute to the frequency to which they can be omitted. The order of the non-core arguments is free and determined by pragmatics, constituents may occur after the verb as afterthoughts, and focused constituents occur sentence initially.

A noun phrase must consist of a noun and may include one or more adjectives. A noun phrase may include either a numeral or quantifier, but not both. The order of these constituents within the noun phrase is relatively free. There are no numeral classifiers in Meitei.

6.1 Semantic role marking and information packaging

Semantic role is indicated by these enclitics: *-nə* for instrument, agent (an argument of two and three-place predicates and morphologically derived causatives) and actor (an argument of stative and copular predicates); *-pu* for patients; *-tə* for experiencer/goal; zero marking for path and theme; *-təgi* for ablative; *-ki* for possessive; and *-kə* for associative. Patients and agents/actors are marked differentially: patients are marked when specific; agents/actors are marked when they participate in an action that is unusual or contrary to expectation. Semantic role is often obscured through the overlay of a system of pragmatic marking which may replace the semantic role marker with one of these enclitics: *-tə* 'exclusive', *-ti* 'delimitative', *-tu* 'distal', *-si* 'proximate' or *-su* 'inclusive'. Pragmatic information can also be signaled by adding one of these enclitics to a semantic role marker and/or manipulating word order, e.g. moving topics to clause initial position. As described in Chelliah (2009) the semantic role

makers agentive, patient, associative, and locative have been repurposed through metonymic extension to function as markers of new information: agentive>contrastive focus; patient>adversative; associative>unanticipative; locative>contrary to expectation.

In some instances, the pragmatic system makes recovery of grammatical relations difficult so that sentences may often have more than one interpretation. As in example (6), the larger discourse context must be used to recover the intended meaning.

- (6) *śydi* *Ramnə* *phúnijjí*
 éy-ti Ram-nə phú-niŋ-i
 1sg-delimitative Ram-contrastive beat-wish-declarative
 ‘It is Ram who wants to beat me (over all of us).’ or
 ‘It is Ram that I (over all of us) want to beat.’

The contrastive *-nə* and the agentive *-nə* are distinguished on the basis of distribution in that *-nə* contrastive can mark any sort of argument, e.g. in (7) where it marks the patient.

- (7) *ə ybunə Ramnə nujsirə bə di*
 ey-pu-nə
 I-patient-contrastive Ram-nə
 Ram-contrastive nupsi-lə-pə-ti
 love-prospective-nominalizer-delimitative

 phágə də wni
 phá-kə-təw-ni
 good-potential-obligation-copula
 ‘If Ram (not Chaoba) loved me (and not Sita), it would be good’.

6.2 Questions

Yes-no questions are formed by the suffixation of the interrogative enclitic *-lə* to a noun as in (8) and to verbs with irrealis modality, i.e. suffixed by one of the following: *-kə* ‘potential’, *-loy* ‘nonpotential’, *-təw* ‘certain future’ and *-ta* ‘strong possibility/obligation’ act as nominalizers.

- (8) *Tomba u kókpə* *míra*
 Tomba u kék-pə mí-lə
 Tomba u cut-nominalizer man-interrogative
 ‘Is Tomba the man who is a woodcutter?’

(9) *čát?əbra*
 čát-tə-pə-lə
 go-negative-nominalizer-interrogative
 ‘(You) didn’t go?’

Tag questions are formed by using the negative form of a positive verb or the positive form of a negative verb as the tag. The tag is followed by the interrogative marker. This is illustrated in (10).

- (10) *učék paybə* *ŋémde* *ŋémbra*
 učék pay-pə ŋém-tə-e ŋém-pə-lə
 bird fly-nominalizer able-negative-assertive able-nominalizer-interrogative
 ‘Birds cannot fly, can they?’

Alternatively one of two invariant tag markers can be used: the lexical item *nəttra* ‘is it not so?’ which requires a response, and the enclitic *-ko* ‘right?’, ‘don’t you agree?’ which can be answered by gesturally expressed agreement or dissent.

There are three ways to form a content question. As shown in (11), a question word can occur with an inflected verb. Second, as in (12), the question word occurs in a sentence that ends with *-no* ‘inquisitive’. Third, *-no* is affixed directly on the question word as in (13).

- (11) *kəna káytheldə* *čátlı*
 kəna kéy-thel-tə čátlı
 who grain-display-locative go-progressive
 ‘Who goes to the market?’
- (12) *kənagi* *yénawno*
 kəna-ki yén-naw-no
 who-possessive chicken-new-inquisitive
 ‘Whose chicken is this?’
- (13) *kənagino*
 kəna-ki-no
 who-possessive-inquisitive
 ‘Whose is (it)?’

6.3 Phrase structure of subordinated sentences

There are five formally distinct types of subordinate clauses: nominalized complements, determiner complements, quotative complements, subordinate adverbial clauses and adverbial participials. The nominalized clause is formed by suffixation of the nominalizer *-pə* to an uninflected verb as in (14). It is used in relative clause formation where the relativized argument occurs to the right of a nominalized verb as in (15). The determiner complement is a nominalized clause followed by the determiner *-sí* ‘this’ or *-tu* ‘that’ as in (16). The quotative complement uses forms of the quotative as the complementizer, as shown in (17).

- (14) *áy čak* *čábə* *həwre*
 áy čak čábə həw-lə-e
 1sg cooked rice eat-nominalizer start-perfect-assertive
 ‘I have started eating rice.’
- (15) *kolom páyrəbə* *nipa*
 kolom páy-lə-pə ni-pa
 pen hold-perfect-nominalizer person-male
 ‘the boy who held the pen’
- (16) *isij thákpədu*
 isij thák-pə-tu
 water drink-nominalizer-determiner
 ‘from that drinking water’
- (17) *məhák-nə* *thoyre*
 mə-hák-nə thoy-lə-e
 3sg-present participant-agentive win-perfect-assertive

she	had won	
<i>háyబəsi</i>	<i>Tombina</i>	<i>kháj yí</i>
háy-pə-si	Tombi-nə	khárj-í
say-nominalizer-determiner	Tombi-agentive	know-declarative
‘Tombi knew that she had won.’		

Adverbial subordinators are derived from the associative, possessive locative, ablative and instrumental markers, following a pattern that is common in Tibeto-Burman languages, as discussed in detail for 26 languages of the family by Genetti (1988). Thus *-kə* ‘associative’ is used to mean ‘at the same time as ving’, *-ki* ‘possessive’ is used to mean ‘for the purpose of ving’, *-tə* ‘locative’ is used to mean ‘after ving’, *-tagi* ‘ablative’ is used to mean ‘resulting from ving’, *-nə* ‘instrumental’ is used to mean ‘by ving, because of ving’. An example is given in (18).

- (18) *áykhoydə* *lakpədə*
 áy-khoy-tə lak-pə-tə
 1sg-plural-dative come-nominalizer-dative
 ‘after coming to our place’

Adverbial participials are suffixed to non-nominalized verbs. The participializers are morphologically complex, lexicalized units. The participial marker *-túna* ‘ving’ is composed of the determiner and instrumental markers, *-tənə* ‘by ving’ is composed of the locative and instrumental markers, *-nəbə* ‘for ving’ is composed of the instrumental marker and the nominalizer *-pə*, *-lágə* ‘after ving’ is composed of the perfect and associative markers, and *-lábə* ‘having ved’ is composed of the perfect marker and the nominalizer *-pə*. Conditional clauses are indicated by the sequence *-rəbədi*, which consists of the prospective marker, a nominalizer and a delimitative marker. Examples are given in (19) and (20).

- (19) *áy mági mətúŋ indúnə laȝí*
 áy má-ki mə-túŋ in-túnə lak-í
 1sg 3sg-possessive nominalizer-behind follow-ing come-declarative
 I his behind following came
 ‘I followed him.’
- (20) *ləyrəbədi*
 ləy-lə-pə-ti
 be-prospective-nominalizer-delimitative
 ‘if it is’

Subordinate clauses can be combined and ordered within sentences in a variety of ways. In the most common orders for determiner–complementizer complement constructions, the complement is sentence initial and followed by the main clause, or is embedded between the verb of the main clause and its arguments. Subordinate clauses can be combined, exhibiting a clause-chaining discourse structure as in (21) where several adverbial clauses are strung together. The ampersand ‘&’ marks the beginning of each clause.

- (21) & *čithi ədo parub ədudə* ‘on reading that letter’
 & *senpannəbə nupá əduñə čithi – purək ɿibə nupá ədudə sen píkho háybənínə*
 ‘since the treasurer was told to give money to the person who brought the letter’
 & *məraybəkphábə dolaypabədo ədudə* ‘and so to that fortunate gatekeeper’
 & *sen píkhəre* ‘(he) gave the money.’

‘On reading the letter, the treasurer, as he was instructed in the letter to do, gave the man that carried the letter, the fortunate gatekeeper, the money.’

6.4 Pre-twentieth century Meitei manuscripts

The Meiteis hold a vast treasure of documents recording all aspects of their history, religion, material cultural, and intellectual cultural practices. The content of these manuscripts includes the history, origin, and migration of the Meitei to the region; their belief systems and the rituals and rites used to observe those beliefs; material culture, botanical, zoological, and astronomic documentation; and other cultural practices such as child and adult games, sports, visual and performing arts, and design and craft (Chelliah and Ray 2002). The colophons of most manuscripts have been lost or are illegible, so data is approximated through the mention of kings and their reign, style of narration, extent of borrowings from Bengali, and whether the indigenous script Meitei Mayek or the Bengali script is used. A sample list of specific manuscripts can be found (Singh 1984). Abbreviation conventions used by different writers include omitting inferable material, such as vowel diacritics, repeated consonants, and short-forming commonly used phrases. The result is a highly condensed and occasionally difficult to read orthographic system. There are currently four institutional holders of old Manipuri manuscripts: the Manipuri State Archive which holds about 4000 manuscripts, of which 620 are in microfiches; the Manipuri State Kala Akademi which hold 615 manuscripts; Manipur University which holds approximately 700 manuscripts; and the Indira Gandhi National Centre for Arts in Delhi holds 254 microfiches, of which duplicates exist at the Manipuri State Archives. In addition to these institutional collections, there are several private traditional curators spread across nine districts in Manipur who hold, according to a survey by the Manipuri State Archives, around 38,000 to 40,000 manuscripts. The earliest manuscripts are hand-written in Meitei Mayek on *agarbak*, a paper derived from the bark of a local tree and written on with bamboo stencils or charcoal-darkened paper written on with a soap-stone pencil. Limited transliterations and translations of these manuscripts exist. See Chelliah and Ray (2002).

FURTHER READING

- Chelliah, S.L. (2011) Introduction. In S.L. Chelliah and G. Hyslop (eds.) *Special Issue on Optional Case Marking in Tibeto-Burman*. *LTBA*. 34.2 (October 2011) and 35.1 (April 2012).
- Chelliah, S.L. (2017) ‘Ergativity and unique differential marking in Tibeto-Burman’, in J. Coon, D. Massam and L. Travis (eds) *Ergativity*, Oxford: Oxford University Press Handbook series.
- Hodson, T.C. (1908) *The Meitheis*, Delhi: B.R. Publishing Corporation.
- Pettigrew, R.W. (1912) *Manipuri (Mitei) Grammar*, Allahabad: Pioneer Press.
- Primrose, A.J. (1887) *Manipuri Grammar*, Manipur: Government Press.
- Singh, C.Y. (1981) ‘Word order in Meiteilon’, in C. Chu, W.S. Coblin and F. Tsao (eds) *Papers from the Fourteenth International Conference in Sino-Tibetan Languages and Linguistics*, Taiwan: Student Book Company, 323–341.
- Singh, C.Y. (1984) ‘Some aspects of Meiteilon (Manipuri) syntax’, PhD dissertation, Jawaharlal Nehru University, New Delhi.
- Singh, C.Y. (1991) ‘Causativization in Meiteilon’, *LTBA* 14.2: 133–144.

- Singh, N.K. (1964) *Manipuri to Manipuri and English Dictionary*, Imphal: Government of Manipur.
- Singh, N.N. (1987) ‘A Meitei grammar of roots and affixes’, unpublished PhD dissertation, Manipur University, Imphal.
- Thoudam, P.C. (1989) ‘Conditioning factors for morphophonemic alternations of manner in Meiteiron’, *LTBA* 12.2: 93–100.

REFERENCES

- Benedict, P.K. (1984) ‘Proto-Sino-Tibetan interrogative *ga(ng) ~ *ka’, *LTBA* 8.1: 1–10.
- Bhat, D.N.S. and Ningomba, M.S. (1997) *Manipuri Grammar*, Munich: Lincom Europa.
- Chelliah, S.L. (1997) *A Grammar of Meitei*, Berlin: Mouton.
- Chelliah, S.L. (2004) ‘Polysemy through metonymy: the case of Meithei pi- “grand-mother”’. *Studies in Language* 28.2: 363–386.
- Chelliah, S.L. (2009) ‘Semantic role to new information in Meithei’, in Jóhanna Barðdal and Shobhana Chelliah (eds) *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case*, Amsterdam: John Benjamins, 337–400.
- Chelliah, S.L. (2013) ‘Predicting reference form: A Pear Story Study of information status, thematic role, and animacy in Meithei (Manipuri, Meiteiron)’, in T. Thornes, E. Andvik, G. Hyslop, and J. Jansen (eds). *Functional-Historical Approaches to Explanation: In Honor of Scott DeLancey*, Amsterdam: John Benjamins, 223–236.
- Chelliah, S. L. and Ray, S. (2002) ‘Early Meitei manuscripts’, in Christopher I. Beckwith (ed.) *Medieval Tibeto-Burman Languages. PIATS 2000: Proceedings of the Ninth Seminar of the International Association for Tibetan Studies*, Leiden: E.J. Brill, 59–71.
- DeLancey, S. (1987) ‘Sino-Tibetan languages’, in B. Comrie (ed.) *The World’s Major Languages*, London: Croom Helm, 797–810.
- Devi, P.R. (1988) ‘A comparative study of Imphal and Phayeng dialects in morphophonemic changes in compounding’, unpublished MPhil thesis, Manipur University, Imphal.
- Genetti, C. (1988) ‘From postposition to subordinator in Newari’, in B. Heine and E. Traugott (eds) *Grammaticalization*, Philadelphia: John Benjamins, 161–191.
- Grierson, G.A. (1919) *Linguistic Survey of India and the Census of 1911*, Calcutta: Superintendent Government Printing, India (reprinted 1967 by Motilal Banarsi das: Delhi, Varanasi, Patna).
- Lodrick, D.O. (2016) ‘Manipur’ *Encyclopædia Britannica Online*. Retrieved 10 January 2016, from www.britannica.com/place/Manipur.
- Matisoff, J.A. (1991) ‘Sino-Tibetan linguistics: present state and future prospects’, *Annual Review of Anthropology* 1991.20: 469–504.
- Singh, H.S. (1988) ‘A comparative study of Meiteiron dialects (compounding)’, unpublished MPhil thesis, Manipur University, Imphal.
- Singh, N.K. (1984) *A catalogue of Manipuri manuscripts*. Imphal: The Manipur State Kala Academy.
- Singh, W.R. (1989) ‘The Kwatha dialect of Meitei’, *LTBA* 12.2: 101–122.
- Thoudam, P.C. (1980) ‘Grammatical sketch of Meiteiron’, PhD dissertation, Jawaharlal Nehru University, New Delhi.
- Voegelin, C.F. and Voegelin, F.M. (1965) ‘Languages of the world: Sino-Tibetan fascicle three’, *Anthropological Linguistics* 7.2.3.1: 1–57.

CHAPTER SEVENTEEN

TANGSA¹

Stephen D. Morey

1 INTRODUCTION

Tangsa, referred to as Tangshang in Myanmar, is one of the languages listed as belonging to the ‘Northern Naga’ (also called ‘Konyak’) grouping within Tibeto-Burman. Tangsa sub-tribes² live on both sides of the India–Myanmar border, in and around the Patkai range. Traditionally mountain people³ mostly in what is now Myanmar, they have been moving into more fertile plains areas on both the India and Myanmar sides for the last several hundred years, so that now some sub-tribes are found only in India, some on both sides of the border, and some only in Myanmar.

The suggestion that there is a linguistic grouping, of which Tangsa was later analysed to be part, was probably first made by Damant (1880: 249–50)⁴ and then further developed by Sten Konow in Grierson (1903: 329), where the group is called the ‘Eastern Subgroup of Naga’. The earliest linguistic description of a Tangsa variety was a word list and some grammatical information about Mueshaungx (Mossang)⁵ and a small amount of information on Shecyü (Shangke) in Needham (1897), where they are termed respectively Môshâng Nâga and Shâng'gê Nâga. In the twentieth century, the most important information about Tangsa were words noted in Weidert (1979, 1987) and notes on different varieties in Das Gupta (1980).

In India, the Tangsa are one of a number of groups categorised as ‘other Naga tribes’ in the Indian Constitution.⁶ The person who was most responsible for the name, Bipin Borgohain, then a Political Officer, in what was then termed the Tirap Frontier Division, wrote:

the once subjugated but now liberated and resurgent lovable Tangsa (Tang = Mountain, sa = person’), a word which was specially coined by the undersigned and accepted by the tribe and the Government for official use ...

(Foreword written by Bipin Borgohain in Barua 2013 (1991): viii)

Prior to the coining of ‘Tangsa’, in India at least, each of the different sub-tribes was known by its own name, often in combination with Naga, as on pre-World War II British maps (Survey of India 1927),⁸ where tribal names include Moklum Naga, Mosang Naga, Yogi Naga, Ponthai, Longchang Naga. If there was a native term that was an overall name for all these groups, or at least a significant number of them, it appears to have been Hawa.⁹

Neighbouring Changlang district is Tirap district, where a number of varieties, probably less diverse than Tangsa, are grouped together under the name Nocte. The distinction between Tangsa and Nocte is not entirely linguistic; some sub-tribes, such as Hakhun and Phong (Ponthai) are grouped under Tangsa when they live in Changlang district or across the border in Assam, and Nocte if they are in Tirap district.

Two other language names are important in this discussion: Tutsa, formerly included under Nocte, but recognised as a separate tribe in around 2000, and Wancho, also a

separate tribe under the Indian constitution. Each of these four groups are listed in *Ethnologue* (Lewis *et al.* 2014) and consequently have ISO codes, respectively ISO 639-3:nst ('Naga, Tase'), ISO 639-3:njb ('Naga, Nocte'), ISO 639-3:tvt ('Naga, Tutsa') and 639-3:nnp ('Naga, Wancho'). Each of the languages within the 'Northern Naga' group has been named by the ISO, following *Ethnologue*, with the element Naga as the first part of the name. The association of Tangsa with 'Naga' is contested by some members of the community, particularly some who have not converted to Christianity, or who have been in India the longest (see Burling 2003: 172 for a discussion of the issue of Naga as a linguistic term). Nevertheless, a majority of those we have spoken to on both sides of the border regard the Tangsa / Tangshang as Naga.

The ISO and *Ethnologue* (Lewis *et al.* 2014) name of Tangsa/Tangshang, 'Naga, Tase', was adopted because /tase/ is the pronunciation of the name in the Chamchang (Kimsing) variety of Tangsa, the first for which a Bible translation was undertaken. In Chamchang, the rhyme -*aj is realised as /-a/ and -*a is usually realised as /-i/, except that in the words for 'child' and 'father' it is irregularly /-e/, hence Tangsa is realised as /tase/.

While the names of the sub-tribes are the same on both sides of the India–Myanmar border, the names for larger groupings are different. The name Rangpan or Rangpang has been used since the nineteenth century (Statezni unpublished), and some of the earlier authors suggested that Rangpan was the name used in Myanmar for the groups called Tangsa in India, which is in the northern part of the Tangshang area (Dewar 1931; Defence Department, Government of Burma 1942).

The name Haimi, Haimye or Heimi first appeared in the *Census of India* (1931), where it refers to groups south of the Rangpan (Statezni unpublished). Saul (2005) produced a language map of all the languages associated with people who identify as Naga, showing an undivided group marked as Tangsa on the India side and Heimi on the Myanmar side, implying that these two are basically the same thing. Within Heimi, one sub-tribe is marked, Yonkon. Our understanding from speakers in Myanmar is that the two groups do indeed include the same sub-tribes.¹⁰ Several consultants have suggested etymologies for Heimi, 'upper people' (Statezni, personal communication) or 'good person'.

In April 2003, the name *Tangshang* was inaugurated at a mass meeting by the Naga Tradition and Culture Committee, Nanyun Township. The name was settled upon after the Government asked for 'a general name for the tribe'. The term Tangshang is *not* cognate with Tangsa, being derived from *Tangnyu Wang* and *Shangnyu Wang*, two siblings in the oral history of the people. All the Tangshang are regarded to be the descendants of these two siblings. Nowadays Shangnyu Wang refers to the groups in the north and Tangnyu Wang to those in the south.¹¹ Within Myanmar, it appears that Tangshang includes groups which in India would be separated as either Tangsa, Nocte, Tutsa or Wancho. (For more information about Tangshang in Myanmar, see Statezni 2013.)

1.1 The Tangsa sub-tribes

Within Tangsa/Tangshang, there are around 80 sub-tribes, each of these having its own name, for example Shecyü /ʃe²tɕeu²/ . These names have several different forms, firstly an autonym (such as Shecyü), and then exonyms, the pronunciation of the name in each one of the other varieties, so that, for example, the Cholim word for Shecyü is /k^ha²k^hja²/ and the Joglei word is /ʃaken/. In addition to these, each group has what we term a 'general name', apparently an outsiders' name that is not known to be in any one of the different varieties. The 'general name' for Shecyü is Shangke. In this paper when referring to sub-tribes, we will use the autonym first and the 'general name' in brackets, if there is a

difference, or if that difference has been recorded.¹² Table 17.1 details some of the variation of sub-tribe names.

Some of these 80 sub-tribes speak linguistic varieties that are almost identical, such as Chamchang (Kimsing) and Shecyü (Shangke), while others are not at all mutually intelligible, such as Hakhun and Tikhak. The most up-to-date available list of these sub-tribes can be found at the Wikipedia site, Tangsa people.¹³ The average population per sub-tribe is probably between 1,000 and 1,500.

The sub-grouping of these 80 linguistic varieties remains a work in progress, since for most of these varieties we only have short word lists (for more discussion of this, see Morey 2015). Almost no research had been done on Tangsa prior to the commencement of our DoBeS project in 2007, and the almost simultaneous commencement of the surveys by the Linguistic Society Naga Survey Team, Myanmar,¹⁴ who have collected around 50 word lists. Our project has collected around 30 word lists on both sides of the border, marked where possible for tone with the assistance of native speakers who have clear intuitions for tonal categories (see the discussion in section 3.3). We have also collected a large number of texts, many of which have been transcribed and translated.

On the India side, there was, until recently, representation of the following distinct groups within Tangsa:

- (1) Tikhak group (Tikhak, Yongkuk, Longchang, Kato, Nokja)
Muklom group (Muklom, Hawoi)
Phong (Ponthai)
Pangwa varieties

Tikhak, spoken by communities that have been in India for several hundred years, is a discrete group definable on linguistic criteria, such as the agreement system (see section 4).¹⁵ Although Kato is almost extinct as a separate variety and Nokja is spoken by very few people, the other three varieties have significant numbers of villages and are above average in population for Tangsa groups. According to Simai (2008: 25), the Tikhak group was the first Tangsa to arrive in India, with the Koto or Kato, Tikhak, Yongkuk, Lungchang and Kengjung migrating into India between 1300 and 1500. There is no independent evidence to confirm such an early date. These groups were certainly well established in the hills of Changlang by the time the British arrived in 1827 (along with Mueshaungx, Joglei and Muklom) and Kato and Yongkuk had already moved into the plains where they are shown in the 1927 map (Survey of India 1927). There are no Tikhak group speakers in Myanmar.

Muklom is also a large sub-tribe, very similar to the smaller group Hawoi. At present the most substantial information about this group is in Das Gupta (1980); Muklom people have told me that they have been in India for several hundred years.

Phong appears to be a distinct variety with no very close relatives within Tangsa, and although bearing similarities to Nocte, it is also distinct from it. It is spoken in around six villages. There are no known Muklom, Hawoi or Phong speakers in Myanmar.

The term Pangwa has been used in India since about the 1970s to refer to sub-tribes who have mostly arrived in India quite recently. One defining feature of Pangwa is that they sing a particular cycle of ritual and historical songs, called the Wiwu song, or Sahwi song (Barkataki-Ruscheweyh and Morey 2013). There are approximately 28 sub-tribes of the Pangwa, on both sides of the India–Myanmar border, and some are not mutually intelligible.

Many of the Pangwa Tangsa people have told us that the languages are ‘paired’, with two very similar varieties being fully mutually intelligible and essentially identical. We have collected the list of these pairings from two different speakers in different language

communities in India (though living less than 1km apart). As we can see in Table 17.1, the two speakers' assessments are almost identical, with only Longkhi/Lungkhai, Shati/ Shangtih and Kochong/Khachal being differently assigned by the two speakers (three varieties most of whose speakers are in Myanmar).

At present, this list represents the best 'subgrouping' within the Pangwa groups that we are able to present. The pairings that we have been able to examine closely, such as Cholim and Longri, are also very similar in terms of vocabulary and morphology. Larger groupings within this list are not yet clear, although one group of four varieties that is a clear linguistic sub-grouping is that including Joglei and Sangwan (no. 6 in Lamsham's

TABLE 17.1 THE 'PAIRINGS' OF PANGWA SUB-TRIBES ACCORDING TO LAMSHAM KHILAK (CHAMCHANG SPEAKER OF KHALAK ANCESTRY) AND NIKAM TAIKHAM JOGLEI (JOGLEI SPEAKER)

Lamsham Khilak ¹	Phonetic (Shecyü) ²	Nikam Taikam Joglei ³	Phonetic	Autonym
1. Chamchang	t̪eam ² t̪eŋ ²	8. Kimsing	kimteŋj	Chamchang
Sheychii	ʃe ² t̪euŋ ²	Shangke	ʃakeŋ	Shecyü
2. Khilak	k ^h i?la?	11. Khalak	k ^h alak	Khalak
Longkhi	lɔŋ ¹ k ^h i ³	Khachal	k ^h ate ^h al	Lungkhi
3. Ronrii	ron ² r̪y ²	7. Ronrang	ronraŋ	Rera
Hiqjii	hi?dʒy ²	Hajang	ha?dʒaŋ	
4. Tiilam	t̪ur ² ləm ²	3. Tonglim	toŋlim	Cholim
Longrai	lɔŋ ¹ rai ²	Longri	loŋrəi	Longri
5. Miiri	mu ² ri ²	10. Mongrang	moŋraŋ	Mungre
Chahi	tea ² hi ²	Shangha	fajxa	
6. Ngimii	ŋj ² mur ²	2. Ngaimong	ŋaimonj	Ngaimong
Haqchung	hi?tsv ²	Hachang	ha?teŋj	Hacheng
7. Miisha	mu ² ʃa ²	14. Mossang	moŋʃaŋ	Mueshaungx
Longphai	lɔŋ ¹ p ^h ai ³	Lungphi	lunp ^h ai	Lungphi
8. Joqlai	jo ² lai ²	1. Joklai	joklai	Joglei
Shawey	ʃa ² ve ²	Sangwal	fajwal	Sangwan
9. Liichang	li ² t̪eŋ ²	9. Langching	laŋt̪eŋj	Lauchang
Diroq	di ¹ rs?	Darok	darok	
10. Dïngi	d̪uŋ ¹ ŋi ²	4. Dongngai	doŋŋai	Jiingi
Miti	mi ² ti ²	Moitai	moŋtai	Maitai
Shati	ʃa ² ti ²			
11. Riise	ri ² se ¹	13. Rangsa	raŋsa	Yvngban Wvng ⁴
Riisi	ri ² si ²	Rangseh	raŋsi	
12.		12. Lungkhai	lung ^h ai	
Shinchang	ʃin ² t̪eŋ ²	Sanching	santeŋj	
13. Tsangkhu	ðan ² k ^h u ²	6. Dingkhu	diŋk ^h u	Rinkhu
Laki	la?ki ²	Lakkih	lakki	Lakki
14. Kochong	kɔ ² t̪eŋ ²	5. Shangtih	fanti	Shangthi
Shorii	ʃɔ? ² r̪y ²	Shokrang	fokraŋ	

¹ As written in the Kimsing orthography by Ninshom Chena.

² List provided by Bynn Kham Lann, in an orthography with tone marking.

³ As written by Nikam Joglei.

⁴ In the orthography used by this sub-tribe, and in several others, <v> represents /ə/. The orthography, which is under development, shows tone marking by means of diacritics, with a dot above the final letter of a syllable marking a high level tone, as in Yvngban Wvng.

list and no. 1 in Nikam's) and Ngaimong and Hacheng (no 8. in Lamsham's and no. 2 in Nikam's).

Not all Pangwa varieties are present in India, but at least two, Joglei and Rera, are now found only in India.

In addition, within the Tangshang in Myanmar, there are a large number of other varieties not present in India at all, or only in small numbers, these can be grouped, tentatively, as follows (the ideas contained here are mostly from Nathan Statezni, personal communication):

- (2) Hakhun¹⁶
 - Bote / Haidley group (Bote, Haidley, Hakhi), similar to Hakhun
 - Ole group (Lumnu, Chamkok, Champang, Nahen and others)
 - Haqchum / Haqman group
 - Kotlum / Aasen / Raqnu group
 - Chuyo / Gaqkat (also present in India)
 - Kaisan group
 - Kon group (including Yongkon)
 - Sansik group
 - Pingku group

Hakhun is quite a large sub-tribe, present in Myanmar, but migrating in significant numbers to India in the last 20 or so years. Some of them are in the Tirap district of Arunachal Pradesh, where they are classified as Nocte, and their language is similar in many respects to other Nocte varieties (both, for example, have hierarchical agreement on verbs, unlike most other Tangsa/Tangshang varieties).

A group of languages that is similar to Hakhun, and perhaps largely mutually intelligible but also clearly distinct, is that represented by Bote (Bongtai) and Haidley [he?ple] (Halang), the latter of which is only found now in India.

Some speakers from the Ole group, particularly Champang speakers, have migrated to India in recent years and set up near Ledo in Assam. Their language is notable in having no verb agreement marking (pronominalisation) unlike the varieties listed in (1) or the Hakhun and Bote group varieties listed in (2). Some speakers in India suggested that Haqchum was very similar to the Ole group, but the Haqchum speaker we were able to interview in Myanmar did have a system of agreement similar to other Tangsa varieties (see section 4), so the latest thinking is that Haqchum belongs in a separate group.

The last five groups listed in (2) are all mostly spoken in Myanmar, and in so far as they have been investigated, none show verb agreement. Kaisan is apparently a large sub-tribe whereas Kon is much smaller, originally spoken in only two villages. Pingku is said to be a mixed language, having been developed when speakers of five Tangsa/Tangshang sub-tribes migrated to a new area, and finding that they could not communicate, created this new variety.

Most Tangsa speakers are multivarietal and multilingual; a typical speaker would know both parents' varieties (even when they are quite distinct), varieties which are dominant in the villages where they live, as well as any varieties that are local lingua francas (such as Shecyü in some places in Myanmar), the local language of wider communication (Assamese or Burmese) plus other languages used as lingua francas in different areas (Jinghpaw, Singpho, English, Hindi) (see Morey forthcoming) There is borrowing of words from Jinghpaw/Singpho and also from Tai languages via Jinghpaw/Singpho, an example of the latter being *na 'paddyfield', a word that has undergone regular sound changes and is exemplified in section 3.3. The presence of these sound changes,

following the same patterns as sound changes applying to words that are clearly not borrowed, such as ‘ear’ and ‘buffalo’, show that *na ‘paddyfield’ was borrowed before the sound changes occurred in these varieties.

Given the large number of varieties, we cannot here give linguistic examples of them all. To show the diversity of the varieties, however, we will present a short set of words in 11 Pangwa varieties¹⁷ in Table 17.2. In this and other tables presenting comparative data, the ordering of the different Pangwa varieties is based on the verbal morphology in section 4 (see further discussion later). Each of the words listed here has been checked for tone category with native speakers. In the case of Mueshaungx, Mungre, Yvngban Wvng and Shecyü, the lists were written by speakers in an orthography recently developed for the language.

The verbs listed in Table 17.2 are, in most cases, presented in two forms, firstly the verbal form, that combines with agreement markers, as /ka¹/ ‘go (downwards)’ in Mueshaungx, and secondly a nominalised form with a prefixal nominaliser /ə-/ or /i-/ (a nasal in Shangthi) as for example /əkai¹/ in Mueshaungx. This kind of stem alternation between the verbal and nominalised form is found to different degrees in different Tangsa/Tangshang varieties (see section 4). Thus the stem varies for the verb ‘go (downwards)’ in Mueshaungx, Ngaimong, Mungre, Yvngban Wvng, Shangthi and Shecyü, but not in Cholim or Lauchang.

2 WRITING

The history of writing in Tangsa can be broadly divided into six stages:

- i Folk stories tell of an ancient form of writing done on the skins of animals that was lost when it was eaten by the ancestors (these stories are widespread in the region).
- ii The writing down of word lists and sometimes texts by scholars, starting in the nineteenth century. The oldest records of this kind are the word lists of Mueshaungx and Shecyü in Needham (1897).
- iii Roman-based orthographies designed for use as ‘common language’, particularly for Bible translation and hymn books and in the medium term for literacy development.
- iv Orthographies designed for single sub-tribes, for example to preserve cultural information or for other purely local uses.
- v More informal Roman-based orthographies.
- vi Non-Roman-based scripts.

Space will not permit a detailed discussion of orthographic standards here, beyond presenting information about two of the orthographies in use on both sides of the border for Bible translation and for community use, namely the Chamchang orthography developed by Rev. Yanger Thungwa and adopted for use by the Tangsa Baptist Churches Association in India, and the orthography developed in Myanmar by Rev. Gam Win for the Mueshaungx language.

The Chamchang orthography has undergone a number of versions, and has recently been further revised for use in Myanmar, where its main promoters have been Mr Tseing Mong and Mr Bynn Kham Lann, who are both speakers of the closely related Shecyü variety. As at August 2015, these revisions are in the process of being adopted by the people in India. Table 17.3 presents some of the key differences between the three orthographies.

The aspirated palatal affricate /t^h/ is regularly written <chh> in India, and the unaspirated equivalent /t^h/ as <ch>. The writing of <chh> is not well accepted in Myanmar, and

TABLE 17.2 COMPARATIVE WORD LIST FOR 11 PANGWA VARIETIES

	Ngaimong	Mueshaungx	Mungre	Jiungi	Cholim	Sheeyü	Lauchang	Yngban	Wvng	Shangthi	Gaqlun	Rinkhu
nose	k ^b oŋ ¹	k ^b uu ¹		fun ¹	k ^b uu ¹	k ^b ju ¹	k ^b jan ¹	n ^a ko ¹	k ^b y po ¹	n ^a ŋku ²	k ^b u kum ¹	
skin	k ^b oŋ ¹	k ^b uu ²		fun ²	k ^b uu ²	k ^b ju ²	k ^b jan ²	k ^b on ²	k ^b raŋ ²	k ^b raŋ ²	ik ^b on ²	
I	ŋa ¹	ŋa ¹	ŋa ¹	ŋei ¹	ŋa ¹	ŋi ¹	ŋi ¹	ŋa ¹	ŋe ¹	ŋa ¹	ŋa ¹	
you (singular)	m ³ ŋɛŋ ⁴	m ³ naŋ ¹		mi ¹	mna ¹	mo ²	m'na?	mu ²	ŋo ¹	ŋuŋ ¹	ŋuŋ ¹	
father	w ^a , ēwa ⁴	əβa ¹		vei ¹	βe ¹	ve ¹	vi ¹	iwa ¹	əve ¹	əva ²	əva ¹	
liquor	teoŋ ¹	teoŋ ¹		teoŋor ¹	lu ³ hu ³	te ^b ai ¹	te ^b au ¹	lawu ³	te ^b ɔŋ ¹	la <u>2</u>	te ^b ɔŋ ¹	
water	k ^b əm ²								lu ³ hu ²			
monkey	vui ²	βin ² sull ²										
elephant	teeŋ ¹											
earth	ha?			hei?	be?ze ¹	te ^b al ¹	te ^b eu ¹	bo? ² la ²	jian ¹	jukvi ²	jukvi ¹	
fire	val ¹			bar ¹	βai ¹	ge?	hi?	boka ²	juŋ ¹	te ^a !	te ^a !	
sun	raŋ ² sal ²			var ¹	raŋ ² far ²	raŋ ² xai ²	raŋ ² e ²	ga?	ge?	ga?	ga?	
one	əfi ⁴			əfi ¹	əfi ¹	əfi ¹	əfi ¹	wa?	vail	vail	vail	
seven	niŋ ¹ ji ²			meŋ ² ai ²	meŋ ² ai ²	meŋ ² si ²	meŋ ² si ²	masau ²	raŋ ² lik	raŋ ² mik	raŋ ² mek	
sleep	kun ¹	əkun ³		kaiso ² , əkatsol ²	kun ³ , ɻep	maʃi ²	maʃi ²	tanik	wa?	wa?	wa?	
fall	dai ¹ , adai ⁴	dai ¹ , adai ¹	/ kur ² , əku ³	dai ¹ , adai ¹	kio ² , i-kiŋ ²	de ¹ / kjan ² , i-kiŋ ²	de ¹ / kjan ² , i-kiŋ ²	ten ³	jip, nijip	ten ³	jip, nijip	
go (downwards)	ka ¹ , əkal ⁴	ka ¹ , əkal ⁴		kei ¹ , əkei ¹	ke ¹ , əke ¹	ki ¹ , i-ki ¹	ki ¹ , i-ki ¹	qat, ngat	lat ja ³	lat ja ³	idit	
fear / frighten	hil ¹ , ahil ⁴	hil ¹ , ahil ⁴		hi ¹ , ahi ¹	hi ²	ha ²	ha ²	ka ¹ , i-ka ¹	ka ¹	ka ¹	ka ¹	
hear	tai ¹ , atai ⁴	tai ¹ , atai ⁴		tei ¹ , atei ¹	tei ¹ , atei ¹	ti ¹ , i-ti ¹	ti ¹ , i-ti ¹	phap ma ³	phap ma ³	phap ma ³	phap ip ^b ap	
								tit na ³	tit na ³	tit na ³	tit na ³	

TABLE 17.3 COMPARISON OF MUESHAUNGX AND CHAMCHANG/SHECYÜ ORTHOGRAPHIES

Phonemic symbol	Mueshaungx (Gam Win)	Chamchang (Yanger Thungwa)	Shecyü (Tseing Mong / Bynn Kham Lann)
t ^c	j	ch	cy
t ^c ^h	ch	chh	ch
j / ʒ	y	j	y
w/β	w	w	v
ə	v	ii (a)	ä
a	a	a	a
low tone	z	x	z
high (falling) tone	x	f	x
high (rising) tone	c	(not found)	(not found)
mid level tone	(not found)	unmarked	unmarked

TABLE 17.4 LAKHUM MOSSANG'S ORTHOGRAPHY

Lakhum's system	Gam Win's system	Tone quality
ø	oz	Low falling
ɔ	oc	High (rising)
ə	oq	Glottal final
ɸ	ox	High falling

so <ch> is used for the aspirated affricate and other symbols, such as <cy>, for the unaspirated. The use of <j> in the Mueshuangx orthography for the unaspirated palatal affricate is then contrasted with the use of the same symbol in Chamchang for the palatal semi-vowel (also realised as voiced fricative).

In the published works using the Chamchang orthography in India (for the Bible and Hymn Book), tone marking is generally not used, although in our online Chamchang dictionary,¹⁸ the Yanger system of tone marking is employed. In Mueshuangx publications, tone marking is generally only found on the last syllable of the word, hence in the first syllable of *wukuiyq* ‘parrot’, the low tone (-z) is not marked, whereas it is on the monosyllabic root *wuz* ‘bird’. In the online Mueshaungx dictionary, we have marked tone on every syllable.¹⁹

Brief mention will also be made here of the script developed by Lakhum Mossang since 1990.²⁰ This has 78 symbols of which four are shown in Table 17.4. Tone marking is incorporated into the symbols for the vowels, so there is no single symbol for /o/, but four different unrelated symbols for /o/ under each of the four tones of Mueshaungx. One by-product of this writing system is that it is impossible not to mark tone contrasts when employing it.

3 PHONOLOGY

3.1 Consonants

Table 17.5 presents the consonants of the Mueshaungx (Mossang) variety, with the orthography developed by Rev. Gam Win followed by phonetic symbols in slashes. The

TABLE 17.5 CONSONANT PHONEMES IN MUESHAUNGX

	Bilabial	Dental	Alveolar	Palatal	Velar	Glottal
Voiceless unaspirated stops	p /p/	th /t/	t /t/	j /te/	k /k/	q /ʔ/
Voiceless aspirated stops	ph /pʰ/	htt /tʰ/	ht /tʰ/	ch /teʰ/	kh /kʰ/	
Voiced stops	b /b/		d /d/			g /g/
Nasals	m /m/		n /n/	ny /ɲ/	ng /ŋ/	
Voiceless fricative			s /s/	sh /ʃ/		h /h/
Voiced fricative					(gh /ɣ/)	
Affricate				ts /ts/		
Semivowel	w /β/				j /ʒ/	
Rhotic approximant				r /ɹ/		
Lateral approximant				l /l/		

voiced velar fricative *gh* /ɣ/ is shown in brackets as it is found as a contrastive sound in the speech only of some individuals. A dental nasal *nh* /ɲ/ is found in the speech of some speakers when followed by /ui/, but it is not contrastive. Note that the realisation of the semivowels varies between semivowel and fricative, [w]~[β] and [j]~[ʒ]. The fricative realisations are more frequent, hence they are shown as the phonemes in Table 17.5.

The 27 consonant phonemes of Mueshaungx can be compared with 24 in Chamchang, where there is no dental series (for most speakers), and no voiced velar fricative. Other Tangsa/Tangshang varieties have a range of different phonemic contrasts, of which we may note that in Shangthi and Rera, there is no contrast between aspirated and unaspirated stops, and in some varieties there are other sounds not found here, such as the voiceless velar fricative [χ] in Cholim and Joglei and the voiced dental fricative [ð] in Ngaimong and Shecyü.

3.2 Vowels

A comparison of the vowel phonemes for Mueshaungx and Chamchang is presented in Table 17.6.

The diphthongs also differ between varieties. Mueshaungx is particularly rich in diphthongs, as we see in (3). Only some of these can occur with codas.

(3.1) Diphthongs that are not found with codas

uey /xi/ (possible variation of /ui/, written by Gam Win in /nxi³/ ‘two’, written /nui³/ by other speakers)

ueu /xu/

oay /ɔəi/

ay /ai/

uy /ui/

oy /oi/ (only recorded in words in the traditional Wiwu song)

vy /əi/ (only recorded in words in the traditional Wiwu song)

vu /əu/ (only recorded in words in the traditional Wiwu song)

(3.2) Diphthongs that are found with codas

uiy /ui/ occurs with final glottal stop

uiu /uu/ occurs with final -?, -k

oa /ɔə/ occurs with final -t, -m

au /au/ occurs with velar finals -k, -ŋ

TABLE 17.6 VOWEL PHONEMES IN MUESHAUNGX AND CHAMCHANG

Mueshaungx			Chamchang		
i /i/ e /e/ v /ə/ a /a/	ui /u/ ue /y/ aw /ɔ/	u /u/ o /o/ aw /ɔ/	i /i/ e /e/ ä /ə/ a /a/	ü /u/ ii /y/ u /u/ o /o/	

TABLE 17.7 APPROXIMATE PITCH AND PHONATION VALUES FOR THE TANGSA TONES

	Tone 1 pitch	phonation	Tone 2 pitch	phonation	Tone 3 pitch	phonation
Ngaimong	/31/126-89	glottal constriction	/344/122-136-137	plain	/11/99-92	plain
Mueshaungx	/21/116-90	plain	/52/152-92	plain	/35/128-161	plain
Mungre	/41/134-97	plain	/44/132-126	plain	/21/116-92	plain
Jiingi	/33/162-159		/21/139-129		/45/175-181	
Cholim	/442/ 198-194	glottal constriction	/32/, (/33/) 180-114	plain	/52/219-139	plain
Chamchang	/21/150-130	plain	/44/209-224	plain	/53/255-208	plain
Shecyü	/42/185-140	slight glottal constriction	/33/174-167	plain	/53/204-217-169	plain
Lauxchangx	/332/ 175-158	glottal constriction	/51/218-102	plain	/44/186-191	plain
Yvngban Wvng	/31/144-120		/33/ ~ /55/ 154-160, 180-189		/14/122-175	
Shangthi	/352/ 130-187	glottal constriction	/22/122-123		/31/140-113	
Gaqlun	/33/131-132		/55/146-150		/31/133-108	
Rinkhu	/552/ 175-173	glottal constriction	/33/149-145	plain	/21/136-122	plain

Chamchang, on the other hand, has only three diphthongs, as in (4). Of these only /ai/ can occur with a coda, and that only with final glottal stop -ʔ, e.g. /maiʔ/ ‘person’.

- (4) ai /ai/
 au /au/
 ea /ɛə/

3.3 Tones

The first to describe Tangsa tones was Weidert (1979, 1987). His analysis of the Joglei variety²¹ concurs with that of the Mueshaungx and Chamchang/Shecyü orthographies in positing three tones on open syllables and a further category of stop finals (-p, -t, -k, -ʔ) in which tone is not distinguished. A comparison of the tone values for the three open tones in the 11 varieties in Table 17.2 and also Chamchang, is presented as Table 17.7. For each tone, the value is given in two ways, in phonemic slashes using tone numerals, followed by an actual pitch value for F_0 in Hz.

Table 17.7 is based on the fact that most words in at least the Pangwa varieties carry the same tone category in all the varieties. Table 17.8 presents a set of correspondences

TABLE 17.8 EXAMPLES OF UNMARKED TONE CORRESPONDENCES BETWEEN EIGHT TANGSA VARIETIES

English	Ngaimong	Mueshaungx	Mungre	Cholim	Chamchang	Sheeyü	Lauchang	Rinkhu	Tikhak
Tone 1									
chicken	v <u>1</u>	-z wu ¹	-x mau ¹	β <u>1</u> maŋ ¹	-x m <u>1</u>	vu ¹ ma ¹	yau ¹ ma ¹	v <u>1</u> maŋ ¹	β <u>1</u> maŋ ¹
dream	maŋ ¹				βe ¹	ve ¹	βi ¹	va ¹	βa ¹
father	v <u>1</u>	ə-w <u>1</u>	v <u>1</u>	βai ¹	v <u>1</u>	va ¹	v in ¹	v in ¹	βu ¹
fire	v <u>1</u>	w <u>1</u>							
Tone 2									
bear	tɛəpbeŋ ²	teəpbø ²	teəp ² ba ²	-Ø	-Ø	teəpbi ²	tɛ ² eʔbo ²	tɛ ² apba ²	tɛapbaatj ²
door	ka ³ lɔŋ ²	ka ³ lɔŋ ²	kei ³ lunj ²	kei ³ lunj ²	ki ³ lur ²	ki ³ lur ²	ka ³ lu ²	ka ³ lu ²	ka ³ lu ²
paddyfield	na ²	na ²	ne ²	ne ²	ni ²	ni ²	na ²	na ²	na ²
younger	no ²	ano ²	ə-no ²	ə-no ²	anɔ ²	anɔ ²	no ²	no ²	no ²
brother	nau ²								
Tone 3									
ear	na ³	-c na ³	-z ŋa ³	-f ne ³	-f ni ³	-f ŋi ³	-f ŋi ³	-f ni ³	-f na ³
buffalo	ŋa ³			je ³	ŋe ³	ŋe ³	ŋa ³	ŋa ³	ŋa ³
four	balai ³	balai ³	balai ³	bal ³	balai ³	balai ³	balai ³	balai ³	(lo ²)
snake	peu ³	peu ³	pao ³	pu ³	pau ³	pau ³	pau ³	pau ³	pu ³

TABLE 17.9 EXAMPLES OF MARKED TONE CORRESPONDENCES BETWEEN EIGHT TANGSA VARIETIES

English	Ngaimong	Mueshaungx	Mungre	Cholim	Chamchang	Shecyü	Lauchang	Rinkhu	Tikhak
	1	1 (-z)	1 (x)	?	2 (Ø)	2 (Ø)	1	stop	stop
blow	mul ¹	əmui ¹	moi ¹	mo?	mer ²	me ²	əmau ¹	(p ^b wut)	əmyt
fall	dəl ¹	dəl ¹	dai ¹	dj?	də ²	dia ²	əde ¹	dit	ək ^b et
ill	ða ¹	ə-ɿuu ¹	atsaa: ¹	ə-de?	tsi ²	adzi ²	ədi ¹	irak	ərak
cloth	kʰəl ¹	kʰə-oi ¹	khai ¹	kʰj?	kʰe ²	khia ²	khe ¹	kʰet	kʰat
trample	na ¹	ənuu ¹	na ¹	ne?	nj ² / ni ²	ni ²			nak

TABLE 17.10 FREE PRONOUNS AND POSSESSIVE PREFIXES IN MUESHAUNGX AND CHAMCHANG

x	Mueshaungx		Chamchang	
x	Free pronouns	Possessive prefix	Free pronouns	Possessive prefix
1st person singular	<i>ngaz</i> /ŋa ¹ /	<i>ix-</i> /i ² -/	<i>ngiz</i> /ŋi ¹ /	<i>i-</i> /i ¹ -/
2nd person singular	<i>mznaungz</i> /m ¹ naun ¹ /	<i>mz-</i> /m ² -/	<i>ümnaq</i> /m ² na?/	<i>üm-</i> /m ² -/
3rd person singular	<i>vpiq</i> /əpi? ¹ /	<i>v-</i> /ə-/	<i>äpaiq</i> /a ¹ pai? ¹ /	<i>ä-</i> /ə-/

for eight varieties. For those varieties which are being written, the orthographical marking of tones is given in the top row of each set. We have also included the non-Pangwa Tikhak variety in this table.

There are however a number of words that do not have the same tones in all varieties. This creates a correspondence pattern such as that seen in Table 17.9. This correspondence set includes words that consistently have tone 1 in Mueshaungx, Mungre, Lauchangx and Ngaimong, tone 2 in Chamchang and Shecyü, a final glottal stop in Cholim and final stops (/t/ or /k/) in Rinkhu and Tikhak.

In Ngaimong, when a verb stem carrying tone 1, or a stop final, is prefixed by the nominaliser, the tone becomes high falling, something like /52/. This in effect creates a new tonal distinction in Ngaimong, with a fourth tone on open syllables, and also creates a tone distinction on stopped syllables. Thus in Table 17.2, the nominalised forms of the verbs ‘fall’, ‘go downwards’, ‘fear’ and ‘hear’ all carry tone 4. The result of these factors means that the correspondence of tone categories between different varieties is not straightforward.

4 MORPHOLOGY

Tangsa languages are largely verb final and isolating, but there are several parts of the grammar that are, or could be, considered examples of agglutination: verbal and nominal prefixing and agreement markers. Nominal prefixes are possessor prefixes, such as those listed for Mueshaungx and Chamchang in Table 17.10, shown together with the free pronouns.

Verbal prefixes can include the following:

- | | | | | |
|--------------------|------------|------------------|-----------|------------------|
| (5) Nominaliser | Mueshaungx | <i>v-</i> /ə-/ | Chamchang | <i>ä-</i> /ə-/ |
| Causative | Mueshaungx | <i>tv-</i> /tə-/ | Chamchang | <i>tä-</i> /tə-/ |
| Reciprocal | Mueshaungx | <i>rv-</i> /rə-/ | Chamchang | <i>rä-</i> /tə-/ |
| Reflexive | Mueshaungx | <i>kv-</i> /kə-/ | Chamchang | <i>rä-</i> /tə-/ |

In some varieties, such as Yvngban Wvng, Cholim, Shangthi, Longri and Khalak, there is another prefix marking attributes (adjectives), the form of which differs from the prefix marking nominalisation. Thus in Shangthi, the nominaliser for active verbs is a nasal, homorganically merged to the place of articulation of the initial consonant of the verb stem, as we see in Table 17.2. With many adjectives, however, the prefix is *ə-*, as *ə-kʰip* ‘bitter’, *ə-səm¹* ‘sweet’. In Chamchang and Mueshaungx, the nominalising and attributive prefix have the same form, *ə-*.

In most varieties so far investigated, the reciprocal and reflexive are marked by a single form, as in Chamchang and also Cholim (see Morey 2011).

In most varieties, for at least some verbs, there is stem alternation between the verbal form and the nominalised form. Some examples of stem alternation are presented in Table 17.10, where the alternate forms are shown in boldface. We have already mentioned that in Ngaimong a fourth tone is created by this alternation, because when prefixed, a stem carrying the low falling tone (tone 1), or stop final, regularly becomes high falling (tone 4). For this reason most Ngaimong verb stems show stem alternation, but it is predictable in these cases. The other examples of stem alternation are not predictable, and it is noticeable that the only varieties in Table 17.10 that have exactly the same patterns of alternation are Chamchang and Shecyü, which are paired, as we saw in Table 17.1.

There is some evidence for the backformation of new verbal stems. Consider the verbs for ‘give’. In most varieties the verbal form has a final glottal stop, as *ku?* in Chamchang and *kʰy?* in Shangthi. In several varieties, the nominalised form has an alternate stem with a final open syllable. In Lauchang, the original verbal stem appears to have been replaced by the nominalised stem re-analysed.

All Pangwa varieties have verb agreement, also termed pronominalisation or indexing, where a particle or suffix marking tense/aspect/modality together with person and number is added to the verb stem. In general it is the verbal form from Table 17.11 that takes this marker. Agreement is a feature of all Tikhak varieties, Muklom and Hawoi, Phong, Hakhun and the Bote/Haidley group. The other varieties listed in (2), in so far as they have been investigated, mostly do not mark person or number agreement on the verb.

There is considerable variation in both the forms and functions of these markers. In some varieties (like Cholim) present time is expressed by a continuous marker (/l-/ combined with agreement particles, whereas in others, like Mueshaungx, the same function is expressed by an invariant particle /jaʔ/. Nevertheless, across many Tangsa varieties, including Pangwa varieties so far investigated, negation, past time and future time/irrealis are marked with agreement. Table 17.11 lists the forms for a number of varieties, with a bold line dividing the Pangwa from non-Pangwa varieties

In a number of varieties, such as Mungre, Ngaimong, Joglei and Mueshaungx, both the past and negative forms have similar structures, the negative has initial *m-* and the past initial *t-*, with the endings being (in the order 1sg, 1pl, 2sg, 2pl, 3), *-ak*, *-i*, *-u*, *-it*, *-a*. Still other varieties like Muklom and Tikhak, which are non-Pangwa, have the same initials but the corresponding non-stop finals for the person endings, for example *-ay*, *-i*, *-u*, *-in*, *-a*. Future forms, on the other hand, always have open endings, and we suggest that historically the future forms were open syllables and the non-future were stopped syllables. In Table 17.12, those varieties which have stop finals are listed first (Ngaimong, Joglei, Mueshaungx, Mungre and in some circumstances Maitai), then a group of varieties that have open finals throughout but with post verbal negation (Cholim, Longri, Chamchang, Shecyü, Lauchang, Lungkhi and Khalak) and a third group with open finals but pre-verbal negation (Yvngban Wvng, Shangthi, Gaqlun, Rinkhu, Rera). This is a tentative grouping for the Pangwa varieties, employed in all the tables in this chapter.

TABLE 17.11 STEM ALTERNATION IN EIGHT TANGSA VARIETIES

	Ngaimong	Mueshaungx	Chamchang	Shecyü	Lauchang	Yngban Wng	Shangthi	Rinkhu
	verbal nominal	verbal nominal	verbal nominal	verbal nominal	verbal nominal	verbal nominal	verbal nominal	verbal nominal
(1)	na ⁱ² tal ¹ run / flee fall	a <i>ni</i> ⁴ ta <i>l</i> ⁴ 3 <i>gu</i> ⁴ da <i>l</i> ¹	na <i>i</i> ¹ ta <i>i</i> ¹ ju ¹ de <i>i</i> ¹	ana <i>i</i> ² ata <i>i</i> ² a <i>gu</i> ¹ ad <i>a</i> ¹	na <i>i</i> ² ta <i>i</i> ² ju ¹ dia ²	na <i>i</i> ² ta <i>i</i> ² ju ¹ dia ²	i-ja <i>i</i> ¹ i-ti ¹ i-za <i>u</i> ¹ ja <i>o</i> ³	ni ² tat ¹ n <i>jwa</i> ² ndat ¹
(2)	roast drink give eat (rice)	fu ¹ mu <i>n</i> ² ko ² sa ² ¹	fe <i>e</i> ⁴ li <i>ne</i> ³ ek <i>u</i> ⁴ se <i>a</i> ² ⁴ / <i>asal</i> ⁴	fu ¹ nu <i>n</i> ² ku ² sa ² ¹	fe <i>e</i> ² a <i>ne</i> ³ ak <i>u</i> ² se <i>a</i> ²	fe <i>e</i> ² a <i>ne</i> ² ku <i>?</i> ² si ²	ja <i>u</i> ¹ ja <i>n</i> ² i-ka <i>u</i> ¹ sa ² ¹	ni <i>u</i> ² i-nan ² ko ² sa ² ¹
(3)	go see wash clothes	ka ¹ k <i>bu</i> ¹ ha <i>l</i> ¹	ek <i>u</i> ³ k <i>bu</i> ³ he ¹	te <i>k</i> ¹ te <i>k</i> ¹ he ³	te <i>k</i> ¹ at <i>eh</i> ³ he ¹	te <i>k</i> ¹ i-hau ² hau ¹	ko ¹ i-hau ² hut	ikot ¹ i-hau ³ hat ¹
(4)	sit look	le <i>l</i> ³ le <i>l</i> ³	ja <i>u</i> ³ le <i>l</i> ³	ja <i>u</i> ³ la <i>u</i> ³	ja <i>u</i> ³ la <i>u</i> ³	i-ja <i>u</i> ² i-lau ²	te <i>o</i> ² ila ³	ja <i>u</i> ³ la <i>u</i> ³
(5)	stand steal bite	te <i>ep</i> ¹ te <i>ep</i> ⁴ ek <i>ak</i> ⁴	te <i>ep</i> ¹ gu <i>u</i> ² ka <i>u</i> ¹	te <i>ep</i> ¹ huk ¹ ka ² ¹	te <i>ep</i> ¹ huk ¹ ka ² ¹	i-te <i>he</i> ² gut ¹ i-ka ²	sep ¹ gut ² gak ¹	te <i>hap</i> ¹ gu ² ikak

TABLE 17.12 TANGSA AGREEMENT MARKERS IN VARIOUS TANGSA VARIETIES

Negative		Past										Future										
1 Sg	IPl	2 Sg	2 Pl	3	1 Sg	IPl	2 Sg	2 Pl	3	1 Sg	IPl	2 Sg	2 Pl	3								
Ngaimong	V muk	V mi?	V mo?	V mit	V mo?	V tuk	V ti?	V to?	V tit	V ta?	V aŋ?	V tu?	V au?	V aŋ?	ə-V aŋ?	ə-V aŋ?	ə-V aŋ?	ə-V aŋ?	ə-V aŋ?	ə-V aŋ?	ə-V aŋ?	
Joglei	V mak	V mai?	V mu?	V mit	V mu?	V tank	V tai?	V tu?	V tit	V ta?	V aŋ?	V tu?	V au?	V aŋ?	ə-V i							
Mueshaungx	V mauk	V mi?	V mu?	V mut	V mu?	V tank	V ti?	V to?	V tit	V ta?	V aŋ?	V tu?	V au?	V aŋ?	V ſauŋ?							
(1) transitive																						
(2) intransitive as (1)																						
Mungre	V (i?) mok	V (i?)	V (i?) mat	V (i?)	V (i?) mok	V (i?) mat	V (i?)	V (i?)	V (i?)	V (i?)	V (i?)	V (i?)	V (i?)	V (i?)	V (i?)	V (i?)	V (i?)					
Maitai	V mu?	V mu?	V mu?	V mi?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?
Cholim	V man?	V mi?	V mu?	V mi?	V mi?	V min?	V mi?	V mu?	V mok	V kyo?	V kyo?	V ku?	V kin?									
Longri	V man	V mi	V mu	V mi	V mi	V min	V mi	V mu	V ko	V ki	V ku	V lau?	V lan?									
Chamchang	V maj?	V mai?	V mau?	V mai?	V mai?	V man?	V mai?	V mak	V kai?	V kai?	V kai?	V lau?	V lan?									
Shecyü	V man?	V mal?	V mau?	V man?	V man?	V man?	V man?	V mak	V kəŋ?	V kai?	V lau?	V lan?										
Lauchang	V mau?	V mai?	V mau?	V man?	V man?	V man?	V man?	V mok	V keŋ?	V kai?	V lau?	V lan?										
Lungkhi	V bi?	V bi?	V bu?	V bi?	V bi?	V bin?	V bi?	V ba?	V dəŋ?	V dəŋ?	V du?	V din?										
Khalak (1)	V ban	V bi	Vbau	V bi	V ba	V bin	V ba	V kan	V kai	V kai	V kau	V kin										
Khalak (2)	n V	n V	n V	n V	n V	n V	n V	mi?	V u?	V aŋ?	V tu?	V təŋ?										
Yngvban	mi?	V aŋ?	mi?	V aŋ?	mi?	V aŋ?	mi?	V aŋ?	mi?	V aŋ?	mi?	V aŋ?	mi?	V aŋ?	mi?	V aŋ?	mi?	V aŋ?	mi?	V aŋ?	mi?	V aŋ?
Wng (1)	i?	i?	i?	i?	i?	i?	i?	i?	i?	i?	i?	i?	i?	i?	i?	i?	i?	i?	i?	i?	i?	i?
(2) intransitive as (1)																						
Shangthi	mi?	V aŋ?	mi?	V u?	mi?	V u?	mi?	V u?	mi?	V aŋ?	V kai?	V ku?										
Gaqlun	i?	V aŋ?	i?	V u?	i?	V aŋ?	i?	V u?	i?	V aŋ?	V tə aŋ?	V tə aŋ?	V tə aŋ?	V tə aŋ?	V tə aŋ?	V tə aŋ?	V tə aŋ?	V tə aŋ?	V tə aŋ?	V tə aŋ?	V tə aŋ?	V tə aŋ?
Rinkhu	mi?	V aŋ?	mi?	V u?	mi?	V u?	mi?	V u?	mi?	V aŋ?	V kai?	V ku?										
Rera	mi	V laj	mi	V laj	mi	V laj	mi	V laj	mi	V laj	mi	V laj	mi	V laj	mi	V laj	mi	V laj	mi	V laj	mi	V laj
Haechum	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V mu?	V tak	V ti?	V tit										
Hakhun	V mi?	V mi?	V mi?	V mi?	V mi?	V mi?	V mi?	V mi?	V mi?	V tə?	V tə?	V tə?	V tə?	V tə?	V tə?	V tə?	V tə?	V tə?	V tə?	V tə?	V tə?	V tə?
Yongkuk	V man	V man	V man	V man	V man	V man	V man	V man	V man	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?
Tikhak	V maj?	V maj?	V maj?	V maj?	V maj?	V maj?	V maj?	V maj?	V maj?	V mo?	V mo?	V mo?	V mo?	V mo?	V mo?	V mo?	V mo?	V mo?	V mo?	V mo?	V mo?	V mo?
Muklom	V mon	V mon	V mon	V mon	V mon	V mon	V mon	V mon	V mon	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?
Phong	V mun	V mun	V mun	V mun	V mun	V mun	V mun	V mun	V mun	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?	V təŋ?

Note: * = not recorded.

Some varieties, Cholim, Longri, Chamchang, Shecyü and Lauchang, have altered initial forms for the past, in non third person. These varieties also all have a preverbal future marker *mV-*. In Cholim and Longri, the non third person forms all have initial *k-*, while in Chamchang, Shecyü and Lauchang second person forms now have initial *l-*. In Cholim and Longri, an apparently more recent change of **aŋ* > /*jo*/~/*o*/ means that the first person singular past tense marker in Cholim is now realised as *kjo*³, rather than a putative proto **tak*. Finally, Cholim and Longri are differentiated by what appears to be an even more recent change in Cholim, final /-n/ > /-ŋ/ in the second person plural.

The future forms in Table 17.11 are complex and varied, but one interesting feature is the distinction in form according to whether the verb is transitive or intransitive found only in Mueshaungx. This may be a recent innovation, because the sentences translated with future tense by Needham (1897: 10), show an invariant form <*khâung*> in the future, as shown in (6)

- (6) ngâ kâ khâung ‘I will go’
 ngâ shong khâung ‘I will sell’

The Mueshaungx speakers that we have worked with confirmed that they would say these sentences as respectively *ya' ka' tauy*² ‘I will go’ and *ya' sauy' sauy*² ‘I will sell’. Some older speakers remember that there was a form /*k^baŋ*/ in the language.

We have done very little work on the Khalak variety, meeting two speakers briefly, one in India and the other in Myanmar. We have included the data about Khalak in this table because two different systems for the negative were recorded. In elicitation, the Indian speaker gave forms with an invariant nasal *n-* followed by the verb, whereas the speaker in Myanmar gave a form which has an agreement marker with initial *b-*. The nasal negative prefix *n-* is also found in the Wihu song, and may represent a survival of the proto nasal form, now hardly used in the spoken languages at all.

NOTES

- 1 Research on Tangsa languages in India, and more recently in Myanmar, has been possible with funds provided first by the Dokumentation der Bedrohten Sprachen (DoBeS) project of the Volkswagen Stiftung, and more recently by an Australian Research Council Future Fellowship. Both of these fellowships were taken up at La Trobe University. I thank my many Tangsa consultants in both India and Myanmar, in particular Lukam Tonglum (Cholim), Rev. Gam Win, Wanglung Mossang and Renman Keluim (Mueshaungx), Ninshom Chena and Rev. Longkhan Yanger Thungwa (Chamchang), Shinyung Ngaimong and Wangkui Ngaimong (Ngaimong), Khithong Hakhun and Phulim Hakhun (Hakhun), Satum Ronrang (Rera), Nongtang Langching, Daniel Mawyio (Lauchang), Shu Maung (Mungre) and Bynn Kham Lann (Shecyü); there have been many others. I am also very grateful to students from Gauhati University Linguistics Department, who have helped with language research, Palash Kumar Nath, Iftiqar Rahman, Niharika Dutta, Asifa Begum, Deepjyoti Goswami and Poppy Gogoi. Others involved in the research have included Jürgen Schöpf and Meenaxi Barkataki-Ruscheweyh. I also thank Nathan Statezni for sharing his data and many useful suggestions.
- 2 This term is used by the communities in India. Within Tangsa there may be several clans within each sub-tribe, and some of those clans are found in more than one sub-tribe. Marriage to a person from the same clan is not permitted, whereas marriage to a person of the same sub-tribe is. As an example, Kalim or Keluim is the name of a single clan in both Maitai and Mossang (and probably others), which are both sub-tribes with different linguistic varieties.

- 3 Migration histories and song cycles of Tangsa people tell of migrations across Myanmar, reaching the Tanai river and dividing and then moving up into the Patkai range. We do not know for how many generations these communities have been in the Patkai ranges.
- 4 Of the groups we now call Tangsa, whom he almost certainly never visited, he says that, after the Namsang Nocte: ‘there are a few broken tribes to the east of them; these are of little note, and are in subjection to the Singpho’. We interpret ‘broken tribes’ as a reference to the astonishing diversity of Tangsa/Tangshang.
- 5 See section 1.1 for a discussion of the conventions in naming the sub-tribes. Mueshaungx is the name of this group using the orthography developed by Rev. Gam Win (see section 2) where the final -x marks tone on the last syllable.
- 6 The Tangsa are listed as Other Nagas in the list of Scheduled Tribes under Articles 342 (i) and 342 (ii) of the Indian Constitution. For further information see Barkataki-Ruscheweyh (2015: 11).
- 7 This is the common Sino-Tibetan root for ‘child’, given as *za or *tsa in Matisoff (2003).
- 8 A modernised version of this map appears in Barkataki-Ruscheweyh (2015: 36).
- 9 Also pronounced Hewa, Hewe, Hiwi depending on the variety. Some speakers have told us that this name refers to all Nagas, encompassing those who speak languages outside ‘Northern Naga’.
- 10 Bradley (1997) distinguished Heimi and Tangsa/Rangpan, grouping the former with Nocte and the latter with Wancho. This classification is not supported by our more recent research.
- 11 These terms have not been heard frequently in our research in India, but Wanglung Mossang (who had not heard the terms Shangnyu Wang and Tangnyu Wang), told us that *shaung* means ‘our own people’ and *taung* means ‘other Naga people’, who are generally living to the west.
- 12 In the Mueshaungx variety, the sub-tribe name is now conventionally written with tone marking -x on the second syllable. We will follow that here in group naming, but for all other groups we will not show tone marking for the group names in orthography.
- 13 See: en.wikipedia.org/wiki/Tangsa_people (accessed 14 August 2016).
- 14 In collaboration with SIL International. These word lists, of around 450 items, are recorded with English and Burmese terms. There are plans to make these available online at sealang.org.
- 15 A PhD on Tikhak is in preparation by Karen Parker at La Trobe University.
- 16 A PhD on Hakhun is being prepared by Krishna Boro at the University of Oregon.
- 17 We have included a list for Shecyü rather than Chamchang; both varieties are almost identical with three regular sound correspondences, Shecyü -ia, -ɔ and ð- corresponding to Chamchang -ɛɔ, -o and ts-.
- 18 Based on a word list prepared by Rev. Yanger Thungwa, revised and recorded by Ninshom Chena, online at sealang.net/chamchang (accessed 14 August 2016).
- 19 Based on a word list prepared by Wanglung Mossang, Renman Mossang and Khumseng Mossang, online at sealang.net/mueshaungx (accessed 14 August 2016).
- 20 A preliminary proposal for the inclusion of this script in Unicode has been prepared by Anshuman Pandey, www.unicode.org/L2/L2013/13231-intro-mossang-tangsa.pdf (accessed 17 July 2015).
- 21 Weidert’s tone 2 is here treated as tone 3, and his tone 3 as tone 2.

REFERENCES

- Barkataki-Ruscheweyh, Meenaxi. 2015. *Performing Ethnicity to Resist Marginalisation: The Tangsa of Assam*. PhD thesis, Vrije Universiteit, Amsterdam.
- Barkataki-Ruscheweyh, Meenaxi and Stephen Morey. 2013. Wiwu song of the Pangwa Tangsa: Poetry and linguistic forms, meaning and the transformation to a symbol

- of identity. In Gwendolyn Hyslop, Stephen Morey and Mark W. Post (eds), *North East Indian Linguistics*, Volume 5, 280–303. Delhi: Cambridge University Press, India.
- Barua, Surendranath. 2013 (1991). *Tribes of the Indo-Burma Border*. Bhabani: Guwahati.
- Bradley, David. 1997. Tibeto-Burman languages and classification. In David Bradley (ed.), *Papers in Southeast Asian Linguistics No. 14: Tibeto-Burman Languages of the Himalayas*, 1–72. Canberra: Pacific Linguistics.
- Burling, Robbins. 2003. The Tibeto-Burman languages of Northeastern India. In Graham Thurgood and Randy J. LaPolla (eds), *The Sino-Tibetan Languages*, 169–191. London: Routledge.
- Census of India. 1931. *Census of India 1931 – Volume XI, Burma*. Rangoon (Yangon): Office of the Superintendent, Government Printing and Stationery.
- Damant, G. H. 1880. Notes on the locality and population of the tribes dwelling between the Brahmaputra and Ningthi rivers. *Journal of the Royal Asiatic Society of Great Britain and Ireland*, New Series, 12.2: 228–258.
- Das Gupta, K. 1980. *The Tangsa Language (A Synopsis)*. Shillong: North East Indian Frontier Agency.
- Defence Department, Government of Burma. 1942. *Brief Notes on the Naga Hills District (Burma) with Details of Routes*. Simla: Government of India Press.
- Dewar, T.P. 1931. Naga tribes and their customs. In *Census of India 1931 – Volume XI, Burma*, 267–295. Rangoon: Office of the Superintendent, Government Printing and Stationery.
- Grierson, Sir George 1903. *Linguistic Survey of India*. Volume III: Tibeto-Burman family, Part II: Bodo-Nāgā & Kachin groups. Calcutta: Office of the Superintendent of Government Printing. Republished in 1966, Delhi: Motilal BanarsiDass. [Much of this volume was written by Sten Konow.]
- Lewis, M. Paul, Gary F. Simons and Charles D. Fennig (eds). 2014. *Ethnologue: Languages of the World* (17th edition). Dallas, TX: SIL International. Available at: www.ethnologue.com (accessed 29 October 2014).
- Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley: University of California Press.
- Morey, Stephen 2011. Transitivity in Cholim Tangsa. In Frantisek Kratochvil, Alexander R. Coupe and Randy J. LaPolla (eds), special issue of *Studies in Language* 35.3: 676–701.
- Morey, Stephen. 2015. The internal diversity of Tangsa: vocabulary and morphosyntax. In Mark W. Post, Scott DeLancey and Stephen Morey (eds), *Language and Culture in Northeast India and Beyond: In Honor of Robbins Burling*, 23–40. Canberra: Pacific Linguistics (Australia-Pacific, Language and Peoples of the Eastern Himalaya Region (series)).
- Morey Stephen forthcoming. The sociolinguistic contexts of the Tangsa languages.
- Needham, F.J. 1897. *A Collection of a Few Moshang Naga Words*. Shillong: Assam Secretariat.
- Saul, Jamie. 2005. *The Naga of Burma*. Bangkok: Orchid Press.
- Simai, Chimo. 2008. *A Profile of the Tikhak Tangsa Tribe of Arunachal Pradesh*. New Delhi: Authorspress.
- Statezni, Nathan. 2013. *Fifty-five Dialects and Growing: Literacy and Comprehension of Vernacular Literature among the Tangshang Naga in Myanmar*. MA thesis, Dallas: Graduate Institute of Applied Linguistics.

- Statezni, Nathan. unpublished. *Names for the Tangshang of Myanmar*.
- Survey of India. 1927. *Map: Assam and Burma, Lakhimpur District, Sadiya Frontier Tract and Tribal Area*. No. 83 M/S.E. Calcutta: Survey of India Office.
- Weidert, Alfons. 1979. The Sino-Tibetan tonogenetic laryngeal reconstruction theory. *Linguistics of the Tibeto-Burman Area* 5.1: 49–127.
- Weidert, Alfons. 1987. *Tibeto-Burman Tonology*. Amsterdam: John Benjamins.

§3.2 Bodish languages

CHAPTER EIGHTEEN

CLASSICAL TIBETAN

Scott DeLancey

1 OLD AND CLASSICAL TIBETAN

Tibetan is attested from the mid seventh century CE, when the nascent Tibetan Empire began to imitate the Chinese fashion of recording contemporary political events. The written language of documents and inscriptions of the Tibetan Empire is referred to as Old Tibetan. Classical Tibetan refers to the language of most documents written after a major language reform and standardization in the ninth century. The term Written Tibetan is also used, typically with specific reference to the phonological and morphological system reflected in the orthography. “Written” or “Classical” Tibetan has been taken in comparative Sino-Tibetan studies as a direct representation of spoken Old Tibetan or Proto-Bodic, but this assumption is problematic (Denwood 2007). The Tibetan writing system is derived from the northwestern variety of the seventh-century Gupta script. The invention of the script, as well as the compilation of the first Tibetan grammar, is traditionally attributed to the legendary minister Thon-mi Sambhota (Miller 1976) in the reign of Srong=btisan sgam-po. The traditional date of the invention of the script is 632 CE. In reality the development of the script was more gradual (van Schaik 2011), but it is clearly in official use by 650 (Takeuchi 2013).

Many early documents were taken from the Dunhuang library a century ago (Thomas 1935–63), but only recently have Old Tibetan materials begun to receive significant scholarly attention; see, for example, Dotson (2009) and the invaluable Old Tibetan Documents project in Tokyo.¹ At last we are beginning to see serious linguistic study of Old Tibetan (e.g. Zeisler 2009; Hill 2010a, 2010b; Takeuchi 2013). This is an important development; previous work both on Tibetan itself and on its relevance to comparative Sino-Tibetan has been based entirely on Classical Tibetan, which is an idealization referring both to over a millennium of written history and to a tradition of prescriptive grammar. Tibetan texts over the last millennium vary in how faithful they attempt to be to the prescriptive grammar. Already in the vernacular *Mi=la Ras-pa'i rNam=thar*, the biography of Milarepa, which probably dates to the fourteenth century, we find a mixture of Classical and more modern features. Western grammars have generally been guides for reading the texts rather than comprehensive grammatical descriptions. Hahn (1996) is an excellent example; Kesang Gyurme (1992) and Beyer (1992) perhaps come closest to a linguistic description of the language. The classic Jäschke (1929, 1954) is useful mainly as a concise summary of the morphographemic alternations, but Francke’s *Addendum* to the 1929 edition contains many useful observations.

Old Tibetan is generally considered to be the ancestor of all modern Tibetic languages (see Chapter 19 ‘Lhasa Tibetan’). It emerged very suddenly in history as a language of

empire (Takeuchi 2013). The resultant intense contact with neighboring and subject languages may have played an important role in shaping the language which we first encounter when the imperial Tibetans adopted the Chinese court practice of recording annals (Takeuchi 2013; Zeisler 2009). Tibetic is closely related to the East Bodish languages (Bumthang, Kurtöp, Dzala, Dakpa) and slightly more distantly to Tamangic. It has long been assumed to be part of a larger Bodic or Western major branch of Tibeto-Burman which includes Western Himalayan and, more distantly, Kham-Magar and Kiranti (Hyslop 2014), but these wider relationships have not yet been uncontroversially demonstrated.

2 PHONOLOGY

There are a handful of Classical Tibetan phonological alternations reflected in the orthography. These are all best interpreted as morphophonemic, and will be dealt with in the next section. In the examples in this chapter I have used hyphens or = to connect morphemes which must be considered part of a single lexical word, but it should be noted that these do not reflect anything in the orthography: the writing system marks syllable boundaries, but gives no indication of word or morpheme boundaries. For this reason I have not used hyphens with plural morphemes, case postpositions, nominalizing morphemes, or anything else which can be given a distinct gloss, although we can assume that these were probably clitics in older stages of the language as in the modern dialects. Hyphens are, however, used to mark off the non-syllabic allomorphs of the genitive, ergative, and terminative case forms, and = to separate the components of compounds.

2.1 Segmental inventory

There are several transliteration systems for Tibetan in current use; the system used here will be clear from the following consonant chart. The Tibetan writing system is generally assumed to directly reflect the phonology of Old Tibetan. There are 30 consonants and five vowels; *i*, *u*, *e*, and *o* are written; a syllable with no vowel symbol has the intrinsic vowel /a/. There is no evidence for tonal distinctions in Classical Tibetan. The consonants are here presented in the traditional alphabetical order, read left to right and top to bottom:

ཀ	ཁ	ག	ང
k	kh	g	ng
ཆ	ຂ	ཇ	Ն
c	ch	j	ny
ဋ	ဋ	ဋ	ဋ
t	th	d	n
Phonetic symbols	Phonetic symbols	Phonetic symbols	Phonetic symbols
ප	Փ	Բ	Մ
p	ph	b	m
Չ	Չ	Շ	Շ
ts	tsh	dz	
Ց	Ց	Ց	Ց
w	zh	z	
Յ	Յ	Յ	
y	r	l	
Ռ	Ռ	Ռ	Ռ
sh	s	h	?

ି ଉ ଏ ଓ
i u e o

The phonetic interpretation of most of the letters is straightforward, the most problematic being ା and ା (transliterated here as ‘and ?; in the Wylie system ା is not represented) (Li 1933; Hill 2005). The letter ା apparently represented a glottal stop, as reflected in its reflexes in modern dialects. The interpretation of ା, often called “a-chung” (‘little a’) in the Western literature, is problematic. The writing system requires a place-holder in root-initial position, and ା occurs there when there is no other consonant: ଽା ‘o ma ‘milk,’ etc. In Central Tibetan these words have initial breathy [h], and low tone. But ା is also a prefix, occurring only before voiced and voiceless aspirated stops and affricates: ାୟ ‘byung ‘emerge,’ etc. In many modern dialects the reflexes of orthographic ା + obstruent onsets are prenasalized obstruents: ଽାୟ ‘gro /ⁿtʃo/. Whether these two functions of the letter represent the same phonological unit remains a matter of controversy. The third orthographic use of ା is to indicate the position of the vowel in an otherwise ambiguous sequence of consonants. For example, the sequence ଽାୟ dg will be read as the onset and coda of a syllable with the default vowel: /dag/. But ଽାୟ d-g is also a possible onset, so the syllable /dga/ must be written ଽାୟା, transliterated as *dga'*.

Though the writing system distinguishes three series of stops, the voiceless unaspirated stops are vanishingly rare initially in lexical stems; in the verb they are in complementary distribution with aspirated stops depending on the prefix (see later). Thus their phonemic status in Classical Tibetan is at best marginal.

2.2 Syllables

The writing system is syllabic: syllable and phrase boundaries are marked, but there is no indication of word boundaries. The orthographic system identifies one member of an onset cluster as the root initial. Consonants which precede the root initial are written above or before it; consonants which follow are written beneath it. For example, *brgyad* ‘eight’ is represented as ଽାୟା, with a reduced form of the ା r superjoined to the g, and a reduced form of the ା y subjoined to it, giving ା. Sometimes an upper-case letter is used in transliteration to indicate the root initial: *brGyat*.

In the European Tibetanist tradition consonants preceding the root initial are referred to as “prefixes”; in fact, while most or all of these may diachronically derive from morphological prefixes or reduced syllables which were originally distinct compounded roots, in many cases they cannot be assigned separate morphological status or even etymologies in the Classical language.

The Classical Tibetan syllable is very close to that of Proto-Tibeto-Burman. The consonants written below the initial are limited to the liquids and glides *r l w y*, the Proto-Tibeto-Burman medial inventory. Combinations of root initials with one of these are probably the only original tautomorphemic onset clusters in the language. The difference between these original clusters and prefix+initial sequences must have had phonological significance, as they have different reflexes in the modern dialects.

There are two series of prefixes: *r l s* written above the initial, and *b d g m* ‘preceding it. Thus the transliterated sequence *gy* could represent either root-initial *y* with prefixed *g* ାୟ, or root-initial *g* with medial *y* ାୟ. Capitalization of the root initial eliminates this ambiguity: ଽାୟା *Gyon pa* ‘wear,’ ଽାୟାୟ *gYon pa* ‘left (side).’ Most transliteration systems indicate the sequence of prefix *g-* and root-initial *y* with some mark of separation: *g.yon pa* or *g-yon pa* ‘left.’

Classical Tibetan provides evidence for one productive inflectional suffix, *-s*, primarily associated with past stems. A *-d* allomorph occurs in some old texts, but is attested primarily by its effect on the sentence-final particle (§4.1.2). Thus the Old Tibetan syllable canon is (C_p)C(G)V(C)-(s/t), where C_p can be a voiced stop *b d g*, a liquid *r l s m*, or *.* Possible root-final codas are the Proto-Tibeto-Burman inventory of *b d g* (presumably pronounced lenis and voiceless), *m n ng r l s*.

2.3 Morphographemic alternations

Several grammatical morphemes, presumably clitics, show alternations in the initial consonant depending on the final of the preceding word:

- The genitive and the ergative/instrumental case markers show the following allomorphs: gen. *gi*, erg. *gis* following velars *g ng*; *kyi*, *kyis* following obstruents *d b s*; *gyi*, *gyis* following sonorants *n m r l*. Following vowels the genitive is *'i*, the ergative *s*, both written as part of the preceding syllable.
- The terminative case marker is *r* or *ru* following vowels, *su* following *s*, *tu* following *b g du* elsewhere.
- The indefinite article *cig* has the allomorphs *zhig* after sonorants and *shig* after *s*. The imperative *cig*, the continuative non-final *cing*, and the quotative *ces* show the same pattern of allomorphy.
- The generic non-final particle (*s)te* appears as *de* after *d*, *te* after other coronals *n r l s*, and *ste* elsewhere.
- The pragmatic particle *yang* has the allomorph *kyang* after obstruents *b d g s*.

3 NOMINALS AND THE NOUN PHRASE

3.1 Nouns

There are a fair number of monosyllabic nouns, especially in basic vocabulary, but, as in other Sino-Tibetan languages, nouns tend to be bimorphic and disyllabic. There is also a large inventory of Noun + Verb constructions (“light verb” constructions) which must be considered fixed lexical items, but we can see from the modern languages, where they are more frequent, that these do not represent formal compounds in the same sense as the compound nouns.

3.1.1 Nominal suffixes

Many nouns consist of only the simple noun stem: *mi* ‘person,’ *bod* ‘Tibet,’ *rta* ‘horse,’ *ja* ‘tea,’ *mgo* ‘door.’ More consist of a stem plus a synchronically meaningless nominal suffix, usually *pa/ba* or *ma*, less commonly *po/bo* or *mo*: *bod-pa* ‘Tibetan (person),’ *sha-ba* ‘deer,’ *o-ma* ‘milk,’ *khra-ma* ‘window.’ In many words with natural gender these formatives specify gender, *pa/po* indicating masculine and *ma/mo* feminine, as, *bod-mo* ‘female Tibetan,’ *rgyal-po* ‘king,’ *rgyal-mo* ‘queen.’ For many animals the stem is interpreted as masculine, with the feminine form requiring a suffix: *khyi* ‘dog,’ *khyi-mo* ‘bitch.’ The significance of these gender-specific suffixes on nouns which do not have natural gender is unclear, but it does appear that they were more frequent in earlier stages of the language, e.g. earlier *mda-mo* ‘arrow,’ for modern *mda*.

3.1.2 Compound nouns

Other nouns are disyllabic compounds, comprised of two noun, verb, or adjective stems. In the native grammatical tradition compounds are categorized according to the Sanskrit system:

- Synonym compounds: *nam=mkha* ‘heaven’ (*gnam* ‘heaven,’ *mkha* ‘heaven’);
- Tatpurusha (“determinative”) compounds: these may be head-final, as *dmag=mi* ‘war=person’ = ‘soldier,’ or head-initial, as *glang=chen* ‘ox-great’ = ‘elephant’;
- Dvandva (“copulative”) compounds: *yab=yum* ‘father=mother’ = ‘parents,’ *rkang=lag* ‘foot=hand’ = ‘limbs’;
- Bahuvrihi (“attributive”) compounds, with one member in a possessive relation to the other: *seng=gdong-ma* ‘lion=face-FEM’ (a lion-faced female deity).

Bahuvrihi proper are the least common type, but occur frequently as titles or personal or place names.

3.2 Adjectives

As in other Tibeto-Burman languages, adjectives are formally nouns. Adjectival forms are bimorphemic, consisting of a stem plus one of the nominal suffixes *pa/ma/po/mo* or *ka* (§3.1.1): *chen-po* ‘large,’ *legs-pa* ‘good,’ *gsha-ma* ‘worthy.’ As with nouns, a few adjectives may express the natural gender of their referent by alternating the masculine *pa/po* and feminine *ma/mo* suffixes, but most adjective forms are fixed.

3.3 Pronouns

Personal pronouns found in the texts include a range of forms, differing probably in dialectal as well as honorific status. The earliest Old Tibetan documents already show evidence of honorific distinctions (Hill 2010c). The common forms are first person *nga*, *nged*, *ngos*, and humilific *bdag*; second person *khyod* and honorific *khyed*. Less commonly *kho bo* (masculine) and *kho mo* (feminine) are used for first person. Plural marking is not obligatory on nouns or pronouns, but all pronominal stems can be pluralized using *cag*, *tsho*, or *rnam* (§3.4). The stem ‘*u* occurs only in first person pluralized forms: ‘*u cag*, ‘*u bu cag* ‘we.’ All pronominal forms with personal reference frequently occur with one of the emphatic/reflexive suffixes *nyid* or *rang*: *nga rang*, *nga nyid* ‘I; *de nyid* ‘that one, s/he.’ Emphatic *rang* occurs alone with first person singular reference; *nyid* for second or third person.

Anaphoric third person reference is commonly zero, for objects as well as subjects (Andersen 1987). When it is explicit, the ordinary form for third person pronominal reference is the distal demonstrative *de*, but the more modern forms masculine *kho*, feminine *mo*, and honorific *khong* also occur. Less common gender-marked forms are masculine *kho ba*, feminine *kho ma*.

3.4 Number words

Number is not an obligatory category, but several number words can be used to specify plurality. The commonest are *tsho*, *rnam*, and *cag*: ‘*di tsho* ‘these,’ *nga cag* ‘we,’ etc. In the modern languages forms like *nga=tsho* ‘we’ are single words, and this is presumably the case as well for the Classical language, though this cannot be demonstrated on the

basis of written evidence. While these were probably phonologically bound in pronominal forms, as in the modern dialects, in a noun phrase they are not bound to the head:

mi chen-po de rnams
person great that PLURAL
'those important men'

3.5 Demonstratives and articles

The demonstratives are *de* 'that,' *di* 'this.' Both occur as specifiers of noun phrases and pronominally. Distal *de* also functions as a definite article, in which use it contrasts with the indefinite article *cig* (see §2.3), an unstressed form of *gcig* 'one.' Another postclitic, *bo*, is occasionally encountered as a definite article. The relation of this form to the gender-specific nominalizing particles *po* and *mo* (§3.1.1) is unclear. This form is still used in the Western languages, but not in the Central region.

3.6 Order within the noun phrase

Demonstratives follow the head, genitive modifiers precede:

bya de- i rgyab-la
bird that-GEN back-LOC
'on the bird's back'

Adjectives, which are categorially nouns, and relative clauses, which are nominalizations, occur both preceding and following the head. Pre-head modifiers are marked as genitive:

gzhan zhig gi yul
other a GEN country
'another country'

sgas pa- i zas rnams
gather NMLZ-GEN food PLURAL
'the food which had been collected'

Post-head modifiers never are:

mtshan g=han-pa
sign other-NMLZ
'other signs'

shing=sdong che ru 'gro ba de
tree great TERMINATIVE become NMLZ that
'the tree which grows great'

The post-head construction can be interpreted as appositive; straightforward nouns also occur in this construction:

nga ma rgan=mo ni bsam=mno btang
1sg mother old.woman TOPIC wishes send
'I, your mother, an old woman, wishing . . .'

This is particularly common with numerals, especially ‘two’ and ‘three’:

yab=yum *gsum*
father=mother three

‘the parents, three’ (referring to two brothers and their wife in a polyandrous marriage)

3.7 Relator nouns

The only spatial postpositions are those listed in the section on case forms (3.8). More explicit spatial reference is accomplished by specialized nouns in a relator noun construction, in which the relator noun serves as the head, with the lexical noun as a genitive dependent:

rgyal=po-’i drung du ‘*dong*
king-GEN vicinity TERMINATIVE go
‘go before the king’

des rtsig-pa zhig gi steng nas mchongs pa
that(ERG) wall a GEN upper ELATIVE leapt NMLZ
‘he jumped off a wall’

Relator nouns may also have more abstract senses:

lta ba-’i ched du
see NMLZ-GEN purpose TERMINATIVE
‘in order to see’

3.8 Noun

3.8.1 Case forms

Postpositions following noun phrases mark case relations. Although some allomorphs of the case markers are phonologically tightly bound to a preceding host, they are clitics rather than inflections, as they attach to the last word of a noun phrase, whether or not it is the head noun. The case forms are (for the distribution of allomorphs see §2.3):

genitive *kyi/gi/gyi/’i*
ergative/instrumental *kyis/gyis/gis/-s*
locative/allative *la*
locative/illative *na*
ablative *las*
elative *nas*
terminative *r ~ ru/su/tu/du*

The locative/ablative forms in *l-* indicate ‘to, at’; the locative/elative forms in *n-* ‘into, in.’ The “terminative” case specifies a goal; this case has a range of somewhat idiosyncratic uses. (For more discussion of these forms and their functions, see Beyer 1992: 267–70; Tournadre 2010.)

4 VERBS

4.1 The verbal system

The Classical Tibetan verb distinguishes four stems. These are generally called **present**, **past**, **future**, and **imperative**, though in terms of their discourse function the present and past are better interpreted as imperfective and perfective. The meanings and uses of the four stems are analyzed in detail in Zeisler (2004). The stems are distinguished by prefixation (sometimes with accompanying change in root-initial consonant), suffixation, and vowel ablaut. For most verbs some combination of ablaut, prefixation, and suffixation distinguishes all four stems. I will summarize the most regular patterns here; more extensive presentations and discussion can be found in Durr (1950); Shafer (1951); Hahn (1996); Beyer (1992) and Coblin (1976). A major concern of linguistic research on Classical Tibetan has been the internal reconstruction of a more coherent verbal system than that which is attested (Li 1933; Durr 1950; Shafer 1951; Uray 1953). Early work in comparative Tibeto-Burman took it for granted that the Classical Tibetan verbal morphology was ancient; it is now clear that it is mostly a Tibetic innovation (Coblin 1976; Jacques 2012).

4.1.1 Ablaut

Most verbs show no ablaut. Those that do fall into several categories. First, many verbs with *a* or *e* have *o* in the imperative (all examples from Jäschke 1881 or Dorje Wangchuk Kharto n.d.):

	Present	Past	Future	Imperative
‘see’	<i>lta</i>	<i>bltas</i>	<i>blta</i>	<i>ltos</i>

This phenomenon must reflect a pre-Tibetan imperative suffix **o*, which is still attested in other Tibeto-Burman languages. The other patterns, all less common, are:

o a a o: ‘overthrow’	<i>rlog</i>	<i>brlags</i>	<i>brlag</i>	<i>rlogs</i>
e a a o: ‘fill up’		<i>gengs</i>	<i>bkang</i>	<i>dgang</i>
i u u u: ‘remove’		<i>byin</i>	<i>phyung</i>	<i>dybung</i>

4.1.2 Suffixation

The only productive inflectional suffix is *-s*. This is characteristically found distinguishing the past and imperative stems from the present and future:

‘finish’	<i>sgrub</i>	<i>bsgrubs</i>	<i>bsgrub</i>	<i>sgrubs</i>
----------	--------------	----------------	---------------	---------------

This *-s* never occurs following coronal finals *d n l r s*. In some older texts a *-d* allomorph occurs after coronal finals (Przyluski and Lalou 1933; Zhang 1981). This is mostly lost in later texts; though some conservative authors continue to write *-d* in these verbs, it must have ceased to be pronounced quite early. In many older and conservative texts, after the past stem of these verbs the final particles (see §3.4) are *to/tam*, e.g. *gyur to* ‘became.’ Since the final particles copy the final consonant of the preceding verb, this must reflect earlier occurrence of the *-d* allomorph.

Other stem alternation patterns show evidence for earlier suffixal morphology which is no longer productive. In some verbs an *-s* occurs in all stems but the future:

‘pursue’	<i>snyegs</i>	<i>bsnyegs</i>	<i>bsnyeg</i>	<i>snyegs</i>
----------	---------------	----------------	---------------	---------------

And in others it occurs only in the present stem (see ‘fill up,’ earlier). It is not clear what connection there is or is not between this and the past *-s*. Other alternation patterns include final *-n* in the present, *-ng* in other stems (see ‘remove’ earlier), and final *-d* in the present stem only:

‘do’ *byed byas bya byos*

4.1.3 *Prefixation*

Most intransitive verbs lack a distinct imperative stem, using the past instead for this function; many also do not distinguish the future from the present. In a smaller number of cases the imperative is identical with the present. In the regular pattern, the present stem has prefixed ‘- where possible (i.e. with root-initial stops and affricates). Some verbs have suffixed *-s* in the past stem:

‘emerge’ *‘byung byung byung*
 ‘whirl’ *‘tshub tshub tshubs*

Where prefixed ‘- is phonologically impossible, the present and past stems are distinguished only by the *-s* suffix:

‘weep’ *ngu ngus*

Transitive verbs fall into several inflectional classes according to the pattern of prefixation which they show. All transitive verbs prefix *b-* in the past stem, usually with the *-s* suffix, which also occurs in the imperative stem. The present stem may be marked either by prefixed ‘-, as in the intransitives, or *g/d* (*g* before coronals, *d* before labials and velars); the future by prefixed *g/d* or *b*. The imperative is typically unprefixed, though it may also show the *g/d* prefix, and generally takes the *-s* suffix. The regular patterns can be seen in Table 18.1.

The distribution of these prefixes is phonologically constrained: ‘- can occur only before stops and affricates; *b* only before fricatives, liquids, and non-labial non-aspirated stops and affricates.

4.1.4 *Initial alternations*

Aspirated root-initials can occur as initials or with the *m* or ‘prefixes. With any other prefix they are replaced by unaspirates:

‘drink’ *‘thung btungs btung thung*

TABLE 18.1 VERBAL PARADIGMS IN CLASSICAL TIBETAN

Prefix alternation	Present	Past	Future	Imperative	Gloss
‘b g	‘debs	btab	gtab	thob	‘hit’
‘b b	‘chos	bzos	bzo	chos	‘make’
g b g	gcod	bcad	gcad	chod	‘cut’
g b b	gtsab	btabs	btabs	gtsabs	‘mince’

There are fewer than half-a-dozen verbs with unprefixed initial unaspirates; the unaspirated series otherwise occurs in the verbal system only following the *b d g r l s* prefixes in alternation with unaspirates.

4.2 Transitivity derivations

The opposition between transitive and intransitive (or control and non-control) verbs is reflected in several frozen and fossilized morphological processes.

4.2.1 *The s- prefix*

Many causative verbs are derived from intransitive stems through an *s-* prefix of Proto-Sino-Tibetan provenance: *nub* ‘sink, decay, decline,’ *snub* ‘destroy, cause to perish,’ *ring* ‘long,’ *sring* ‘stretch, extend.’ With stop initials the *s-* prefix alternates with ‘-’ in the intransitive form: *khol* ‘boil (intr.)’ *skol* ‘boil (tr.)’ (*s* cannot occur before aspirates), *byor* ‘stick, adhere to,’ *sbyor* ‘affix, fasten to.’ Some pairs are further distinguished by a coronal final in the transitive form, lacking in the intransitive: *na* ‘sick,’ *snad* ‘injure.’

4.2.2 ‘- prefix

Intransitive verbs may be derived, usually from noun stems, by a prefixed ‘-: *grib* ‘shade,’ *grib* ‘grow dim,’ *grogs* ‘friend,’ *grogs* ‘be associated.’ This has been compared to intransitive derivations by initial voicing found in other Sino-Tibetan languages, but see §4.2.3.

4.2.3 Ancient alternations

An older pattern involves unconditioned alternation in the voicing of the initial consonant. Most typically we find pairs in which the intransitive member has a voiced initial, while the corresponding transitive has a voiceless initial in the past and imperative stems:

	Present	Past	Future	Imperative
‘die’	‘ <i>gum</i>	<i>gum</i>		
‘kill’	‘ <i>gums</i>	<i>bkums</i>	<i>khums</i>	

Similar alternations are found elsewhere in Sino-Tibetan; the correlations are complicated by the fact that in some branches (including Central Tibetic) cluster simplification has produced similar alternations out of morphological derivations apparently cognate to those described in the preceding two sections. Moreover, as Hill (2014) notes, there are more complex patterns as well, in which distinct related intransitive stems differ in voicing:

	Present	Past	Future	Imperative
‘fill (intr)’	<i>gang</i>	<i>gang</i>		
‘be full’	<i>kengs</i>	<i>kengs</i>		
‘fill (tr)’	‘ <i>gengs</i>	<i>bkang</i>	<i>dgang</i>	<i>kong</i>

5 CLAUSE AND SENTENCE

Word order is generally SOV. Objects may precede subjects for discourse pragmatic reasons, but nothing but a sentence operator—a final or subordinating particle or a nominalizer—follows the highest verb.

5.1 Transitivity and alignment

Classical Tibetan is prescriptively consistently ergative, i.e. the subject of a transitive verb has overt case marking (identical to the instrumental case) contrasting with the zero marking of intransitive subjects:

<i>dbyig-pa-can de 'bros pa-r</i>	<i>brtsams te</i>
<i>dByig-pa-can that flee</i>	<i>NMLZ-TERMINATIVE begin</i>
'dByig-pa-can tried to flee'	
<i>dbyig-pa-can gyis bdag gi bu bsad do</i>	
<i>dByig-pa-can ERG 1sg GEN child killed FINAL</i>	
'dByig-pa-can killed my child.'	

Some texts show an aspectual split pattern, with ergative marking only in perfective clauses but most do not (see Andersen 1987; Tillemans and Herforth 1989; Saxena 1989; Dempsey 1993). Some texts show ergative marking on agentive intransitive subjects, as in modern Central Tibetan (Saxena 1989: 37–8):

<i>de-s rtsig-pa zhig gi steng</i>	<i>nas mchongs pa</i>
<i>that-ERG wall a GEN upper.surface</i>	<i>ELATIVE leapt NMLZ</i>
'he jumped off a wall'	

Takeuchi and Yoshiharu (1995) report examples of this from the Old Tibetan Chronicle. (They suggest that this may originally have been a distinct construction from the true ergative, but the prevalence of this kind of variable case marking in modern languages throughout the family suggests that we should analyze the Old and Classical patterns the same way.) Ergative NPs are typically, though not exclusively, definite (Andersen 1987).

The recipient argument of a trivalent verb is always marked with the locative *la*. Objects of transitive verbs are either unmarked or marked with *la*. This is not the pragmatically governed “antiergative” or “primary object” pattern found in many other languages. The case marking of objects is lexically determined: verbs representing a change of state in the object (‘kill,’ ‘cut,’ etc.) take unmarked objects, while those representing contact (‘hit,’ etc.) require locative marking:

<i>shing la sta-re gzhus pa</i>
<i>tree LOC axe hit NMLZ</i>
'hit the tree with an axe'

<i>sta-re-s shing 'chad pa</i>
<i>axe-INSTR tree cut NMLZ</i>
'cut down the tree with an axe'

No alternation is possible; i.e. *'chad-pa* can never take a locative-marked object, and *gzhus-pa* cannot take an unmarked object.

Tibetan has no morphologically or syntactically explicit voice alternation. The functional load of an active/passive alternation is divided among zero anaphora, NP-fronting, and the alternation of related intransitive/transitive verb pairs.

5.2 Sentence particles

A CT sentence ends with a final particle: declarative '*o*', interrogative '*am*', imperative *cig* (see §2.3):

dbyig-pa-can rgyal-bar gyur bla 'o
dByig-pa-can victory gain better FINAL
'Better dByig-pa-can should win.'

ma smras pa-s lce chod cig
NEG speak NMZ-INSTN tongue cut IMPERATIVE
'Since [he] didn't speak, cut out [his] tongue!'

The '-initial of the declarative and interrogative particles copies a previous final consonant:

rab-tu dbul-phongs so
very poor-poor FINAL
'[He was] very poor'

mi de dgra yin nam
person that enemy be INTERR
'Is that man an enemy?'

Like other Asian SOV languages, Tibetan has a clause-chaining structure; the major function of the declarative final particle is to mark the end of a clause chain. (See the example in §6.3).

5.3 Questions and negation

Yes–no questions are marked by the final particle '*am*’ (§5.2):

khyod kyis glang brnyas sam
2sg ERG ox borrow INTERR
'Did you borrow [his] ox?'

Content question words occur in their syntactically appropriate position; content questions have no final particle:

nga-s khyod kyi khyo ci-ltar sbyin
1sg-ERG 2sg GEN husband how give
'How can I give you back your husband?'

The negative marker precedes the highest verb. It is *ma* with past and imperative stems, *mi* with present and future: *ma bsgrubs* 'didn't finish,' *mi bsgrub* 'won't finish.' There is no distinct prohibitive form. The interrogative final particle '*am*' clearly represents a reduction of an earlier balanced question construction, probably: **V 'o ma-V 'V* (or) not *V?* > *V 'am 'V?*

The copulas *yin* and *yod* have special fused negative forms *min* and *med*. *Med* often occurs in the sense of ‘without, not having’:

rab-tu dbul-phongs-pa bza'-ba dang bgo-ba
very poor-poor-NMLZ food and clothes

med-pa zhig go
not.have-NMLZ a FINAL
‘[He was] a very poor one, one without food or clothing.’

5.4 Pragmatic particles

Two particles with pragmatic force are quite common in texts. *Ni* marks a continuative or resumptive topic:

khyim-bdag gi bu de ni ba-she-cir yin no
householder GEN son that TOPIC Ba-she-cir be FINAL
‘That son of the householder was Ba-she-cir.’

In this example the householder’s son has been mentioned briefly several lines previously, and is now being reintroduced as the central figure in the action of the story.

Yang (see §2.3) has the general sense of ‘also,’ ‘again,’ ‘even,’ and is also often used to mark direct contrast:

khyim=bdag gi bu-s di skad ces gsol to . . .
householder GEN boy-ERG this word thus ask FINAL
rgyal=po-s kyang de bzhiN du gnang nas
king-ERG also that like TERMINATIVE do(HON) ABL
‘The householder’s son made this request . . . The king, in turn, assented to that’

5.5 Complementation

The basic complement construction is simple nominalization with *pa*, but a range of other constructions are found. The most common are clauses marked with the terminative case, and complements of verba dicendi marked by *ces* (*zhes/shes*, see §2.3).

Purpose clauses, and complements in general which refer to a future event or state attendant upon the action of the main verb, are marked by the terminative case, either on the nominalized verb or directly on the verb stem:

dbiyig-pa-can de 'bros pa-r brtsams te
dByig-pa-can that flee NMLZ-TERMINATIVE begin NON-FINAL
‘dByig-pa-can tried to flee’

yul pha=rol=po-r ltad=mo lta-r song nas
country other.side-TERMINATIVE sights see-TERMINATIVE went NON-FINAL
‘[they] went to another country to see the sights’

Quotations are marked by *ces* (*zhes, shes*):

nga-s khyod kyi khyo ji-ltar sbyin zhes smras nas
1sg-ERG 2sg GEN husband how give QUOT spoke NON-FINAL
‘“How can I give you [back] your husband?” [he] said,’

The use of *ces* is reminiscent of Early Modern English *thus*, in that it does not always occur directly contiguous to the quoted material:

khyim=bdag gi bu-s 'di skad ces gsol to
householder GEN boy-ERG this word QUOT ask FINAL
'the householder's son made this request . . .'

5.6 Subordination and clause chaining

There are two morphemes which function only as subordinating particles: *(s)te*, the generic non-final, and *c^{ing}*, which indicates temporal overlap between the clause where it occurs and the next. Since it contrasts with *c^{ing}*, *(s)te* implies temporal succession where that concept is applicable, but, as with conjunctions like English *and*, this is a matter of pragmatic inference rather than part of the semantic content of the morpheme. (For the allomorphy of these particles see §2.2.)

The other subordinators are case postpositions, a common source for subordinators in Tibeto-Burman. The characteristic of subordination, as opposed to complementation, is that this case marker is attached to the bare verb stem, with no mark of nominalization. The semantic equations are:

elative	<i>nas</i>	> temporal succession, 'then'
locative	<i>na</i>	> if/when
instrumental	<i>gyis</i>	> cause, or logical inference

As is typical of Tibeto-Burman and other Asian SOV languages, Classical Tibetan discourse is organized into often lengthy clause chains. The primary function of the final particle '*o*' is to mark the end of a clause chain; non-final clauses are marked with various subordinators. The example below illustrates most of the characteristic features of Classical Tibetan sentence organization; the subordinating particles, which will be discussed below, are boldfaced, as is the final particle:

khyim=bdag gis bya=phrug de bu la byin nas
householder ERG birdling that boy DAT give ELATIVE
'The householder gave the baby bird to his son'

cher=skyes pa dang gcig la gcig chags te
grow.up NMLZ and one DAT one love NON-FINAL
'[they] grew up and loved one another'

khyim=bdag gi bu phan=chun du ltad-mo la 'gro na
householder GEN boy here.and.there TERMINATIVE sights LOC go LOC
'when the householder's son went around to see the sights'

bya de-'i rgyab la 'dug nas bya-s khyer zhing
bird that-GEN back LOC sit ELATIVE bird-ERG carry SIMULTANEOUS
'sitting on the bird's back, the bird carrying [him]'

nam=mkha' la 'phu te
sky LOC fly NON-FINAL
'flying in the sky'

'gro 'o
go FINAL
'[he] went'

The discourse unit here is roughly equivalent to an English paragraph, rather than a sentence. Some of the subordinating particles marking the non-final clauses correspond to categories familiar from European languages, but the (*s*)te particle, glossed as “non-final,” represents a category characteristic of clause-chaining languages, in that it carries no semantic information other than that this is not yet the final clause of the clause chain. In Western works on Tibetan it is often compared to a participial construction (primarily because that is its closest translational equivalent in Sanskrit), but in many cases its discourse function is closer to that of a conjunction like *and*.

NOTE

1 See <http://otdo.aa-ken.jp/> (accessed 12 August 2016).

REFERENCES

- Andersen, Paul Kent (1987) ‘Zero-anaphora and related phenomena in Classical Tibetan’, *Studies in Language* 11.2:279–312.
- Beyer, Stephan (1992) *The Classical Tibetan Language*. Albany: State University of New York Press.
- Coblin, W. South (1976) ‘Notes on Tibetan verbal morphology’, *T'oung Pao* 62:45–70.
- Dempsey, Jakob (1993) ‘“Ergativity in Milarepa’s *Rnamthar*”: another viewpoint’, *LTBA* 16.2:113–120.
- Denwood, Philip (2007) ‘The language history of Tibetan’, in Roland Bielmeier and Felix Haller (eds), *Linguistics of the Himalayas and Beyond*, 47–69. Berlin: Mouton de Gruyter.
- Dorje Wangchuk Kharto (n.d.) *Thumi'i dGongs gTer (The Complete Tibetan Verb Forms)*. Delhi: C.T. Khrto.
- Dotson, Brandon (2009) *The Old Tibetan Annals: An Annotated Translation of Tibet's First History*. Vienna: Verlag der Österreichischen Akademie der Wissenschaften.
- Durr, Jacques (1950) *Morphologie du verbe tibétain*. Heidelberg: Carl Winter.
- Hahn, Michael (1996) *Lehrbuch der klassischen tibetischen Schriftsprache*. Bonn: Indica et Tibetica Verlag.
- Hill, Nathan (2005) ‘Once more on the letter ཡ’, *LTBA* 28.2:107–137.
- (2010a) ‘An overview of Old Tibetan synchronic phonology’, *Transactions of the Philological Society* 108.2:110–125.
- (2010b) *A Lexicon of Tibetan Verb Stems as Reported by the Grammatical Tradition*. Munich: Bayerische Akademie der Wissenschaften.
- (2010c) ‘Personal pronouns in Old Tibetan’, *Journal Asiatique* 298.2: 549–571.
- (2014) ‘A note on voicing alternation in the Tibetan verbal system’, *Transactions of the Philological Society* 112.1:1–4
- Hyslop, Gwendolyn (2014) ‘Bodic’, *Language and Linguistics Compass* 8.6:243–270.
- Jacques, Guillaume (2012) ‘An internal reconstruction of Tibetan stem alternations’, *Transactions of the Philological Society* 110.2:212–224.
- Jäschke, H. A (1881) *A Tibetan–English Dictionary*. London: Routledge and Kegan Paul.
- (1929) *Tibetan Grammar*. Berlin: Walter de Gruyter. [Contains very useful Addenda by A.H. Francke.]
- (1954) *Tibetan Grammar*. New York: Frederick Ungar. [The 1929 edition with added reading material and vocabulary, but without Francke’s Addenda.]

- Kesang Gyurme [sKal-bZang ‘Gyur-Med] (1992) *Le clair miroir: Grammaire tibétaine*, translation, adaptation, and commentary by Nicolas Tournadre. Paris: Editions Prajña.
- Li, Fang-kuei (1933) ‘Certain phonetic influences of the Tibetan prefixes upon the root initials’, *BIHP* 4:135–157.
- Miller, Roy (1976) *Studies in the Grammatical Tradition in Tibet*. Amsterdam: John Benjamins.
- Przyluski, Jean, and Marcelle Lalou (1933) ‘Le *da-drag* tibétain’, *BSOAS* 7:87–89.
- Saxena, Anju (1989) ‘Ergative in Mi=la=ras=pa’i rnam thar’, *LTBA* 12.2:35–39.
- Shafer, Robert (1951) ‘Studies in the morphology of Bodic verbs’, *BSOAS* 13:702–724.
- Takeuchi, Tsuguhiro (2013) ‘Formation and transformation of Old Tibetan’, in Tsuguhiro Takeuchi and Norihiko Hayashi (eds), *Historical Development of the Tibetan Languages*, 3–18. Kobe: Research Institute of Foreign Studies.
- Takeuchi, Tsuguhiro, and Yoshiharu Takahashi (1995) ‘Split ergative patterns in transitive and intransitive sentences in Tibetan: a reconsideration’, in Yoshio Nishi, James Matišoff and Yasuhiko Nagano (eds), *New Horizons in Tibeto-Burman Morphosyntax*, 277–288. Osaka: National Museum of Ethnology.
- Thomas, Frederick (1935–63) *Tibetan Literary Texts and Documents concerning Chinese Turkestan*. Four volumes. (Oriental Translation Fund. New series, v. 32, 37, 40, 41). London: Royal Asiatic Society.
- Tillemans, Tom, and Derek Herforth (1989) *Agents and Actions in Classical Tibetan*. (Wiener Studien zur Tibetologie und Buddhismuskunde, Heft 21). Vienna: Arbeitskreis für Tibetische und Buddhistische Studien Universität Wien.
- Tournadre, Nicolas (2010) ‘The Classical Tibetan cases and their transcategoriality: from sacred grammar to modern linguistics’, *Himalayan Linguistics* 9.2:87–125.
- Uray, Géza (1953) ‘Some problems of the Ancient Tibetan verbal morphology: methodological observations on recent studies’, *Acta Linguistica Hungarica* 3:37–62.
- van Schaik, Sam (2011) ‘A new look at the source of the Tibetan script’, in Yoshiro Imaeda, Matthew T. Kapstein and Tsuguhiro Takeuchi (eds), *New Studies of the Old Tibetan Documents: Philology, History and Religion*, 45–96. Old Tibetan Documents Online Monograph Series vol. III. Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Zeisler, Bettina (2004) *Relative Tense and Aspectual Values in Tibetan Languages*. Berlin: Mouton de Gruyter.
- (2009) ‘Reducing phonological complexity and grammatical opaqueness: Old Tibetan as a lingua franca and the development of the modern Tibetan varieties’, in Enoch Aboh and Norval Smith (eds), *Complex Processes in New Languages*, 75–95. Amsterdam: Benjamins.
- Zhang, Liansheng (1981) ‘The puzzle of *da-drag* in Tibetan’, *LTBA* 9.1:47–64.

CHAPTER NINETEEN

LHASA TIBETAN

Scott DeLancey

1 TIBETIC LANGUAGES

Modern varieties of Tibetan are spoken across a broad area reaching from northern Pakistan to Qinghai, Sichuan, and Yunnan. Although the literature refers to Tibetan “dialects,” “Tibetan” is a family comprising six to eight very distinct languages, and at least two dozen or so varieties that could reasonably be considered distinct languages. Tournadre (2008, 2013) suggests speaking of “Tibetic languages” rather than “Tibetan dialects,” since there are Tibetic languages which on non-linguistic grounds are not considered Tibetan. For example, Dzongkha, the national language of Bhutan, is never referred to as Tibetan, although it is closely related to the Central Tibetan group to which Lhasa belongs, while the Amdo languages, which are genetically much more distant from Lhasa, are always called Tibetan by their speakers and their neighbors, and in official Chinese lists.

The subclassification of Tibetic is not established. There are four established nuclei: Amdo, spoken mostly in Qinghai and Sichuan, Khams in Yunnan and the eastern Tibetan Autonomous Region, Central or Ü-tsang (*dBu-gTsang*), including Lhasa and Shigatse, and Western, including Balti and Ladakhi in northwest India and Pakistan. Nishi (1986) and Tournadre (2013) each add other major branches, and there are a large number of recently described Tibetic varieties which do not fit easily into any of these schemes (Tournadre 2013; Sun 2014).

This chapter describes the dialect of educated Lhasa speakers, which is the basis for two widespread standard varieties, the standard Tibetan used for public purposes in the Tibetan Autonomous Region and to some degree in other Tibetan-speaking areas of China, and a semi-standardized variety used by the Tibetan government in exile in India, and used as a lingua franca among various diaspora communities and groups. These varieties are not clearly distinct, and each of the terms “Lhasa Tibetan,” “Standard Tibetan,” and “Modern (Spoken) Tibetan” can be found applied to each variety. Some useful analytical work on the language has been most accessibly presented in the form of teaching materials, for example Chang and Shefts 1964, Goldstein and Nornang 1970, Goldstein 1991, and Tournadre and Dorje 2003, 2005; we also have a reliable reference grammar (Denwood 1999) and dictionaries (Goldstein 1975; Goldstein with Ngawangthondup 1984).

2 PHONOLOGY

2.1 Transcription and transliteration

Examples in this chapter will be presented in orthographic transcription (see Chapter 18), as well as phonemic transcription where relevant. There is not yet a consensus on the phonological analysis of the language, and thus some inconsistency in the literature in phonemic transcription, though the analysis of Chang and Chang (Chang and Shefts 1964; Chang and Chang 1978–81) is accepted by many American scholars. The system

used in this chapter is close to that of Goldstein with Ngawangthondup (1984), with some modifications: /k^y kh^y/ instead of /k kh/, /k kh/ instead of /q qh/, /ä/ instead of /e/, and a substantively different representation of tone, which will be explained later.

2.2 Segmental inventory

The consonant phonemes of Lhasa Tibetan [LT] are given in Table 19.1.

The retroflexes/t̪ th/ reflect original stop + r clusters; /k^y kh^y/ reflect original velar stop + y clusters. Some speakers also have a series of prenasalized stops, reflecting both older nasal + stop clusters and sequences of “a-chung” + stop (see Chapter 18). Some speakers have all the voiceless nasals on the chart, others only [mh] and only in the allomorph of the negative prefix preceding aspirated stops, some none at all.

The vowel phonemes are given in Table 19.2.

The five original vowels, /i e a o u/ occur both short and long, in open and closed syllables. Fronted /ü ö ä/ reflect a lost coronal coda d n s or l, and occur only in long rhymes (oral and nasalized) and in glottal/falling tone syllables. Nasalization is contrastive for all vowel positions. Nasalized vowels are always long. In non-final position nasalization surfaces as a nasal consonant homorganic with a following obstruent: *bon* /phöön/ ‘the Bon religion,’ *bon-po* /phöömpo/ ‘an adherent of Bon.’ Nasalization will be represented as raised /n/.

The phonemic status of [ɔ] is debatable. It reflects original *og* and *or* rimes. Since both final consonants are still pronounced in some circumstances, it is more economical to treat [ɔ] as a conditioned allophone of /o/, although there are forms where only [ɔ] ever occurs, with no phonological conditioning environment in the surface representation (for example, the dative/locative form of the third person pronoun /kho/, orthographically *khor* is always [khɔɔ]). /a/ is realized as [ə] when unstressed, under vowel harmony, or before tautosyllabic /p, m/.

TABLE 19.1 THE CONSONANT PHONEMES OF LHASA TIBETAN

	Labial	Dental	Retroflex	Palatal	Palatalized	Velar	Glottal
stop: plain	p	t	t̪		k ^y	k	(?)
asp	ph	th	h		kh ^y	kh	
affricate: plain		ts		c			
asp		tsh		ch			
nasal: voiced	m	n		ñ		ŋ	
voiceless	(mh)			(ñh)		(ŋh)	
fricative		s		š			h
liquid: voiced		l	r				
voiceless		lh	rh				
glide	w			y			

TABLE 19.2 THE VOWEL PHONEMES OF LHASA TIBETAN

	Front	Central	Back
high	i ü		u
mid	e ö		o
low	ä	a	(ɔ)

2.3 Vowel harmony

LT shows dominant/recessive vowel harmony—all vowels are raised in any word which contains a high vowel. (For more detailed descriptions, including exceptions to this generalization, see Sprigg 1961; Chang and Chang 1968, 1978–81.) Central vowels are raised to their high counterpart; /a/ is raised to [ə]. Harmony applies both progressively and regressively, within compounds and from affixes to stems and vice versa:

'gro	/tʂo/ 'go'
'gro=gvi yin	/tʂu kiyiin/ 'go=FUTURE'
dkar=po	/kaapo/ 'white'
zhim=po	/ʂimpu/ 'delicious'

For some speakers /ä/ and /ɔ/ raise to /e o/ under vowel harmony; for other speakers raised /ä ɔ/ are slightly higher than /e o/ (Chang and Chang 1978–81: vii–xix).

2.4 Syllable structure

Phonetically, all syllables have a consonantal onset. In word-initial syllables with no phonological initial, there is an automatic laryngeal gesture. High tone syllables begin with [?], and low tone syllables with an approximation of the glottis which produces a weak breathy [h]. All consonants can occur as onsets; only /p m k ɳ/ occur as codas. If the palatalized velars /k^y/ and /kh^y/ and the prenasalized stops are analyzed as unitary segments, there are no consonant clusters. The inventory of syllable patterns is:

CV	short open rime
CVV	long open rime
CVN	nasal-final rime
CVV ⁿ	long nasalized rime
CVC	obstruent-final rime
CV?	glottalized rime ~ CV [^] long rime with falling tone

CVV rimes reflect original liquid codas /r/. CVVⁿ rimes reflect original nasal, and CV? and CV[^] original obstruent finals. Labial codas are retained in word-final position. Velar codas are variably present: *chang* ‘beer’ [chaaⁿ] or [chan], *lug* ‘sheep’ [luk], [lu[?]] or [luü]. All original coronal codas are lost in LT, but are reflected in fronting of non-front vowels.

2.5 Tone

Every syllable is intrinsically high or low tone; for most speakers low tone vowels have breathy articulation. High tone reflects original syllables with voiceless root initial (see Chapter 18) or sonorant initial with prefix; low tone syllables reflect original voiced obstruent or unprefixed voiced sonorant initials. Low tone is indicated by underlining of the vowel; high tone is unmarked.

LT has a word- rather than a syllable-tone system: high vs low tone is distinguished only on the first syllable of a word. Compounds and certain derived and inflected verb forms show a characteristic tone melody: H H if the first syllable is intrinsically high, L followed by an H (often pronounced as a mid rather than truly high pitch) if the first syllable is intrinsically low. In this chapter the morpheme boundary in words showing

these tonal patterns is represented by = in the transliteration. Suffixes and clitics separated by a hyphen show drastic phonological reduction and do not participate in the word tone pattern. Thus /CV/ represents a high tone syllable, /CV/ low tone, /-CV/ an atonal clitic, and /=CV/ a suffix or second member of a compound which will show the second part of the word tone contour determined by the preceding stem.

All falling-tone syllables alternate with a glottal-final rime, and always behave as closed syllables with respect to all phonological rules. These will be consistently represented here as glottal-final, i.e. *stag* /taʔ/ rather than /taà/ as in much modern work. Both the final glottal and falling tone can occur only in a final syllable. Non-finally these rimes are usually realized as long vowels: *bod* /phöʔ/ ‘Tibet,’ *bod=skad* /phökäʔ/ ‘Tibetan language’; irregularly in some words as short rimes: *bod-pa* /phöpa/ ‘Tibetan person.’

While the high/low distinction is universally recognized, other aspects of Tibetan tonal phonology are analyzed quite differently by various scholars. For discussion of the issues see Mazaudon 1977, 1984; Hu 1982; Duanmu 1991; Spragg 1993; Sun 1997, 2003.

3 NOMINALS AND THE NOUN PHRASE

As in other Tibeto-Burman languages, nouns tend to be bimorphemic and disyllabic. See section 3.1 and Chapter 18 for discussion of noun formation. Genitive modifiers (including one variety of relative clause) precede the head noun; all other NP elements follow.

3.1 Nouns

A substantial number of common nouns are monomorphemic and hence monosyllabic: *mi* /mi/ ‘person,’ *khyi* /kh̥i/ ‘dog,’ *gri* /thi/ ‘knife,’ *bod* /phöʔ/ ‘Tibet.’ But the majority of Tibetan nouns are disyllabic. A handful of disyllabic nouns are unanalyzable, but nearly all are bimorphemic. These are of two kinds: lexical noun or verb stems with a noun suffix, and compounds of two noun stems or, less commonly, of a verb stem and a noun.

The common noun suffixes are *-pa*, *-po*, and *-mo*. The latter two often indicate natural gender, as in *rgyal-po* ‘king,’ *rgyal-mo* ‘queen’; others are semantically empty. The *-pa* formative may be derivational, as in *bod-pa* ‘Tibetan person,’ but many such nouns are opaque, e.g. *lag-pa* ‘hand.’

Some nouns consist of a disyllabic stem, always at least etymologically bimorphemic, with the *pa* nominal suffix. Typically the final nominalizing morpheme is absorbed into the second syllable, so that the resulting noun is still disyllabic: *slob=gra-ba* /lapṭāā/ ‘student.’

A large number of nouns are compounds, e.g. *bod=skad* ‘Tibetan language,’ *nyal=khang* ‘bedroom’ (“lie down-room”). Tibetan has a strong predilection for disyllabic nouns. Thus when compounds are formed from disyllabic constituents, only one syllable of the constituent is used. In the case of stems which in isolation require a nominal suffix, only the stem enters a new compound: *smyug=ma* ‘bamboo,’ *zam-pa* ‘bridge,’ *smyug=zam* ‘bamboo bridge.’

3.2 Adjectives

Adjectives in LT show morphological similarities to both nouns and verbs. An adjective occurs in several forms. The bare adjective stem occurs only in the comparative construction, where it is inflected as a verb:

- (1) NP₁ NP₂-*las* *yag=gi red*
 -ABL good-IMPF
‘NP₁ is better than NP₂.’

In predicate or modifying position the adjective stem requires a nominal suffix, generally *-po*: *deb yag=po* ‘good book,’ *deb di yag=po red* ‘this book is good.’ A handful of adjectives take other nominal suffixes in this construction, e.g. *gsar=pa* ‘new.’ Earlier texts show a larger range of nominal suffixes here, with the feminine forms *-ma* and *-mo* attested as well.

3.3 Pronouns

The common pronominal forms are: first person *nga*; second person plain *khyod=rang*, honorific *khyed=rang*; third person masculine *kho*, feminine *mo*, honorific *khong*. All form the dual by adding *gnyis* ‘two,’ and the plural by adding *tsho*: *nga=gnyis /ŋañi?/, nga=tsho /ŋatsho/, khyed=rang=gnyis /kh'eranñi/*, etc.

3.4 Demonstratives

The demonstratives are ‘*di /ti/* proximal ‘this,’ *de /the/* distal ‘that,’ *pha=gi /phaki/* far distal ‘yon.’ These follow the head noun and any modifying adjective, and precede numerals in the unmarked construction. The distal demonstrative *de* functions as a definite determiner. The indefinite determiner is *cig /cik/*, from *gcig* ‘one’:

- (2) *rnam=lha zer=mkhan-gyi 'brog-pa cig*
Namla call=NMZ-GEN nomad a
‘a nomad named Namla’
- (3) *lo kha=she cig-gi rjes-la*
year few a-GEN after-LOC
‘after a few years’

Numerals occur as bare stems following demonstratives: *deb de gnyis* ‘those two books.’ They can also occur in the form and position of adjectives, preceding the demonstrative: *deb gnyis=po de* ‘those two books.’

3.5 Case

LT has four enclitics which encode case. These cliticize or suffix to the last element in the NP.

- **Genitive:** *gi/gyi/kyi* after consonants, realized as unstressed, often voiceless [ki] or [ki?]; *i* after vowels, realized as lengthening and fronting of a final vowel.
- **Ergative/instrumental:** *gis, gyis, kyis* after consonants, realized as unstressed, often voiceless [ki]; *-s* after vowels, realized as final glottal/falling tone, with fronting of a final vowel.
- **Dative/localative:** *la /la/* after consonants; *-r* after vowels, usually realized as lengthening of a final vowel, and lowering of /o/ to [ɔ].
- **Ablative:** *nas /nä/* or */nä?/*.

The fifth surface case category is the unmarked nominative/absolutive.

The stop-initial allographs of the genitive and ergative collapse into modern /ki/; the ergative has falling tone if pronounced in citation form, but this is usually not audible in running speech. Both are sometimes pronounced with an aspirated initial in careful speech: /thubtääⁿ khi/ ‘Thubten’s.’

Remnants of the somewhat more elaborated Classical Tibetan case system are found in restricted contexts in LT. The ablative *las* no longer occurs in a spatial sense, but is preserved in the comparative construction (section 3.2). The terminative *du* is still found in a number of frozen constructions, including relator noun constructions such as *NP-gi ched-du* ‘for the benefit of NP,’ and adverbs such as *sger-du* ‘privately, personally,’ *lhag=par-du* ‘especially.’

3.6 Relator noun constructions

Spatial location is typically marked by a combination of locative case and a relator noun forming the head of an NP of which the lexical noun is a genitive dependent:

- (4) *rkub=kyag-gi rgyab-la*
chair-GEN behind-LOC
'behind the chair'

Relator noun constructions occur in other than spatial senses as well, e.g. *blo=bzang-gi don=dag-la* ‘for Lobsang’s benefit.’ Several of the commoner relator nouns, such as *nang* ‘in,’ *sgang* ‘on,’ and *og* ‘under,’ do not allow genitive marking on the lexical noun:

- (5) *zim=chung(*-gi) nang-la*
bedroom(*-GEN) in-LOC
'in the bedroom'

3.7 Nominalization and relative clauses

The Tibetan nominalization and relativization systems are essentially the same; relative clauses are nominalized clauses used as genitive modifiers or appositives. This results in an unusually complex system of relativization (see DeLancey 1999). There are four nominalizers in the system: *mkhan* for actor nominalizations, *sa* locative/dative, and the default nominalizers *yag*, used for patients and instruments in non-perfective relative clauses, and *pa*, used in perfective relative clauses when the head noun is not the actor:

- (6) *mog=mog zhim-po bzo=mkhan*
momo delicious cook=NMLZ
'one who makes good momos'
- (7) *mog=mog zhim-po bzo=mkhan bu=mo de*
momo delicious cook=NMZ girl that
'the girl who makes good momos'
- (8) *mo-s bzos-pa*
she-ERG cooked-NMLZ
'what she cooked'
- (9) *mo-s bzos-pa-'i mog=mog*
she-ERG cooked-NMLZ-GEN momo
'the momos which she made'

- (10) *mo-s bzo-yag-gi mog=mog*
she-ERG cook-NMLZ-GEN momo
'the momos which she makes'

- (11) *mo-s mog=mog bzo=sa*
she-ERG momo cook-NMLZ
'the place where she makes momos'

Genitive marking of a preposed relative clause is obligatory with *-pa* and *=yag*, and optional with *=sa*; it does not occur with *=mhan* in colloquial speech.

Relative clauses may also follow the head noun, with no genitive marking possible:

- (12) *mog=mog kho-s bzos-pa*
'the momos which she made'

4 THE VERB

4.1 The stem

The Classical Tibetan verb is marked for certain derivational (i.e. transitivity) and inflectional (tense/aspect) categories by a complex system of prefixation, suffixation, and ablaut. (See Chapter 18.) The merger of unprefixed voiced obstruent initials with the voiceless aspirated series, and the subsequent simplification of all initial and final clusters, has eliminated some distinctions, and greatly reduced the already cloudy transparency of the system, but many verbs still exhibit alternations in tone, stem vowel, and initial consonant which constitute relict reflexes of this older system.

In Classical Tibetan a verb could have up to four distinct stems, traditionally called present, past, future, and imperative. Although the morphological markers which distinguished these stems have been lost, their phonological reflexes—tone and the aspiration of the initial—still distinguish two or three different stems for many verbs. A good summary of the stem alternation patterns in LT can be found in Goldstein and Nornang (1970: 53–6) or Goldstein with Ngawangthondup (1984: xvii–xviii). In this chapter present and past stems will be glossed with present and past forms of English verbs. There is some variation among speakers in which verbs retain alternate stems, which ones they retain, and sometimes even in the form of the stem.

Since verbs ordinarily occur with tense/aspect suffixes, the stem alternation no longer carries the weight of indicating tense or other categories. However, any auxiliary or verb suffix of any kind (including nominalizers and subordinators) requires a particular verb stem. For example, any form based on the nominalizer *-pa* requires the past stem, and any form based on *=ki* requires the present stem:

- (13) *bsad-pa yin /sä?payiiⁿ/* 'killed (PERFECTIVE PERSONAL)'

- (14) *gsod=gi yin /sü?kiyiiⁿ/* 'will kill' (FUTURE PERSONAL)'

Remnants of an old system of derivation of transitive from intransitive stems and vice versa by change in the stem-initial consonant are found throughout the Sino-Tibetan family. The consonantal and tonal alternations still extant in Central Tibetan reflect this ancient process as well as various later, Tibetan-internal derivation processes. The

following examples illustrate the commonest alternations: initial stops aspirated with intransitives and unaspirated with transitives, and sonorant-initial stems with low tone intransitive and high tone transitive:

Intransitive	Transitive
<i>khol</i> /khöö/	<i>bkol</i> /köö/ ‘boil’
<i>chag</i> /chaʔ/	<i>bcag</i> /caʔ/ ‘break’
<i>ril</i> /rii/ ‘fall’	<i>dbril</i> /rii/ ‘knock s.o. down’

4.2 Finite suffixes

Finite verbs take one of a set of finite suffixes which undergo drastic phonological reduction in normal speech. These morphemes are defined positionally by their final position in the verb complex.

The core of the tense/aspect/evidentiality system consists of historically nominalized verb stems in construction with one of the copular verbs *yin* /yiiⁿ/ and *red* /reʔ/ or the existentials *yod* /yöʔ/ and *'dug* /tuʔ/. This core paradigm, with rough glosses, is as follows (for further discussion of the use and meaning of several of these forms, see Chang and Chang 1981; DeLancey 1990; Jin 1979, 1983a, 1983b):

- <i>pa yin</i> /pa-yii ⁿ /	‘perfective personal’
- <i>pa red</i> /pa-reʔ/	‘perfective generic’
- <i>ki yin</i> /ki-yii ⁿ /	‘future personal’
- <i>ki red</i> /ki-reʔ/	‘future generic’
- <i>ki yod</i> /ki-yöʔ/	‘imperfective personal’
- <i>ki 'dug</i> /ki(-tuʔ)/	‘imperfective immediate’

(The /-tuʔ/ of the imperfective immediate form is regularly omitted in spoken LT.)

Other members of the finite suffix category can be identified, orthographically and/or by intuitive identification by native speakers, with a Written Tibetan verb stem. The perfective -*song* is an obsolete past stem of the verb ‘*gro* ‘go,’ replaced in modern LT by *phyin*; the other source verbs are still in use as main verbs. The others are:

- <i>song</i> /-soo ⁿ / direct-evidential perfective	(obsolete past stem of ‘go’)
- <i>bzhang</i> /-šaʔ/ indirect-evidential perfective	(< ‘put’)
- <i>byung</i> /-chuu ⁿ / perfective, speaker as Goal	(< ‘appear, come to pass’)
- <i>yong</i> /-yoo ⁿ / gnomic, future grounded in inherent tendency	(< ‘come’)

4.3 Conjunct/disjunct marking and evidentiality

Tibetic languages in general have unusual and complex systems of evidentiality, mostly variations on the same basic system which distinguishes three categories of knowledge: generic knowledge, which one knows because it is something that people know, personal knowledge, which one knows because it is part of one’s own experience, and immediate knowledge, which one knows only because one has perceived the fact or evidence of it (Tournadre 1996b; Zeisler 2004; DeLancey 2012b).

4.3.1 Evidentiality in the copulas

These are distinguished in the LT copular system:

- (15) *bod-la g.yag yog-red* (or *yod-pa red*)
Tibet-LOC yak exist-GENERIC
'There are yaks in Tibet.'
- (16) *nga-'i ama bod-la yod*
1sg-GEN mother Tibet-LOC exist.PERSONAL
'My mother is in Tibet.'
- (17) *bod-la moṭa mang-po 'dug*
Tibet-LOC auto many exist.IMMEDIATE
'There are lots of cars in Tibet.'

In (15), *yog-red* indicates that this statement is based on general knowledge, rather than personal experience or direct perception. In (16), with *yod*, is based on personal knowledge. In (17), '*dug* indicates that the statement is based purely on direct perception; it might be spoken by someone making a visit to Tibet after several decades' absence, who was not aware until their visit how much things have changed.

4.3.2 Evidentiality and person

There is an obvious connection between the personal category and first person, and a strong association of the immediate category with non-first person—since it is rare that one suddenly perceives something previously unknown about oneself. But in appropriate contexts statements about other persons can be expressed with the personal forms, as in (15), which I could not say if, for example, my mother is visiting Tibet as a tourist and I know this only because she told me. An ideal context would be one in which I grew up in Tibet with my mother, and she is still there in the old family home, and I Skype with her regularly, so that her presence in Tibet is something I know as part of my past and present life experience. There are also perfectly ordinary uses of the immediate category in a first person context:

- (18) *nga-r dngul tog=tsam yod*
1sg-DAT money some exist.PERSONAL
'I have some money.'
- (19) *nga-r dngul tog=tsam 'dug*
1sg-DAT money some exist.IMMEDIATE
'I have some money.'

Use of immediate '*dug*' in (19) indicates that the speaker has just discovered the fact, e.g. has just reached into his pocket and discovered some money that he had not known he had. The semantics of the distinction is discussed further in DeLancey (1990, 1992, 2012b), Zeisler (2004), Hill (2012), and further references found there.

4.3.3 Evidentiality in the verbal system

Verb suffixes based on the copulas have the same evidential value as the copula:

- (20) *kho(s) thang=ka 'gel=gyi yog-red*
he(-ERG) thangka hang=IMPF.GENERIC
'He hangs up thangkas.' [as, that is his job or regular function]
- (21) *nga(s) thang=ka 'gel=gyi yod*
1sg(-ERG) thangka hang=IMPF.PERSONAL
'I am hanging up thangkas.'
- (22) *kho(s) thang=ka 'gel=gyis (= gyi-dug)*
he(-ERG) thangka hang=IMPF.IMMEDIATE
'He is hanging up thangkas.' [based on the speaker's direct perception]

The relation between evidential categories in the verbal system and person is strict. In affirmative statements the personal knowledge forms can occur only with volitional verbs with a first person actor. A non-volitional verb with first person actor takes the immediate suffixes:

- (23) *nga mthong-gis (*gi yod)*
1sg see=IMPF.IMMEDIATE (*PERSONAL)
'I see [it].'

4.3.4 Conjunct/disjunct or egophoric patterning

The personal and the other two categories show a syntactic patterning which is referred to as CONJUNCT/DISJUNCT OR EGOPHORIC. In this pattern, found in several other Himalayan languages, "conjunct" or "egophoric" forms occur with first person in statements and second person in questions, "disjunct" or "allophoric" forms elsewhere:

- (24) *nga bod=pa yin*
1sg Tibetan be.PERSONAL>CONJ
'I am a Tibetan.'
- (25) *kho bod=pa red*
he Tibetan be.GENERIC>DISJ
'He is a Tibetan.'
- (26) *khyed=rang bod=pa yin-pas*
2sg Tibetan be.PERSONAL>CONJ-INTERROGATIVE
'Are you a Tibetan?'
- (27) *nga rgya=mi red-pas*
1sg Chinese=person be.GENERIC>DISJ-INTERROGATIVE
'Am I a Chinese?'

With verbs of speaking, conjunct forms occur in the copula when its subject is coreferential with the higher subject, disjunct forms when they are not:

- (28) *kho-s kho bod=pa yin zer=gyis*
he-ERG he Tibetan be.PERSONAL>CONJ say=IMPF
'He_i says that he_i is a Tibetan.'

- (29) *kho-s kho bod=pa red zer=gyis*
he-ERG he Tibetan be.GENERIC>DISJ say=IMPF
'He says that he_i is a Tibetan.'
- (30) *kho-s nga bod=pa red zer=gyis*
he-ERG 1sg Tibetan be.GENERIC>DISJ say=IMPF
'He says that I am a Tibetan.'

The existentials likewise show a distinction between conjunct *yod* and disjunct *'dug*. This and related phenomena in Lhasa and other Tibetic languages are the subject of an extensive literature which cannot be summarized here (see DeLancey 1990, 2012b; Zeisler 2004; Tournadre 2008, and references therein).

4.3.5 Further evidential distinctions

In the perfective system we find further evidential distinctions. In disjunct clauses *-pa red* contrasts with a direct evidential perfective *-song*, indicating a completed event which the speaker witnessed, and an indirect inferential perfect *-zhag*, which marks a clause as a report of an event whose occurrence the speaker infers from present traces. Thus the sentence below with either *-song* or *-zhag* can be glossed as 's/he broke the cup,' but *-song* is used when the speaker witnessed the event, while the sentence with *-zhag* could be used to report inference from cup shards found in a kitchen where the subject had been moments before:

- (31) *kho-s dkaryol bcag-song*
he-ERG cup broke-PERFECTIVE.DIRECT
'He broke the cup.'
- (32) *kho-s dkaryol bcag-zhag*
he-ERG cup broke-PERFECT.INFERENTIAL
'He broke the cup.'

One last perfective form, *byung*, occurs only in clauses in which a first person in statements, or a second person in questions, plays the role of a Goal or Experiencer:

- (33) *khong phebs-byung*
he(HON) went(HON)-PERFECT.PERSONAL_GOAL
'He came (to where I was).'
- (34) *kho-s nga-r gzhus-byung*
he-ERG 1sg-DAT hit-PERFECT.PERSONAL_GOAL
'He hit me.'
- (35) *nga-s kho=tsho mthong-byung*
1sg-ERG 3pl see-PERFECT.PERSONAL_GOAL
'I saw them.'

4.4 Auxiliaries

Auxiliary verbs can occur directly after a bare verb stem, but do not undergo phonological reduction, and are themselves conjugated with a final or non-final verb suffix. There

are about a dozen of these in LT. Many of these also occur as main verbs. A given auxiliary requires either the perfective or the imperfective stem.

Auxiliaries requiring the present stem:

<i>'gro</i>	/t <u>o</u> /	'distal motion'	(‘go’)
<i>yong</i>	/yoo ⁿ /	'proximal motion'	(‘come’)
<i>bsdad</i>	/tä ⁿ /	'progressive'	(‘sit, stay’)
<i>ran</i>	/rään/	'be time to, almost'	(‘appropriate’)
<i>thub</i>	/thup/	'can'	
<i>nus</i>	/nü ⁿ /	'dare'	
<i>shes</i>	/seen/	'know how to'	(‘know’)
<i>'dod</i>	/t <u>ö</u> /	'want to'	

Auxiliaries requiring the past stem:

<i>bzhag</i>	/š <u>a</u> ?/	'do with deleterious effect'	(‘put’)
<i>tshar</i>	/tsha <u>a</u> /	'completive perfect'	(‘finish’)
<i>myong</i>	/ňoo ⁿ /	'experiential perfect'	(‘taste’)
<i>chog</i>	/cho ⁿ /	'may'	

4.5 Reduplication

LT uses a reduplicated verb stem in several constructions. For example, the reduplicated past stem with perfective conjugation has a sense of ‘already for some time’ or ‘regularly, repeatedly’:

- (36) *nga-’i yi=ge ’di bris=bris-pa yin*
 1sg-GEN letter this wrote=wrote-PERF.PERSONAL
 ‘This letter of mine was written long since.’

- (37) *kho phyin=phyin-pa red*
 he went=went-PERF.GENERIC
 ‘He comes and goes; keeps going back and forth.’

For other uses of reduplication, see Wang (1988).

4.6 Nominal tense/aspect/modality constructions

Some verb suffixes require a copular or existential construction, for example *rtsis /tsi?* / ‘plan to,’ as in:

- (38) *nga ’gro=rtsis yod*
 1sg go=plan exist.PERSONAL
 ‘I’m planning to go.’

Phonologically *rtsis* shows compound phonology, and induces vowel harmony in the verb. Syntactically these constructions differ from ordinary inflected verbs in that they are made finite with the simple existentials *yod* and ‘dug’, rather than taking inflectional suffixes. Other members of this group include *long /loōⁿ/*, ‘have time to’ and *grab(s) /tap/* ‘about to, ready to.’

5 WORD FORMATION

5.1 Light verbs

There are no true compound verbs, but Tibetan has a very large set of “light” verbs, consisting of a noun and a semantically empty verb. There is a small set of common verbs which form these constructions. For example:

<i>skad gtong</i>	‘call, invite’	(<i>skad</i> ‘speech’)
<i>gru gtong</i>	‘sail’	(<i>gru</i> ‘boat’)
<i>skad=dpar gtong</i>	‘play records’	(<i>skad=dpar</i> ‘phonograph’)
<i>skad rgyag</i>	‘shout’	(<i>skad</i> ‘speech’)
<i>gom-pa rgyag</i>	‘walk’	(<i>gom-pa</i> ‘step (n.)’)
<i>rgyugs=śad rgyag</i>	‘comb’	(<i>rgyugs=śad</i> ‘a comb’)
<i>chams-pa rgyag</i>	‘catch cold’	(<i>chams-pa</i> ‘a cold’)

5.2 Honorifics

Central Tibetan, and other Tibetic languages such as Ladakhi and Dzongkha which have historically been spoken in a monarchical context, have a system of honorific vocabulary called *zhe=sa*. Honorific speech is used with and among people of relatively high social standing, reflecting the absolute social status of the addressee, rather than the relative status of speaker and addressee.

Besides the honorific pronouns, there is an honorific suffix *lags* /laâ/ suffixed to names and terms of address in second or third person reference. There are distinct honorific forms for a few verbs and a considerable number of nouns, and any verbal expression can be made honorific by the addition of the honorific *gnang* /naa/ ‘do’:

- (39) *kho-s yi=ge bris-song*
he-ERG letter write-PERFECTIVE.DIRECT
'He wrote a letter (plain).'
- (40) *khong-gis phyag=bris bris=gnang-song*
s/he(HON)-ERG letter(HON) write=HON-PERFECTIVE.DIRECT
'S/he wrote a letter (honorific).'

6 SYNTAX

6.1 Clause and sentence

Unmarked order within the clause is SOV, but any preverbal constituent can be fronted:

- (41) *phru=gu-s mog=mog sgan=ga zas-zhag*
child-ERG momo all ate-PERFECT.INFERENTIAL
'The children ate all the momos.'
- (42) *mog=mog phru=gu-s sgan=ga zas-zhag*
'idem.' (e.g. in answer to the question "Are there any momos left?")

A sentence in LT is defined by the occurrence in final position (or directly preceding a final question particle) of a verb with marking for a complex t/a/e category. An inflected LT verb complex is either non-final, and thus followed by a subordinating particle of some kind, or final, and thus followed by one or more finite verb suffixes. Like other Tibeto-Burman languages, Tibetan has a clause-chaining discourse structure. Sentences consisting of several clauses with only one finite verb are common, and two-clause chains are the commonest sentence type in connected discourse. It is quite common to find two-clause chains in which all arguments of the second clause are present in the first, and thus zeroed in the second:

- (43) *stag-gis gyag-la so brgyab-byas bsad-pa red*
 tiger-ERG yak-DAT bit-NF kill-PERFECTIVE.GENERIC
 ‘The tiger bit the yak and killed it.’
- (44) *kho-s kha=lags zas-byas phyin-pa red*
 he-ERG meal ate-NF went-PERFECTIVE.GENERIC
 ‘He ate and left.’

The non-final marker (glossed “NF”) is obligatory in such constructions, and thus serves as a diagnostic to distinguish fortuitous concatenation from grammaticalized serial verbs (section 3.4).

6.2 Alignment

The case role marked by the ergative has been extensively discussed in recent literature (DeLancey 1990, 2012a; Tournadre 1991, 1996a). In elicited examples it marks transitive subjects in perfective clauses and optionally in non-perfective clauses, and active intransitive subjects in perfective clauses and under limited conditions in non-perfective clauses:

- (45) *blo=bzang-gi nga-r mthong-byung*
 Lobsang-ERG 1sg-LOC see-PERF/CONJUNCT-GOAL
 ‘Lobsang saw me.’

In connected discourse, while ergative case is more common in these syntactic contexts, there is no syntactic or semantic context where it is obligatory. Rather, the occurrence of the ergative always indicates some emphasis or contrast, either emphasizing the agentivity of an argument or indicating pragmatic contrast (Saxena 1991; Tournadre 1991; DeLancey 2012a).

The same case form marks instruments, with no added pragmatic force:

- (46) *blo=bzang-gis me=mda-s stag bsad-pa red*
 Lobsang-ERG gun-INST tiger killed-PERF/DISJUNCT
 ‘Lobsang killed a tiger with a gun.’

The *la* case, sometimes referred to as “dative/locative” in grammars, marks not only oblique locatives, but also recipients in trivalent constructions, experiencer arguments of predicates such as ‘like’ and ‘need,’ and possessors in ‘have’ constructions:

- (47) *nga-s blo=bzang-la ngul-tsam sprad-pa yin*
 1sg-ERG Lobsang-LOC money-some give-PERF/CONJUNCT
 ‘I gave some money to Lobsang’

- (48) *blo=bzang-la ngul dgos=kyi*
Lobsang-LOC money need-IMPF/DISJUNCT
'Lobsang needs money'
- (49) *blo=bzang-la ngul-tsam yod-pa red*
Lobsang-LOC money-some exist-DISJUNCT
'Lobsang has some money'

It also marks the non-Agent argument of certain transitive predicates:

- (50) *thub=bstan-gyis blo=bzang-la gzhus-song*
Thubten-ERG Lobsang-LOC hit-PERF
'Thubten hit Lobsang.'
- (51) *thub=bstan-gyis blo=bzang(*-la) bsad-pa red*
Thubten-ERG Lobsang(*-LOC) kill-PERF
'Thubten killed Lobsang.'

This is not the “primary object” or “antidative” construction familiar from Romance, Indic, some Tibeto-Burman, and many other languages. The occurrence of *-la* is absolutely determined by the verb: surface contact verbs like *gzhus* ‘hit’ always require a *-la*-marked argument; change of state verbs like *gsod* ‘kill,’ *gcod* ‘cut,’ *bzo* ‘cook,’ etc., require an unmarked argument. There can be no alternation between the two constructions with the same verb.

6.3 Adverbial clauses and subordinators

The simplest subordinating construction consists of the bare verb stem with a clitic subordinating particle: *na* ‘if,’ *tsang* ‘because,’ *nas /nä?/* and *byas /cä?/* ‘non-final,’ among others. There are also subordinators similar in structure to relator noun constructions, which likewise follow a bare verb stem, e.g. *bar-du* /phaatu/ and *ring-la* /riŋla/ ‘while,’ *rjes-la* /ceela/ ‘after,’ and *khong-la* /khoŋla/ with negated verb ‘before’:

- (52) *kho rgya=gar-la ma-'gro khong-la*
he India-to NEG-go before-LOC
dngul mang-po bsags-pa red
money much save-PERFECTIVE.GENERIC
'Before he went to India, he saved a lot of money.'

The nominalized past stem with *-r* indicates ‘in order to’ or, with a negated verb, ‘without’:

- (53) *kho-s nga ma-bsugs-pa-r phyin-zhag*
he-ERG 1sg NEG-wait-NMLZ-LOC went-PERFECT.INFERENTIAL
'He left without waiting for me.'

6.4 Complementation

Most complement clauses are nominalizations. The nominalizers most commonly used as complementizers are *-pa*, *=yag*, *=rgyu*, and *=gag*. The complement of a verb of perception or cognition can be a finite clause nominalized with *-pa*:

- (54) *nga-s kho las=ka byed=kyi yod-pa mthong-byung*
1sg-ERG he work do=IMPF.PERSONAL-NMLZ see-PERFECT.PERSONAL_GOAL
'I see he is working.'

Otherwise complement constructions involve an uninflected verb stem plus a nominalizer. The general-purpose complementizers are *-pa*, *=yag*, and *=rgyu*.

- (55) *nga-s kho-r bshad=rgyu khas=len byas-pa yin*
1sg-ERG he-LOC tell-NMLZ promise did-PERFECTIVE.PERSONAL
'I promised to tell him.'
- (56) *kho-s las=ka byed=yag 'go btsugs-song*
he-ERG work do=NMZ start-PERFECTIVE.DIRECT
'He started to work.'
- (57) *nga bod-la phyin-pa kho-s shes=kyi*
1sg Tibet-LOC went-NMLZ he-ERG know=IMPF.IMMEDIATE
'He knows I went to Tibet.'

=yag and *=rgyu* are interchangeable in many constructions. They sometimes contrast with *-pa* in aspect, *-pa* marking a perfective, *=yag*/*=rgyu* an imperfective complement:

- (58) *'di- i skad=cha dris=rgyu / dris-pa*
this-GEN question ask=NMLZ
gus=zhabs med-pa red
polite not.be-PERFECTIVE.GENERIC
'It's not polite to ask / have asked about this.'

The *=gag* construction generally indicates some sort of purpose or intention:

- (59) *nga-s 'di-r khyed=rang-la skad=cha bshad=gag*
1sg-ERG this-LOC 2sg-LOC speech talk=NMZ
yongs-pa yin
came-PERF/CONJUNCT
'I came here to talk to you.'

Finite complements occur with verbs of speech. These may be unmarked or marked with the complementizer *ze*, usually pronounced as a semi-syllabic [s]:

- (60) *kho-s nga-s bcag-pa red zer=gyis*
he-ERG 1sg-ERG broke-PERF/DISJUNCT say=IMPF/DISJUNCT
'He says I broke [it].'
- (61) *kho-s nga-s bcag-pa red-ze zer=gyis*
'idem.'

This *ze* is a grammaticalized form of *zer* 'say,' but now freely co-occurs with it, as in the above example.

6.5 Negation and questions

The negative prefix is *ma-* in perfective and future constructions, *mi-* in imperfective forms. It is attached to the highest verbal element in the sentence:

- (62) *kha=lag za=dus skad=cha ma-shod*
food eat=TIME speech NEG-speak
'Don't talk while you're eating!'

Typically this is a tense/aspect marker rather than the main verb:

- (63) *nga gnyid 'khugs ma-byung*
1sg sleep fall NEG-PERF.PERSONAL_GOAL
'I couldn't get to sleep.'

In tense/aspect forms based on copulas, the copula carries the negation:

- (64) *nga=gnyis thug=gi ma-red*
1sg=two meet=NONPERF NEG-be
'We two won't [be able to] meet.'
- (65) *kho za=gi mi-'dug*
he eat=NONPERF NEG-exist.IMMEDIATE
'He isn't eating.'

The conjunct copulas *yod* and *yin* have special negative forms, *med /meʔ/* and *min /miŋ/*.

Polar questions are marked by the final particle *pas /-pää/*, which assimilates to a preceding velar:

- (66) *kho bod-pa red-pas*
he Tibetan be.GENERIC-INTERROGATIVE
'Is he a Tibetan?'
- (67) *khyed=rang bod-la phebs myong-ngas*
2sg Tibetan-LOC go(HON) EXPERIENTIAL-INTERR
'Have you ever been to Tibet?'

The interrogative words are *su /su/* 'who' and a series based on *ga*: *ga=re /khaře/* 'what,' *ga=pa-r /khapaa/* 'where,' *ga=dus /khatü/* 'when,' *ga='dra /khanṭa/* 'how.' Information questions have the interrogative word in its expected sentence position, and are optionally marked with the final particle *ga /kha/*:

- (68) *slob=gra tshar-na khyed=rang ga=re gnang-ga*
school finish-when 2sg what do(HON)-INTERR
'What will you do when you finish school?'

REFERENCES

- Chang, Betty Shefts, and Kun Chang (1981) 'Perfective and imperfective in Spoken Tibetan', *BIHP* 52.2:303–321.
- Chang, Kun, and Betty Shefts (1964) *A Manual of Spoken Tibetan (Lhasa Dialect)*. Seattle: University of Washington Press.
- Chang, Kun, and Betty Shefts Chang (1968) 'Vowel harmony in Spoken Lhasa Tibetan', *BIHP* 40:53–124.
- Chang, Kun, and Betty Shefts Chang (1978–81) *Spoken Tibetan Texts*. Taipei: Institute of History and Philology, Academia Sinica.
- DeLancey, Scott (1990) 'Ergativity and the cognitive model of event structure in Lhasa Tibetan', *Cognitive Linguistics* 1.3:289–321.
- DeLancey, Scott (1992) 'The historical status of the conjunct/disjunct pattern in Tibeto-Burman', *Acta Linguistica Hafniensia* 25:39–62.
- DeLancey, Scott (1999) 'Relativization in Tibetan', in Yogendra Yadava and Warren Glover (eds) *Studies in Nepalese Linguistics*, 231–249. Kathmandu: Royal Nepal Academy.

- DeLancey, Scott (2012a) ‘“Optional” “ergativity” in Tibeto-Burman languages’, *LTBA* 34.2:9–20.
- DeLancey, Scott (2012b) ‘Still mirative after all these years’, *Linguistic Typology* 16.3:529–564.
- Denwood, Philip (1999) *Tibetan*. Amsterdam: John Benjamins.
- Duanmu, San (1991) ‘An autosegmental analysis of tone in four Tibetan languages’, *LTBA* 15.1:65–91.
- Goldstein, Melvyn (1975) *Tibetan–English Dictionary of Modern Tibetan*. Kathmandu: Ratna Pustak Bhandar.
- Goldstein, Melvyn (1991) *Essentials of Modern Literary Tibetan*. Berkeley: University of California Press.
- Goldstein, Melvyn, and Nawang Nornang (1970) *Modern Spoken Tibetan: Lhasa Dialect*. Seattle: University of Washington Press.
- Goldstein, Melvyn, with Ngawangthondup Narkyid (1984) *English–Tibetan Dictionary of Modern Tibetan*. Berkeley: University of California Press.
- Hill, Nathan W. (2012) ‘“Mirativity” does not exist: *ḥdug* in “Lhasa” Tibetan and other suspects’, *Linguistic Typology* 16.3:389–433.
- Hu, Tan (1982) ‘Recherches sur les tons du tibétain (dialecte de Lhasa)’, *Cahiers de Linguistique Asie Orientale* 9.1:11–46.
- Jin, Peng (1979) ‘Lun Zangyu Lasa-kouyu dongci-de tedian yu yufa jiegou-de guanxi (On the relations between the characteristics of the verb and the syntactic structure in Spoken Tibetan (Lhasa dialect))’, *Minzu Yuwen* 1979.3:173–181.
- Jin, Peng (1983a) ‘Zangyu Lasahua dongci-de shi ji qi biaoda fangfa (The moods of Tibetan (Lhasa dialect) verbs and the ways to represent them)’, *Minzu Yuwen* 1983.1:9–18.
- Jin, Peng (1983b) *Zangyu jianzhi (A Brief Description of Tibetan)*. Beijing: Minzu Chubanshe.
- Mazaudon, Martine (1977) ‘Tibeto-Burman tonogenetics’, *LTBA* 3.2:1–123.
- Mazaudon, Martine (1984) ‘Rev. of Hari, *An Investigation of the Tones of Lhasa Tibetan*’, *LTBA* 8.1:91–99.
- Nishi, Yoshio (1986) ‘Gendai Chibetto-go hōgen-no bunrai (A classification of Tibetan dialects)’, *Bulletin of the National Museum of Ethnology* 11:837–901.
- Saxena, Anju (1991) ‘Pathways of the development of the ergative in Central Tibetan’, *LTBA* 14.1:109–116.
- Sprigg, R.K (1961) ‘Vowel harmony in Lhasa Tibetan’, *BSOAS* 24:116–138.
- Sprigg, R.K (1993) ‘Controversy in the tonal analysis of Tibetan’, *BSOAS* 56:470–501.
- Sun, Jackson T.-S. (1997) ‘Typology of tone in Tibetan’, in Tseng Chiu-yu (ed.), *Chinese Languages and Linguistics IV: Typological Studies of Languages in China; Symposium Series of the Institute of History and Philology, Academia Sinica*; Number 2, 485–521. Nankang, Taipei: Academia Sinica.
- Sun, Jackson T.-S. (2003) ‘Variegated tonal developments in Tibetan’, in David Bradley, Randy LaPolla, Boyd Michailovsky and Graham Thurgood (eds), *Language Variation: Papers on Variation and Change in the Sinosphere in Honour of James A. Matisoff*, 35–51. Canberra: Department of Linguistics, Australian National University.
- Sun, Jackson T.-S. (ed.) (2014) *Phonological Profiles of Little-Studied Tibetic Varieties (Language and Linguistics monograph series)*. Taipei: Institute of Linguistics, Academia Sinica.
- Tournadre, Nicolas (1991) ‘The rhetorical use of the Tibetan ergative’, *LTBA* 14.1: 93–107.

- Tournadre, Nicolas (1996a) *L'ergativité en tibétain: approche morphosyntaxique de la langue parlée*. Paris: Peeters.
- Tournadre, Nicolas (1996b) ‘Comparaison des systèmes médiatifs de quatre dialects tibétains’, in Zlatka Guentchéva (ed.) *L’Énonciation médiatisée*, 195–211. Paris: Peeters.
- Tournadre, Nicolas (2008) ‘Arguments against the concept of “conjunct”/“disjunct” in Tibetan’, in Brigitte Huber, Marianne Volkart and Paul Widmer (eds), *Chomolangma, Demawend und Kasbek: Festschrift für Roland Bielmeier zu seinem 65. Geburtstag*, 281–308. Halle: International Institute for Tibetan and Buddhist Studies (German for ‘Ltd’).
- Tournadre, Nicolas (2013) ‘The Tibetic languages and their classification’, in Tom Owen-Smith and Nathan Hill (eds), *Trans-Himalayan Linguistics*, 105–129. Berlin: de Gruyter.
- Tournadre, Nicolas, and Sangda Dorje (2003) *Manuel de tibétain standard: langue et civilisation*. Paris: Langues et Mondes, L’asiathèque.
- Tournadre, Nicolas, and Sangda Dorje (2005) *Manual of Standard Tibetan: Language and Civilization* (trans. of Tournadre and Dorje 2003). Boston: Shambhala.
- Wang, Huiyin (1988) ‘Zangyu Lasahua dongci de chongdie xingshi (Reduplicated forms of verbs in Lhasa Tibetan)’, *Minzu Yuwen* 1988.3:36–43.
- Zeisler, Bettina (2004) *Relative Tense and Aspectual Values in Tibetan Languages*. Berlin: Mouton de Gruyter.

CHAPTER TWENTY

KURTÖP

Gwendolyn Hyslop

1 INTRODUCTION

Kurtöp is a Tibeto-Burman language spoken by approximately 15,000 people. Within Tibeto-Burman, Kurtöp belongs to the East Bodish branch, a group of seven languages or dialect groups. The relationship of the East Bodish languages within Tibeto-Burman remains subject to debate (e.g. Hyslop 2013a, 2014a) though the general consensus is that the East Bodish languages are not Tibetan languages but are likely close relatives. This chapter presents an overview of Kurtöp grammar, focusing on the following topics: phonology, case-marking, verbal affixes, nominalization, clause-chaining and serialization, and epistemic modality. For a full description of Kurtöp, refer to Hyslop (2017).

2 BACKGROUND

Kurtöp is spoken in Lhüntse district in Northeastern Bhutan. In addition to Tangmachu (where Chocangaca is also spoken), Kurtöp is spoken between the village of Gorgan, south of Lhüntse, and up until Naling, near the border with Tibet. Some of the villages included in this area are Gangzur, Shawa, Zhamling, Dungkar, Tabi, Jasabi, Tünpe, Cakzom, and Nê, among several other villages. In total, there are probably around 15,000 speakers of Kurtöp.

The Kurtöp language community spans the approximate geographic coordinates of $27^{\circ} 35' 38.90''$ N on the southern end of the region to $27^{\circ} 53' 01.50''$ N on the northern edge, just south of the border with Tibet. The lowest point of the river valley in the area is around 3,800 feet and the highest mountain peak in the region is approximately 16,500 feet; however, the villages are located at elevations ranging from approximately 4,000 to 8,500 feet. The approximate location of the Kurtöp homeland is shown in Map 20.1.

2.1 Name of language

Scholars in Bhutan have offered a few potential hypotheses regarding the origin of the name *kur* or *kuri*. Drâsho Sangay Dorji (personal communication)¹ observes that the name of the glacier from which the *Kur* river originates is *Kulagangri* and thus states it is possible that *Kur* may have a shared history with *Kula*. However, Dorji also points out that most rivers in Bhutan are named after the valley, rather than the source. Geshe Tenzin also reports that he saw <*skuristod*> and <*skurismad*> used as place names in a few Classical Tibetan texts, without any other reference to meaning.

Kurtöp is the word used in Dzongkha to refer to people from the Kurtö region; this term is not used in Bhutan to refer to a language, but only to people. The terms *Kurtöpkha* or *Kurtöbikha* are Dzongkha words which refer to the language spoken by *Kurtöps*, or people from the Kurtö region in Bhutan. Within the Kurtö region in Bhutan, a number of languages are actually spoken. In the village of Khoma and surrounding area, the *Kurtöps*



MAP 20.1 APPROXIMATE LOCATION OF THE KURTÖP HOMELAND

speak a language known as Dzala in the Western literature. South of Lhüntse town, *Kurtöps* speak another language, closer in relation to Dzongkha, often called *Chocangackha* or '*Matpaikha*'. In Lhüntse itself, and the villages directly north of Lhüntse along the Kurichu up to the border of Tibet, and by some people in the village of Tangmachu *Zháke Kurtöp*, the focus of this chapter, is spoken.

Because *Kurtöp* literally means 'one from Kurtö' in Dzongkha, and several distinct languages are actually spoken in the Kurtö region, the term *Kurtöp language* in Bhutan is actually ambiguous. In order to specify which particular *Kurtöp* language is at question, speakers often identify *Zháke Kurtöp* by the way they ask 'where are you going?', which is *au gemale?* The use of the translated question 'where are you going' is a common way to identify other languages in Bhutan as well. For example, Dzala is often referred to as *i ga brok*, which is how one asks 'where are you going?' in the language. Following Western linguistics literature, I retain the name *Kurtöp* for the language discussed here.

2.2 Ethnographic practices

Speakers of *Kurtöp* are by and large subsistence farmers, with each family raising cows for dairy and cultivating grains and vegetables for individual use. Dairy from cows is used primarily for making cheese and butter; their byproducts whey and buttermilk are also consumed.

Rice has become a staple grain only in recent years; previously maize, ground into small pieces called *kharang*, was probably the most important grain, but wheat (*go*), sweet buckwheat (*cara*), millet (both finger millet, *Eleusine coracana*—called *thre* locally, and foxtail millet, *Setaria italica*—called *ran* locally), and bitter buckwheat (*brama*) have also been in use. Historically, taro (*byo*) and other tubers were also used, though these days people have the means, and indeed prefer, to cultivate the grains described above instead.

Common crops are potatoes (*ki*), which grow particularly well in the village of Shawa, green onion (*tsong*), beans of several varieties (*shepen*), eggplant (*dolom*), daikon radish (*muya*), squash (*laushar*) and, more recently, cabbage (*banda kopi* < Hindi), cauliflower (*meto kopi* < *Kurtöp* + Hindi). Chiles (*banggala*) are a very important crop and they are

used in several different ways. In addition to being used in their fresh, green state, chiles may be dried in the sun after turning red (*banggala kam*), dried green ('*ngokam*) or boiled and then dried, becoming white.

Several foods are found growing in the wild, including a wide variety of mushrooms (*mu*), and fiddlehead (*zhiwa*). Fruits are citron (*kapula*), banana (*cela ~ ceya, ngala* in some dialects), guava ('*andre*), fig (*khongdi*), orange (*tshalu*), a sweet tomato that grows on trees ('*lambenda*), and various berries (*mrip*). Spices are cilantro (*wesi*), ginger (*saga*), garlic (*chacu*), and salt (*tsha*), which is used generously in nearly all cooking. People rarely drink water, but tea (*ja*) in a variety of forms (e.g. *suja* ‘butter tea’ *'ngaja* ‘sweet tea’), whey (*shurkhu*) and buttermilk (*tarwa*) are common. Kurtöps also make alcohol out of a variety of grains, such as rice, corn, and millet. Distilled alcohol is called *zhor*. Potatoes are a common cash crop.

The official religion of the Kurtöps, like mainstream Bhutan, is Buddhism. However, also like mainstream Bhutan, many of the Buddhist practices are actually interlaced with various pre-Buddhist practices. Within the Kurtöp speech community this pre-Buddhist religion primarily took the form of Bon (see Huber 2013, 2014 for more details about the ‘East Bodish’ Bon religion).

As elsewhere in Bhutan, marriage among the Kurtöps involves little ceremonious ritual. People become ‘married’ once they’ve moved in together. In Dungkar *geok* both polyandry and polygamy are practiced. We are aware of one instance of each. In the case of the former one woman has three husbands who are brothers and in the case of the latter the husband’s two wives are sisters. In this case the husband is a former monk, a status which affords him social prestige and added financial gain. Villagers report that the norm is for men to move in with their wives’ families but it was not possible to obtain much data about this in practice. In the house which served as the base for the researchers in Tabi village, it was the case that it was the woman’s family home and the husband had moved in.

2.3 Language endangerment and sociolinguistic situation

A growing trend at present is for younger villagers to leave the area for education and eventually find themselves settled in the capital, marrying people from other regions in Bhutan, and speaking Dzongkha to their children. For example, in the family which housed the researchers, there were five children but none of them resided in the village, despite the fact that the youngest was only approximately ten years of age. The children were either living/studying in Thimphu or studying elsewhere.

Even in a community as small as the Kurtöp-speaking community, there are different registers of speech, depending primarily on education, time spent in the village, and exposure to Dzongkha. The highest register of Kurtöp involves a high level of Classical Tibetan and Dzongkha borrowings and is characterized by the use of the honorific particle *la* and honorific vocabulary. Interestingly, the pronunciation of these words varies drastically, depending on education and experience of the speaker. For example, front-rounded vowels are only found in the speech of the most educated speakers, or those who have grown up in Thimphu. As in Dzongkha, Hindi borrowings are also characteristic of the ‘cool’ speech of the younger generation, though some words, such as *thrika* ‘okay; good,’ appear to have filtered down to all registers.

In addition to these different registers, most Kurtöp speakers are also highly multilingual. Without exception, every Kurtöp speaker we met spoke at least one other language (Dzongkha or Chocangaca, minimally) in addition to Kurtöp. More common was for

speakers to also be familiar with at least one or more of the following languages: Tshangla, Dzala, Bumthap, Khengkha, Tibetan, Nepali, Hindi, English.

One result of these different forces—tendency toward migration out of the village and exogamous marriages plus rampant multilingualism—is that the language is endangered. Children outside of the village are, by and large, no longer learning the language. Even among those who are learning, the language change—influence from neighboring languages—is happening at such a fast rate that there are several obvious generational differences.

3 PHONOLOGY

Kurtöp phonology has been described in depth in several publications, first in Michailovsky and Mazaudon (1994) and then later in Hyslop (2006, 2008, 2009). The current discussion on phonology summarizes these previous findings and includes an updated analysis. The data presented in this section are in the practical orthography (Hyslop 2014b).

3.1 Consonant phonemes

Kurtöp contrasts 15 stops, three fricatives, two affricates, two laterals, one rhotic, four nasals, two glides and a glottal aspirate, as shown in Table 20.1.² Only some of these may serve as syllable codas; these are summarized in Figure 20.1. A subset of the Kurtöp consonant phonemes may be combined to make complex onsets; these are illustrated in Figure 20.2. See Hyslop (2017) for more details about the phonetics, synchronic distribution, and historical development of these sounds.

3.2 Vowels

All varieties of Kurtöp have five vowels: /i, e, ə, o, u/ but educated speakers also produce the front-rounded vowels /y/ and /ø/, under influence from Dzongkha and Tibetan.³ Kurtöp has four diphthongs: /iu, ui, oi, eu/, as shown in Table 20.2.

3.3 Suprasegmental contrasts

Kurtöp minimally contrasts vowel length and tone. Vowels may be long or short in open syllables. The contrast is not terribly robust; the difference between long and short is usually less than 50ms and there is not an abundance of minimal pairs. Nonetheless, a few can be found, as shown in Table 20.3.

TABLE 20.1 KURTÖP CONSONANT PHONEMES, REPRESENTED IN LOCAL ORTHOGRAPHY

	Labial	Dental	Retroflex	Palatal	Velar	Glottal	
stops	p, ph, b	t, th, d	tr, thr, dr	c, ch, j	k, kh, g	(?)	
affricates		ts, tsh					
fricatives		s, z		sh		h	
nasals	m	n		ny	ng		
laterals		l, lh					
rhotics		r					
glides	w			y			

-p	-t	-k
(-s)		
-m	-n	-ng
-r		(-l)

FIGURE 20.1 KURTÖP CODA CONSONANTS

Note: Coda *-s* is found word-internally in all varieties but occurs word-finally in only the southern dialects. Coda *-l* is only found in borrowings or due to synchronic loss of final vowels in some *l*-initial suffixes. See Hyslop (2017) for more details.

pr-	py-	pl-
p ^h r-	p ^h y-	
br-	by-	bl-
mr-	my-	
kw-	k ^h w-	gw-

FIGURE 20.2 KURTÖP COMPLEX ONSETS

Note: Kurtöp appears to be in the process of simplifying complex onsets and as such some of these are only now found in a few words for a few speakers. For example, for many speakers *bl-* and *br-* have merged and the labial-glide series may be pronounced as palatal stops. See Hyslop (2017) for more details.

TABLE 20.2 KURTÖP VOWEL PHONEMES

	Front	Back	Diphthong
high	i <i>	u <u>	iu <iu>
	(y <ü>)		ui <ui>
mid	e <e>	o <o>	oi <oi>
	(ø <ö>)		
low		e <a>	eu <au>

TABLE 20.3 CONTRASTIVE VOWEL LENGTH IN KURTÖP

Kurtöp	Gloss
lha	‘deities’
lhâ	‘excess’
she	‘over.pour’
shê	‘glass’
tsi	‘sticky’
tsî	‘calculation’
ko	‘door’
kô	‘hoe’
mu	‘mushroom’
mû	‘COP.EXIS.NEG’

TABLE 20.4 CONTRASTIVE (HIGH/LOW) TONE FOLLOWING SONORANT CONSONANTS OF INITIAL SYLLABLES AND PALATAL FRICATIVES

Phoneme	High tone	Gloss	Low tone	Gloss
/m/	'mang	'community; crowd; everyone'	mang	'be.excessive'
/n/	'nam	' <i>Perilla frutescens</i> '	nam	'sky; weather'
/ɲ/	'nyu	'be.crazy'	nyu	'borrow'
/ŋ/	'ngap	'dry.out'	ngap	'be.thin'
/ɾ/	'rung	'make.stand; get up'	rung	'small.storage.basket'
/l/	'lem	'flat.spoon'	lem	'be.delicious'
/w/	'wang	'blessing'	wang	'pit'
/j/	'yap	'awning'	yap	'wear.on.shoulders'
/ç/	sham	'shoes'	zham	'man's.length.measurement'

Tone in Kurtöp is contrastive on initial syllables following sonorant consonants and the palatal fricative, as evidenced by the minimal pairs in Table 20.4.

Following obstruents, tone is predictably high if the obstruent is voiceless and predictably low if the obstruent is voiced. This synchronic picture can be understood fairly clearly in the diachronic process of tonogenesis, as outlined in Hyslop (2009).

3.4 The syllable

The maximum syllable in Kurtöp is CCVC, though the majority of syllables are CV or CVC. The number of possible complex onsets is limited (see Figure 20.2) as are the possible codas (see Figure 20.1). There is a tendency toward simplification of syllable structure in the language, with complex onsets merging (for example *bl-* and *br-* are *br-* for many speakers) or reducing (for example, *py-* is *c-* for many speakers).

3.5 The word

Phonological words in Kurtöp are fairly easily identified, primarily by the obligatory presence of stress and tone. Each word must contain one stressed syllable and one tone. Stress is always root-initial and tone is always word-initial. The only instances in which these don't match up are when the negative prefix is used; in these cases the root is the second syllable. Tone is moved to the negative prefix (but determined by the root) and stress stays on the root. The primary correlate of stress in Kurtöp appears to be length. Another feature of phonological words is the tendency to delete post-*l* and word-final vowels. There is a strong tendency for mono- and disyllables in the language but words can be three or four syllables long as well.

4 MORPHOSYNTAX

Kurtöp morphosyntax allows us to distinguish several distinct word classes, including nouns, verbs, adjectives, numerals, and adverbs, in addition to other minor categories. There are no classifiers in the language. Finite clauses consist minimally of a finite verb, verbal construction (nominalized verb plus copula or non-final marked verb as part of a clause-chain), or copula; overt arguments are optional. The Kurtöp verb is made up of a verb stem (almost always monosyllabic) plus one optional verbal suffix and/or one or two optional enclitics. Negation is done by way of the prefix *ma*.⁴

- (1) *ma-gewalarisa*
me-ge-pala=ri=sa
 NEG-GO-PFV=REP=CEXP
 ‘I heard he didn’t go (counter to expectation).’

Finite verbal suffixes encode future tense, perfective aspect and imperfective aspect, each of which also encodes an epistemic value (see §5). There are also several non-finite verbal suffixes which perform a range of functions, including subordination of various types and nominalization. There are three verbal enclitics; *=mi* is an inclusive tag, *=sa* marks counter expectation, and *=ri* is an evidential marker encoding oral source of knowledge. Transitivity is not overtly coded and the number of arguments a verb can take is not always obvious. Serial verb constructions (see §4.3) are the only way to adjust the valency of a verb. Verbal arguments are not encoded directly on the verb. Clauses have SV/AOV word order and like other South Asian languages Kurtöp makes ample use of light verb constructions (for example, *friend do* = ‘help,’ *swimming do* = ‘swim,’ *eye do* = ‘see,’ etc.). Constituent order in the NP is usually: (Demonstrative) (Genitive Phrase) Noun (Adjective) (Numeral).

Some aspects of Kurtöp grammar which are particularly interesting in a typological or comparative perspective are described in greater detail later. Case marking is discussed in §4.1, nominalization in §4.2, and clause-chaining/serialization in §4.3. For additional details in Kurtöp grammar, see Hyslop (2017).

4.1 Case marking

Case marking in Kurtöp is handled by means of phrasal enclitics. These are shown in Table 20.5. The difference between the two locative markers is subtle and remains a matter of ongoing research. They have an almost entirely overlapping distribution, with a few exceptions, such as only the forms with the back rounded vowel occurring in directional adverbs.

The Kurtöp case system is best described as ‘pragmatic ergative’ and illustrated in greater detail in Hyslop (2010). A simplified description of the system is as follows. Verbs in Kurtöp are either monovalent (taking maximally one overt core argument) or bivalent (taking maximally two overt core arguments). Monovalent verbs can be divided into two categories: those which tend to take an absolutive-marked argument and those which readily allow for an ergative-marked argument in order to encode various semantic or pragmatic values. Bivalent verbs can also be divided into two categories: those whose A argument tends to be marked as ergative, or those whose A argument tends to be absolute but may be ergative for pragmatic or semantic reasons. In the instances in which both overt arguments of a bivalent verb are absolute, word order will determine the difference between A and O; A appears first (or postponed to the end of the clause) while the O argument is adjacent to the verb.

TABLE 20.5 KURTÖP CASE MARKERS

Gloss	Kurtöp
LOC	<i>=na~nang</i>
LOC	<i>=to~ro~ko~ngo</i>
ABL	<i>=ni~ning</i>
GEN	<i>=gi~i~li</i>
ERG	<i>=gi~i~li</i>

4.2 Nominalization

As elsewhere in Tibeto-Burman, nominalization is an important aspect of both the synchronic Kurtöp grammar and diachronic development of that grammar. There are two nominalizers (*-sa* ‘NMZ:LOC’ and *-thang* ‘NMZ:MNR’) which apply directly to the verb and four which can be used on the clausal level and which differ in terms of aspectual value given to the clause. Given the tendency for Tibeto-Burman languages to innovate main clause grammar via clausal nominalizations, this is perhaps not surprising. The four clausal nominalizers (*-pala* ‘NMZ:PFV’, *-khan* ‘NMZ:IPFV’, *-male* ‘NMZ:IRR’, *-sang*) are discussed in turn below. Two of these (*-pala*, *-male*,) also occur as finite suffixes.

The nominalizer *-pala* can relativize A, S, and O arguments; an example of a nominalized monovalent verb is shown in (2).⁵

- (2) *zikorna byonpala gapo*
zikor=na byon-pala gapo
tour=LOC go.HON-NMZ:PFV PL.FOC
'Those who went on tour and all'

Examples (3) and (4) show *-pala* integrating into the finite verbal system. In (3) the nominalized verb *zatpala* is followed by the copula *wenta* but the reading of the sentence is completely finite. We see that in (4) if the copula is removed the reading of the sentence is essentially the same. Note that in (2–4) *-pala* is also associated with perfective aspect.

- (3) *khitya Kurtötpa jur zatpala wen*
khit=ya Kurtöt-pa jur zat-pala wen
3.ABS=also Kurtöt-DEN become finish-NMZ:PFV COP.EQ
'He had also turned into a Kurtöp'.
- (4) *khitya Kurtötpa jur zatpala*
khit=ya jur zat-pala
3.ABS-also become finish-PFV
'He had also turned into a Kurtöp'.

The nominalizer *-male* ‘NMZ:IRR’ is like *-pala* in that it can relativize an A, S, or O argument and is also integrating into the finite verbal system. As a relativizer, *-male* brings with it irrealis aspect but as a finite suffix encodes future tense. Example (5) shows *-male* relativizing an S argument. Examples (6) and (7) show *-male* primarily functioning as a tense marker; in (6) it co-occurs with a copula while in (7) it is fully finite. Again, it is not clear what—if any—functional difference exists between (6) and (7).

- (5) *phya ngaksi nimale khepo*
phya ngaksi ni-male khepo
Bon.festival QUOT stay-NMZ:IRR FOC
'the one that remains being called *phya*. . . '
- (6) *lojuthe zhuhale wen*
loju=the zhu-male wen
talk=INDEF say.HON-NMZ:FUT COP.EQ
'(I) will give a talk.'

- (7) *lojuthe zhumble*
loju=the zhu-male
 talk=INDEF say.HON-FUT
 '(I) will give a talk.'

The nominalizers *-khan* and *-sang* also perform relativizing functions as well as tense/aspect functions. However, unlike *-pala* and *-male*, they must be followed by a copula in these contexts. Nominalizer *-khan* also marks imperfective aspect. An example of *-khan* relativizing an A argument is shown in (8). Example (9) shows *-khan* co-occurring with a copula in a finite clause. Note, however, that removal of the copula results in an ungrammatical utterance.

- (8) *sungchot 'nangkhan khepo...*
sungchot 'nang-khan khepo
 ceremony.HON give.HON-NMZ:IPFV FOC
 '(those) who offer the religious ceremony...'
- (9) *bosaya gekhan wen*
bosa=ya ge-khan wen
 orphan=also go-NMZ:IPFV COP.EQ
 'The orphan is also going.'
 **bosaya gekhan* (with intended meaning as finite clause)

The nominalizer *-sang* can relativize an oblique argument and also functions with a copula to encode future tense. Example (10) shows *-sang* relativizing an oblique argument, though note the interpretation is more of an instrument than an agent. In (11) we see *-sang* co-occurring with a copula to encode future tense. Again, the interpretation is finite.

- (10) *'ipa zusang*
'ipa zu-sang
 food eat-NMZ:OBL
 '(thing) for eating (i.e. "plate")'
- (11) *'langpochegi yung rasang wen*
'langpoche=gi yung ra-sang wen
 elephant=ERG get come-NMZ:FUT COP.EQ
 'The elephant will come to get (you').
 **'langpochegi yung rasang* (with intended meaning as finite clause)

In summary, nominalization is a highly productive process in Kurtöp, serving the function of relativization but also, when used in conjunction with a copula, plays a role in the tense/aspect system. As one might suspect, Kurtöp also has a large number of copulas (into the 20s, at least) which encode a wide range of finite values. (These are briefly discussed in §5.)

4.3 Clause-chaining and verb serialization

As in many Asian languages, clause-chaining and verb serialization are also important aspects of Kurtöp grammar. The clause-chaining construction is among the most productive syntactic constructions in the language. Kurtöp finite sentences are often very long and complex, consisting of several chained clauses within one finite sentence. Prototypically, the Kurtöp clause-chaining construction consists of at least two verbs. The first verb is a converb suffixed with the converb marker *-si*⁶ and the second verb will be finite or part of a finite construction (such as a nominalization). An example of the Kurtöp clause-chaining construction is shown in (12) below, with the converb in bold font.

-
- (12) *depa kutsi tsawai 'lama ngak tanpal wenta*
- | | | | | | | |
|----------|---------|----------|-------|------|----------------|------------|
| depa | kut-si | tsawa=gi | 'lama | ngak | tan-pala | wenta |
| devotion | send-NF | root=GEN | DM | do | adhere-NMZ:PFV | COP.EQ.MIR |
- 'Being very devoted, (she) made (him her) root lama'.

Arguments may or may not be shared across the verbs and there may or may not be other interceding material, such as noun phrases or adverbs, between the clauses. The non-final marked converb is always unmarked for tense/aspect/evidentiality; this information is provided by the finite verb.

If the final verb in the chain is a copula or auxiliary there are different formal and functional consequences. When the final verb is a copula the interpretation of the chain can be monoclausal, encoding durative aspect. There is no interceding material allowed between the non-final marked converb and copula. An example is (13). Note that the construction is potentially ambiguous; although durative aspect is the preferred reading, some speakers can still interpret the construction as a clause chain encoding two separate events.

- (13) *chorten kora thungzi nawala*
- | | | | |
|---------|----------------|----------|----------|
| chorten | kora | thung-si | nawala |
| stupa | circumambulate | do-NF | COP.EXIS |
- '(S/he) was circumambulating the stupa.'
- '(S/he) circumambulated the stupa and stayed there'.

When the final verb is an auxiliary, there is a tendency for no material to occur between the verbs and further for speakers to drop the non-final marker, making two directly adjacent verbs: the first a bare stem and the second a finite-marked verb. The interpretation is of one event, with the semantics of the event being provided by the lexical verb and added grammatical value being provided by the auxiliary. In other words, in these contexts Kurtöp can be said to have a serial verb construction. An example is (14), in which the auxiliary *blek* 'keep' seems to encode habitual aspect.

- (14) *phini woso tok blekta*
- | | | | |
|------------|-----------|------|---------------|
| phi-ni | woso | tok | blek-ta |
| flour-CFOC | like.this | beat | keep-IPFV.MIR |
- 'Flour is ground like this'.

A list of Kurtöp auxiliaries and their functions as auxiliaries in a serial verb construction is summarized in Table 20.6. Further details about clause-chaining in Kurtöp and its diachronic trajectory can be found in Hyslop (2013b).

TABLE 20.6 FUNCTIONS OF AUXILIARY IN SERIAL VERB CONSTRUCTIONS

	Argument Structure	Direction	Aspect	Unknown
<i>bi</i> 'give'	✓			
<i>ni</i> 'stay'			✓	
<i>blek</i> 'keep'	✓		✓ (?)	
<i>ra</i> 'come'		✓		
<i>zat</i> 'finish'			✓	
<i>thung</i> 'do'				✓
<i>ge</i> 'go'		✓	?	?
<i>zon</i> 'send'				(lexical?)
<i>ngak</i> 'do'				?

5 EVIDENTIALITY AND RELATED CATEGORIES

Kurtöp has an incredibly rich system of evidentiality, mirativity, and related categories. The system makes use of finite verbal suffixes, copulas, clitics and copulas in order to create an elaborate set of contrasts. This system of evidentiality, mirativity, egophoricity, and epistemicity is part and parcel of the verbal system. These categories are not optional; rather they are required in any finite clause, including questions. The discussion here is a brief summary of Hyslop (2014c), with some updates as reported in Hyslop (2014d). See also Hyslop (forthcoming) for relevant discussion.

Kurtöp makes a five-way contrast in perfective aspect, as summarized in Figure 20.3. For contexts in which speakers are unsure of their knowledge, for example if they are making an assumption, they will use the form *-para*. There are four separate forms used when the speaker is certain of their knowledge. If they gained it through indirect evidence or inference, the form *-mu* is used. If the speaker has direct evidence but was not previously aware of the knowledge—in other words in mirative contexts—the form *-na* is used. Finally, two forms, *-shang* and *-pala*, are both used for certain, direct, non-mirative contexts. The difference between the two is nuanced and the focus of other studies (e.g. Hyslop 2014d). It will suffice for the present chapter to state that for contexts in which the speaker has exclusive access to knowledge—that is, egophoric contexts—the form *-shang* will be used, while in non-exclusive contexts the form *-pala* will be used.

Kurtöp marks a two-way contrast in imperfective aspect between clauses which are mirative, or unexpected to the speaker (or actor), and those that are not unexpected, as illustrated in Figure 20.4.

As mentioned briefly in §4.3, Kurtöp has a rich set of copulas. Abstract bases can be identified for affirmative and negative existential and affirmative and negative equational copulas. Within those sets, several forms are found, encoding a wide range of evidential, mirative, and epistemic contrasts. The forms and their functions are summarized in Tables 20.7 and 20.8. For more details, including examples of use, see Hyslop (2014c).

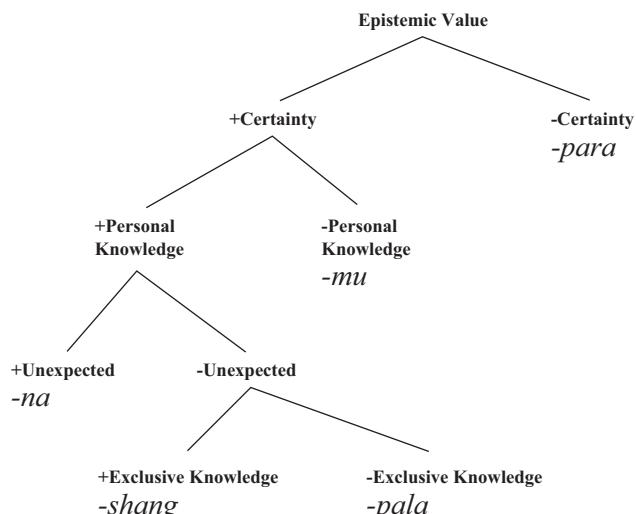


FIGURE 20.3 CONTRASTS MADE IN PERFECTIVE ASPECT (HYSLOP 2014D: 204)

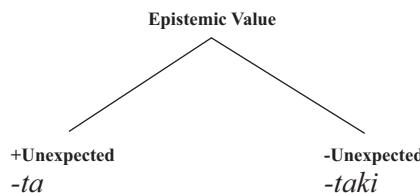


FIGURE 20.4 KURTÖP IMPERFECTIVE ASPECT SUFFIXES (HYSLOP 2011)

TABLE 20.7 EXISTENTIAL COPULAS

Form	Source		Expectation	Certainty
	Direct	Indirect		
<i>nawala</i>	+			
<i>mû</i>	+			
<i>nâ</i>			-	
<i>mutna</i>			-	
<i>nawara</i>				-
<i>mutpara</i>				-
<i>mutle</i>		+		
<i>mutlera</i>		+		-

Note: Affirmative forms begin with *na-* while negative forms begin with *mu-*. Here and in Table 20.8, a greyed cell indicates that the particular contrast is not relevant, while ‘+’ indicates a form positively encodes a given contrast while ‘-’ indicates a given form negatively encodes a given contrast.

TABLE 20.8 EQUATIONAL COPULAS

Form	Source		Expectation	Certainty
	Direct	Indirect		
<i>wen</i>	+			
<i>min</i>	+			
<i>wenta</i>			-	
<i>minta</i>			-	
<i>wenpara</i>				-
<i>minpara</i>				-
<i>minle</i>		+		
<i>minlera</i>		+		-
<i>weni</i>			- (?)	- (?)
<i>wenim</i>			- (?)	- (?)

Note: Affirmative forms begin with *we-* while negative forms begin with *mi-*. A ‘?’ indicates that it is not as yet entirely clear what the core contrast is.

NOTES

- 1 I am grateful to Namgay Thinley for discussing the etymology of Kurtö with scholars in Bhutan on my behalf.
- 2 The use of parentheses with the glottal stop indicates this segment has not been found to be phonemically contrastive; it only precedes vocalic high-toned initials and occurs occasionally in place of coda /k/.

- 3 Speakers who are fluent in Dzongkha and/or well-studied with a monastic education tend to produce front-rounded vowels when preceding a coronal consonant. They also tend to produce the diphthong /ui/ as [y] and /oi/ as [ø].
- 4 The form of the negative differs according to several factors. A difference in tense is denoted by the use of *ma-* versus *me-* ~ *mi-*; *ma-* denotes past tense while *me-* ~ *mi-* denotes non-past. In non-past, *me-* is used when the vowel of the stem is non-high and *mi-* is used when the vowel of the stem is high; that is, the negative prefix exhibits assimilation of height. Both past and non-past negatives also agree with the tone of the verbal stem. Verb stems obligatorily have either high or low tone (contrastive following sonorants but predictable following obstruents—voiceless conditions high tone while voiced conditions low tone; see §3.3. This tone spreads to the prefix, so that a negative prefix has high tone if the verbal stem has a high tone, while the negative prefix has a low tone if the verbal stem has a low tone.
- 5 Relativized clauses tend to end with a focus particle—*khepo* if the argument is semantically singular or *gapo* if the argument is semantically plural. However, it is not a requirement of a relativized clause to do so and both *khepo* and *gapo* readily occur in all types on NPs; as such, it would be a mistake to attribute a relativizing function to *khepo* or *gapo*.
- 6 The non-final suffix *-si* has allomorph *-zi* following vowels (i.e. open stems), and voiced consonants (i.e. *-m*, *-n*, *-r*, *-ng*).

REFERENCES

- Huber, Toni. 2013. The iconography of gshen priests in the ethnographic context of the extended Eastern Himalayas, and reflections on the development of Bon Religion. In *Nepalica-Tibetica: Festgabe für Christoph Cüppers, Band 1*, edited by Franz-Karl Ehrhard and Petra Maurer, 263–294. Andiast: International Institute for Tibetan and Buddhist Studies.
- Huber, Toni. 2014. Hunting for the cure: a Bon healing narrative from Eastern Bhutan. In *Tibetan and Himalayan Healing: An Anthology for Anthony Aris*, edited by Charles Ramble and Ulrike Roesler, 371–382. Kathmandu: Vajra Books.
- Hyslop [Lowes], Gwendolyn. 2006. Kurtöp phonetics and phonology. University of Oregon MA Thesis. Eugene, OR.
- Hyslop, Gwendolyn. 2008. Kurtöp phonology in the context of North East India. In *North East Indian Linguistics*, edited by Stephen Morey and Mark W. Post, 3–25. New Delhi: Cambridge University Press.
- Hyslop, Gwendolyn. 2009. Kurtöp tone: a tonogenetic case study. *Lingua* 119: 827–845.
- Hyslop, Gwendolyn. 2010. Kurtöp case: the pragmatic ergative and beyond. *Linguistics of the Tibeto-Burman Area* 33(1): 1–40.
- Hyslop, Gwendolyn. 2011. Mirativity in Kurtöp. *Journal of South Asian Linguistics* 4: 43–60.
- Hyslop, Gwendolyn. 2013a. On the internal phylogeny of the East Bodish languages. In *North East Indian Linguistics* 5, edited by Gwendolyn Hyslop, Stephen Morey and Mark W. Post, 91–109. New Delhi: Cambridge University Press/Foundation.
- Hyslop, Gwendolyn. 2013b. The Kurtöp clause-chaining construction: conversbs, clause chains, and verb serialization. In *Functional-Historical Approaches to Explanation: In Honor of Scott DeLancey*, edited by Timothy Thornes, Erik Andvik, Gwendolyn Hyslop and Joana Jansen, 155–178. *Typological Studies in Language* 103. Amsterdam: John Benjamins.
- Hyslop, Gwendolyn. 2014a. A preliminary reconstruction of East Bodish. In *Trans-Himalayan Linguistics*, edited by Nathan Hill and Thomas Owen-Smith, 155–179. Berlin: Mouton de Gruyter.

- Hyslop, Gwendolyn. 2014b. Kurtöp orthography development in Bhutan. In *Developing Orthographies for Unwritten Languages*, edited by Michael Cahill and Keren Rice, 211–230. Dallas, TX: SIL International.
- Hyslop, Gwendolyn. 2014c. Grammar of knowledge in Kurtöp: evidentiality, mirativity, and expectation of knowledge. In *The Grammar of Knowledge*, edited by Alexandra Y. Aikhenvald and R. M. W. Dixon, 108–131. Oxford: Oxford University Press.
- Hyslop, Gwendolyn. 2014d. On the category of speaker expectation of interlocutor knowledge in Kurtöp. In *Proceedings from the 40th Annual Meeting of the Berkeley Linguistics Society*, edited by Herman Leung, Zachary O'Hagan, Sarah Bakst, Auburn Lutzross and Jonathan Manker, 201–214. Berkeley: University of California.
- Hyslop, Gwendolyn. 2017. *A Grammar of Kurtöp*. Languages of the Greater Himalayan Region, 18. Leiden: Brill.
- Hyslop, Gwendolyn. Forthcoming. Mirativity and egophoricity in Kurtöp. In *Egophoricity*, edited by Elisabeth Norcliffe, Lila San Roque and Simeon Floyd. Amsterdam: John Benjamins.
- Michailovsky, Boyd, and Martine Mazaudon. 1994. Preliminary notes on languages of the Bumthang group. In *Tibetan Studies: Proceedings of the 6th Seminar of the International Association for Tibetan Studies*, edited by Per Kværne, 2:545–557. Fagernes: Institute for Comparative Research in Human Culture.

CHAPTER TWENTY-ONE

TSHANGLA

Erik Andvik

1 INTRODUCTION

Tshangla is a Tibeto-Burman language of the Bodish subgroup, close to but just outside of the inner nucleus of Tibetic languages. While Tshangla has been referred to as *Monpa*, it is distinct from both *Northern Monpa* spoken in Tawang of Arunachal Pradesh, and *Cuona (nTsho-sna) Monpa* of Tibet.

Tshangla is the mother tongue of between 150,000 and 200,000 people in eastern Bhutan. Speakers are also known as *Sharchop*, or *Sharchokpa*, ‘Easterners.’ Their settlements range from the subtropical plains near the border with India upwards into the Himalayan foothills to roughly 3,000 meters above sea level. They live primarily on subsistence agriculture and small holdings of livestock. Bilingualism in Nepali is high among Tshangla speakers living in the border areas or near to major towns, and English and Dzongkha are taught in the primary schools. However, use of Tshangla is vigorous, and intergenerational transmission of the language remains robust. Tshangla serves, in fact, as a language of wider communication throughout much of the country, most Bhutanese having at least some rudimentary knowledge of Tshangla (van Driem 1998: 27–9). The regional variety described in this chapter is spoken in the Trashigang District in eastern Bhutan.

Tshangla is also spoken in India by perhaps 5,000–10,000 people just across the border from eastern Bhutan in Khalagtang in Arunachal Pradesh, also in and around Dirang in the Kameng Region, where it has been called Central Monpa (Das Gupta 1968).

Tshangla is also spoken in Tibet, in a cluster of communities geographically separated from Bhutan by several hundred miles, in the region formerly known as Padma-bKod (Pemakö), located just north of the point where the Tsangpo River (Siang), crosses the McMahon Line separating south-eastern Tibet from the Indian state of Arunachal Pradesh. Tshangla spoken in this region has been called *Cangluo Monpa* (Zhang 1986) and earlier *Motuo Monba* (Sun *et al.* 1980). A population here of perhaps 5,000 speakers may extend onto the Arunachal Pradesh side as well. The geographical separation of the Bhutan and Padma-bKod groups is apparently due to migration from Bhutan during the nineteenth century (Aris 1980: 9).

2 PHONOLOGY

Table 21.1 shows all syllable-initial consonant phonemes. Alternative orthographic representations used in the present chapter are given in curly brackets. Asterisked phonemes occur syllable-finally as well. Items in parentheses are not native but occur in the large number of lexemes borrowed from Dzongkha or Tibetan. The non-coronal aspirated stops (*/p^h/*, and */k^h/*) are realized as fricatives ([ɸ] and [χ] or [h]) intervocally.

There are five monophthong vowels in Tshangla, as shown in Table 21.2.

Possible syllable structures are *v*, *c(r)v*, *c(r)vc*, and *c(r)vv*. Possible monomorphemic *vv* sequences are */ai/* (*phai* ‘house’), and */au/* (*tau* ‘pot’). Tshangla has retained, as its only

TABLE 21.1 TSHANGLA CONSONANTS

	Labial	Alveolar	Retroflex	Palatal	Velar	Glottal
vl. stop	p*	t*	t̪ {tr}		k*	
+ asp.	p ^h {ph}	t ^h {th}	t̪ ^h {thr}		k ^h {kh}	
vd. stop	b	d	d̪ {dr}		g	
vl. aff.		ts		te {c}		
+ asp.		ts ^h {tsh}		te ^h {ch}		
vd. aff.		(dz)		dz {j}		
vl. fric.		s*		ɛ {sh}		
vd. fric.		z		(z) {zh}		
nas.	m*	n*		ɲ {ny}	ŋ* {ng}	
lat.		l*				
lat. fric.		(l) {lh}				
flap		r* {r̪}				
approx.	w			y		h

TABLE 21.2 TSHANGLA VOWELS

	Front	Central	Back
high	i		u
mid	e		o
low		a	

consonant cluster, the initial cr (preserved in Written Tibetan (WT)) for labial initials (e.g. *brangtong* ‘chest’). The WT velar-initial Cr clusters are pronounced with a retroflex coronal in Trashigang but retained in some dialects (e.g. *krame* ‘to distribute’; Trashigang: *tame*). The combination /peɪ/ occurs as well, but in only a few lexemes, including *pshi* ‘four.’

While lexical tone does not play any significant role in the Bhutan dialects of Tshangla, there is evidence of tonogenesis in progress. In the Padma-bKod dialect, the voicing contrast in initial consonants of the Bhutan dialects has apparently been replaced by a two-way tone contrast, with formerly voiced initials now voiceless with low tone, and high tone on originally voiceless initials. A high/low tone contrast is evident on some sonorant initials even in some Bhutan dialects.

3 MORPHOPHONEMICS

Tshangla verb roots may be divided into four classes according to their phonological conditioning of the initial consonant of a suffix. Representative suffixes are shown in Table 21.3. Vowel-final roots take an additional non-morphemic stem-extender /n/ in certain contexts. When these roots take the copula *la*, the initial /l/ of the copula is elided, leaving only the stem-extender (e.g. *pha-n-a*).

Two groups of verb roots do not pattern according to the generalizations presented in Table 21.3. First, certain nasal-final and *r*-final verb roots pattern together with the obstruent-final roots, as shown in Table 21.4 for the exceptional roots *lay* ‘mount’ and *sor* ‘exchange’. The exceptional nasal-final roots also show an alternation in the final consonant of the verb root itself, with *lak* ‘mount’ in the copula *la* and non-final clause forms,

TABLE 21.3 TSHANGLA VERB CLASSES

Form of root	Root	Copula tea	Copula la	Non-final clause
obstruent-final	<i>yek</i> ‘speak’	<i>yektea</i>	<i>yekla</i>	<i>yek</i>
nasal-final	<i>lay</i> ‘sit’	<i>laytea</i>	<i>layla/layna</i>	<i>lay</i>
liquid-final	<i>eor</i> ‘lose’	<i>cortea</i>	<i>corla</i>	<i>eor</i>
vowel-final	<i>pha</i> ‘bring’	<i>phantea</i>	<i>phana</i>	<i>phan</i>
Form of root	Nominalized	Infinitive	Imperative	Dubitive
obstruent-final	<i>yekpa</i>	<i>yekpe</i>	<i>yekteo</i>	<i>yektu</i>
nasal-final	<i>layma</i>	<i>layme</i>	<i>layeo</i>	<i>laydu</i>
liquid-final	<i>corba</i>	<i>corbe</i>	<i>coreo</i>	<i>cordu</i>
vowel-final	<i>phawa</i>	<i>phale</i>	<i>phayo/phai</i>	<i>phadu</i>

TABLE 21.4 EXCEPTIONAL NASAL/LIQUID-FINAL VERB ROOTS

Form of root	Root	Copula tea	Copula la	Non-final clause
excep. nas-final	<i>lay</i> mount	<i>laytea</i>	<i>lakla</i>	<i>lak</i>
excep. liquid-final	<i>sor</i> exchange	<i>sortea</i>	<i>sor²la</i>	<i>sor²</i>
Form of root	Nominalized	Infinitive	Imperative	Dubitive
excep. nas-final	<i>laypa</i>	<i>laype</i>	<i>layteo</i>	<i>laytu</i>
excep. liquid-final	<i>sorpa</i>	<i>sorpe</i>	<i>sorteo</i>	<i>sortu</i>

contrasting with *lay* ‘sit’ from Table 21.3. The exceptional *r*-final roots alternate between final /r/ and glottalized /r²/.

The second group of exceptional verb roots are those which end in /e/ or /i/ but pattern together with the liquid-final roots, shown in Table 21.5 for exceptional *ke* ‘send’ and *qi* ‘write.’ There is evidence to suggest that these are historically derived from a syllable-final /l/ (see Andvik 2010: 60–9 for a complete account).

4 NOUN PHRASES

The elements of the noun phrase and the general order in which they may occur in relation to the head noun (*N*) are as follows:

Poss., Dem., rc, Adj., N, (Adj), (rc), Quant., Indef., Topic, ‘ALL’, Case, prt.

The personal pronouns are shown in Table 21.6. The agentive pronouns are grammaticalized from the personal pronoun plus the agent case marker. The possessive and locative/dative pronouns are grammaticalized mergers of the personal pronoun plus the locative marker. There is also a dual form of the pronouns, transparently based on the lexeme *nyiktsing* ‘two’:

A reflexive construction is formed by the addition to the personal pronoun of the reflexive suffix *-ten*, possibly derived from the verb *tenme* ‘to rely on.’ The reflexive ordinarily indicates co-reference between subject and another argument of the clause, as in example (1) below. It may also be used with emphatic meaning, as in example (2).

- (1) *Ro ro-ten sokjap a-le ma-r-ba*
 3 3-REFL save do-INF NEG-can-INF
 ‘He is unable to save himself.’

TABLE 21.5 EXCEPTIONAL VOWEL-FINAL VERB ROOTS

Form of root	Root	Copula tea	Copula la	Non-final clause
excep. e-final	<i>ke</i> send	<i>ketea</i>	<i>kela</i>	<i>ke</i>
excep. i-final	<i>qi</i> ‘write’	<i>ditea</i>	<i>qila</i>	<i>qi</i>
Form of root	Nominalized	Infinitive	Imperative	Dubitive
excep. e-final	<i>keba</i>	<i>kebe</i>	<i>keeo</i>	<i>kedu</i>
excep. i-final	<i>qiba</i>	<i>qibe</i>	<i>qieo</i>	<i>qidu</i>

TABLE 21.6 PERSONAL PRONOUNS

	Unmarked	Agentive	Locative (possessive, dative)	Ablative
1s	jang	jigi (ji, jinggi)	jaga (ja, jangga)	jagai (janggai)
2s	nan	nangi	naga	nagai
3s	ro	roki	roka	rokai
1/2d	aching	achiki	achika	achikai
2d	naching	nachiki	nachika	nachikai
3d	roktsing	roktsiki	roktsika	roktsikai
1p	ai	ashi	aha	ahai
2p	nai	nashi	naha	nahai
3p	rokte	rokteki	rokteka	roktekai

- (2) *Jang-ten la Thimpu u-phe sem-rang manji*
 1s-REFL TOP Thimpu come-INF mind-EMPH NEG.COP
 ‘I myself have never planned to come to Thimpu!’

Adjectives may either precede or follow the head noun. Pre-nominal adjectives are restrictive; post-nominal adjectives are non-restrictive and descriptive of the head:

- (3) *dukpu waktsa khepa*
 poor child TOP
 ‘the poor child’ (identification)
- (4) *waktsa dukpu khepa*
 child poor TOP
 ‘the child, who is poor’ (description)

Other than a very small set of apparently non-derived adjectives, most adjectives in Tshangla appear to be diachronically derived from verbs. Some of these verbs are still extant, such that the derived adjective is homophonous with the stative form of the verb. However, even in these cases, it is possible to distinguish the adjectives from verbs by means of morphosyntactic tests. (See Andyik 2010: 39–42 for details.)

Relative clauses, like adjectives, occur both before and after the NP-head, the pre-nominal position being restrictive (5) and the post-nominal descriptive (6):

- (5) *Onya phai cot-khan songo ja-ga charo gila.*
 DEM house make-REL person 1SG-LOC friend COP
 ‘That person building the house is my friend.’

- (6) *Onya songo phai cot-khan khepa ja-ga charo gila.*
 that person house build-REL TOP 1sg-LOC friend COP
 ‘That person, who is building a house, is my friend.’

Demonstratives always precede the noun. Both the locative and non-locative sets may be used as demonstrative pronouns or demonstrative adjectives. The distinction between the locative and non-locative sets is semantic (Table 21.7).

A quantifier occurs after the nominal head and any postposed adjectival modifier or relative clause. The quantifier may be a numeral, a quantifying expression such as *tshe-bang* ‘some,’ *mangpu* ‘many,’ *nyungpu* ‘few,’ or the definite plural clitic *-ba*. There is no other number marking on nominals:

- (7) *ngam zum*
 day seven
 ‘seven days’
- (8) *pecha chilu-ba*
 book great -PL
 ‘the large books’

A bleached and grammaticalized sense of the numeral quantifier *thur* ‘one’ may occur as an indefinite marker. As such it follows other quantifiers, and may be collocated with them; in example (9) indefinite *thur* occurs with *mangpu* ‘many’:

- (9) *Jang songo mangpu thur-gi-rang kha+tang-nyi . . .*
 1sg person many one-AGT-EMPH criticize-NF
 ‘I am criticized by many people.’

A content question word together with the indefinite marker *thur* is the common way of forming an indefinite relative clause, ('whatever . . . , 'whoever . . . ' etc.):

- (10) *Ji-gi pura hang tshat-pa thur nan-ga bi-wa.*
 1sg-AGT completely what need-NOM one 2sg-LOC give-NOM
 ‘Whatever (you) needed I gave you.’

The contrastive topic marker *khepa* follows the NP head, and any postnominal modifier, quantifier, or indefinite marker:

- (11) *Unyu waktsa khepa ji-gi she-le khe-le dang.*
 DEM child TOP 1sg-AGT kill-INF must-INF EVID
 ‘That child I must kill. . . ’

The case particles follow the head, quantifiers, and topicalizer. As clitics, they are phonologically bound to the immediately preceding constituent:

- (12) *Jang nangpa cho khepa-ga ten-nyi . . .*
 1sg Buddhist faith TOP-LOC adhere-NF
 ‘I adhered to the Buddhist faith . . . ’

TABLE 21.7 DEMONSTRATIVES

	Non-locative	Locative
Proximate	<i>uthu</i> ‘this’	<i>otha</i> ‘this/here’
Distal	<i>unyu</i> ‘that’	<i>onya</i> ‘that/there’

There is a diverse class of focus particles that follow the case marker. Examples are the contrastive emphatic marker *-bu*, an emphatic marker *-rang*, and marked topic particles *-cho*, and *-la*.

- (13) *Tshebang-gi waktsa-ba-ka-bu ngalong-ga chas-rang phi-na.*
some-AGT child-PL-LOC-FOC Ngalong-LOC speech-EMPH do-COP
'Some even speak the Ngalong language to their children.'

5 SYNTACTIC ROLES

Tshangla constituent order is sov. Deviations from this occur in pragmatically marked contexts. While Tshangla lacks monoclausal morphosyntactic alternations (passivization, raising, etc.), a syntactic subject may be defined (independently of semantic or pragmatic features) as the controller of zero-anaphora in multi-clause constructions such as conjoined clauses and complement clauses. In example (14), the agent-subject of the non-final clause, being in subject position, controls (is coreferent with) the zero subject of the final clause, while in example (15), the semantic patient subject of the non-final clause controls the final clause subject:

- (14) *Gopen-gi tsonpa kong-nyi, 0 phiska di-wa.*
chief-AGT prisoner strike-NF 0 outside go-NOM
'The chief beat the prisoner and went outside.'
*'The chief beat the prisoner and he (the prisoner) went outside.'
- (15) *Tsonpa gopen-gi brak-nyi, 0 phiska di-wa.*
prisoner chief-AGT scold-NF 0 outside go-NOM
'The prisoner was scolded by the chief and went outside.'
*'The prisoner was scolded by the chief and he (the chief) went outside.'

Previously mentioned participants with high topicality, both subjects and objects, are frequently referred to by means of zero-anaphora. In addition to this, however, an argument of any multivalent verb may be omitted when it is regarded as unimportant, as in example (16) below:

- (16) *Drowan thong-ma.*
thief see-NOM
'The thief was seen.'

A large number of 'compound verbs' consist of an inflectable verb preceded by some other uninflectable lexical item (e.g. *ha + gole* 'heart + put' = 'to understand'). Where the meaning of the preverbal element is not completely obscure, the meaning of the combination is idiomatic and the two function as a single lexeme. They are, however, independent words grammatically and may be separated by certain particles or adverbials.

6 CASE MARKING

6.1 Agentive

The presence of the 'agentive' case particle *-gi* (*-ki*)¹ is determined by a complex combination of syntactic, semantic, and pragmatic factors such as control, volition, consequence, directed or purposive activity, and violation of expectations. Tense/aspect marking also has some effect on agentive case marking; the agentive marker is more

likely in the past tense and perfective aspect. There is also evidence that pragmatic considerations such as topicality and focus are factors in determining agentive case marking, (see Andvik 2010: 124ff. for discussion). However certain predicates, such as verbs of speech, perception, and cognition, regularly take agentive-marked subjects regardless of other potential determining factors.

The Tshangla agentive marker has a secondary instrumental function, which appears when two or more nominals marked in *-gi* are present in the clause. One, the agent, is typically a human or animate actor; the other, the instrument, is some inanimate object which the agent uses in order to carry out the action:

- (17) *Ji-gi nan chowang-gi za-me!*
 1sg-AGT 2s sword-AGT slash-INF
 'I will slash you with a sword!'

The agentive marker may also occur on oblique agent/causers in an intransitive clause:

- (18) *Ro don-gi nyos-pa.*
 3s demon-AGT crazy-NOM
 'He went crazy because of the demon.'

The agentive case marker is also used to mark an adverbial clause as a cause of the main clause event (see section 14).

6.2 Locative/dative

The locative/dative case marker *-ga* (*-ka*) marks the recipient or goal of a ditransitive verb such as *yekpe* 'speak, tell,' *genme* 'show,' or *bile* 'give,' and also locative arguments of monotransitive verbs of position or motion, such as *tsepe* 'lean' or *tanme* 'stand.' The locative/dative is also used to mark experiencer or goal objects of many monotransitive verbs such as *trokpe* 'bother,' *brangpe* 'scold,' *rele* 'depend,' and *rumpe* 'help':

- (19) *Ro-ki ja-ga brang-pa.*
 3sg-AGT 1sg-LOC scold-NOM
 'He scolded me.'

Oblique locative or temporal adverbials, which often precede the core subject, may also be marked with *-ga*:

- (20) *Thimpu-ga phom khe-na.*
 Thimpu-LOC snow fall-COP
 'It's snowing in Thimpu.'

Locative/dative case marking occurs on the subject to encode possession in a copular clause:

- (21) *Jelpo-ga waktsa sam cho-wa.*
 king-LOC child three stay-NOM
 'The king had three children.'
- (22) *Ja-ga gari man-ca.*
 1sg-LOC car NEG-COP
 'I don't have a car.'

The locative/dative case marker also encodes the dependent constituent in an NP, most commonly but not restricted to possessors, kinship terms, etc.:

- (23) *Na-ga mi gapthang lekpu ma-la.*
2sg-LOC bow shooting good NEG-COP
'Your shooting is not good.'

There is evidence in some dialects for an earlier distinction, now largely neutralized, between a genitive *-ga* and locative/dative *-gu*, the latter still extant on fossilized constructions such as *na-gu thale* (ear-LOC lay) 'to hear,' and as an alternant in free variation with *-ga* in locative/dative but not genitive contexts.

6.3 Ablative

The ablative case marker *-gai* (*-kai* ~ *ke*) marks oblique locative or temporal arguments as the source or reference point away from which the action of the proposition is directed:

- (24) *Ro-ki yek-khan nyan-pe-ga songo-ba throm-gai zo-ma-la.*
3sg-AGT speak-REL listen-INF-LOC person-PL town-ABL gather-NOM-COP
'The people gathered from the town to listen to him speak.'
- (25) *Chutse gu-gai ja-ga sungjapa a-le kor cho-wa.*
hour nine-ABL 1sg-LOC duty do-INF turn stay-NOM
'From nine o'clock it was my turn to be on duty.'

The ablative may also mark a constituent as a pathway along which the action of the proposition takes place:

- (26) *A-ching nyiktsing ung nang-kai leng jong-khe.*
1pl-DUAL two field in-ABL move go-HOR
'Let's us two take a walk through the fields.'

The meaning of the ablative case may extend beyond a concrete physical or temporal notion to encode the notion of source or pathway in a more abstract sense, such as previous state, cause, reason, instrument, means to an end, logical reference point, or point of comparison, as in example (27):

- (27) *Jang uthu tapthur di-wa-gai kukhaila-kap di-wa drik-pe.*
1sg DEM with go-NOM-ABL tiger-with go-NOM fit-INF
'I would rather go with the tiger than with that one.'

7 TENSE/ASPECT

A Tshangla sentence, i.e. a minimal finite context-independent utterance, consists of a single final clause optionally preceded by one or more non-final clauses. Non-final clauses are not specified for any grammatical category. Final clauses are inflected for tense, aspect, and mirativity. These categories are morphologically represented by a periphrastic combination of verbal participles plus a copula or auxiliary verb. The chart in Table 21.8 shows the entire final verb paradigm for the verb *dile* 'to go' (non-mirative forms only; see section 9).

The nominalized or infinitive participle (see Table 21.3) functions as the main verb, though it is structurally embedded under the auxiliaries *chole* 'to stay,' *uphe* 'to come,' and the existential/descriptive copula *ca*. The perfective verbal phrase is formally analogous to a complement clause structure (see section 13), while the imperfective verbal phrase is formally analogous to a serial verb construction (see section 16). Note that the

TABLE 21.8 AFFIRMATIVE VERB PARADIGM

	Past	Present	Future
Simple pftv.	<i>di-wa</i> ‘went’		<i>di-le</i> ‘will go’
Simple impf.	<i>din-cho-wa</i> ‘was going’	<i>din-ca</i> ‘is going’	<i>din-cho-le</i> ‘will be going’
Perfect pftv.	<i>di-wa-cho-wa</i> ‘had gone’	<i>di-wa-ca</i> ‘had gone’	<i>di-wa-u-phe</i> ‘will have gone’ <i>di-wa-u-phe</i> ‘would have gone’
Perfect impf.	<i>din-cho-wa-cho-wa</i> ‘had been going’	<i>din-cho-wa-ca</i> ‘has been going’	<i>din-cho-wa-u-phe</i> ‘will have been going’ <i>di-le-u-phe</i> ‘would have been going’
Prospective pftv.	<i>di-le-cho-wa</i> ‘was going to go’ ‘would have gone’	<i>di-le-ca</i> ‘is going to go’	<i>di-le-u-phe</i> ‘will be about to go’ ‘may go’

simple perfective forms are composed of the nominalized or infinitive participle alone (for past or future time respectively). These are at first glance exceptions to the verb-plus-copula pattern. However, not shown in the chart is the copula *gila*, an uninflected equative copula which may be postposed to the various inflections as a stylistic variation but with no semantic consequence. This gives for example *di-le gila* as an alternative for the future perfective *di-le*, and *di-wa gila* for the past perfective *di-wa*.

Note here, and also for the negative below, that some of the periphrastic constructions have alternative modal readings, such as contrafactual, permission, and obligation (e.g. *di-le-cho-wa* ‘would have gone,’ *di-le-u-phe* ‘may go’).

8 NEGATION

Negation is marked by means of the prefix *ma-*, but in indicative final clauses its interaction with the elements of the verbal phrase is complex, sometimes occurring phrase-initially on the verb, sometimes as a prefix on the auxiliary, as shown in Table 21.9. With the exception of the prospective forms, the negative prefix is attached to the verb root in perfective constructions, and to the auxiliary in imperfective. In the negative constructions, the place of the existential copula is occupied by the auxiliary *chi*. The verb suffix *-lu* occurs only in the simple imperfective negative past and present forms, where the negative prefix occurs on the auxiliary.

Negative constructions with *gila* are based on the affirmative forms, with the addition of the negative prefix to the verb root (e.g. past perfective *ma-di-wa gila* ‘did not go’), or, if the negation itself is emphatic, with a negative form of *gila* itself (past perfective *di-wa manggi*).

All non-indicative, non-final, or embedded nominalized clauses are negated simply by means of the prefix *ma-* on the verb root. Prohibitives are composed of a root with an imperative suffix (see section 11) and a negative prefix.

9 MIRATIVITY

The Tshangla final-clause verbal phrase also encodes ‘mirativity,’ an evidential-like distinction which marks an utterance as conveying information which is new or unexpected

TABLE 21.9 NEGATIVE VERB PARADIGM

	Past	Present	Future
Simple pftv.	<i>ma-din-chi</i> ‘did not go’		<i>ma-di-la</i> ‘will not go’
Simple impf.	<i>di-lu-man-chi</i> ‘was not going’	<i>di-lu-ma-n-ca</i> ‘is not going’	<i>din-ma-cho-la</i> ‘will not be going’
Perfect pftv.	<i>ma-di-wa-cho-wa</i> ‘had not gone’	<i>ma-di-wa-ca</i> ‘has not gone’	<i>ma-di-wa-u-phe</i> ‘will not have gone’ <i>ma-di-wa-u-phe</i> ‘would not have gone’
Perfect impf.	<i>di-wa-man-chi</i> ‘had not been going’	<i>di-wa-ma-n-ca</i> ‘has not been going’	<i>di-wa-ma(ng)-pha</i> ‘will not have been going’ <i>di-wa-ma(ng)-pha</i> ‘would not have been going’
Prospective pftv.	<i>ma-di-le-cho-wa</i> ‘was not going to go’	<i>ma-di-le-ca</i> ‘is not going to go’	<i>ma-di-le-u-phe</i> ‘will not be about to go’ <i>ma-di-le-u-phe</i> ‘may not go’ alt: <i>di-le-ma(ng)-pha</i> ‘should not, ought not go’
		‘would not have gone’	

to the speaker. Mirativity is marked in both the existential and equative copulas, whether they occur in copular clauses or as auxiliaries in verbal inflections. The mirative alternant of the equative copula *gila* is *giwala*, while the mirative alternant of the existential copula *ca* is *la*:

- (28) *Ama khamung zik-ca.*
 mother clothes wash-COP
 ‘Mother is washing the clothes.’
- (29) *Ama khamung zik-la.*
 mother clothes wash-COP
 ‘Mother is evidently washing the clothes.’

The non-mirative utterance in (28) would be spoken by a person with prior knowledge of the proposition. The mirative utterance (29) would be spoken by someone who had just now learnt the matter, whether by hearsay, or by first-hand observation, for example by walking around the corner and seeing mother in the process.

In the verbal phrase, only the simple perfective and imperfective forms show a mirative contrast; the distinction is unavailable for the perfect or prospective. For the simple perfective forms (*di-wa* ‘went’ *di-le* ‘will go’), which lack a copula, the mirative is encoded by addition of the mirative copula to the nominalized verb, i.e. *di-wa-la* ‘evidently went’; here the mirative copula *la* contrasts with zero-marking, rather than with the non-mirative *ca*.

10 COPULAR CLAUSES

Both final and non-final clauses may contain a copular predicate. Copular predicates are of two types, those containing the existential copula *ca* (or its mirative counterpart *la*) and those containing the equative copula *gila* (or its mirative counterpart *giwala*).

- (30) *Otha phai chilu ca.*
 DEM house great COP
 ‘That house is big.’

- (31) *Ja-ga dung Wamrong gila.*
1sg-LOC village Wamrong COP
'My village is Wamrong.'

Copular clauses built on *ca* encode notions such as description, possession, location, existence, while copular clauses built on *gila* normally encode identity and set membership, but may also encode description, possession, location, and existence when special emphasis is being added:

- (32) *Unyu to zhimpu gila!*
DEM food delicious COP
'This food is delicious!'

In non-final clauses, for descriptive and other predictions, the grammaticalized verb *chole* 'to stay' takes the place of the existential copula *ca*. For non-final equative clauses, the equative *gila* takes non-final inflection.

11 MODALITY AND SENTENCE-TYPES

Final clauses may be marked for subjunctive or hortative moods, in which case they are not inflected for tense, aspect, or mirativity. The subjunctive mood is encoded by a verb root with the suffix *-du/-tu*, or with the infinitive verb followed by a subjunctive copula *gidu*. It presents a proposition as a potentiality:

- (33) *Om toka sha tsong-me gi-du?*
now bull meat sell-INF COP-SUB
'How about selling the meat?'

The subjunctive frequently marks the complement of a cognitive act, such as 'to hope':

- (34) *Ser nyong-tu dak-pa rewa-gi rok-tsing thur thur-gi kholong phi-wa.*
gold receive-SUB say-NOM hope-AGT 3-DUAL one one-AGT fight do-NOM
'In the hope of getting the gold, the two of them fought with one another.'

An additional subjunctive form *gisa* of the equative copula may occur either in an equative predicate or as a grammatical auxiliary in a verbal predicate. There is no corresponding subjunctive final verb suffix *-sa*. (However see section 12 for a homophonous relativizer suffix *-sa*.)

- (35) *Topda-gi khe-wa gisa.*
gun-AGT strike-NOM COP
'(He) may have been shot.'

The first person plural hortative suffix *-khe* marks an utterance as an invitation for the listener to join the speaker in the activity indicated, as in English 'let us . . .' or 'we ought to . . .':

- (36) *A-ching ja-ga ajang-ga brang-ka din-than chas a-khe.*
1p-DUAL 1sg-LOC uncle-LOC place-LOC go-NF talk do-HOR
'Let's the two of us go to my uncle's place and talk.'

The second person hortative suffix *-sho/-co/-i(yo)* encodes the imperative mood:

- (37) *Ja-ga tam hur shik-co!*
1sg-LOC story one explain-HOR
'Tell me a story!'

The third person hortative suffix *-chen* expresses an injunction by the speaker upon a third party, as in English ‘Let him . . .’ ‘He should . . .’, or ‘He must . . .’:

- (38) *Songo mar-khan-gi man zan-chen.*
person sick-REL-AGT medicine eat-HOR
‘The sick person should take his medicine.’

The interrogative sentence-type is distinct from the moods described above, in that it consists of a finite clause followed by a sentence-final question particle. The different moods may cross-code with the interrogative and declarative sentence-types. Interestingly, when the hortative mood is logically ‘nested’ within an interrogative speech act, the deixis of the hortative marker reflects the situation of the ‘embedded’ hortative proposition, rather than that of the interrogative utterance. So, in example (39), the second person imperative *-sho* places the obligation upon the first person subject of the interrogative sentence:

- (39) *Jang thar-sho mo?*
1sg release-HOR QUES
‘Must I release (him)?’

Interrogatives are of two types, the polarity question with the sentence-final particle *mo*, as in (39) and (40), and the content question with the sentence-final particle *ya*, as in (41). In content questions, the constituent which represents the unknown information is substituted for by a question word, such as *hang* ‘what,’ *ibi* ‘who,’ *o* ‘where,’ *hala* ‘when,’ *hanyi* ‘why,’ and *hangten* ‘how.’ The question word occurs in the preverbal focus position:

- (40) *Unyu wa daza brang-ma mo?*
DEM cow young bear-NOM QUES
‘Has this cow born a calf?’
- (41) *Lopen-gi hang yek-pa ya?*
teacher-AGT what speak-NOM QUES
‘What did the teacher say?’

The existential copular clause may occur in all of the non-declarative sentence-moods. Here as well the grammaticalized verb *chole* ‘to stay’ takes the place of the existential copula *ca*. The equative copula does not occur in the hortative mood (**gikhe*, **gisho*, **gichen*).

12 RELATIVE CLAUSES

Relative clauses are formed by a ‘gapping’ strategy, in which the place of the head noun within the relative clause is left empty. The relative clause verb is either nominalized or infinitive, or it takes one of two relativizing suffixes. Relative clauses may be formed on subjects, objects, or (somewhat rarely) on an oblique argument. Relative clauses may be headless, or externally headed, i.e. the head noun occurring outside the relative clause itself. The most common relativizing suffix is *-khan*:

- (42) *Yu jam-khan songo khepa namesame shonang phe-le.*
wine drink-REL person TOP very happy feel-INF
‘A person who has drunk wine becomes very happy.’

- (43) *Ro-ki gadang-gi tsung-khan songo thamchen-rang rolong ri-le.*
3sg-AGT hand-AGT seize-REL person all-EMPH zombie become-INF
'The people that he seizes with his hands all become zombies.'

The suffix *-sa* is used to relativize on a locative argument, as in 'the place where':

- (44) *Nyi lok-nyi jang cho-sa phai-ga shek-pa.*
And return-NF 1sg stay-REL house-LOC arrive-NOM
'So returning, we arrived at the house where I was staying.'

Relative clauses may be formed on the nominalized or infinitive verb, followed by an optional locative/dative case marker:

- (45) *Nang-ka shek-pa(-ga) songo-ba shadar phi-nyi ca giwala.*
inside-LOC arrive-NOM (-LOC) person-PL shout do-NF COP COP
'The people who had arrived inside were shouting.'
- (46) *Dangpo shing khuk-pe (-ga) zogo khepa thur cho-wa-la.*
ancient wood carve-INF (-LOC) carpenter TOP one stay-NOM-COP
'Once upon a time, there was a carpenter who carved wood.'

Relative clauses may also be headless, whether formed from *-khan* (47), *-sa* (48), or a nominalized or infinitive verb, (49) and (50):

- (47) *Ngang se-khan-gi ngang jang-ca.*
song know-REL-AGT song sing-COP
'Those who know the songs are singing.'
- (48) *Nan-gi yek-sa cho-le gi-du.*
2sg-AGT speak-REL stay-INF COP-SUB
'I suppose we will live where you say ("... live at your saying").'
- (49) *Nyi rang-gi re-ba (-ga) thur a-n-cho-wa.*
And self-AGT can-NOM-LOC one do-NF-stay-NOM
'And they were doing whatever they could.'
- (50) *Na-shi zemu-ba za-le ja-me (-ga) bi-le khe-le.*
2p-AGT small-PL eat-INF drink-INF-LOC give-INF must-INF
'You must give the small ones (something) to eat and drink.'

13 COMPLEMENTATION

Tshangla complement clauses contain a nominalized or infinitive verb followed by an optional locative case marker:

- (51) *Jang yu ja-me (-ga) thup tha-wa.*
1s wine drink-INF (-LOC) throw leave-NOM
'I left off drinking wine.' (*thup thale* = 'abandon, leave')
- (52) *Jang waktsa rokha di-wa (-ga) thong-ma.*
1s child fall go-NOM(-LOC) see-NOM
'I saw the child fall down.'

Deontic modal notions of ability, obligation, etc., are encoded by means of complement-taking verbs such as *rebe* 'can,' 'to be able to,' *khele* 'must,' 'to be required to.'

A paraphrastic causative construction is formed with the matrix verb *bile* ‘to give’ and an infinitival complement:

- (53) *Ro-ki jelpo zi-me bi-wa.*
3sg-AGT king lie-INF give-NOM
'He had the king lie down.'

There is case-marking evidence (see Andvik 2010) that suggests that at least one of the arguments of the complement clause has a role in the argument structure of the matrix verb as well. A distinct type of complement clause, the ‘information’ complement in (54), by contrast, does not necessarily share any argument with the matrix clause verb. It occurs only with certain verbs of perception, cognition, or utterance. The complement contains a finite clause with a content question word, followed, interestingly enough, not by the content question particle *ya*, but by the polarity question particle *mo*:

- (54) *Ibi u-na mo got-co!*
who come-COP QUES look-HOR
'See who is coming.' (i.e. 'Look to find out who is coming.')

Infinitive and nominalizing verbs are also used in a number of constructions where a clause is embedded as an argument of constituents other than verbs, for example noun complements (55) and objects of postpositions like *saken/sakpu* ‘until,’ *korgai* ‘about,’ and *dabu* ‘as,’ ‘like’ (56). The locative/dative case marker occurs optionally on the embedded noun-complement, but not on the complement of a postposition:

- (55) *Nyi dan shi-wa-ga tshang khepa songo thamcen-gi na tha-wa.*
And 3sg die-NOM-LOC news TOP person all-AGT ear leave-NOM
'And all the people heard the news that he had died.'
- (56) *Uthu cho jang kawa chilu cat-pa dabu thong-mu man-chi.*
DEM TOP 1s hardship great suffer-NOM like see-PTCP NEG-COP
'This I do not see as suffering a great hardship.'

14 ADVERBIAL CLAUSES

Tshangla adverbial clauses are non-final clauses which limit, expand upon, or otherwise modify the meaning of the matrix clause. As opposed to complement clauses, their arguments do not have a grammatical role in the final clause:

- (57) *Shama di-nyi-bu, jang dangsanken yitka u-na.*
long.time go-NF-FOC 1sg clearly memory come-COP
'Even though a long time has passed, I still remember clearly.'
- (58) *Choga phi-le-ga, ro tormu katang cos-pa.*
ritual do-NP-LOC 3sg tormu big make-NOM
'In order to do the ritual, he built a large *tormu*'

The rhetorical relationship between the adverbial and matrix clauses is determined by a combination of verb suffixes, particles, and case particles:

- <i>n-than</i>	sequence ('after . . .' < <i>tha</i> 'to leave')
- <i>deke</i>	result ('because . . .' < <i>den</i> 'meaning' + 'Ablative')
- <i>nyi-la</i>	conditional ('if . . .')
- <i>nyi-sha</i>	exclusive conditional ('only if . . .')

-nyi-bu	concessive ('even though . . .')
-nyi	unspecified non-final verb
-la	concurrent ('while . . .')
-la-gai (ABL)	temporal anteriority ('until . . .', 'before . . .')
-le(-ga) (LOC)	purpose/goal ('in order to . . .')
-wa(-ga) (LOC)	attendant circumstance ('with/without . . .')
-wa-gi (AGT)	cause ('because . . .')
-wa-gai (ABL)	temporal posteriority ('after . . .'); comparison ('rather than . . .'); cause/effect ('because . . .')

15 CLAUSE CHAINS

Clause chains are concatenations of non-final clauses which do not merely modify the main clause event, but participate on a par with the final clause in the main 'event-line' of the discourse:

- (59) *Nyi gisa na gisa dak-nyi, bozong zong-nyi,*
And maybe EMPH maybe say-NF cassava boil-NF
khoptang khop-nyi, laga-gi chom-nyi, nyi sa nang-ka
skin peel-NF leaf-AGT wrap-NF and ground inside-LOC
che-nyi, onya wang thur tsuk-nyi tha-wa.
plant-NF thus hole one put-NF leave-NOM.
'And thinking, "Well, maybe", boiling the cassava, peeling it, wrapping it with a leaf, and planting it in the ground, they put it in a hole.'

The most common subordinating marker in a clause-chain construction is *-nyi* by itself. Adverbial clauses formed on the infinitive and nominalizer suffixes are not given clause-chain readings.

16 CONCATENATION

In the clause-chain construction (above), each of the clauses consists potentially of a verb with a complete set of grammatical arguments. With the unspecified non-final marker *-nyi*, however, the conjoined constituents may be something less than a complete clause. Andvik (2010: 338ff.) shows that clause chains, 'serial predicates,' and 'serial verbs' represent three points on a 'concatenation continuum,' representing the degree of syntactic and semantic integration.

- (60) *Dorji tiru chum-nyi u-phä.*
Dorji money finish-NF come-NOM
'Dorji used up the money and came.'
- (61) *Ro ro-ten jap a-nyi chas mangpu zhu-wa-la.*
3s 3sg-REFL save do-NF talk much offer-NOM-COP
'He said many things in his own defense (lit. 'Saving himself, he offered much talk').'
- (62) *Kha thur phai nang-ka phur-nyi u-phä.*
bird one house inside-LOC fly-NF come-NOM
'A bird flew into the house.'

Example (60) is a clause chain, where all arguments of both non-final and final clauses, although they *may* be coreferential, are at least *potentially* distinct. Each clause in the chain encodes a distinct event. Example (61) is a so-called ‘serial predicate,’ where the non-final and final clause may have distinct object arguments, but where subject and all peripheral arguments are necessarily shared. Semantically, the verbs in a serial predicate construction represent distinct events which are nonetheless construed by the speaker as facets or components of some ‘macro event.’ Finally, example (62) is a ‘serial verb’ construction, defined here as one in which non-final and final verbs must occur adjacently, and necessarily share all arguments. In the serial verb construction, the two verbs together represent a single event.

Many of the final verbs which occur in a serial verb construction are well on their way to becoming grammaticalized auxiliaries. The verb *bile* ‘to give,’ for example, is semantically bleached and generalized to the point where it may occur with any other verb as a benefactive/malefactive marker, as in example (63).

- (63) *Nyi ro thong-ma-kap-ny songo thur-gi ngar-nyi bi-wa-la.*
And 3s see-PTCP-with-NF person one-AGT laugh-NF give-nom-COP
'And upon seeing him, one person laughed at him.'

As seen in section 7 earlier, the serial verb construction with *chole* ‘to stay’ is unique in having grammaticalized to the point where *cho* has become an inflectional (imperfective) marker.

17 GRAMMATICALIZED NON-FINAL VERBS

Certain non-final verbs in a serial predicate construction are grammaticalized to postpositions or complementizers. The verb *khonme* ‘to follow,’ ‘chase,’ for example, in non-final position and carrying its own object but no independently specifiable subject, is bleached to a postposition meaning ‘along’ or ‘around’:

- (64) *Jang das dung khon kor-be di-le.*
1s bit village follow go.around-INF go-INF
'I'm going to walk around the village a bit.'

In example (64), the erstwhile object argument of *khonme* (i.e. *dung* ‘village’) has now become the nominal argument of the postposition (a similar course of development to, for example, the ‘co-verbs’ of Mandarin Chinese).

Two grammaticalized non-final verbs which play a significant role in Tshangla discourse are the verbs *dakpe* ‘to say’ and *ale* ‘to do,’ which, uniquely among this class, may take not only a nominal argument, but an adjective, a verb, a nominalized clause, and even a complete utterance as their complement. The verb *dakpe*, in its non-final inflection *dak-nyi*, functions as a complementizer, acting as head of the constituent containing the embedded quotation. *Daknyi* may take a complement of speech (65) or cognition/intention (66):

- (65) *Nan shong la-i dak-nyi, brumsha-gi yek-pa-la.*
2s breath take-HOR say-NF pumpkin-AGT say-NOM-COP
'The pumpkin spoke, saying, "You take a rest!"'
- (66) *Lhangpoche mang-Ø-chen dak-nyi, ro shu a-nyi lhak-cho-wa.*
elephant NEG-come-HOR say-NF 3s strong do-NF read-stay-NOM
'Thinking/intending, "May the elephant not come," he was reading out loud.' (i.e.
'In order to keep the elephant from coming, he was reading out loud.')

The verb *ale* ‘to do,’ in its non-final form *anyi*, like *daknyi*, in addition to thought, cognition, or intention complements, may also take adjectives, instruments, and non-finite clausal complements as well. An adjective embedded under *anyi* is the common way of forming manner adverbial expressions:

- (67) *Phama-gi waktsa-ba thamcen dolo a-nyi phang-ca.*
parent-AGT child-PL all equal do-NF love-COP
'Parents love their children equally.'

An instrumental argument may be embedded under *anyi*:

- (68) *Ro-ki nai-ba ri-gi a-nyi thrisor phi-wa.*
3sg-AGT 2p-PL water-AGT do-NF cleansing do-NOM
'He cleansed you with water.'

A non-finite clause containing a nominalized or infinitive verb may be embedded under *anyi*; the embedded clause event being interpreted as an ‘attendant circumstance’ to the final clause event. This construction is most common with a negative complement. Its function is similar to an English clause embedded under a preposition such as ‘with/without’:

- (69) *Songo mangki ma-se-wa a-nyi ro ihup tha-le khe-le-la.*
person public NEG-know-NOM do-NF 3s throw leave-INF must-INF-COP
'I must leave her without the public knowing.'

Grammaticalized postpositional *anyi* with a non-finite clausal complement is often reduced to *-an* or *-n*:

- (70) *Jang thola to hang-rang ma-ga-la-n, ri thur-rang*
1s up.there food what-EMPH NEG-give-PTCP-NF water one-EMPH
ma-ga-la-n, otha nang-ka dap tha-nyi . . .
NEG-give-PTCP-NF DEM inside-LOC insert keep-NF
'... keeping me there, without giving me any food, without giving me any water . . .'

Temporal expressions such as ‘while’ and ‘before’ are commonly encoded by means of a syntactically complex construction where the clause containing the participle in *-la* is embedded under the postposition *-(a)n*, and in turn under the verb *chole* ‘to stay’:

- (71) *Ro ma-shi-la-n cho-la-kap, khamung tshok-pa cho-wa.*
3s NEG-die-PTCP- stay-PTCP- clothing sew-NOM stay-NOM NF with
'Before she died, she had made clothing.' (lit. 'While she was staying not dying,
she had made clothing.')

FURTHER READING

Egli-Roduner, Susanna. 1987. *Handbook of the 'Sharchhokpa-lo/Tshangla'*. Thimpu: Helvetas. (A wordlist containing hundreds of words with short glosses, especially medical terms.)

NOTE

- 1 Unlike the alternations in the initial consonant of the verbal suffixes described above (see Table 21.3), the voicing alternations (k/ g) in case markers resist a straightforward morphophonemic explanation.

REFERENCES

- Andvik, Erik. 2010. *A Grammar of Tshangla*. Leiden: Brill. (A comprehensive reference grammar with illustrative examples.)
- Aris, Michael. 1980. Notes on the history of the Mon-Yul corridor. In Michael Aris and Aung San Suu Kyi (eds), *Tibetan Studies in Honour of Hugh Richardson*, 9–20. New Delhi: Vikas.
- Das Gupta, Kamalesh. 1968. *An Introduction to Central Monpa*. Shillong: North-East Frontier Agency.
- Sun, Hongkai, Lu Shaozun, Zhang Jichuan, and Ouyang Jueya. 1980. *The Languages of the Monpa, Luoba, and Teng Peoples*. Beijing: Chinese Academy of Social Sciences.
- van Driem, George. 1998. *Dzongkha*. Leiden: Research School of Asian, African and Amerindian Studies.
- Zhang, Jichuan. 1986. *A Brief Description of Cangluo Monpa*. Beijing: Chinese Academy of Social Sciences.

CHAPTER TWENTY-TWO

DOLAKHA NEWAR

Carol Genetti

1 INTRODUCTION

Dolakha Newar is a Tibeto-Burman language spoken in the village of Dolakha, approximately 145 kilometres to the east of Kathmandu, Nepal.¹ The language has been referred to as Dolakhali, Dolakha (Newar), and Dolakhae. The latter form contains the genitive morpheme = *e*, thus means ‘of Dolakha’, and is the form used by the speakers to refer to themselves and their language. I will use the terms Dolakhae and Dolakha Newar interchangeably.

Dolakhae is clearly in the Newar family and is closely related to the dialects of Newar spoken in the Kathmandu Valley and other locations throughout Nepal. I prefer to refer to Dolakhae as a ‘language’ as opposed to a ‘dialect’ of Newar, as it is mutually unintelligible with the Newar spoken in the Kathmandu Valley. There are a number of features that contribute to this, most notably the retention of final consonants, the loss of breathy voice, the presence of a full subject-agreement paradigm on the verb, and different morphology for nominalization, a ubiquitous construction in both languages. Historical records indicate that the split of Dolakhae from the other dialects occurred a minimum of 700 years ago, perhaps much earlier.

The number of speakers of this language is unknown, as many Dolakha Newars have migrated to other parts of Nepal and beyond to seek employment. The number is likely to be between 5,000 and 10,000, probably towards the lower end of this range. The current trend is for young adults from the village to move out to Kathmandu or other cities for business and educational opportunities, and to marry within the wider Newar community. Consequently, many of the current children of Dolakha Newars are not learning to speak the language. On the other hand, the people of Dolakha in the outlying areas have a strong civic commitment to the village and sense of community. Several groups have grown up dedicated to maintaining the language and culture of the Dolakhae people.

It is not known if there are any true dialectal divisions within the Dolakhae language. The village of Dolakha is relatively small and homogeneous, thus unlikely to give rise to significant dialectal development. However, while the relevance of dialect is questionable, the notion of idiolect is clearly important. My two primary consultants were two young female cousins who grew up in the same household until they were about eight years old. They always conversed freely with each other and never noticed any differences in their speech. However, as I worked with the two of them I found linguistic differences in a number of areas, including different vowel harmony patterns, different inflections of the copula, and different patterns of use of the nominalizers. Later I worked with another, younger, cousin, and found yet another pattern of vowel harmony. These differences seem to be the result of different analyses of occurrent patterns during language acquisition, probably reflecting variation already present in the speech of the preceding generations.

The goal of the current chapter is to provide information on the most central aspects of the grammar of Dolakhae. The final section (6) contains a brief narrative which allows the reader to obtain a truer sense of the language. When possible, the description of the language will be indexed with references to the text, with line numbers enclosed in angled brackets, e.g. <10>. Space limitations require brevity; for further work on the language, see Genetti (2007) and works cited therein (pp. 22–3), especially Shrestha (1996).

A list of abbreviations specific to this chapter is given in an appendix.

2 PHONOLOGY

The consonant phonemes of Dolakhae are given in (1):

(1)	p, ph, b	t, th, d	t̪, t̪h, d̪	c, ch, j	k, kh, g
	m	n		ɳ	
	s			h	
	l	r			
	w		y		

Of particular interest are the presence of a retroflex series of consonants (not attested in other Newar dialects), and the absence of breathy voice. The former is attested by the minimal pairs *tār-* ‘hear’ ‘hear’ and *tār-* ‘fix’, *thon-* ‘hide’ and *thon-* ‘wake’, and *dōn-* ‘finish’, *dōn-* ‘stand’. The dental/retroflex distinction was probably a feature of Proto-Newar, and has been lost in other dialects. Note that the symbols /c, ch, j/ represent alveo-palatal affricates.

There are a few words in Dolakhae with breathy-voiced stops; however, there is no evidence of breathy-voiced nasals or approximants, as found in Kathmandu Newar. Most of the vocabulary items with breathy voice are clearly borrowings from Nepali or Sanskrit. There are a few Newar words which speakers pronounce with *dh* or *jh* in careful pronunciation, such as *dhū* ‘tiger’ and *jhi* ‘ten’; however, the breathy voice is variable across speakers and it is not heard in connected speech. It is notable that some Dolakhae speakers pronounce some Nepali breathy consonants with plain phonation, e.g. *ghar* ‘house’ is pronounced by some as *gar*.

One other note on the consonant inventory is that two of these ‘phonemes’ are actually in complementary distribution. The retroflex stop /d̪/ occurs in word-initial position, while the tap /r/ occurs between vowels and at the end of a syllable. On purely distributional grounds, these two are allophones of one phoneme and should be represented as such. However, the two sounds are different phonetically, and speakers insist that they are distinct. This is probably attributable to these sounds being phonemically distinct in Nepali (an Indo-Aryan language, the national language of Nepal, which all Dolakhae speakers are bilingual in), and to their being represented by different characters in the Devanagari alphabet. So although distributional criteria argue for a single phoneme, I prefer to analyse them as two, which I take to be a more accurate reflection of speaker knowledge.

The vowel phonemes of Dolakhae are given in (2):

(2)	i, ī	u, ū
	e, ē	o, ō
	ā, ā̄	a, ă

The vowel /ā/ is low and central, while /a/ ranges from the low, back [a] to [ʌ] and [ə]. The vowel /o/ has a wide range of phonetic variants, some of them predictable, including

[wa], [wo], [o] and [?o]. The phoneme /e/ also has a range of predictable variants: [e], [ɛ], [ʌ], [je].

There are some interesting patterns of vowels in combination. In a sequence of two identical high vowels across a morpheme boundary, one vowel elides, e.g. *bi-i* → [bi] ‘give-INF’. By contrast, if two repetitions of /e/ come together across a morpheme boundary, both vowels are pronounced, and each has a distinct pulse, indicating two syllables: *ye-e* [ye.e] ‘come-NR2’. However if the /e/ of the suffix is followed by another element, then the /e/ changes to [ɛ]: *ye-eu* [ye.eu] ‘come-3SFUT’. When a suffix beginning with /e/ follows another vowel, disyllabic sequences result: *tā-eu* [tā.eu] ‘hear-3SFUT’. If the preceding vowel is high, then a glide is inserted: *si-eu* [si.yeu] ‘die-3SFUT’.

The three prefixes of the language all exhibit vowel harmony. The harmony patterns do not seem to be regularized among the community. The most elaborate pattern found among the three speakers I have worked with in depth is illustrated here with the negative prefix *ma-*:

(3)	<i>ma-ta-u</i>	NEG-put-3SPAST
	<i>ma-ye-ŋ</i>	NEG-bring-3SPAST
	<i>ma-tul</i>	NEG-fall.3SPAST
	<i>ma-sit</i>	NEG-die.3SPAST
	<i>mā-tā-u</i>	NEG-hear-3SPAST
	<i>mo-cō</i>	NEG-stay.3SPAST
	<i>mo-tul</i>	NEG-fall.3SPAST
	<i>mwā-mwāl</i>	NEG-search.3SPAST

Consonant clusters with the glide /y/ block the harmony process, thus *ma-ŋyāt* [NEG-buy.3SPAST] ‘don’t buy!’.

Dolakhae does not have tone. Words pronounced in isolation have stress on the initial syllable of a root; however, this pattern is overridden by intonational contours in connected speech (Genetti 2007: 64, 68).

3 MORPHOLOGY

3.1 Nouns and pronouns

Dolakhae has separate pronouns for first (*ji*), second (*chi*) and third (*ām*) person. An inclusive/exclusive distinction is found with first person plural; compare *isi* ‘we’ (exclusive) with *thiji/chiji* ‘we’ (inclusive). The latter is transparently formed from a combination of the first- and second person pronouns. There is also a second person honorific pronoun *thamu*. Second- and third person plurals are formed with the plural morpheme *-pen*; they are *chipen* and *āpen* respectively. Case inflection of pronouns involves idiosyncratic morphophonemics, although the case elements are always transparent. Pronouns in the accompanying text are found in <55, 69, 70, 75, 76>.

Nouns and pronouns inflect for number and case. Number marking is limited to a single morpheme: *-pen*. This morpheme can be used to indicate plurality and it can also be used with a collective sense <51>. There is very little nominal derivational morphology in Dolakhae. A single diminutive suffix *-cā* is used to indicate the young of animals, e.g. *sā* ‘cow’, *sā-cā* ‘calf’. (The use of this suffix is more widespread in the Kathmandu dialect, where it is a general diminutive.) There is no nominal marking for gender.

The Dolakhae casemarkers are morphological clitics (except for the associative, which is a postposition). They are listed in (4):

(4)	Ergative	= na, = n	<28, 29, 55, 61, 69>
	Instrumental	= na, = n	<6, 26, 27>
	Dative	= ta	<33, 54, 69>
	Genitive	= e	<43, 64, 75>
	Locative	= ku	<13, 17, 20, 31, 32, 70>
	Associative	nāpa	
	Allative	= ke	
	Ablative	= lān	<14, 30>

The use of the ergative is obligatory on all subjects of transitive verbs. There is no split ergativity, by aspect, person or NP type, and neither does ergative casemarking depend on semantic or discourse-pragmatic patterns. Instead, all Dolakhae verbs are classified according to whether they are transitive (may take an object) or intransitive (may not), and subjects of all transitive verbs are ergative. Instrumental case is syncretic with the ergative.

The dative marks all recipients of ditransitive verbs. It is also found on patient arguments of transitive verbs which are human and whose referent is given or accessible. Patients of transitive verbs and recipients of ditransitive verbs are grammatical objects. The dative is also found marking subjects but only with dative-subject predicates borrowed from Nepali (e.g. <33–34> Nepali *samjhanu* ‘remember’), or calques on Nepali dative-subject predicates using Dolakhae vocabulary. In addition, the dative is used on objects of exchange and on benefactive objects. In the latter case, the predicate must contain the auxiliary verb *bir-* ‘give’.

The oblique cases require little comment. The locative and allative differ primarily in that the allative references an animate goal. Inanimate goals are marked with the locative.

3.2 Numerals and classifiers

There are both decimal- and vigesimal-based systems of numerals. The systems are identical up until thirty-nine. The numerals for one to twenty are given in (5). The suffix *-gur* is the general numeral classifier that is always used when counting.

(5)	<i>thi-gur</i>	‘one’	<i>j(h)im-thi-gur</i>	‘11’
	<i>nis-gur</i>	‘two’	<i>j(h)im-nis-gur</i>	‘12’
	<i>sō-gur</i>	‘three’	<i>j(h)im-sō-gur</i>	‘13’
	<i>pe-gur/ pē-gur</i>	‘four’	<i>j(h)im-pe-gur</i>	‘14’
	<i>ŋā-gur</i>	‘five’	<i>j(h)im-ŋā-gur</i>	‘15’
	<i>khu-gur</i>	‘six’	<i>j(h)im-khu-gur</i>	‘16’
	<i>nas-gur</i>	‘seven’	<i>j(h)im-nas-gur</i>	‘17’
	<i>cyā-gur</i>	‘eight’	<i>j(h)im-cyā-gur</i>	‘18’
	<i>gu-gur</i>	‘nine’	<i>j(h)im-gu-gur</i>	‘19’
	<i>j(h)i-gur</i>	‘ten’	<i>ni(i)-gur</i>	‘20’

‘Thirty’ is *swi*. The decimal and vigesimal systems diverge in the multiples of ten from ‘40’ to ‘90’:

(6)	<u>Decimal</u>	<u>Vigesimal</u>	
	<i>pi-gur</i>	<i>ne-ni-nis-thā̄i</i>	‘40’
	<i>ŋāi-gur</i>	<i>ne-ni-nis-thā̄i-o-j(h)i-gur</i>	‘50’
	<i>khwi-gur</i>	<i>ne-ni-sō-thā̄i</i>	‘60’

<i>nai-gur</i>	<i>ne-ni-sō-thā̄i-o-j(h)i-gur</i>	‘70’
<i>cyai-gur</i>	<i>ne-ni-pi-thā̄i</i>	‘80’
<i>gwi-gur</i>	<i>ne-ni-pi-ṭhā̄i-o-j(h)i-gur</i>	‘90’

The decimal system is constructed by adding a simple *-i* suffix to the numeral. This can then be followed by another numeral indicating the one’s place, e.g. *ya-i-pe* ‘54’. The vigesimal system takes as its base the sequence *ne-ni*, which speakers translate as ‘twenty-twenty’. These collocations end in the noun *ṭhā̄i* ‘place’, which replaces the classifier, so ‘40’ and ‘60’ translate as ‘twenty-twenty-two-places’ and ‘twenty-twenty-three-places’ respectively. To designate the odd multiples of ten, *o(j)im* ‘and ten’ is added; one can also add 11 to 19, e.g. *ne-ni-pi-ṭhā̄i-o-j(h)im-ŋā-gur* ‘95’. As these collocation end in numerals, the classifier is again required.

‘Hundred’ is indicated by the mensural classifier *sar* – e.g. *nas-sar* ‘700’ – and ‘thousand’ is indicated by the mensural classifier *dol* (*nis-dol* ‘2,000’).

Dolakhae numerals never occur as bare stems, but are always followed by a bound classifier. The numeral–classifier combination generally precedes the modified noun. There are 30 classifiers that I know of. Classifiers only occur with numerals; they are not found with demonstratives or other nominal modifiers. Some classifiers are transparently derived from nouns, but none are identical to nouns which they classify (i.e. there are no ‘echo classifiers’ in the language). The majority of members of a class are usually amenable to semantic classification. A full description of the classifiers can be found in Genetti (2007: 220–7). Of particular interest and centrality are *-mā* for animals, *-gar* for roundish things, *-kā* for implements, *-pā* for relatively rigid long things, *-pu* for relatively flexible long things, *-darpā* for bowls of yogurt, *-kto/koto* for pieces or chunks of food, and *-gur* as a general classifier. Auto-classifiers include *-nu* for days, *-chi* for nights, *-lā* for months, *-da* for years, and *-oti* for words. Classifiers in the text are found in <1, 2, 16>.

3.3 Adjectives and deverbal modifiers

Most modifiers of nouns that denote property concepts (those commonly coded by a lexical class of adjectives in the world’s languages) are derived from verbs in Dolakhae. For example *bāla-ku* ‘beautiful’ is a nominalized form of the verb *bālat-* ‘become beautiful’ <75>. Similar examples are *hāka-u* ‘black’ from *hākar-* ‘become black’ and *gāy-gu* ‘dry’; compare *gān-a* ‘it dried’. The suffix used in the derivation of nominal modifiers is NR1 (see section 5.3), the form used to relativize verbs when the head noun is coreferential with the relative clause subject. Thus all deverbal modifiers may be analysed syntactically as relative clauses, e.g. *tuj-gu guābā* ‘ripe guava’; ‘guava that has ripened’; *hēga-u parsi* ‘red sari’; ‘sari that is red’.

Although the majority of property concepts in Dolakhae are deverbal, there are a small number of true adjectives as well, that is, property concepts that form a separate lexical class and do not have the inflectional possibilities of either nouns or verbs. These include *lyās-mā/lyās(i)-misā* ‘young (male/female)’ <33>, *nullu* ‘new’, *wē/wini* ‘crazy (male/female)’ and at least 15 others (Genetti 2007: 207). When used attributively, these adjectives precede the noun, although *lyās-mā* and *lyās(i)-misā* may be used as nominal compounds, hence function referentially, e.g. <33>. When used predicatively, these adjectives are followed by one of the copulas or the verb *jur-* ‘become’, depending on tense-aspect.

3.4 Clitics and particles with discourse function

Dolakha Newar has five morphemes – three particles and two clitics – that together form a central subsystem of discourse structuring. The particles are phonologically independent and can be stressed; the clitics are unstressed and phonologically bound to their host. The five morphemes are in rough paradigmatic alternation and do not co-occur.

The topic particle *wā* follows a noun phrase to highlight a referent as being the entity to which the clause especially pertains, e.g. *āu wā mo-oŋ-gi. libi* [now TOP NEG-go1SPAST later] ‘Now I didn’t go. (I’ll go) later’. This morpheme also occurs after temporals and non-finite clauses <15, 19, 61>. The focus particle *tuj* marks elements that constitute new and important information that is explicitly profiled against knowledge that is presupposed. For example, in the phrase *nāpā tuŋ ū-i ale* [together FOC go-1FUT then] ‘let’s go together then’, there is a presupposition that the two referents of the subject will be going; the new focused information is that they will go together as opposed to separately. The third particle in this set is rather ordinary: *jukun* ‘only’.

The two clitics have nearly opposite functions and form an interesting and unusual grammatical subsystem.

The clitic =(*u*)*ri* has an individuating function, so specifies one element out of a group that it contrasts with, e.g. *thābi* = *ri* ‘the top one’ (<*thābi* ‘top’); *thi-mā* = *ri* ‘one of the group’. This clitic is commonly used with pronouns and kin terms <76>, in which case it may precede the casemarker, e.g. *kehē* = *uri* = *n* ‘younger sister (ERG)’. Interestingly, there are cases where this clitic occurs twice in a single word, on both sides of the casemarker: *mā* = *uri* = *n* = *uri*. ‘mother (ERG)’. Presumably, the first occurrence represents a lexicalization of the kin term and clitic, and the later occurrence has discourse-pragmatic force.

In contrast to =(*u*)*ri* is =(*u*)*ŋ*, which delineates a group of which the host is part; it is glossed as ‘extensive’. When bound to a noun phrase with a singular referent, it translates as ‘also’, e.g. *ji=ŋ ū-i* [1s=EXT go-1FUT] ‘I also will go’. When bound to a numeral, it translates as ‘both’ or ‘all’; compare *nis-mā* ‘two’ with *nis-mā* = *ŋ* ‘both’. It also occurs with adverbials, where its interpretation is variable. Common examples are *lita* = *ŋ* ‘again’ (from *lita* ‘next’) and *ānθi* = *ŋ* ‘just like that’ <33> (from *ānθi* ‘like that; in that manner’).

4 VERBS

4.1 Copula and existential

The copula in Newar is *khyāŋ*, while the intransitive existential verb is *dam(u)*. These verbs are morphologically irregular and independent of the four inflectional verb classes. The existential *dam* has a number of finite and non-finite inflectional forms. Especially common are *da-u*, the past stative form, and *ma-da*, the negative. The copula *khyāŋ* may also be pronounced *khyāū*; the negative is *ma-khe*.

The primary function of *dam* is existential, as in *gār dam* ‘there is a wound’. It is also used in locational constructions with a locative noun phrase: *gār lāhā* = *ku dam* ‘the wound is on the arm’. Note that when the location is animate, the verb *con-* ‘stay’ is used <17>. The existential is also used in possessive constructions in which the possessor is realized as a genitive-marked dependent of the possessed noun, thus *jana mica dam* ‘I have a daughter’, where *jana* is the genitive form of the first person pronoun; literally the meaning is ‘my daughter exists’.

The copula *khyaj* is used equationally, thus *misā ma-khe* ‘this is not a woman’, *twāysona sona khyaj* [rhododendron flower cop] ‘the rhododendron is a flower’. The negated form *ma-khe* is often used as a strong denial, meaning ‘it is not so! And with the question particle, the form *khe rā?* is commonly used to create a tag, i.e. ‘isn’t it so?’

4.2 Inflectional verb classes

With the exception of the two verbs just described, all Dolakhae verbs belong to one of four inflectional verb classes. The four classes are most simply characterized by the stem-final consonant which occurs in the third person singular past form of the verb. The four classes are thus n-stems (*on-a* ‘went’), r-stems (*yer-a* ‘came’), l-stems (*lipul-a* ‘returned’) and t-stems (*gyāt-a* ‘feared’). Each stem class has distinctive morphophonemic variation that is regular for all members of the class.

Verbs borrowed from Nepali are incorporated into the inflectional system in two ways. Transitive verbs are put in stem form and followed by the verb *yet-* ‘do’, e.g. *tār yet-* ‘to cross’ (from Nepali *tār-nu* ‘cross (river)’). Intransitive verbs are suffixed by *-ai*, glossed as bvs for ‘borrowed verb suffix’, then followed by the intransitive verb *jur-* ‘become’; ‘happen’ <31, 70>.

4.3 A complete verb paradigm

A paradigm of the affirmative finite verb *nar-* ‘eat’ is given in Table 22.1. Only one affirmative finite suffix is not illustrated here; it is the third person singular past suffix *-a*, which is found on intransitive verbs, as opposed to *-cu/ju*, which is found on transitive verbs.

The Dolakhae finite verb inflects for four tenses: the past anterior, simple past, present, and future. Tense markers follow the verb directly. The past anterior is *-u*, the past has no marker, the present is *-a*, and the future is *-i/-e*. The verb also reflects the person (first, second, third) and number (singular, plural) of the subject. There is no separate verbal inflection to distinguish inclusive and exclusive forms of the first person plural; this distinction is maintained only in the pronouns (see above). Second person honorific forms are identical to those of the first person plural, except in future tense.

In addition to the affirmative, the Dolakhae verb has separate inflectional paradigms for the negative, imperative, prohibitive and optative. The negative has separate but related agreement morphology, but only inflects for three tense categories, the past and present not being distinguished. The optative and prohibitive are marked by prefixes, *tha-* and *da-* respectively. The verbs in these moods and the imperative also inflect for number, each category demonstrating slightly different inflectional patterns.

TABLE 22.1 AFFIRMATIVE PARADIGM OF *NAR-* ‘EAT’

	Past Anterior	Past	Present	Future
1s	<i>na-u-ī</i>	<i>nar-gi</i>	<i>nar-a-gi</i>	<i>na-i</i>
1p	<i>na-u-pe</i>	<i>nar-gu</i>	<i>nar-a-gu</i>	<i>na-i</i>
2s	<i>na-u-n</i>	<i>nar-mun</i>	<i>nar-a-n</i>	<i>na-i-na</i>
2p	<i>na-u-min</i>	<i>nar-min</i>	<i>nar-a-min</i>	<i>na-i-nan</i>
2hon	<i>na-u-pe</i>	<i>nar-gu</i>	<i>nar-a-gu</i>	<i>na-i-ta</i>
3s	<i>na-u(-ju)</i>	<i>nar-ju</i>	<i>nar-a-i</i>	<i>na-e-u</i>
3p	<i>na-u-tan</i>	<i>nar-hin</i>	<i>nar-a-hin</i>	<i>na-e-u</i>

There are four non-finite forms of the Dolakhae verb. Illustrating again with *nar-* ‘eat’, they are the infinitive (*na-i*), the participle (*na-en*) and the two nominalizers, NR1 (*na-u*) and NR2 (*na-e*). The functions of these forms are discussed in section 5.3.

4.4 Auxiliary verbs

There are a number of verbs which may be analysed as auxiliaries, and which convey aspect, direction or the presence of a benefactive argument in the clause. While a few of these auxiliaries follow a main verb in the infinitive (e.g. *na-i don-ju* ‘finish eating’, *na-i ten-a-gi* ‘about to eat’), the most common auxiliaries follow a verb in participial form. The three which occur most frequently are *con-* ‘sit’, ‘stay’, ‘reside’, which conveys a progressive aspect <18, 75>, *tal-* ‘put’ which indicates an extended state or perfect aspect, and *bir-* ‘give’, which indicates the presence of an affected participant (beneficiary or maleficiary); the NP that refers to the participant may or may not be a grammatical argument of the clause; thus, this is not a strict applicative construction.

5 SYNTAX

5.1 Word order

A noun phrase may consist of the following elements: demonstrative <33, 54>, adjective/relative clause <75>, genitive modifier <3, 11, 12, 43, 65, 75>, quantifier <16>, noun.

Demonstratives are generally phrase-initial <33, 54>. Prenominal demonstratives and independent demonstratives are of identical form. Genitive NPs and relative clauses precede the noun <43, 75>. Nouns are generally final in the phrase, although quantifiers may be positioned postnominally. Head nouns may be omitted if reference is clear from context. The clitic postpositions are bound to the last element in the phrase, even if not nominal.

Dolakha Newar is clearly a verb-final language, although it is possible to postpose arguments or adverbials after a finite verb. When this happens, these elements are in separate intonation units <13, 14>. No postverbal positioning of elements is found with non-finite verbs. Nominal arguments are often omitted if clearly understood from context <11, 12, 15, 19, etc.>, or are of very general or non-specific reference. When subject and object NPs are both present, it is more common for the subject to precede the object <69>. The opposite order is also attested, typically when the object noun phrase is the discourse topic and/or the subject noun phrase is the focus. Adverbial modifiers always precede the verb they modify.

5.2 Grammatical relations

There is clear evidence for subject and object as grammatical categories in Dolakhae. The subject may be readily identified as the NP which controls verb agreement. Also, the suffix glossed NR1 is used when the head of a relative clause is coreferential with the subject (A or s) of the relative clause. There are no passive constructions which derive subjects in this language. Objects are arguments which: (a) are not subjects; (b) are subcategorized for by the semantic structure of transitive and ditransitive verbs; (c) may be marked with the dative = *ta*; and (d) take the suffix NR2 when relativized. There is no evidence that indirect objects and direct objects form distinct syntactic classes (Genetti 1997).

Transitivity is a strong organizational principle in Dolakhae syntax. Any verb which takes a notional object is grammatically transitive, regardless of whether or not the object

referent is important or specified as an NP. Thus *āmun jā nar-ju* ‘he ate the rice’ and *āmun nar-ju* ‘he ate’ are both grammatically transitive, as can be seen from the ergative form of the third person singular pronoun *āmun*, and the consistent use of the transitive 3s suffix *-ju*; with intransitive verbs, the suffix is *-a*. There is no way to detransitivize a transitive clause, although there are a number of transitive/intransitive verb pairs, the transitive containing an aspirated initial, the intransitive containing a voiced initial, e.g. *gyāt-/khyāt-* ‘be afraid’/‘scare’. (The aspirated form is clearly a reflex of the old causative **s-* prefix reconstructed for Proto-Tibeto-Burman, see Benedict 1972: 105.) Intransitive verbs may be made transitive only through the addition of the causative suffix *-ker*.

5.3 Clause combining

One of the most interesting areas of Dolakhae grammar is the manner in which clauses are woven together to form long sentences of intricate structure (Genetti 2011a). Sentence breaks in narrative discourse may be identified by the presence of a non-embedded finite verb, sometimes followed by particles, and typically produced with a rising or high level intonation contour. The final clause usually ends in a verb with finite morphology and is produced with falling intonation. Example (7) is illustrative of sentence complexity; note that quoted speech is shown in bold:

- (7) *dwā-ku = ri iri oho: | ji yā githi yer-i,*
 senior-NR1 = IND daughter.in.law EXCL 1s EMPH how do-1FUT
lās cār-a-gi hay-an, ithi: yen-an, mikhā ti-en
 shy feel-PR-1s say-PART like.this do-PART eyes close-PART
on-yasin; dirtarāstra janm-ai jur-a.
 go-when Dhirtarāsta be.born-BVS be-3SPST
 ‘The eldest daughter-in-law, saying “Oh, how shall I do this? I am feeling shy”,
 and going like this, when she closed her eyes and went, Dhirtarastra was born.’

The three underlined italic verbs in the sentence have the ubiquitous participial suffix, which may be considered a ‘converb’ similar to those described for Altaic languages (e.g. Johanson 1995; for a full analysis of the construction as a converb, see Genetti 2005). In these examples, the suffix links non-final clauses in a chain with a final clause. If the final clause is finite, chained clauses will depend on it for the specification of tense. Functionally, participial clauses either list events in a causal or temporal sequence (as in the first example above), and/or give adverbial modification to a following verb (as in the second example). These functions are often difficult to differentiate, as the line between them is not clear <e.g. 1, 4>. Note that the first of the chaining verbs, *hay-an* in (7) earlier, is a quotative. It functions here as a lexical predicate, but this verb form may also occur as a quotative complementizer <57, 58>. The embedded direct quote is a syntactic object (Genetti and Slater 2004), thus entire sentences may be embedded within sentences in this construction. (For prosodic features of embedded speech, see Genetti 2011b.)

In the last line of (7), the verb *on-yasin* carries a temporal suffix that creates an adverbial clause. Adverbial suffixes are distinct from the participial suffix because of their semantic specificity; they denote temporal or logical relationships which hold between the adverbial clause and the following clause <15, 19, 20, 38, 55, etc.>. Genetti (2007: 483–4) compares and contrasts the syntax of adverbial and participial clauses.

Returning to example (6), note that the subject of the chain, *iri* in the first line, is not marked with the ergative, so is presumably casemarked for the verb *on-yasin*, the final verb in the series and the only intransitive one. The final and finite clause of the sentence has a new subject, Dhirtarastra.

Dolakhae has several types of complementation. Verbs in complement clauses are suffixed with either the infinitive or one of the nominalizers, e.g. *na-i mal-a* ‘must eat’ (with the infinitive), *na-e khon-ju* ‘saw him/her eat’ (with NR2) <55, 70>. Complements of cognition verbs have an abstract noun incorporated into the structure, which may be analysed either as a nominal constituent of the higher clause, or as a complementizer (e.g. *khā* ‘talk’, ‘matter’, *saykā* ‘suspicion’). The complement verb takes the infinitive plus the emphatic = *uri* if the clause is unrealis, or NR2 if the clause is realis. Thus the phrase *na-i = uri khā*, with the infinitive, may be roughly translated as ‘the fact that I’ll eat’, while *na-e khā*, with NR2, means ‘the fact that I ate’. The realis/unrealis distinction is not overtly reflected elsewhere in the morphology of the language.

The two suffixes which are used in nominalized and relative clauses are glossed NR1 and NR2. NR1 has the forms *-gu*, *-ku*, or *-u*, while NR2 is pronounced *-a* or *-e*. The allomorphy depends on the inflectional class of the verb. While in some languages of the Tibeto-Burman family relative clause structures might be argued to be syntactic nominalizations which are in an appositive relationship with the head noun, in Dolakhae the two structures are syntactically distinct (see Genetti *et al.* 2008: 127). The two suffixes NR1 and NR2 are in a paradigmatic relationship, and the distribution of these suffixes *vis-à-vis* each other is complicated and varies by speaker (Genetti 2007: 403–7). A few general patterns can be mentioned here. When these two suffixes are used to form relative clauses, NR1 is used when the head noun is coreferential with the subject of the relative clause, while NR2 is used when the head noun is coreferential with the object. When the head noun is coreferential with an oblique NP, either suffix can be used, and speakers seem to base their judgments on the person of the subject or the tense of the clause. When the two suffixes are used to nominalize complements of perception verbs, NR1 is preferred when the verb is intransitive, and NR2 when the verb is transitive. And in nominalized questions or emphatic constructions, NR1 is the preferred form when the subject is third person, and NR2 when the subject is first or second.

5.4 Other constructions

The comparative in Dolakhae is a bit unusual. The object of comparison appears in the absolute (unmarked case), followed by *so-en*, the participial form of the verb *sor-* ‘look’. This is followed by the attributive clause, e.g. *ji so-en chi bāla-ku khyāñ* ‘You are more beautiful than me’. The construction is not a simple case of parataxis. The verb *sor-* normally takes an ergative subject and dative-marked object. The object of comparison, by contrast, is never casemarked. Thus the construction has taken on idiosyncratic properties in the process of grammaticalization.

The reflexive is formed with the reflexive pronoun *thau*. This pronoun necessarily refers back to the subject of the clause. It is most commonly used with a possessive sense, e.g. *āpen thau chē oj-an ...* ‘they went to their own house’. The reflexive pronoun may be repeated to convey a distributive meaning, e.g. *chipe thau thau chē o-n* ‘you all go each to your own houses’. The reflexive is generally not used to indicate objects being coreferential with subjects, paratactic expressions being preferred for such cases. The reflexive pronoun can also be used pronominally, if it is coreferential with the subject of the preceding clause.

6 NARRATIVE TEXT

The following text is a portion of a story told by Mrs Sanu Laxmi Joshi in Dolakha, in January 1989. A small portion has been ellipted, as it contains a summary of the preceding text. The story has been divided into prosodic units, and each unit is set on a separate line and marked for its intonation type by line-final punctuation. The six intonation types are marked by the following conventions:

- (8) . prototypical final, generally falling pitch
 | narrative final intonation, high-level pitch throughout unit
 ? interrogative final, high rise in pitch at end of unit
 ! exclamatory final, steep rise-fall of pitch on final syllable
 ; anticipatory continuing, strong rise on nucleus of final syllable
 , non-anticipatory continuing, some rise in pitch on final syllable

In addition to breaking the text into prosodic units, with one on each line, the text has also been broken into sentences, based on morphosyntactic and prosodic cues. Embedded quotation is in bold to aid readers in understanding the different levels of structure.

The story concerns Viśnu, who is visiting Śiva and Parvati. In an earlier portion of the story, Śiva has expressed interest in seeing Viśnu in the guise of a beautiful young maiden, which Viśnu had earlier used to deceive demons. Viśnu has asked Śiva if he would be able to resist falling in love with the young maiden who is actually Viśnu in disguise. Śiva has promised three times that he would.

6.1 Sentence 1

- 1 *nis-nu so-nu li,*
 two-day three-day after
- 2 *so-nu li;*
 three-day after
- 3 *sora barsa = e umer,*
 sixteen year = GEN age
- 4 *yārling phi-en;*
 earring put.on-PART
- 5 *sāri dalkyāñ-an;*
 sari wear-PART
- 6 *gājal = na uj-an;*
 eyeliner = INST draw-PART
- 7 *sā þhu-en;*
 hair braid-PART
- 8 *bakundo u-en;*
 ball dance-PART
- 9 *harararararañ,*
 EXP
- 10 *bisnu bāgabāñ on-a.*
 Viśnu god go-3SPAST

‘Two or three days later, three days later, as a sixteen-year-old, putting on earrings, wrapping a sari, applying eyeliner, and dancing with a ball, *hararararay*, God Viśnu went.’

6.2 Sentence 2

- 11 *mohini = e abatār kā-en;*
maiden = GEN form take-PART
- 12 *sibaji = e bagainca on-a.*
Śiva = GEN garden go-3SPAST
- 13 *kailās tarpha-ku.*
Kailaś direction = LOC
- 14 *baikuntha = lān.*
Baikhuntha = ABL

‘Taking the form of a young maiden, he went to Śiva’s garden. In Kailaś. From Baikuntha.’

6.3 Sentence 3

- 15 *āle āku on-γasin wā;*
then there go-when TOP
- 16 *nis-mā siba pārbati,*
two-CL Śiva Pārbati
- 17 *atāli = ku coŋ-an;*
balcony = LOC stay-PART
- 18 *musukka yil-en coŋ-gu.*
smiling smile-PART stay-3SPH

‘Then when he went there, the two of them, Śiva and Parbati, sitting on the balcony, were smiling prettily.’

6.4 Sentence 4

- 19 *āle caṭṭa mikhā tar-γasin wā;*
then EXP eye put-when TOP
- 20 *gulpa = γ lāhā = ku bokunda metha-e-thē,*
when = EMPH hand = loc ball play-NR2-as.if,
- 21 *gulpa = γ bokunda metha-e-thē,*
when = EMPH ball play-NR2-as.if
- 22 *sārpuli halyāŋ-an;*
ribbons swing-PART
- 23 *yārlīŋ dalkyāŋ-an;*
earring wear-PART
- 24 *mikhā tarkyāŋ-an;*
eye flirt-PART

- 25 *miu pita kā-en;*
?? out take-PART
- 26 *ebam prakār = na pyākhan u-en;*
every type = INST dance dance-PART
- 27 *camcamcamcam = na yer-ŋasin;*
EXP = INST come-when
- 28 *caṭṭa sibaji = n khoy-a ju-en con-a.*
EXP Śiva = ERG see-NR2 be-PART stay-3SPAST

‘Then when he put his eyes on her, at one time as if she were playing with a ball in her hand, at another time again as if playing with a ball, her ribbons swinging, her earrings dangling, her eyes flirting, taking out [? unknown], dancing every type of dance, when she came with a *camcamcamcam*, Śiva saw her.’

6.5 Sentence 5

- 29 *sibaji = n khō-i-sāt;*
Śiva = ERG see-INF-as.soon.as
- 30 *sibaji puklukka aṭāli = lān,*
Śiva EXP balcony = ABL
- 31 *kotha bō = ku jar-ai ju ju-en con-a.*
room ground = LOC fall-BVS be.NR1 be-PART stay-3SPAST

‘As soon as he saw her, Śiva fell – *puklukka* – from the balcony onto the floor of the room.’

6.6 Sentence 6

- 32 *bō = ku ye-i doy-an li;*
ground = LOC come-INF finish-PART after
- 33 *āle āmu bisnu bāgabān = ta lyāsmisā = ta = y;*
then that Viśnu god = DAT young.woman = dat = EXT
- 34 *saṃj-ai ju-en,*
remember-BVS be-PART
- 35 *tin lok;*
three realms
- 36 *caudha bhuban;*
fourteen worlds
- 37 *ḍulyeṇ-an;*
lead-PART
- 38 *jō-i-thē jō-i-thē ju-i-ho,*
catch-INF-as.if catch-INF-as.if be-INF-when
- 39 *ŋil-en;*
smile-PART

- | | | | |
|----|-----------------|---------------|--|
| 40 | <i>phisikka</i> | <i>ŋil-en</i> | <i>thi-khana,</i> |
| | EXP | smile-PART | one-side |
| 41 | <i>phisikka</i> | <i>ŋil-en</i> | <i>thi-khana,</i> |
| | EXP | smile-PART | one-side |
| 42 | <i>phisikka</i> | <i>yil-en</i> | <i>on-ŋa</i> <i>on-ŋa</i> <i>on-ŋa;</i> |
| | EXP | smile-PART | go-when go-when go-when |
| 43 | <i>āme</i> | <i>jati</i> | <i>sibaji = e</i> <i>jati,</i> |
| | 3SGEN | possessions | Śiva = GEN possessions |
| 44 | <i>swāl</i> | <i>hā-u,</i> | |
| | EXP | fall-NR1 | |
| 45 | <i>bijan,</i> | | |
| | seed | | |
| 46 | <i>lū,</i> | | |
| | gold | | |
| 47 | <i>oho,</i> | | |
| | silver | | |
| 48 | <i>liuri,</i> | | |
| | bronze | | |
| 49 | <i>pitāl,</i> | | |
| | copper | | |
| 50 | <i>ānāgu</i> | <i>ānāgu,</i> | |
| | like.that | like.that | |
| 51 | = <i>pen</i> | <i>ju-en</i> | <i>dhātu = pen</i> <i>khas-ai</i> <i>ju-en</i> <i>ðŋ-gu</i> <i>ju-en</i> <i>con-a </i> |
| | = PL | be-PART | metal = PL fall-BVS be-PART go-NR1 be-PART stay-3SPAST |

'Then after he fell to the ground, he again remembered Viśnu, the maiden, and she led him through the three realms (heaven, earth, hell) and the fourteen worlds, and when he made as if to catch her, she would smile, smiling on one side, smiling on the other side, going on and on and on smiling, his possessions all fell from him, Śiva's possessions, his seed, gold, silver, bronze, copper, all things like that, metals, they all fell away from him.'

6.7 Sentence 7

- | | | | | | | | |
|----|---------------|------------------|--------------------|-----------------|------------------|------------|-------------------|
| 52 | <i>āle</i> | <i>tin</i> | <i>lok,</i> | | | | |
| | then | three | realms | | | | |
| 53 | <i>caudha</i> | <i>bhuban</i> | <i>duler-i</i> | <i>doy-ani,</i> | | | |
| | fourteen | world | lead-INF | finish-PART | | | |
| 54 | <i>guli</i> | <i>u</i> | <i>sibaji = ta</i> | <i>dulet-ki</i> | <i>dukha</i> | <i>bi </i> | |
| | how.much | this | Siva = DAT | lead-1SPAST | trouble | give.INF | |
| 55 | <i>'chin</i> | <i>māyā = ku</i> | <i>jit-ai</i> | <i>ju-i</i> | <i>phar-a-gu</i> | <i>rā'</i> | <i>hat-yasin,</i> |
| | 2SERG | love = LOC | win-BVS | be-INF | able-PR-2h | Q | say-when |

- 56 '*ma-pha-gi*' |
 NEG-able-1s
- 57 *hay-an* *ha = ku.*
 say-PART say-2hPAST
- 58 '*phar-a-gi*' *hay-an* *ha-ku.*
 able-PR-1s say-PART say-2hPAST
- 59 *āu* *ākhiri = ku* *ota* *hay-anē,*
 now end = LOC this say-PART
- 60 *phyāṭṭa* *ithi,*
 EXP like.this
- 61 *sibaji = n* *jō-i-sāt* *wā,*
 Śiva = ERG catch-INF-as.soon.as TOP
- 62 *simā = e* – (false start)
 tree = GEN
- 63 *lyāsimisā* *hay-a* *jon-ηasin;*
 young.woman say-NR catch-when
- 64 *simā = e* *thuthā* *jur-a.*
 tree = GEN log become-3SPAST

‘Then, having led him through the three realms and fourteen worlds, Viśnu said to himself, ‘I led this Śiva so much to give him trouble. When I said ‘Can you win in love?’, he said ‘I cannot’, he said ‘I can’.’ When he spoke like this in the end, suddenly like this, as soon as Śiva caught him, when he caught the one said to be a young woman, he transformed himself into a log.’

6.8 Sentence 8

- 65 *simā = e* *thuthā* *ju-i-ho,*
 tree = GEN log become-INF-when
- 66 *lās* *cār-a.*
 embarrassment feel-3SPAST

‘After (Viśnu, disguised as a young maiden) became the log of a tree, (Śiva) became embarrassed.’

6.9 Sentence 9

- 67 *sibaji* *lās* *cār-a;*
 Śiva t embarrassment feel-3SPAST
 ‘Śiva was embarrassed.’

6.10 Sentence 10

- 68 *āle* *lās* *cā-enī;*
 then embarrassment feel-PART

69	<i>jin</i>	<i>thaeta</i>	<i>pahilā</i>	<i>tuy</i>	<i>ha-ku-i</i>	<i>le.</i>
	1SERG	2hDAT	before	TOP	say-PA-1S	PRTCL
70	<i>'thamu</i>	<i>māyā = ku</i>	<i>jit-ai</i>	<i>ju-i</i>	<i>phar-a-gu</i>	<i>rā </i>
	2HON	love = LOC	win-BVS	happen-INF	able-PR-2h	Q
71	<i>ma-phar-a-gu</i>	<i>rā'</i>	<i>hay-an</i>	<i>'phar-a-gi'</i>	<i>ha-kupe.</i>	
	NEG-able-PH-2h	Q	say-PART	able-PR-1S	say-PI-2h	
72	<i>gwā</i>	<i>le</i>	<i>hat-cu </i>			
	where	PRTCL	say-3SPAST			

‘He became embarrassed, then Viśnu said: “I told you before, ‘Are you able to win in love or not?’ and you said: ‘I am able to’. Where is your ability then?”’

6.11 Sentence 11

73	<i>āle</i>	<i>ānthi</i>	<i>har-i-ho;</i>
	then	like.that	say-INF-when
74	<i>phisikka</i>	<i>pārbati = y</i>	<i>yil-en,</i>
	EXP	Parbati = TOP	smile-PART
75	<i>ji</i>	<i>ulistule</i>	<i>bāla-ku</i>
	1s	like.this	beautiful-NR1
	<i>pārbati</i>	<i>thae</i>	<i>sāt-ku</i>
	Parbati	2hGEN	side = LOC
			stay-PART
76	<i>thamu = ri</i>	<i>ānthi = y</i>	<i>yil-a-gu</i>
	2h = TOP	like.that = EMPH	smile-PR-2S
			Q

‘Then when he spoke like that, smiling Parbati also smiled and said “I, Parbati, who am beautiful like this, sit at your side. You smile in just that manner (at some other girl)?”’

(Small portion ellipted here.)

6.12 Sentence 12

77	<i>sibaji</i>	<i>summake</i>	<i>lās</i>	<i>mā-cā-en –</i>
	Śiba	silent	embarrassment	NEG-feel-PART
78	<i>lās</i>	<i>cā-en</i>	<i>con-a.</i>	
	embarrassment	feel-PART	stay-3SPAST	

‘Siva silently sat feeling embarrassed.’

ADDITIONAL ABBREVIATIONS

1,2,3	person indices
s,p,h	singular, plural, honorific
BVS	borrowed verb suffix
EXP	expressive vocabulary
EXT	extensive
IND	individuating
NR1~2	nominalizer/relativizer 1 or 2
PA	past anterior
PART	(conjunctive) participle / converb

NOTE

- 1 Funds for this study were provided by the National Science Foundation, grant BNS-8811773.

REFERENCES

- Benedict, Paul K. (1972) *Sino-Tibetan: A Conspectus*, Cambridge: Cambridge University Press.
- Genetti, Carol. (1997) ‘Object relations and dative case in Dolakha Newari’, *Studies in Language* 22.1: 37–68.
- (2005) ‘The participial construction in Dolakha Newar: Syntactic implications of an Asian converb’, *Studies in Language* 29.1: 35–87.
- (2007) *A Grammar of Dolakha Newar* (Mouton Grammar Library 40), Berlin: Mouton de Gruyter.
- (2011a) ‘The tapestry of Dolakha Newar: Chaining, embedding, and the complexity of sentences’, *Linguistic Typology* 15: 5–24.
- (2011b) ‘Direct speech reports and the cline of prosodic integration in Dolakha Newar’, *Himalayan Linguistics (Special issue in memory of Michael Noonan and David Watters)* 10.1: 55–76.
- Genetti, Carol, and Keith Slater. (2004) ‘Syntax/prosody interactions in a Dolakhae Nepāl Bhāsā rendition of the Mahābhārata’, *Himalayan Linguistics* 3: 1–91.
- Genetti, Carol, Ellen Bartee, Alexander R. Coupe, Kristine Hildebrandt, and You-Jing Lin. (2008) ‘Syntactic aspects of nominalization in five Tibeto-Burman languages of the Himalayan area’, *Linguistics of the Tibeto-Burman Area* 31.2: 97–144.
- Johanson, Lars. (1995) ‘On Turkic converb clauses’, in Martin Haspelmath and Ekkehard König (eds) *Converbs in Cross-Linguistic Perspective*, Berlin: Mouton de Gruyter, 313–347.
- Shrestha, Rudra Laxmi. (1996) *Dolakha Nepāl Bhāsāyā Varnātmak Adhyana*, (in Kathmandu Newar: *A Descriptive Study of Dolakha Nepāl Bhāsā*) doctoral dissertation, Tribhuvan University.

CHAPTER TWENTY-THREE

KATHMANDU NEWAR (NEPĀL BHĀŚĀ)

David Hargreaves

1 INTRODUCTION

Nepāl Bhāśā (Kathmandu Newar), the ancestral language of the Newars of the Kathmandu Valley, is spoken by roughly 846,557 speakers (Central Bureau of Statistics 2011). The term *Nepāl Bhāśā* and the colloquial term /newa bhæ:/ are the traditional Newar terms, although Western and Newar linguists publishing outside Nepal have usually used the term Newari, and more recently, Newar. The language is spoken primarily in the Kathmandu valley, in Kathmandu, Patan, and Bhaktapur and the villages surrounding these urban areas. With two notable exceptions (Dolakha Newar and Badikhel Pahari), the dialects of the surrounding villages and the smaller communities of Newar speakers outside the Kathmandu valley are directly related to Kathmandu, Patan, and Bhaktapur. Primarily on the basis of patterns of verbal morphology, Shakya (1992) argues for two major dialect groupings. The Kathmandu/Patan/Bhaktapur grouping is distinguished by the occurrence of egophoric (conjunct/disjunct) verb conjugations (see section 3.2.1), while a second grouping, consisting of Badikhel Pahari and Dolakha Newar (Genetti 2007, this volume; Shrestha 2010), is characterized by the presence of subject/verb agreement morphology. While the egophoric (conjunct/disjunct) inflectional morphology among the Kathmandu, Bhaktapur, and Patan related dialects is clearly cognate, the agreement morphology in Badikhel Pahari has not yet been shown to be cognate with Dolakha Newar, and neither appears cognate with the egophoric (conjunct/disjunct) morphology. Thus, the historical processes leading to different verb conjugation patterns remain opaque.

Locating the Newar languages within the Tibeto-Burman family is problematic. They most likely belong in the Himalayish branch of Tibeto-Burman (Matisoff 2003), although their position relative to further subgroupings remains unresolved. Based on his analysis of verb morphology in Dolakhā Newar, van Driem (1993) argues for placing the Newar family close to Kiranti, although Kansakar (1997: 20) wisely notes that at the heart of the sub-grouping problem lies ‘the fact that the Newar language is a language evolving from mixed ethnic/linguistic influences that do not lend easily to a neat classification’.

Throughout history, the Kathmandu Valley has been a crossroads of cultural exchange and Newars have had extensive interaction with northern Indic cultures and languages. A large percentage of the Newar lexicon can be traced to Indic sources. Increasing Nepali/Nepāl Bhāśā bilingualism continues to influence the lexicon and grammar, and the status of Nepali, as the national lingua franca and medium of education (along with English), exerts a powerful force towards language shift away from Nepāl Bhāśā (Shrestha 1990).

Written forms of Nepāl Bhāśā date back as early as 1114 AD [NS 235] and a classical literary tradition flourished up through the late eighteenth century (Jørgenson 1931, 1941; Malla 1982, 2000; Kansakar 1997), making Nepāl Bhāśā (along with Tibetan and

Burmese) one of the oldest literary languages in the Tibeto-Burman family. Genres from the classical tradition include poetry and drama, as well as technical, religious and historical prose manuscripts (Malla 1982). Many of the early manuscripts are translations of Sanskrit works; later works are more likely to be composed in Nepāl Bhāṣā. The oldest manuscripts are written in the *bhujimmola*, or fly-headed script. More commonly, the older texts are written in the *Nepālākṣara* or *Pracalit* script (Malla 2000). In modern Nepal, the Devanagari script predominates.

2 PHONOLOGY

2.1 Syllable structure

Syllables of non-Indic origin are constructed from the template: (C) (G) V (V). The most elaborated syllable type will consist of a complex onset, consonant plus glides /y/ or /w/, and a complex nucleus, a diphthong, as in *khwaū* ‘cold’. Syllables without onsets occur in a small number of monosyllabic lexical items, *i*: ‘time’, *uī*: ‘madwoman’, and occur commonly as the first syllable in bisyllabic words: e.g. *o.no* ‘there’, *i.ku* ‘dizzy’.

Given both the long history of contact with Indic languages and the increasing Nepāl Bhāṣā/Nepali bilingualism, words with Indic syllable patterns occur frequently in a wide range of sociolinguistic contexts. These expanded templates include syllable onset clusters, e.g. /pr/ or /sth/, and closed syllables, e.g. *bhut* ‘ghost’. Closed syllables may also occur with phonological reductions *mca-tɔ* ‘child-plural’ > [mɔs.tɔ] and with an adverbial suffix with a geminate onset cluster: *gya-* ‘be afraid’ + *kkɔ* adverbial suffix > [gyak.kɔ] ‘fearfully’.

2.2 Consonants

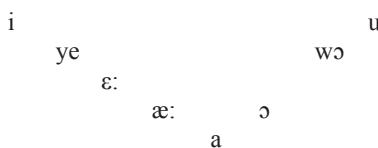
The Kathmandu Newar consonant inventory is outlined below:

<i>Labial</i>	<i>Alveolar</i>	<i>Alveo-palatal</i>	<i>Velar</i>	<i>Glottal</i>
p	t	c	k	h
ph	th	ch	kh	
b	d	j	g	
bh	dh	jh	gh	
m	n			
mh	nh			
	s			
	l			
	lh			
	(r)			
w	y			

With the exception of the glides /w/ and /y/ and the fricatives /s/ and /h/, all consonants come in unaspirated/aspirated pairs. The segment /r/ is not phonemic but does occur frequently in the large number of Indic loans.

2.3 Vowels

Length is contrastive for all vowels, except [ɛ:] and [æ:], which occur only as long vowels. Nasalization is contrastive for all vowels.



There are four simple vowels: /i/, /a/, /u/, /ɔ/. Of these four, the vowel /ɔ/ exhibits the most allophonic variation, occurring as [ɔ], [ə], and [ɑ]. The mid back rounded vowel, represented as /wo/, occurs with a labial onset. The mid front vowel, represented as /ye/, occurs with a palatal onset. A confusing variety of Devanagari and romanized spellings for these two vowels can be found in both the Newar and Western literature: ‘e’, ‘ye’ and ‘ya’ for /ye/ and ‘o’, ‘wo’ and ‘wa’ for /wo/ (see Manandhar 1986; Kölver and Shresthacarya 1994: x–xi). The long vowels [ε:] and [æ:] are represented as [œ] and [ae] following conventions in Devanagari transliteration, but are not true diphthongs. Diphthongs end in high vowels, the most common of which are /ai/, /au/, /i:/, /u:/, less frequently /ei/, and /eu/.

3 INFLECTIONAL MORPHOLOGY

Word classes can be identified via the inflectional possibilities of the stem. Nominal roots are those that directly take case marking; verbal roots are those that take tense/aspect morphology. One set of adjectival predicates are actually verbal roots in their imperfective form (see section 3.2), mainly colour terms (*hyāñ* ‘red’) and sensory attributes (*lumu* ‘warm’). The remaining adjectives do not take verbal inflections. Both sets occur with attributive suffixes when occurring as modifiers in attributive constructions (see section 5.1). Non-inflecting morphemes comprise a heterogeneous set including temporal, locative, and manner adverbials, discourse and speech act particles, and conjunctive clause-linking morphemes.

3.1 Noun phrases

Noun phrase morphology includes number, case, and numeral classifiers. Gender and diminutive marking is discussed in section 4.2.

3.1.1 Number

Number is marked on ordinary nouns with animate referents via the suffix *-tɔ*, i.e. *khi-ca-tɔ* ‘dogs’; plural nouns with honoured referents are marked with *-p̄i*: i.e. *pasa-p̄i*: ‘friends’. Nouns representing inanimate referents do not normally take plural suffixes (Hale 1986: xxxiii).

3.1.2 Case

Case is marked by clitics on noun phrases (Kölver 1976; Hale and Manandhar 1980). Nouns with a final open syllable and long vowel in their absolute form often exhibit a final consonant with suffixed case forms: *lɔ:* ‘water’ vs *lɔkh-e* ‘in the water’, *lha:* ‘hand’, vs *lhat-ɔ* ‘by hand’. Which final consonants occur in a particular word is lexically idiosyncratic.

Case forms include unmarked absolute, genitive *-ya*, animate locative/associative *-ke*, and inanimate locative suffix *-e*. In addition, the suffix *-nɔ* (regular allomorph as

nasalization and vowel lengthening) marks an abstract source category, including locative source (ablative), inanimate causal sources (instrumental), transitive agents (ergative), and causal subordinate clauses. Conversely, the suffix *-(ya)tɔ* marks an abstract goal category, including (dative) recipients, benefactives, and experiencers, as well as animate affected patients in transitive clauses, and purpose infinitive clauses. The irregular ergative forms *-sāt̪*, *-sāt̪:* and *-syāt̪:* occur in plural, honorific pronominal paradigms and with classifiers. Case paradigms for common nouns are given below:

	friend		eye
	<i>Singular</i>	<i>Plural</i>	<i>Singular</i>
Absolutive	<i>pasa</i>	<i>pasapīt̪:</i>	<i>mikha</i>
Ergative	<i>pasāt̪:</i>	<i>pasapisāt̪:</i>	
Dative	<i>pasayat̪o</i>	<i>pasapīt̪:tɔ</i>	
Associative	<i>pasayake</i>	<i>pasapīt̪:ke</i>	
Genitive	<i>pasaya</i>	<i>pasapini</i>	<i>mikhaya</i>
Locative			<i>mikhae</i>
Instrumental/Ablative			<i>mikhāt̪:</i>

Case forms for first person pronouns are given below:

	<i>Singular</i>	<i>Plural inclusive</i>	<i>Plural exclusive</i>
Absolutive	<i>ji</i>	<i>jhi:/jhi:pīt̪:</i>	<i>jipīt̪:</i>
Ergative	<i>jīt̪:</i>	<i>jhi:sāt̪:</i>	<i>jimisāt̪:</i>
Dative	<i>jit̪o</i>	<i>jhi:tɔ</i>	<i>jimit̪o</i>
Associative	<i>jike</i>	<i>jhi:ke</i>	<i>jimike</i>
Genitive	<i>ji(-gu)</i>	<i>jhi:(-gu)</i>	<i>jimi:(-gu)</i>

Second and third person pronouns index familiar, respected, and high honorific addressees. Familiar and respect forms for second person are given below:

	<i>Familiar</i>		<i>Respect</i>	
	<i>Singular</i>	<i>Plural</i>	<i>Singular</i>	<i>Plural</i>
Absolutive	<i>chɔ</i>	<i>chi:pīt̪:</i>	<i>chi</i>	<i>chikpīt̪:</i>
Ergative	<i>chɔt̪:</i>	<i>chimisāt̪:</i>	<i>chīt̪:</i>	<i>chikpisāt̪:</i>
Dative	<i>chɔnt̪o</i>	<i>chimit̪o</i>	<i>chit̪o</i>	<i>chikpinto</i>
Associative	<i>chɔnke</i>	<i>chimike</i>	<i>chike</i>	<i>chikpinke</i>
Genitive	<i>chɔt̪:(-gu)</i>	<i>chimi(-gu)</i>	<i>chi(-gu)</i>	<i>chikpini(-gu)</i>

The third person pronominal forms function as part of a set of demonstratives that index proximity relations between speaker, hearer, and referent object: *wɔ* ‘that’, *thwɔ*, ‘this’, *amɔ* ‘addressee proximate’, *hū* ‘far distant’. Complete pronominal and demonstrative paradigms can be found in Malla (1985), Hale and Shrestha (2006) and Josī (1992 [NS 1112]). The familiar and respect forms for animate referents are given below:

	<i>Familiar</i>		<i>Respect</i>	
	<i>Singular</i>	<i>Plural</i>	<i>Singular</i>	<i>Plural</i>
Absolutive	<i>wɔ</i>	<i>ipīt̪:</i>	<i>wɔekɔ:</i>	<i>wɔekɔ:pīt̪:</i>
Ergative	<i>wɔt̪:</i>	<i>imisāt̪:</i>	<i>wɔekɔlɔt̪:</i>	<i>wɔekɔ:pisāt̪:</i>
Dative	<i>wɔyat̪o</i>	<i>imit̪o</i>	<i>wɔekɔ:yat̪o</i>	<i>wɔekɔ:pinto</i>
Associative	<i>wɔyake</i>	<i>imike</i>	<i>wɔekɔ:yake</i>	<i>wɔekɔ:pinke</i>
Genitive	<i>wɔya(-gu)</i>	<i>imi(-gu)</i>	<i>wɔekɔ:ya(-gu)</i>	<i>wɔekɔ:pini(-gu)</i>

The interrogative pronouns (*su* ‘who’, *chu* ‘what’, *chae* ‘why’, *gōnō* ‘where’, *gōbole* ‘when’, *gōthe* ‘how’, *gu* ‘which’, *guli* ‘how much’, *gwɔ*-classifier ‘how many’) and the reflexive pronoun *thɔ:* also participate in case marking paradigms. Given appropriate discourse contexts, reflexive *thɔ:* may occur pronominally in all argument positions, including both ergative and absolute ‘subject’, as well as in emphatic adnominal position.

3.1.3 Numeral classifiers

Numerical expressions require bound classifiers that are suffixed to numerals and selected as a function of the head noun, whose semantic class they index (Malla 1985). They include: (1) general classifiers such as: *-hmo* ‘animate’, and *-gu(l)* ‘general inanimate’; (2) shape classifiers such as: *-gɔ(l)/-gwɔ(l)* ‘rounded’, ‘containers’, *-pu* ‘long, thin objects’, *-pa* ‘flat objects’; (3) unit classifiers such as: *-ku*: ‘piece (of meat)’, *-pe* ‘fingerful/mouthful (of rice)’, *-thu* ‘bundle (of vegetable stalks)’; (4) measure and measure-like classifiers such as: *-mōna* ‘measure of liquid’, *-pi*: ‘measure of land’; (5) unique classifiers such as: *-kha* ‘houses (only)’, *-ta* ‘pastry (only)’; (6) repeater classifiers such *hɔ:* in *hɔ: chɔ-hɔ:* ‘leaf one-leaf (classifier)’.

3.2 Verb phrases

Verbal morphology indexes two categories – tense/aspect and intentionality/egophoricity – but does not directly encode person/subject. Following Hale (1980), the terms most widely used for the intentionality/egophoricity contrasts are *Conjunct* (CJ) for egophoric and *Disjunct* (DJ) for non-egophoric, although Newar language scholarship has also used the terms *ātmā* ‘self’ and *para* ‘other’ (Josī 1992 [NS 1112]: 83). Although the function of the verb morphology is best characterized (following Tournadre 2008) as ‘egophoric’, rather than ‘conjunct/disjunct’, the latter terms are retained here for glossing purposes since they maintain continuity with earlier scholarship.

The most important morphophonemic contrasts occur with stem final consonant alternations whose distributions define the verb classes given below (Hale 1973, 1986; Sresthacharya 1981; Kansakar 1982; Malla 1985; Josī 1992 [NS 1112]):

Class 1	n-class	<i>wɔn-</i>	go
Class 2	t-class	<i>sya(t)-</i>	kill
Class 3	l/y class	<i>bi(l)-</i>	give
Class 4	l class	<i>hal-</i>	call, sing
Class 5	p,t,k class	<i>bhalɔp-</i>	think
		<i>so:t-</i>	call, invite
		<i>penk-</i>	kick

The verbal template is ROOT-consonant-suffix, where the verbal root minimally consists of a vocalic nucleus, usually with an initial consonant or consonant plus glide onset. The diagnostic stem consonant alternations occur across the finite inflectional paradigm. Below, a consonant-vowel (CV) root template is used to illustrate the inflectional paradigms.

<i>Class</i>	<i>Past</i>	<i>Perfective</i>	<i>Imperfective</i>	<i>Non-past</i>	<i>Non-past</i>
	<i>conjunct</i>	<i>disjunct</i>	<i>disjunct</i>	<i>conjunct</i>	<i>disjunct</i>
1	<i>CVn-a</i>	<i>CVn-ɔ</i>	<i>CV:</i>	<i>CVn-e</i>	<i>CVn-i</i>
2	<i>CVn-a</i>	<i>CVt-ɔ</i>	<i>CV:</i>	<i>CV-e</i>	<i>CV-i</i>
3	<i>CVy-a</i>	<i>CVl-ɔ</i>	<i>CV:</i>	<i>CV-e</i>	<i>CV-i</i>
4	<i>CVl-a</i>	<i>CVl-ɔ</i>	<i>CV:</i>	<i>CVl-e</i>	<i>CVl-i</i>

5	<i>CVp-a</i>	<i>CVp-ɔlɔ</i>	<i>CVp-yu:</i>	<i>CVp-e</i>	<i>CVp-i</i>
	<i>CVt-a</i>	<i>CVt-ɔlɔ</i>	<i>CVt-u:</i>	<i>CVt-e</i>	<i>CVt-i</i>
	<i>CVk-a</i>	<i>CVk-ɔlɔ</i>	<i>CVk-u:</i>	<i>CVk-e</i>	<i>CVk-i</i>

The regular marking of imperfective (IMPF) with classes 1, 2, 3, and 4 involves lengthening of the root vowel minus any stem final consonant; however, if the root vowel is /i/, then /i/ > /y/ and the imperfective form appears as the long vowel /-u:/. Similarly, if the root vowel is /e/, then /e/ > /y/ and the imperfective suffix will appear as /-ɔ:/. In addition to the inflectional forms listed above, there is an imperative (IMP) form identical to the imperfective disjunct, except all vowels are short.

Earlier descriptions of Newar verbal morphology in the English-language scholarship contrast a ‘past’ with a ‘non-past’ for both conjunct and disjunct categories (past conjunct /-a/, non-past conjunct /-e/, past disjunct /-ɔ/, non-past disjunct /-i/) and then identify a separate ‘habitual’ or ‘stative’ form, with the lengthened stem vowel (see Hale 1973, 1986; Malla 1985). However, the temporal semantics in finite clauses shows a more complex relationship between the egophoric (conjunct/disjunct) system and temporality. The conjunct category contrasts tense; the disjunct category contrasts a mixed tense and aspect system. For example, the two conjunct forms, /-a/ (past) and /-e/ (non-past), mark a binary tense opposition within the conjunct category. In contrast, in the disjunct category, there exists a trinary mixed tense/aspect contrast: perfective /-ɔ/, imperfective /-v:/, and non-past /-i/ (see Hargreaves 2005; Hale and Shrestha 2006).

3.2.1 Intentionality/egophoricity (conjunct/disjunct)

In addition to encoding tense/aspect contrasts, the verbal morphology exhibits a set of oppositions that index intentional action/egophoric contrasts in finite clauses. In egophoric contexts – whenever the action is construed as intentional, and the actor/agent is also the evidential source reporting the action – conjunct forms will occur (Bendix 1974, 1992; Hargreaves 1991, 2005, forthcoming). More specifically, the verb will have a conjunct form whenever:

- 1 the verb is finite; *and*
- 2 the event is construed as involving an intentional action by the actor; *and*
- 3 the speech act is:
 - a declarative with a first person subject; *or*
 - b interrogative with a second person subject; *or*
 - c reported speech where the main clause subject and reported speech complement clause subject are coreferential.

In non-egophoric contexts – all other finite environments except those outlined above – disjunct suffixes occur.

The construal of intentionality is, for the most part, lexically governed. The distribution of egophoric markers reveals a three-way categorization of verb types:

- a Control verbs, such as intransitive motion (*wɔn-* ‘go’), body posture (*dɔn-* ‘stand up’) and prototypical transitive verbs with affected objects (*jwɔn-* ‘grab’), entail intentionality with animate subjects, and require conjunct forms under the conditions outlined above. Non-intentional interpretations with these verbs are marked uses, typically indicated via adverbial or evidential markers.
- b Non-control verbs, such as *gya-* ‘be afraid’, or *si(t)-* ‘die’, entail no potential for intentional action and disallow egophoric inflection.

- c Fluid verbs, such as *sən-* ‘move’ or *thi-* ‘touch’, allow for alternative interpretations of intentionality. With this small set of verbs, the occurrence of either set of forms may function as the sole index of egophoric or non-egophoric interpretations.

3.2.2 Causatives

There are two morphological processes marking causativization. One process, confined to a restricted set of verbal forms, is a non-productive reflex of the Proto-Tibeto-Burman causative prefix **s*- (Malla 1985: 99). Examples include: *den-* ‘lie’ > *then-* ‘lay’; *gya-* ‘afraid’ > *kyha-* ‘frighten’.

A second causative process, fully productive, is marked by the causative morpheme *-k-/kol-* suffixed to the verb stem. The causative stem plus suffix is then inflected via the tense/aspect and egophoric (conjunct/disjunct) system. The paradigm with the verb *ya-* ‘do’ is given below:

Past	Perfective	Imperfective	Non-past	Non-past
<i>conjunct</i>	<i>disjunct</i>	<i>disjunct</i>	<i>conjunct</i>	<i>disjunct</i>
<i>ya-k-a</i>	<i>ya-kol-ɔ</i>	<i>ya-k-u:</i>	<i>ya-k-e</i>	<i>ya-k-i</i>

3.2.3 Non-finite inflections

There is an inflectional form /-a/ appearing in serial-like verb concatenations (see section 5.6) that follows the same stem final patterns as the past conjunct inflection, and indeed, on morpho-phonological grounds appears identical with the past conjunct inflection. Nevertheless, morpho-syntactically, it is non-finite, and does not index any of the tense/aspect/person/intention/egophoric contrasts characteristic of a finite morpho-syntactic environment. Similarly, the infinitive form /-e/ patterns morpho-phonologically exactly like the non-past conjunct, although again it is morpho-syntactically a non-finite inflection, and does not index any contrasts characteristic of a finite morpho-syntactic environment. A core group of sensory adjectival verbs takes the suffix /-se/ in a non-finite construction with the auxiliary *cwɔn-* ‘stay’.

4 WORD FORMATION

Both nominal and verbal roots undergo a variety of word formation processes, including compounding, affixation, and reduplication.

4.1 Compounding

Compounds resulting in nominal heads occur with a variety of roots:

noun-noun: *ja-ti* ‘rice-broth’ *jya-khū* ‘work-thief’ > ‘idler’, ‘slacker’

adjective-noun: *paū-khwa* ‘sour-face’ > ‘a sourpuss’, ‘grump’; *kwa-ti* ‘hot-broth’

verb-noun: *nɔye-wa* ‘eat-tooth’ > ‘molar’, ‘*bwɔ-solɔ*’ ‘flying-horse’ > ‘type of mythical animal’.

4.2 Suffixation

Productive suffixes include the feminine marker, e.g. *jyapu* ‘labourer’/‘peasant’ > ‘*jyapu-ni*’ ‘female labourer’, the diminutive marker, e.g. *dugu* ‘goat (male)’ > *dugu-ca*

‘small goat (male)’, and the human/animacy marker: *jya-mi* ‘worker’ < *jya* ‘work’; *panga-mi* ‘resident of Panga village’.

There are also a variety of derivational-like suffixes, although many are not fully productive (Sresthacharya 1981; Malla 1985). They include: *gyan-pu* ‘fearful’, ‘dangerous’ < *gya-* ‘be afraid’; *nɔ-sa* ‘food’ < *nɔ-* ‘eat’; *nhi-su* ‘cheerfulness’, ‘laughter’ < *nhil-* ‘to laugh’; *œla-gulu* ‘drunkard’ < *œla* ‘distilled liquor’.

Manner adverbs may be derived from verbal predicates with the suffix *-kɔ/-kkɔ* as in *gya-* ‘be afraid’ > *gyakkɔ* ‘fearfully’.

Demonstrative forms may be derived from a demonstrative root plus suffix, as in *thwɔ* ‘this’ > *thu-li* ‘this much’, *thu-khe* ‘this direction’, *thɔ-nɔ* ‘this place’, *thɔ-the* ‘this manner’, *thu-bɔle* ‘this time’, *thu-paecwɔ* ‘this extent/degree’.

4.3 Prefixation

One common set of prefixes is derived from a directional system. Directional morphemes, such as *kwɔ* ‘down’, can function both as a root morpheme, as in *kwɔ-ne* ‘under/below’, *kwɔ-hã*: ‘downward self-induced motion’, *kwɔ-tɔ* ‘downward causative motion’, and as a verbal prefix: *kwɔ-ka(l)-* ‘take down’ < *ka(l)* ‘take/grab’; *kwɔ-ca(l)-* ‘close, end, get to the bottom of’ < *ca(l)* - ‘close’. Other directional morphemes functioning both as root forms and verbal prefixes include: *du* ‘in’; *pi* ‘out’; *li* ‘behind’; *nhyɔ* ‘front’; *tha* ‘up’.

Some verbal prefixes, such as the prefix *svo-* ‘into’, ‘inside’, have direction meanings but do not occur as free direction morphemes: *svo-ca(l)-* ‘get thrust into’, be ‘stuck into’; *svo-tu(t)-* ‘be bogged down’, ‘stuck in’; *svo-thɔn-* ‘put in safe keeping’, ‘thrust in pocket’, ‘stuff, pack’. There are a wide range of other non-productive verbal prefixes, many of which are synchronically opaque (Malla 1985: 23; Hargreaves 2004).

4.4 Reduplication

Reduplication occurs with a variety of stems, partly productive and partly lexically idiosyncratic (Sresthacharya 1976; Malla 1985: 26). Reduplication of nominal roots indicates a plural meaning elusive in its actual semantics but roughly ‘bunch of *x*, with a potential pejorative sense of “motley”’. Typically the reduplicated stem undergoes a vowel change in which high vowels [i] and [u] become [a] in the reduplicated form; all others become [i]. The plural suffix *-tɔ* may co-occur with reduplication: *khica* ‘dog’ > *khica-khaca*; *khūtɔ* ‘thieves’ > *khūtɔ-khātɔ*; *kwɔ* ‘crow’ > *kwɔ-ki*; *ja* ‘rice’ > *ja-jí*; *mɔstɔ* ‘children’ > *mɔstɔ-mistɔ*. Some forms of plural reduplication involve an idiosyncratic change in the initial consonant of the reduplicated stem as in *jhyatu* ‘heavy’ > *jhyatuyatu*. Interrogative pronouns undergo reduplication (without vowel changes) to mark plural: *su* ‘who’ (singular) > *susu* ‘who’ (plural); also *chu* ‘what’ > *chuchu* and *gu* ‘which’ > *gugu*.

Non-finite verb forms followed by auxiliaries may undergo reduplication with a stem vowel change as in *twɔna wɔn-a* ‘drink go’ > *twɔna-tina wɔn-a* and *bwɔna hɔya* ‘accompany bring’ > *bwɔna-bina hɔya*. The semantic characterization is again elusive, but often takes on the aspectual colouring of an imperfective, roughly ‘doing *x* in a prototypical manner’ (see Malla 1985: 27). Reduplication of verbal stems accompanied by nasalization marks a progressive aspect commonly used in narration: *ya-* ‘to do’ > *ya-yã* ‘doing and doing’. Verbal prefixes may be reduplicated for intensification: *licile* ‘to retreat’ > *lilicile* ‘retreat way back’.

5 SYNTAX

5.1 Noun phrase syntax

Noun phrases may be expanded as follows:

(demonstrative) (genitive) (attributive) head (numeral/classifier) (postposition)

Demonstratives function both pronominally and as determiners: *wɔ* ‘that’, *thwɔ*, ‘this’, *amɔ* ‘addressee proximate’, *hū* ‘far distant’. Possessive nouns or pronouns are marked with genitive case forms (see section 3.1.2). In addition to the genitive case forms, genitive nouns may be optionally marked with the attributive (ATR) suffixes as follows: *-gu* ‘inanimate head noun’; *mhɔ* ‘animate head noun’. In fact, optional attributive marking is possible with demonstratives, genitive nouns and head nouns for kin terms (Kölver 1977; Hale 1985, 1994).

As noted in section 3 earlier, there are two sets of adjectives occurring in noun phrases, both of which are marked with the ‘attributive’ (ATR) suffixes relative to the head noun: *-gu* ‘inanimate’, *mhɔ* ‘animate singular’, and *-pt̪*: ‘animate plural’. The numeral-classifier will normally follow the head noun, but may occur in the position preceding the attributive adjectives. Postpositional elements occur outside the case marked item and include directional markers, *chē tɔkkɔ* (house up.to), or particles (see section 5.13). Adverbial modifiers precede adjectival modifiers. An expanded noun phrase is given below:

- (1) *thwɔ ji-gu tɔsɔkɔ bāla:-gu sɔphu ni-gu*
DEM 1-GEN/ATR very good-ATR book two-CL
'These two really good books of mine . . . '

When a classifier follows the head, it will be marked with the appropriate case markers.

- (2) *wɔ-yā pasa chɔ-mhɔe-syā:*
3-GEN friend one-CL-ERG
'By a friend of his . . . '

5.2 Relative clauses

Relative clauses are a category of nominalization preposed as modifiers to a head noun and marked with one of the three attributive suffixes: *-gu* ‘inanimate’, *mhɔ* ‘animate singular’ and *-pt̪*: ‘animate plural’. The attributive morphemes are suffixed to the finite verbal inflectional morphology, although the perfective disjunct form does not occur in these constructions. Thus, attributive suffixes mark constituents that function in noun phrases as determiners or modifiers to a head noun, or in anaphoric contexts as the head. The pairs of sentences below illustrate ‘headed’ and ‘headless’ noun phrases, where attributive suffixes function as heads and take case marking.

- (3) *ji-gu chē ji-gu*
1-ATR house 1-ATR
'my house' 'mine' (inanimate)
- (4) *hyāñ: -gu chē hyāñ: -gu*
red:IMPF-ATR house red:IMPF-ATR
'red house' 'red (one)'

- (5) *wɔ̃:-gu* *chɛ* *wɔ̃:-gu*
 go:IMPF-ATR house go:IMPF-ATR
 'The house (s/he) has gone to . . .' '(the place) gone to . . .'

(6) *wɔ̃:-mhɔ* *manu* *wɔ̃:-mhɔ*
 go:IMPF-ATR person go:IMPF-ATR
 'The person who went . . .' '(The one) who went . . .'

(7) *wɔ̃:-pt̪i:* *manu-tɔ* *wɔ̃:-pt̪i:*
 go:IMPF-ATR person go:IMPF-ATR
 'The people who went . . .' '(Those) who went . . .'

Although relative clauses are a category of nominalization, the label ‘nominalization’ (NMLZ) will be reserved for clauses occurring as arguments of copular or perception verbs (see section 5.7) below, or non-embedded nominalizations. Note the formal similarity between the headless relativization *wɔ̄:-gu* in (5) above and the nominalization *wɔ̄:-gu* in (8) below. Relative clause constructions and nominalizations are syntactically distinguished by the presence of a gapped constituent in relative clauses corresponding to the head noun:

- (8) *w̪ɔ:* *sita ch̪e* *w̪ɔ:-gu* *khɔnɔ*
 3:ERG Sita house go:IMPF-NOM see-PFV.DJ
 'S/he saw Sita go home.'

5.3 Basic sentence types

The canonical word order for basic sentences is sov. Manner and related adverbials appear in a variety of preverbal positions, but not postverbally. Basic sentence types include: copular (with/without copular verb), intransitive, transitive and ditransitive clauses:

- (9) *Manoj guru* (*khɔ:*)
 Manoj teacher be:IMPF
 'Manoj is a teacher.'

(10) *Sita wɔl-ɔ*
 Sita come-PFV.DJ
 'Sita came.'

(11) *j̥i: ja nɔy-a*
 1:ERG rice eat-PST.CJ
 'I ate rice.'

(12) *imi-sɔɔ: dyɔ:-yato me syat-ɔ*
 3:PLUR-ERG god-DAT buffalo kill-PFV.DJ
 'They sacrificed a buffalo to the god.' (Malla 1985: 83)

Some word order variation with transitive and ditransitive clauses does occur in marked contexts, although object-verb contiguity is highly preferred and zero nominals are frequent enough to make ditransitive clauses with all three arguments rare in actual discourse contexts.

5.4 Infinitive

Infinitive marked verbs, most commonly complements to a set of complement-taking verbs, take the suffix *-e/*. Emphatic purpose clauses appear with the infinitive plus the dative case marker. A less emphatic purpose clause occurs with a reduced stem.

- (13) *jī: ja nɔ-e mɔ-phu:*
1:ERG rice eat-INF NEG-able:IMPF
'I'm not able to eat rice.'
- (14) *ji den-e-tɔ wɔy-a*
1:ABS lie-INF-DAT come-PST.CJ
'I came in order to lie down.'
- (15) *wɔ nɔ won-ɔ*
3:ABS eat:PURP go-PFV.DJ
'S/he went and ate / went to eat.'

5.5 Verb concatenation

Verb concatenation constructions consist of a non-finite verb form, marked with /-a/, followed by small set of verbs which carry directional and aspectual meanings, functioning as 'auxiliaries' in concatenations, but maintaining independent main verb functions in other contexts (Hargreaves 1986).

- (16) *Sita wɔy-a cwaŋ-ɔ*
Sita come-CM stay-PFV.DJ
'Sita is coming.'
- (17) *pwa syan-a wɔl-ɔ*
stomach ache-CM come-PFV.DJ
'(My) stomach has begun to ache.' (Malla 1985: 76)

5.6 Nominalized clauses

Nominal clauses are marked with the nominalizing suffix /-gu/ and occur with copular verbs and a set of cognition/perception verbs (see section 5.2). Non-embedded or bare nominalized clauses without copula verbs occur frequently in conversation.

- (18) *sita wɔy-a cwɔ̄:-gu du*
Sita come-CM stay:IMPF-NOM be:IMPF
'Sita has come.'
- (19) *wɔ̄: sita wɔy-a cwɔ̄:-gu khɔn-ɔ*
3:ERG Sita come-CM stay:IMPF-NOM see-PFV.DJ
'S/he saw Sita coming.'
- (20) *chɔ gɔnɔ wɔn-e-gu*
2:ABS where go-NPST.CJ-NOM
'Where are you going?'

5.7 Clause chaining constructions

Chaining constructions, characteristic of narrative portions of discourse, consist of strings of non-finite clauses with shared arguments (Genetti 1988). The final clause in the sequence is marked with the finite verb form; all non-final verbs are marked with the suffix /-a:/.

- (21) *wɔ̄: ja chɔ̄-mɔ̄na thuy-a: nɔ̄y-a: pihã: won-ɔ*
 3:ERG rice one-CL cook-NF eat-NF out go-PFV.DJ
 ‘S/he cooked a mana of rice, ate it, and went out.’

5.8 Quotative complementation

Reported speech complement clauses are marked via the quotative complementizer *dhɔ̄ka:*, or occur as main clauses marked with the quotative (QT) evidential particle *hɔ̄*. Conjunct forms index co-reference between the actor/subject of the verb of speaking and actor/subject of the reported clause verb. Disjunct forms occur in all other environments.

- (22) *pine swɔ̄ wɔ̄n-a hɔ̄:*
 outside look:PURP go-PST.CJ QUOT
 ‘(S/he_i) said “(I_j) went outside to look”.’
- (23) *pine swɔ̄ wɔ̄n-ɔ hɔ̄:*
 outside look:PURP go-PFV.DJ QUOT
 ‘(S/he_i) said “(s/he_j) went outside to look”.’
- (24) *jii: ɔ:pwɔ̄ twɔ̄n-ɔ hɔ̄*
 1:ERG much drink-PFV.DJ QUOT
 ‘(S/he_i) said I_j drank too much.’

The complementizer *dhɔ̄ka* also occurs with mental states.

- (25) *pine swɔ̄ wɔ̄n-e jiu: dhɔ̄ka pihã won-ɔ*
 outside look:PURP go-INF ok:IPFV.DJ COMP out go-PFV.DJ
 ‘Thinking it’s OK to go outside to look, s/he went out.’

5.9 Subordination/coordination

Coherence relations between clauses are marked via a set of clause linking morphemes, including *-sa* ‘if’, *-sā* ‘even if’/‘though’, *-bole* ‘when’, *ki* ‘but’, ‘either’/‘or’, ‘if’, ‘when’, *-gulī* ‘because/since’.

- (26) *ya:-sā jiu*
 do-IPFV-COND be.OK:IPFV.DJ
 ‘(Even) if (s/he) does (it), (that’s) OK.’

The verb *dha-* ‘to say’ may function as a coordination marker: *dhae-wɔ̄* ‘as soon as’, *dha-sa* ‘if’, and *chae dha-sa* literally ‘why say-if’, i.e. ‘because’:

- (27) *dyɔ̄ won-ɔ dhae-wɔ̄ nhɔ̄ wɔ̄l-ɔ*
 sleep:PURP go-PFV.DJ say-AND sleep come-PFV.DJ
 ‘As soon as he lay down to sleep, he fell asleep.’

Other morphemes marking coordinate relations include *ale* ‘then’, *aesa* ‘if so’/‘therefore’, *aesā*, ‘nevertheless’, *tɔrɔ* ‘but’.

- (28) *dyɔ̄ wɔ̄n-ɔ tɔrɔ nhɔ̄ wɔ̄-e mɔ̄-phɔ̄t-ɔ*
 lie:PURP go-PFV.DJ but sleep come-INF NEG-able-PFV.DJ
 ‘(S/he) went to lie down, but didn’t fall asleep.’

5.10 Comparatives

Comparative clauses are formed with postpositional morphemes: *swɔya:*, alternatively *sibe*.

- (29) *chɔ swɔya:/sibe wɔ tɔ:dhī*
2:ABS COMPARATIVE 3:ABS tall
'Compared to you, he's tall.' (Kölver and Shresthacarya 1994: 327)

5.11 Negative clauses

Negation is marked with the verbal prefix *mɔ-*.

- (30) *jī ja mɔ-nɔy-a ni*
1:ERG rice NEG-eat-PST:CJ yet
'I haven't eaten rice yet.'

Negative imperatives are constructed via an infinitive verb plus the negative imperative form, *mɔte*, a frozen lexical form consisting of the negative *mɔ-* and the verb stem *te* 'be time to'/'be OK to'.

- (31) *nɔ-e mɔte*
eat-INF NEG.IMP
'Don't eat!'

5.12 Declarative/interrogative

Declarative sentences are generally unmarked. Yes–no interrogative clauses are marked with the suffix *la* (Hargreaves 1996).

- (32) *ja nɔ-e dhun-ɔ la*
rice eat-INF finish-PFV.DJ Q
'Have you already eaten?'

Substance (wh-) questions are marked with interrogative pronouns inflected with case marking morphology.

- (33) *su wɔl-ɔ le*
who:ABS come-PFV.DJ PRT
'Who has come?'

- (34) *sunā: yat-ɔ*
who:ERG do-PFV.DJ
'Who did (it)?'

5.13 Particles

Speech act particles appear clause finally and include: *la* 'Yes–no interrogative', *le* 'pre-suppositional' (often co-occurs with wh- questions), *re/le* 'persuasive', *hɔ* 'reported speech', *ka* 'assertive'. Other discourse particles appear clause internally with local scope, including: *la* 'topic', *nɔ* 'also', *jɔkɔ* 'just', *tū*: 'emphatic/particular', *he* 'emphasis', and *thē* 'similar/like'.

REFERENCES

- Bendix, E. (1974) ‘Indo-Aryan and Tibeto-Burman contact as seen through Nepali and Newari verb tenses’, *International Journal of Dravidian Linguistics* 3.1: 42–59.
- (1992) ‘The grammaticalization of responsibility and evidence: interactional potential of evidential categories in Newari’, in J. Hill and J.T. Irvine (eds) *Responsibility and Evidence in Oral Discourse*, 226–247. Cambridge: Cambridge University Press.
- Central Bureau of Statistics (2011) National Population and Housing Census. Government of Nepal. Available at: cbs.gov.np/nada/index.php/catalog/54 (accessed 14 August 2016).
- Genetti, C. (1988) ‘A syntactic correlate of topicality in Newari narrative’, in S. Thompson and J. Haiman (eds) *Clause Combining in Grammar and Discourse*, 29–48. Philadelphia: John Benjamins.
- (2007) *A Grammar of Dolakha Newar*. Mouton Grammar Library 40. Berlin: Mouton de Gruyter.
- Hale, A. (1973) ‘On the form of the verbal basis in Newari’, in H.R. Kahane, R. Kahane and B.B. Kachru (eds) *Issues in Linguistics: Papers in Honor of Henry and Renee Kahane*, 279–299. Urbana, IL: University of Illinois Press.
- (1980) ‘Person markers: finite conjunct and disjunct forms in Newari’, in R. Trail (ed.) *Papers in Southeast Asian Linguistics* 7 (Pacific Linguistics Series A, no. 53), 95–106. Canberra: Australian National University.
- (1985) ‘Noun phrase form and cohesive function in Newari’, in U. Piepel and G. Stickel (eds) *Studia Linguistica Diachronica et Synchronica*, 303–321. Berlin: Mouton de Gruyter.
- (1986) ‘Users’ guide to the Newari dictionary’, in T. Manandhar (ed.) *Newari–English Dictionary*, xxii–lxvi. Delhi: Agam Kala Prakashan.
- (1994) ‘Entailed kin reference and Newari -mha’, paper presented to the 27th International Conference on Sino-Tibetan Languages and Linguistics, Paris, France.
- Hale, A. and Manandhar, T. (1980) ‘Case and role in Newari’, in R. Trail (ed.) *Papers in Southeast Asian Linguistics* 7 (Pacific Linguistics Series A, no. 53), 79–93. Canberra: Australian National University.
- Hale, A. and Shrestha, K. (2006) *Newār (Nepāl Bhāsā)*. Munich: Lincom Europa.
- Hargreaves, D. (1986) ‘Independent verbs and auxiliary functions in Newari’, *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society* 12: 401–412.
- (1991) ‘The conceptual structure of intentional action: data from Kathmandu Newari’, *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society* 17: 379–389.
- (1996) ‘From interrogation to topicalization: PTB *la in Kathmandu Newar’, *LTBA* 19.2: 31–44.
- (2004) ‘Directional prefixes in Kathmandu Newar’, in A. Saxena (ed.) *Himalayan Languages: Past and Present, Trends in Linguistics: Studies and Monographs* 149, 273–284. Berlin: Mouton de Gruyter.
- (2005) ‘Agency and intentional action in Kathmandu Newar’, *Himalayan Linguistics* 5: 1–48.
- (forthcoming) “Am I blue?”: Privileged access constraints in Kathmandu Newar,’ in S. Floyd, E. Norcliffe and L. San Roque (eds) *Egophoricity* [edited volume in preparation]. Amsterdam/Philadelphia: John Benjamins.
- Jørgenson, H. (1931) ‘A dictionary of the Classical Newari’, *Det. Kgl. Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser* 23.1: 1–178.

- (1941) ‘A grammar of the Classical Newari’, *Det. Kgl. Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser* 27.3: 1–108.
- Josī, Š.K. (1992) [NS 1112] *Nepāl bhāṣāyā bhāṣāvaijñānika vyakaraṇa (A Linguistic Grammar of Nepāl Bhāṣā)*. Kathmandu: Lacoul Publications.
- Kansakar, T.R. (1982) ‘Morphophonemics of the Newari verb’, in T.R. Kansakar (ed.) *Occasional Papers in Nepalese Linguistics, No. 1*, 12–29. Lalitpur, Nepal: Linguistic Society of Nepal Publication.
- (1997) ‘The Newar language: a profile’, *Newāḥ Viññāna: Journal of Newar Studies* 1.1: 11–28.
- Kölver, U. (1976) ‘Satztypen und Verbsubcategorisierung der Newari’, *Structura* 10, Munich: Fink.
- (1977) ‘Nominalization and lexicalization in Newari’, *Arbeiten des Kölner Universalen-Projekts* 30 working paper. Allgemeine Sprachwissenschaft, Institut für Linguistik, Universität zu Köln.
- Kölver, U. and Shresthacarya, I. (1994) *A Dictionary of Contemporary Newari*. Bonn: VGH Wissenschaftsverlag.
- Manandhar, T. (1986) *Newari–English Dictionary*. Delhi: Agam Kala Prakashan.
- Malla, K.P. (1982) *Classical Newari Literature: A Sketch*. Kathmandu: Educational Enterprise.
- (1985) *The Newari Language: A Working Outline*, Monumenta Serindica No. 14, Institute for the Study of Languages and Cultures of Asia and Africa. Tokyo: Tokyo University of Foreign Studies.
- (ed.) (2000) *A Dictionary of the Classical Newari: Compiled from Manuscript Sources*. Kathmandu: Cwasā Pāsā, Nepal Bhasa Dictionary Committee.
- Matisoff, J. (2003) *Handbook of Tibeto-Burman: System of Philosophy of Sino-Tibetan Reconstruction*. University of California Publications in Linguistics: 135. Berkeley: University of California Press.
- Shakya, D.R. (1992) ‘Nominal and verbal morphology in six dialects of Newari’, Master’s thesis, University of Oregon.
- Shrestha, O. (2010) ‘A grammar of Pahari’ unpublished PhD dissertation, Tribhuvan University, Nepal.
- Shrestha, U. (1990) ‘Social networks and code-switching in the Newar community of Kathmandu City’, PhD dissertation, Ball State University, USA.
- Sresthacharya, I. (1976) ‘Some types of reduplication in the Newari verb phrase’, *Contributions to Nepalese Studies* 3.1: 117–127.
- (1981) *Newari Root Verbs*, Bibliotheca Himalayica 2.1. Kathmandu: Ratna Pustak Bhandar.
- Tournadre, N. (2008) ‘Arguments against the concept of “conjunct”/“disjunct” in Tibetan’, in B. Huber, M. Volkart and P. Widmer (eds) *Chomolangma, Demawend und Kasbek, Festschrift für Roland Bielmeier, zu seinem 65 Geburtstag*, Vol. I, 281–308. Halle (Saale): International Institute for Tibetan and Buddhist Studies.
- van Driem, G. (1993) ‘The Newar verb in Tibeto-Burman perspective’, *Acta Linguistica Hafniensis* 26: 23–43.

§3.3 Tamangic

CHAPTER TWENTY-FOUR

TAMANG

Martine Mazaudon

1 INTRODUCTION

Tamang is spoken by about 1,350,000 speakers in central Nepal, according to the 2011 Nepal Census. Many different regional varieties exist, to the point where intercomprehension is difficult (Varenkamp 1996). The particular dialect described here is spoken in the village of Risiangku, Sindhupalchok District, Bagmati Zone. It can be considered typical of Eastern Tamang.

Tamang is a close relative of Gurung, Thakali (including the dialects of Marpha, Thini and Syang), Manange, the Nar-Phu dialects, Chantyal, and the dialect of Tangbe, in Mustang zone. Together they form the TGTM group of Tibeto-Burman, a sister branch to the Tibetan group in what Robert Shafer (1955) identified as the Bodish Section of the Bodic Division of Tibeto-Burman. Shafer named this branch the ‘Gurung branch’ after the name of the best known member of the family at the time. Since most of the speakers of these languages designate themselves in their own languages by a word cognate to the word ‘Tamang’, Mazaudon (1978a) proposed to call the branch the ‘Tamang branch’, or to be completely explicit ‘TGTM’, an acronym coined from the initials of the largest ethnic groups in the branch: Tamang-Gurung-Thakali-Manange.¹

Tamang was included under the name ‘Murmi’ in Grierson’s (1908) *Linguistic Survey of India*. It was not further studied until the late sixties. Among early studies we can cite Taylor (1973) and Everitt (1973) on Western Tamang, and Mazaudon (1973, 1976) on Eastern Tamang. Some texts in Western Tamang have been published in Hale and Pike (1970: pt. III, 132–64) and a large collection in Hoefer (1981–1997). Texts in the Risiangku dialect can be found in the Pangloss electronic archive (Mazaudon 2000).

Until recently Tamang was an unwritten language. The situation has changed in the last 40 years, and an important literature is now produced in Tamang, as well as linguistic studies written in Nepali, by native speakers of Tamang, e.g. Yonjan-Tamang (1997, 2014).

For other languages of the TGTM group see Hildebrandt’s and Noonan’s contributions on Manange, Nar-Phu and Chantyal (this volume). For TGTM languages not included here, the reader can consult Glover (1974) on Gurung, the first full grammar of any language of the TGTM group, Georg (1996) on Marphali Thakali, and the grammatical sketches in Hale and Watters (1973). The next closest language on which grammatical information is available is Tibetan.

2 PHONOLOGY

A detailed account of the synchronic phonology of Risiangku Tamang can be found in Mazaudon (1973).

2.1 Initial consonants

The system of syllable initial consonants is the same in word initial and word internal positions (see Table 24.1). Voicing is not distinctive in either position. In word internal position unaspirated stops are pronounced fully voiced if intervocalic or if preceded by a voiced nasal or continuant, but they are voiceless if preceded by a voiceless stop or by *s*. In word initial position partial voicing of unaspirated stops is tone-dependent (section 2.7). The transcription of the sibilant affricate [ts] as /c/ follows the traditional usage in Indian linguistics. The retroflex series is only slightly retroflexed and the stop portion is followed by a trilled appendage, as first described in related Gurung by Burton-Page (1955).

2.2 Initial consonant clusters

Clusters are allowed with initial labials and velars, absent or very restricted with any other initials.

	p ^h	p	m	t ^h	t	n	l	c ^h	c	s	t ^h	t	r	k ^h	k	ŋ	h	
r	+	+	+	–	–	–	–	–	–	–	–	–	–	–	+	+	–	+
l	+	+	+	–	–	–	–	–	–	–	–	–	–	–	+	+	–	+
j	+	+	+	–	–	–	–	+	+	+	–	–	–	–	+	+	+	(+)
w	(+)	(+)	(+)	–	(+)	–	(+)	+	+	+	+	+	(+)	+	+	–	(+)	
rw	–	(+)	(+)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	(+)

+ means that the cluster is permitted, – that it is not, (+) that there are restrictions on the following vowel.

2.3 Final consonants

p t k m n ŋ l r (s)

2.4 Vowels

i i: e e: a a: o o: u u:

Vocalic length is distinctive in open initial syllables. Nasality is very marginally distinctive. It is transcribed by the tilde [~].

TABLE 24.1 INITIAL CONSONANTS

	Labial	Dental	Sibilant affricate	Trilled retroflex	Velar	Glottal
aspirated	p ^h	t ^h	c ^h	t ^h	k ^h	
unaspirated	p	t	c	t	k	
nasal	m	n			ŋ	
continuant		l	s	r		h
semi-vowel	(j, w)					

2.5 Vowel clusters or diphthongs

ui	oi	ai	au	iau
iu	io	ia	ua	uai
iu:	io:	ia:	ua:	iua:
iui	ioi	iai		

An *i* or *u* vowel followed by another vowel is pronounced as a semivowel. Thus, in the chart of initial consonants and consonant clusters *j* and *w* are not different from *i* and *u* as first elements of a vocalic cluster.

2.6 Tones

There are four tones, which apply to phonological words. They are transcribed by a raised number preceding the word. A word can contain several morphemes which are separated by a hyphen in the transcription. The first morpheme determines the tone of the whole word.

Particles which carry their own tone are considered separate words, and are not hyphenated with the preceding word.

The phonetic pitch of the tones is approximately as follows: ¹a, high falling; ²a, mid-high level; ³a, mid-low level; ⁴a, very low – falling if the word is monosyllabic, falling-rising-falling if the word has two syllables or more.

Words carrying the two higher tones, ¹ and ², have a clear voice quality; words in the low tones have a breathy voice quality to varying degrees (Mazaudon 1973; Mazaudon and Michaud 2008).

2.7 Tones and initials

The main point of interest in the phonology is the correlation between the tones *cum* voice quality and the initial consonants. The opposition between aspirated and non-aspirated consonants is not found under the two lower tones: only unaspirated stops are found in that context, and that archiphonemic series is pronounced with a slight optional voicing. This neutralization of aspiration under the two low tones, as well as the breathy voice quality of these tones derive from the historical origin of the tonal system.

Proto-Tamang can be reconstructed with a two-tone system, and three series of initial stops: voiced, voiceless and aspirated. The devoicing of the old voiced series in Proto-Tamang caused its merger with the old voiceless unaspirated series, and this produced a split in the old tonal system, by which two lower tones were phonologized after what used to be voiced initials (Mazaudon 1977, 1978a). This is a classical process of merger and compensatory split, which is exemplified in numerous languages of East Asia (Haudricourt 1961, forthcoming).

2.8 Canonical forms

The canonical form of the syllable in Risiangku Tamang is the richest in the TGTM group:

(Tone) (Initial Consonant) (Liquid) (Semivowel) Vowel (Final Consonant)

All verb roots are monosyllabic. About half of the nominal roots are monosyllabic. Morphemes of more than two syllables are exceptional. All bound morphemes are mono-syllabic.

3 TYPOLOGICAL SUMMARY

- Basic word order: sov: Tamang conforms in most respects to the patterns established by Greenberg (1966) for an sov language (except for the place of the negation before the verb and of numbers after the noun) although there is great freedom to move the arguments about, for rhetorical purposes.
- In complex verb phrases modals follow the verb root. Subordinate clauses precede the main clause.
- Case marking: ergative.
- Morphology: limited, exclusively suffixing. Bound morphemes include case markers on the NPs and aspect marking on the verb. There is no agreement pattern.
- Clause subordination pattern: mostly through participial and gerundive constructions. Topic–comment structure is frequent.
- Information structure markers play a major role in the expression of logical relations.

4 BASIC SENTENCES

4.1 Intransitives

4.1.1 Verbal predicate

s-ABS (\pm PERIPHERAL ARGUMENTS) v

4.1.1.1 Active

- (1) $^2ai\text{-}\emptyset \quad ^3naykar \quad ^1ni\text{-}pa ?$
2sg-ABS tomorrow go-IPFV
'Are you going tomorrow?'

4.1.1.2 Stative

Stative verbs carry verbal suffixes. They can be used in the perfective with an inchoative meaning.

- (2) $o, \quad ^3ce\text{-}ci \quad ^4tamo$
oh beautiful-PFV now
'Oh, it's really pretty this time.'

But they are mostly used in the participial form in -pa in the same construction as adjectival predicates (with or without a copula): $^3ce\text{-}pa \; ^1mula$ 'it is pretty'.

4.1.2 Non-verbal predicate

s-ABS PREDICATE (\pm COPULA)

4.1.2.1 Adjective predicate

- (3) *²cu ³mento ³caca*
 this flower small
 ‘This flower is small.’

The absence of a copula is very frequent. But there exist two main copulas and some others. *'mu-la* is mostly used for existence and attribution, and *³hin-pa* for identification.²

Several other verbs can be used as copulas:

- (4) *¹p^hjukpa-la ³caca ²kola-m ²os-pa ³a ¹k^ha-pa-¹ka*
 rich-GEN small child-TOP thus-NMLZ NEG come-IPFV-FOC
 ‘The children of the rich are not (*lit.* do not come) like that.’

Modifications of the adjective occur only when the adjective is used as predicate. A change of copula can be used for this purpose, for instance *¹ta-ci* expresses excess.

- (5) *³caca ¹ta-ci*
 small reach.completion-PFV
 ‘It is too small.’

Intensification is usually expressed by repetition of the adjective (29). Another frequent form of intensification is an intonation pattern. One of the syllables of the emphasized word is pronounced on a very high pitch with a falsetto voice. This procedure is an areal feature in Nepal.

Adjectives are rarely modified by adverbs.

4.1.2.2 Noun predicate

The meaning difference between the two basic copulas when used with an adjectival predicate is clear, but it does not seem to hold when the predicate is a noun.

- (6) *¹ya-la ¹apa ¹c^honypa {Ø / ³hin-na / ¹'mu-la}*
 1sg-GEN father merchant {Ø / be-NPST / exist-NPST}
 ‘My father is a merchant.’

4.1.2.3 Locative predicate

- (7) *⁴kle ¹it ³nay-ri ¹'mu-la*
 king this inside-LOC be-NPST
 ‘The king is inside this [bag].’

4.1.2.4 Predicate of possession

The predicate of possession, either of existence ('I have X') or of identification ('X is mine') is generally expressed by an NP in the genitive case.

- (8) *¹ya-la ⁴came ⁴ni: ¹'mu-la*
 1sg-GEN daughter two be-NPST
 ‘I have two daughters.’

- (9) *¹ya-la ¹ka*
 1sg-GEN FOC
 ‘It is mine.’

The genitive is the most commonly used case for possession, but the dative is also used:

- (10) *⁴tap ¹mu-la ²ani-ta ?*
needle be-NPST aunt-DAT
'Do you have a needle, Auntie?' (*lit.* 'Does Auntie have a needle?')

The dative is the usual case when there is a transfer or attribution of possession (11), although, even in that situation, the genitive can be used (12).

- (11) *¹ya-ta ³pa-u*
1sg-DAT give-IMP
'Give it to me!'
- (12) *⁴came ⁴ca ¹t^he-la ¹am-la-n ¹pin-ci*
daughter TOP 3sg-GEN mother-GEN-INT give-PFV
'The daughter, they gave to her mother.' (*lit.* 'of her mother')

4.1.3 Weather expressions

Natural events are usually expressed by a noun + a weak verb:

- (13) *²nam ¹k^ha-ci*
rain come-PFV
'It's raining.'

Similar constructions are used to express the way in which we experience external events. (See section 6.5.)

4.2 Transitives

Transitive verbs construct their subject in the ergative and their object either in the absolute or the dative. Some verbs allow only one or the other construction and can thus be sub-categorized as direct ((14), (15)) or indirect transitives (16).

s-ERG o-ABS v

- (14) *²naka-se ³tap-Ø ¹ca-ci*
chicken-ERG vegetable-ABS eat-PFV
'The chicken is eating the vegetable.'

Perception predicates follow the direct transitive pattern.

- (15) *¹ya-i ⁴jo:-Ø ¹mray-ci*
1sg-ERG thief-ABS see-PFV
'I saw the thief.'

s-ERG o-DAT v

- (16) *¹mam-se ²kol'-kat'-ta ³pay-pa*
Grandma-ERG children-DAT scold-IPFV
'Grandma is scolding the children.'

Many verbs allow variation in the case marking of the patient for semantic or rhetorical purposes. With those verbs, it is statistically more frequent that animate objects be put

in the dative, and inanimate objects in the absolute. Definite/indefinite, the degree to which the object is affected, and the information structure of the sentence also play a part.

Ditransitives

s-ERG io-DAT o-ABS v

- (17) *'am'-se ²kol'-ta ¹kan-Ø ²k^hwa:-ci*
mother-ERG child-DAT rice-ABS feed-PFV
'The mother fed the child rice.'

4.3 Experiencer constructions

n-DAT n-ABS v

A number of verbs, especially of feeling or experience, put the experiencer in the dative:

- (18) *²aya-ta ¹'pet-pa ?*
sister-DAT shy-IPFV
'Do you feel shy?' (said to a young girl addressed as 'sister')

With some other verbs, the experiencer is in the absolute and the source, or object of the feeling, is in the dative, yielding a construction similar to the indirect transitive construction except for the reversed word order.

- (19) *¹ya-ta ¹t^he-Ø ¹'nai-pa*
1sg-DAT 3sg-ABS disgust-IPFV
'I disgust him.'

Inanimate objects of feelings require a subordinate clause:

- (20) *¹kli ¹mraŋ-si ¹ya-Ø ¹nai-ci*
faeces see-ANT 1sg-ABS disgust-PFV
'Excrement disgusts me.' (lit. 'Seeing faeces I feel disgust.')

A construction which can also be used with animate objects as in:

- (21) *¹ya ¹mraŋ-si ²ai-Ø ¹'pet-pa ?*
1sg see-ANT 2sg-ABS shy-IPFV
'Are you shy with me?'

Note the case variation on the experiencer subject of 'shy' as compared to (18).

4.4 Peripheral arguments

Aside from io, some do, and experiencer subjects, the dative is used for many arguments loosely bound to the verb:

4.4.1 Beneficiary

- (22) *¹t^he-ta ³ro ¹kik ¹por-pa ¹ce-n*
3sg-DAT friend one take-IPFV only-INT
'He took only one travelling companion for himself.'

4.4.2 Instrument

Ergative, instrumental and ablative, three cases which share the general semantic notion of source, are homophonous.

4.5 Variation in subject case marking

There is ample variation in the case marking of the subject as well as the object (see section 4.2). This sometimes reflects the fact that the same verb can be used as a transitive or an intransitive or that an expression is in the process of lexicalization.

Information structure is also an important factor. In (24) topicalization of ‘you’ is accompanied by a shift to the dative case,³ as opposed to the expected ergative (23), although topicalization does not necessarily imply such a shift in case marking: compare (91).

- (23) *inam bakas⁴ k^hup 'jay-ci 'jay-la p^halana sipai-se*
reward a lot find-PFV us-GEN so-and-so soldier-ERG
'He got a big reward, our soldier so-and-so.'
- (24) *²os 'ta-m' ²ai-ta inam bakas ³a 'jay*
thus happen-SIM 2sg-DAT reward NEG find
'In that case you don't get a reward.'

5 THE NOUN PHRASE

5.1 Word order

The general word order is Modifier–Modified. Genitive phrases and relative clauses precede their heads, and the structure of the simple NP is as follows:

Demonstrative, Adjective₁, Adjective₂, Noun, Numeral, {Case + Particle/Particle + Case}⁵

5.2 Form classes

Noun modifiers: demonstrative, adjective (4), numeral (27), genitive NP (25) or headless relative clause (32), can function as nouns.

Noun modifiers are not morphologically distinguishable from nouns. The main difference between them and nouns is that they modify a noun without an intervening genitive marker.

5.3 Pronouns and pronominalization

The most common form of pronominalization is by deletion (a.k.a. zero-pronominalization). Any definite argument of the verb, the head of a relative clause (32), or the head of a complex NP (25) can be pronominalized by deletion. On the other hand, an indefinite element cannot. Deleted elements are referential.

The set of personal pronouns consists of *'ŋa* ‘1sg’, *'ai* ‘2sg’, *'t^he* ‘3sg’, *'jay* ‘1pl inclusive’, *'in* ‘1pl exclusive’, *²ai-ni* ‘2pl’, *'t^he-n* ‘3pl’. The two plural pronouns of the first person are often used instead of the singular. *'ai* is often replaced, for politeness, by a kinship term ((10), (18)), or by the intensifier *³ray* ‘oneself’, in person’.

The demonstratives ²cu ‘this’ and ¹oca ‘that’, and the nouns ³mi ‘man’, and ³ro ‘companion’ are also often used as pronouns.

- (25) ³mi-la Ø ⁴tamom ³rap-si ³cin-ci
person-GEN [thing] now play-ANT finish-PFV
'Her [tape-recorder] has now stopped playing.'

Personal or demonstrative pronouns take the same case markers as nouns, except for ¹ya '1sg' with which a variant of the ergative is found ¹ya-i 'by me'.

5.4 Reflexives, reciprocals and intensifiers

There are no special reflexive pronouns. Usually one of the pronouns, or both, are deleted. See (22).

The intensifier pronoun ³ray ‘in person’, ‘one’s own’ (König 2001) can be used for the reflexive.⁶

- (26) ³ponpo-se ³ray-ta-n ³kjay-si ¹cja:-pa
shaman-ERG self-DAT-INT hit-ANT see-IPFV
'The shaman checks [that it is not too hot for the patient] by hitting himself [with the hot broom].'

Reciprocity is expressed either by a verb with a conjoined subject:

- (27) ²o-te-ma ⁴ni:.nun ²c^hat-cim, ³rempa-then ¹mriy
that-much-upon both fight-PFV husband-and wife
'And thereupon they (both) fought, husband and wife.'

or by repetition of the arguments with reversed case marking:

- (28) ³rempa-ta ¹mriy ¹mriy-ta ³rempa ¹nai-sam . . .
husband-DAT wife wife-DAT husband disgust-if
'If husband and wife are disgusted with each other . . .'

5.5 Adjectives

Adjectives are few and rarely used as modifiers. They comprise ‘big’, ‘small’, ‘elder’, the colours and a few others. Most notions commonly expressed by adjectives in other languages are expressed by participial forms of stative verbs. Adjectives differ from stative verbs in that they do not carry verbal suffixes (¹tar ‘white’, ³caca ‘small’) although some of them end in a non-alternating final syllable -pa (¹k^heppa ‘big’) which is probably etymologically related to the nominalizer -pa.

Adjectives used as modifiers usually precede the noun. They follow it in kinship expressions, and sometimes for emphasis.

5.6 Definite/indefinite

The numeral ‘one’ and the demonstratives can be used as definite/indefinite markers:

- (29) ⁴kle ⁴kik **ra:ni** ⁴kik ¹mu-pa ²cim
king one queen one be-PAST EVID
¹oca ⁴kle-^hen **ra:ni-mi** ¹p^hjukpo-¹p^hjukpo ¹mu-pa ²cim
that king-and queen-TOP rich-rich be-PAST EVID

‘[Once upon a time] there was a king and a queen. That king and queen were extremely rich.’

Their use is optional:

- (30) *³mi ¹k^ha-ci*
man come-PFV
‘Someone came./There is someone at the door./The man [whom we expected]
came./People (plural) came.’

5.7 Plural, numbers and quantifiers

Plural, or collective, markers and numbers occupy the same slot, after the head noun. Quantifiers have a different status.

There is no compulsory marking of plurality on nouns. An indefinite plurality marker, *-kate*, or a numeral can follow the noun. They are mutually exclusive. *-kate* is adjoined to the plural form of pronouns: *²ai-ni-kate* ‘you all’. *-kate* is a collective rather than a plural; it extends the meaning of the NP to related, not necessarily identical, objects, like English ‘etc.’.

The number system is vigesimal. Numbers are built according to the general syntax of the NP: name of the base followed by multiplicator: *⁴pokal ⁴ni*: |twenty.two| ‘40’. Addition is indicated by the ablative *-se*: *⁴pokal ⁴ni:-se ⁴ni*: |twenty.two-from two| ‘42’ (Mazaudon 2010: 135).

The suffix *-te* ‘as much as’, added to a number or measure (93) gives it an imprecise meaning.

Quantifiers like ‘all’, ‘many’, etc. are either used as full nouns (31) or as adverbs (23) (see section 6.3). They are not adjectives.

- (31) *³mokko-se-n ²se:-pa*
all-ERG-INT know-IPFV
‘Everybody knows.’

5.8 Case markers and particles

Case markers are all toneless suffixes. They include: *-se* ‘ergative’, ‘instrumental’, ‘ablative’, *-la* ‘genitive’, *-ta* ‘accusative’, ‘dative’, *-ri* ‘locative’ (used for time, space, purpose), *-t^hen* ‘sociative’. Temporal *-ma* is bound to some words, like *¹tu-kjuŋ-ma* ‘last year’ (33).

Information structure particles end the NP. They can also apply to VPs or whole clauses (see section 10).

5.9 Possessor NP

The possessive construction marks the dependent NP, or possessor, with the genitive; the head noun is unmarked (4, 6, 38).

This construction is often used to link postpositions to the noun they govern (54, 97).

The genitive is also used to indicate price: *²sika ‘prat-la man’-naj ⁴mar* |rupee eight-GEN *mana*-ful butter| ‘butter at eight rupees a *mana*'; but expressions of amounts using a measure word do not use an overt genitive marker: *makai mana ‘pli* |corn *mana* four| ‘four *manas* of corn’.

5.10 Relative clauses

The verb of the relative clause (RC) is in the participial form in *-pa*, and the RC precedes its head without any intervening connecting marker. This is a difference with Western Tamang and Tibetan, which use a genitive marker on the RC (Mazaudon 1978b). There is no relative pronoun, and the relativized argument is simply deleted from the relative clause: ⁴*tot* ²*pwi-pa* ³*mi* [load carry man] ‘a porter’.

Like any other NP, the head of a RC can be pronominalized by zero. Case markers are suffixed directly to the nominalized verb of a headless RC (32).

- (32) ¹*k^ha-pa-ta* ¹*pin* ¹*to:-pa*, ³*a* ¹*k^ha-pa-ta* ¹*pin* ³*a* ¹*to:*
come-NMLZ-DAT give must-IPFV NEG come-NMLZ-DAT give NEG need
‘We must give [food] to the [people who] come, to [those who] don’t come, we
don’t have to give.’

Tamang can relativize on very peripheral elements in terms of Keenan and Comrie’s accessibility hierarchy (Keenan and Comrie 1977): locative, genitive expressing the material, and even, especially when the head of the relative clause is accompanied by a demonstrative, on elements which have no clear function in the RC (33).

- (33) ¹*tukjuyma* ²*maca-ri* ²*ai-ni-kate* ³*to:-pa* ¹*oca* ¹*mam-¹k^hwi*
last.year down-LOC 2-PL-COLL reach-NMLZ that old.woman
‘That old woman [whose house] you guys went to last year down [i.e. in the city].’

Corelatives and free relatives

Corelatives, where the relativized element is represented both in the RC and in the main clause, are not usual. A structure of this type is found only when relativizing on an adverbial, represented inside the RC by an interrogative adverb, and taken up again in the main clause by a deictic adverb. It conveys an indefinite meaning ‘wherever, however much’. Compare (81).

- (34) ¹*sun* ⁴*ka:te* ¹*to:-pa-cim* ²*ote* ¹*tat-si* ²*to:-ko*
rice how.much need-IPFV-EVID that much take-ANT pound-IMP
‘Take out and polish as much rice as will be required!'

This can also be expressed without a resumptive word in the main clause, and such structures might better be analysed as free relatives:

- (35) ¹*sun* ⁴*ka:te* ¹*ni-pa-cim* ²*to* ¹*to:-ci*
rice how.much go-IPFV-EVID pound must-PFV
‘We must pound the amount of rice that will get used up.’

6 THE VERB PHRASE

6.1 Agreement pattern

Whether the verb be main or subordinate there is no agreement with any argument of the verb, except marginally in the optative/hortative *-ke* ‘first person’ *-kai* ‘third person’. The conditional *-sai* also functions as a sort of ‘first person plural inclusive’ or ‘indefinite third’ of the optative: ¹*ni-ke* ‘let’s go!’ ¹*pin-sai* ‘let us give (alms)’ ¹*tik* ¹*la-sai* ‘what can we do?’

6.2 Word order

The verb phrase comprises the following elements, in order:

Adverb, Negation, Verb, Directional, Modal, {TAM marker or subordinating suffix}.

6.3 Adverbs

Like adjectives, adverbs are rare. Time and place adverbs or prepositional phrases are not part of the VP and are placed freely in the sentence, with a preference for time to precede place (33).

Manner and quantity adverbs are the only ones which modify the verb (as opposed to the sentence) and which are potentially part of the VP or the predicate phrase.

6.3.1 Manner adverbs

Simple words are restricted to ‘fast’ and ‘slow’ and onomatopoeic words (62). They precede the verb or predicate phrase as closely as possible, allowing only the negation or a suffixless subordinated verb to come between them and the main verb. In closely knit expressions like ‘get wood’, ‘eat rice’, ‘speak words’ the object can sometimes come between the adverb and the verb.

Manner is most often expressed by a subordinate clause in *-si* (69) or in *-na* (64).

6.3.2 Quantity adverbs

Quantity words, if they are not NPs (see section 5.7), always modify the predicate and occur in the same position as manner adverbs: between NP arguments and the verb ((23), (36)) or before a complex predicate (37).

- (36) *'sun* ⁴*ca* ⁴*la:mo* ³*a* ¹*k^ha* ¹*mu-la*
rice TOP much NEG come be-NPST
'Rice does not grow well (*lit.* much) [here].'

- (37) *apai*, ⁴*la:man* ¹*suy* ¹*k^heppa*
EXCL much mouth big
'Oh! He is very much a big-mouther (= he is very greedy).'

6.4 Negation

Negation precedes the verb. It has two forms, ⁴*ta* for imperatives and optatives, and ³*a* for all other verb forms.

6.5 Verb

The verb is usually a simple monosyllabic root. A limited number of verbs can be followed by the directional suffix *-ka* which derives from the verb ‘to come’ (¹*k^ha-pa*): ²*pit-ka-pa* ‘to send hither’. Compounds are exceptional (²*pit-hur-pa* ‘to throw to someone’ from ²*pit-pa* ‘to send’ and ¹*hur-pa* ‘to throw’).

A verbal expression consisting of a noun and weak verb sequence (¹*sem* ¹*ta-pa* |feelings happen| ‘to like’, ³*prot* ¹*k^ha-pa* |taste come| ‘to be tasty’) can replace the verb and be constructed like a simple verb ((38), (40)).

-
- (38) *'ya-ta* *³twa-la* *'sja* *³prot* *¹k^ha* *¹mu-la*
1sg-DAT pork-GEN meat taste come be-NPST
'I like pork meat.'

6.6 Modals

Modals are fully toned words, which also exist as full verbs with nominal complements.

Whether a succession of two fully toned verb roots is to be interpreted as a serial construction, as a sequence of main verb + modal, or as a sequence of subordinated verb + main verb is open to debate. In such constructions, the first verb appears as a bare root and the second, which carries the tense-aspect markers, expresses some kind of modality. The details of modal construction are treated later under 'Complex Sentences' (see section 7.3).

6.7 Tense-aspect, mood and sentence final particles

Basic tense-aspect and mood distinctions are expressed by a small set of suffixes which follow the last verb, directional or modal of the clause: *-pa*, present/imperfective; *-ci*, past/perfective; *-la* (*-na* after dentals), non-past/irrealis; *-u/-ko* imperative; *-kel/-kai* optative; *-sai* conditional.⁷ Only main clause or completive clause verbs (complements of verbs of thinking and saying) carry tense-aspect markers.

Negation is incompatible, in Risiangku Tamang, with the perfective suffix. A negated perfective is expressed with the irrealis marker *-la*, followed by the intensifier *-i*: *¹k^ha-ci* [come-PFV] 'he came'; *³a-¹k^ha-la-i* [NEG-come-IRR-INT] 'he did not come' (Mazaudon 1988).

The perfective suffix *-ci* is used with an inchoative sense, as in Nepali ((2), (13), (14)).

Other aspects of the verb are expressed through complex constructions consisting of a subordinate clause followed by a copula or by a small set of verbs like 'finish', 'stay' used as auxiliaries: *V-pa-n* *'mula*, progressive; *V-si* *²cipa*, durative (73); *V-sai* *'mupa*, conditional past (77); etc.

The reported speech particle *-ro* is added to the tensed form of the verb ((55), (93)).

6.8 Subordination suffixes

The logical relationship of the subordinate clause to the main clause is indicated by a set of suffixes added to the last verb root of the subordinate verb phrase: *-sam/-sai/-saka*, condition; *-si*, temporal succession; *-ma*, temporal simultaneity; *-na*, manner; *-pa-ri*, purpose; *-si/-pa-se/-pa-se* *'la-ma*, cause.⁸

7 COMPLEX SENTENCES

7.1 Formal types of subordination

Subordinated verbs or clauses precede the main verb or clause. There are four ways a verb can be subordinated to another.

7.1.1 The bare root

The bare root of the subordinated verb immediately precedes the main verb (except for the negation which comes in between). This construction is used with the modals 'need to', 'want to', 'know how to', 'can', half the 'be about to' sentences, and half of the causatives.

-
- (39) *'ni ⁴ta ²puy-o*
go don't let-IMP
'Don't let him go.'

7.1.2 *Participial form*

The subordinated verb can be in a participial form, or it can be in a finite form followed by a weak verb ('do', 'say'), itself in the required participial form.

- (40a) *²nam ¹k^ha-sam ¹ŋa ³a ¹k^ha*
rain come-if 1sg NEG come
- (40b) *²nam ¹k^ha-ci ³pi-sam ¹ŋa ³a ¹k^ha*
rain come-PFV say-if 1sg NEG come
'If it rains, I won't come.'

7.1.3 *Nominalized form*

The subordinated verb can be in the nominalized form followed by a suffix. Most of these suffixes are homophonous with nominal case markers.

- (41) *²ai-ni-kate ⁴pu ²mra ²sjot-pa-ri ¹ni-ci*
2-PL-COLL field weeds pull.out-NMLZ-LOC go-PFV
'You guys had gone to weed the fields.'

7.1.4 *Juxtaposition*

Two clauses can be simply juxtaposed, and be clearly subordinated in meaning.

- (42) *'ni-la ¹ta-la*
go-FUT happen-FUT
'He might go.'

7.2 Sentential subjects and compleatives

7.2.1 *Sentential subjects*

The sentential 'subject' appears as a topic (or an anti-topic) for the main clause and can carry the topic marker.

- (43) *'ti:la ⁴ni:-ra ²som-ra ¹ta-ci ¹ŋa ¹c^hjai-pa-m*
yesterday two-days three-days happen-PFV 1sg ready-IPFV-TOP
'Yesterday [already], I had been ready for two or three days.'

Formally, a modal notion like epistemic possibility is expressed in the same way as a sentential subject. Example (42) can be paraphrased 'It may happen [that] he will go'.

7.2.2 *Compleatives: verbs of thinking, saying, feeling and wanting*

The most usual way of introducing the complement of a verb of opinion is as a quotation, using a dummy verb 'to say' in the participial form before the main verb 'to say',

‘to think’, etc. The construction ‘having said, he said’ is not specific to Tamang but is an areal feature, found also in Nepali among other languages of the area.

- (44) *paisa* ³*naykar* ³*pa-u* ³*pi-si* ³*pi-m*
money tomorrow bring-IMP say-ANT say-EMPH⁹
‘He said that you should bring the money tomorrow.’ (*lit.*: “Bring the money tomorrow” saying he said.)

Verbs of saying, intention or desire are often constructed without any overt mark of subordination by two juxtaposed finite clauses.

- (45) ¹*ni-la* ⁴*man-pa*
go-FUT like-IPFV
'I want to go.'

7.3 Modality

The most frequent expression of modality is a v-v construction, which could be analysed as a complex verb phrase belonging to a single clause. Since this construction alternates with clearly subordinated structures we treat it as a complex sentence.

Case marking on the participant shared by the two clauses is normally the same as in the corresponding simplex sentence ((47), (79)). With the modal of necessity, if the embedded verb is transitive, the case can shift from ergative to dative (46).

- (46) ¹*ya-ta-m* ¹*tamo* ²*aru-la* ²*p^ha:* ¹*to:-ci*
1sg-DAT-TOP now aunt-GEN repay need-IPFV
'Now I have to repay my aunt's [loan].'

The three main modals of ability are: ²*k^ham-pa* ‘be (physically) able to’ ²*se:-pa*, ‘know how to’, ²*mjaŋ-pa* ‘have time to’.

With these modals the more analytic ‘purpose’ clause construction (see section 7.5) can be used with the same meaning as the v-v construction:

- (47) ²*ai-se* {⁴*pwi* / ⁴*pwi-pa-ri*} ⁴*k^ham-la* ?
2sg-ERG {carry / carry-NMLZ-LOC} can-FUT
'Will you be able to carry it?'

Immediate future is expressed either by a v-v construction with the modal ⁴*tam-pa* or, more commonly, by a purpose clause with the verb ¹*c^hjai-pa* ‘be ready to’ (61).

- (48) ¹*ya-ta* *sarpa-se* ¹*sat* ⁴*tam-pa*
1sg-DAT snake-ERG kill be.about.to-IPFV
'A snake was about to kill me.'

In a modal construction, it is not possible to negate the first (or embedded) verb independently. A subordinated construction has to be used instead.

- (49) *³*a* ¹*ca* ²*k^ham-la* ?
NEG eat can-FUT
'*Can you refrain from eating?'

- (50) ³*a* ¹*ca-na* ²*ci* ²*k^ham-la* ?
NEG eat-ing¹⁰ stay able-FUT
'Will you be able to stay without eating?'

7.4 Causatives

7.4.1 *v²punj-pa*

One of two ways of expressing causation is a v-v construction with the verb ²*punj-pa* ‘to let, allow, order, cause’ as the second verb. The causee appears in the dative if the embedded verb is transitive (51), in the absolute if it is intransitive (52). In this particular construction the embedded verb has to be voluntary or at least active.

- (51) *'t^hen-kat'-ta-n* *'klu* *²punj* *'to:-nam*
3pl-PL-DAT-INT buy make must-INFR
'We have to get them to buy [it] [for us].'
- (52) *sikar- se* *'nak^hi-Ø* *'jar* *²punj-pa*
hunter-ERG dog-ABS run let-IPFV
'The hunter lets his dog run/makes his dog run.'

With negation, *'jar* ^{3a-}*²punj*, the meaning is ‘prevent, not let’ and not a simple negation of the causation. The lower verb cannot be negated independently of the higher verb.

7.4.2 *v-na* *'la-pa*

Causation is often expressed by a ‘result’ alias ‘manner’ clause followed by the verb ‘to do’; hence ‘to act in such a way that v’. The case marking in the subordinate clause remains what it would be in an independent clause:

- (53) *²ani-kat'-se-nun* *'t^hai-na* *'la-u*
nun-PL-ERG-INT hear-so.that do-IMP
'Tell the nuns themselves.' (lit. 'Do so that they hear.')

This construction can always replace the v ²*punj* construction, except for the meaning ‘let’, ‘allow’.

Negation of causation (as opposed to denial of permission) uses the v-*na* *'la-pa* construction:

- (54) *³ca-ta-mi* *³mi-la* *'yon-ti* *icat* *'ni-na* *⁴ta* *'la-u*
son-DAT-TOP man-GEN front-LOC face go-so.that don't do-IMP
'Don't make your son lose face in front of people.'

7.5 Purpose clauses

The most usual way to express purpose is by adding the suffix *-ri* to the nominal form of the subordinated verb (41). The subject of the embedded clause has to be subject (41) or object (55) of the main clause verb.

- (55) *³ca* *¹cjanpa* *⁴ca* *³kruj* *²k^hru-pa-ri* *²pit-cim-ro*
son youngest TOP guts wash-NMLZ-LOC send-PFV-RS
'The story says that they sent the youngest son to wash the guts.'

‘The youngest son’ is the object (ABS) of ‘sent’ and the (deleted) subject (which would be ERG) of ‘wash’.

The embedded verb must be active. Failing this condition an optative construction connected by *³pi-si* ‘having said’ used as a conjunction is employed.

- (56) ³*key* ¹*me-ri* ⁴*po-kai* ³*pi-si* ⁴*cay-pa*
bread fire-LOC rise-OPT say-ANT insert-IPFV
'I put the bread in the fire so that it may puff.'

- (57) ^{*1}*me-ri* ⁴*po-pa-ri* ⁴*cay-pa*

A negative aim has to be expressed by something like 'for fear of' or a negative optative followed by the quotation formula 'saying'.

- (58) ¹*cjan-se* ¹*ya-ta* {¹*ca-la* / ³*a-¹ca-kai*} ³*pi-si* ⁴*jar-pa*
tiger-ERG 1sg-DAT {eat-FUT / NEG-eat-OPT} say-ANT escape-PFV
'I ran away {for fear the tiger would eat me/so the tiger would not eat me}.'

The participle ³*pi-si* used as a conjunction can also be added to the usual *v-pa-ri* construction. It indicates that the aim is not going to be reached:

- (59) ⁴*pu* ³*su-pa-ri* ³*pi-si* ¹*apa* ¹*k^ha-ma* . . .
field plant-NMLZ-LOC say-ANT father come-SIM
'When the father came thinking that he would plant the field . . .'

Using a conditional form ⁴*pu* ³*su-sai* ³*pi-si* [field plant-COND say-ANT] makes the outcome even more doubtful.

The purpose construction is also used for 'try to v' and 'start to v', and optionally with the modals of ability (47).

- (60) ²*nam* ¹*ka* ¹*k^ha-pa-ri* ⁴*mai-ci*
rain FOC come-NMLZ-LOC search-PFV
'It's going to rain.' (lit. 'It's rain that is trying to come.')

(61) ¹*ni-pa-ri* ¹*c^hjai-ci*
go-NMLZ-LOC be.ready-PFV
'He is about to go/he has started.'

7.6 Synchronic relationship: time and manner

Temporal simultaneity is expressed by the suffix *-ma* 'while, when'.

- (62) ¹*sjoŋ* ³*kuj-ri* ³*to:-ka-ma* ¹*sjoŋ-se* ⁴*lap* ⁴*por-ci*
river middle-LOC reach-DIR-SIM river-ERG ADV carry-PFV
'As he was reaching the middle of the river, the river carried him off.'

The addition of a topic marker facilitates a causal reading:

- (63) ³*a-¹k^ha-ma-m*, ⁴*mai* ¹*ni-ci*
NEG-come-SIM-TOP search go-PFV
'As he did not arrive, they went to look for him.'

Manner or result clauses are formed with the suffix *-na* 'in such a way that', 'with the result that'. This construction is rare in the affirmative, except as a way to express causation with the head verb 'to do' (see section 7.4.2).

- (64) ³*rit-si* ⁴*pra-pa* ¹*pa:ci* ²*mren-na* ¹*ca-ci*
beg-ing walk-NMLZ beggar satiated-so.that eat-PFV
'He, a beggar, who lived on begging, ate to his full.'

Terminus ad quem

The notion of ‘until’ or ‘before’ is expressed via a grammaticalized use of the verb ‘to reach’, in the manner-participial form ⁴*to:-na*, or the Nepali loan equivalent *samma*, following a negative form of the verb bearing the suffix *-te*, which indicates extent.

- (65) *'jay* ³*maima-m* ³*a* ¹*si-te* *samma-m* ³*rit-sai* ¹*ca-sai*
 1pl women-TOP NEG die-as.long.as until-TOP beg-COND eat-COND
‘As for us, women, until we die, let us beg for our food.’

Logical implicature

A logical implicature (without causal or temporal connection) is expressed by the simultaneity participial form of the verb *v-ma* followed by the postposition ⁴*ŋjancʰem* ‘from’.

- (66) *'oca* ¹*kan* ³*a* ²*min-ma* ⁴*ŋjancʰem* . . .
 that rice NEG cook-SIM since . . .
‘Since that rice did not cook . . . (reporting on an experiment where the fire was three yards away from the pot.)’

7.7 Logical or temporal succession

Temporal antRIORITY is expressed by a participial clause in *v-si*.

- (67) *'nana* ¹*kʰa-si* ¹*capasai* ³*ŋjo-si* ¹*ca-ci*
 sister come-ANT food cook-ANT eat-PFV
‘Nana came, we cooked food, and we ate.’

This is also the most common way of expressing a cause.

- (68) ²*nam* ¹*kʰa-si* ⁴*pu-ri* ¹*ni* ³*a* ²*mjay-lai*
 rain come-ANT field- LOC go NEG can-IRR
‘Because it rained, we could not go to the field.’

If the action of the verb of the main clause can be construed as expressing a terminal point for the whole process, the action of the subordinate clause in *-si* need not be anterior, it can express a manner.

- (69) ¹*sja-si* ¹*sja-si* ¹*kʰa-ci*
 dance-ANT dance-ANT come-PFV
‘She came dancing all the way.’

Like the finite suffix *-ci* ‘perfective aspect’, *-si* indicates that the action has or will definitely happen. It cannot be negated. A negated manner clause is used instead (50).

- (70) ²*nam* ³*a* ¹*kʰa-na* ⁴*pu* ³*su* ³*a* ²*mjay-lai*
 rain NEG come-ing field plant NEG can-IRR
‘Because it did not rain, we are not able to plant the fields.’

Auxiliary constructions

With the head verbs ‘give’, ‘finish’ and ‘stay’ used as auxiliaries, the *v-si* construction expresses the notions of ‘doing something for someone’ (71), completion (72) and duration or protracted action (73). Comparable constructions are found in Nepali.

-
- (71) *'am'-se ¹*kwan-si ¹*pin-pa***
mother-ERG dress-ANT give- IPFV
'His mother dresses [him].'
- (72) *'t^he-jen ¹*t^hai-si ³*cin-ci***
they-too hear-ANT finish-PFV
'They too have already heard [the news].'
- (73) *³*mi ⁴*pra-si-n* ²*ci-pa***
people walk-ANT-INT stay-IPFV
'There are always passers-by./People keep walking by.'

7.8 Conditionals

Hypothetical verb forms are built on the verb root augmented with the suffix *-sa*, always followed by a second suffix which differentiates several kinds of conditions. The origin of these second suffixes as discourse markers is discussed in Mazaudon (2003). These forms have been grammaticalized and are not synchronically analyzable.

7.8.1 Simple conditionals

Simple conditionals are etymologically formed with the addition of the topic marker *-m(i)*.

- (74) *²*nam ¹*k^ha-sa-m* ¹*ŋa ³*a ¹*k^ha*****
rain come-if-TOP 1sg NEG come
'If it rains, I won't come.'

The verb of the main clause can be in the indicative mood as in (74). It can also be in the conditional mood, expressing greater doubt as to the certainty of the event.

- (75) *¹*iampu-ri ¹*mar ¹*ni-pa ³*mi ¹*mu-sam,*****
Kathmandu-LOC down go-NMLZ man be-if
*¹*ŋa-m ³*cakir ¹*ca-pa-r'****
1sg-TOP salary eat-NMLZ- LOC go-COND
'If someone was to go down to Kathmandu, I would go [with him] to look for a job
(lit. to earn a salary).'*

A necessary condition 'only if' is expressed by a disjunction of the opposite conditions.

- (76) *^{t^h}*unta ²*pit-ka-sa-m* ¹*ŋa ²*so-la, ^{t^h}*unta ³*a-²*pit-ka-sa-m* ¹*ŋa ¹*si-la******
food send-DIR-if-TOP 1sg live-FUT food NEG-send-DIR-if-TOP 1sg die-FUT
'If you send me supplies I will survive, otherwise I will die.'**

7.8.2 Counterfactuals

If the verb of the main clause is in the conditional past tense (*V-sai ¹*mupa**), the meaning becomes counterfactual.

- (77) *⁴*kjarca ²*som ¹*la-sai ¹*mu-pa, ²*arku-se ³*hin-sa-m*****
hundred three make-COND be-PAST¹¹ other-ERG be-if-TOP
'They would have made it 300 Rs, had it been someone else (= if someone else had fixed the fine).'**

With counterfactuals, the subordinate clause usually replaces the topic marker *-m* by the focus marker *-ka*, which corresponds well to the pragmatic use of counterfactuals: the new, important, information is usually in the subordinate clause.

- (78) *'sar 'ju:-sa-ka ⁴mrai-sai*
manure pour-if-FOC swell-COND
'If you'd put MANURE on, it might have grown!'

Even in the absence of any main clause, the reading of *-sa-ka* is counterfactual: *²se:-sa-ka* 'if only I had known!'

7.8.3 Emphatic conditionals and concessive conditionals

If we add to the verb carrying the hypothetical suffix *-sa* an intensifying suffix *-i* 'also', yielding a conditional form similar to that used in main clauses, the meaning is more hypothetical, whence 'if ever' (79).

- (79) *²c^hjai ¹ta-sa-i, ²ai-se ²p^ha: ¹to:-pa ³ahin*
debt occur-if-INT 2Sg-ERG pay must-IPFV isn't it?
'And if ever he incurred a debt, you would have to repay it, right?'

The emphatic conditional *-sa-i* is also used to express a concessive condition 'even if'. In this meaning it is often strengthened by a second intensifier *-nun*.

- (80) *¹t^he ³a-¹k^ha-sai(-nun) ¹ya ¹ni-la*
he NEG-come-even.if(-INT) 1sg go-FUT
'Even if he won't come, I will go.'

A 'universal concessive conditional' (see Haspelmath and König 1998) is expressed by an interrogative/exclamative word with the emphatic conditional *-sai*.

- (81) *⁴kjat ⁴kate-n ¹la-sai ³a ³cin-pa*
work how.much-even do-even.if NEG finish-IPFV
'However much work I may do, it is n[ever] finished.'

8 COMPARISON

Word order: standard + marker + predicate.

- (82) *¹ya banda ²ti:-tir-la ³puriy*
1sg than one-year- GEN younger.sister
'She is one year younger than me. (she is not necessarily my sister).'

The morpheme *banda* is from Nepali *bhanda*, meaning 'saying'. A Tamang calque *³pi-ma* is also used.

Whole clauses can appear as the compared terms.

9 QUESTIONS

9.1 Yes–no questions

Yes–no questions are most usually marked by intonation only (1). A particle *wa*, or the loan word *ki* can occur in sentence final position:

- (83) *⁴me ¹mama-i ¹ni-ci wa ?*
cow female-also go- PFV Q
'Did the mother-cows also go?'

The particle *wa* can follow a focalized word, phrase or clause in a yes–no question.

The question marker *ki* at the end of the sentence can co-occur with *wa* after the term questioned. The question is made even more emphatic by adding *'tik* ‘what’ after *ki*.

- (84) *³plo:-si-wa ¹t ²an-pa ki ¹tik ?*
boil-ANT-Q put-IPFV or what
'Is [that egg] boiled or what [that you handle it so roughly]?'

The verb can be the term emphatically questioned. In that case it is repeated:

- (85) *²cʰat-wa ²cʰat-pa-ki ²to-ca-kate-m ?*
fight-Q fight-IPFV=Q up-that-COLL-TOP
'Are they fighting for good, the people upstairs?'

Alternative questions are frequently used.

- (86) *²na ⁴yacʰa ³to:-ci ²na ¹licʰa ²ci-ci*
Q ahead reach-PFV Q behind stay-PFV
'Were they in front or had they stayed behind?'

9.2 Interrogative word questions

The interrogative word remains in the position of the questioned NP.

- (87) *²osem ²kʰacjupa guru ²sjap-ci ?*
thus which guru meet-PFV
'Then, which guru did you get to see?'

9.3 Interrogatives as exclamatives and indefinites

Exclamations are made with ‘wh-’ question words: *'kate ³ce:pa!* |how.much pretty| ‘how pretty!?’; indefinites are made with reduplicated ‘wh-’ words.

- (88) *²kʰaima ²kʰaima ¹klay-pa*
when when play-IPFV
'He gambles occasionally.'

Negative polarity indefinites

In all the languages of the area, including Nepali, the most common way to say ‘no one’, ‘nothing’, ‘never’, etc. is to use an interrogative word suffixed with the intensifier ‘also’/‘even’ and a negated verb ((89), (98)).

- (89) *'tik-la-i ¹tik-la-i ³a ¹pin-nai-ro*
what-GEN-INT what-GEN-INT NEG give-IRR-RS
'She says they did not give her anything at all.'

Note that the reduplication of the indefinite intensifies it (89), as it does an adjective (29), while the repetition of the ‘wh’-word makes it indefinite (88).

Negative polarity indefinites can also be used in questions.

- (90) *'tik-la-i ¹cjana ³raypa ¹mula?*
what-GEN-INT strainer like is?
'Do you have anything like a strainer?'

10 INFORMATION STRUCTURE

Tamang possesses two topic markers, *-mi/-m* for simple topic and *⁴ca* for contrastive topic, one focus marker *-ka/¹ka*; and a set of intensifiers *-i/-e/-ja* ‘even’, ‘also’ and *-n/-nun* ‘own’, ‘self (intensifier, not reflexive)’. These markers code the information structure of a sentence without interfering with its grammatical structure; case marking and word order can remain unchanged when information structure markers are added.

10.1 Topic and topicalization

If we define ‘topic’ (following Lambrecht 1994) as ‘the matter of current concern’ about which new information is being added in an utterance, the most common surface expression of topic in Tamang, whether in conversation or in running text, is zero. Identifiable, currently topical arguments are not mentioned.

The next most common form a topic can take is as an anti-topic – several of them if need be – placed at the end of the sentence, after the verb, accompanied by lowered intonation (23).

If an active topic has to be mentioned in its proper grammatical place in the sentence, it is marked with the topic marker *-m(i)*. This structure occurs in story telling where the cohesion of discourse is ensured by the formal repetition of the topic from one sentence to the next (29).

A topic which is accessible but not active (present in the slightly more distant context) is reactivated by being mentioned with the topic marker *-m(i)*. This is the case for a demonstrative summing up a situation, or for a pronoun referring to the speaker (91).

- (91) *'oca-m 'ya-i-mi ⁴ta: ³are, mai*
that-TOP 1sg-ERG-TOP knowledge not.be Mai
‘That, as far as I am concerned, I don’t know, Mai.’

A contrastive topic is marked by the tonal particle *⁴ca*, possibly a loanword from Nepali *cahī*. ((12), (55)).

Clauses also can be topicalized with either topic marker. See (43).

- (92) *'si-pa ⁴ca 'ya-i patta ³are*
die-NMLZ TOP 1sg-ERG knowledge not.be
‘That she had died, I did not know.’

Topic marker on subordinate clauses

On subordinate clauses, *-m(i)* explicitly sets the clause as background ((63), (65)). In (93) it prepares thus a dramatic effect expressed in the main clause. Compare with the unemotional reporting in (62).

- (93) *³kuy-te-ri ³to:-ka-ma-m ¹apa ¹sjoy-se ⁴lap ⁴por-ci-ro*
middle-about-LOC reach-DIR-SIM-TOP father river-ERG ADV carry-PFV-RS
‘But as he was reaching the middle of the river, the father was carried off by the river.’

On plain conditionals, it has become grammaticalized (Mazaudon 2003), corresponding to the now well accepted idea that ‘conditionals are topics’ (Haiman 1978) (see section 7.8).

10.2 Focus

Plain predicate focus and sentence focus structures contain no focus marking morpheme. Strong narrow focus (or contrastive focus) is marked with the suffix *-ka*. It can occur on all elements of the sentence including the predicate, nominal or verbal.

10.2.1 Argument focus

- (94) *²ai- la ⁴mar-ka ⁴ni:-nun ²cunj-o ³pi-pa ¹ŋa-i-mi*
2sg-GEN gold-FOC two-INT sell-IMP say-IPFV 1sg-ERG-TOP
'It is your gold [earrings] that I said to sell both of [I did].'

10.2.2 Predicate and sentence focus

- (95) *¹ŋa-ka*
1sg-FOC
'It's me!' (typical answer at the door)

The marker *-ka* on a verbal predicate does not necessarily indicate a contrastive narrow focus on the verb. It can simply make the assertion strongly affirmative (or negative as the case may be), indicating that the assertion is contrary to expectation (contrary to the presupposition the hearer is thought to hold, or contrary to his or her wish). Its domain in that case is the sentence or the predicate.

- (96) *¹ŋa-i-ja ¹nipa-ka, ¹apa*
1sg-ERG-also go-FOC father
'I want to go too, Daddy.'

The focus marker which semantically applies to the whole predicate can be affixed to the object of the verb (97) or to its subject in the case of sentence focus (60).

- (97) *²ai-se-mi ¹ŋa-la ¹tʰo- ri ¹cʰa:mo-ka ³pa-ci*
2sg-ERG-TOP 1sg-GEN top-LOC co.wife-FOC bring-PFV
'You brought a CO-WIFE on me.'

10.2.3 Focus marker on subordinate clauses

When used on a subordinate clause, the narrow focus marker *-ka* has a contrastive value comparable to clefting in English:

- (98) *¹ti:la ¹ŋa ¹oray ³ŋjot-la ³pi-si-ka ⁴pay*
yesterday 1sg that.way drunk-FUT say-ANT-FOC strength
³are-pa-ri ²kʰana-i ³a-²wan-pa
not.be-NMLZ-LOC where-INT NEG-enter-IPFV
'Yesterday it was for fear of getting drunk, like that, in my weak condition, that I did not enter anywhere.' (Because any host would have offered drinks.)

As we have seen, the focus marker is almost systematically used on counterfactual conditionals (78).

ADDITIONAL ABBREVIATIONS

Words in ***bold italic*** are loans from Nepali. Abbreviations specific to this section are:

ANT	participle: anteriority/simultaneity/cause, ‘V-ing’ ‘having V-ed’
COND	conditional mood
EMPH	emphatic
EXCL	exclamation
INT	intensifier
IRR	irrealis
<i>lit.</i>	literally
NMLZ	nominalizer (infinitive or participle)
NPST	non-past
RS	reported speech
SIM	participle: simultaneity/cause, ‘when’ ‘as’ ‘since’
Ø	zero marker
*	ungrammatical structure
[]	word added in translation
()	word deleted from translation

NOTES

- 1 M. Noonan (this volume) uses the term ‘Tamangic’ for this group, an improvement over Shafer in as much as the endonym used by most members of the group is ‘Tamang’ rather than ‘Gurung’, but problematic if we want to respect Shafer’s sensible proposition to reserve the ending ‘-ic’ for large families (Indic, Sinitic). If a name derived from the name of one language of the group is desired, Shafer’s system would call for ‘Tamangish’ as a language-level, not family-level grouping.
- 2 In Risiangku Tamang the copulas *'mu-la* and *'hin-pa* are only found with non-perfective suffixes. In this context the suffix *-la* does not have its usual future/irrealis value, but is plain non-past or imperfective. *'mu-la* cannot be negated; a substitutive root *'are* is used instead.
- 3 These variations are the closest Tamang gets to voice: (24) could be analysed as a sort of ‘impersonal passive’: ‘There is no reward finding, as far as you are concerned’, whence the dative.
- 4 Words in ***bold italic*** are poorly assimilated (toneless) loans from Nepali.
- 5 The order of these elements depends on the tonal or non-tonal status of the particle.
- 6 *'ray* is not a reflexive *per se*. Sentence (26) could, although semantically surprisingly, be construed to mean ‘The shaman checks by hitting *you* (politely expressed as ‘[your] own person’).
- 7 This refers to the conditional mood (‘would’, glossed COND), used in the main clause, not to the participial form of the verb in the subordinate ‘if’ clause, described in section 6.8.
- 8 This very rough categorization, and the glossing labels *-si* ‘ANT’ (‘anterior’), *-ma* ‘SIM’ (‘simultaneous’), do not do justice to the complexity of the meanings and uses of the subordinating suffixes which are explained in the rest of this chapter. See for instance (69) and (71) to (73) for non-anterior usages of *-si*.
- 9 The emphatic marker *-m* should not be confused with the topic marker *-mi/-mi*; we believe that the emphatic marker *-m* derives from the copula *'mu(-la)* ‘to exist’.
- 10 The suffix *-na* is used as the negative-context variant of *-si* (see section 7.7 and (70)).
- 11 On the special conjugation of the auxiliary/copula *'mula*, *'mupa* see section 4.1.2 (note 2).

REFERENCES

- Burton-Page, John (1955) ‘Two studies in Gurungkura: I. Tone; II. Rhotacisation and retroflexion’, *Bulletin of the School of Oriental and African Studies* 17(1): 111–119.
- Everitt, Fay (1973) ‘Sentence patterns in Tamang’, in Ronald Trail (ed.) *Patterns in Clause, Sentence, and Discourse in Selected Languages of India and Nepal*, Vol. 1. SIL publications in linguistics and related fields 41. Norman: SIL, University of Oklahoma, 197–234.
- Georg, Stephan (1996) *Marphatan Thakali*. Munich: Lincom Europa.
- Glover, Warren (1974) *Semantic and Grammatical Structures in Gurung (Nepal)*. SIL publications in linguistics and related fields 49. Norman: SIL, University of Oklahoma.
- Greenberg, Joseph (1966) *Language Universals*. The Hague, Paris: Mouton.
- Grierson, George A. (1908) *Linguistic Survey of India*, Vol. 3, Tibeto-Burman family, reprint Delhi: Motilal BanarsiDass 1967. Calcutta: Office of the Superintendent of Government Printing.
- Haiman, John (1978) ‘Conditionals are topics’, *Language* 54: 565–589.
- Hale, Austin and Kenneth Pike (eds) (1970) *Tone Systems of the Tibeto-Burman Languages of Nepal*, Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics 3. Urbana, Illinois: Publications of the Department of Linguistics University of Illinois.
- Hale, Austin and David Watters (eds) (1973) *Clause, Sentence and Discourse Patterns in Selected Languages of Nepal*. SIL publications in linguistics and related fields 40. Norman: SIL, University of Oklahoma.
- Haspelmath, Martin and Ekkehard König (1998) ‘Concessive conditionals in the languages of Europe’, in J. Van der Auwera in collaboration with Dónall P. Ó. Baoill (eds) *Adverbial Constructions in the Languages of Europe*, Berlin: Mouton de Gruyter, 563–640.
- Haudricourt, André-Georges (1961) ‘Bipartition et tripartition des systèmes de tons dans quelques langues d’Extrême-Orient’, *Bulletin de la Société de Linguistique de Paris* 56(1): 163–180.
- Haudricourt, André-Georges (forthcoming) *Studies in the Evolution of Languages and Techniques*. Trends in Linguistics Series Monograph 270. Berlin: de Gruyter-Mouton.
- Hoefer, Andras (1981–1997) *Tamang Ritual Texts*, Beiträge zur Südasiensforschung 65. Wiesbaden: Franz Steiner, Vol. 1, 1981, Vol. 2, 1997.
- Keenan, Edward L. and Bernard Comrie (1977) ‘Noun phrase accessibility and universal grammar’, *Linguistic Inquiry* 8: 63–99.
- König, Ekkehard (2001) ‘Intensifiers and reflexive pronouns’, in Martin Haspelmath, Ekkehard König, Wulf Oesterreicher and Wolfgang Raible (eds), *Language Typology and Language Universals. An International Handbook*, Vol.1. Berlin, New York: de Gruyter. 747–760.
- Lambrecht, Knud (1994) *Information Structure and Sentence Form*. Cambridge Studies in Linguistics 71. Cambridge: Cambridge University Press.
- Mazaudon, Martine (1973) *Phonologie du Tamang*. Paris: SELAF.
- Mazaudon, Martine (1976) A typological sketch of Tamang. mimeogr. 36pp. Seminar on linguistic typology conducted by Professor Edward L. Keenan at the UCLA Linguistic Department, Fall 1976.
- Mazaudon, Martine (1977) ‘Tibeto-Burman tonogenetics’, *Linguistics of the Tibeto-Burman Area* 3(2): 1–123.
- Mazaudon, Martine (1978a) ‘Consonantal mutation and tonal split in the Tamang sub-family of Tibeto-Burman’, *Kailash* (Kathmandu) 6(3): 157–179.

- Mazaudon, Martine (1978b) ‘La formation des propositions relatives en tibétain’, *Bulletin de la Société de Linguistique de Paris* 73(1): 401–414.
- Mazaudon, Martine (1988) ‘Temps aspect et négation en tamang’, in Nicole Tersis and Alain Kihm (eds) *Temps et Aspects*, Paris: Péeters/SELAF, 179–185.
- Mazaudon, Martine (2000) Tamang recorded texts, with synchronized transcription and translation. Available at: lacito.vjf.cnrs.fr/pangloss/languages/Tamang_en.php (accessed 13 August 2016).
- Mazaudon, Martine (2003) ‘From discourse to grammar in Tamang: topic, focus, intensifiers, and subordination’, in David Bradley, Randy LaPolla, Boyd Michailovsky and Graham Thurgood (eds) *Language Variation: Papers on Variation and Change in the Sinosphere and in the Indosphere in Honour of James A. Matisoff*. Canberra: Australian National University, 145–157.
- Mazaudon, Martine (2010) ‘Number-building in Tibeto-Burman languages’, *North East Indian Linguistics* 2: 117–148.
- Mazaudon, Martine and Alexis Michaud (2008) ‘Tonal contrasts and initial consonants: a case study of Tamang, a “missing link” in tonogenesis’, *Phonetica* 65: 231–256.
- Shafer, Robert (1955) ‘Classification of the Sino-Tibetan languages’, *Word* 11(1): 94–111.
- Taylor, Doreen (1973) ‘Clause patterns in Tamang’, in Austin Hale and David Watters (eds) *Clause, Sentence, and Discourse Patterns in Selected Languages of Nepal*, Vol. 2. Norman: SIL, University of Oklahoma, 81–174.
- Varenkamp, Bryan K. (1996) *Tamang Tam. A Sociolinguistic Study of Eastern Tamang Dialects (in Nepal)*. Kirtipur (Nepal): Center for Nepal and Asian Studies, Tribhuvan University.
- Yonjan-Tamang, Amrit (1997 [2054 B.S.]) *Tamang vyakaran [Tamang Grammar]*. Kathmandu: Royal Nepal Academy (in Nepali).
- Yonjan-Tamang, Amrit (2014 [2070 B.S.]). *Tamang br̥hat śabdakoś [Tamang Large Dictionary]*. Tamang–Nepali, Nepali–Tamang, with a Tamang–Nepali wordlist by semantic category. Kathmandu: Yamburi Book Point.

CHAPTER TWENTY-FIVE

CHANTYAL¹

Michael Noonan and Kristine A. Hildebrandt

1 INTRODUCTION

The Chantyal language is spoken by approximately 2,000 of the 10,000 ethnic Chantyal. The Chantyal live in the Baglung and Myagdi Districts of Nepal; the villages where the Chantyal language is spoken are all located in the eastern portion of the Myagdi District and include the villages of Mangale Khāni, Dwāri, Ghya Kharkā, Caura Khāni, Kuine Khāni, Thārā Khāni, Pātle Kharkā, Mālāmpāhār and Malkābāng. There is relatively little linguistic variation among these villages, though where differences exist, it is the speech of Mangale Khāni, that is represented here.

The Chantyal language is a member of the Tamangic group (along with Gurung, Thakali, Nar-Phu, Manange and Tamang, the last three of which are discussed in this volume). Within the group, it is lexically and grammatically closest to Thakali. Assessment of the internal relations within the group is complicated by a number of factors, among which is the fact that shared innovations may be the product of geographic contiguity (language contact) as much as shared genetic background. At the moment, the most likely classification is as in Figure 25.1.

Chantyal, however, is in many respects the most deviant member of the group, lacking a tone system and having borrowed a large portion of its lexicon from Nepali. In addition, there appears to be a layer of Tibeto-Burman vocabulary that is not Tamangic.

2 PHONOLOGY

2.1 Vowels

The vowel system of Chantyal consists of a set of six vowel phonemes and their nasalized counterparts. Distinctive vowel length is a marginal part of the system, however, and long nasal vowels are quite rare: in general, vowel length is the product of fairly recent – and still rather unstable – processes of syllabic coalescence.

- /i/: pronounced [i].
- /e/: pronounced ordinarily at a point roughly midway between [ɛ] and [e].
- /ə/: when stressed, is either [ʌ] or [ʊ], sometimes becoming rounded and sounding almost like [ɔ]. When unstressed, it is pronounced [ʌ]. /əə/ (i.e. long /ə/) is pronounced [ɜ].
- /a/: pronounced [ɑ].
- /o/: pronounced ordinarily at a point roughly midway between [ɔ] and [o].
- /u/: pronounced [u].

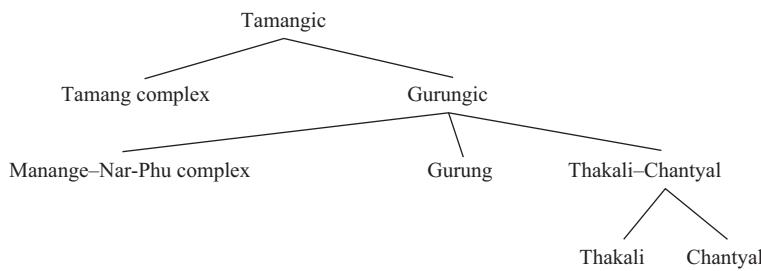


FIGURE 25.1 RELATIONS AMONG THE TAMANGIC LANGUAGES

All possible combinations of on-glide /y/ and /w/ and vowel are attested. For off-glides, the following are attested:

- /iw/
- /ew/ /ey/
- /əw/ /əy/
- /aw/ /ay/
- /oy/
- /uy/

2.2 Consonants

Chantyal contrasts four points of articulation: bilabial, dental, alveolar/alveolo-palatal and velar. The dental point of articulation is lamino-dental. The alveolar/alveolo-palatal series consists of a set of oral affricates whose stop portions are alveolar and whose fricative portions are alveolo-palatal before front vowels (/ci/ = [tei]) and alveolar elsewhere (/ca/ = [tsa]); the fricatives show a similar distribution in being alveolo-palatal before front vowels and alveolar elsewhere. The nasals in this series are postalveolar, whereas the tap approximants are apico-alveolar.

The Chantyal consonant inventory is rich in contrasts involving voice onset time and murmur. Typical of the South Asian speech area, Chantyal contrasts voiceless, voiceless aspirated, voiced and murmured stops. In addition, Chantyal has stops with voiceless and voiceless aspirated onsets followed by murmur.² Nasals, approximants, fricatives and glides also contrast murmured and non-murmured phonemes. In the transcription used here, <h> indicates aspiration, <f> murmur.³

Unaspirated stop	p	t	k
Aspirated stop	ph	th	kh
Voiced stop	b	d	g
Murmured stop	bɦ	dɦ	gɦ
Murmured stop with voiceless onset	pɦ	tɦ	kɦ
Murmured stop with voiceless aspirated onset		thɦ	khɦ
Voiced nasal stop	m	n	ny
Murmured nasal stop	mɦ	nɦ	nɦy
Unaspirated affricate			c
Aspirated affricate			ch
Voiced affricate			j

Murmured affricate	jfi
Murmured affricate with voiceless onset	chf
Voiced lateral approximant	l
Murmured lateral approximant	lh
Voiced tap approximant	r
Murmured tap approximant	rhi
Voiceless fricatives	s
Murmured fricatives with voiceless onsets	shf
Glides	y w
Murmured glides	hy fw

As the chart above shows, there are gaps in this system: /phfi/ and /chfi/ are unattested, and while /thfi/ and /khfi/ are attested, they are rare. The lack of attested /ŋfi/ likely reflects the absence of /ŋ/ morpheme initially: murmured consonants are generally restricted to morpheme-initial position, except in some borrowings.

Geminate consonants occur and are distinctive, but they are found only intervocally within morphemes.

Retroflex consonants are heard very occasionally in a few Nepali borrowings.

2.3 Tone and stress

Chantyal is the only Tamangic language which is not tonal. Most likely, the massive influx of Nepali vocabulary contributed to the loss of the tonal system. In native vocabulary, primary stress is on the first syllable. In borrowed Nepali vocabulary, stress follows Nepali stress rules.

2.4 Phonotactics and phonological alternations

Murmured consonants occur only word-initially, except in some recent Nepali and English borrowings, wherein borrowed murmured consonants (/bf/, /df/, /jf/, and /gf/) occur word internally in words of Nepali origin and /bf/ occurs finally in words of Nepali and English origin: /bf/ is used in borrowings from English where it corresponds to English /v/.

All consonantal segments may occur word-initially save /ŋ/. Word-initially, in native vocabulary, Chantyal permits only clusters of consonant + glide. Borrowings from Nepali and English permit also initial clusters of stop + liquid; however, initial clusters of /tl/ and /dl/ do not occur and the alveolar/aveolo-palatal affricates do not form clusters with liquids.

Word medially, clusters of moderate complexity may occur. The following sorts are attested (where G = glide, N = nasal, K = obstruent (stop, affricate or fricative), L = liquid):

KG	<i>əbyala</i>	‘late [N]’	GK	<i>bāwso</i>	‘hoe’
KL	<i>bakhra</i>	‘goat [N]’	LK	<i>arko</i>	‘next [N]’
KK	<i>bəkselu</i>	‘larva’	KK	<i>tuktuk</i>	‘hacking’
KN	<i>thutno</i>	‘snout [N]’	NK	<i>əncəl</i>	‘province’
NG	<i>kajjyo</i>	‘comb [N]’	GN	<i>bhyawni</i>	‘ghost’
NL	<i>ayla</i>	‘joint’	LN	<i>jurni</i>	‘joint’ [N]

Prior to the massive influx of Nepali and, recently, English words, allowable word-final segments only included vowels, nasals, liquids and plain stops (i.e. not affricated and not

aspirated). However, with the recent borrowings, all segments but murmured stops (with the exception of /bf/) occur word-finally.

All vowels may occur initially, medially and word-finally.

There are few phonological or morphophonological processes commonly encountered in moderately careful speech. Of these, the most important are the reduction of /wa/ in suffixes to [o], and /wā/ and /ma/ to [ō], the weakening of morpheme initial /s/ to [h] in suffixes, and the voicing of the alveolo-palatal affricates [/c/, /ch/, /ch/] to [j/ in intervocalic position. There is also the phenomenon of ‘emphatic’ gemination, whereby the last intervocalic consonant is geminated.

3 MORPHOLOGY

3.1 Generalizations

Chantyal is overwhelmingly suffixing and agglutinative. There are only two sorts of native prefixes: the negative prefixes *a-* and *tha-* on verbs and adjectives, and the deictic prefixes *yi-* ‘this’, *hi-* ‘that’ and *wu-* ‘yonder’. A small number of prefixes may be found with Nepali borrowings, but these are not productive in Chantyal.

3.2 Nouns

Nouns are inflected for number: the plural is marked with the suffix {-ma}; the singular is unmarked. Number is not obligatorily marked on notionally plural count nouns, but it is commonly so marked and can be found even with nouns quantified by numeral or non-numeral quantifiers: even the presence of a classifier does not rule out overt number marking, though not many such cases have been recorded:

tin-ta jəmməy naku-ma
three all dog-PLURAL
'all three dogs'

The plural suffix is often found with the first person plural pronoun, e.g. *nhi-ma*, a form which is already fully specified for plurality; other expressions fully specified for plurality may also contain the plural morpheme:

nə-ye ama bəw-ma
1sg-GEN mother father-PL
'my parents'

The plural suffix is also used collectively: *Ram-ma* means ‘Ram and his family/companions’.

Case is marked on nouns by means of a large number of case enclitics. At this stage in the history of the language, these forms are clearly clitics since, for example, only the last NP in a set of conjoined NPs need be specified for case. The clitics, however, form a tight phonological bond with the nouns to which they are affixed and in a very few cases condition idiosyncratic changes on the nouns:

na '1sg' *nə-ye* 'my' [1sg-GENITIVE]
thim 'house' *them-əŋ* 'homeward' [house-LOCATIVE]

The case clitics are listed below:

Absolutive	unmarked
Ergative/Instrumental	-sə
Ablative	-gəm-sə
Elative	-n̪ha-ri-gəm-sə
Abessive-proximal	-nasa-w
Dative	-ra
Allative	-nas
Locative	-ri
	-əŋ [with a few common nouns, deictics, case clitics]
	-chə [with the noun <i>mfiun</i> ‘night’]
	-ra
Inclusive	-muwa
Adessive	-muwa-ri
Inessive	-n̪ha-ri
Genitive	-ye, -i
Allative/Comitative	-siŋ, -səŋ, -səŋə, -səŋgə (N, except? -siŋ)
Comitative	-ru
Circumulative	-mar
Subessive	-phiri
Sublative	-phiri-ŋ-sə
Superessive	-phyara-ŋ-sə
Temporal	-ma (N)
Comparative/Temporal	-bhənda (N)
Comparative/Temporal	-d̪hin, -d̪ikhin (N)
Comparative/Temporal	-rasə
Essive/Comparative	-thɔy

(The forms followed by ‘(N)’ are borrowings from Nepali.) As shown above, the absolute case is indicated by the absence of any case clitic. The case clitics can be combined within a single NP, and frequently are:

thim-n̪hari-gəmsə
house-INESSIVE-ABLATIVE
'out from inside the house'

Indeed, many of the case clitics presented in the chart above originated as combinations of case clitics which have become grammaticalized. I have inserted hyphens in these forms to show the historical components.

Within the nominal word, the order is:

NOUN — PLURAL — CASE

Chantyal is consistently ergative in case marking (transitive subjects are ergative; intransitive subjects and direct objects are absolute), but does not demonstrate syntactic ergativity; that is, grammatical processes that refer to subjects refer to ergative and absolute marked subjects equally and no grammatical processes also refer to absolutes

regardless of grammatical role. Ergative subjects may be agents, but they may also represent other semantic roles, for example experiencers, as in the following:

<i>khi-sə</i>	<i>uttər</i>	<i>thāya-m</i>
3sg-ERGATIVE	answer	know-NON.PAST
‘He knows the answer.’		

Animate direct objects are typically marked as datives, i.e. Chantyal evidences ‘anti-agentive’ marking. The use of the genitive (as opposed to simple juxtaposition of an absolute NP) to mark genitival relations is used in the majority of cases but is not obligatory. Partitives are formed by juxtaposition of the measure word with the substance measured:

<i>yek</i>	<i>gilas</i>	<i>cha</i>
one	glass	tea
‘one glass of tea’		

The locative case is used with both static (locational) and dynamic (allative) senses: *hyunda-ri* ‘in the winter’, *Bini-ri* ‘to Beni’. The names of the other cases are intended to provide a general indication of their meaning and use.

Comparative constructions are formed using one of the comparative case forms:

<i>nə-ye</i>	<i>naku</i>	<i>khi-ye</i>	<i>naku-bhənda</i>	<i>thyo</i>	<i>mu</i>
1sg-GENITIVE	dog	3sg-GENITIVE	dog-COMPARATIVE	big	be+NON.PAST
‘My dog is bigger than his dog.’					

There is also a topic/focus particle *nə*, which occurs always in NP-final position following case and number morphology. The interpretation as topic or (contrastive) focus is entirely contextual. It can be used more than once per clause and is commonly found with orienting information, whether locative or temporal:

<i>cəŋ</i>	<i>nə</i>	<i>ram-sə</i>	<i>nə</i>	<i>bənnu</i>	<i>chij-ji</i>
then	TOPIC	Ram-ERGATIVE	FOCUS	gun	hold-PERFECTIVE
‘Then, it was Ram who held the gun.’					

However, its use with episode initial temporal clauses (‘When he returned to the village . . .’) is virtually obligatory.

There are no concord classes (genders) in Chantyal. In rare instances, speakers may use the feminine form of an adjective borrowed from Nepali to modify a feminine noun borrowed from Nepali, but such usage, rare enough in the local dialect of Nepali, is used only when speakers are trying to sound sophisticated and never happens in spontaneous speech.

Chantyal does not have special sets of honorific nouns and verbs.

3.3 Locational expressions

In addition to the case clitics, Chantyal has an extensive vocabulary for describing location. Some of these are nouns and bear an obvious relation to the case clitics. Others are built off of independent elements, often in combination with case clitics. Still others

contain traces of an earlier set of demonstratives. Only some of these forms can be discussed here.

Direction on a vertical scale is expressed by means of the following set of forms:

	<i>Location</i>	<i>Direction</i>
down	<i>məŋ</i>	<i>mar</i>
up	<i>tuŋ</i>	<i>tor</i>

The forms indicating location contain locative *-əŋ*; those indicating direction contain locative *-ri*.

A number of locative expressions attest an earlier set of demonstratives; in these forms *ca-* denotes the proximal relation and *te-* the distal. Some examples follow:

<i>ca-ŋ</i>	‘this place, over here’	<i>te-ŋ</i>	‘that place, over there’
<i>ca-r</i>	‘in this way/direction’	<i>te-r</i>	‘in that way/direction’
<i>ca-jam</i>	‘this side (of river or gorge)’	<i>te-jam</i>	‘that side (of river or gorge)’

Note that, with these forms too, locative *-əŋ* indicates location and locative *-ri* direction.

A number of locative nouns have assumed the function of postpositions and take genitive complements. A few instances of this construction are illustrated below:

<i>Ram-ye lele</i>	‘after Ram’	<i>Ram-ye lesəŋ</i>	‘behind Ram’
<i>tin məyna-ye liga</i>	‘for 3 months’	<i>Dasə-ye ligəm</i>	‘after Dasain’
<i>Ram-ye wən</i>	‘before Ram’	<i>Ram-ye wənənən</i>	‘in front of Ram’

3.4 Classifiers

There is no native system of nominal classifiers. There is, however, a system of classifiers borrowed imperfectly from Nepali. This consists of a two-way classification into humans and non-humans. This system is realized as a pair of suffixes attached to the numerals ‘one’, ‘two’ and ‘three’: the suffixes are obligatory for ‘one’ and ‘two’, but not for ‘three’:

	<i>Non-human</i>	<i>Human</i>
One	<i>yew-ta</i>	<i>yeg-jana</i>
Two	<i>duy-ta, dwi-ta</i>	<i>dwi-jana</i>
Three	<i>tin-ta</i>	<i>tin-jana</i>

In practice, however, the human classifier is seldom used: it is found mostly in very self-conscious speech where the speaker is imitating the forms he or she would use in formal Nepali. In less guarded speech, the human classifier is rarely encountered. The non-human suffix *-ta*, however, is habitually used whenever the numerals ‘one’ and ‘two’ are used to directly quantify nouns.

3.5 Numerals

The entire system of numerals, both cardinal and ordinal, has been borrowed from Nepali.

3.6 Pronouns and demonstratives

The personal pronouns, with their major variants, are listed below:

1sg	<i>na</i>
1dl	<i>nagi</i> [< <i>na khi</i> ‘I you’; inclusive sense only]
1pl	<i>nhi, nhī-ma</i> [we-PLURAL]
2sg	<i>khi</i> [INFORMAL], <i>nūwə</i> [FORMAL]
2pl	<i>nūwə</i>
3sg	<i>khi</i> [humans only], <i>cə</i> [‘that’: humans and non-humans]
3pl	<i>thəwə, thəwə-ma</i> [they-PLURAL], <i>cə-ma</i> [that-PLURAL]

The second and third person plural forms derive from forms with overt plural marking, though as *thjwə-ma* ‘they’ shows, these forms have ceased being analysed as such.

The ‘informal’ second person singular pronoun is the one used in almost all speech situations; the use of the second person plural pronoun as a ‘formal’ second person singular pronoun is in imitation of Nepali usage and is not done consistently. As for the third person singular pronouns, reference to humans may be affected either by use of *khi* or *cə*, though *cə* seems to be restricted to cases where a true deictic reference is intended: otherwise *khi* is used. A few instances have been recorded where *khi* has a non-human referent.

The case clitics used with nouns are also found with the personal pronouns. A few irregularities occur: the first person singular has an irregular genitive *nə-ye*, and the second and third person singular have commonly encountered irregular datives, *khyā-ra* and *khya-ra*; the regular forms *khi-ra* and *khi-ra* are also possible.

There are two sorts of demonstratives in Chantyal: the ‘root’ demonstratives and the ‘prefixal’ demonstratives.

The root demonstratives are used both adnominally and pronominally. When they are used pronominally, they occur with the same case clitics as nouns and personal pronouns; they also occur with the plural suffix. In their adnominal use, they are not inflected; that is, they do not agree with the head noun in case or number.

The root demonstratives form a two-member deictic system consisting of proximal *cu* ‘this’ and distal *cə* ‘that’. The distal form is used as a third person singular personal pronoun and, with plural marking, a third person plural personal pronoun.

The prefixal demonstratives form a three-member deictic system: proximal *yi-* ‘this’, distal *ha-* ‘that’ and remote *wu-* ‘yonder’. These forms may be prefixed onto only a limited number of roots. These include the root demonstratives, measure terms, some locational nouns and the adjective *thyawa* ‘big’. Examples of the last two sorts follow:

<i>yi-sər</i>	‘this way’	<i>yi-jəgər</i>	‘this much’
<i>ha-sər</i>	‘that way’	<i>ha-jəgər</i>	‘that much’
<i>ha-tuj</i>	‘up there’	<i>yi-ddyo</i>	‘this big’
<i>wu-tuj</i>	‘up there yonder’	<i>ha-ddyo</i>	‘that big’

The prefixal and root demonstratives together form a system of deictic reference which can be used adnominally or pronominally. The following combinations are attested:

<i>yi-cu</i>	‘this (near speaker, closer of two alternatives)’
<i>ha-cu</i>	‘this (near speaker, further from speaker of two alternatives)’

<i>yi-cə</i>	‘that (distant from speaker, nearer of two alternatives)’
<i>hə-cə</i>	‘that (distant from speaker, more distant of two alternatives)’
<i>wu-cə</i>	‘yonder (distant from speaker and hearer)’

In these forms, the root establishes the primary deictic focus and the prefix a secondary focus.

Interrogative pronouns may also occur with case clitics and the plural morpheme. The main forms are listed below:

<i>su</i>	‘who’	<i>ta</i>	‘what’
<i>khənə</i>	‘which, whose’	<i>khəni</i>	‘where’
<i>khawa</i>	‘how, what kind’	<i>khəre</i>	‘how many, how much’
<i>tala</i>	‘why’		

The indefinite pronouns *su-i* ‘someone’, ‘anyone’ and *tə-y* ‘something’, ‘anything’ consist of the corresponding interrogative pronouns and a suffix *-i* which is likely the same as a suffix meaning ‘too’, ‘also’; ‘even’.

3.7 Verbs

Verbs are inflected for tense, aspect and mood. They may also be nominalized (becoming nominalizations, which may occur with case clitics and which may be used in adnominal functions) and adverbialized (becoming converbs, that is non-finite verbals having adverbial functions). Verbs are not inflected for agreement with arguments, for direction or for voice; they do not demonstrate a conjunct/disjunct distinction. In the native vocabulary, verbs are not marked for transitivity, but in the borrowed Nepali vocabulary, distinctions in valence may be marked morphologically.

The various verbal suffixes are listed below:

Tense-aspect:

Non-past	<i>-m, -mu</i>
Non-past interrogative	<i>-m-č̄, -č̄</i>
Perfective	<i>-ji, -i</i>
Perfective interrogative	<i>-la</i>
Imperfective	<i>-ma, -wā, -ð</i>
Negative anterior imperfective	<i>-s-are, -s-ere</i>
Past anterior	<i>-sč̄</i>
Emphatic, mirative	<i>-wa</i>

Mood:

Hypothetical	<i>-i</i>
Suppositional	<i>-ndə, -nnə, -n</i>
Suppositional	<i>-t</i>
Hypothetical	<i>-rð, -r</i>
Desiderative	<i>-to</i>
Imperative	<i>-o</i>
Polite imperative	<i>-ne</i>
Hortative	<i>-ye</i>
Optative	<i>-kay, -gøy, -ge</i>

Subordinate:

Anterior	-si-
Nominalizer	-wa
Infinitive	-nu (N?) ⁴
Cotemporal	-khiri, -khir, -khi (N)
Progressive converb	-kay, -gay, -ge
Sequential converb	-si-rə, -si-r, -si
Resultative	-nə
Conditional	-la
Negative conditional	-khare, -ghare
Remote conditional	-la-i

The anterior suffix *-si-* combines with the non-past, imperfective, emphatic (really a form of the nominalizer suffix), suppositional, hypothetical and nominalizer suffixes, adding an anterior (secondary past) sense.

Only one verb, *hya-* ‘go’, has an honorific counterpart: this honorific verb *ba-* is defective in having only imperative forms.

A large number of periphrastic verbal constructions exist and are used to express a wide variety of tense, aspect and mood senses. The principal auxiliary verbs are:

<i>hin</i>	‘be’	<i>mu</i>	‘be’
<i>ta-</i>	‘become’	<i>la-</i>	‘do’
<i>pin-</i>	‘give’	<i>yā-</i>	‘find’

Of these, the two *be*-copulas are the two most commonly used auxiliaries. In these periphrastic constructions, the semantic main verb can occur in a number of non-finite forms, the most common of which are illustrated below:

<i>-wa</i>	NOMINALIZER
<i>-si-wa</i>	ANTERIOR-NOMINALIZER
<i>-si</i>	ANTERIOR
<i>-gay</i>	PROGRESSIVE
<i>-la</i>	CONDITIONAL
<i>-nə</i>	RESULTATIVE

The basic tense-aspect system contrasts a perfective with an imperfective in the past, but does not oppose these aspects in the future. The perfective has past perfective and immediate future senses; the non-past is used to express present and future senses: the future sense is neutral with regard to aspect. The forms are given in Figure 25.2.

The progressive aspect can be expressed by means of various periphrastic constructions in both the past and non-past. Secondary tense distinctions, perfect and prospective, are also formed peripherastically.

There are no reflexive pronouns nor is there any special verbal form used to express reflexive senses. When ambiguity would otherwise arise, ordinary personal pronouns (marked as datives if direct objects) are used to translate English reflexive pronouns.

<i>na-sə</i>	<i>na-ra</i>	<i>jhi-i</i>
1sg-ERGATIVE	1sg-DATIVE	bite-PERFECTIVE
‘I bit myself.’		

	Past	Present	Immediate Future	Future
Perfective	- <i>ji</i>		- <i>ji</i>	
Imperfective	- <i>ma</i>		- <i>m</i>	

FIGURE 25.2 THE TENSE-ASPECT SYSTEM

There is, however, a special reciprocal form:

Ram	rə	Pirəm-sə	lhi-si	khum	la-i
Ram	and	Piram-ERGATIVE	hit-ANTERIOR	RECIPROCAL	do-PERFECTIVE
'Ram and Piram hit each other.'					

In the vocabulary borrowed from Nepali – but not in the native vocabulary – a pair of derivational affixes is used to indicate differences in transitivity *vis-à-vis* the basic, inherent valency of the verb. The lower degree of transitivity is indicated by the derivational affix *-i-*; the higher degree of transitivity is indicated by *-ə-*. The lower degree of transitivity may either be intransitive or transitive; the higher degree is transitive or causative (three implicit arguments). While this system is based on morphological distinctions made in Nepali, it is not entirely congruent with the Nepali system, which is richer, both morphologically and semantically.

As noted, the *-i-* suffix indicates relatively low transitivity, *-ə-* relatively high. A couple of examples are given below:

<i>Root</i>	<i>Intransitive</i>	<i>Transitive</i>	<i>Causative</i>
pətk-	pətk- <i>i-</i>	pətk- <i>ə-</i>	
'explode, burst'	'explode, burst (intr)'	'explode, burst (tr); fire a gun'	
dħəl-	dħəl- <i>i-</i>	dħəl- <i>ə-</i>	
'fall'	'fall over'	'knock over'	
phutk-	phutk- <i>i-</i>	phutk- <i>ə-</i>	
'loose'	'escape, untangle, untie (intr)'	'let loose, untie (tr)'	
tər-		tər- <i>i-</i>	tər- <i>ə-</i>
'cross'		'cross'	'take across'
buj-		buj- <i>i-</i>	buj- <i>ə-</i>
'understand'		'understand'	'explain'
səmjħ-		səmjħ- <i>i-</i>	səmjħ- <i>ə-</i>
'remember'		'remember'	'remind'

In a few cases, the *i*-form is associated with a root vowel /a/ and the *ə*-form with a root vowel /ə/; in almost all such cases, the alternation was borrowed directly from Nepali: *pak-i-* 'cook' (intr), *pək-ə-* 'cook' (tr); *nacc-i-* 'dance', *nəcc-ə-* 'make (someone) dance'. In a very few cases, the two forms may signal different kinds of objects rather than differences in transitivity: for example, *bħər-i-* 'fill' takes substances as objects, whereas *bħər-ə-* 'fill' takes containers as objects. (Note that *bħər-i-* also has an intransitive sense: 'be full'.)

The case marking associated with verbs having these derivational affixes is entirely predictable and expected: transitive subjects are marked with the ergative; intransitive subjects are absolute. There is also a periphrastic causative, illustrated below:

<i>ram-sə</i>	<i>pirəm-ra</i>	<i>nhaka</i>	<i>rhe-nə</i>	<i>la-i</i>
Ram-ERGATIVE	piram-DATIVE	chicken	steal-RESULTATIVE	do-PERFECTIVE
'Ram made Piram steal the chicken.'				

This is the only native valence-changing process, morphological or syntactic.

3.8 Copular verbs

There are two stative copular verbs in Chantyal, and one active. The active copula is *ta-* ‘become’; the stative copulas are *hin* and *mu*, both of which are morphologically irregular:

<i>Non-past</i>	<i>Interrogative</i>	<i>Negative</i>	<i>Past</i>	<i>Negative</i>
	<i>non-past</i>	<i>non-past</i>		<i>past</i>
<i>hin</i>	<i>nhe</i>	<i>hayn</i>		
<i>mu</i>	<i>mu-ē</i>	<i>hare</i>	<i>mu-ma</i>	<i>hare-ma</i>

mu is regular except for the suppletive root [*rhe*] found in the negative and in the non-past, which lacks the suffix *-m*. (The latter is understandable given that the copula *mu* is the origin of the non-past suffix.) *hin* is defective in not having past forms: the past forms of *mu* are used instead. *hayn* is analysable simply as the negative suffix *a-* plus *hin*; *nhe* is simply a contraction of *hin* and the non-past interrogative suffix *-ē*.

In general, *hin* is used to indicate identity; *mu* is used for location and attribution: it is also used in presentative constructions. *mu*, however, is clearly the unmarked form, and is occasionally found in contexts normally associated with *hin*. Both copular verbs are commonly used as auxiliaries.

The presence of a copula in clauses with a predicate adjective or a predicate nominal is virtually obligatory: in a corpus of over 5,000 analysed clauses, only two clear instances of such clauses without a copular verb have been recorded.

3.9 Adjectives

The category of adjective is defined syntactically, not morphologically, in Chantyal. Prior to the massive influx of vocabulary from Nepali, adjectives all contained the nominalizer suffix *-wa*. (As will be discussed below, the nominalizer has adnominal, attributive functions.) In modern Chantyal, all but one native adjective contains *-wa* (or a fossilized version of it); however, the great majority of adjectives in contemporary Chantyal are Nepali borrowings and such words follow Nepali rules in not being marked as nominalizations. The following NP contains a native and a borrowed adjective:

<i>thya-wa</i>	<i>kalce</i>	<i>naku</i>
big-NOMINALIZER	black	dog
'big, black dog'		

3.10 Expressive vocabulary

The Chantyal lexicon contains a large number of items which I refer to collectively as ‘expressive’ vocabulary. These words describe sounds (often onomatopoeically), the

appearance of things, modes of action, physical sensations or some combination of these factors. The words are often full reduplications and sometimes partial reduplications; those forms that are not reduplicated typically employ ‘emphatic’ gemination of the last consonant.

<i>rhay rhay</i>	sound of something dry and crunchy
<i>cwakta cwakta</i>	sound of cutting up meat
<i>jhəm jhəm</i>	burning, tingling sensation (as produced by alcohol, nettle, Szechuan pepper)
<i>khwak khwak</i>	sound of choking, gagging
<i>burruk burruk</i>	action of jumping
<i>phwak phwak</i>	coming off in chunks
<i>phwakkə</i>	coming off in chunks
<i>cakal cikil</i>	eating sloppily, noisily
<i>phappə phuppa</i>	scattered carelessly, easily picked up and/or stolen
<i>khəllə bəllə</i>	action of quarrelling
<i>lətə pətə</i>	appearance of scattered things
<i>khwassa</i>	inserting quickly and neatly
<i>lwarra</i>	pulling something, which offers some resistance, smoothly

The expressive vocabulary can form verbal expressions with the verb *la-* ‘do’:

<i>burruk burruk la-wa</i>	‘jump’
<i>ghadi gudi la-wa</i>	‘be silent’

These verbal expressions can then be used adverbially:

<i>cakre makre</i>	<i>la-i</i>	<i>ci-wa</i>
sit-cross-legged	do-ANT	sit-NOM
‘to sit cross legged’		

3.11 Verb particles

There are about a dozen commonly occurring verbal particles, so called because they attach as enclitics to the verb. These particles express a variety of senses, ranging from evidentiality to emotional reaction to the state of affairs described in the sentence. They are found only with finite clauses. The most common particles and their meanings are listed below:

<i>tə</i>	speaker asserts the statement is true
<i>ro</i>	speaker asserts that he or she cannot guarantee the truth of the statement (for example, because the event described was not directly witnessed)
<i>nu</i>	the sentence should serve as a reminder of information the hearer should already be in possession of
<i>sə</i>	expresses a sense of disappointment or frustration with regard to the state of affairs associated with the statement
<i>n̪i</i>	used to elicit a response indicating agreement with the truth of the statement or an indication that the statement was understood

The semantics and pragmatics of these forms are complex and so the senses given above should only be taken as approximations. The use of the evidential particles *tə* and *ro* is not obligatory; they are used primarily for emphasis.

4 WORD FORMATION

4.1 Derivation

Chantyal has relatively little in the way of native derivational morphology. With the exception of verbs, Nepali vocabulary is borrowed into Chantyal with its derivational morphology intact, but there is little evidence that the morphological patterns instantiated in this vocabulary are used productively. (As discussed earlier, verbs, too, are borrowed so as to reflect Nepali derivational categories, though the inventory of categories in Chantyal represents a subset of those found in Nepali.)

An important piece of derivational morphology in Chantyal is the nominalizer suffix *-wa*. This suffix is available for verbs, adverbs and other word classes. Its functions are described in some detail later. An important set of derivational morphemes are the con-verbal suffixes, which adverbialize verbs.

4.2 Compounding

Noun-noun compounding is an active process in Chantyal, but it does not have the prominence in Chantyal that it does in some other Tamangic languages, such as Nar-Phu. One likely reason for this is the frequency and ease with which Chantyals borrow vocabulary from Nepali and, more recently, English. Many Chantyal compounds, however, are composed of Nepali elements and have their origins in the local Nepali dialect.

Noun-noun compounds involve simple juxtaposition of the two nouns with the last noun constituting the semantic head. A few examples follow:

<i>Compound</i>	<i>Meaning of parts</i>	<i>Meaning of compound</i>
<i>buri aŋjula</i>	old woman digit	'thumb'
<i>dhiopini mimi</i>	washerwoman bird	'brown dipper'
<i>har khor</i>	bone shelter	'skeleton'
<i>bhiūy puləm</i>	ground berry	'mock strawberry'

In a few cases, however, the relation between the component nouns is coordinative, in which case the last noun is not the semantic head.

<i>Compound</i>	<i>Meaning of parts</i>	<i>Meaning of compound</i>
<i>nhe thara</i>	milk buttermilk	'dairy products'
<i>ama bəw</i>	mother father	'parents'

A few fossilized noun-noun compounds exist, forms which derive historically from noun-noun compounds but which are no longer analysed as such as their origins have ceased to be transparent. A few examples are given below:

<i>Compound</i>	<i>Origin</i>	<i>Meaning</i>
<i>pʰəkən</i>	<i>pʰara</i> 'flour' + <i>kan</i> 'food'	'porridge'
<i>khele</i>	<i>khə</i> 'mouth' + <i>le</i> 'tongue'	'tongue'
<i>nagi</i>	<i>na</i> '1sg' + <i>khi</i> '2sg'	'we two'

Noun-verb compounds, however, are also common. With noun-verb compounds, the noun-verb unit is independently lexicalized and, as a result, may take on senses not directly inferrable from the component parts: this is true also for non-noun compounds.

Further, the noun component in noun-verb compounds does not count as an object for purposes of case assignment since it is part of the predicate expression. These nouns are always in the absolute case – that is, they are unmarked. A few noun-verb compounds are given below:

<i>Compound</i>	<i>Meaning of parts</i>	<i>Meaning of compound</i>
<i>bənnu lhi-</i>	gun hit	‘shoot’
<i>chutti ta-</i>	release become	‘dismiss’
<i>səncə la-</i>	health do	‘treat, cure’
<i>sunne la-</i>	emptiness do	‘be silent, desert’

Many noun-verb compounds have *la-* ‘do’ as their verb.

5 SYNTAX

5.1 Structure of the noun phrase

Word order within the NP is fairly rigid. The canonical order is as follows:

DETERMINER/GENITIVE + RELATIVE CLAUSE + NUMERAL + ADJECTIVE + HEAD NOUN

Very few deviations from this arrangement have been recorded in a corpus of over 5,000 analysed clauses.

5.2 Structure of the clause

In the overwhelming majority of cases, the verb complex, by which I mean the verbal predicate together with any auxiliaries and verb particles, comes last in the clause. When it does not – afterthoughts aside – the effect is usually to focus attention on the verb, often signalling surprise or astonishment that the particular action took place:

<i>bhuluy-sə</i>	<i>ca-i</i>	<i>bura-ra</i>
leopard-ERGATIVE	eat-PERFECTIVE	old.man-DATIVE
‘The leopard actually ate the old man.’		

As for the other clausal constituents, the order usually follows the empathy hierarchy:

SPEECH ACT PRONOUNS [FIRST AND SECOND PERSON]
THIRD PERSON PRONOUNS
PERSONAL NAMES
OTHER HUMAN REFERENCES
ANIMATE NON-HUMANS
INANIMATES

In general, the higher a referent is on the hierarchy, the more likely it is to occur early in the clause.

<i>na-ra</i>	<i>bhuluy-sə</i>	<i>ca</i>
1sg-DATIVE	leopard-ERGATIVE	eat-HYPOTHETICAL
‘The leopard might eat me.’		

Where there are two participants of equal rank on the empathy hierarchy, the subject will precede the object, and the object will precede any obliques. Orienting information, temporal or locative, usually occurs first:

<i>cəŋ</i>	<i>gāw-ri</i>	<i>na-sə</i>	<i>bənnu</i>	<i>bhar-ri</i>
that-LOCATIVE	village-LOCATIVE	1sg-ERGATIVE	gun	rack- LOCATIVE
<i>kar-ji</i>				
put-PERFECTIVE				

‘Then, in the village, I put the gun on the rack.’

There is a special sentence topic slot which, when it is filled, comes first in the sentence, preceded only by orienting information. Such sentence topics, always marked with the topic/focus particle *nə*, are typically accompanied by a special intonation and may lack expected subject (ergative) or direct object (dative) case marking:

<i>cəŋ</i>	<i>nə,</i>	<i>pirəm</i>	<i>nə,</i>	<i>lata</i>	<i>mənchi,</i>	<i>dula</i>	<i>kho-i</i>
then	TOPIC	Piram	TOPIC	stupid	person	hole	dig-PERFECTIVE
‘Then Piram, that stupid person, dug a hole.’							

The notional subject, Piram, lacks the expected ergative case marking because it is occupying the sentence topic slot.

Chantyal is remarkably consistent in employing the ergative case with transitive subjects, regardless of tense, aspect or mood; regardless of whether the subject was acting volitionally; or, indeed, regardless of whether the subject was acting at all. As long as the situation is expressed transitively, the subject will be coded in the ergative case:

<i>khi-sə</i>	<i>na-ra</i>	<i>cini-m</i>
3sg-ERGATIVE	1sg-DATIVE	know-NON-PAST
‘She knows me.’		

Conversely, the ergative is never used with intransitive subjects, for example as an emphatic marker.

Direct objects may be expressed in the absolute case or in the dative. While there are a number of factors involved in the choice of case marking in such cases, the primary factor is, again, the empathy hierarchy: the higher a referent expressed as a direct object is on the empathy hierarchy, the more likely it is to be coded as a dative. Human referents are generally coded as datives; non-human referents generally are not. Other factors include the degree to which the entity is perceived as registering sensation as a result of the action expressed in the clause and the degree of empathy felt towards the entity. Compare:

<i>khi-sə</i>	<i>nhaka</i>	<i>tha-i</i>
3sg-ERGATIVE	chicken	cut-PERFECTIVE
‘He sacrificed/killed the chicken.’		

with:

<i>khi-sə</i>	<i>nhaka-ra</i>	<i>tha-i</i>
3sg-ERGATIVE	chicken-DATIVE	cut-PERFECTIVE
‘She cut the chicken [so that it bled].’		

The dative case is also used with indirect objects (recipients in transactions) and with experiencers in constructions where the verb is basically not transitive, e.g. a copula or an intransitive verb of motion:

na-ra joro kha-si-m
1sg-DATIVE fever come-ANTERIOR-NON.PAST
'I have a fever.'

As noted above, experiencers are coded as ergatives when the verb is transitive.

5.3 Copular clauses

As noted above, Chantyal has two stative copular verbs and one active copula. It was further noted that a copular verb is obligatory in clauses with a non-verbal predicate (i.e. a predicate nominal, predicate adjective or an oblique case-marked NP functioning as the predicate).

In clauses with a predicate nominal or predicate adjective, the verb complex containing the copular verb occurs in clause final position in the same manner as other verb complexes. The verb complex is preceded immediately by the non-verbal predicate:

PREDICATE NOMINAL
nhi bidyarthi hin
1pl student be+NON.PAST
'We're students.'

PREDICATE ADJECTIVE
na khusi mu
1sg happy be+NON.PAST
'I'm happy.'

OBLIQUE CASE-MARKED NP FUNCTIONING AS PREDICATE
cə kitab tebəl-phyaray mu
that book table-SUPERESSIVE be+NON.PAST
'That book is on the table.'

5.4 Negative clauses

The negative morpheme is the prefix *a-*, attached to the verb.⁵ There is a special imperative negative *tha-*. Apart from these forms, there are no morphemes that can be used to negate a clause and only one other specifically negative form, *khamməy* 'never', though this form requires the negative prefix on the verb:

khi-sə kan khamməy a-ca-m
3sg-ERGATIVE rice never NEGATIVE-eat-NON-PAST
'He never eats rice.'

Chantyal thus has no specifically negative indefinite pronouns, such as English 'no one' or 'nothing', which can independently negate a verb, and further does not require, like

Russian, negative agreement of indefinite pronouns with a negative verb, as shown in the following two examples:

nhi-sə *su-i-ra* *pəni* *a-pin-ji*
we-ERGATIVE who-even-DATIVE also NEGATIVE-give-PERFECTIVE
'We didn't give it to anyone.'

na-sə *tɔy* *a-yā-i*
1sg-ERGATIVE anything NEGATIVE-find-PERFECTIVE
'I didn't find anything/I found nothing.'

(If the negative morpheme were not present on the verb, the first sentence would mean 'we gave it to whomever' and the second would mean 'I found something'.)

5.5 Interrogative clauses

Yes-no questions can be formed: (1) with an appropriate interrogative suffix on the finite verb within the verb complex; (2) with an interrogative tag; or (3) by means of an appropriate interrogative intonation on a sentence whose grammatical form is the same as the corresponding statement.

Interrogative suffixes are available only for the perfective and non-past and for the periphrastic forms built off them. For other tense/aspect/mood forms, intonation or interrogative tags are used to form questions. The perfective suffix is *-la*, which replaces the declarative perfective suffix. The non-past suffix is *-č*, which is added to the non-past suffix *-m*.

khi-sə *bħalu* *sar-la*
you-ERGATIVE bear kill-PERFECTIVE.INTERROGATIVE
'Did you kill a bear?'

khi-sə *bħalu* *sar-m-č*
you-ERGATIVE bear kill-NON.PAST-INTERROGATIVE
'Will you kill a bear?'

Interrogative tags have a pragmatic sense similar to tag questions and tag particles in English: they are typically used in situations where the speaker is not sincerely trying to acquire information, but rather is trying to get the hearer to confirm an opinion the speaker already has.

khi-sə *bħalu* *sar-m* *nħi*
you-ERGATIVE bear kill-NON-PAST OK
'You'll kill a bear, OK?'

Information questions are formed with an interrogative pronoun which is placed in the usual position within the clause that a corresponding non-interrogative form would take, i.e. the pronoun is not obligatorily fronted. In the perfective and non-past, the interrogative forms of the finite verb are used.

khi-sə *su-ra* *mara-la*
you-ERGATIVE who-DATIVE see-PERFECTIVE.INTERROGATIVE
'Whom did you see?'

5.6 Complement (nominal) clauses

Complement clauses – clauses occupying nominal slots – are typically expressed as nominalizations in *-wa*.

<i>na-ra</i>	<i>syaw</i>	<i>ca-wa</i>	<i>mən</i>	<i>kha-m</i>
1sg-DATIVE	apple	eat-NOMINALIZER	desire	come-NON-PAST
'I want to eat an apple.'				

There are no finite subordinate clauses in Chantyal, except as direct quote complements of *bhi*- ‘say’. There are numerous instances of clauses which are presented as though they were direct quotes, but are in fact not literally quotes: these clauses function as complement clauses. The verb *bhi*- thus takes on many of the characteristics of a complementizer, though it is not yet fully grammaticalized in this role.

<i>khi</i>	<i>tisuy</i>	<i>kadmandu-ri</i>	<i>hyā-i</i>	<i>bhi-wa</i>
3sg	last.year	Kathmandu-LOCATIVE	go-PERFECTIVE	say-NOMINALIZATION
<i>na-sə</i>	<i>səmjhī-i</i>			
1sg-ERGATIVE	remember-PERFECTIVE	'I remembered that she went to Kathmandu last year.'		

5.7 Adjectival (relative) clauses

Adjectival clauses, clauses that modify nouns, are grammatically nominalizations in *-wa*. These clauses do not employ a relative pronoun, nor do they employ a resumptive pronoun within the clause: the role of the referent within the clause that is coreferential to the head must be inferred from context.

<i>kyo-ma-ru</i>	<i>duli-si-wa</i>	<i>kyata</i>		
friend-PLURAL-COMITATIVE	wander-ANTERIOR-NOMINALIZER	boy		
<i>them-əy</i>	<i>hyā-i</i>			
house-LOCATIVE	go-PERFECTIVE	'The boy who was wandering around with his friends went home.'		

5.8 Adverbial clauses

Adverbial clauses in Chantyal are grammatically of two types: they are either nominalizations with an appropriate case marker which provides their semantic interpretation, or they are converbs, specialized adverbial clause types.

Case marked nominalizations are used for a number of adverbial functions, including the expression of purpose and cotemporality.

<i>səŋlal-ma</i>	<i>məə</i>	<i>tara-wa-ri</i>	<i>hyā-i</i>
Sanglal-PLURAL	honey	gather-NOMINALIZER-LOCATIVE	go-PERFECTIVE
'Sanglal and some others went in order to gather honey.'			

There are a number of converbs in Chantyal and they are used to express a wide variety of senses, including condition, temporal sequence and so on.

<i>pirəm</i>	<i>kha-lanə</i>	<i>a-kha-m</i>
Piram	come-CONDITIONAL	NEGATIVE-come-NON-PAST
'If Piram comes, I won't come.'		

Overt conjunction of clauses is little used in Chantyal; instead, other devices, most particularly the sequential converb, are used:

<i>ram-sə</i>	<i>gāw-ri</i>	<i>hya-si-rə</i>	<i>thim-nhari</i>
Ram-ERGATIVE	village-LOCATIVE	go-ANTERIOR-SEQUENTIAL	house-INESSIVE
<i>wō-si-rə</i>	<i>nhaka-ye</i>	<i>sya</i>	<i>ca-i</i>
enter-ANTERIOR-SEQUENTIAL	chicken-GENITIVE	meat	eat-PERFECTIVE
'Ram went to the village, entered the house, and ate the chicken.'			

The anterior converb, which is found mostly in a set of grammaticalized constructions, is used to form an (anti-)benefactive construction:

<i>na-ra</i>	<i>pir-si</i>	<i>pin-o</i>
1sg-DATIVE	let.loose-ANTERIOR	give-IMPERATIVE
'Let me loose!'		

5.9 ‘Zero’ anaphora

Chantyal does not normally express referents overtly if their identity is inferrable from context. The result, relative to languages like English, is that discourses appear highly abbreviated, assuming an almost telegraphic style, and require for their interpretation a greater degree of familiarity with the physical setting and the social and historical circumstances of the participants than a similar discourse would in a language whose grammar and discourse conventions required a greater degree of ‘copiousness’ with regard to information. Consider, for example, the following mini-discourse:

<i>mənchi-sə</i>	<i>thar-ra</i>	<i>bənnu</i>	<i>lhi-si-rə</i>
person-ERGATIVE	mountain.goat-DATIVE	gun	hit-ANTERIOR-SEQUENTIAL
<i>sar-ji</i>	<i>chala</i>	<i>tar-si-rə,</i>	<i>ənnarabħūri</i>
kill-PERFECTIVE	skin	extract-ANTERIOR-SEQUENTIAL	viscera
<i>thuti-si-rə,</i>		<i>them-əŋ</i>	<i>kha-si-rə,</i>
pull.out-ANTERIOR-SEQUENTIAL	house-LOCATIVE		bring-ANTERIOR-SEQUENTIAL
<i>dewri-ri</i>	<i>pəkə-ysi-rə,</i>		<i>ca-i</i>
pot-LOCATIVE	cook-ANTERIOR-SEQUENTIAL		eat-PERFECTIVE

'The man shot a mountain goat and killed it. He skinned it, cleaned it, took it home, put it in a pot, cooked it and ate it.'

In this discourse, the referents ‘man’ and ‘mountain goat’ are introduced in the first clause and are not repeated in any of the clauses that follow since the referents in these clauses are clear; by contrast, the English translation contains one instance of ‘he’ and seven instances of ‘it’.

NOTES

- 1 Work on Chantyal was supported by the National Science Foundation, grant No. DBC-9121114. Editors' note: Because Professor Noonan unfortunately passed away in 2009, we asked Kristine Hildebrandt, an expert on the Tamangic languages (see her contributions on Manange and Nar-Phu in this volume), to update this chapter for us. We are grateful for her assistance.
- 2 An alternative view of this situation is that there are three sorts of oral stops – voiceless, voiceless aspirated and voiced – and two sorts of syllabic nuclei: murmured and non-murmured. The view presented in the main body of the text reflects the intuitions of Chantyals, though this 'intuition' has doubtless been influenced by Nepali orthographic tradition. The transcription system employed here, with {ħ} representing murmur, is compatible with either view.
- 3 Note from Kristine Hildebrandt: Some comments are in order. First, in the 2003 edition and in a 2005 publication, Noonan describes /c, ch, j, jħ, cħ/ as stops. These are very likely affricates. Many words with these consonants are Nepali loans, and in Nepali they are palatal affricates (e.g. Chantyal *cak* 'buttocks, lower back', Nepali *cāk* 'buttocks'; Chantyal *jal* 'net', Nepali *jāl* 'net'). Additionally, the perfective suffix *-ji* in Chantyal has equivalents in other Tamangic languages that are affricates, e.g. *-tsi* in Manange and *-ce* in Nar-Phu. Second, Noonan transcribes the murmured palatal nasal as *nyħ*, but we represent it here as *nyħ*. At any rate, it is rarely attested, with only two observed forms from Noonan (1999): *nyħakəm* 'firefly', *nyħar* 'spill'.
- 4 The status of this form as a Nepali borrowing would appear to be self-evident were it not for the fact that the use of this form does not correspond to the use of the Nepali infinitive in *-nu*, even in dialectal usage. The Chantyal suffix is used in a construction whose sense is 'about to V':

<i>thū-nu</i>	<i>thū-nu</i>	<i>la-gøy</i>	<i>a-thū</i>
drink-INFINITIVE	drink-INFINITIVE	do-PROGRESSIVE	NEGATIVE-drink
'She was about to drink, but didn't.'			

- 5 It may also be prefixed onto adjectives, though generally not when they are used predicatively.

BIBLIOGRAPHY

With the exception of works by Michael Noonan, very little has yet been published on Chantyal. Collections, articles and chapters that focus exclusively or primarily on Chantyal include the following, all of which were written by Noonan:

- Noonan, Michael (1996) 'The fall and rise and fall of the Chantyal language', *Southwest Journal of Linguistics* 15.1–2: 121–136, (Also: *Milwaukee Studies on Language* 9, 1995.) (Describes the social and historical setting for the decline of the Chantyal language.)
- (1997) 'Versatile nominalizations', in Joan Bybee, John Haiman and Sandra Thompson (eds) *Essays on Language Function and Language Type: In Honor of T. Givón*, Amsterdam/Philadelphia: John Benjamins, 373–394. (Describes the range of functions of nominalizations in Chantyal and other languages.)
- (1999) 'Converbal constructions in Chantyal', in Yogendra P. Yadava (ed.) *Topics in Nepalese Linguistics*, Kathmandu: Royal Nepal Academy, 401–420. (Describes the syntax and semantics of the sequential and progressive convers.)
- with Ram Prasad Bhulanja, Jag Man Chhantyal and William Pagliuca (1999) *Chantyal Dictionary and Texts*. Berlin: Mouton de Gruyter.

- (2001) ‘The “double demonstratives” of Chantyal’, *Linguistics of the Tibeto-Burman Area* 24.2: 173–188. (Describes the spatial and discourse uses of the combinations of prefixal and root demonstratives.)
- with Karen Grunow-Hårsa (2002) ‘Posture verbs in two languages of Nepal’, in John Newman (ed.) *The Linguistics of Sitting, Standing, and Lying*, Amsterdam/Philadelphia: John Benjamins, 79–101. (Discusses the syntax and semantics of posture verb expressions [sit, stand, lie] in Chantyal and Magar, the latter being another Tibeto-Burman language of Nepal.)
- (2003a) ‘Motion events in Chantyal’, in Erin Shay and Uwe Siebert (eds) *Motion, Direction, and Location in Languages*, Amsterdam: John Benjamins, 211–234.
- (2003b) ‘Recent language contact in the Nepal Himalaya’, in David Bradley, Randy LaPolla, Boyd Michailovsky and Graham Thurgood (eds) *Language Variation: Papers on Variation and Change in the Sinosphere and in the Indosphere in Honour of James A. Matisoff*, Canberra: Pacific Linguistics, 65–87.
- (2005) ‘Spatial reference in Chantyal’, in Yogendra P. Yadava (ed.) *Contemporary Issues in Nepalese Linguistics*, Kathmandu: Linguistic Society of Nepal, 155–188.
- with Ram Prasad Bhulanja (2005) ‘Chantyal discourses’, in *Himalayan Linguistics: Himalayan Linguistics Archive 2*. Available at: escholarship.org/uc/item/4221169b?query=chantyal (accessed 12 August 2016).
- (2006a) ‘The rise of ethnic consciousness and the politicization of language in West-Central Nepal’, in Anju Saxena and Lars Borin (eds) *Globalization, Technological Advances, and Minority Languages in South Asia*, Berlin: Mouton de Gruyter, 161–174.
- (2006b) ‘Direct speech as a rhetorical style in Chantyal’, *Himalayan Linguistics* 6: 1–32.
- (2008) ‘Contact-induced change in the Himalayas: the case of the Tamangic languages’, in Peter Siemunds and Noemi Kintana (eds) *Language Contact and Contact Languages*, Amsterdam: John Benjamins, 81–106.
- (2011) ‘Aspects of the historical development of nominalizers in the Tamangic languages’, in Foong Ha Yap, Karen Grunow-Hårsa, Karen and Janick Wrona (eds) *Nominalization in Asian Languages: Diachronic and Typological Perspectives [Typological Studies in Language 96]*, Amsterdam/Philadelphia: John Benjamins, 195–214. (Addresses development of nominalizers in Chantyal that demonstrate tense distinctions, and the development of periphrastic constructions involving nominalized verbs in the verbal complex.)

CHAPTER TWENTY-SIX

MANANGE¹

Kristine A. Hildebrandt and Oliver Bond

1 INTRODUCTION AND TYPOLOGICAL SUMMARY

Manange (endonyms: Nyeshang, Nyeshante, Nyangmi) is a Tamang-Gurung-Thakali-Manange sub-group (Tamangic) language spoken in eight villages of the upper Manang District in central-northern Nepal. Other Tamangic languages, to which Manange will occasionally be compared in this description, include Tamang, Gurung, Thakali, Nar-Phu and Seke/Tangbe. The name Manange is composed of *Manang* ‘the people below’ and *kje* ‘voice, language.’ Published reports on speaker populations are conflicting, with the Central Bureau of Statistics (2012) reporting under 400 speakers while speaker self-reporting indicates between 3,000 and 5,000. In other cases, Manange is lumped in with Gurung (Gurung 1998; Tumbahang 2012). More recent research estimates some 2,000–3,000 active speakers distributed across upper Manang, with roughly the same number of ethnic group members living in Kathmandu and abroad. Not all of these diaspora Mananges are active users of their language, and as such, Manange could be classified as between “threatened” and “shifting” (to Nepali and English) according to the *Ethnologue*, due to a dwindling number of active younger speakers.

The basic word order of Manange in elicited structures and in most discourse-embedded clauses is verb-final, with post-positions and with post-nominal modification. Relative clauses are pre-nominal and negation is signaled by either a prefix or by suppletive forms of the copula. Case marking is reliably ergative-absolutive in elicited sentences, but the frequency of overt realization of case marking in discourse is also tied to pragmatic factors. Like other Tamangic languages, Manange lacks agreement. With the exception of the negative prefix, noun and verb morphology is exclusively suffixing or enclitic. Verbal affixes code aspect and modality, and nominalization of main verbs is frequently encountered in discourse.

2 PHONOLOGY

The consonant phoneme inventory in word-initial onset position is in Table 26.1.

The labio-velar approximant /w/ is extremely rare, with only one attested example: *jawa*²² (*ja*²² ‘hand’ and *wa*²² ‘clap’) ‘clap hands.’ The glottal plosive /?/ is idiosyncratic in occurrence in all positions: *?u/iu*⁵² ‘apple,’ *pe?*⁴⁴ ‘really/very,’ *tsʰa?ray*⁴⁴ ‘all, every.’ Minimal sets for other consonant phonemes are provided in Table 26.2.

The syllable canon is: (C1)(C2)V(C3). Consonants in C2 onset position are restricted to the tap and the approximants and are not uniformly attested in combination with C1 consonants from all places of articulation. For example, there are no consonant clusters in Manange when C1 is a retroflex. In the coda, the distribution of the consonants is even more restricted. The most commonly encountered coda consonant is the velar nasal, but final liquids and plosives are found in some words, as in *tuk*⁵² ‘poison,’ *sup*⁵² ~ *sûp*⁵² ~ *su*⁵² ‘body,’ *tsʰoy*²² ‘business,’ *kʰjor*⁴² ‘copper,’ *kʰol*⁴⁴ ‘boil.’ There is not a great amount of

TABLE 26.1 THE MANANGE CONSONANTS

	Bilabial	Dental	Alveolar	Post-Alveolar	Retroflex	Palatal	Velar	Glottal
Plosive	p p ^h	t t ^h			t t ^h		k k ^h	(?)
Labialized Plosive	p ^w p ^{hw}						k ^w k ^{hw}	
Affricate		ts ts ^h				tʃ tʃ ^h		
Nasal	m	n				ɲ	ŋ	
Labialized Nasal	m ^w						ŋ ^w	
Fricative			s	f	ʃ			
Tap				r				h
Approximant	(w)						j	
Lateral			l					
Approximant								

TABLE 26.2 MINIMAL SETS

	Initial	Medial
p	p ⁱ⁵² ‘say’	k ^h japs ⁴² ‘king’
p ^h	p ^h ⁱ²² ‘wine’	
t	ti ⁴⁴ ‘pluck, pull’	koto ²² ‘walnut’
t ^h	thi ⁱ⁴² ‘teapot’	
t ^l	tu ⁱ⁵² ‘cereal grain’	tonja ²² ‘thousand’
t ^h	thu ^{u42} ‘six’	
k	ki ²² ‘feces’	p ^h joko ²² ‘tree bark’
k ^h	k ^h i ²² ‘s/he, it’	
p ^w	p ^w e ²² ‘corn husk’	
k ^w	k ^w e ⁴⁴ ‘lift’	
k ^{hw}	k ^{hw} e ⁴² ‘song’	
ts	tsaq ⁵² ‘bed’	p ^h tsi ⁴⁴ ‘knee’
ts ^h	ts ^h 3 ²² ‘shine/glare’	
tʃ	tʃe ⁵² ‘tea’	tʃhotʃe ²² ‘time’
tʃ ^h	tʃhe ⁴² ‘book’	
m	mi ⁵² ‘person’	ŋimʒŋ ⁵² ‘ear’
n	ne ²² ‘knead dough’	kini ⁵² ‘quickly’
ɲ	ne ²² ‘obey’	
ŋ	ŋi ⁵² ‘two’	tanji ⁴⁴ ‘flea’
m ^w	m ^w i ⁴⁴ ‘money’	
ŋ ^w	ŋ ^w o ⁴⁴ ‘fry’	
r	r ³ ²² ‘goat’	tore ⁵² ‘graveyard’
s	s ³ ²² ‘ground’	nese ⁵² ‘tomorrow’
ʂ	ʂo ⁵² ‘wheat’	
ʃ	ʃ ³ ²² ‘flesh’	ʃoʃo ⁴⁴ ‘paper’
h	hani ⁴⁴ ‘where’	
j	ja ²² ‘hand’	mʒji ⁴⁴ ‘buffalo’
l	la ⁴⁴ ‘flee’	p ^h oli ⁴² ‘spider’

morpho-phonological alternation in Manange, but one regular pattern is a floating nasal feature in verb stems, which is phonetically realized as a homorganic nasal consonant, depending on the place of articulation of the initial consonant of a following suffix. For example: *priⁱ²²* ‘hit,’ ~ *prin-tsi²²* ‘hit-PFV,’ *prim-p^h3²²* ‘hit-NMLZR,’ *prij-ko²²* ‘hit-IMPER.’ Additionally, obstruent consonants are regularly voiced in intervocalic position, both

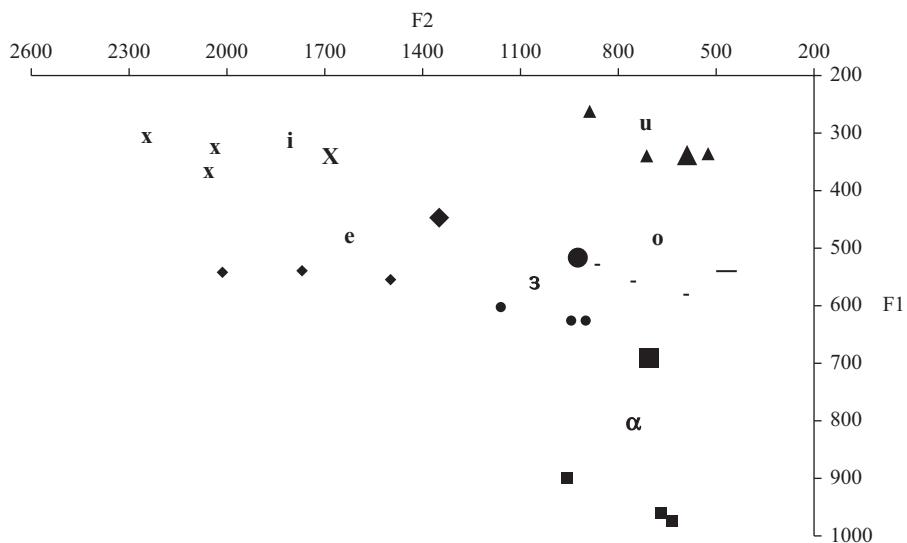


FIGURE 26.1 F1~F2 PLOTS, MANANGE VOWELS

²² word-medially and also between stems and suffixes/enclitics, e.g. *amtsay* ~ [ɬamdzaŋ] ‘maternal aunt.’ This is discussed in more detail in Hildebrandt (2004).

Manange has six oral and five nasal vowels. The acoustic distribution of the oral vowels in the vocal tract is illustrated in Figure 26.1 with data from 30 monosyllabic words elicited from three female speakers and one male speaker.² The values plotted in the chart represent average F1 (height) and F2 (backness) values for each speaker. In general, with the exception of the close-mid and close-back vowels, the male speakers' vowels are more centralized than the females.

A minimal set for oral and nasal vowels is provided in Table 26.3.

Nasalization in Manange is diachronically emergent. In this and in sister languages, older vowel-plus-nasal sequences are becoming reanalyzed as open syllables with a nasalized vowel. This is variable across speakers and even across pronunciations. This appears to be part of an overall historical trend towards coda consonant loss in final position in Tamangic.

¹ Manange has a system of four tones (for stopped and non-stopped syllables) that aligns mainly along the parameter of contrastive vowel fundamental frequency (F_0), realized as ‘pitch.’³ There are two level tones: a low-level and a high-level, as with the words *mje*²² ‘fire’ and *mi*⁴⁴ ‘eye,’ and there are two falling contour tones: a very high falling and a

TABLE 26.3 ORAL AND NASAL VOWELS

i	ti ⁴² ‘teapot’	ĩ	tĩ ⁴² ‘house’
u	ku ²² ‘chest’	ũ	ku ⁴⁴ ‘expensive’
e	le ²² ‘warm’	ẽ	lẽ ⁴⁴ ‘tongue’
o	ko ‘definite article’	õ	kõ ²² ‘get dressed’
a	tʃa ²² ‘search’	ã	tʃã ²² ‘small’
ɔ	lɔ ²² ‘do’		

TABLE 26.4 TONES

Tone number	Pitch contour	Example
/1/	22	<i>mje²²</i> ‘fire’
/2/	44	<i>mi⁴⁴</i> ‘eye’
/3/	52	<i>mi⁵²</i> ‘person’
/4/	42	<i>mje⁴²</i> ‘cow’

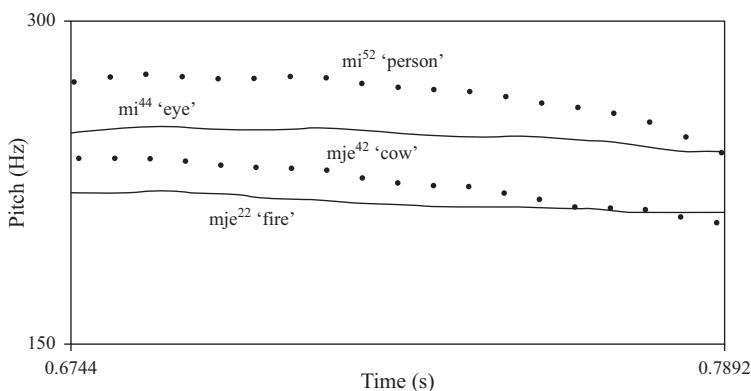


FIGURE 26.2 F_0 TRACES, SONORANT-INITIAL MONOSYLLABIC WORDS

mid-low falling, as with the words *mi⁵²* ‘person’ and *mje⁴²* ‘cow.’ This is represented in Table 26.4.

Figure 26.2 shows F_0 -plots for the starting, mid-point, and ending fundamental frequency values for the near-minimal set of nasal-initial words. The words belonging to the two contour (falling) tones are speckled. All words were uttered in a frame-medial sentence context of ‘I saw the X,’ where X is the target word illustrated in this figure, and the sentence is verb-final.

A secondary parameter of onset aspiration is revealed in words with initial plosives and affricates. In these cases, tones /3/ and /4/ show a split; tone /3/ words begin with only unaspirated consonants, and tone /4/ words begin with only aspirated consonants. This split is not present for words belonging to tones /1/ or /2/, where the initial consonant may be aspirated or unaspirated. This aspiration split is not relevant for words with initial nasal, fricative or approximant consonants. In other Tamangic languages, this split is realized phonetically as breathy phonation, but in Manange tone has phonologized further such that this phonation difference is found only as an aspiration feature on voiceless obstruents (Hildebrandt 2003, 2004; Mazaudon and Michaud 2008: 254). This is shown in Figure 26.3.

The tone-bearing unit in Manange is the phonological word. Those F_0 contrasts that are found on monosyllabic, monomorphemic words are also found, and expand across disyllabic words, whether they are mono- or polymorphemic (e.g. a stem plus an inflectional or derivational suffix). Figure 26.4 shows four disyllabic monomorphemic words (*poli²²* ‘shoe,’ *pʒtʃi⁴⁴* ‘knee,’ *pʒlɪ⁵²* ‘leg,’ *pʰɔli⁴²* ‘spider’) showing the same F_0 pattern and trajectory through time as is found with monosyllabic words.

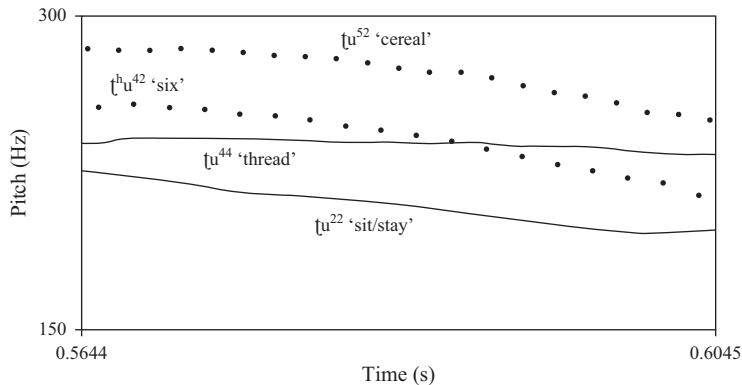


FIGURE 26.3 F_0 TRACES, OBSTRUENT-INITIAL MONOSYLLABIC WORDS

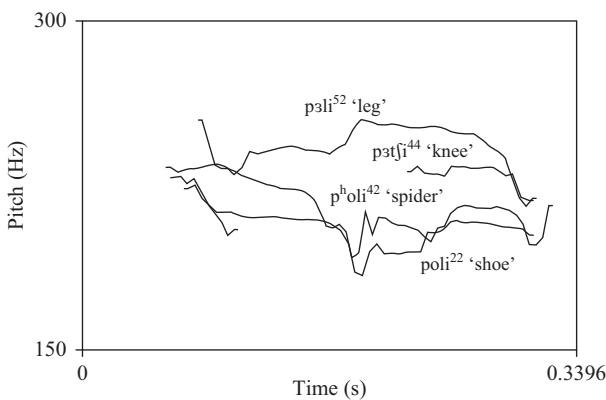


FIGURE 26.4 F_0 TRACES, DISYLLABIC WORDS

Figure 26.5 shows four bi-morphemic words (verb stems plus the a-tonal nominalizer/citation suffix $-p_3$) showing the same F_0 pattern and trajectory through time as is found with monosyllabic words.

3 WORD FORMATION, WORD CLASSES AND NOMINAL MORPHOLOGY

Nouns are the largest, most productive lexical class in Manange. Most nouns (and verbs) are monosyllabic and monomorphemic. New words are added via compounding or borrowing and more rarely via derivation. Some older loanwords that have been phonologized into the native system carry tone. This is shown in Table 26.5.

Nouns may be marked as plural by means of the clitic $=tse$. The plural enclitic regularly occurs with animate and inanimate nouns, and these forms have been found in elicited as well as connected speech: $njukju=tse^{22}$ ‘dogs,’ $kola=tse^{22}$ ‘children,’ $so=tse^{52}$ ‘friends,’ $juyjp_3=tse^{22}$ ‘rocks,’ $siyto=tse^{44}$ ‘fruit=PL,’ $k3η=tse^{44}$ ‘mountains,’ $phra=tse^{42}$ ‘hills,’ and even on loans, $up3h3r=tse$ ‘gifts.’ We observe that number marking by the plural enclitic

exhibits differential behavior and that its occurrence on semantically plural NPs is conditional based on pragmatic factors.

Nouns may also take the definite article =*ko* or the indefinite article =*ri* and they host case enclitics. Verbs are primarily suffixing for a range of aspects and modalities and there is a limited range of evidential particles. Some aspectual and modality distinctions are marked by periphrasis.

Like Nar-Phu (see Chapter 27), Manange has sets of honorific nouns and verbs that are used when speaking to or about those to whom special respect is due (primarily religious figures such as lamas). A partial listing of these words is provided in Table 26.6. These are

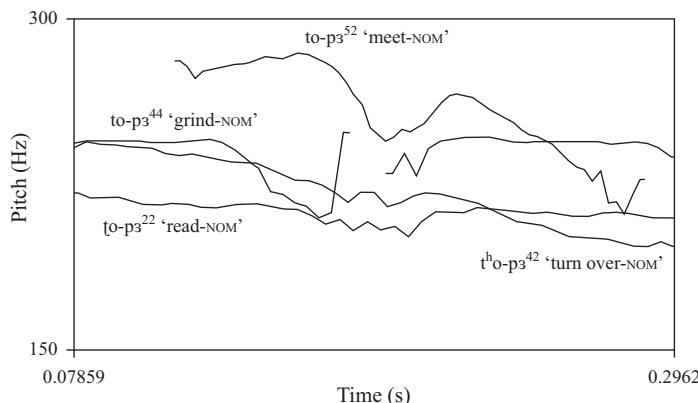


FIGURE 26.5 F_0 TRACES, SUFFIXED VERBS

TABLE 26.5 WORD FORMATION

Compounding	Loanwords	Derivation
<i>jafu</i> ²² ‘hand’ + ‘cover’ ⇒ ‘glove’	<i>kot̪a</i> ‘pasture’ (< Nepali <i>gotha</i>)	<i>n3-p3</i> ⁴⁴ be sick-NMLZR ‘sickness/disease’
<i>kjelu</i> ²² ‘voice’ + ‘teach’ ⇒ ‘recorder’	<i>tauli</i> ‘towel’ (< English <i>towel</i>)	<i>tay-p3</i> ⁵² be ancient-NMLZR ‘a long time ago’
<i>kjepʰra</i> ²² ‘barley’ + ‘flour’ ⇒ ‘tsampa flour’	<i>si</i> ²² ‘cotton’ (< Nepali <i>ril</i>)	

TABLE 26.6 HONORIFIC NOUNS AND VERBS

Nouns			Verbs		
Non-Honorific	Honorific	Gloss	Non-Honorific	Honorific	Gloss
<i>pakli</i> ²²	<i>?u</i> ²²	head	<i>f̪i</i> ²²	<i>p^be</i> ²² , <i>t^boŋ</i> ²²	die
<i>so</i> ⁵²	<i>ton</i> ²²	corpse	<i>t̪u</i> ²²	<i>fu</i> ²²	sit/stay
<i>li</i> ²²	<i>kuntu</i> ⁴⁴	face	<i>t̪jɔ</i> ²²	<i>fe</i> ²²	be big
<i>ti</i> ²²	<i>t̪u</i> ⁴⁴	heart	<i>pin</i> ²²	<i>nan</i> ²²	give
<i>siki</i> ⁴⁴	<i>sol</i> ²²	food	<i>jɔ</i> ²²	<i>p^be</i> ²²	go
<i>ara</i> ²²	<i>ʃet̪a</i> ⁴⁴	liquor	<i>njo</i> ²²	<i>si</i> ⁵²	look

infrequently encountered in discourse; speakers do not use them categorically when speaking about respect figures.

The personal pronouns are displayed in Table 26.7.

For some speakers, first person inclusive is expressed by a combination of the first person plural pronoun plus *tsʰaraj⁴⁴* ‘all/every.’ Other speakers use *ŋi²²* to indicate exclusivity. The proximal demonstrative pronoun is *tsu⁴⁴* and the distal is *u²²*.

Numerals in Manange operate on a decimal system, in contrast to Tamang’s vigesimal system (see Chapter 24). The cardinal numbers are given in (1):

(1)	1 <i>fɪ⁴²</i>	11 <i>tsukre²²</i>	30 <i>sumtʃu⁴²</i>
	2 <i>ŋɪ⁴²</i>	12 <i>tsunj²²</i>	40 <i>pʰlitʃu⁴²</i>
	3 <i>sen⁴⁴</i>	13 <i>tsupsen²²</i>	50 <i>ŋʒtʃu²²</i>
	4 <i>pʰli⁴²</i>	14 <i>tsupʰli²²</i>	60 <i>tʰukʒtʃu²²</i>
	5 <i>ŋʒ²²</i>	15 <i>tsunjʒ²²</i>	70 <i>ŋitʃu²²</i>
	6 <i>tʰu⁴²</i>	16 <i>tsutʰu²²</i>	80 <i>pʰretʃu⁴⁴</i>
	7 <i>ŋi⁵²</i>	17 <i>tsongi²²</i>	90 <i>kutʃu²²</i>
	8 <i>pʰre⁴²</i>	18 <i>tsapʰre²²</i>	100 <i>pʰrʒ²²</i>
	9 <i>ku²²</i>	19 <i>tsuku²²</i>	1000 <i>tonra²²</i>
	10 <i>tʃu²²</i>	20 <i>ŋitʃu⁴²</i>	

In compounded numerals, intrusive nasals and plosives are observed in some cases, along with limited suppletion. In elicitation, speakers use one classifier with inanimate nouns that is likely borrowed from Nepali: *-tʰa*. In discourse, another classifier is sometimes encountered: *-ti ~ -t̪i*, as in (2).

(2)	<i>ŋʒ²²</i>	<i>lo⁴²</i>	<i>ŋʒtʃu²²</i>	<i>epʒ=ko=rɪ²²</i>	<i>lo⁴²</i>	<i>ŋʒ-ti²²</i>	<i>bahira</i>	<i>tu-tst²²</i>
	1.SG	year	fifty	age=DEF=LOC	year	five=CLASS	outside	stay-PFV

‘As I was approaching 50 years of age, I lived outside (Manang) for five years.’
[Pisang_M20132007]

Ordinal numerals are borrowings from Tibetan: *tanpʒ⁵²* ‘first,’ *njipʒ²²* ‘second,’ *sumpʒ²²* ‘third,’ *ʃipʒ⁴²* ‘fourth,’ *ŋʒpʒ⁴²* ‘fifth,’ *tʰupʒ⁴²* ‘sixth,’ *tinpʒ⁴²* ‘seventh,’ *kjepʒ⁴²* ‘eighth,’ *kupʒ⁴⁴* ‘ninth,’ *tsupʒ⁴⁴* ‘tenth’ (Hoshi 1986a: 202). Quantifiers like ‘all,’ ‘many,’ and ‘some’ do not occur as the sole head of a noun phrase. Rather, they are post-nominal, both within the NP and post-NP, as in (3).

(3)	<i>ʒntse</i>	<i>tʃu=ko=rɪ⁴⁴</i>	<i>kju⁴⁴</i>	<i>tʃimɪ=ko</i>	<i>le⁴²</i>	<i>are²²</i>	<i>mo²²</i>
	and.then	PROX=DEF=LOC	water	spring=DEF	many	NEG.COP	COP

‘There are not many spring water sources here.’ [Braga_M22013006]

Manange has two enclitics to mark definiteness and indefiniteness. The enclitic *=ko* marks definiteness of previously introduced referents or referents that are assumed as known between interlocutors, while *=ri* marks indefinite or newly introduced referents. In (4) the

TABLE 26.7 PERSONAL PRONOUNS

	1SG	1PL	2SG	2PL	3SG	3PL
ABS	<i>ŋʒ²²</i>	<i>ŋjan²²</i>	<i>kjʒ⁵²</i>	<i>kimi²²</i>	<i>kʰi²²</i>	<i>kʰim²²</i>
GEN	<i>ŋʒlʒ²²</i>	<i>ŋjanljʒ²²</i>	<i>kjʒlʒ⁵²</i>	<i>kimilʒ²²</i>	<i>kʰilʒ²²</i>	<i>kʰimilʒ²²</i>
ERG	<i>ŋʒtse²²</i>	<i>ŋjanjtse²²</i>	<i>kjʒtse⁵²</i>	<i>kimitse²²</i>	<i>kʰitse²²</i>	<i>kʰimitse²²</i>
DAT/LOC	<i>ŋʒri²²</i>	<i>ŋjanjr²²</i>	<i>kjʒri⁵²</i>	<i>kimir²²</i>	<i>kʰir²²</i>	<i>kʰimir²²</i>

narrator introduces a boy at the start of a story. In (5) she then references the boy again (as ‘given,’ with *=ko*) and immediately introduces two new referents: a dog and a frog.

- (4) *tayp³*⁵² *tayp³*⁵² *fi*⁴² *ale=r²²* *mo*²² *mu*²² *ro*²²
 ancient ancient one boy=INDEF COP EVID REP
 ‘Once upon a time, there was one/a boy.’ [Eden_BDF001]⁴
- (5) *u*²² *ale=jun=ko*²² *njukju*²² *r₃* *fi*⁴² *p^holp³=r⁴²* *mo*²² *mu*²² *ro*²²
 dist boy=COMIT=DEF dog CONJ one frog=INDEF COP EVID REP
 ‘With that boy, there was a dog and one frog.’ [Eden_BDF002]

Yes-no questions in Manange are formed syntactically identically to declaratives, with the verb in-situ in final position, as in (6).

- (6) *naray*²² *kj³*⁵² *u*²² *t^hl=r⁴²* *j₃-tsi*²² ?
 before 2.SG DIST house=LOC go-PFV
 ‘Did you go to the house earlier/before?’

A response may consist of a repetition of the verb (*j₃-tsi*²²) for the affirmative or a negated version (*a-j₃*²²) for a refutation. Another way of expressing ‘no’ is via a suppletive negated copula: *a-jin*²².

The interrogative pronouns are listed in (7).

- (7) *ta*²² ‘what’
*halo*⁴⁴ ‘which’
*hani*⁴⁴ ‘where’
*su*⁴⁴ ‘who’
*hatsu*⁴⁴ ‘how manner’
*hatfun*⁴⁴ ‘how state’
*hajunj*⁴⁴ ‘when, never’
 (*phunj*⁴⁴) *k3tti*⁴⁴ ‘how many’ (lit. ‘(egg) how.many’)
*ta*²² *pi-tse*⁵² ‘why’ (lit. ‘what say-cc’)

In interrogative clauses, *hajunj*⁴⁴ translates as ‘when,’ but it can also be used as a response meaning ‘never.’ In content questions, the interrogative is regularly clause-final in pre-verbal position. Some examples are in (8)–(11).

- (8) *kj₃=l³*⁵² *min*²² *ta*²² ?
 2.SG=GEN name what
 ‘What is your name?’
- (9) *u*²² *mi=ko*⁵² *su*⁴⁴ *mo*²² *mu*²² ?
 DIST person=DEF who COP EVID
 ‘Who is that person?’
- (10) *muntse*⁴² *siki*⁵² *hajunj*²² *ts₃*²² ?
 night food when eat
 ‘When is dinner?’
- (11) *kj₃*⁵² *pwal=r⁵²* *ta*²² *pi-tse*⁵² *j₃-tsi*²² ?
 2.SG Kathmandu=LOC what say-CC go-PFV
 ‘Why did you go to Kathmandu?’

Manange has two classes of adjective: simple adjectives and verb-like adjectives. Simple adjectives constitute a small and closed class and include most color terms and some

words expressing semantic dimensions like human states, speed, and value. In attributive function in the NP, simple adjectives are post-nominal and they host inflectional enclitics such as plural, case, and definiteness clitics, as in (12).

- (12) *njukju²²* *k^hjokro=ko=tse⁴²* *ale=r²²* *pju-tsi⁵²*
 dog old=DEF =ERG boy=LOC chase-PFV
 ‘The old dog chased the boy.’

Simple adjectives cannot be the single (head) element of an NP, so a structure like *k^hjokro⁴²* ‘the old one’ is not acceptable to speakers. Additionally, simple adjectives in predicate position do not directly inflect for aspect or mood distinctions, nor do they occur before an evidential marker, nor are they prefixed with negative *a-*.

In contrast, ‘verb-like’ adjectives constitute a larger and open class. They share some, but not all, formal properties with verbs. In attributive function, verb-like adjectives follow the noun (like simple adjectives), as in (13). Also like verbs, they are suffixed with the nominalizer *-p3*. In the predicate, verb-like adjectives show some, but not all, of the morphological properties of verbs, as in (14). Both are marked for aspect and modality but while verbs can show negative prefixing, the vast majority of verb-like adjectives negate via a suppletive negative copula *a-re*²², but a small set of verb-like adjectives do show direct negation, as in (15).

- (13) *kju⁴⁴ t^hj₃-p₃=ri²²* *t^hɛ²² t₃-tsi²²*
 water be.big-NMLZR=LOC throw become-PFV
 '(The ashes) were thrown in big water (like a river).'

(14) *khi²² t^hj₃-tsi²²*
 3.SG be.big-PFV
 'He was big.'

(15) *zts_e tsu⁴⁴ mi⁵² a-s₃-p₃⁵²,* *kjaj-p₃⁵² mi-tse³²*
 and.then PROX person NEG-be.good-NMLZR be.far-NMLZR person-PL
 'And then, these bad people, the ones who remain far away...' [Pisang 2013M2022]

Also unlike verbs, verb-like adjectives in attributive function are post-nominal, while relativized verbs are pre-nominal. Manange also has some phrasal adjectives, comprised of compounds and other collocations, e.g. *t²² t²²j²²* [heart + big] ‘brave,’ *ara²² mre²²* [alcohol + forget] ‘drunk.’ Genetti and Hildebrandt (2004) provide a fuller treatment of adjectives in Manange.

Case in Manange is marked by a relatively small set of enclitics, listed below. Examples of ergative, genitive, and general locative marking may be found in (12), (8) and (13), respectively. Examples of instrumental, ablative, and comitative may be found in (16) through (18).

Absolutive	unmarked
Ergative, instrumental	= <i>tse</i>
General locative, dative	= <i>ri</i>
Genitive	= <i>l₃</i>
Ablative	= <i>t₃r</i> , = <i>ro</i>
Comitative	= <i>juj</i>

- (16) *k3r=tse⁵²* *ŋja-p3-ni⁴²* *zle²²* *k'hol-p3⁴⁴*
fabric=INSTR **wrap-NMLZR-ADV** like.this boil-NMLZR
 ‘After wrapping (the pot) with fabric, (the liquor) boils like this.’ [Ngawal_M1999F2030]

-
- (17) *s3-le⁴⁴* *t3n=ro⁴⁴* *nay-p3⁴⁴*
 be.nice-ADV god-ABL give.HON-NMLZR
 '(Crops) are given as a blessing from the gods.' [Pisang_2013M2043]
- (18) *tiŋi⁴⁴* *ŋ3²²* *so=jun⁵²* *mo²²*
 today 1.SG friend=COMIT COP
 'Today I am with (my) friend.'

There are two ablative markers, but the functional distinction between these is not clear. Similarly to what is observed in Nar-Phu, there is yet another option: =*tse* also occasionally occurs in texts as a marker of ablative, but it is not frequent (Noonan 2003b; Noonan and Hildebrandt, this volume). This may be an older form, which has been extended to instrumental and agentative uses (see, for example, LaPolla 1995).

Inanimate patient arguments are unmarked for case, but animate patients are marked with =*ri*, which functions not only as a general locative as in (19), but also to mark both animate objects and recipients of ditransitive verbs, as in (20).

- (19) *kʰwe⁴²* *napraŋ²²* *pʰute=ko=tse²²* *njukju=ko=ri²²* *pju-p3⁵²* *ro²²*
 honey fly swarm=DEF=ERG DOG=DEF=PDAT chase-NMLZR REP
 'The swarm of honey bees chased/was chasing the dog.' [BDF_Gyalpo]
- (20) *mriŋ=ko=tse²²* *ufu=ko²²* *kola=ko=ri⁵²* *pin-tsi²²*
 woman=DEF=ERG apple=DEF child=DEF=LOC give-PFV
 'The woman gave the apple to the boy.'

In elicited sentences, the ergative marker is strongly preferred on the A arguments of transitive verbs. In connected speech, however, a different pattern emerges. In a corpus of 129 clauses taken from connected speech in Manange, a total of three ergative-marked A arguments of transitive verbs have been observed (Bond *et al.* 2015). In other words, most of the time in connected speech, A arguments are unmarked for ergativity. An example of this can be found in (21) and (22). In (21) the first person pronoun A argument *ŋ3²²* is unmarked in a perfective clause. In (22), the third person lexical noun *tsay²²* is unmarked for ergativity (but does show the definite enclitic) with a non-perfective, nominalized verb.

- (21) *ŋ3²²* *nofu⁵²* *k3ru=ko⁵²* *g3u-tsi⁴⁴*
 1.SG first pot=DEF wash-PFV
 'First, I washed the pot.' [Ngawal_1999F2006]
- (22) *tsay=ko²²* *p3le⁵²* *ke-p3⁵²*
 bride=DEF leg touch-NMLZR
 'The bride (differentially) touches (the) leg (of the groom's relative).' [Tengki_2013M1021]

This extreme paucity of ergative marking is compounded by an additional phenomenon of low referential density: Core arguments themselves are also vanishingly rare in connected speech. In this same corpus only 20 percent of intransitive verbs show any overt S argument (pronoun or lexical noun) and only 9 percent of transitive verbs show any overt A argument. However, the presence (or lack thereof) of arguments and ergative marking is not entirely random in Manange. Bond *et al.* (2013) observe that ergative marking in Manange discourse correlates with information structure: Ergative-marked transitive subjects indicate a switch in reference in clauses where objects (or other verb complements) are also overtly realized. Unmarked transitive subjects are used for maintenance

of reference, or else they are the subject of a chained clause where the A argument of that clause is the same as the referent of the main clause.

Manange has several word-like morphemes that occur in post-nominal position in an NP. They encode specific spatial information, and these host the locative enclitic =ri. They are word-like in that they carry referential meaning and independent tone, but most of these obligatorily follow another noun in the NP. Some examples are given in (23). The proximal and distal demonstratives participate in this spatial encoding.

- | | | | |
|---|------------------------------|--|--------------------------------|
| (23) p ^{b1} i ⁴⁴ | ‘up, on top of objects’ | tso ⁴⁴ | ‘on top of/summit’ |
| k ^o ⁴⁴ | ‘high up from’ | naj ⁵² | ‘inside, downward, below’ |
| ti ²² | ‘nearby’ | t ^b acun ²² | ‘far from’ |
| p ^{sr} ²² | ‘in between’ | kun ²² | ‘in the middle’ |
| ko ⁴⁴ | ‘around’ | t ^b aj ²² | ‘out, flat on the ground’ |
| tsi ⁴⁴ | ‘this side’ | u ²² | ‘that side’ |
| j ^a ²² torts ³⁴⁴ | ‘left side’ lit. ‘hand left’ | ja ²² kets ³⁴² | ‘right side’ lit. ‘hand right’ |
| ku ²² | ‘around the corner’ | ts ^b aran ⁴⁴ k ^b ja ⁴² | ‘everywhere’ lit. ‘all place’ |
| ŋonts ³²² | ‘in front of’ | lits ³²² | ‘behind’ |

Manange does not have robust evidence for a distinct lexical class of “adverbs,” but a number of morphologically simple and phrasal elements that cross-cut lexical categories express time information and speaker attitude, position or evidence, shown in (24).

- | | | | |
|---|------------------------|--|----------------------|
| (24) tʃu ⁴⁴ ~ tʃutʃu ⁴⁴ | ‘after, later’ | lake ²² | ‘again’ |
| ts ^b e ⁴⁴ | ‘always’ | nosu ⁵² | ‘(at) first’ |
| lakelake ²² | ‘sometimes’ | toso ⁵² | ‘still, yet’ |
| tan-pz ²² [old-NMLZR] | ‘a long time ago’ | toso ⁵² ~ tosoŋ ⁵² | ‘now’ |
| hi l3-le ²² [when do-ADV] | ‘(at) last’ | hajun ⁴⁴ | ‘never’ |
| je ²² | ‘also’ | tosaran ⁵² | ‘already’ |
| tini ²² | ‘today’ | munse ⁴⁴ | ‘tonight’ |
| nese ⁴² | ‘tomorrow’ | tele ²² | ‘yesterday’ |
| hunji ⁴⁴ | ‘day before yesterday’ | raji ²² | ‘day after tomorrow’ |
| tsoko ti ²² | ‘only’ | ts ^b aran ⁴⁴ | ‘completely, all’ |
| ŋoto ⁵² | ‘truly, certainly’ | ta ⁴⁴ | ‘maybe’ |

Adverbial modification (manner) is expressed with a general clause-chaining structure on a verb, as in (25).

- | | | |
|--|----------------------------|------------------------------------|
| (25) ale ²² | tu-tse²² | ts ³ -tsi ²² |
| boy | sit-CC | eat-PFV |
| ‘The boy ate while sitting/in a sitting manner.’ | | |

4 VERB COMPLEX AND SIMPLE CLAUSE STRUCTURE

Verbs are morpho-syntactically distinct from nouns and other lexical categories in Manange. They host aspect, modality, and evidential suffixes and particles, they are negated via the prefix *a-*, and they (along with inflectional markers) are typically in sentence-final position. Like other Tamangic languages, there is no agreement or person-marking system in Manange.

The copula, *mo-pz²²* serves all existential, equative, attributive, possessive, and locational functions. Some examples follow in (26) through (29).

-
- (26) *ja⁵² k₃t_i⁴⁴ mo²² ro²²*
 yak many COP REP
 ‘There were many yaks.’ [YakBuff_002]
- (27) *jul=ri⁵² ke²² mo²² t^hti⁴² mo²²*
 village=LOC field COP house COP
 ‘(Yes), in (my) village, there are fields, there are houses.’ [Pisang_2013M1055]
- (28) *kola²² p^hli⁴² mo²², nani ij⁴² mo²²*
 child four COP daughter two COP
 ‘(I have) four children, two daughters.’ [Ngawal_2013053]
- (29) *ηjaj²² jul⁴² kompa ts^he⁴⁴ mo²²*
 1.PL village gompa be.good.quality COP
 ‘Our village gompa is good.’ [Ngawal_M1060]

This copula is multifunctional, occurring with a lexical verb in imperfective and in irrealis mood clauses. Despite its prominence in Manange grammar, it is frequently absent or else optional in copula clauses in discourse. The copula does not host the full range of aspect and modality morphology that lexical verbs do.

Verbs are negated directly, while copula verbs, similarly to most verb-like adjectives, show a negated suppletive form, as in (30) and (31).

- (30) *k^hi²² amtsi²² a-t₃-tsi²²*
 3.SG doctor NEG-become-PFV
 ‘S/he did not become a doctor.’
- (31) *k^hi²² amtsi²² a-jin²²*
 3.SG doctor NEG-COP
 ‘S/he is not a doctor.’

Manange has a small set of evidential particles aligning very loosely with verbal aspect and person differences. The non-first person, perfective evidential is *mi²²*. Another evidential is *mu²²*, a general non-past, imperfective evidential with a wider person alignment distribution. Both of these evidentials follow lexical verbs, but only *mu²²* can follow the copula, and even in these cases its appearance is optional. Examples are found in (32) and (33).

- (32) *t^hu₁ p₃r-ti mi²² t₃r⁵² ηe⁴⁴ mi⁴⁴ pi-tse⁵²*
 impure affect-LOAN EVID religious.object offer EVID say-CC
 ‘It is said that because something impure was done, religious objects were offered.’
 [Pisang_2013M2053]
- (33) *bides mi=tse⁵² pisanj pik=r_i k_{re}-p₃-ri²² k₃t_i⁴⁴*
 foreigner person=PL Pisang peak=LOC climb-NMLZR-PURP many
pi-le⁵² k^h3²² mu²²
 say-ADV come EVID
 ‘Foreigners, whatever number (of them), come here to climb Pisang Peak.’
 [Pisang_2013M2020]

There is an irrealis/future evidential *z²²* that may be a reduced form of *ta⁴⁴* ‘maybe’ (34). Another similar post-verbal particle, *ns²²*, expresses uncertainty and can be used with both first and non-first person actors (35). Another irrealis/future evidential is *ko²²*, which has a ‘checked and confirmed’ sense (36), and there is a reported speech particle, *ro²²*, seen in (4), (5), and (19) earlier.

- (34) *tʃutʃu⁴⁴ lo⁵² k^himi=juy²² ñt⁴⁴ kola=tse²² t³²² 3²²*
after year 3.PL=COMIT seven child=PL become EVID
'Maybe next year they will have seven children.'
- (35) *ñz²² j³-p³²² n³²²*
1.SG go-NMLZR EVID
'I'll go (now), I think.'
- (36) *ztse²² ja⁵² k³tti⁴⁴ mo-p³²² ko²²*
like.this yak many COP-NMLZR EVID
'Like this, there were many yaks.' (follows from clause in example (26)) [YakBuff_003]

Lexical verbs host aspect morphology, which is expressed either concatenatively through suffixation (perfective), or via a periphrastic structure (progressive and non-progressive imperfectives). The copula does not host aspect morphology; rather, the lexical verb *t³²²* 'become' is used for this. Within aspect distinctions, time adverbials provide necessary temporal information.

Some temporal distinctions in Manange are expressed via nominalization, which encodes a future temporal meaning. For many (but not all) speakers, what would be called "future tense" (as well as other modal voices like desiderative, deontics, and immediates) is in fact part of a larger category of irrealis mode evidenced via an ergativity split. Transitive A arguments of irrealis verbs do not host the ergative enclitic, which suggests a split in how modality is viewed with regard to unrealized vs realized actions and events. Examples for future-time, a deontic, and an immediate are shown with a single transitive verb *gu⁴⁴* 'wash' in (37) through (39).

- (37) *ñz²² kola²² gu-p³⁴⁴*
1.SG clothing wash-NMLZR
'I will wash the clothing.'
- (38) *ñz²² kola²² gu⁴⁴ l³²² jan⁵²*
1.SG clothing wash do DEONTIC
'I should wash the clothing.'
- (39) *ñz²² kola²² gu-pi⁴⁴ l³-ts²²*
1.SG clothing wash-IMM do-PFV
'I was about to/prepared to wash the clothing.'

Durative or continuously repetitive actions are encoded with two different strategies in Manange: with the continuous suffix *-tse* plus the copula, or else with a serial verb construction, shown in (40) and (41).

- (40) *k^himi²² juŋp³=tse²² k^hjē=ri⁴² k^hja-tse²² mo²²*
3.PL stone=PL road=LOC throw=CONT COP
'They continue to throw stones in the road.'
- (41) *k^himi²² juŋp³=tse²² k^hjē=ri⁴² k^hja²² tu-p³²²*
3.PL stone=PL road=LOC throw stay-NMLZR
'They continue to throw stones in the road.'

The nominalizer is in fact multifunctional in Manange. In addition to encoding a general future/irrealis mode, it suffices to verbs in citation/elicitation form, it combines with the evidential *ko* to provide a 'checked and confirmed' meaning of an action or event, regardless of aspect (example (36) earlier), it serves a de-verbalizing/nominalizing

function ((33) earlier), it combines with the suffix *-ri* to form purposive clauses and with the suffix *-ni* to form adverbial clauses (section 5), it is suffixed to the main verb of relative clauses (section 5), and in discourse it is often observed on the verb of a main clause, indicating that the event/action is generally true, for example when speakers state their names, ages, or locations of residence in monologues, or when they describe a general fact ((42) and (43)). This multifunctionality is a common feature of Tamangic languages (see, for example, Noonan 2003a, for a discussion of nominalization in Chantyal).

- (42) *lo⁴⁴ tuktū⁴² t^hu⁴² je-pz⁴⁴*
 year sixty six return-NMLZR
 '(I) am sixty-six years old.' [Braga_2013M1038]
- (43) *himal k^hre²² l^hz²² a-pim-pz²²*
 Himalaya climb do NEG-give-NMLZR
 '(The gods) do not allow climbing (of) the Himalaya.' [Pisang_2013M2025]

5 CLAUSE COMBINING

Manange makes use of a wide range of strategies for clause combining. They include verb nominalization for relativization, purposives, and certain modal expressions, and also clause chaining for adverbial subordination of various kinds, and causation.

Verbs are nominalized for relativization, and relative clauses are pre-nominal. The head noun of the relative clause is always marked for its semantic role in the larger sentence. This is shown in (44), where the A argument in the larger sentence *ja⁵²* 'yak' is ergative marked, despite being the S argument of an intransitive verb in the relative clause. In most cases, relativized verbs do not inflect for any aspect or modality, but a limited degree of aspectual information may be found in the relativized clause, expressed through verb serialization, as in (45) and (46).

- (44) *t^hutu⁴⁴ k^hη=ri⁴⁴ tu-pz²² ja=tse=tse⁵² k^hi²² mlayt^ha mi²²*
 after mountain=LOC sit-NMLZR yak=ERG=PL 3.SG curse EVID
 'After, the yaks who stayed on the hill cursed them (their friends).' [YakBuff_022]
- (45) *η³=tse²² g³²² se-pz²² mi=ko=ri⁵² m^w i⁴² p^hr³⁴² pin-tsi²²*
 1.SG=ERG goat kill-NMLZR person=DEF=LOC money hundred give-PFV
 'I gave 100 rupees to the man who killed the goat.'
- (46) *t^he⁴⁴ p^ho⁴⁴ tu-pz²² kola=ko=tse²² η³=ri²² mwi⁴² p^hr³⁴² pin-tsi²²*
 always beat stay-NMLZR child=DEF=ERG 1.SG=LOC money hundred give-PFV
 'The boy who is always beaten gave me 100 rupees.'

Verbs are also nominalized for certain types of complementation, particularly purposive clauses and certain types of modal structures. In purposive clauses, the nominalized verb is suffixed with *-ri*, and in modal clauses like desideratives and abilitatives, the nominalized verb is followed by a specific modal verb.

- (47) *afan=ko²² k^himi²² k^h3-pz-ri²² sjo-pz⁴⁴ tē*
 relatives=DEF 3.PL come-NMLZR-PURP dance-NMLZR and
sle prin-tse⁴⁴ sjo-pz⁴⁴
 Ale.song affect-CC dance-NMLZR
 'In order for the relatives (of the bride) to come (to the wedding festival), there is dancing and the "Ali" song being sung, there is dancing.' [Tengki_20131016]

- (48) η^3^{22} $nu-p3^{42}$ say^{52} k^h3^{22} mo^{22}
1.SG sleep-NMLZR desire come COP
'I want to sleep.'

- (49) η^3^{22} $k^{wh}e^{42}$ $pri-p3^{44}$ $l3^{22}$ $k^h\tilde{e}^{42}$ mo^{22}
1.SG song affect-NMLZR do able COP
'I am able to sing.'

Adverbial clause modification is expressed in Manange through a variety of different morpho-syntactic structures and strategies, shown in (50).

- (50) Conditional $(kj3-)\eta_3$
Concessive $-t\eta\etaj, l\bar{e}$
Causation $-tse$, serialization, periphrasis
Purposive $-p3-ri$
Simultaneity $-tse, k^h3-ni$
Sequential (before) *pili naraŋ*
Sequential (after) $-tse, -p3-ni, -ni$

There is no clear constructional distinction between adjectives and adverbs in Manange.

With conditionals, the syntax is clause juxtaposition, with the adverbial element in between the two clauses, prosodically aligned with the first clause.

- (51) $kj3^{52}$ $kini^{52}$ $a-j3^{22}$ $kj3n3$ $kj3^{52}$ p^hi-p3^{22}
2.SG fast NEG-go COND 2.SG late-NMLZR
'If you do not go quickly, you will be late.'

With causation, multiple strategies are used. Clause chaining is frequently encountered in elicitation and in discourse, with an occasional emphasis more on the temporal relationship between two events rather than a true causal relation.

- (52) η^3^{22} $dzua$ $kjan-tse^{22}$ η^3^{22} $tu-p3^{52}$ $t3-tsi^{22}$
1.SG gamble play-CC 1.SG poor-NMLZR become-PFV
'Because I gambled, I became poor/I, gambling, became poor.'

Another strategy for expressing causation is with structures conveying both "why" and "because" preposed before the clause of causation.

- (53) hai pi^{52} $kin3$ $tek3$ kre^{22} $l3^{22}$ $a-pi\tilde{r}^{22}$
why say why.loan because climb do NEG -give
'... Because (those who stay) are not allowed to climb mountains (granted by the ancestral gods).' [Pisang_2013M2023]

Causation is also indicated by use of $l3^{22}$ 'do' juxtaposed with a verb-like adjective or another verb. Compare (54) and (55).

- (54) $t\beta^{52}$ $le-tsi^{22}$
tea warm-PFV
'The tea was warm.'

- (55) $\eta^3=tse^{22}$ $t\beta^{52}$ le^{22} $l3-tsi^{22}$
1.SG=ERG tea warm do-PFV
'I warmed the tea/I made the tea warm.'

Sequential actions are indicated by different strategies. One is nominalization followed by the suffix *-ni*, as in (56). But just as frequently, sequential actions or events are indicated by the clause-chaining suffix *-tse*, as in (57).

- (56) *p^{hi}22 t^hay⁴² k^h₃-p₃-ni²²* *pu⁴⁴ nay⁵² ts^han-tsi⁴⁴*
 chang odor come-NMLZR-ADV clay.pot inside put-PFV
 After a chang smell has come, I put (it) in a clay pot.' [Ngawal_1999F2010]
- (57) *m₃r⁵² ken-tse⁴⁴ l₃-tse²² prin-tse⁴⁴ k^h₃ta kom-p₃²²*
 butter apply-CC do-CC put-CC prayer.scarf wear-NMLZR
 'And when the butter is applied (ceremoniously), doing this, then placing the prayer scarves, (the wedding participants) wear them (as a form of blessing).' [Tengki_2013M1023]

The clause chainer is also frequently encountered with simultaneous actions and events, as in (58).

- (58) *ŋⁱ22 ker-tse²² pisay pik pt⁵²*
 1.PL.EXCL identify-CC Pisang peak say
 'This is called Pisang Peak in our own language.' [Pisang_2013M2018]

A specific clause-chained construction very frequently encountered in discourse is the *3tse²² l₃-tse²²* structure, where chained 'do' follows a discourse marker translated roughly as 'like this.' The function of this construction is to recapitulate the event of the previous clause (or even previous actions or events indicated many clauses prior in a larger stretch of discourse), illustrated in (59).

- (59) *3ni²² 3tse²² l₃-tse²² u=ri²² tinj⁴⁴ tsu⁴⁴ uqaiti l₃²²*
 and.then like.this do-CC DIST=LOC day PROX fly.loan do
a-t₃²² mi²²
 NEG -become EVID
 'And (people) doing like this, there these days, this (building) is not swept away (by avalanche).' [Pisang_2013M2051-052]

As (59) shows, there is no "same subject constraint" at work in these chained structures, as the (implied) subject referents in prior actions are people, and the subject of the post-chained clause is an (implied) force of nature.

An additional function of these recapitulation structures is to signal a perfective aspectual relation between the actions or events in a stretch of discourse; the prior action or event is treated as a whole in relation to the event articulated in the post-chained clause. With example (60a–b) below, the action of the boy telling his dog to be quiet is viewed as a complete event in relation to the ensuing action of going to search (for the frog) in (60c).

- (60) a. *3ne²² k^himi=ko²² u²² ale=ko=tse²² njukju=ko=ri²²*
 and.then 3.PL=DEF DIST boy=DEF=ERG dog=DEF=DAT
 b. *t_{fu}-p₃⁴⁴ l₃²² kje²² a-te-ro⁴⁴ pi-tsi⁵²*
 keep-NMLZR do "sound NEG -take.out-IMPER" say-PFV
 c. *3tse²² l₃-tse²² ηjo⁵² j₃²² mi²²*
 like.this do-CC search go EVID
 'After, the boy told the dog "Be quiet, do not make a sound." Having done this, they went looking.' [BDF_Eden]

FURTHER READING

In addition to the two grammar sketches (Hoshi 1986a; Hildebrandt 2004), published materials on Manange include lexicons and an analysis of lexical categories (Hoshi 1986b; Nagano 1984; Genetti and Hildebrandt 2004), phonetic and phonological descriptions (Hildebrandt 2005), considerations of language contact (Hildebrandt 2007a, 2009, 2013), and typological comparisons (Hildebrandt 2007b; Genetti *et al.* 2009). A large-scale language documentation project ongoing in Manang between 2012 and 2017 has resulted in an online text archive housed at the University of Virginia: audio-video.shanti.virginia.edu/collection/manange# (accessed 18 November 2015).

NOTES

- 1 Ongoing work on Manange is supported by a National Science Foundation grant 1149639 and a British Academy Small Grant. We would like to thank the members of the Manange speech community for teaching us about their language.
- 2 The male speaker's values are represented by larger symbols of the same shape: [i] x; [u] ▲; [e] ♦; [o] –; [ɔ] •; [ɑ] ■.
- 3 Many, but not all, Nepali loans carry one of these four tones, and there are some words in this account (words only encountered in discourse) where the tone representation has not yet been established.
- 4 Note that in example (4) the numeral *ʃɪ⁴²* ‘one’ precedes the noun *ale²²* ‘boy.’ This story was told by a Manange speaker born and raised in Kathmandu, whose language of wider communication is Nepali (Manange is used at home, with family). Among other changes observed in her speech is a change in numeral–noun ordering. See Hildebrandt (2003, 2004, 2007a) for more discussion on contact-induced changes in Manange.

REFERENCES

- Bond, Oliver, Kristine A. Hildebrandt and Dubi Nanda Dhakal. 2013. Optional ergative case marking: What can be expressed by its absence? Presentation at the 10th Biennial Meeting of Association for Linguistic Typology. University of Leipzig, Germany. 16 August 2013.
- Bond, Oliver, Kristine A. Hildebrandt and Dubi Nanda Dhakal. 2015. Conditions on differential ergative case marking, Presentation at the 89th Annual Meeting of the Linguistics Society of America. Portland, OR. 9 January 2015.
- Central Bureau of Statistics. 2012. *National Population and Housing Census 2011*. Thapathali, Kathmandu, Nepal: Government of Nepal National Planning Commission Secretariat.
- Genetti, Carol and Kristine A. Hildebrandt. 2004. The two adjective classes in Manange. In R.M.W. Dixon and A.Y. Aikhenvald (eds), *Adjective Classes*. Oxford: Oxford University Press, 74–96.
- Genetti, Carol, Alec Coupe, Ellen Bartee, Kristine Hildebrandt and You-Jing Lin. 2009. Syntactic aspects of nominalization in five Tibeto-Burman languages of the Himalayan area, *Linguistics of the Tibeto-Burman Area* 31.2: 97–143.
- Gurung, Harka. 1998. *Nepal: Social Demography and Expressions*. Kathmandu: New ERA.
- Hildebrandt, Kristine A. 2003. *Manange tone: Scenarios of retention and loss in two communities*. UC Santa Barbara. Dissertation.
- Hildebrandt, Kristine A. 2004. A grammar and dictionary of the Manange language. In Carol Genetti (ed.), *Tibeto-Burman Languages of Nepal: Manange and Sherpa*. Canberra: Pacific Linguistics, 2–189.

- Hildebrandt, Kristine A. 2005. A phonetic analysis of Manange segmental and suprasegmental properties, *Linguistics of the Tibeto-Burman Area* 28.1: 1–36.
- Hildebrandt, Kristine A. 2007a. Manange. In Yaron Matras and Jeanette Sakel (eds), *Grammatical Borrowing in Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter, 283–300.
- Hildebrandt, Kristine A. 2007b. Tone in Tibeto-Burman languages: Typological and socio-linguistic approaches. In Matti Miestamo and Bernhard Wälchli (eds), *New Trends in Typology: Young Typologists' Contributions to Linguistic Theory*. Berlin: Mouton de Gruyter, 67–90.
- Hildebrandt, Kristine A. 2009. Loanwords in Manange, a Tibeto-Burman language of Nepal. In Martin Haspelmath and Uri Tadmor (eds), *Loanword Typology*. Berlin: Mouton de Gruyter, 447–470.
- Hildebrandt, Kristine A. 2013. Manange (South Asia): A language sketch. In Carol Genetti (ed.), *How Languages Work*. Cambridge, UK: Cambridge University Press, 404–423.
- Hoshi, Michiyo. 1986a. An outline of the Prakaa grammar: A dialect of the Manange language. *Anthropological and Linguistic Studies of the Gandaki Area in Nepal*. Tokyo: Institute for the Study of Language and Cultures of Asia and Africa, 187–317.
- Hoshi, Michiyo. 1986b. A Prakaa vocabulary: A dialect of the Manang language. *Anthropological and Linguistic Studies of the Gandaki Area in Nepal* 2. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, 133–202.
- LaPolla, Randy J. 1995. On the utility of the concepts of markedness and prototypes in understanding the development of morphological systems, *Bulletin of the Institute of History and Philology* 66.4: 1149–1185.
- Mazaudon, Martine and Alexis Michaud. 2008. Tonal contrasts and initial consonants: A case study of Tamang, a ‘missing link’ in tonogenesis. *Phonetica* 65: 231–256.
- Nagano, Yasuhoshi. 1984. A Manang glossary. *Anthropological and Linguistic Studies of the Gandaki Area in Nepal* 12. Tokyo: Institute for the Study of Language and Cultures of Asia and Africa, 202–206.
- Noonan, Michael. 2003a. Chantyal. In Graham Thurgood and Randy J. LaPolla (eds), *The Sino-Tibetan Languages*. London: Routledge, 315–334.
- Noonan, Michael. 2003b. Nar-Phu. In Graham Thurgood and Randy J. LaPolla (eds), *The Sino-Tibetan Languages*. London: Routledge, 336–352.
- Tumbahang, Govinda Bahadur. 2012. Linguistic pluralism in Nepal. *Contributions to Nepalese Studies* 39: 77–104.

CHAPTER TWENTY-SEVEN

NAR-PHU¹

Michael Noonan and Kristine A. Hildebrandt

1 INTRODUCTION

The Nar-Phu language is spoken by approximately 600 people of the villages of Nar (*Ethnologue*: NPA, endonym: *tʰypruy*) and Phu (endonym: *Nartwe*), located in the valley of the Nar Khola in Manang District of Nepal. The territory they inhabit is very high (the lowest point in their territory is approximately 3,500 metres in altitude). Traditional main occupations are yak herding and small-scale farming, although *yarsagumba* (a caterpillar-fungus fusion valued for its medicinal qualities) harvesting has become an increasingly major (and controversial) economic activity in recent years.

By the standards of other languages in Nepal, the influence of Nepali on the Nar-Phu language has been relatively small. A study of Nepali (and English) lexical borrowings revealed fewer than 6 percent of loanwords in Nar-Phu, compared to 9 percent in Manange and 12 percent in the Manang variety of Gurung (Hildebrandt 2004). However, contact with Tibet and Tibetans has been fairly extensive over the years—there is a large monastery in Nar which houses a number of Tibetans—and as a result there are a large number of Tibetan borrowings in Nar-Phu. Contact with Tibetans continues and there is some literacy in Tibetan: the people of the Nar and Phu villages are adherents of a version of Tibetan Buddhism, and literacy in Tibetan is mainly achieved for the purpose of reading religious texts. The influence of Nepali is growing, however, as the school established 12 years ago and operated intermittently since introduces Nepali literacy to the population. Further, increasing numbers of people spend at least part of the year in Nepali-speaking areas and the association of competence in Nepali and economic betterment has grown in the minds of the people.

The Nar-Phu language is a member of the Tamangic group (along with Chantyal, Gurung, Manange, Tamang, and Thakali). There are a number of phonological and lexical differences between the dialects of Nar and Phu. This chapter describes the Nar dialect only. The dialects of the two villages are part of a dialect continuum with the dialects of the Manange language and from a purely linguistic perspective it is not obvious that the dialects of Nar and Phu should be accorded the status of a separate language. The primary justifications for doing so are sociological: the people of the two villages see themselves as being a group apart from the Manangis (and the local Gurungs as well). They even share a “secret language,” the point of which is to confound Manangis and Gurungs who might otherwise understand their conversations. The Manangis, apparently, share the view that the people of Nar and Phu are not Manangis, so in deference to local feelings we will consider the speech of the villages of Nar and Phu as an independent language. The decision to call the language Nar-Phu (as opposed to “Narpa,” which is perceived as the Manange designation) was made in consultation with native speakers.

2 PHONOLOGY

2.1 Vowels

The vowel system of Nar is somewhat unusual relative to the Tamangic norm. There are, for example, an unusually large number of front vowels, nasal vowels which are marginal to the system and may be resolvable to vowel + nasal consonant, and long vowels which seem always to come about via some phonological process and thus appear not to be lexical. The system of simple vowels is displayed in Table 27.1, with minimal and near-minimal sets in Table 27.2.

The acoustic distribution of the oral vowels in the vocal tract is illustrated in Figure 27.1 with data from the initial (prominent) syllables of approximately 100 mono- and disyllabic words elicited from the Swadesh word list for one female and one male speaker of Nar.

Vowels are somewhat fronted for the female in comparison to the male, and for the male, /e/ and /ɛ/ show almost total acoustic overlap, while for the female they are acoustically more distinct. The status of [œ] as an independent vowel is not clear. In what follows, it is always written <we>, a transcription that reflects, in part, its pronunciation—it is typically pronounced [wœ] or [yœ]—and the fact that the sequence [we] otherwise does not occur.

With /ɛ/ a preceding glide [j] is often heard in open, stressed syllables. Word final /ɛ/, especially in the suffix /-pe/, sounds similar to /a/ in very slow speech. /e/ and /ɛ/ contrast in open syllables, but are neutralized to [ɛ] in closed syllables: <e> is written in these cases as this corresponds to what is heard. This /ɛ/, however, never has the [jɛ] alternative pronunciation. (This reflects the fact that /ɛ/ in closed syllables derives historically from *e whereas /ɛ/ in open unstressed syllables derives historically from *a.). The vowel /æ/ is low and front-central.

Nasal vowels seem always, with three exceptions so far, to be resolvable into a v plus nasal c, the latter pronounced as a consonant (usually /ŋ/) in careful speech. The two native exceptions are [hrâre] ‘millet’ and [khēro] ‘uphill’; the other word is a borrowing from Nepali, [bāsi] ‘bamboo.’ These words have been transcribed with a sequence of vowel + consonant: /hranre/, /khenro/, and /bansi/. (Notice that the two native words have the nasal vowel before /r/: perhaps there’s a rule that deletes nasals and nasalizes vowels in such cases. No other instances of /nr/ have been found and only one instance of /ns/, /mfiunse/ ‘all night,’ has been recorded.)

TABLE 27.1 ORAL VOWELS

i	u
e	o
ɛ	œ
æ	a

TABLE 27.2 VOWEL MINIMAL/NEAR-MINIMAL SETS

ts ^b i ‘teapot’	tse ‘one (pronoun)’	kâj ‘wall’
ts ^b e ‘fight (of animals)’	tsœ ‘nerve’	kōj ‘urine’
ts ^b ɛ ‘graze’	nfiæ ‘ear’	kûn ‘center’
	nñâ ‘forest’	

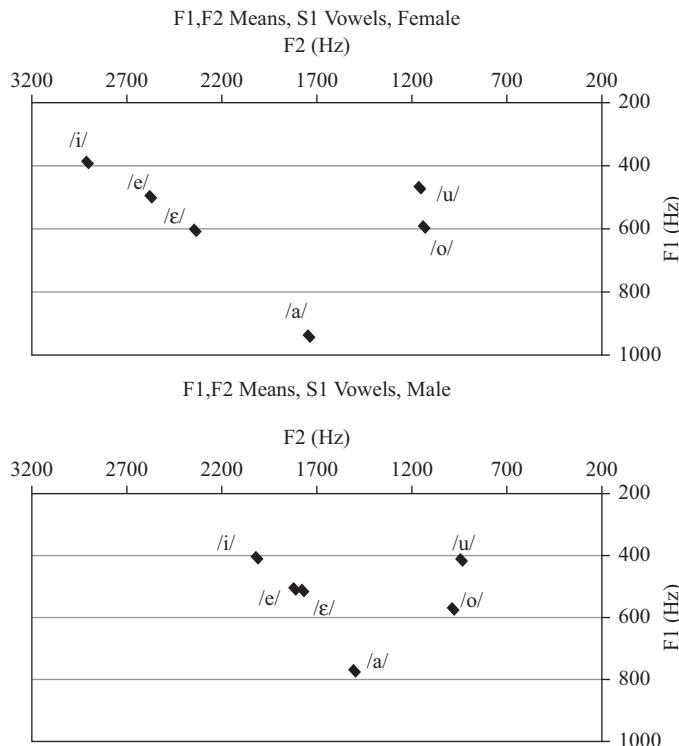


FIGURE 27.1 F1~F2 PLOTS, NAR-PHU VOWELS

TABLE 27.3 DIPHTHONGS AND SIMPLE VOWELS

Diphthong	Simple Vowel
ts ^b wɛj ‘color’	tʃwɛ ‘load,’ ts ^b ɛ ‘graze’
sów ‘apple’	só ‘tooth’
tʃuij ‘time’	tʃu ‘snap off’
ts ^b ae ‘bamboo strip’	ts ^b â ‘heat’
fiaj ‘yawn’	—

Long vowels come about as a result of phonological or morphological processes: the loss (always [?], restorable in slow pronunciation) of syllable final /k/, /p/, or /r/ in Tibetan borrowings, and the coalescence of a final front vowel in nouns with the genitive: *lame=e* ‘lama’s,’ *yé=e* ‘my.’ In a very few cases, long vowels were recorded which do not appear to be the product of any phonological process, e.g. *khee-* ‘move further away’: in these cases the long vowel may be lexical. A few borrowed words have long vowels: *amrikaa* ‘America; English (language).’

The following diphthongs have been recorded—leaving out of consideration here onglides with /j/ and /w/ (recall the special status of /wɛ/, discussed earlier), and lexical examples, contrasted with simple vowels when possible, are in Table 27.3:

ae, aj, ow, uj, wɛj

With /ae/, the first element is hard to hear and seems to vary in quality: sometimes it is a back, unrounded vowel, either [ʌ] or [ɤ]; in some cases it appears to be a velar fricative [ɣ]. This diphthong is distinct from the rare (syllable-final) /aj/, wherein the first element is a low back vowel and the glide a trajectory toward the high front position: this diphthong seems to occur only in Nepali borrowings and onomatopoetic words. /wej/ represents [œø], [œj], [yœø], or [yœj]. The conditions under which [y] appears in /wey/ and /we/ are not clear since it is sometimes present and sometimes not, likely being conditioned by the preceding consonant.

2.2 Consonants

There does not appear to be distinctive voicing among stops and fricatives (though see later), though there is with liquids: [l] and [r] contrast with [ɿ] and [ɻ]. Murmur is distinctive, but is best considered part of the tone system (see later). Seven points of articulation are attested: bilabial, lamino-dental, alveolar, post-alveolar, retroflex, palatal, and velar. The glottal stop appears to be marginally distinctive only in a few Tibetan borrowings. Aspiration is distinctive for [oral] stops and affricates. In the rest of this chapter the transcription on the right below will be used (see Table 27.4).

A set of minimal or near-minimal contrastive forms for most of these consonant phonemes is in Table 27.5, with further observations below.

The rhotic consonants are sometimes produced as taps [ɾ], sometimes as approximants [ɹ]. The alveolar-lateral fricative is variably the fricative, or more commonly, a voiceless lateral approximant [ɻ].

TABLE 27.4 CONSONANTS

	Bilabial	Dental	Alveolar	Post-Alveolar	Retroflex	Palatal	Velar	Glottal
Plosive	p p ^h	t t ^h			t (t ^h)		k k ^h	(?)
Affricate		ts ts ^h				tʃ tʃ ^h		
Nasal	m	n				ŋ	ŋ	
Fricative			s	ʃ	ʂ			
Alveolar-Lateral			ɬ					
Fricative								
Rhotic (tap)				r				
Approximant	w					j		
Voiceless Rhotic			ɻ					
Lateral Approximant			ɬ					

TABLE 27.5 MINIMAL/NEAR-MINIMAL SETS¹

pi ‘go fast, rush’	tʃ ^h i ‘grease’	kil ‘stop a herd of animals’
p ^h i ‘peel’	ʃhi ‘precious stone from Tibet’	k ^h i ‘lend’
tile ‘yesterday’	ʃi ‘die, spoil’	mi ‘eye’
t ^h i ‘break, crack’	tʃi ‘skin’	ni ‘go’
tsipe ‘coarse wool from a yak’s belly’	-ri ‘but’	nippe ‘old’
ts ^h i ‘hay’	li ‘back of something’	nj ^h i ‘seven’
tʃin ‘little’	lhipe ‘heavy’	

¹ ‘hooked h’ [ɦ] is a tonal feature, as discussed in section 2.3.

[u] replaces [w] in some speech styles for some speakers. [u] or [y] also occurs in the diphthong /ae/, e.g. in /tʰāe/ [tʰyē] ‘load.’

[?] seems to occur non-redundantly only in some Tibetan forms, where it alternates with /k/ and /p/ or a lengthened vowel in morpheme final position: /k/ and /p/ are written when this is the case. In such cases it seems to reflect the Tibetan falling tone. (In the word /kʰēpe/ ‘eighth,’ a [?] has been recorded—though not consistently: [kʰē?pe]. This does not alternate with /k/, /p/, or /t/ (though the last is what occurs in the Written Tibetan), a long vowel may be heard instead.)

Distinctively voiced stops are found in initial position in recent borrowings only. These words may contain other atypical features, e.g. distinctively nasalized vowels: [bāsī] ‘bamboo.’ In words with the murmured tones 3 and 4, the murmured voice extends, typically, to the initial consonant, so that /kʰie/ ‘work’ is often [gʰie]. In this respect, Nar-Phu differs from its relative Chantyal, in which initial voiceless stops and affricates remain voiceless in murmured syllables.

The aspirated retroflex /t^h/ is heard only in unassimilated Nepali words (*t^hik* ‘right, correct’), but such words are not yet fully integrated into Nar-Phu, and [t^h] is frequently heard instead.

There is no voiceless glottal fricative in Nar-Phu (unlike Manange), and [h] in transcription (see below) represents the two voiceless approximants. /r/ is rare, but can be contrasted with /r/: /ru/ ‘sweat, thread’ and /ru/ ‘close the eyes, squint.’ Although there are no true minimal pairs yet found, /rōŋpe/ ‘parcel, package’ may be compared with /lō/ ‘read, study, teach’ to demonstrate the two voiceless approximants.

The labio-velar and palatal approximant are also comparatively rare in the Nar-Phu lexicon, but their contrastive status can be established with the sets in Table 27.6.

For the rest of this sketch, the following transcription will be used:

p	t	t̪	k	?
ph	th	t̪h	k	
c	č			
ch	čh			
s	š			
m	n	ny	ŋ	
l				
hl				
r				
hr				
w		y		

2.3 Tone

Four tones are distinguished. Two of the tones contain murmur as a distinctive part of the realization of the tone: murmur is transcribed with the ‘hooked h’ character <ḥ>. Two of the

TABLE 27.6 LABIO-VELAR AND PALATAL APPROXIMANTS

jâ ‘hand, arm’	wan ‘go in’
tâ ‘lion’	t̪an ‘stink’
tâ ‘anything’	taj ‘pot’
râ ‘fence’	saj ‘clean something’
mâ ‘not full to the top’	paj ‘rhubarb’

tones have falling pitch: falling pitch is transcribed with the character <^>. So as to facilitate comparison, the numbering of the tones corresponds to that used for etymological sets for other Tamangic languages by Martine Mazaudon. Examples are given in Table 27.7.

Figure 27.2 shows F_0 -plots for the starting, mid-point and ending fundamental frequency values for the minimal set of nasal-initial words, uttered by a female Nar speaker. All words were uttered in a frame-medial sentence context of ‘I saw the X/I said X,’ where X is the target word illustrated in this figure, and the sentence is verb-final.

Figure 27.3 shows F_0 -plots for the starting, mid-point, and ending fundamental frequency values for the minimal set of obstruent-initial words, uttered in the same type of frame.

The fall-vs-rise difference between tones 1 and 2 and tones 3 and 4 is better appreciated with sonorant-initials. Tone 2 is distinguished by its clear, high quality. One ongoing question about the nature of tone in languages like Nar has to do with the phonetic manifestation of breathy or murmured qualities for the two low-register tones (tones 3 and 4). In this sub-grouping, Manange is unique in having phonologized tone as aligning primarily along vowel fundamental frequency (F_0 /pitch) and not consonant/vowel phonations. Conversely, Tamang and Gurung are said to manifest tone as a combination of vowel F_0 and differing consonant and/or vowel phonation types. In low tones in these languages, initial obstruents display an optional, slightly or semi-voiced quality, which also corresponds with a general laxness, evidenced by breathy, whispery or murmured voicing to the vowel/nucleus.

In Nar, the role of breathiness in distinguishing the two low-register tones is currently not clear. Acoustic phonetic measurements show that words in tones 3 and 4 are tentatively

TABLE 27.7 TONES

Tone number	Pitch contour	Nasal-initial example	Obstruent-initial example
/1/	53	<i>nâj</i> ‘reciprocal obligation’	<i>pâj</i> ‘true’
/2/	44	<i>naj</i> ‘full’	<i>paj</i> ‘animal pen’
/3/	12	<i>nħaj</i> ‘planted in rows’	<i>pħaj</i> ‘thatch’
/4/	21 or 31	<i>nħāj</i> ‘inside’	<i>pħāj</i> ‘fight’

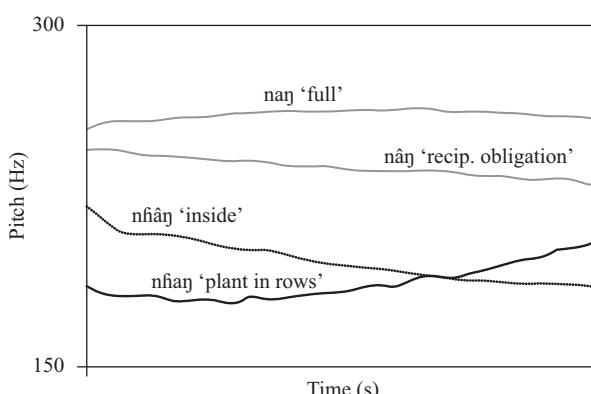


FIGURE 27.2 NASAL-INITIAL PITCH TRACES

distinguished from tone 1 only in terms of vowel jitter (a-periodic or irregular glottal pulses) and vowel spectral tilt, also known as Fast-Fourier Transform (FFT) spectrum analysis, which measures the difference in energy between the first two harmonics of a voiced sound ($F_0 - H_2$). But this distinction shows interspeaker variation. Similarly to what has been observed for Tamang (Mazaudon and Michaud 2008), word-medial inter-vocalic plosives tend to be produced with weak voicing in low-register tones in Nar, but again, this is not totally regular.

Acoustic cues to breathy/murmured phonation can be appreciated somewhat more in sonorant-initial words, as Figure 27.4 shows.

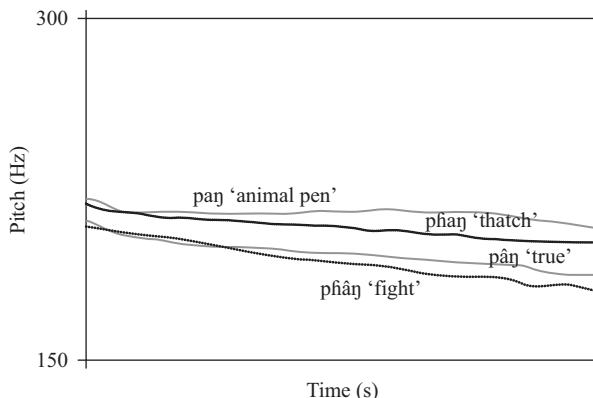


FIGURE 27.3 OBSTRUENT-INITIAL PITCH TRACES

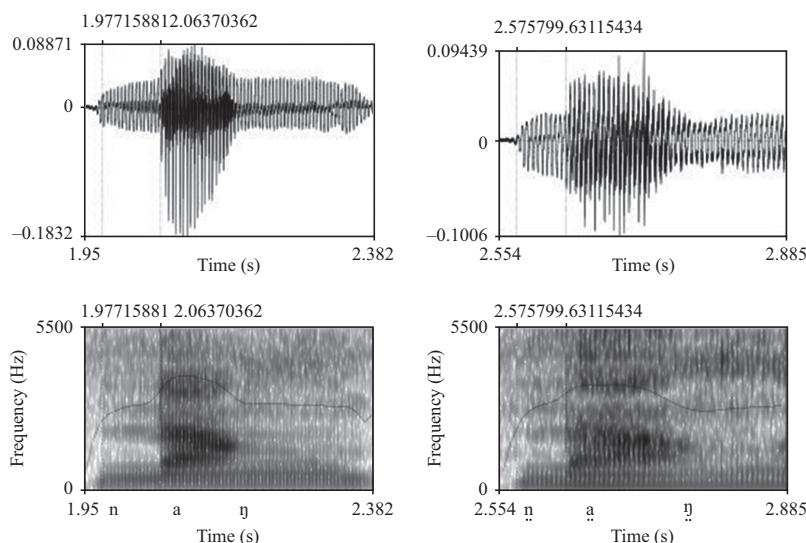


FIGURE 27.4 WAVEFORM AND SPECTROGRAM OF *Nâd* 'OBLIGATION' AND *Nħâd* 'INSIDE'

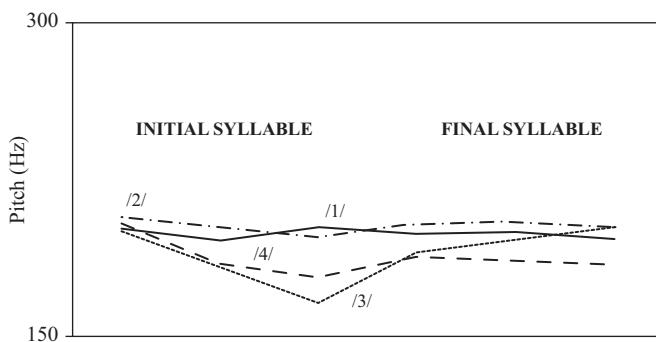


FIGURE 27.5 PITCH TRACES FOR DISYLLABIC WORDS

nħāy ‘inside,’ on the right side of the image, shows comparatively weaker voicing cycles in the waveform and weaker or more ill-defined formant separation for both the nasal and the vowel (even with the coda velar nasal) in the spectrogram.

The pitch contours are ordinarily distributed over entire words; that is, they distribute over a root and any affixes it may take. Figure 27.5 shows average pitch-traces for disyllabic words representing the four tones, recorded from the same female Nar speaker.

Tone 1 and 2 words carry higher pitches across both syllables than do tones 3 and 4, and tone 3 words show the fall-rise contour as it carries across the disyllabic word. In general, affixes have no independent tone, but a few, recently developed from verbal or nominal roots, do possess independent tones. In compounds, however, each element has its own tone.

2.4 Stress

Words are stressed on the first syllable of the root; this is true also of compounds, which are stressed on the first element of the compound.

2.5 Phonotactics and phonological alternations

While all consonants in Table 27.4 are attested in word-initial position, C2 onset position is restricted to r, l, w, y, and the cluster possibilities in this position are limited. These possibilities are shown in Table 27.8.

TABLE 27.8 ONSET CONSONANT CLUSTERS

pfri ‘female yak’	phrâŋ ‘rosary’	mriŋ ‘woman’
tħron ‘slope’	kħri ‘wheat’	kħru ‘wash’
plâ ‘get loose’	phlâ ‘slip’	klê ‘field’
mfilân ‘black’	pŵi ‘husk’	tħiwina ‘grain scattered in worship’
thwîme ‘joint’	cwêsôm ‘thirteen’	chwe ‘color’
čħwe ‘drink.HON’	kħwɛ ‘honey’	khwoton ‘ditch’
q̊wɔnṭe ‘front’	mħiwise ‘all night’	nwē ‘mat for sitting/sleeping’
kfirwē ‘eagle’	ħfwêy ‘slide down on posterior’	pyuŋ ‘male, husband’
phyan ‘over’	kyu ‘water’	khjô ‘put in’
mye ‘arrow’	nyä ‘call’	

TABLE 27.9 CONSONANTS IN FINAL POSITION

čfin ‘finish something’	khop ‘all’	ik ‘hiccup’
kfilám ‘road’	pñräj ‘Buddhist rosary’	
kfyor ‘copper’	rol ‘escort’	

There are a few gaps, most notably with the alveolar and retroflex series. Also, the affricates have fewer combinatorial possibilities, and the palatal nasal /ny/ is not found in clusters. It is also worth noting that Cw clusters primarily precede the front vowels. Also of note is that Nar-Phu is different from Manange in that [w] is analyzed in Manange as part of a labio-velar series (Hildebrandt, this volume) rather than as an independent consonant that can occur in clusters. In Manange, word-initial [w] is virtually unattested (but it is commonly found in Nar-Phu). Additionally, in Manange it follows only bilabial and velar consonants, while in Nar-Phu it occurs in clusters with consonants from a wider range of positions of articulation.

In word final position, a smaller range of nasals, approximants and unaspirated stops may occur, and this is in general less frequent in the Nar-Phu lexicon, as in Table 27.9.

Affricates and retroflex stops do not occur in final position in native words, but the retroflex stop occurs word-finally in a few borrowings from English and Nepali, as do /s/ and /ʃ/. In general, these loans stand counter to the phonotactic generalizations of Nar-Phu: *bâksa* ‘large box,’ *amrika* ‘foreigner, American, English,’ *not* ‘paper money,’ *estil* ‘steel,’ *čækte* ‘jacket,’ *ijliš* ‘English.’

Medial clusters of consonants occur in a few (probably derived) words (*hlâmpe* ‘scrap of cloth,’ *čarpe* ‘sticky,’ *čhanju* ‘sweetened chang,’ *kalpe* ‘alike,’ *chunta* ‘finger’) and in compounds. In these cases, any combination of final and initial consonant (including consonant clusters) is possible. Words synchronically analyzable as polysyllabic—if native—are derived historically from polymorphemic words. As a result, all the cluster types allowable in compounds are allowable in such words.

Unaspirated consonants are voiced intervocally (including internally within compounds), and, intervocally, aspirated stops are pronounced with aspiration in slow, careful pronunciation, but lose their aspiration in casual speech.

Except in some recent borrowings from English, /æ/ occurs only in word final position. In compounds, when it is in the first element it becomes variously [a] or [o] (/nfiæ/ ‘ear,’ [nfaci tñtñj] ‘area in front of the ear,’ [nfokli] ‘earwax’); when it is the final element it usually becomes [ɛ] (/tæ/ ‘horse,’ [phôrta] ‘gelding,’ [môrte] ‘mare’), though there are apparent exceptions. In inflection, /æ/ becomes [a] (/ñâ/ ‘I,’ [ñâ=se] ‘1SG=ERG.’) All other vowels occur initially, finally, and medially, though some vowel changes occur (apparently) irregularly in the first component of noun compounds.

In compounds, consonants which have been lost in free standing forms may surface: /tæ/ ‘horse,’ [phôrta] ‘gelding,’ [môrte] ‘mare’ (Written Tibetan has *rta* ‘horse’); /nôw/ ‘snot,’ [nopšu] ‘handkerchief.’

3 MORPHOLOGY

3.1 Generalizations

Nar-Phu is overwhelmingly suffixing and agglutinative. The only prefix is negative /a-/: *a-câ-w* [NEG-eat-IMPER] ‘don’t eat it!’

3.2 Nouns

Nouns in Nar-Phu are marked as plural by means of the clitics *=cuke* and *=ce*. The form *=ce* behaves as an ordinary clitic and has no independent tone; *=cuke*, on the other hand, always has tone 2 and generally behaves like the second element of a compound. Notionally plural NPs with count noun heads are not obligatorily marked for plurality, but usually are. NPs whose heads are quantified by numerals, however, are not marked for plurality.

The plural morphemes are phrase final clitics and thus attach to the last word of the NP:

- (1) *nôkyu=cuke* ‘dogs’
nôkyu mhilay=cuke ‘black dogs’

Pluralization is not obligatory, and is frequently unexpressed in discourse if plural referents are quantified by a numeral:

- (2) *êle som-pe=ri* *khæ-ce*
boy three-NMLZR-INDEF come-PFV
‘Three boys have arrived (on the scene).’ [Pear Story]

Case is marked with a relatively small set of case clitics. Like the plural clitics, these forms are NP-final. When a plural clitic and a case clitic are found in the same NP, the case clitic is last:

- (3) *nôkyu=cuke=re* ‘to the dogs’

The case clitics are listed below:

Absolutive	unmarked
Ergative, instrumental, ablative	<i>=se</i>
Dative, locative	<i>=re</i>
Genitive	<i>=(y)e, =i</i>
Independent genitive	<i>=nê</i>
Comitative	<i>=ten</i>

The absolute case is unmarked: it is used with intransitive subjects, (many) direct objects, and predicate nominals.

The ergative/instrumental/ablative serves a variety of functions: it marks transitive subjects (fairly consistently: see §5.2) as well as NPs understood as instruments and those indicating source (ablative). It is possible to have more than one NP marked with this suffix in a given clause provided that each instance is understood as coding a different relation.

The dative/locative is used with indirect objects, with certain direct objects (most human referents), with NPs having allative (motion toward) senses, and with locatives with a static ‘location at’ sense.

The genitive is used to code any sort of attributive relation subordinating one NP to another; it is often omitted. The independent genitive is used to code headless genitival relations (e.g. ‘John’s beat mine’), including those functioning as predicate nominals.

The comitative is used to code the comitative ‘with’ relation, as in (4).

- (4) *êle ce saikul=ten ni-ce*
boy DEF bicycle=COMIT go-PFV
‘The boy goes with his bicycle.’ [Pear Story]

Partitive relations are expressed without any special case marking: the substance measured is followed by the measure noun, which in turn is followed by any numerals, as in (5).

TABLE 27.10 GENDER SUFFIXES

Ungendered noun	Male	Castrated male	Female
<i>r̥ho</i> ‘bond friend’	<i>r̥hope</i>		<i>r̥home</i>
<i>nôkyu</i> ‘dog’	<i>nôkyupho</i>		<i>nôkyumo</i>
<i>râ</i> ‘goat’	<i>râpho</i>	<i>râpe</i>	<i>râmo</i>
<i>tæ</i> ‘horse’	<i>photen</i>	<i>phôrte</i>	<i>môrte</i>

- (5) *čhæ* *šô* *som*
 tea cup three
 ‘three cups of tea’

The case of objects of postpositions is governed by the postposition and is, depending on the postposition, the absolute, the genitive, or either.

There are no concord classes (genders) in Nar-Phu. There is, however, a system of suffixes used to create nouns which refer explicitly to male or female people or animals and even, in a few cases, to castrated males. A few examples are shown in Table 27.10.

3.3 Honorific vocabulary

Nar-Phu has sets of honorific nouns and verbs which are used when referring to people to whom special respect is due, such as lamas, important government officials, etc. In addition, there are two verbs which are specifically ‘humble,’ i.e. are used by the speaker as a means of showing special deference to an honored addressee. In many cases, the honorific vocabulary bears no phonological (or etymological) relation to the ordinary term.

Honorific noun counterparts of ordinary nouns are found only for names of food items, body parts, and items of clothing. When no special honorific exists, a set of morphemes may be employed to create new forms: *šhe* (food items), *čha* (body parts), *šhep* (clothing). The resulting honorific may still display considerable phonological idiosyncrasy. Some examples follow in Table 27.11.

Honorific verbs are illustrated in Table 27.12.

TABLE 27.11 HONORIFICS

	Ordinary	Honorific
‘beard’	<i>kyôw</i>	<i>šalcham</i>
‘body’	<i>čhû</i>	<i>kîsuk</i>
‘boot’	<i>khyô</i>	<i>šhepkhyô</i>
‘butter’	<i>mhar</i>	<i>šhemhar</i>
‘chang (Tibetan beer)’	<i>phow</i>	<i>čhwečhâj</i>
‘eye’	<i>mi</i>	<i>čen</i>
‘face’	<i>yotoj</i>	<i>šhal</i>

TABLE 27.12 HONORIFIC VERBS

	Ordinary	Honorific	Humble
‘be sick’	<i>na-</i>	<i>nyûñ-</i>	
‘buy, take’	<i>kîn-</i>	<i>šhî-</i>	
‘get up, stand’	<i>re-</i>	<i>šhâñ-</i>	
‘give’	<i>pîñ-</i>	<i>nâñ-, nâñ kê-</i>	<i>phûl-</i>
‘smell (tr.)’	<i>nay-</i>	<i>nay-</i>	<i>sûl-</i>

3.4 Numerals

Nar-Phu cardinal numbers have been greatly influenced by Tibetan, especially the higher numbers. The numerous irregularities are largely the product of borrowing from Central Tibetan: CT low tone is borrowed as the murmured tone 4. The numerals are shown in Table 27.13.

Internal to the number phrase itself, numbers follow the order of larger-smaller when additive ($40 + \text{seven} = 47$), smaller-larger when multiplicative ($\text{four} \times 100 = 400$), except after *toŋta* ‘1,000’, which is treated as a noun and therefore followed by the numerals that quantify it. The only other complication with numbers lies with the use of the ergative/instrumental/ablative suffix *-se* following hundreds and thousands. Some examples are in (6):

- (6) ‘47’ *p̪hliču ɻi*
 ‘439’ *šipkye-se sômču ku*
 ‘1,996’ *toŋta khři-se kupkye-se kûču thûk*
 ‘67,735’ *toŋta thukču ɻi-se thuŋkye-se sômču ɻhâ*
 ‘347 dogs’ *nôkyu som khřeč-se p̪hliču ɻi*

The ordinal numbers are entirely borrowed from Central Tibetan. They evidence numerous complexities and only the first three decades are given in Table 27.14.

Nar-Phu does not have a system of classifiers.

3.5 Pronouns, demonstratives and articles

The paradigms for the personal pronouns are displayed in Table 27.15.

TABLE 27.13 CARDINAL NUMBERS

1	<i>kři</i>	11	<i>cûkhři</i>			100	<i>khyâ</i>
2	<i>ɻhî</i>	12	<i>cûjhři</i>	20	<i>ɻhyûšu</i>	200	<i>ɻhî khyâ</i>
3	<i>som</i>	13	<i>cwêysôm</i>	30	<i>sômču</i>	300	<i>som khyâ</i>
4	<i>p̪hlî</i>	14	<i>cûlhři, cûlp̪hlî</i>	40	<i>p̪hliču</i>	400	<i>šipkyc</i>
5	<i>ɻhâ</i>	15	<i>cûjhřa</i>	50	<i>ɻhaču</i>	500	<i>ɻhapkye</i>
6	<i>thûk</i>	16	<i>cûthřuk</i>	60	<i>thukču</i>	600	<i>thupkye</i> (sic)
7	<i>ɻi</i>	17	<i>cwêyñi</i>	70	<i>ɻjču</i>	700	<i>thuŋkye</i>
8	<i>p̪hře</i>	18	<i>côphře</i>	80	<i>p̪hreču</i>	800	<i>p̪hrekyc</i>
9	<i>ku</i>	19	<i>cûrkú</i>	90	<i>kûču</i>	900	<i>kupkye</i>
10	<i>ču</i>					1,000	<i>toŋta</i>
100,000	<i>lak, p̪hum</i>						
10,000,000	<i>thuŋčur, thuykur</i>						

TABLE 27.14 ORDINAL NUMBERS

1st	<i>thayp�</i>	10th	<i>čup�</i>	20th	<i>ɻhyûšupe</i>
2nd	<i>ɻh�pe</i>	11th	<i>čupc�k�pe</i>	21st	<i>ɻhyûšu cakc�k�pe</i>
3rd	<i>s�mp�</i>	12th	<i>čuŋ�k�pe</i>	22nd	<i>ɻhyûšu caŋh�k�pe</i>
4th	<i>š�pe</i>	13th	<i>čuksump�</i>	23rd	<i>ɻhyûšu caks�mp�</i>
5th	<i>ɻ�pe</i>	14th	<i>čupsh�pe</i>	24th	<i>ɻhyûšu capsh�pe</i>
6th	<i>th�k�pe</i>	15th	<i>ceŋh�k�pe</i>	25th	<i>ɻhyûšu ceŋh�k�pe</i>
7th	<i>th�n�pe</i>	16th	<i>čuŋh�k�pe</i>	26th	<i>ɻhyûšu ceʃh�k�pe</i>
8th	<i>kh�pe</i>	17th	<i>čupt�n�pe</i>	27th	<i>ɻhyûšu capth�n�pe</i>
9th	<i>kh�p�</i>	18th	<i>cepkh�pe</i>	28th	<i>ɻhyûšu capkh�pe</i>
		19th	<i>curk�pe</i>	29th	<i>ɻhyûšu capk�pe</i>

TABLE 27.15 PERSONAL PRONOUNS

	1st SG	1st EMPH	1st PL INCL	1st PL INCL	1st PL EXCL	1st PL EXCL	1st PL PARTN
ABS	<i>ŋâ</i>	<i>ŋhâ</i>	<i>ŋhî</i>	<i>ŋhî-cuke</i>	<i>ŋhyâj</i>	<i>ŋhyâj-cuke</i>	<i>kuj</i>
GEN	<i>ŋê-e</i>	<i>ŋhâ-e</i>	<i>ŋhî-i</i>	<i>ŋhî-cuke-e</i>	<i>ŋhyâj-e</i>	<i>ŋhyâj-cuke-e</i>	<i>kuj-e</i>
IGEN	<i>ŋê-e-nê</i>	<i>ŋhâ-nê</i>	<i>ŋhî-i-nê</i>	<i>ŋhî-cuke-nê</i>	<i>ŋhyâj-nê</i>	<i>ŋhyâj-cuke-nê</i>	<i>kuj-nê</i>
E/I/A	<i>ŋâ-se</i>	<i>ŋhâ-se</i>	<i>ŋhî-se</i>	<i>ŋhî-cuke-se</i>	<i>ŋhyâj-se</i>	<i>ŋhyâj-cuke-se</i>	<i>kuj-se</i>
D/L	<i>ŋâ-re</i>	<i>ŋhâ-re</i>	<i>ŋhî-re</i>	<i>ŋhî-cuke-re</i>	<i>ŋhyâj-re</i>	<i>ŋhyâj-cuke-re</i>	<i>kuj-re</i>
COM	<i>ŋâ-ten</i>	<i>ŋhâ-ten</i>	<i>ŋhî-ten</i>	<i>ŋhî-cuke-ten</i>	<i>ŋhyâj-ten</i>	<i>ŋhyâj-cuke-ten</i>	<i>kuj-ten</i>
	2nd SG FAM	2nd PL FAM	2nd SG	2nd PL			
ABS	<i>nujay</i>	<i>nujay-cuke</i>	<i>khyay</i>	<i>khin-cuke</i>			
GEN	<i>nujay-e</i>	<i>nujay-cuke-e</i>	<i>khi-i/ye</i>	<i>khin-cuke-e</i>			
IGEN	<i>nujay-nê</i>	<i>nujay-</i> <i>cuke-nê</i>	<i>khyay-nê,</i> <i>khi-i-nê</i>	<i>khin-</i> <i>cuke-nê</i>			
E/I/A	<i>nujay-se</i>	<i>nujay-cuke-se</i>	<i>khyay-se</i>	<i>khin-cuke-se</i>			
D/L	<i>nujay-re</i>	<i>nujay-cuke-re</i>	<i>khyay-re</i>	<i>n-cuke-re</i>			
COM	<i>nujay-ten</i>	<i>nujay-</i> <i>cuke-ten</i>	<i>khyay-ten</i>	<i>khin-</i> <i>cuke-ten</i>			
	3rd SG FAM	3rd SG	3rd RESP	2nd PL FAM	3rd PL	3rd GEN	
ABS	<i>nuy</i>	<i>cû</i>	<i>hota</i>	<i>nuyye</i>	<i>cû-cuke</i>	<i>ce</i>	
GEN	<i>nwe-ye</i>	<i>cû-i</i>	<i>hota-e</i>	<i>nuyye-ye</i>	<i>cû-cuke-e</i>	<i>ce-ye</i>	
IGEN	<i>nuy-nê</i>	<i>cû-i-nê, cû-nê</i>	<i>hota-nê</i>	<i>nuyye-nê</i>	<i>cû-cuke-nê</i>	<i>ce-nê</i>	
E/I/A	<i>nuy-se</i>	<i>cû-se</i>	<i>hota-se</i>	<i>nuyye-se</i>	<i>cû-cuke-se</i>	<i>ce-se</i>	
D/L	<i>nuy-re</i>	<i>cû-re</i>	<i>hota-re</i>	<i>nuyye-re</i>	<i>cû-cuke-re</i>	<i>ce-re</i>	
COM	<i>nuy-ten</i>	<i>cû-ten</i>	<i>hota-ten</i>	<i>nuyye-ten</i>	<i>cû-cuke-ten</i>	<i>ce-ten</i>	

In the first person, inclusive and exclusive forms are distinguished in the plural. Both plural forms can combine with the plural morpheme *-cuke*, which seems to be used in cases of three or more where the individuality of the members of the group is emphasized, as opposed to the membership in a collectivity. In the first person there is also a special ‘emphatic’ form, used primarily when the first person referent is in contrastive focus: it is neutral between singular and plural readings. Further there is a special ‘partnership’ pronoun, which implies action as a partnership or collectivity.

In the second person, there is a truly ‘familiar’ form, used only with family members. It is not, for example, used with servants or others for whom little social respect is due: in such cases the noun would substitute for the pronoun if an overt reference is required contextually and the ordinary second person pronoun is thought inappropriate. The familiar pronoun has a plural counterpart, formed with the regular plural suffix. There is also a non-familiar second person pronoun, used for most other second person referents, i.e. for people with whom one is not related and for whom special respect or low respect is due: this form also has a plural formed with the regular plural suffix. Where high respect is due, as with people entitled to honorific nominal and verbal forms, the appropriate noun (and, therefore, third person reference) is used instead of a second person pronoun.

In the third person, a familiar form is distinguished from a neutral form and an honorific form. The last has no special plural: there is no form **hota-cuke*. The other third

person pronouns have plurals formed with the regular plural morpheme. The generic third person pronoun translates as ‘one’: it can be used to refer to both human and non-human referents.

The interrogative pronouns are listed in (7):

- (7) ‘what’ *tæ*
 ‘who’ *sū*
 ‘where’ *khana*
 ‘which’ *khoncu*
 ‘how much/many’ *khate*

These forms are used with the interrogative forms of the verb:

- (8) $\check{c}hæ$ *khate* *mu-pe*
 tea how.much be-INTERROG
 ‘How much tea is this?’

Interrogative pronouns are not fronted within their clause, being placed ordinarily in the slot appropriate to their grammatical role or the preverbal focus position:

- (9) *lakpe=se* *sū=re* *mraj-pe*
 Lakpa=ERG who=DAT/LOC see-INTERROG
 ‘Who did Lakpa see?’

Some negative pronouns and adverbs have been observed in discourse contexts: *tajay*, *cikay*, *cilki*, *lalemhi* all meaning something akin to ‘nothing/nobody,’ *kate-maj* ‘never,’ lit. Nepali ‘how much’-never (Bond and Hildebrandt 2011).

Determiners constitute a simple set from an areal perspective: proximal *cū* contrasts with distal *tēta* (fast speech) and *theta* (slow, careful speech). These forms are used attributively and pronominally. The generic third person pronoun *ce* also has a definite article-like function and lacks any sense of spatial deixis; it follows the nouns it determines: *phālpe ce* ‘the toad,’ *the-pe ce-se* ‘the big one.’

There is an indefinite singular article *-ri*, derived historically from the numeral *khri* ‘one.’ It is an NP-final clitic. It is not used with all indefinite singular NPs, however. Rather, its use parallels the use of ‘this’ as an indefinite specific article in spoken English, introducing topical referents into discourse:

- (10) *ŋæ* *mhi=ri* *mraj-cin*
 1SG person=INDEF see-PST
 ‘I saw a certain person’ ~ ‘I saw this guy.’

The implication is that the referent so introduced will be the topic of the following discourse.

3.6 Verbs

Verbs are inflected for tense, aspect, mood, evidentiality. Negative verbal forms involve the negative prefix *a-* and, often, the suffix *-i* substituting for a positive tense-aspect suffix. There is a special interrogative suffix *-pe*. Periphrastic verbal constructions are common: the auxiliary verb is always a copular verb. Verbs may be nominalized (becoming verbal nouns) or adverbialized (becoming conversbs, that is non-finite verbs having adverbial functions). Verbs are not inflected for agreement with arguments, for direction, or for voice; they do not demonstrate a conjunct/disjunct distinction (but see the next paragraph). Verbs are not morphologically marked for transitivity.

Central to the organization of the verbal paradigm is the direct/indirect evidential distinction. Direct forms are used to report situations that the speaker has witnessed; indirect forms are used to report situations that the speaker has indirect knowledge of, that is, has come to know of the situation secondhand through hearsay, through inference, and so on. Sentences with first person subjects are ordinarily direct, though if the speaker were, for example, sleepwalking and were told of his/her actions by another, an indirect verb form would be used.

Nar-Phu has a past–present tense distinction and perfective–imperfective aspect, and this is shown in Table 27.16.

In general, the indirect forms are identical to the direct forms with the addition of a copula (see later for discussion of the copular verbs): if a copula is already associated with a direct form, a sort of double copula is used with the corresponding indirect.

The ‘past’ is used to refer to situations in the non-immediate past. It ordinarily is associated with perfective aspect but speakers recognize these as distinct forms with distinct functions. The suffix *-ce* (and its allomorph *-se*) is a gnomic (universal) perfective, and *-te* is imperfective in the sense that it codes events or actions with some duration (Hildebrandt 2011).

Main clause modal affixes include the following:

Imperative: singular subject	<i>-(t)aw</i>
Imperative: plural subject	<i>-(t)ow</i>
Hortative	<i>-so</i>
Potential	<i>-ki</i>

Subordinate clause affixes include the following:

Perfective converb	<i>-ce ~ -se</i>
Imperfective converb	<i>-te</i>
Conditional	<i>-reme, -teme</i>
Indeterminate nominalizer	<i>-pe</i>
Potential nominalizer	<i>-ne</i>
Past relative	<i>-pi</i>
Non-past relative	<i>-pe</i>

The converb *-ce* codes the bounded temporal nature of the non-finite verb in relation to the finite verb, while *-te* codes the durative nature of the non-finite verb in relation to the

TABLE 27.16 TENSE-ASPECT

	Direct		Indirect	
	Positive	Negative	Positive	Negative
Past	<i>v-čin</i>	<i>a-v-i</i>	<i>v-či(n) mû</i>	<i>a-v-i mû</i>
Perfective	<i>v-ce</i>	<i>a-v-i</i>	<i>v-ce mû</i>	<i>a-v-i mû</i>
Present	<i>v (mû)</i>	<i>a-v mû</i>	<i>v(-pe/te) mûmu</i>	<i>a-v(-pe/te) mûmu, v(-pe/te) hârmu</i>
Imperfective	<i>v-pe/te mû</i>	<i>a-v-i</i>	<i>v-pe/te mûmu</i>	<i>a-v-i mûmu</i>

Note: In this table, ‘v’ is used for any verbal root.

finite verb. Examples of these different temporal relations are shown in (11) and (12) respectively.

- (11) *hota=se tû-ce phi tû mumû*
3SG=ERG stay-CONV speak stay be:PRES
'Having sat down, he sits speaking.'
- (12) *phjay=re ni-te hli ce krê-ce*
up=LOC go-CONV ladder DEF climb-PFV
'He then climbs up the ladder.' [Pear Story]

The potential nominalizer as an indeterminate temporal sense, where the completion of the nominalized action or event is not specified, as in (13).

- (13) *ŋâ=se lakpe=re hîke phri-pe mray-čin*
1SG=ERG Lakpa=DAT/LOC letter write-NMLZR see-PST
'I saw Lakpa writing the letter.'

There is also an unaffixed subordinate verb form which is used in constructions with a range of meanings, including ingressive and egressive senses and causative senses. The generic name for this sort of construction is 'verb concatenation.' The last verb in the concatenation chain is inflected; there may be up to four verbs in the concatenation chain. These are shown in (14) through (17):

- (14) *p̥hra ni-w*
walk go-IMPER
'Go for a walk!'
- (15) *kyây lâ-w*
reach do-IMPER
'Reach for it!'
- (16) *târ kyû pi ni-w*
be.in.order run go.fast go-IMPER
'Run in order!'
- (17) *nôkyu ce=se thuy nâpraj čhây ce pi tê lâ-čin*
dog DEF=ERG bee nest DEF go.fast fall do-PST
'The dog unwittingly knocked down the beehive.' [Frog Where Are You]

3.7 Copular verbs

There are two copular verbs, one of which, *nhe*, is defective in lacking past forms and in lacking a contrast between direct and indirect forms. The two copular verbs are listed in Table 27.17.

In terms of use, *mû* is the unmarked form and is the form used in verbal periphrasis. Its basic use is in clauses asserting identity, but it has encroached on the territory of *nhe*, whose basic use is the expression of location. In the past tense, *mû* is used exclusively. The nominalized copular verb also occurs preceding the reported speech particle *ro* in situations of reported or in narrative genres. Examples of these forms are given in (18) through (21), and in §5.4.

TABLE 27.17 COPULAR VERB FORMS

	<u>mû</u>	<u>n̄fe</u>	
	Direct	Indirect	
Declarative	<i>mû</i>	<i>mûmu</i>	<i>n̄fe</i>
Negative	<i>hare</i>	<i>hârmu</i>	<i>a-yin</i>
Interrogative	<i>mu-pe</i>	<i>mu-pe</i>	<i>hin-pe</i> (> <i>himpe</i> > <i>me</i>)
Negative interrogative	<i>har-pe</i>	<i>har-pe</i>	<i>ha-in-pe</i>
Declarative past	<i>mû</i>	<i>mû-i</i>	
Negative past	<i>hare</i>	<i>hare</i>	

- (18) *phâlpe ce har-ce*
 toad DEF NEG.be-PFV
 ‘The toad wasn’t there.’ [Frog Where Are You]
- (19) *kfjû phô a-yin mwo*
 sheep male NEG-be female
 ‘It is not a male sheep, it’s female.’ [Sheep Organs]
- (20) *che yorčhen prânce mû-pe ro*
 child many weak be-NMLZR REP
 ‘Many children are weak.’ [The Strongest Man in the World]
- (21) *khrôy tel pho thop cê lâ-pi mfi hârmu*
 die after belly stone ceremony do-NMLZR person NEG.be
 ‘After he died, there is no one to do the Belly Stone Ceremony.’ [The Belly Stone Ceremony]

3.8 Adjectives

Adjectives can be distinguished from verbs in that, when they are predicates, they do not take verbal inflections and require a copular verb, and, when they are attributive, they follow their heads rather than precede them as relative clauses do.

A few adjectives have the nominalizing suffix *-pe* as a fixed component, but the suffix is not a regular part of Nar-Phu adjectival morphology.

3.9 Expressive vocabulary

Like its relative Chantyal, Nar-Phu has an extensive set of lexical items which I will refer to as ‘expressive vocabulary.’ These words describe sounds (often onomatopoetically), the appearance of things, modes of action, physical sensations, or some combination of these factors. The words are usually full reduplications and are overwhelmingly in tones 2 and 3. Some examples follow:

- charcar* ‘falling down’
thokthok ‘rapping, pounding’
khusti ‘in a joking manner’

Example (22) shows the use of the expressive vocabulary in a sentence.

-
- (22) *thō charcar tæ-pe mü*
saliva falling.down become-NMLZR be:PRES
'He's slobbering.'

3.10 Verb particles

Nar-Phu has a set of particles which express senses ranging from evidentiality to emotional reaction to the state of affairs expressed in the sentence. They are referred to as verb particles because they attach as clitics to the finite verb. The most common are *kâ*, which has a contrastive emphatic sense, *ro*, which indicates reported speech, and *ri*, which indicates that the statement is factual, i.e. is not hearsay.

4 WORD FORMATION

4.1 Derivation

Nar-Phu is not particularly rich in derivational morphology. Apart from the nominalizer and converbal affixes, there is relatively little. The gender affixes on nouns have been discussed earlier. There is also the suffix *-ten*, which means 'one from': *čhupruŋ-ten* 'one from Nar,' *pħälpe-ten* 'one from Kathmandu.'

4.2 Compounding

Compounding is productive in Nar-Phu and an important morphological process. Both noun-noun compounds and noun-verb compounds exist. Here are a few examples of noun-noun compounds:

<i>mħilaykhuŋ</i>	'throat'	[< <i>mħilay</i> black + <i>khuŋ</i> hole]
<i>mēphra</i>	'ash'	[< <i>mē</i> fire + <i>pħrā</i> flour]
<i>kħrēmso</i>	'molar'	[< <i>kħrem</i> cheek + <i>sô</i> tooth]
<i>hlachħwe-</i>	'pray, worship'	[< <i>ħla</i> god + <i>čħwe-</i> pray]

As these examples show, a number of phonological processes are associated with the compounding process, though these are not yet fully understood.

5 SYNTAX

5.1 The structure of the noun phrase

The canonical order is as follows:

RELATIVE CLAUSE + DEMONSTRATIVE/GENITIVE + HEAD NOUN + ADJECTIVE + NUMERAL

This order is fairly consistently maintained, with one exception: in a very few cases adjectives precede their heads, e.g. *pħutlu mħi* 'short person.' These cases may, on closer inspection, turn out to be compounds since the order within compounds is always head-final. Plural and case markers are all NP-final clitics: where both are present, the plural marker precedes the case marker. Some examples follow:

- (23) *cū nāwar čhi-pi cū nōkyu aŋye thepe*
this cat bite-REL.PST this dog very big
'This very big dog that bit this cat.'

- (24) *lame thepe=cuke=e che=cuke*
 lama big=PL=GEN child=PL
 'the important lamas' children'

5.2 The structure of the clause

In the overwhelming majority of cases, the verb complex (the semantic predicate and any auxiliaries and verb particles) will come last in the clause. As for the other constituents, orienting information (locative or temporal) generally comes first, followed by the rest of the constituents arranged according to their rank on the nominal hierarchy, which is reproduced below:

Speech act participant pronouns (first and second person) > Third person pronouns >
 Personal names > Other human referents > Animate non-humans > Inanimates

In general, the higher a referent is on the hierarchy, the more topical the referent is likely to be, and the closer to the beginning of the clause it will appear. When there are two referents of equal rank, the subject will precede the object, and the object will precede obliques. Deviations from this arrangement are usually attributable to considerations like focus: topicalized items may be placed first in the clause.

Nar-Phu is rather less consistent than Chantyal in marking subjects of transitive clauses as ergative. In elicited sentences, the ergative marker is strongly preferred on the A arguments of transitive verbs. In connected speech, however, this is rare. In a corpus of 86 transitive clauses taken from connected speech from Nar speakers, only 32 clauses had overtly marked subjects, and only two of these were ergative-marked (Bond *et al.* 2013). In other words, most of the time in connected speech, A arguments are unmarked for ergativity, when they occur at all. It is observed that lexical nouns tend to correlate with switches in reference, and ergative marking (when it occurs) does so if more than one argument is realized. The matter requires further study.

Direct objects are, by default, in the absolute case, but direct objects high on the empathy hierarchy are marked as datives. This sort of dative marking (aka 'anti-agentive') is obligatory for personal pronouns, virtually so for proper nouns, but is very unusual for NPs representing inanimates. Two examples follow:

- (25) *ŋâ=se lakpe=re mraŋ-čin*
 1SG=ERG Lakpa=DAT/LOC see-PST
 'I saw Lakpa.'
- (26) *ŋâ=se hikε mraŋ-čin*
 1SG=ERG letter see-PST
 'I saw the letter.'

5.3 Copular clauses

Copular verbs are, for the most part, obligatory in clauses with non-verbal predicates, though examples with no copular verb have been encountered in simple identification clauses:

- (27) *ŋê=e mîn lakpe*
 1SG=GEN name Lakpa
 'My name is Lakpa.'

Ordinarily, though, a copular verb is present and comes last in the clause:

PREDICATE NOMINAL

- (28) *khyāŋ hâpe mûmu*
2SG novice.monk be:PRES
'You're a novice monk [I've been told].'

PREDICATE ADJECTIVE

- (29) *ŋâe khay mû*
1SG cold be:PRES
'I'm cold.'

OBLIQUE CASE-MARKED NP FUNCTIONING AS PREDICATE

- (30) *ŋâe thosor phâlpε=re mû*
1SG now Kathmandu=DAT/LOC be:PRES
'I'm in Kathmandu right now.'

5.4 Negative clauses

The negative prefix is *a-*, which assimilates to the tone of the verbal root, becoming murmured if the root has murmur. When a verb is negated, finite indicative suffixes are replaced by *-i*, as is shown in the paradigms above. Other verbal suffixes—either subordinate or non-indicative—are not replaced by *-i*:

- (31) *lama a-kha-kî*
lama NEG-come-POT
'The lama may not come.'

5.5 Interrogative clauses

Yes-no questions can be formed in several ways. Most simply, they can be formed from the corresponding statement through the substitution of an appropriate interrogative intonational contour. They can also be formed by the addition of the suffix *-pε* to the verb. This suffix usually replaces finite indicative verbal morphology (32), but sometimes is used in addition to these suffixes (33).

- (32) *tile lakpε čhæ̰ thuy-pε*
yesterday Lakpa tea drink-INTERROG
'Did Lakpa drink tea yesterday?'

- (33) *ni-ce-pε*
go-PFV-INTERROG
'Did he go?'

A construction involving positive and negative interrogative verbs is commonly encountered:

- (34) *cû kyû thuy-pε a-thuy-pε*
this water drink-INTERROG NEG-drink-INTERROG
'Did he drink water?'

Information questions are formed with an interrogative pronoun which is placed either in the usual position within the clause that a corresponding non-interrogative form would

take or in the preverbal focus position, i.e. the pronoun is not obligatorily fronted. The interrogative suffix is found in such sentences:

- (35) *čhǣ sū=se thuȳ-pe*
 tea who=ERG drink-INTERROG
 ‘Who drank the tea?’

5.6 Complement [nominal] clauses

The nominalizer suffix can be used to nominalize clauses to fill nominal slots, for example, subject:

- (36) *hlekc hlō-pe cū=i kheyaj kā̄*
 book read-NMLZR this=GEN work EMPH
 ‘Reading books is his profession.’

In addition, there is also the suffix *-ne*, referred to as the ‘potential’ nominalizer. Verbs marked with this suffix also fill nominal slots and have an interpretation that marks them as coding unfulfilled, potential events, or states:

- (37) *ŋé̄ ŋhún-ne tēn mū̄*
 1SG sleep-NMLZR desire be:PRES
 ‘I want to sleep.’

Nominalizers are used extensively in verbal periphrasis. Also of note is that there are no finite subordinate clauses except as direct quote complements of *phi-* ‘say.’

5.7 Adjectival (relative) clauses

Adjectival clauses, clauses that modify nouns, are clearly related to nominalizations. There are two relative suffixes, the present *-pe* and the past *-pi*: the present is no doubt identical to the indeterminate nominalizer *-pe*, the past is probably *-pe* plus an additional suffix, *-i*. *-i* could be identical to the *-i* suffix found in the verbal paradigm in negative clauses; it is also possible that it is a reflex of genitive suffix *-(y)e*, reflecting a regionally familiar pattern of genitive-marked nominalizations functioning as adjectival clauses. While tempting, the difficulty with the latter lies in explaining why the genitive should persist only in clauses with a past tense interpretation. Examples of relative clauses follow:

- (38) *ŋâ̄=se ſí-pi tħulthun mraŋ-čin*
 1SG=ERG die-PST.REL snake.body see-PST
 ‘I saw a dead snake.’
- (39) *mîn te-ne a-tâ-pe pħuluj*
 name call-NMLZR NEG-become-PRES.REL insect
 ‘centipede’ (‘the insect whose name isn’t called’)

5.8 Adverbial clauses

Adverbial clauses in Nar-Phu are grammatically of two types: they are either nominalizations with an appropriate case marker which provides their semantic interpretation, or they are ‘converbs,’ specialized adverbial clause types.

Case marked nominalizations and converbs are used for a number of adverbial functions, including the expression of purpose.

- (40) *yâ=se inliš hlô-te-re, ni-čin*
1SG=ERG English teach-CONV-DAT/LOC go-PST
'In order to teach English, I went.'

Converbal clauses are used to code a variety of senses, including condition:

- (41) *khi=i laki=re nhô câ-reme tano mü*
you=GEN sake=DAT/LOC garlic eat-COND good be:PRES
'If you eat garlic, it's good for you.'

And temporal sequence:

- (42) *mhyê cä-pe=re ni-se, mhyê a-yây-se, hyul=re*
cow eat-NMLZR=DAT/LOC go-CONV cow NEG-find-CONV village=DAT/LOC
ye ni-čin
return go-PST
'The cow having gone in order to eat, the cow having not been found, the man returned to the village.'

Overt conjunction of clauses is hardly used in Nar-Phu; instead, other devices, most particularly the sequential converb, are used.

FURTHER READING

- Mazaudon, Martine. 1997. An outline of the historical phonology of the dialects of Nar-Phu (Nepal). *LTBA* 19.1: 103–114.
Pohle, Perdita. 1990. *Useful Plants of Manang District: A Contribution to the Ethnobotany of the Nepal-Himalaya*. Stuttgart: Franz Steiner.

NOTE

- 1 Work done by Noonan on Nar-Phu was supported by the National Science Foundation, grant no. SBR-9600717 and some of these data are drawn from his unpublished notes. Work done by Hildebrandt was supported by a Hans Rausing Endangered Languages Documentation grant, SG-0025, and by the National Science Foundation, DEL-1149639.

REFERENCES

There is a small but growing body of publications on the Nar-Phu language. Additionally, audio-video archives of translated and transcribed texts of both Nar and Phu are available at the Endangered Languages Archive and at the University of Virginia (elar.soa.ac.uk/deposit/0103 and audio-video.shanti.virginia.edu/collection/nar-phu#, both accessed 30 November 2015).

- Bond, Oliver, and Kristine Hildebrandt. 2011. Negation across Nar discourse genres. In Peter K. Austin, Oliver Bond, David Nathan and Lutz Marten (eds), *Proceedings of the 3rd Workshop on Language Documentation & Linguistic Typology*. London, UK: School of Oriental and African Studies, 141–150.

- Bond, Oliver, Kristine Hildebrandt and Dubi Nanda Dhakal. 2013. Optional ergative case marking: What can be expressed by its absence? Presentation at the 10th

- biennial Meeting of Association for Linguistic Typology. University of Leipzig, Germany. 16 August 2013.
- Hildebrandt, Kristine. 2004. Nepali Loanwords in Three Bodish Languages. Loanword Typology Workshop, Leipzig Germany: Max Planck Institute for Evolutionary Anthropology, May 2004.
- Hildebrandt, Kristine. 2011. Converb and aspect-marking polysemy in Nar. In Elena Mihas, Bernard Perley, Gabriel Rei Doval and Kathleen Wheatley (eds), *Responses to Language Endangerment: In Honor of Mickey Noonan. New Directions in Language Documentation and Language Revitalization*. Amsterdam/Philadelphia: Benjamins, 97–117.
- Mazaudon, Martine and Alexis Michaud. 2008. Tonal contrasts and initial consonants: A case study of Tamang, a ‘missing link’ in tonogenesis. *Phonetica* 65: 231–256.

§3.4 Rung languages

§3.4.1 Rgyalrong languages

CHAPTER TWENTY-EIGHT

TSHOBUND RGYALRONG

Jackson T.-S. Sun

1 BACKGROUND

Tshobdun (aka Caodeng) is a member under the Rgyalrong language group of northwest Sichuan, which includes three other distinct languages: Situ, Japhug, and Showu (aka Zbu). Rgyalrong, in turn, is closely affiliated with two neighboring language clusters, Horpa and Lavrung, to form a Rgyalrongic unit in Sino-Tibetan (Huang 1991, 2001; Sun 2000a, 2000b; Jacques forthcoming). The Tshobdun data in this chapter, representing the speech of Kakhoris Village, are from my fieldwork conducted over the past decades.

2 PHONOLOGY

The structure of the Tshobdun syllable is (c)(c)(c)v(c)(c). The language shows a preference for complex syllable-initial clusters, while the nucleus and coda positions are normally filled by only one segment. The following system of simple initial consonants can be established (items enclosed in parentheses are marginal or non-native segments):

(1)	p	t	c	k	q
	p ^h	t ^h	c ^h	k ^h	q ^h
	b	d	ɟ	g	
	ⁿb	ⁿd	ⁿɟ	ⁿg	ⁿG
		ts	tʂ	tʃ	
		ts ^h	tʂ ^h	tʃ ^h	
		(dz)	(dʐ)	(dʒ)	
		ⁿdz	ⁿdʐ	ⁿdʒ	
	m	n	ɳ	ɳ	
	(f)	s	(ʂ)	ʃ	x
	v	z		ʒ	χ
		r		y	ʁ
		(t̪)			
	w	l	j		

An enormous number of initial clusters are attested, including many three-member clusters. Phonological evidence shows nasal + voiceless stop combinations as consonant clusters, but voiced prenasalized stops are unitary phonemes. The consonants permitted as syllable codas are -t, -v, -z, -y, -m, -n, -ŋ, -r, -l, -j. The coda -r is often devoiced; the

lateral coda *-l* is distinctly laminodental. Syllables may also take a glottal-stop coda which may combine with continuant codas, creating the only kind of cluster coda in the language. Syllables carrying a glottal coda behave phonologically like checked syllables with regard to accent placement.

Nine vowel phonemes are distinguished: *a, e, i, u, e, ε, o, ɔ, and ə*. Only two intrinsic complex vocoids *ej* and *oj* are noted in the native vocabulary. These cannot take consonant codas other than the glottal stop, and are treated herein as vowel + consonantal glide sequences.

Several types of pitch-related phenomena are noted. First, checked syllables (including those checked by a glottal stop) and non-checked ones are normally spoken in a high level and a high-falling pitch, respectively. Second, these allophonic pitch variations have taken on an ancillary tense-aspect marking function. Glottalized verb roots regularly drop the glottal coda in the past and one of the continuative verb forms (§3.4.2), resulting in a predictable shift of pitch from level (marked by a macron) to falling (marked by a circumflex). Significantly, this pitch-pattern alternation is also extended to verb roots checked by the stop coda *-t* which, unlike the glottal stop, does not undergo apocope under stem alternation.

- (2) a. *rō?* ‘fetch’
esr-rō ‘She/he is fetching.’
je-rō ‘She/he fetched.’
- b. *fkōt* ‘carve’
esr-fkōt ‘She/he is carving.’
ne-fkōt ‘She/he carved.’

Unlike in the closely related Showu Rgyalrong language where all syllable types show a two-tone contrast (level vs falling), Tshobdun syllables with *-t* normally do not carry the falling pitch barring a small number of exceptions (e.g. *tsʰēt* ‘goat’; *xtât* ‘be short’). The phenomenon in (2), therefore, represents a grammatically conditioned type of tonality alternation. Tshobdun phonology is further characterized by a functionally more substantial system of pitch accent in which a pitch drop (H-L) inside a phonological word plays a distinctive role at both lexical and morphosyntactic levels. The locus of the accented syllable, i.e. the high-pitched syllable immediately preceding the pitch drop, is highly restricted. The default (unmarked) stem-final accent is in effect the same as lack of accent, as suffixes are generally low-pitched. As a matter of fact, the only position in which a marked accent occurs in this language is the penult of the stem. Interestingly, monosyllabic words contrast in accent even though the citation pitch remains identical; the latent accent (marked here by an apostrophe) materializes only when a preceding morphological element is added, resulting in a pitch drop, e.g. *χser* ‘gold,’ *é-χser* ‘my gold;’ *rjul* ‘silver,’ *e-rjúl* ‘my silver.’ A number of segmental phonological processes are found, including vowel assimilation (e.g. *rō-”di* → *ri”di* ‘as soon as she/he rides’), vowel fusion (e.g. *mə-tə-o-”du-aŋ* → *motov”doy* ‘She/he did not beat me.’), consonant assimilation (e.g. *n-rewə* → *”dʒewə* ‘hope’; *sə-rem* → *səyrem* ‘make dry’), and syncope (e.g. *kətʃov”-sqe* → *kətʃosqe* ‘sixty’; *esr-mərkə* → *esmərkə* ‘he is stealing it.’). For productive phonological alternations at the morphosyntactic level, see §3.2.

3 MORPHOLOGY

3.1 Lexical categories

The major lexical categories in Caodeng are nominals, verbs, and ideophones; there are a number of indeclinable minor form classes also, such as adverbials, conjunctions,

interjections, classifiers, and particles. As is usual in Tibeto-Burman, attributive adjectives are nominalized stative verbs in form, e.g. *pafì?* *kə-snoj?* (crabapple NMLZ-be.deliciously.ripe) ‘deliciously ripe crabapples.’ Few true classifiers (e.g. *b'bye?* ‘classifier for any long object, such as hairs, sticks, logs, and trees’) are used; numerals in most cases directly quantify their head nouns. A range of illocutionary-force meanings are coded by various sentence-final particles, as in (3a) through (3c):

- (3) a. *kəko?* *kəru?* *fte?* *ko*
 3SG Rgyalrong be:EMPH SFP
 ‘So (I realize now that) she/he is Rgyalrong!’
- b. *kəko?* *kəru?* *fte?* *mu*
 3SG Rgyalrong be:EMPH SFP
 ‘She/he is Rgyalrong, isn’t she/he?’
- c. *kəko?* *kəru?* *fte?* *ta?*
 3SG Rgyalrong be:EMPH SFP
 ‘I am certain she/he is Rgyalrong.’

3.2 Morphological processes

A typical Caodeng word comprises a root plus a string of clearly segmentable affixes, especially prefixes.

The derivation of reciprocal and intensive process verbs involves a partially reduplicated syllable consisting of the initial of the root plus the vowel *-ə*, e.g. *(-mə)-"gə-*^h*ge* ‘call each other’ from root *"ge* ‘call’; *nə-vdá-vde* ‘become better and better’ from root *vde* ‘be good.’ Full reduplication is drawn upon to form verbal iterative aspect, e.g. *nə-spér-spér* ‘keep moving (objects)’ with prefix *nə-* and root *spér* ‘move (an object).’ Reduplication applies to lexical categories other than verbs; examples are certain temporal phrasal idioms (e.g. *k^he-syi-syi* ‘day after day’ and *syi-ku-syi* ‘everyday,’ from *syi* ‘day’), complex nominals (e.g. *kə^hdʒə(-lŋə)-lŋə?* ‘parent and child,’ from *te-lŋə?* ‘child’), as well as derived ideophones (e.g. *nə-buy-buy?* ‘in a swarming fashion’ from root *buy*; *χc^ho-və-χc^he ~ χc^ho-pe-χc^ho-le?* ‘in an unshapely or messy way’ from root *χc^hev*).)

Stem modification by means of consonantal alternation is utilized in the derivation of a few lexical causatives, e.g. *prei* ‘cause (a rope) to break’ vs *"bret* ‘(as of a rope) break.’ A notable example of productive stem alternation in inflection is past stem formation by glottality inversion, e.g. *p^hjəz?* ‘wipe [STEM 1],’ *p^hjəz* ‘wipe [STEM 2];’ *sroy* ‘guard [STEM 1],’ *sroj?* ‘guard [STEM 2].’ Equally remarkable is the phenomenon of ablaut. Playing a role in compound-noun (e.g. *k^he* ‘house,’ *k^hv-rmi?* ‘house-name’) and compound-numeral (e.g. *sqe?* ‘ten,’ *sqe-myo* ‘fifteen’) formation, ablaut is at full play in verb inflection (Sun 2004). One type of ablaut derives the second (or past) verb stem from the first (or base) stem,¹ e.g. *rp^haltʃəm* ‘be in estrus [STEM 1],’ *rp^haltʃəm?* ‘be in estrus [STEM 2];’ *mde?* ‘be level [STEM 1],’ *mde* ‘be level [STEM 2].’ Another type builds the third stem occurring in certain specific transitive contexts, such as the singular imperative, e.g. *néji* ‘wait,’ *nə-néje* ‘You [sg] wait for him/her!'; *pe* ‘do,’ *nə-pə* ‘You [sg] do it! Stems also undergo modification via accent adjustment, which assigns or removes accent depending on the specific construction. Accent modification also occurs sporadically in lexical derivation, e.g. *nənə?* ‘that,’ *nánəz* ‘there;’ *tʃor?* ‘be sour,’ *kátʃor* ‘turnip-leaf pickles.’ Systematic application of accent modification in Caodeng grammar also occurs in the formation of vocatives (§3.4.1) and polar interrogatives (examples (22) to (23)).

Compounds, the majority of which are nominal, may be composed entirely of unaffixed roots, e.g. *qnej-yve?* ‘wheat awn.’ As a rule, nominal prefixes of the initial compound components are retained, resulting in affixed compounds, e.g. *te-se-ro?* ‘hemp stalk,’ cf. *te-se?* ‘hemp,’ *te-ro?* ‘stalk.’ Compounds containing more than two components are rare, e.g. *rje-pye-lɔ?* ‘hen house,’ literally ‘Chinese-fowl-nest.’

3.3 Derivational morphology²

Generous use is made of derivational processes to form morphologically related words. With nominal roots, suffixation yields gender (male *-pʰo*, female *-mo*) and diminutive (*-pu*) forms, while denominalization is achieved by means of the verbalizing prefixes *nə-/nə-*, e.g. *tʃane* ‘loach,’ *nə-tʃane* ‘catch loaches’; *té-ye* ‘sun,’ *né-ye* ‘bask in the sun’; *tá-myna* ‘eye,’ *ná-myna* ‘aim [v.]’ Sometimes actors or instruments are also denominalized in this way, yielding transitive verbs, e.g. *rtʃaxpe* ‘robber,’ *nə-rtʃaxpe* ‘rob.’ The other major type of derived word formed by altering lexical categories is the deverbal noun. Action nominals take one of the two nominalizing prefixes, *kə-* (dynamic/human) and *kə-* (stative/non-human), e.g. *kə-gu* ‘be poor,’ *kə-ŋu* ‘become poor,’ *kə-xsor?* ‘give birth (as of bovines);’ *kə-sa-sce* ‘give birth (as of humans).’ Participant nominalization (§4.1), on the other hand, derives nouns referring to arguments of the source verbs. The prefix *kə-* produces nouns denoting actor or undergoer arguments of intransitive verbs, e.g. *kə-n"gi?* ‘patient’ (<*n"gi?* ‘ill’), *kə-tim?* (<*tim?* ‘be rich’) ‘rich person,’ as well as actors of transitive verbs, e.g. *kə-mərku* ‘thief’ (<*mərku* ‘steal’). The prefix *kə-* turns transitive verb roots into undergoer deverbal nouns ‘that which is verb-ed,’ e.g. *kə-ŋdze* ‘food’ (<*"dze* ‘eat’); *v-kə-rge* ‘one loved by me’ (<*rge* ‘love’). There is in addition a manner nominalizer *tə-*, e.g. *o-tə-rtʃone?* ‘the way she/he dances’ (<*rtʃone* ‘dance’), and also an oblique (instrument/location/time) nominalizer *sv-*, e.g. *sv-sce?* ‘birthplace; birthday’ (<*sce* ‘give birth’); *sv-ŋdzenthi* ‘dining place,’ ‘instrument with which to eat a meal’ (<*ŋdzenthi* ‘dine’), both of which are attached to past verb stems. Another important type of derivational morphology changes verb valence. Lacking a true analytical causative, Tshobdun achieves causativization predominantly by adding a number of valence-increasing causative prefixes *sə-*, *sv-*, *wə-*, or *f-*, e.g. *smi* ‘be cooked,’ *wé-smi* ‘cook [v.]’; *tʰi* ‘drink,’ *f-tʰi* ‘give to drink; suckle.’ On the other hand, a detransitivizing prefix turns a transitive verb into a corresponding intransitive, e.g. *ntʰe* ‘kill,’ *ré-ntʰe* ‘do slaughtering.’ Reflexive verb formation (involving the reflexive prefix *je-*) can also be considered a type of detransitivizing derivation, as reflexive verbs are marked morphologically as low in transitivity. Example (4) shows the low-transitivity continuous marker as well as inverse marking on the reflexive verb:

- (4) *oji?* *tʰə-o-je-ŋdu=cə*
 3SG CONT:LTR-INV-REFL-beat=MED
 ‘She/he is beating her-/himself.’

3.4 Inflectional morphology

3.4.1 Nominal inflection

Nominals are inflected for number (dual =*ni?*; plural =*re?*), case, and possession. The use of the two basic case enclitics, the ergative-instrumental =*kə* and the locative =*z*, is obligatory only when needed for disambiguation. Nominals bearing the undergoer roles

(patient, recipient, theme), furthermore, are generally not case-marked. Case-marking morphology is heterogeneous, employing enclitics, agreement prefixes, as well as stem modification by accent readjustment. A split ergative system obtains where the ergative marker *=kə* is mandatory on an actor argument only if it is outranked by the undergoer argument on an empathy hierarchy: speaker > hearer > non-participant > non-human animate > inanimate (Kuno 1976; DeLancey 1981b; Sun and Shi 2002). With an abundance of instrument-incorporating verbs such as *kə-r̥tov* ‘beat with a rod’ and *kə-ȝnov?* ‘beat with fists,’ the occurrence of case-marked instruments is also infrequent. Importantly, the presence of an instrumental argument requires the verb to be marked with a causative prefix:

- (5) *"dʒəŋyo=kə tə-sə-nə"dzétʰi*
chopsticks=INST IMP-CAUS-eat.meal
'Eat with chopsticks!'

The local case marker *=z* expresses a generalized locational meaning, occurring in locative, ablative, as well as allative contexts. More precise locative senses are conveyed through the construction possessive prefix + relator noun (+ local case), e.g. *jəye ó-ta(=z)* ‘on the book,’ literally ‘book its-top=LOC.’ The most common relator nouns are *'ta* ‘space on/over/above,’ *pʰi* ‘space under/below/beneath,’ *nay ~ nay?* ‘space inside,’ *pʰi?* ‘space outside,’ *vzər* ‘space at the side of,’ *vz(r)* ‘space on the body/vertical side of,’ *ꝝóri* ‘space in front of,’ and *qʰu* ‘space at the back of.’ Several other oblique case markers are possessed relator nouns in form, including *krv* (comitative), *χte/fcz* ‘on behalf of,’ *tənka* ‘for the sake of,’ and *'pʰa* (animate goal/source), the last of which is illustrated by these examples:

- (6) a. *sonəm=ópʰa tékʰu v-nə-tə-kʰε*
Sonam=to cigarette IRR-IMP-2-pass
'Pass the cigarette to Sonam!'

b. *sonəm=ópʰa tékʰu v-ʃə-nə-tə-roj?*
Sonam=from cigarette IRR-go-IMP-2-fetch
'Go and get the cigarette from Sonam!'

Standards in comparative constructions take the comparative case markers */ʃxtenə?* or */ʃkrenə?*, as in:

- (7) *vji? kréši /ʃxtenə? "bri?-ay=cə*
1SG Krashi than be.tall-1SG=med
'I am taller than Krashi.'

Standards of equative comparison, on the other hand, take the semblative *jermər*. The vocative case is marked suprasegmentally by accent; e.g. *džomə* (a woman’s name), *džómə* (vocative). Arguments bearing the beneficiary role are indicated by possessive prefixes on the head noun representing the object being created, moved, or manipulated for the person’s benefit, as in:

- (8) *jə-leju kə-pe tə-cʰv*
1PL:POSS-song NMLR-do 2:Q-be.able
'Can you sing songs for us?'

Possession is another major nominal inflectional category. Given two juxtaposed nouns in a possessive relation, the possessor noun is cross-referenced on the possessed noun by a head-marked possessive prefix (see Table 28.1).

TABLE 28.1 TSHOBDUN POSSESSIVE PREFIXES

	Sg	Du	Pl
1	<i>v-</i>	<i>nv-</i>	<i>o-</i>
2	<i>tsə-</i>	<i>"dzə-</i>	<i>"dzə-</i>
3	<i>jə-</i>	<i>nə-</i>	<i>nə-</i>
Generic person		<i>tə-</i>	

Pronominal possessors are usually omitted, except where emphasis is intended. An inalienable noun (including kinship terms, body-part terms, and names of certain intimate personal belongings) must drop its noun prefix before taking a possessive prefix. Contrast:

- (9) *tv̥-lo* ‘milk’
é-lo ‘my own (mother’s) milk (inalienable)’
v-tv̥-lo ‘milk owned by me (alienable)’

In another important usage, possessive prefixes represent actor or undergoer arguments in various types of syntactic structures, such as purposives (see example (27)), participant nominals (10a–b), and complex transitive predicates (10c):

- (10) a. *v-ké-j"dzu=cə* *to?*
 1SG:POSS:U-NMLZ:A-accuse=INDEF exist
 ‘There is someone who accuses me.’
- b. *v-ké-j"dzu=cə* *to?*
 1SG:POSS:A-NMLZ:U-accuse=INDEF exist
 ‘I have someone to accuse.’
- c. *nənə?* *vlerme=kə* *v-mtʰu* *tʰe-lēt*
 that monk=ERG 1SG:POSS:U-spell PFV:TR-cast₂
 ‘That monk has cast a spell on me.’

3.4.2 Verbal inflection

Verbs are highly conjugated. Verb inflectional categories include person-number, direction, orientation, transitivity, tense-aspect, evidentiality, and passive voice.

Core arguments get cross-referenced on the verb by means of person-number markers. The various person-number indexes in intransitive sentences are displayed in the following paradigm (where V = verb stem; see Table 28.2).

**TABLE 28.2 TSHOBDUN PERSON-NUMBER
 MARKERS ON INTRANSITIVE VERBS**

	1	2	3
Sg	V- <i>aŋ</i>	<i>ta-</i> V	V
Du	V- <i>tsə</i>	<i>ta-</i> V- <i>"dzə</i>	V- <i>"dzə</i>
Pl	V- <i>jə</i>	<i>ta-</i> V- <i>nə</i>	V- <i>nə</i>

The same paradigm applies to transitive verbs with inanimate undergoers. In transitive sentences involving animate undergoers, person-marking is dictated by the aforementioned empathy hierarchy. Specifically, the various argument configurations are subclassified into local ($1>2$ and $2>1$), parallel ($3>3$), and disparate ($1, 2$ interacting with 3) types. In local and parallel configurations, the argument cross-referenced on the verb is always the undergoer and actor, respectively. In disparate scenarios, however, the verb indexes the higher-ranking argument on the hierarchy, irrespective of semantic role. If and only if the undergoer participant is the speaker her-/himself, the verb may optionally carry a second index, always placed after the first person singular index, to represent the actor. In this construction, more emphasis is put on the actor, as shown by:

- (11) *kəko?*=*ni?*=*kə* *vji?*
 3=DU=ERG 1SG
tʰo-sə-wə?-*ay-**dzə* *ŋo?*
 pfv:downriver:INV-CAUS-come₂-1SG-3DU be
 ‘It was the two of them who made me come downriver.’

The category direction marks the relative place of actor and undergoer arguments on the empathy hierarchy (DeLancey 1981a). Scenarios where the actor is lower on the hierarchy than the undergoer take the inverse marker *o-*. The two local configurations also receive distinct marking: $1>2$ is represented by *ta-*; $2>1$ is marked as inverse as well as by the prefix *kə-*, interchangeable with *ta-*.

The Tshobdun person system is typologically interesting in distinguishing a fourth person, namely the generic person (Sun 2014a, 2014b). The generic-person markers, which occur in gnomic, generic sentences, comprise a paradigm with the other inflectional person prefixes in the same morpheme slot. The two generic-person indexes *ka-* and *kə-*, representing respectively *volitional* and *non-volitional* generic human subjects, are illustrated below.³

- (12) *ko?* *tewa?* *kə-ldeñ* *v-nə-ka?*-*tʰi=nə?*
 this liquor NMLZ-be.much IRR-PFV-GP;VOL-drink=DET
ʃla *kə-lde?* *ʃte?*
 quickly GP:NONVOL-be.drunk be:EMPH
 ‘If one drinks too much of this liquor, one quickly gets drunk.’

Orientation, or spatial grounding, is a salient trait in Rgyalrongic grammar. Three distinct subsystems are at work, each of which comprises two opposing terms: vertical (up–down), riverine (upriver–downriver), and solar (east–west). By virtue of metaphorical extension, the riverine and solar subsystems have also acquired an inside–outside and centripetal–centrifugal opposition, respectively. In addition to orientational adverbials and pronouns, there is a whole array of verbal orientation prefixes, a basic set of which is displayed in Table 28.3.

Required on all perfective and imperative verbs, the same forms of orientation prefixes do triple duty, coding aspect, imperativity, as well as spatial orientation *per se*. With non-motion verbs, the selection of collocating orientation prefixes is not semantically transparent, but a matter of convention (e.g. ‘downward’ with the verb ‘eat,’ but ‘westward’ with the verb ‘drink’).

The verb is highly sensitive to transitivity. Apart from ablaut, additional morphological devices are available for indicating transitivity. In perfective sentences with a third person actor, transitivity is marked by shifting the vocalism (*e* or *a*) in the orientation prefixes uniformly to *e*. In direct imperfective sentences with a singular non-first person actor, moreover, a transitivity marker *-ja* is added to verb stems containing no coda other than

TABLE 28.3 BASIC TSHOBDUN ORIENTATION PREFIXES

up	down	upriver	downriver	eastward	westward
<i>tɔ-</i>	<i>nv-</i>	<i>lv-</i>	<i>t^hv-</i>	<i>kɔ-</i>	<i>nɔ-</i>

the glottal stop (e.g. *jɔ-tɔ-q^hvɔ-jɔ* ‘You will dislike it.’); for ablauting verbs such as *t^hi* ‘drink,’ this transitive marking is optional (e.g. *jɔ-tɔ-t^he(-jɔ)* ‘You will drink it.’). For yet another transitivity-related distinction, see (14).

Tshobdun has grammaticalized absolute tense, distinguishing two non-past stems (STEM 1 and also STEM 3 for certain transitive verbs) and a past stem (STEM 2) in all verbs. The unmarked tense is the non-past which, as in many languages, has a wide range of uses, including gnomic, habitual, as well as future. There is also a relative future or *prospective* prefix *jɔ-*, which is not bound to the moment of speech but is also appropriate for indicating ‘future in the past’ and, with the past stem, the meaning ‘almost.’ The basic aspectual contrast, on the other hand, is between imperfective and perfective, marked by different orientation/aspectual prefixes. The interaction between tense and aspect yields two past tenses built on STEM 2: an aorist (perfective past) denoting events viewed as a whole and processes or resultant states thereof, and an imperfect (past imperfective) denoting ongoing situations existing prior to the moment of speech, marked with *nv/ne-*. The latter is opposed to the present imperfective, marked by adding to a non-past stem an orientation/aspectual prefix with a uniform vocalism *-e* (see example (15)). The contrast between the aorist and the imperfect is shown in (13a–b):

- (13) a. *for?* *ko?* *tɔ-n^hgi=cɔ*
 yesterday 3SG PFV-be.ill₂=MED
 ‘She/he became ill yesterday (AORIST).’
- b. *for?* *ko?* *nv-n^hgi=cɔ*
 yesterday 3SG IPFV:PST-be.ill₂=MED
 ‘She/he was ill yesterday (IMPERFECT).’

The aorist form can also co-occur with temporal adverbials referring to the present; in other words, it can be used as a perfect, which relates completed situations to the moment of speech. The past imperfective is also pressed into service to denote another type of perfect, namely *perfect of persistent situations* (e.g. *nv-t^hi-ay* ‘I have been drinking’). Several other aspectual categories figure in verb inflection, including two continuatives and a habitual. The continuatives are differentiated by relative transitivity: the high-transitivity continuative prefix *esv-* is attached to a variant of STEM 1, whereas the low-transitivity continuative prefix *t^hv-* requires STEM 2. Consider:

- (14) a. *stat^her* *leju* *esv-pe/*t^hv-pe?*
 Stathar song CONT:HTR-do
 ‘Stathar is singing songs.’
- b. *stat^her* *t^hv-rleju?*/**esv-rleju*
 Stathar CONT:LTR-sing₂
 ‘Stathar is singing.’

Events which recur with regularity can be expressed by the bare non-past verb or by adding a present imperfective prefix. This marked habitual form also applies to recurring events in the past:

- (15) *syjonts^hez kréfi=kə ts^hemgon "ge-mte? =nə? te-remúru=cə*
formerly Krashi=ERG Tshemgon IPFV-see₃=DET IPFV-get.angry=MED
'Krashi used to get angry whenever he saw Tshemgon.'

An experiential meaning is conveyed by the complement-taking verb *rni?* 'taste; experience.'

The Tshobdun evidential system consists basically of a mediative *=cə* and a hearsay *=tétsə*. Illocutionary force is expressed partly through verb inflection and partly through sentence-final particles. Examples of inflectionally coded illocutionary forces are polar questions formed via accent adjustment, distal imperatives ((6); (27)), and jussives and optatives formed via the irrealis prefix *v-* (Sun 2007).

In addition, Tshobdun distinguishes a type of *agentless* passive verb form, made up of the prefix *v-* attached to STEM 1. This formation predicates of a past transitive event with an unspecified agent, from the perspective of the affected patient. A pertinent example is (16), uttered by a person who walks into a household shrine and sees a Buddha figurine draped in uncanny attire:

- (16) *ko? sku tʃʰə v-ʒge yo?*
this buddha.figurine what PASS-put.on be
'What got put on this Buddha figurine?'

4 SYNTAX

4.1 Syntactic relations

Neither nominal case marking nor argument cross-referencing, which operate on semantic and pragmatic principles, can give reliable clues regarding syntactic relations in Tshobdun. Participant nominalization (§3.3), however, provides evidence for the neutralization of the actor and undergoer roles in an accusative pattern; specifically, the single intransitive argument (s) and the transitive actor argument (a) are marked by the same nominalizing prefix *kə-*, as opposed to the transitive undergoer argument (o) marked by a different nominalizing prefix *kr-*. Furthermore, the existence of a privileged syntactic relation subject is borne out by the fact that this is the only NP position relativizable by finite nominalization (§4.5.2).

4.2 Constituent order

The sequence of morpheme slots in any verb form is: irrealis + negator + go and + low-transitivity continuative + orientation/aspect/imperative + nominalizer + direction + inverse + reflexive + high-transitivity continuative + spontaneous + stem + person-number + evidential, whereas the morpheme sequence in a nominal form is: possessive prefix + nominal prefix/nominalizer + stem + number + case. The normal word order inside a nominal phrase is: demonstrative + possessor + head noun + adjective (formally a relativized stative verb) + numeral + classifier + determiner. The adjective can also precede the head noun, denoting a partitive meaning. When a relative clause is added, it must follow the demonstrative; the relative verb, however, can also occur disjunctively after the adjective or even before the determiner at the end of the nominal phrase.

Within a clause, the verb occupies final position. The word order of core arguments is pragmatically determined, with the most topical element appearing first. Temporal adverbials normally precede locational ones.

4.3 Major sentence types

The indicative sentence is the unmarked sentence type. Fully conjugated equational copulas (positive *yo?*, negative *ma?*, emphatic/contrastive *sie?*) occur obligatorily in equative clauses:

- (17) *c^he?=z jeni? vverge ne-yo-ⁿdzo*
former.time=LOC 3DU monk IPFV:PST-be₂-3DU
‘The two of them used to be monks.’

Copulas are also employed to form presentational sentences (i.e. those with sentence-focus) in which, not surprisingly, copulas show tense-aspect but not person-number distinctions.

The main existential verbs are *to?* (STEM 2 *to*) and its negative form *me?* (STEM 2 *me?*). In addition to predicationg the existence of entities, they express possession via the construction (possessor noun) + possessor prefix – possessed noun + existential verb, as in:

- (18) *neji? ne-lja? á-to?*
2SG 2SG:POSS-child Q-exist
‘Do you have children?’

A separate verb *smo* must be used with animate subjects to express the locational meaning ‘be present at a certain place’:

- (19) *sonam pə/k^ho k^hénay t^hv-smo*
Sonam now home CONT:LTR-be.present
‘Sonam is at home now.’

However, Tshobdu does not differentiate its existential verbs further for the purpose of nominal sub-classification, a salient morphosyntactic trait found in many Qiangic and Burmic languages in the area.

Topic–comment sentences are prevalent. The most widely used topicalizing marker in this language is *na?*:

- (20) *na-nə? o-let^he=na? fə-kə-ro-na?*
that 3SG:POSS-object=TOP go.and-IMP-fetch-2PL
‘That stuff of his, go and fetch it!’

There are three other topicalizers, *rcənə*, “der, and *yónə*, which are used to express emphasis or contrast.

Negation is expressed by prefical negators, of which there are three: *me-* (imperfective), *ma-* (perfective and prohibitive), and *me-* (high-transitivity continuative, habitual, and non-finite). The copula *yo?* and existential verb *to?* have distinct negative forms, respectively *ma?* and *me?*. The latter also serves as a propositional negator (i.e. ‘it is not true at all that’):

- (21) *kəko? rervv? "bri? me=cə*
this mountain be.high exist:NEG=MED
‘This mountain is not high at all!’

The following types of interrogative sentences have been recorded: content questions, polar questions, alternative questions, and questions with positive/negative presupposition. Content questions contain various question words, which are not fronted. Unlike in the better-known Situ dialect (Lin 1993: 391–2), pragmatically neutral polar questions are formed out of the corresponding statements simply by marked accent on the verb:

- (22) *for?* *sonam* *tewa?* *né-t^{hi}*
 yesterday Sonam liquor PFV:Q-drink₂
 ‘Did Sonam drink liquor yesterday?’

Polar questions with sentence-focus take an accented copula with an interrogative prefix *ə-*:

- (23) *for?* *sonam* *tewa?* *ne-t^{hi}* *ə-yo*
 yesterday Sonam liquor PFV-drink₂ Q-be
 ‘Is it true that Sonam drank liquor yesterday?’

Alternative questions show an interrogative particle *so?* placed after the first clause, in addition to a particular intonation pattern. Questions carrying positive or negative presuppositions are formed by adding various sentence-final particles. An imperative verb form consists of a non-past stem and an orientation prefix with an imperative-marking function and, where appropriate, a second person number suffix (dual *-dzə*; plural *-nə*). The negative imperative (prohibitive) takes the imperative negator *mə-*. The prefix *v-* marks imperatives presented for action by a third person, e.g. *v-tə-”dze* ‘Let him/her eat it!’, as well as distal imperatives where the action is not to be executed immediately at the speech-act location; for an example see (27). The second person index *tə-*, as well as the 2 > 1 configuration marker *kə-o-*, show up only in prohibitive and distal imperatives. Hortatives are formed out of non-past stems with inclusive first person dual *-tsə* or plural *-jə* suffixes.

4.4 Clause coordination

Tshobdun makes no use of coordinating conjunctions at the clausal level. Instead, one finds sequences of finite clauses loosely connected by the sequence linker *q^ho?* or *q^hónə*:

- (24) *kəko?* *qa* *te-”di* *q^ho?*
 3SG hoe PFV:up-carry₂ LINK
kom? *te-cu* *q^ho?*
 door PFV-open₂ LINK
təjtse-nay *kə-reme* *jv-”sə?*
 field-in PURP-do.labor PFV-go₂
 ‘He picked up the hoe, opened the door, and went off to do labor in the field.’

Also recorded are serial verbs, which denote simultaneous verbal events (25) or those in immediate temporal sequence (26):

- (25) *pa* *ŋemgon* *χsəm* *kə-nl̩sa* *kə-nl̩s?* *c^hv*
 pig river three NMLZ-swim NMLZ-cross be.able
 ‘Pigs are able to swim across three rivers (at a stretch).’
- (26) *o-kórnə* *tə-zγi* *nv-vldv=cə*
 3SG:POSS-brain PFV-spin₂ PFV-fall.down₂=MED
 ‘His head spun and he fell down.’

4.5 Clause subordination

4.5.1 Non-finite subordinate clauses

The two nominalizing prefixes *kə-* (dynamic/human) and *kə-* (stative/non-human) yield non-finite nominalized clauses which function as sentential subjects or

complements to modal and other complement-taking verbs. The following example exemplifies the purposive use of the nominalized clause in *kə-*, serving as an adjunct to a motion verb:

- (27) *v-kə-qʰro* *v-jə-tə-wi*
 1SG-PURP-welcome IRR-IMP-2-come
 'Come and welcome me.'

There is also the converb (verbal adverb), composed of the prefix *sv-* and a reduplicated STEM 2, which modifies the situation predicated by the main verb:

- (28) *ó-ⁿge* *sv-nəc^hi^{ci}?* *noofsféfsət* *te-ⁿge?*^{=cə}
 3SG:POSS-clothes CVB-wet₂:REDUP in.that.way PFV-put.on₂=MED
 'She/he put on her/his clothes while they were still wet.'

4.5.2 Finite subordinate clauses

Finite dependent clauses function as arguments (complement clauses),⁴ argument modifiers (relative clauses), or sentential adjuncts (adverbial clauses).

Dependent clauses can occur in non-nominalized, fully finite form taking a determiner =*nə?*, as is the case of the complement clause below serving as an object of the cognition verb 'to know':

- (29) */yvez=kə* *kʰeze?* *o-nne?* *te-ⁿdze?*^{=nə?} *vji?* *'siz-aŋ*
 hog.badger=ERG dog 3sg:POSS-nose PFV-bite₂=DET 1SG know-1SG
 'I know that the hog badger bit the dog on the nose.'

Relative clauses are built on nominalized clauses. The most common relativizing strategy is head-marked gapping. The following example exemplifies a head-internal relative clause, where the relative head *támə* 'woman' appears in the relative clause, leaving a gap in the main clause:

- (30) *[for?* *támə* *nə-mti-aŋ]* *Ø=nə?* *jə-wə?*^{=cə}
 yesterday woman PFV-see₂-1SG =DET PFV-come₂=MED
 'The woman I saw yesterday has come.'

Dependent-marked gapping is also permitted, as shown in the following syntactic alternative to (30):

- (31) *[for?* Ø *nə-mti-aŋ]* *támə=nə?* *jə-wə?*^{=cə}
 yesterday PFV-see₂-1SG woman=DET PFV-come₂=MED

Relativization by means of finite nominalization is accessible only to the syntactic pivot (subject) of a clause. In this construction, the two nominalizers *kə-* and *kv-* mark distinct semantic roles of the relative-clause subject, with the latter denoting only derived patientive subjects. The subject of the following (headless) relative clause bears the actor role, hence the nominalizer *kə-*:

- (32) *[paʃi?tʃətʃen* *kə-və-pʰətʃ]* *o-vzər=nə?* *lv-zjōt-nə*
 pear NMLZ:Sbj-CONT:HTR-pick 3SG:POSS-SIDE=DET PFV:upriver-arrive₂-3PL
 'They arrived upriver to the side of the one who was picking pears.'

On the other hand, the nominalizer in the example below must take the form *kv-* (fused from *kə-* and agentless passive prefix *v-*), as the subject of the relative clause 'child' is the semantic undergoer of the verb *rqɔ?* 'to embrace':

-
- (33) [t₁elj₂a? t₂-k₂-r₂q₂]=n₂? n₂-n₂re₂?
 child PFV:up-NMLZ:SBJ:U-embrace₂=DET PFV-laugh₂
 ‘The child that got embraced laughed.’

Transitive events predicated from the viewpoint of an undergoer argument, however, may be expressed as self-caused ones, in which case the undergoer is promoted to pivot status and thus relativizable by the nominalizer *k₂-*:

- (34) [ó-pe k₂-k₂-o-o-ja-s₂-r₂q₂] t₁elj₂a?=n₂? n₂-n₂re₂?
 3SG:POSS-father PFV-NMLZ:SBJ-INV-REFL-CAUS-embrace₂ child=DET PFV-laugh₂
 ‘The child who got himself embraced by his father laughed.’

Adverbial clauses expressing temporal relations are marked by morphemes showing various temporal meanings in combination with the determiner *n₂?*, e.g. *orjajn₂?z*) (literally: *o-rjaj?₂-n₂?=z* 3SG:POSS-time-that-LOC) ‘when,’ *ʃɔxten₂* ‘before,’ *oq^hun₂?* ‘after,’ *oʒor₂* ‘while.’ The meaning ‘as soon as’ is however coded in a special way by an accented non-past stem plus the prefix *ro-*:

- (35) pye o-kér^hdzem ró-qet t^hv-mel^hbjəm
 bird 3SG:POSS-wing as.soon.as-stretch.open PFV:downriver-fly₂
 ‘The bird flew away downriver as soon as it opened its wings.’

Conditional clauses are marked by the irrealis prefix *v-* in combination with a perfective prefix attached to the non-past stem, in addition to the determiner *=n₂?*. Alternatively, the conditional form of the copula *v-né-yo* may follow the verb:

- (36) a. osto v-né-t₂-vzjaj=n₂? kac^hi? te-^hbi?
 hard IRR-PFV-2-study=DET candy 1>2-give
 b. osto t₂-vzjaj v-né-yo=n₂? kac^hi? te-^hbi?
 hard 2-study IRR-PFV-be=DET candy 1>2-give
 ‘I will give you candy if you study hard.’

The instrumental-ergative *=k₂* expresses a loose logical connection between two clauses:

- (37) a. kənm̥rémsu=k₂ pén̥tjaj ó-ta ke-mdzu=k₂ ne-c^hov
 fat.person=ERG chair 3SG:POSS-top NMLZ-sit=INST PFV:TR-break₂
 ‘The fat person broke the chair by sitting on it.’
- b. jle o-k₂-ni?=n₂? tʃu? mtʃoldən te-ké-tsə
 hybrid.cattle 3S:POSS-NMLZ:SBJ-drive=DET PN PN IPFV-GP-say
 tərm̥v=c₂ to?=k₂
 man=INDEF exist=INST
 nənə?=k₂ ke-né? nə-yo
 ANA=ERG IPFV:EAST-drive₃ IPFV:PST-be₂
 ‘The one who drove the hybrid cattle was a Zhulin man named Mcholdan, and he was driving (it) from behind.’

Cause and purpose clauses can both take the complex markers *otánk₂*, while concession clauses are marked with *ntʃon(n₂)*.

ADDITIONAL ABBREVIATIONS

GP	generic person
HTR	high-transitivity

INV	inverse
IRR	irrealis
LTR	low-transitivity
MED	mediative
NONVOL	non-volitional
SFP	sentence-final particle
VOL	volitional

NOTES

- 1 The derived (i.e. second and third) stems are indicated with subscript numerals.
- 2 For a thorough treatment of derivational morphology in Tshobdun and other Rgyalrong languages, see Sun (2014a), Xiang (2008: §3).
- 3 A third GP prefix *sr-* occurs with copular verbs.
- 4 Full details of complement clauses and complementizing strategies in Tshobdun Rgyalrong are provided in Sun (2012).

REFERENCES

- DeLancey, Scott (1981a) ‘The category of direction in Tibeto-Burman’, *LTBA* 6.1: 83–101.
 —— (1981b) ‘An interpretation of split ergativity and related patterns’, *Language* 57: 626–657.
- Huang, Bufan (1991) ‘Qiangyuzhi (The Qiangic branch)’, in Ma Xueliang (ed.) *Hanzangyu Gailun (Introduction to Sino-Tibetan Languages)*, Beijing: Beijing University Press, 208–369.
 —— (2001) ‘Guanyinqiaohua yushu wenti yanjiu’ (A study on the genetic affiliation of the Guanyinqiao language), *Language and Linguistics* 2.1: 69–92.
- Jacques, Guillaume (forthcoming) ‘Rgyalrong’, in Rint Sybesma, Wolfgang Behr, Gu Yueguo, Zev Handel, C.-T. James Huang and James Myers (eds) *The Encyclopedia of Chinese Languages and Linguistics*, Leiden: Brill.
- Kuno, Susumo (1976) ‘Subject, theme, and the speaker’s empathy’, in Charles N. Li (ed.) *Subject and Topic*, New York, San Francisco, London: Academic Press, 417–444.
- Lin, Xiangrong (1993) *Jiarongyu Yanjiu (A Grammar of Rgyalrong)*, Chengdu: Sichuan Nationality Press.
- Sun, Jackson T.-S. (2000a) ‘Parallelisms in the verb morphology of Sidaba rGyalrong and Lavrung in rGyalrongic’, *Language and Linguistics* 1.1: 161–190.
 —— (2000b) ‘Stem alternations in Puxi verb inflection: toward validating the rGyalrongic subgroup in Qiangic’, *Language and Linguistics* 1.2: 211–232.
 —— (2004) ‘Verb-stem variations in Showu rGyalrong’, in Ying-chin Lin, Fang-min Hsu, Chun-chih Lee, Jackson T.-S. Sun, Hsiu-fang Yang and Dah-an Ho (eds) *Studies on Sino-Tibetan Languages: Papers in Honor of Professor Hwang-cherng Gong on his Seventieth Birthday (Language and Linguistics Monograph Series W-4)*, Taipei: Institute of Linguistics, Academia Sinica, 269–296.
 —— (2007) ‘The irrealis category in rGyalrong’, *Language and Linguistics* 8.3: 797–819.
 —— (2012) ‘Complementation in Caodeng rGyalrong’, *Language and Linguistics* 13.3: 471–498.
 —— (2014a) ‘Typology of generic-person marking in Tshobdun Rgyalrong’, in Richard VanNess Simmons and Newell Ann Van Auken (eds) *Festschrift to Honor South Coblin: Studies in Chinese and Sino-Tibetan Linguistics: Dialect, Phonology, Transcription and Text*, Taipei: Institute of Linguistics, Academia Sinica, 225–248.

- (2014b) ‘Sino-Tibetan: Rgyalrong’, in Rochelle Lieber and Pavol Štekauer (eds) *The Oxford Handbook of Derivational Morphology*, Oxford: Oxford University Press, 630–650.
- and Shi, Danluo (2002) ‘Caodeng jiarongyu yu rentongdengdi youguan de ruogan yufa xianxiang’ (Empathy hierarchy in some areas of Caodeng rGyalrong grammar), *Language and Linguistics* 3.1: 79–99.
- Xiang, Bolin (Guillaume Jacques) (2008) *Jiarongyu Yanjiu (A Grammar of Rgyalrong)*, Beijing: Nationality Press.

CHAPTER TWENTY-NINE

COGTSE RGYALRONG

Yasuhiko Nagano

1 INTRODUCTION

rGyalrong (*rGyal rong* in Written Tibetan (WT); 嘉戎 in Chinese) is a Tibeto-Burman (TB) language spoken in the northwestern part of Sichuan Province, China. This language has long attracted the attention of scholars, because of the striking similarity of some of its lexical items to those of WT as well as its complicated system of affixation, which could be regarded as reflexes of Proto-TB (PTB) morphology.

rGyalrong has numerous words similar or sometimes identical to WT, but almost all of those are cultural words likely to be loans. In contrast, the basic lexical items such as body parts, adjectives, and verb roots are more comparable to Benedict's PTB forms. Third, rGyalrong should not be sub-classified on the basis of vocabulary with the Tibetan branch.

As for affixation, its structure is reminiscent of Jinghpaw, Rawang and PTB, while its pronominalization system is somewhat parallel to certain Himalayish languages. However, we take no position about its genetic relationship to specific TB branches yet. As far as the very sophisticated system of pronominalization is concerned, it does not look like a direct reflex of the PTB system but instead is probably a later development.

1.1 Distribution

The majority of rGyalrong people inhabit the Aba (=阿坝: WT *rn̥ga ba*) Tibetan and Qiang Autonomous Prefecture and the Gantse (=甘孜: WT *dkar mdzes*) Tibetan Autonomous Prefecture of Sichuan Province. The exact population is unknown since they are not officially recognized as an independent nationality in China but are categorized as being of the Tibetan nationality. But the number of native speakers of rGyalrong is estimated to be at least 150,000. Unlike Tibetan, their native language is not used in education in this region, and consequently, bilingualism with rGyalrong at home and Chinese in public places has become rather dominant.

1.2 Dialects

It is traditionally said that there are 18 dialects. This non-linguistic classification came from the historical division of the area into 18 administrative zones in the twelfth century, with a separate dialect spoken in each zone. On the basis of modern linguistic data, however, Lin claims that the language has three dialects; northern, eastern and western (Lin 1993: 411–14). Datsang (WT *da tshang*; 大藏) is in the centre of the northern area where about 10,000 native speakers are found. The western area includes the Dzatang (WT *'dzam thang*; 壤塘) and Tanpa (WT *bstan pa, dam pa*; 丹巴) areas, in which there are 50,000 people. The eastern area covers a relatively vast region including Cogtse (WT *lcog rtse*; 卓克基), Barkham (WT *'bar khams*; 馬尔康), Suomo (WT *so mang*; 梭磨),

Tshakunao (= Tshako, WT *bkra shis gling*; 雜谷脑), Lishan (WT *lis rdzong*; 理县), Shaojin (WT *btsan lha*; 小金川), Jinchuan (WT *rab brtan*; 大金川), Heishui (WT *khro chu rdzong*; 黑水) and Mawo (WT *bha dbo*, 麻窝). There are around 80,000 rGyalrong people in this area.

These dialects are classified according to the initial consonant clusters and the behaviour of pronominals in the predicate. However, not all the descriptions of dialects are complete, so this classification remains incomplete and thus tentative. In contrast, J. Sun (2000) has posited a different sub-classification of *rGyalrongsic* dialects – West, North and East. And, recently, Marielle Prins divided rGyalrong proper into three sub-groups – Northern, Central and Southern – directing her attention to the exclusive existence of uvular consonants in the Northern group and of tonality in the Southern group (Prins 2011: 20–4).

Among these dialects, the Cogtse (WT *lcog rtse*; 卓克基) dialect conservatively keeps a set of affixes and is considered the standard. The following description is based on that dialect.

2 OUTLINE OF PHONOLOGY

2.1 Consonant phonemes

	labial	alveolar	post-alveolar	retroflex	palatal	velar	glottal
plosive	p p ^h b	t t ^h d		t̪ t̪ ^h d̪		k k ^h g	?
affricate		ts ts ^h dz	tʃ tʃ ^h dʒ		tc tc ^h dz		
fricative		s z	fʒ				h
nasal	m	n			n̪	n̪	
tap		r					
approximant	w	l		j			

In addition to the above, there is phonemic nasalization, /N/, which occurs before stops and affricates. If it occurs at the C_f position, it nasalizes the preceding vowel.

[c, c^h, j] are free variations of /tɕ, tɕ^h, dz/. /r/ usually appears as a tap [ɾ].

/ʃ/ has two allophones, [ʂ] before plosives and [f] before affricates.

[Kj], [f] and [h] may appear in the loanwords from Tibetan or Chinese.

2.2 Vowels

The vowel phonemes are /a, i, u, e, o, ə, ə̄/. /i/ usually appears as [ɪ] and /e/ as [ɛ].

2.3 Tone

Tone is not distinctive. Several minimal pairs have been reported, but when they were checked with my elder informant, who passed away at the age of 85 in 1986, they were homophones in his speech, which formed the basis for this study. With my younger informant, however, several pitch patterns were observed. I checked the minimal pairs that Lin (1989) and Dai and Yanmucha (1990) listed and found that the following minimal pairs appear to be distinctive:

- | | | | |
|--|------------|-------------------------------------|------------------------|
| (1) kə ²² tʃor ⁴⁴ | ‘narrow’ | kə ²² tʃor ⁵² | ‘sour’ |
| ka ²³ rma ⁵⁵ | ‘to sleep’ | ka ²² rma ⁴² | ‘white eared pheasant’ |

<i>kə²³</i> jam ⁵⁵	‘sun’	<i>kə²³</i> jam ⁵³	‘broad’
<i>te²³</i> ro ⁵⁵	‘head of clan’	<i>ta²</i> ro? ⁴³	‘breast’
<i>ta²³</i> wu ⁵⁵	‘felt carpet’	<i>te²³</i> wu ⁴⁴	‘grandfather’
<i>ta²³</i> mŋam ⁵⁵	‘deaf person’	<i>te²³</i> mŋam ⁵²	‘pain’

Among these examples, falling pitch is distinctive in the first three pairs but not in the other pairs. Further, the second pair, *ka² rma⁴²* ‘white eared pheasant’ is probably a loan from WT *rma bya* ‘peacock’. We must await further data and analysis to see whether this phenomenon is an instance of incipient tonogenesis in rGyalrong.

2.4 Syllable structure

The syllable structure is (c)c_i(G)v(c_f)(s), where the parenthesized portions are optional. (c) can be occupied by p-, t-, k-, r-, l-, s-, ſ-, m- or n-. Of these, p-, t-, k-, s- and ſ- are voiced when followed by a voiced c_i. All the consonants shown in 2.1 can occur at c_i. (G) stands for glides, -r-, -l-, -w- and -j-. C_f is either -p-, -t-, -k-, -?, -tʃ-, -s-, -m-, -n-, -ŋ-, -l-, -r-, -w-, -j- or -n-. The syllables that begin with a glottal stop plus a vowel are treated as ?v here.

3 MORPHOLOGY AND MORPHOSYNTAX

The following description focuses on the most characteristic features of rGyalrong morphology and morphosyntax, the productive affixes, which mark a variety of grammatical functions.

3.1 Nouns

The majority of nouns are marked by the prefix *tə-*, *ta-*, or *te-*, glossed in the examples as NFP ‘noun-forming prefix’: for example, *tə-rmi* ‘man’, *te-jək* ‘hand’. This prefix also appears before classifiers, indicating a unit. Thus, *tə-pa* ‘one year’, *tə-rgi* ‘one (unit) of’, *tə-lpek* ‘a piece of’. *tə-rgi* is used when the number of ‘one’ must be emphasized, such as in *pak tə-rgi* ‘one pig’, while *tə-pak* ‘a pig’ is non-emphatic. For example, ‘three pigs’ is *tə-pak kə-sam*, where *sam* ‘three’ takes the adjectival prefix *kə-*.

3.1.1 Syntactic order

The syntactic order of noun and noun-qualifier(s) is as follows:

(2)	this umbrella	<i>ʃtə wə-dek</i> (this 3SGPOSS-umbrella)
	two rooms	<i>kʰo kə-nes</i> (room two)
	these four pens	<i>ʃtə wə-spəwə kə-wdi</i> (this 3SGPOSS-pen four)
	a big tent	<i>sgar (kə-)kte</i> (tent big)
	this tasty gruel	<i>ʃtə wə-pepe kə-mem</i> (this 3SGPOSS-gruel tasty)
	three pretty girls	<i>tə-mi kə-snaya kə-sam</i> (woman pretty three)
	these three black pencils	<i>ʃtə wə-ʒasnu kə-nek kə-sam</i> (this 3SGPOSS-pencil black three)

3.1.2 Nominalizers

The *?a-* prefix nominalizes direction markers: *?a-ta* ‘the above; the upper place’, versus *ta-* ‘upward’, *?a-na* ‘down(n), the lower place’ versus *na-* ‘downward’. *?a-* also implies

proximity to the speaker; in contrast, *ha-* nominalizes direction markers distant from the speaker, thus, *ha-ta* ‘an upper place over there’.

The prefix *to-* added to an infinitive verb marks a patient; thus, *to-ka-ʒu* is an accused person versus *ka-ʒu* ‘to accuse’.

The nominalizer *sv-* indicates a place to do something or a utensil to do it with. For example, *sv-dzup* ‘bedroom’ vs *ka-dzup* ‘to sleep’ and *sv-top* ‘a hammer’ vs *ka-top* ‘to hit’.

3.1.3 The gender/number markers

The (natural) gender markings on nouns are *-pʰo* (male) and *-mo* (female). The number markings on all nouns are *-ndʒas* (dual) and *-ne* (plural).

3.1.4 Relative clauses

The general structure of relative clauses is a non-final verb complex + *wə*-noun, where the non-final verb complex consists of (aspect marker/directive)-*v_{inf}*-pronominal suffix. See §§3.4.2 and 3.4.3 on the aspect markers and directives. Pronominal affixes are described under §3.4.5. *v_{inf}* is a verb root prefixed by *ka-*. Thus,

- (3) *ta-pu Ø-ka-ndza ma nə-rga-w wə-dza*
child Ø-INF-eat NEG APPLIC-like-3SG 3SGPOSS-food
'the food that a child does not like to take.'
- (4) *miser pʰendzokʰan wu-ŋuj to-ka-nə-tfara-ŋ wə-tʰa*
yesterday library 3SGPOSS-in PFV(up)-INF- APPLIC-read-1SG 3SGPOSS-book
'the book that I read in the library yesterday.'
- (5) *miser ta-ki-ŋ wə-tʰa tə-na-pfi-ŋ.*
yesterday PFV(up)-buy-1SG 3SGPOSS-book NFP-PFV(down)-lose-1SG
'I lost the book that I bought yesterday.'

3.2 Pronouns

3.2.1 Personal pronouns

The independent personal pronouns are as listed in §3.2.1.1. The possessive prefixes, given in §3.2.1.2, and the pronominal affixes in verb phrases, given in §3.4.2, are reduced forms of the free pronouns.

Among these, only older people use *jo* for first person plural, *no* for second plural, and *mə* for third singular, while *nə-ŋa* is sometimes used for first person dual inclusive, which seems to be a newer formation, ‘you and I’.

3.2.1.1 Independent personal pronouns

When it is necessary to emphasize the concept of possession or distinguish particular possessors (in the cases of duals and plurals), the independent personal pronouns are added before nouns affixed with the possessive prefix (Figure 29.1).

	Singular	Dual	Plural
First	<i>ŋa</i>	<i>tʃʰi-dzo</i>	<i>ji-dzo</i>
		<i>ji-ndʒo</i> (exclusive)	<i>ji-no</i> (exclusive)
			<i>jo</i>
Second	<i>nə-dzo</i>	<i>dʒi-dzo</i>	<i>ju-dzo</i>
	<i>nə-jo</i> (honorifics)		<i>juo</i>
Third	<i>wu-jo</i>	<i>wu-jo-dʒis</i>	<i>wu-dzo-ne</i>
	<i>wi-jo-ne</i> (honorifics)		<i>wu-jo-ne</i>
	<i>mə</i>		<i>wi-jo-ne</i>

FIGURE 29.1 INDEPENDENT PERSONAL PRONOUNS

Thus,

(6) <i>ta-pe</i>	father
<i>ŋa tʃʰe-pe</i>	my father
<i>na nə-pe</i>	your father
<i>wu-jo wə-pe</i>	his/her father
<i>ji-dzo je-pe, jo jə-pe</i>	our father
<i>ji-dzo je-pe, jo jə-pe</i>	your father
<i>wu-jo-ne jə-pe</i>	their father
<i>tʃʰi-dzo ndʒə-pe</i>	the father of us two
<i>dʒi-dzo ndʒə-pe</i>	the father of you two
<i>wu-jo-dʒis ndʒə-pe</i>	the father of them two

3.2.1.2 Possessive prefixes:

(7) <i>tə-mo</i>	mother
<i>ŋe-mo</i>	my mother
<i>nə-mo</i>	your mother
<i>wə-mo</i>	his/her mother
<i>jə-mo</i>	our mother (exclusive and inclusive)
<i>ju-mo</i>	your/their mother
<i>ndʒə-mo</i>	the mother of ours/yours/theirs (duals)

3.2.2 Demonstrative pronouns

Principally, *ʃtə* ‘this’ and *wəta* ‘that (over there)’ are distinguished. The form *ʃtə* seems to originate from **ʃətə*, where **ʃə-* means ‘near’. The component *-tə* in both pronouns is treated as a referential indefinite demonstrative pronoun.

3.2.3 Interrogatives

The main interrogative pronouns are shown in (8). They almost always appear immediately before verbs or auxiliary verbs.

(8) <i>sə</i>	who
<i>tʰə</i>	what

- kə-rti* when
kə-tci where
tʰə-ni how
tʰə-ste how many

3.3 Adjectives

3.3.1 Marking of adjectives

Adjectives are marked by *kə-*, such as *kə-mbro* ‘high’, *kə-mo* ‘empty’, *kə-tʃ^l/em* ‘fine (meshes)’, *kə-ktsi* ‘small’, *kə-pram* ‘white’. Property words borrowed from WT do not bear *kə-*. Examples: *ldʒan-ku* ‘green’, *sar-pa* ‘new’. Numerals (§3.6) are also prefixed by *kə-*.

Some adjectives require another affix between *kə-* and the root, for example, *kə-mə-stak* ‘cold’, *kə-sə-mo* ‘wicked’ and *kə-mə-skru?* ‘pregnant’. These affixes seem to behave the same as those in the verb phrases. See §3.4.4.

3.3.2 Reduplication

Reduplication of the adjective roots generally means ‘very much’. Thus, *kə-kte-kte* ‘very big’ against *kə-kte* ‘big’, *kə-pra-pram* ‘very white’ against *kə-pram* ‘white’, *kə-mə-sta-stak* ‘very cold’ against *kə-mə-stak* ‘cold’. If the root ends with a consonant, it disappears in the first component of the reduplication.

3.3.3 Comparative and superlative

The prefix *ndzok* is used on adjective roots to indicate the comparative and *stuy-* indicates the superlative. For instance, *kə-skren* ‘long’, *ndzok-kə-skren* ‘longer’ and *stuy-kə-skren* ‘longest’.

3.3.4 Modification

When an adjective modifies a noun, it follows the noun; *me-tok kə-wə-rne* '(flower- red) red flower'

3.3.5 Composite adjectives

Composite adjectives are formed in two ways: noun + adjective and verb root + adjective, e.g. *tə-wə kə-mo* ‘to be hungry’, where *tə-wə* means ‘stomach’ and *kə-mo* ‘empty’.

3.3.6 Adjectives in the predicate

When an adjective is the predicate, the adjectival prefix *kə-* is dropped and it behaves just as verbs do in terms of person marking and aspect. Instances of person marking with adjectives are shown below. The shape in the braces shows the underlying form of the preceding word.

- | | | |
|-----|--|------------------------|
| (9) | <i>kə-mʃor</i> | beautiful |
| | <i>ŋa mʃɔŋr {Ø-mʃor-ŋ}</i> | I am beautiful. |
| | <i>nə-dzə tə-mʃor {tə-mʃor-Ø}</i> | You are beautiful. |
| | <i>dʒi-dzə tə-mʃor-ntʃ {tə-mʃor-ntʃ}</i> | You two are beautiful. |

3.4 Verb phrases

rGyalrong sentences are either simple or compound. Simple sentences have one verb complex that is necessarily the final one, while compund sentences have any number of non-final verb complexes and a final one. The structure is illustrated schematically as (10):

- (10) [(NP) + VPnon-final]ⁿ (particle) [(NP) + Vpfinal] (AUX) (n is 0, 1 or 2)

3.4.1 *The structure of verb phrases*

The following description mainly deals with simple sentences and the morphological structure of the Vpfinal, which indeed is of puzzling complexity, so much so that the genetic affiliations of the language are somewhat controversial. A VPfinal has the following general structure and it constitutes a word:

- (11) VPFINAL → (*ka*, *kə* or *ke*)-(*ke*)-P1-P2-(P3)-ROOT-(*s*)-S1 (the parenthesized parts are optional)

The prefix *ka*, *kə* or *ke* generally signals the beginning of a VP, being mandatory in a VPnon-final while optional in a VPFinal.

(*ke*) indicates either future or past, in combination with P1. However, ‘tense’ is quite foreign to this language, and actually this prefix does not point to a specific time but refers to a relatively remote period.

P1 stands for the aspect marker or direction marker, while P2 and S1 represent pronominal affixes. They specify agreement, with agent, patient, goal and beneficiary.

P3 is an adverbial affix, which indicates the ‘manner’ of action. If a progressive marker appears at P3, P1 does not occur.

A morpheme *-s*, the derivative suffix to the root, may appear between the root and S1. This suffix appears only with ‘process’ verbs and marks at the same time that the viewpoint aspect is perfective.

There is no variation in the order of the affixes.

3.4.1.1 The morphosyntax of prefixes in verb phrases

What lies beneath the regularity of the prefix ordering are the semantic functions and classes of prefixes. They are summarized as follows (Figure 29.2).

Morphological component	Function	Semantic class
Syntactic order	<i>ka/ə/v</i>	signals VP
	<i>ke</i>	tensifies aspect
	P1	tells if it is done
		or
		tells direction of act
	P2	tells who to whom
	P3	tells manner of act

FIGURE 29.2 MORPHOSYNTAX OF PREFIXES

3.4.1.2 Layers of prefixation

Thus, the prefixes occur as follows (Figure 29.3).

On the semantic level, the chart can be interpreted as follows (Figure 29.4).

It can be claimed, therefore, that the more remote from the root, the less concrete the meaning of the prefix, or that the more remote from the root, the more abstract the semantic function.

3.4.2 Aspect markers

Aspect markers appear in the P1 position, indicating either imperfective or perfective. Imperfective is marked by Ø- and perfective by *nə-*, as shown in the following examples:

- (12) *ŋa diŋ {Ø-dit-ŋ}*
 1SG Ø-give-1SG
 'I am going to give (it).'
- (13) *ŋa nə-diŋ {nə-dit-ŋ}*
 1SG PFV-give-1SG
 'I have given (it).'
- (14) *ŋa ŋə-myŋak ro {Ø-ro}*
 1SG 1SG-eye Ø-wake
 'I wake up.'
- (15) *ŋa ŋə-myŋak nə-ro-s {nə-ro-s}*
 1SG 1SG-eye PFV-wake-PFV.
 'I have awokened.'
- (16) *ŋidzo tə-rdzap tə-sarŋ mə ŋos {Ø-tə-sar-ŋ}*
 2PL(HON) NFP-marriage 2PL-marry-2PL Q AUX
 'Are you going to marry?'

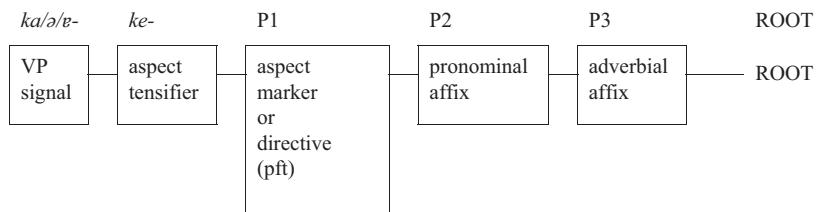


FIGURE 29.3 LAYERS OF PREFIXES

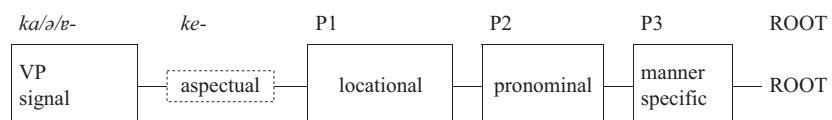


FIGURE 29.4 SEMANTIC LAYERS OF PREFIXES

- (17) *nidzo tə-rdzap nət-sarŋ mə ɲos {nə-tə-sar-ŋ}*
 2PL(HON) NFP-marriage PFV-2PL-marry-2PL Q AUX
 ‘Have you got married?’

The prefactive marker *nə-* is often replaced by a direction marker. But the following verbs conventionally require *nə-*: *te^hop* ‘to cut’, *k^hak* ‘to peel’, *ki* ‘to borrow, to buy’, *krok* ‘to scratch’, *krot* ‘to cut’, *tea* ‘to untie’, *mʃ^hi lat* ‘to bite’, *mʒit* ‘to fall’, *p^hot* ‘to break’, *p^hjis* ‘to wipe’, *p^hit* ‘to drop’, *pja* ‘to take’, *sat* ‘to kill’, *steo* ‘to write’, *ta* ‘to take off’, *jo* ‘to rob’.

3.4.2.1 The aspect tensifier *ke-*

The prefix *ke-* ‘tensifies’ the aspect, moving the time of action to a more remote stage. Compare the following four examples:

- (18) *ŋa pjay {pjə-ŋ}*
 1SG take-1SG
 ‘I am going to take (it).’
- (19) *ŋa ke-pjay {ke-pjə-ŋ}*
 1SG TSF-take-1SG
 ‘I will take (it).’
- (20) *ŋa nə-pjay {nə-pjə-ŋ}*
 1SG PFV-take-1SG
 ‘I have taken (it).’
- (21) *ŋa ke-nə-pjay {ke-nə-pjə-ŋ}*
 1SG TSF-PFV-take-1SG
 ‘I had taken (it).’

3.4.3 Direction markers

The P1 position is occupied either by an aspect marker or by a direction marker. In the imperfective, aspect is marked by Ø- (zero), and no directives appear. Therefore, P1 is always left empty in the imperfective. When the direction of action must be specified in the imperfective, an adverb of time or of direction appears before VPfinal. In the perfective, on the other hand, a variety of affixes occur, specifying the aspect and the direction of the action of the verb. As is mentioned in §3.4.2, *nə-* primarily marks the perfective, but direction markers not only show direction but also function as the perfective markers. If one of them appears, *nə-* disappears.

Each direction marker has two forms, one based on ‘direct information’ and the other on ‘indirect information’. This distinction seems to reflect a psychological nearness or distance of the speaker to the referent. It is also interesting that all the ‘indirect information’ markers end in *-a*.

3.4.3.1 The vertical contrast

There are three markers for vertical contrast. In the items below, the one on the left of the slash is a marker based on ‘direct information’ while the one on the right is ‘indirect information’.

up, uphill	<i>to/ta</i>	↔	<i>no/na</i>	down, downhill, downstream
upstream	<i>ko/ka</i>			

A straightforward contrast is the following:

- (22) *wu-jo-dʒis to-tʰəntʃ* {*to-tʰəl-ntʃ*}
3DL up-go-3DL.
'They two have ascended.'

- (23) *wu-jo-dʒis no-tʰəntʃ* {*no-tʰəl-ntʃ*}
3DL down-go-3DL
'They two have descended.'

Statistically, the distribution of *to-* and *no-* is rather fixed, depending upon the meaning of verbs. Some verbs contain by nature the meaning of 'upward', such as *rwas* 'to get up', *mpʰat* 'to vomit', *kte* 'big, to grow up'. Some others contain the concept of 'to accomplish', requiring *to-*; *sə-jok* 'to finish', *pka* 'to become full', *kʰram* 'to dry', *pa* 'to collect, to make'. In English too, you have 'to eat up', 'to write up', 'to finish up' and so on. Those verbs almost always require *to-* or *no-* in an unmarked situation.

rGyalrong often shows an interesting meaning contrast when a prefix other than the one normally expected is used. For instance:

- (24) *ya mə-ʃʰit no-pʃɪŋ* {*no-pʃɪt-η*}
1SG saliva down-spit-1SG
'I spat.'

- (25) *ya mə-ʃʰit to-pʃɪŋ* {*to-pʃɪt-η*}
1SG saliva up-spit-1SG
'I spat upward.'

Example (24) is a normal statement, whereas (25) is a proverb that means 'The wheel has come full circle'.

- (26) *ya ɳə-nгла to-teʰeŋ* {*to-teʰe-η*}
1SG my-step up-walk-1SG
'I walked.'

- (27) *ya ɳə-nгла no-teʰeŋ* {*no-teʰe-η*}
1SG my-step down-walk-1SG
'I walked step by step.'

Example (26) is a normal utterance, while in (27), much more attention has been paid to the speaker's steps when he/she walked.

The upstream/downstream contrast is similar to that of up/down. *ko-* 'upstream' covers the semantic area of 'coiling up, wringing up'.

- (28) *wu-jo-dʒis ya-ɳə-mki kow-ptsirʃ* {*ko-wu-ptsir-tf*}
3DL my-neck coiling.up-INV-wring-3DL
'They (two) wrung up my neck.'

- (29) *tʰi-i-dzo tə-tak ko-patʃ* {*ko-pa-tʃ*}
1DL NFP-weaving coiling.up-do-1DL
'We (two) have woven.'

- (30) *ŋa tə-dʒi ko-wa-sts^hey {ko-wa-sts^he-ŋ}*
 1SG water coiling.up-CAUS-hot-1SG
 'I have boiled water.'

3.4.3.2 The horizontal contrast

There are two horizontal contrasts: one is front vs behind, and the other is seat of honour vs lower seat.

front	<i>ro-/ra</i>	↔	<i>re /ra</i>	behind
seat of honour	<i>ku/ka</i>	↔	<i>ni/na</i>	lower seat

3.4.3.2.1 THE FRONT/BEHIND CONTRAST

Compare the following:

- (31) *wu-jo-ne ŋa-ŋɔ-rpak rew-nt^hey {re-wu-nt^hen-ŋ}*
 3PL my shoulder back-INV-pull-1SG
 'They have pulled my shoulder.'
- (32) *wu-jo-ne ŋa-ŋɔ-rpak row-nt^hey {ro-wu-nt^hen-ŋ}*
 3PL my shoulder front-INV-pull-1SG
 'They have pulled my shoulder.'

In (32), the agents and the speaker are in a face-to-face position and the speaker's shoulder was pulled towards the agents' noses. In (31), on the other hand, the speaker is located behind the agents, and they stretched their hands to pull the speaker's shoulder towards them. Thus, the location of agent, patient and referent can be predicted from the affix.

- (33) *ʃtə wu-rni-tə re-diŋŋ {re-dit-ŋ}*
 this red-NOMZR back-give-2PL
 'Please give (me) that red one.'

This sentence shows that the speaker is talking to the shop assistant behind whom the merchandise is displayed, and he asks the assistant to take something for him from behind the assistant.

3.4.3.2.2 THE SEAT OF HONOUR/LOWER SEAT CONTRAST

rGyalrong people are so sensitive to the social hierarchy that they distinguish the seat of honour, which is usually located in the eastern part of the room, and the lower seat (the host's seat), which is across the hearth from the seat of honour. *ku-* stands for 'towards the seat of honour' while *ni-* for 'towards the lower seat'. Nowadays this socio-cultural distinction is not always observed in rGyalrong in general. Lin (1993) describes still another system of direction markers.

3.4.3.3 Other markers

There are two more markers, *ne-* and *ji-*. *ne-* implies the movement of 'to go and return', 'to get back'. For instance, *ne-ja* 'to return' is a compound of *ne-* 'to get back' and *ja* 'to go home'. *ji-* stands for a general movement. The expression of 'to go' and 'to come' require *ji-* unless a specific direction of going and coming has to be indicated.

- (34) *wu-jo jik-t^hel {ji-kə-t^hel-Ø}* *ŋos*
3SG general.movement-3SG-go AUX
'He has gone.'

A similar, but a slightly extended, usage of *ji-* is observed in an elegant expression for 'to die'.

- (35) *no-fis*{no-ʃi-s}
down-die-PFV
'He/She died.'
- (36) *ni-fis*{nə-ji-ʃi-s}
PFV-general movement-die-PFV
'He/She passed away.'

3.4.4 Adverbial affixes

The P3 position is occupied by an adverbial affix that specifies the manner of action. Adverbial affixes include progressive markers, causative markers, verbalizers, repetitive act markers and others. 'Adverbial' is the name Wolfenden used, but 'manner specifier or modalizer' seems to me a better description, except for the causative, which is too grammatical to be a manner, and the progressive, which is too aspectual to be a modal.

3.4.4.1 Causative markers

The **s*- prefix, a widespread morpheme in TB languages, functions to mark causativity or goal-oriented directionality. Some innovative languages lost the prefix, retaining only the vestiges of it in other forms. In others, however, it survives in orthography or still functions productively. rGyalrong not only preserves vestiges of the old **s*- but also has ways of converting verbs into causative ones by putting particular morphemes at the P3 position. In this section, only the productive devices at P3, *sə-*, *ʃə-*, *rə-* and *wa-*, will be described. As for the old vestiges, see §3.4.7.

sə- is the most frequent prefix that converts verbs into causative ones. The vowel in the affix harmonizes with that in the root; if the root has a front/unrounded vowel, /ə/ becomes [ɛ]; if the root has a low/back/rounded vowel, it becomes [ʊ]; otherwise, it remains -ə-.

The following pairs illustrate the typical function of *sə-*.

- (37) *miser tə-rmi ke-ta-kej-dzu* *{ke-ta-kə-ji-dzu}.*
yesterday NFP-men TSF-up(PFV)-3PL-general.movement-gather
'People gathered yesterday.'
- (38) *ŋa miser tə-rmi ke-to-sej-dzuyŋ* *{ke-to-sə-ji-dzu-ŋ}.*
1SG yesterday NFP-men TSF-UP(PFV)-CAUS-gather-1SG
'I assembled people yesterday.'
- (39) *ʃtə wə-t^ha wu-ngu-j tə-dok ta-ŋa-teolo {ta-ŋa-teolo} no-to.*
this tea 3SGPOSS-in-LOC NFP-poison up(PFV)-mutual.act-mix EVI-AUX
'Poison has been mixed in this tea.'

- (40) *fθə wə-sman tə-dʒi wu-nGU-j tə-sə-tco-low {tə-sə-tcolo-w}*.
this 3SGPOSS-drug NFP-water 3SGPOSS-IN-LOC 2SG-CAUS-mix-2SG
'Mix this drug in the water.'

Examples of the conversion of transitive verbs to causative are as follows:

- (41) *ŋa ŋə-nGA ke-nə-taŋ {ke-nə-ta-ŋ}*
1SG 1SGPOSS-cloth TSF-PFV-take off-1SG
'I took off my clothes.'
- (42) *ŋa wu-nGA ke-nə-sə-taŋ {ke-nə-sə-ta-ŋ}*
1SG 3SGPOSS-cloth TSF-PFV-CAUS-take.off-1SG
'I undressed him.'
- (43) *ŋa wu-nGA ŋə-ndi nə-sə-taŋ {nə-sə-ta-ŋ}*
1SG 3SGPOSS-cloth 1SGPOSS-servant PFV-CAUS-take off-1SG
'I made my servant undress him.'

Besides this productive *sə-*, rGyalrong has a transitive/intransitive contrast at the initial position of verb roots.

- (44) *ka-ndzop* to burn (vi) *ka-ntcop* to burn (vt)
ka-nbak to split *ka-pʰak* to tear
ka-nglak to fade *ka-klak* to scour off

The affix *fθ-* serves not only to make verbs causative but also to add the meaning of 'to help'. Compare the following three sentences:

- (45) *ŋa ke-rwas {ke-rwas-ŋ}*
1SG TSF-rise-1SG
'I will rise.'
- (46) *ŋa wu-jo ke-sə-rwas {ke-sə-rwas-ŋ}*
1SG 3SG TSF-CAUS-rise-1SG
'I will raise him.'
- (47) *ŋa wu-jo ke-fθ-rwas {ke-fθ-rwas-ŋ}*
1SG 3SG TSF-CAUS-rise-1SG
'I will help him rise.'

The affix *rə-* is the third causative marker. *ka-kfut* means 'to get out' while *ka-rə-kfut* means 'to expel'. Another example is *ka-rə-tfʰak* 'to decrease' from *kə-tfʰak* 'few'.

The last causative marker, *wa-*, functions to transitivize adjectives and nouns. Examples are:

- (48) *nɪdzo tə-dʒi ke-wa-stsʰəŋ mə ŋos {ke-wa-stsʰə-ŋ}*
2PL NFP-water TSF-CAUS-HOT-2PL Q AUX
'Will you boil water?'
- (49) *wa-rdzap dza-roŋ na-tfʰe na-wa-rmow {na-wa-rmo-w}*
3SGPOSS-wife rGyalrong PFV-go PFV-CAUS-dream-3SG
'He dreamt that his wife went to rGyalrong.'

3.4.4.2 Mutual action marker

When it occurs before the root, *ŋə-* serves to indicate the act is mutual.

-
- (50) *wu-jo-džis kew-top* {*ke-wu-top*}
3DL TSF-3DL-hit
'They two will hit (somebody).'
- (51) *wu-jo-džis kew-ŋə-top* {*ke-wu-ŋə-top*}
3DL TSF-3DL-mutual.act-hit
'They two will hit each other.'

3.4.4.3 Repetitive act markers

A repetitive act is marked by *ra-* or *na-*. Jin Peng *et al.* (1957/1958) says that *na-* is followed by reduplicated roots, but, in my data, it is not necessarily so.

- (52) *ŋa nə-ra-kroy* {*nə-ra-kro-ŋ*}
1SG PFV-repetitive.act-scratch-1SG
'I have scratched and scratched.'
- (53) *ʃtə wakej koho-ke məma ra-stcop* {*ra-steo-n*}
this than nice-ADVERBIALIZER polite.demand repetitive.act-write-2PL
'Would you please write more nicely than this?'
- (54) *ʃtə wə-rmi-jo ke-kə-na-riŋ* {*ke-kə-na-ri-n*}
this 3SGPOSS-man-PL TSF-3PL-repetitive act-laugh-3PL
'These guys will laugh.'

3.4.4.4 Automatic/uncontrollable act marker

To express that the act is automatic or uncontrollable, *mə-* appears at P3.

- (55) *ŋa to-mə-mpʰatŋ* {*mpʰat-ŋ*}
1SG up-automatic.act-vomit-1SG
'I have vomited.'

The act 'to vomit' is non-volitional; the verb root *mpʰat* usually requires *mə-*. The following pair of imperative sentences beautifully illustrates the function of *mə-*:

- (56) *to-mə-mpʰat!*
up-automatic.act-vomit
'Vomit!'
- (57) *to-mpʰat!*
up-vomit
'Vomit intentionally.'

Sentence (56) is of a neutral sense, where the addressee feels sick and the speaker tells him not to counteract his natural physiology. In (57), where *mə-* is omitted, on the other hand, the addressee does not feel like vomiting but the speaker thinks that he had better vomit (probably because the speaker knows that the addressee has swallowed something poisonous).

3.4.4.5 Objectivizer

The prefix *sa-* serves to establish psychological distance from a mental action of the agent.

- (58) *na-rmo ke-no-sa-pap* {ke-no-sa-pa-n}.
 2SGPOSS-dream TSF-down-OBJECTIVIZER-make-2PL
 ‘Please dream.’
- (59) *ŋa wu-mi ke-no-sa-nəŋaŋ* {ke-no-sa-nə-ŋa-ŋ}.
 1SG 3SGPOSS-daughter TSF-PFV-OBJECTIVIZER-love(stative)-1SG
 ‘I loved his daughter.’

The verb root in (59) is *nə-ŋa*, which consists of an applicative marker and a root but behaves as a single root.

3.4.4.6 Progressive marker

Progressive aspect is marked by *nə-* at the P3 position. The shape of this affix is identical to that of perfective marker, but little ambiguity occurs because of its position.

- (60) *wu-dzo-ne ŋa-myok wu-dza* {wu-dza}
 3PL 1SGPOSS-grain 3PL-eat
 ‘They are going to eat my grain.’
- (61) *wu-dzo-ne ŋa-myok wu-nə-dza* {wu-nə-dza}
 3PL 1SGPOSS-grain 3PL-PROG-eat
 ‘They are eating my grain.’
- (62) *wu-dzo-ne ŋa-myok tu-dza* {to-wu-dza}
 3PL 1SGPOSS-grain PFV-3PL-eat
 ‘They have eaten my grain.’
- (63) *ji-no nɪdzo nə-myok no-nə-dzej* {no-nə-dza-j}
 1PL(EXC) 2PL 2POSS-grain PFV-PROG-eat-1PL
 ‘We were eating your grain.’

3.4.4.7 Reflexive marker

The prefix *na-* marks reflexives and middles when it appears at P3. For instance, we have *ka-nə-top* ‘to hit oneself’ against *ka-top* ‘to hit’ and *kə-nə-ngri* ‘to collapse by itself/from inside’ against *kə-ngri* ‘to collapse’ (see Jin Peng *et al.* 1957/1958: 81).

3.4.5 Pronominal affixes

Pronominalization is a widespread phenomenon among the TB languages, in the sense that personal pronouns or their remnants are crucial elements in the verb complex. The ways of participation differ greatly from language to language: Lolo-Burmese is really the extreme where pronominalization is almost completely lacking, while the other pole is represented by rGyalrong, Rawang, Lushai, Qiang and some Himalayish languages, in which pronominal components are indispensable constituents of the verb complex. ‘Pronominalization’ in this chapter refers to morphological affixes which appear in the verb complex and reflect agent and patient/goal/beneficiary arguments. Pronominal affixes appear at the P2 and S1 positions as a set. Morphologically, however, those at the P2 position are of a non-pronominal origin (presumably of a demonstrative origin), while those at the S1 position are the remnants of independent personal pronouns, which reflect their person and number. The paradigms differ in their intransitive and transitive structures.

3.4.5.1 The affixes of the intransitive paradigm

	P2	S1
1SG	(<i>kə-</i>)	- <i>ŋ</i>
1DL	(<i>kə-</i>)	- <i>tʃ</i>
1PL	(<i>kə-</i>)	- <i>j</i>
2SG	<i>tə-</i>	- <i>n</i>
2DL	<i>tə-</i>	- <i>ntʃ</i>
2PL	<i>tə-</i>	- <i>n</i>
3SG	(<i>kə-</i>)	-Ø
3DL	<i>kə-</i>	-Ø (or - <i>ntʃ</i>)
3PL	<i>kə-</i>	-Ø (or - <i>n</i>)

The affixes at S1 are recognized to be the remnants of independent personal pronouns (§3.2.1). Those are further analysed as -*ŋ* being for the first person, -*n* for the second person, -*tʃ* for dual and -*j* for plural. In the third person, S1 is marked by zero. The zero marking for the third person category seems to be a universal tendency. The P2 position is occupied by *kə-* or *tə-*. These two prefixes are derived from a non-pronominal origin. As was exhaustively studied by Bauman (1975), *kə-* stands for the first person category and *tə-* for that of the second person. The reason why *kə-* appears for third person arguments is unclear, but this phenomenon may be parallel to that observed in some TB languages in Assam (Tableng, Lhota and Angami, for instance) where the first and third person pronominals merge partly (Bauman 1975: 162–4).

3.4.5.2 The affixes of the transitive paradigm

When there is a set of a human agent and a human patient (or goal or beneficiary), the following sets of pronominal affixes (P2-R[oot]-S1) appear:

PTT/BNF/GOA	AGT		
	1	2	3
1SG		<i>kəw-R-ŋ</i>	<i>wu-R-ŋ</i>
1DL		<i>kəw-R-tʃ</i>	<i>wu-R-tʃ</i>
1PL	<i>ka-R-j</i>	<i>kəw-R-j</i>	<i>wu-R-j</i>
2SG	<i>ta-R-n</i>		<i>təw-R-n</i>
2DL	<i>ta-R-ntʃ</i>		<i>təw-R-ntʃ</i>
2PL	<i>ta-R-n</i>		<i>təw-R-n</i>

PTT	AGT									
	1SG	1DL	1PL	2SG	2DL	2PL	3SG	3DL	3PL	
3	<i>Ø-R-ŋ</i>	<i>Ø-R-tʃ</i>	<i>Ø-R-j</i>	<i>tə-R-n</i>	<i>tə-R-ntʃ</i>	<i>tə-R-n</i>	<i>Ø-R-w</i>	<i>wu-R-Ø</i>	<i>wu-R-Ø</i>	

In the first and second person patient series, the S1 is occupied by the affixes of pronominal origin and agrees with the patient. The affixes at the P2 position are *kəw-* (2>1), *təw-* (3>2), *ka-* (1>1PL), *ta-* (1>2), and *wu-* (3>1). These are underlyingly *kə-wu-*, *tə-wu-*, *kə-a-*, *tə-a-*, and *Ø-wu-* respectively, where *kə-* and *tə-* are the affixes for the first and second person categories of non-pronominal origin, -*wu-* is an inverse marker, and -*a-* is a direct marker. Therefore, P2 also shows patient agreement. The inverse marker -*wu-* occurs only in the 2>1, 3>1 and 3>2 configurations, and -*a-* only in the 1>1PL and 1>2

configurations. This *-a-* seems related to *?a-*, a nominalizer of direction markers (§3.1.1), that always implies proximal deixis.

Unlike the first and second person patient series, the third person patient series shows agent agreement. The inverse prefix, *wu-*, occurs in the $3_{DL}/3_{PL}>3$ configurations. In $3_{SG}>3$ configurations, *-w* is not the inverse affix but the third person marker; *-w* or *-u* is a widespread marker for the third person category in TB that appears primarily in the transitive structure.

If a human patient is absent, the affixing system is as follows:

AGT	P2	S1
1SG	\emptyset	<i>-η</i>
1DL	\emptyset	<i>-tf</i>
1PL	\emptyset	<i>-j</i>
2SG	<i>tə-</i>	<i>-w(u)</i>
2SL	<i>tə-</i>	<i>-ntf</i>
2PL	<i>tə-</i>	<i>-n</i>
3SG	\emptyset	<i>-w</i>
3DL	<i>wu-</i>	\emptyset
3PL	<i>wu-</i>	\emptyset

These components are identical to those of the third person patient agreement, except for 2SG. The reason why *-w(u)* occurs at S1 with the 2SG agreement is unknown.

3.4.6 Suffix *-s*

This suffix indexes ‘perfective’. It is, however, much less productive than other affixes and occurs only with a limited number of verbs. Unlike in WT, where *-s*, which is in complementary distribution with *-d*, is generally employed in the perfective roots, the *-s* in rGyalrong marks the perfective of intransitive process verbs.

- (64) *wu-jo lasa-s no-kə-stce-s {no-kə-stce-s}*
 3SG Lhasa-ABL PFV-3SG-be.born-PFV
 ‘He was born in Lhasa.’
- (65) *wu-jo-ne dzagar-s no-kfis {no-kə-fi-s}*
 3PL India-LOC PFV-3PL-die-PFV
 ‘They (went to India and) died.’

If the subject is 1SG in (64), the VP is of the shape *no-stce-η*, and if it is 1PL in (65), *nəp-fi-j* appears as the VP. Both examples show that the first person pronominals have a higher rank than the suffix *-s*.

Some auxiliary verbs, such as *kə-ra* ‘to need’, *ndo* ‘to exist’, *kə-te^ha* ‘to be able’ and *kə-sa-k^ha* ‘to be difficult’, may take *-s* for their perfective. For instance, *ŋa ka-che no-ra* ‘I need to go’ vs *ŋa ka-che no-ra-s* ‘I needed to go’.

3.4.7 Lexicalization of prefixes

rGyalrong has a complicated VPfinal structure, in which a good number of affixes behave regularly and productively. We have already seen that particular verbs and affixes are naturally connected to each other depending on the meaning of the verb. In more frequent combinations, particular affixes drop vowels and the prefix becomes part of the verb root.

rGyalrong had several waves of this kind of lexicalization. The following examples are the result of that phenomenon:

(66) to change (vt)	<i>s-dzur</i>	to change (vi)	<i>n-dzur</i>
to turn around (vt)	<i>s-kor</i>	to turn around (vi)	<i>n-kor</i>
to wind (vt)	<i>s-kru</i>	to wind (vi)	<i>n-kru</i>
to show	<i>s-rong</i>	to see	<i>Ø-rong</i>
to lend	<i>s-ki</i>	to borrow	<i>Ø-ki</i>
to rise	<i>Ø-was</i>	to get up	<i>r-was</i>

On the other hand, however, we know that in some other dialects a vowel is inserted between a prefix and the root. This sort of de-lexicalization restores the productivity of affixes. This dynamism seen in present-day rGyalrong seems quite reminiscent of what happened in the remote stages of TB.

3.5 Auxiliary verbs

The auxiliary verbs frequently used are: *ka-te^ha* ‘can’, *ka-spa* ‘can’, ‘to be able’, *k^hut* ‘may, to be ready’, *ra* ‘need’, *ka-jok* ‘may, to be allowed,’ *ka-sajok* ‘to finish’, *ndo* ‘to exist’. These require the infinitive form of verbs (normally *ka*-ROOT) before them.

- (67) *ŋa dʒundʒak ka-pa tɛ^hay* {*tɛ^ha-ŋj*}
 1SG swimming INF-do can-1SG
 ‘I can swim.’
- (68) *ŋa kuruskat ka-pa ſpanj* {*ſpa-ŋj*}
 1SG Tibetan INF-do can-1SG
 ‘I can speak Tibetan.’
- (69) *tɔ-dzim ka-ni ma nə-k^hut*
 NFP-house INF-live NEG EVI-ready
 ‘The house is not ready to be lived in.’
- (70) *ŋa tɔ-dzim wu-nGU-j ka-nGO mə jOK*
 1SG NFP-house 3SGPOSS-in-LOC INF-enter Q may
 ‘May I enter the house?’
- (71) *semdə ka-pa ma ra*
 worry INF-do NEG need
 ‘You don’t need to worry.’
- (72) *nədzo ka-nə-ndza mə tɔ-sajok*
 2SG INF-PFV-eat Q 2SG-finish
 ‘Have you finished eating?’
- (73) *nədzo chamdo-j ka-che mə no-ndo-s*
 2SG Chamdo-LOC INF-go Q PFV-AUX-PFV
 ‘Have you ever been to Chamdo?’

Besides these auxiliary verbs above, there are auxiliary verbs of statement (attribute markers), *ŋos* (affirmation) and *mak* (negation), and existential auxiliary verbs, *ndo* (affirmation) and *me* (negation). They function as main verbs, but also bear the role of auxiliary verbs.

- (74) *so-sni tə-mu no-lat dʒe-lat dʒi, ya ka-che yos*
tomorrow NFP-rain down-fall NEG;PFV-fall also 1SG INF-go AUX
'Whether it rains or not, I will go.'

In (74), the main clause without *yos*, *ya ka-chey* {*ka-che-ŋ*}, is fully grammatical. The combination of aspect marker + *to* may appear in place of *ndo*. The auxiliary does not take any pronominal affixes.

3.6 Numerals

3.6.1 Basic numerals

The basic numerals are given in (75):

(75)	one	<i>kə-tek</i>	two	<i>kə-ŋes</i>
	three	<i>kə-sam</i>	four	<i>kə-wdi</i>
	five	<i>kə-myø</i>	six	<i>kə-ŋok</i>
	seven	<i>kə-ŋføs</i>	eight	<i>wə-ŋjat</i>
	nine	<i>kə-ŋgu</i>	ten	<i>sdze</i>
	11	<i>sdze tek</i>	12	<i>sdze ŋes</i>
	20	<i>kə-ŋes sdze</i>	22	<i>kə-ŋes sdze kə-ŋes</i>
	10	<i>pə-ŋja</i>	1,000	<i>stoŋ-tso</i>

3.6.2 Ordinals

The ordinals are loans from WT.

(76)	rGyalrong	WT
	first	<i>tay-bo</i>
	second	<i>ŋes-pa</i>
	third	<i>səm-ba</i>
	fourth	<i>bʒə-ba</i>
	fifth	<i>rja-pa</i>
	sixth	<i>tək-pa</i>
	seventh	<i>bdən-ba</i>
	eighth	<i>rdʒat-pa</i>
	ninth	<i>rgu-ba</i>
	tenth	<i>ptʃu-pa</i>

The names of the months require ordinals; for instance, February is *zla-wa ŋes-pa* 'month second'. For comparison, 'two months' is *kə-ŋes tsə-la*.

3.6.3 Frequency

To express the number of times an action has happened, *tə-lok* is used, e.g. *kə-sam tə-lok* 'three times'.

3.6.4 Fractions

Fractions are expressed by *tə-ŋok*, e.g. *sdze tə-ŋok wu-ŋgūj kə-sam tə-ŋok* 'three tenths'.

3.6.5 Classifiers

Classifiers are fairly abundant. For example, *pʰjar* ‘a piece of (paper, leather)’, *rgi* ‘a grain of, a drop of’, *lpek* ‘a piece of (meat, cloth)’, *pjam* ‘a suit of (clothes)’, *rzək* ‘a bundle of’, *ntʰak* ‘a drop of liquid’.

3.7 Case marking particles

Nominal case markers seem to have been quite alien to rGyalrong because of its intricate but sophisticated system of pronominalization. The language has two locative markers, an instrumental marker, a genitive marker and an ergative marker.

3.7.1 *-j(i)*

-j(i) is a locative marker in a most general sense, which means ‘in’, ‘at’, ‘to’, ‘towards’.

- (77) *bi-sni-so pʰot-pa wə-tʰa tseng-du-j par wu-nə-lat*
yesterday-tomorrow-day Tibetan 3SGPOSS-book Chengdu-LOC photo 3PL-PROG-hit
'Nowadays Tibetan books are being printed in Chengdu.'

Another function of *-j(i)* is to link verbs to mean ‘in order to’.

- (78) *ŋa tə-tʰa kə-ki-j (kə-)tʰeŋ {kə-tʰe-ŋ}*
1SG NFP-book INF-buy-LOC (1SG-)go-1SG
'I go to buy a book.'

3.7.2 *-s* as locative marker

Another locative marker is *-s*, which represents ‘stationary location’, instead of ‘shifting or moving location’ as implied by *-j(i)*.

- (79) *wu-jo lasa-s no-kə-st̩es {no-kə-st̩e-s}*
3SG Lhasa-LOC PFV-3SG-be.born-PFV
'He was born in Lhasa.'

3.7.3 *-s* as ablative and instrumental marker

The marker *-s* also functions as an ablative and instrumental case marker. This *-s* is in complementary distribution with *-ki*, the ergative marker (see §3.7.5); when the preceding syllable ends in a vowel, *-s* appears; otherwise, *-ki*.

- (80) *ŋa teomko-s nə-pay {nə-pa-ŋ}*.
1SG Kyomkyo-ABL PFV-come-1SG
'I came from Kyomkyo.'

- (81) *kə-zor wə-ji-s*
painful 3SGPOSS-reason/cause-INST
'because of pain'

3.7.4 *-i* as a genitive marker

The genitive marker is *-i*. This seems to be a loanword from WT 'i-.

- (82) *wutə-n nə-jo-i yos*
 that.over.there-PL 2SG(HON)-GEN COP
 ‘Those are yours.’

3.7.5 *-ki*

The ergative marker is *-ki*, which may be another loanword from WT. It is suffixed to the transitive agent of any person. It is true that *-ki* and *wu-*, an inverse affix, usually co-occur, but *-ki* can appear when *wu-* is absent. And a transitive agent without *-ki* is also grammatical. This fact may imply that the split ergativity in this language is rather a matter of discourse prominence.

- (83) *fə wə-rmi-tə-ki fə wə-dzat na-nə-mfɔr*
 this 3SGPOSS-man-NFP-ERG this 3SGPOSS-woman PFV-PROG-love
 ‘The man was loving the woman.’

Just like in WT, this *ki-* also marks the instrumental case (see §3.7.3).

3.8 Negation and question marker

Negation is expressed by *ma-* and question marker by *mə-*, both of which are followed by VPfinal, VP non-final or auxiliary verb. Another negation marker, *dʒa/dʒi (dʒe)* appears in perfective aspect in the younger generation’s utterance. See Nagano (2003b).

BIBLIOGRAPHY

- Bauman, James John (1975) *Pronouns and Pronominal Morphology in Tibeto-Burman*. PhD dissertation, University of California, Berkeley.
- Benedict, Paul K. (1972) *Sino-Tibetan: A Conspectus*. Cambridge: Cambridge Univeristy Press.
- Benedict, Paul K. (1991) A note on Proto-Sino-Tibetan-level morphosyntax. *LTBA* 14.1, 137–141.
- Chang, Kun (1968) The phonology fo a Gyarong dialect. *Bulletin of the Institute of History and Philology (Academia Sinica)* 38, 261–275.
- Chang, Kun and Betty Shefts Chang (1975) Gyarong historical phonology. *Bulletin of the Institute of History and Philology (Academia Sinica)* 42.4, 623–765.
- Chirkova, Katia (2012) The Qiangic subgroup from an areal perspective: A case study of the languages of Muli. *Language and Linguistics* 13.1, 133–170.
- Dai, Qingxia and Yanmucha (戴庆厦 & 严木初) (1990) Are there tones in Suomo variety of Gyarong? (in Chinese: 嘉戎语梭磨话有没有声調?). Handout at the 23rd International Conference of Sino-Tibetan Languages and Linguistics, Beijing, China.
- DeLancey, Scott (1980) *Deictic Categories in the Tibeto-Burman Verb*. PhD dissertation, Indiana University.
- DeLancey, Scott (1989) Verb agreement in Proto-Tibeto-Burman. *Bulletin of the School of Oriental and African Studies*, University of London 52.2, 315–333.
- Edgar, J. Huston (1932) The Giarung language. *Journal of West China Border Research Society* 5 (suppl.), 1–63.
- Edgar, J. Huston (1934) Language changes in west China. *Journal of West China Border Research Society* 6, 258–262.

- Gatehouse, David (2011) *Qiang peoples and languages of Sichuan's Ethnic Corridor*. Available at: sichuanzoulang.com (accessed December 2013).
- Gates, Jesse (2011) *One and some in rGyalrongic: exegeting the imperfections*. Term paper, Canada Institute of Linguistics.
- Hodgson, Brian Houghton (1849) On the aborigines of the eastern frontier. *Journal of the Asiatic Society of Bengal* 18, 238–246.
- Hodgson, Brian Houghton (1850) On the aborigines of the north-east frontier. *Journal of the Asiatic Society of Bengal* 19, 309–316.
- Hodgson, Brian Houghton (1853) Sifan and Horsok vocabularies. *Journal of the Asiatic Society of Bengal* 22, 117–151.
- Hsieh, Feng-fan (謝豐帆)(1999) *Theoretical Aspects of Zhuokeji rGyalrong Phonology* (嘉戎語四土話卓克基方言音韻學的理論諸面向). MA thesis, National Tsing Hua University.
- Huang, Bufan (黃布凡) (2007) *Study of Lavrung Language* (『拉坞戎语研究』). Beijing: Minzu Chubanshe (北京：民族出版社).
- Jacques, Guillaume (see also Xiàng Bólín 向柏霖) (2004) *Phonologie et morphologie du japhug (rGyalrong)*. PhD dissertation, University of Paris. Available at: tel.archives-ouvertes.fr/tel-00138568/en (accessed July 2008).
- Jacques, Guillaume (2010) The inverse in Japhug Rgyalrong. *Language and Linguistics* 11.1, 127–157.
- Jin, Peng (co-authored by Dan Kerang, Qu Aitang and Ling Xiangrong) (1957/1958) The phonology and morphology of the Somo dialect of Gyarong language (in Chinese: 嘉戎語梭磨話的語音和形态). *Yuyan Yanjiu* 2, 123–151 and 3, 71–108.
- Kin, P'eng (金鵬) (1949) Étude sur le Jyarung. *Han Hiue* 3, 211–310.
- LaPolla, Randy J. (1992) On the dating and nature of verb agreement in Tibeto-Burman. *Bulletin of the School of Oriental and African Studies*, University of London 55.2, 298–315.
- Li, Fanwen and Lin Xiangrong (李范文 & 林向荣) (1981) On the relationship between rGyalrong and Daofu (in Chinese: 試論嘉戎与道孚語的關係). 西夏研究學術討論會論文 39, 1–19.
- Lin, Xiangrong (林向荣) (1982) On word-formation in Gyarong (in Chinese: 嘉戎语构词法研究). Handout at the 15th International Conference of Sino-Tibetan Languages and Linguistics, Beijing, China.
- Lin, Xiangrong (林向荣) (1983) On word-formation in Gyarong (in Chinese: 嘉戎语构词法研究). *Minzu Yuwen* 1983.3, 47–58.
- Lin, Xiangrong (林向荣) (1988) The phonology of the Zhoukeji dialect of the Jiarong language in Barkam County (in Chinese: 嘉戎语马尔康卓克基话音系). *Yuyan Yanjiu* 1988.2, 168–189.
- Lin, Xiangrong (林向荣) (1989) On the tones of the Jiarong language (in Chinese: 关于嘉戎语的声调問題). *Journal of the Central Institute for Nationalities (Bimonthly)* 1989.5, 64–68.
- Lin, Xiangrong (林向荣) (1993) *Study of the Gyarong Language* (in Chinese: 『嘉戎语研究』). Chengdu: Sichuan Minzu Chubanshe (成都：四川民族出版社).
- Lin, You-Jing (林幼青) (2000) *Tense, Aspect and Modality Inflection in the Zhuōkèjī rGyalrong Verb* (嘉戎語卓克基話動詞的時、體、與模態之屈折形態). MA thesis, National Tsing Hua University.
- Lin, You-Jing (林幼青) (2003) Tense and aspect morphology in the Zhuōkèjī rGyalrong verb. *Cahiers de Linguistique—Asie Orientale* 32.3, 245–286.
- Lin, You-Jing (林幼青) (2009) *Units in Zhuōkèjī rGyalrong Discourse: Prosody and Grammar*. PhD dissertation, University of California, Santa Barbara.

- Nagano, Yasuhiko (長野 泰彦) (1978a) Preliminary remarks to rGyarong dialectology. Handout at the 11th International Conference of Sino-Tibetan Languages and Linguistics, Tucson, Arizona.
- Nagano, Yasuhiko (長野 泰彦) (1978b) *A note to the rGyarong Tsangla body part terms*. Term paper, University of California, Berkeley.
- Nagano, Yasuhiko (長野 泰彦) (1979a) A historical study of rGyarong initials and prefixes. *Linguistics of the Tibeto-Burman Area* 4.2, 44–67.
- Nagano, Yasuhiko (長野 泰彦) (1979b) A historical study of rGyarong rhymes. *Linguistics of the Tibeto-Burman Area* 5.1, 37–47.
- Nagano, Yasuhiko (長野 泰彦) (1984a) *A Historical Study of the rGyarong Verb System*. Tokyo: Seishido.
- Nagano, Yasuhiko (長野 泰彦) (1984b) The rGyarong directives (in Japanese). *Kikan Jinruigaku* 15.3, 1–52.
- Nagano, Yasuhiko (長野 泰彦) (2001) An outline of Gyarong grammar (in Japanese). *Bulletin of the National Museum of Ethnology* 26.1, 131–164.
- Nagano, Yasuhiko (長野 泰彦) (2003a) Cogtse Gyarong, in: Graham Thurgood and Randy J. LaPolla (eds), *The Sino-Tibetan Languages*, 469–489. London and New York: Routledge.
- Nagano, Yasuhiko (長野 泰彦) (2003b) Preliminary remarks on Gyarong negation particles, in: David Bradley, Randy J. LaPolla, Boyd Michailovsky and Graham Thurgood (eds) *Language Variation*, 159–172 (Pacific Linguistics No.555). Canberra: Australian National University.
- Nagano, Yasuhiko (長野 泰彦) (2008) A preliminary note to the Gyarong color terms. *Revue d'Études Tibétaines* 14, 99–106.
- Nagano, Yasuhiko and Marielle Prins (eds) (2013) *The rGyalrongic Languages Database*. Osaka: National Museum of Ethnology. Available at: htq.minpaku.ac.jp/databases/rGyalrong/ (accessed October 2013).
- Prins, Marielle (2011) *A Web of Relations: A Grammar of rGyalrong Jiǎomùzú (Kyom-kyo) Dialects*. PhD dissertation, Leiden University.
- Qu, Aitang (瞿靄堂) (1983) The person category of the Gyarong verb (in Chinese: 嘉戎语动词的人称范畴). *Minzu Yuwen* 1983.4, 31–34.
- Qu, Aitang (瞿靄堂) (1984) A brief description of Gyarong (in Chinese: 嘉戎语概观). *Minzu Yuwen* 1984.2, 67–80.
- Qu, Aitang (瞿靄堂) (1990) The dialects of Gyarong (in Chinese: 嘉戎话的方言). *Minzu Yuwen* 1990.4, 1–8.
- Shirai, Satoko (2010) Directive affixes in nDapa, rGyalrong and the surrounding languages. Handout at the Tibeto-Burman Linguistics Seminar 22 at Kobe City University of Foreign Studies.
- Sun, Hongkai (孙宏开) (1990) Languages of the ethnic corridor in Western Sichuan. *Linguistics of the Tibeto-Burman Area* 13.1, 1–31.
- Sun, Hongkai (孙宏开) (2000a) Parallelisms in the verb morphology of Sìdàbà rGyalrong and Lavrung in rGyalrongic. *Language and Linguistics* 1.1, 161–190.
- Sun, Hongkai (孙宏开) (2000b) Stem alternations in Pǔxī verb inflection: Toward validating the rGyalronic subgroup in Qiangic. *Language and Linguistics* 1.2, 211–232.
- Sun, Hongkai (孙宏开) (2001) On the Qiangic branch of the Tibeto-Burman language family (in Chinese: 论藏缅语族中的羌语支语言). *Language and Linguistics* 2.1, 157–181.
- Sun, Jackson T.-S. (孫天心) (1994) Caodeng rGyalrong phonology. *Linguistics of the Tibeto-Burman Area* 17.2, 29–47.

- Sun, Jackson T.-S. (孫天心) (1998) Nominal morphology in Caodeng rGyalrong. *Bulletin of the Institute of History and Philology (Academia Sinica)* 69, 103–149.
- Sun, Jackson T.-S. (孫天心) (2000) Parallelisms in the verb morphology of Sidaba rGyalrong and Lavrungr in rGyalrongic. *Language and Linguistics* 1.1, 161–190.
- Sun, Jackson T.-S. (孙天心) (2004) Verb-stem variations in Showu rGyalrong. *Studies on Sino-Tibetan Languages*, 269–296. Available at: www.ling.sinica.edu.tw/eip/FILES/publish/2007.9.19.6519716.91287595.pdf (accessed 17 August 2016).
- Sun, Jackson T.-S. (孫天心) (2007) The irrealis category in rGyalrong. *Language and Linguistics* 8.3, 797–819.
- Sun, Jackson T.-S. (孫天心) and Lin You-Jing (林幼青) (2007) Constructional variation in rGyalrong relativization. Handout at the International Workshop on Relative Clauses, Taipei.
- Sun, Jackson T.-S. (孙天心) and Shi Danluo (石丹罗) (2007) Expressives in the Caodeng dialect of rGyalrong (in Chinese: 草登嘉戎语的状貌词). *Minzu Yuwen* 2004.5, 1–11.
- Suzuki, Hiroyuki (2004) Consonant clusters of the languages in the ethnic corridor of West Sichuan (in Japanese). Handout at the Tibeto-Burman Linguistics Seminar 3 at Kyoto University.
- Suzuki, Hiroyuki (2004) (2010) Numerals in Sidaba rGyalrong. Handout at the Tibeto-Burman Linguistics Seminar 22 at Kobe City University of Foreign Studies.
- Thomas, Frederick W. (1948) *Nam*. Publications of the Philological Society 14. London: Oxford University Press.
- Thomas, Frederick W. (edited by T. Takeuchi, B. Quessel and Y. Nagano) (2011) *Research Notes on the Zhangzhung Language by Frederick W. Thomas at the British Library*. Senri Ethnological Reports No. 99. Osaka: National Museum of Ethnology.
- Thurgood, Graham (1985) Pronouns, verb agreement systems, and the subgrouping of Tibeto-Burman, in: Graham Thurgood, James A. Matisoff and David Bradley (eds) *Linguistics of the Sino-Tibetan Area*, 376–400 (Pacific Linguistics Series C No. 87). Canberra: Australian National University.
- bTsan-lha Ngag-dbang Tshul-khrims (贊拉 阿旺措成) (1997) On the Old Tibetan words in the rGyalrong language (试论嘉戎语话中的古藏语). Handout at the Beijing Tibetology Seminar.
- bTsan-lha Ngag-dbang Tshul-khrims (贊拉 阿旺措成) (2009) *Lexicon of the rGyalrong bTsanla Dialect*. Senri Ethnological Reports No. 79. Osaka: National Museum of Ethnology.
- van Driem, George (1993) The Proto-Tibeto-Burman verbal agreement system. *Bulletin of the School of Oriental and African Studies*, University of London 56.2, 292–334.
- von Rosthorn, A. (1897) Vokabularfragmente ost-tibetischer Dialekte. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 55, 524–533.
- Watters, David E. (1975) The evolution of a Tibeto-Burman pronominal verb morphology. *Linguistics of the Tibeto-Burman Area* 2.1, 45–79.
- Wei, Jiewu (韦介武) (2001) *Verbal Prefixation in (Zhoukeji) rGyalrong*. MA thesis, Tsing Hua University.
- Wen, Yu (聞宥) (1943) Verbal directive prefixes in the Jyarong language and their Ch'iang equivalents. *Studia Serica* 3, 11–20.
- Wen, Yu (聞宥) (1944) Personal endings of the verb in the Gyarong language (in Chinese: 論嘉戎語動詞之人稱尾詞). *Bulletin of Chinese Studies* 4, 79–94.
- Wolfenden, Stuart N. (1929) *Outlines of Tibeto-Burman Linguistic Morphology*. London: Royal Asiatic Society.

- Wolfenden, Stuart N. (1936) Notes on the Jyarong dialect of Eastern Tibet. *T'oung Pao* 32, 167–204.
- 向柏霖 (Xiàng Bólín = Guillaume Jacques) (2008) *Study of the rGyalrong Language* (『嘉绒语研究』). 中国新发现语言研究丛书. 北京: 民族出版社.
- Yamaguchi, Zuiho (1968) rTsang yul and Yan lag gsum pa'i ru (in Japanese). *Toyo Gakuho* 50.4, 1–69.
- Yamaguchi, Zuiho (1969) Pai-lan and the rLangs of the Sum pa (in Japanese). *Toyo Gakuho* 52.1, 1–61.
- Yanmucha (严木初) (2013) A study of the Tibetan and Chinese influences upon the rGyalrong language (in Chinese: 藏语和汉语的嘉戎语影响变迁情况研究). *Bulletin of the National Museum of Ethnology* 38.1, 1–16.

CHAPTER THIRTY

STAU (ERGONG, HORPA)¹

Guillaume Jacques, Lai Yunfan, Anton Antonov and Lobsang Nima

1 INTRODUCTION

The cluster of languages variously referred to as Stau, Ergong or Horpa in the literature are spoken over a large area from Nd Zamthang county (in Chinese Rangtang 壤塘县) in Rongba prefecture (Aba 阿坝州) to Rtau county (Daofu 道孚) in Dzarmdzê prefecture (Ganzi 甘孜州), in Sichuan province, China. At the moment of writing, it is still unclear how many unintelligible varieties belong to this group, but at least three must be distinguished: the language of Rtau county (referred to as ‘Stau’ in this paper), the Dgebshes language (Geshizha 格什扎话) spoken in Rongbrag county (Danba 丹巴), and the Stodsde language (Shangzhai 上寨) in Nd Zamthang. The people speaking these languages are all classified as Tibetans by the administration.

Research on these languages is still limited. No dictionaries or text corpora have yet been published. Stodsde is only known through a few articles (Sun 2000a, 2007), Dgebshes by a grammatical sketch (Duoerji 1998). The dialects spoken in Rtau county have been investigated by several teams of scholars (Huang 1991; Sun and Tian 2013; Jacques *et al.* 2014), but no detailed description of this language is available yet.

There is no consensus as to how these languages should be named. The present paper adopts the Tibetan names of these languages rather than the Chinese-based ones, since the Chinese names are transcriptions of the Tibetan ones.

The native speaker among the authors (Lobsang Nima) favours the name *rəsnəske* for his language, but this name is not used by all speakers and we prefer the geographically based name ‘Stau’. The spelling with *St-* rather than *Rt-*, aside from being more pronounceable for the average Western reader, reflects better the local pronunciation of the county name *stewu*.

This spelling has already been used in English (see Wang 1970–1). Most of the data in this paper come from the Khang.gsar *qʰɔrye* dialect of Stau (in Chinese 孔色 Kongse), except for some Khroskyabs and G.yu.rong Horpa examples from Lai Yunfan’s fieldwork.

For the group comprising all three languages (Stau, Dgebshes and Stodsde), we could adopt either Jackson Sun’s (2000a) term ‘Horpa’, which has indeed been used for this area in the past, or the more specific ‘Tre-Hor’ (an unusual Tibetan name with the rare *tr-* initial cluster). The Chinese term Ergong 尔龚 used by scholars such as Sun Hongkai (1983), on the other hand, appears to lack any basis in the local languages.

As shown in section 8, there is evidence from verbal morphology and lexicon that Horpa languages form a subgroup within Rgyalrongic with Khroskyabs (previously known as ‘Lavrung’), as these two branches present common innovations which are unlikely to represent parallel developments, and which are not shared with the core Rgyalrong languages (Japhug, Tshobdun, Zbu and Situ).

2 PHONOLOGY

Unlike Khroskyabs and Rgyalrong languages, tonal contrasts have not been reported in any variety of Stau.

2.1 Onsets

Table 30.1 presents the consonantal inventory posited for Stau. Unlike in Japhug or Tshobdun, there is no evidence for treating the prenasalized voiced stops as single phonemes in Stau.

Some dialects of Stau have contrastive aspirated fricatives (see Sun 2000b; Jacques 2011a). In the Khang.gsar dialect, the voiceless fricative phonemes are realized as aspirated in syllable-initial position without a cluster, and as unaspirated when they are part of a cluster. Thus, *nə-sow* ‘I killed him’ is realized as [nəs^how], while no such aspiration is observed in *nə-fse* ‘he killed him’ due to being in a cluster.

Voiced stops are almost absent word-initially in the Khang.gsar dialect of Stau. In the case of verbs, the voiced stop or affricate resurfaces when a prefix is added, as in the examples in Table 30.2 (the bare stem appears in non-finite and in factual forms).

Voiceless stop and affricate initials do not show any alternation, as for instance *tei* ‘wear (hat)’ → *rə-teu* ‘I wore it’.

Nouns do not have any productive prefixes, and therefore apart from *bələs* ‘cheek’, the only native word whose voicing is preserved word-initially, we have no way of ascertaining which nouns have underlyingly voiced initials.

The status of the voiceless labiodental fricative *f* is unclear. This sound is never found as a simple onset in native words (only in borrowings from Chinese). However, some

TABLE 30.1 CONSONANTAL PHONEMES IN STAU

		Bilabial	Labiodental	Dental/ Alveolar	Retroflex	Alveolo- palatal	Palatal	Velar	Uvular
Stop	voiceless	p		t		c	k	q	
	aspirated	p ^h		t ^h		c ^h	k ^h	q ^h	
	voiced	b		d		j	g		
Affricate	voiceless			ts	tʂ	tɕ			
	aspirated			ts ^h	tʂ ^h	tɕ ^h			
	voiced			dz	dʐ	dʑ			
Nasal		m		n			jŋ	ŋ	
Fricative	voiceless		f	s	ʂ	ɕ	x	χ	
	voiced		v	z	ʐ	ʑ	y	ɣ	
Approximant		w					j		
Rhotic					r				
Lateral	sonorant			l					
	voiceless			ɺ					
	voiced				ɻ [ɻd]				

TABLE 30.2 NEUTRALIZATION OF VOICED STOPS IN WORD-INITIAL POSITION

Bare stem	Meaning	Prefixed verb	Meaning
<i>pərje</i>	burn (it)	<i>tə-bərje-sə</i>	it burnt
<i>te</i>	do	<i>nə-dej</i>	do it!
<i>tɕə</i>	meet	<i>kə-dʐə</i>	I met (him)
<i>kə</i>	wear	<i>rə-gu</i>	I wore it

clusters such as *gf[lɸʷ]* (contrasting with *rv*), where *f* is the element closest to the vowel, are difficult to account for without positing a voiceless labial fricative phoneme.

In clusters with a nasal as a first element, only stops and affricates are attested, except for the cluster *nl*, which is in free variation with *ntʰ*.

2.2 Rhymes

There are ten vowels in the Khang.gsar dialect of Stau, six plain vowels (*-i, -e, -a, -ə, -ə, -u*), two velarized vowels (*-oʳ, -aʳ*) and two nasal vowels (*-ō, -ā*).

The velarized vowels are almost exclusively attested in Tibetan loanwords (*-oʳ* and *-aʳ* correspond to Tibetan *-og* and *-ag/-eg*, respectively).

Only three codas are possible in Stau: *-v*, *-r* and *-m* (the latter attested only in Tibetan loanwords). In native words, the coda *-r* sometimes appears to correspond to *-r* in other Rgyalrongic languages (as in *χteʰar* ‘sour’, Japhug *teur*), but in other cases, such as *spar* ‘be thirsty’ (Japhug *epaʂ*) or *zdər* ‘cloud, be cloud’ (Japhug *zdum* ‘cloud’), there is no *-r* coda anywhere else in the family. A possible explanation is that the sensory evidential suffix *-rə*, whose vowel tends to be elided, has been reanalysed as part of the stem in third person forms.

2.3 Vowel fusion

In order to account for vowel alternations observed in the verbal and nominal systems, Jacques *et al.* (2014) postulate a series of vowel fusion rules, summarized in Table 30.3 (the symbol stands for either final *-r* or final *-v*).

Vowel fusion cannot occur with the vowels *-u, -oʳ, -aʳ, -ō* and *-ā* and the rhymes ending in *-m*.

3 VERBAL MORPHOLOGY

3.1 Intransitive conjugation

In the intransitive conjugation, the second and third person singular forms are in the bare stem, while first person (singular and plural) forms have a suffix *-ā*.

TABLE 30.3 VOWEL FUSION IN STAU

Stem	Suffix	<i>-w</i>	<i>-ā</i>	<i>-j</i>
<i>i</i>		<i>u</i>	<i>ā</i>	<i>i</i>
<i>e</i>		<i>ow</i>	<i>ā</i>	<i>ej</i>
<i>a</i>		<i>ow</i>	<i>ā</i>	<i>ej</i>
<i>ə</i>		<i>u</i>	<i>ō</i>	<i>i</i>
<i>ə</i>		<i>ow</i>	<i>ō</i>	<i>ej</i>
<i>əv</i>		<i>u</i>	<i>ō</i>	<i>iv</i>
<i>ər</i>		<i>u</i>	<i>ō</i>	<i>i</i>
<i>aC</i>		<i>ow</i>	<i>ā</i>	<i>ej</i>
<i>eC</i>		<i>ow</i>	<i>ā</i>	<i>ej</i>
<i>əC</i>		<i>ow</i>	<i>ō</i>	<i>ej</i>

TABLE 30.4 VOWEL ALTERNATIONS IN OPEN-SYLLABLE INTRANSITIVE VERBS IN STAU

Meaning	look at	move	like	be full	be ill	be hot
1	<i>scəqā</i>	<i>mbəcā</i>	<i>rgā</i>	<i>fkō</i>	<i>ŋō</i>	<i>cʰu</i>
1 (underlying)	<i>scəqi-ā</i>	<i>mbəce-ā</i>	<i>rga-ā</i>	<i>fkə-ā</i>	<i>ŋə-ā</i>	<i>cʰu-ā</i>
2/3	<i>scəqi</i>	<i>mbəce</i>	<i>rga</i>	<i>fkə</i>	<i>ŋə</i>	<i>cʰu</i>

TABLE 30.5 IRREGULAR INTRANSITIVE VERBS IN STAU

meaning	go	say
1	<i>ɛā</i>	<i>jā</i>
2/3	<i>ɛə</i>	<i>jə</i>

Following the rules of vowel fusion in Table 30.3, vowel fusion between the first person suffix *-ā* and the verb stem has different results depending on the vowel (and coda) of the verb stem.

Six classes of alternations are found in verbs with open syllables; class 6 includes verbs without alternation, whose rhyme can be any of *-u*, *-o^y*, *-a^y*, *-ō* and *-ā* (see Table 30.4).

There are two irregular verbs with intransitive morphology which have *-ā* in the first person instead of expected *-ō* (see Table 30.5).

3.2 Transitive conjugation

The basic structure of transitive conjugations in Khang.gsar Stau can be illustrated by Table 30.6 (the columns represent the P and the rows the A). In addition to the first person suffix *-ā*, we find two additional suffixes (1sg→3 *-w* and 2→3 *-j*) and the inverse prefix *f-/v-* (the allomorphs are conditioned by the voicing of the onset). The only unmarked form in the paradigm is the 1→2 slot, a curious fact which however can be accounted for historically (see Jacques *et al.* 2014).

The vowel fusion rules in Table 30.3 apply to all suffixed forms, as in the intransitive paradigms. No example of irregular vowel fusion has yet been discovered with transitive verbs.

The *f-/v-* prefix appears in 2/3→1, 3→2 and 3→3 forms. Its presence in 2→1 precludes an analysis as a third person agent marker, and it is best to treat it as an inverse marker (see section 8). The inverse prefix does not occur with verb stems containing an initial cluster, or with monosyllabic verbs with the coda *-v* such as *kʰev* ‘scoop (water)’ or *cev* ‘take out (of a pile)’.

3.3 Other dialects

Some dialects of Stau have person marking systems that are remarkably different from the one described above. In the G.yu.rong (*ɛjurō*) variety of Stau, for instance, we find the paradigms indicated in Table 30.7.

In comparison with Khang.gsar, the G.yu.rong dialect presents both innovative and conservative features. Khang.gsar is more conservative in having a special marker *-w* for

**TABLE 30.6 THE TRANSITIVE CONJUGATION
IN STAU**

P A \	1SG	1PL	2	3
1SG	v-Σ-ã		Σ	Σ-w
1PL				Σ-ã
2				Σ-j
3				v-Σ

**TABLE 30.7 PERSON INDEXATION IN THE
G.YU.RONG DIALECT**

P A \	1SG	1PL	2	3
1SG	v-Σ-ŋ	v-Σ-j	Σ-n	Σ-ŋ
1PL				Σ-j
2				Σ-n
3			v-Σ-n	v-Σ
INTR	Σ-ŋ	Σ-j	Σ-n	Σ

TABLE 30.8 ORIENTATION PREFIXES IN STAU

Direction	Perfective/Imperative	Interrogative/Irrealis
Up	rə-	rí-
Down	nə-	ní-
North	kə-	kí-
South	γə-	γí-
No direction	tə-	tí-

1SG→3, distinct from 3→1SG, whereas G.yu.rong, like Khroskyabs and Rgyalrong languages, has the suffix in 1SG intransitive, 1SG→3 and 2/3→1SG forms. G.yu.rong on the other hand preserves a distinction between first singular (-ŋ) and plural (-j) in all forms.

3.4 Orientation and negative prefixes

In Stau, orientation prefixes come in two series (Table 30.8), the *-ə* series for imperative and perfective/evidential, and the *-i-* series for interrogative and irrealis forms. A few verbs such as *vde* ‘see’ *ste* ‘finish’ or *si* ‘know’ never appear with any orientation prefix even in the perfective.

There are three negative prefixes in Stau: the past negative *ma-*, non-past *mí-*, and the prohibitive *di-*. The prohibitive is used with an orientation prefix, as in *nə-dif-se-ã* (*nədifšã*) PFV-PROHIB-INV-kill-1 ‘don’t kill me’. The past negative can be used with the orientation prefix (see (1)) or without it. In both cases, the negative appears after the orientation prefix, the reverse order of the one found in Rgyalrong languages.

- (1) *e-cʰe nə-ma-və-w*
 one-CL PFV-NEG-do-1SG→3
 ‘I did not do anything.’

3.5 Derivational morphology

The only denominal prefix in Stau is *s-/z-*, cognate to Japhug *sui-/sr-/eu-* (on which see Jacques 2014: 14–17). It derives transitive verbs, illustrated by the examples in Table 30.9. While the original semantics of this prefix was likely ‘use X’ and ‘cause sb. to have X’ as in Japhug, in Stau the semantics of the denominal verbs are largely unpredictable, and result from semantic shifts (‘use a staff’ → ‘hit with a staff’ → ‘hit’).

The only denominal verb shared by Stau, Khroskyabs and Japhug is *smi* ‘give a name’ (in Japhug *sr̥mi*). When the stem of the base noun has an initial cluster, addition of the *s-/z-* denominal prefixes causes dropping of the preinitial consonant to make the word conform to the phonotactics of the language (thus *r̥mi* → *smi*, not **sr̥mi*, an impossible cluster in Stau).

In at least one case (*smi* ‘hurt’), in addition to the *s-* causative, we observe unexplained vowel alternation.

Stau has a few examples of anticausative verbs with voiced onsets (Table 30.10). The anticausative verbs derive from the transitive ones (not the opposite direction of derivation; see Jacques 2012a). There are two irregular forms in this table. The first one concerns the pair *fteə* vs *dza* ‘melt (tr/it)’, as the f- element of the transitive verb has no equivalent in the anticausative one. Interestingly, the same irregularity is found in the cognate pair in Japhug (*ft̥si* vs *ndz̥i*). As for the second one, *fɪsə* vs *brə* (instead of *(v)*dzə*) ‘wake (tr)/wake up (intr)’, it appears to be unique to Stau.

Unlike Rgyalrong languages, anticausative derivation in Stau is attested in verbs with fricative initial consonants (as in *zəla* ‘fall’).

There is a causative *s-/z-* prefix in Stau attested in a few verbs, but unlike in Rgyalrong languages, it is not productive. Table 30.11 provides a representative list of verb pairs. The phonological alternations attested with the *s-/z-* prefix are much less complex than those attested in Stödsde (Sun 2007) or in Khroskyabs (Lai 2014). However, the causative forms are not always predictable from the underlying base form: for instance, the causative form of voiced initial verbs can be either voiced (*zgə* ‘put clothes on’) or voiceless (*spərje* ‘burn (tr)’), and that of voiceless initial verbs can be aspirated (*sqʰə* ‘extinguish’).

Some Tibetan loan verbs, borrowed in pairs, should be distinguished from the native causative pairs (*mbjer* ‘be pasted on’ vs *zgwer* ‘paste’ from Tibetan *nb̥iar* and *sb̥iar*).

TABLE 30.9 DENOMINAL VERBS IN STAU

Base noun	Meaning	Denominal verb	Meaning
<i>pəcʰa</i>	staff	<i>zbəcʰa</i>	hit
<i>r̥mi</i>	name	<i>smi</i>	give a name
<i>y̥me</i>	wound	<i>smi</i>	hurt

TABLE 30.10 ANTICAUSATIVE VERBS IN STAU

Base verb	Meaning	Anticausative verb	Meaning
<i>zəla</i>	cause to fall	<i>zəla</i>	fall
<i>pʰre</i>	break (tr)	<i>bre</i>	break (it)
<i>fkʰe</i>	cut down	<i>vge</i>	break away, off
<i>fɪsə</i>	wake (tr)	<i>brə</i>	wake up
<i>fteə</i>	melt (tr)	<i>dza</i>	melt (it)

TABLE 30.11 CAUSATIVE VERBS IN STAU

Base verb	Meaning	Causative verb	Meaning
<i>lə</i>	boil	<i>zʒdə</i>	boil (tr)
<i>cʰu</i>	hot	<i>scʰu</i>	cook
<i>rŋi</i>	borrow	<i>sŋi</i>	lend
<i>qə</i>	go out (fire)	<i>sqʰə</i>	extinguish
<i>tʰi</i>	drink	<i>stʰi</i>	give to drink
<i>kə / -gə</i>	wear	<i>zgə</i>	put clothes on
<i>nə</i>	burn (it)	<i>snə</i>	burn (tr)
<i>pərje / -bərje</i>	burn	<i>spərje</i>	burn (tr)

In addition to the prefix *s-/z-*, there is one example of a causative *y* prefix in the pair *ndzi* ‘learn’ → *yzi* ‘teach’. This may be a fossilized allomorph of the causative prefix (in Khroskyabs the corresponding pair is *ndzé* ‘learn’, *ldzé* ‘teach’ with l- allomorph of the causative prefix).

The *s-/z-* has semantic effects that are sometimes better described with terms other than ‘causative’. There is one example of applicative *s-* prefix: *qʰe* ‘laugh’ → *sqʰe* ‘laugh at’. In addition, we find *nənə* ‘smell (intr.)’ → *snənə* ‘smell (tr.)’ (not ‘cause to have a smell’; notice how the causative prefix takes part in the reduplication of the base form), whose semantics can be described as ‘tropative’ (X → ‘find, consider to be X’; see Jacques 2013).

The causative prefix *s-/z-*, unlike its cognate in Core Rgyalrong languages, is not productive anymore, and the only productive way to express causation is a synthetic construction with the auxiliary *xtsʰə* ‘let’ (see example (17)).

There is little evidence of incorporation in Stau, unlike in Japhug (Jacques 2012b) or Khroskyabs (Lai 2013b, 2015). Only two examples have been found: *rvatca* ‘carry on shoulders’ (which incorporates the noun *rva* ‘shoulders and upper back’) and *mbarji* ‘stride over’ (with *pa* ‘a step’).

3.6 Existential verbs

As previously noticed by Huang Bufan (1991: 38), as in most languages of the area (Huang Chenglong 2013), Stau has several existential verbs depending on the nature of the S.

The verb *stʰə* is used to refer to movable things, that have been put at a particular place, as in example (2).

- (2) *ŋi xtse əzi stʰə?*
 1SG:GEN soup where exist
 ‘Where is my soup?’ (The louse and the flea, 08)

The verb *xi* is used with objects that are fixed in place or plants that have grown on the ground.

- (3) *akəstəmba-j zgərju qʰi nə-xi-sə ŋə-rə*
 Akhu.stonba-GEN windows near PFV-exist-IFR be-SENS
 ‘(The big watnut tree) was near Akhustonba’s window.’ (Akhustonba and the walnut tree, 08)

For animates (humans and animals), *ci/ji* is employed. It is the only existential verb to be compatible with person marking, in particular in the progressive construction, which combines a non-finite verb form with the existential *ci/ji*, as in (4).

- (4) *ŋa təq^{hi} tea zɻdə-sə ci-ã*
 1SG here tea boil-NMLZ exist-1SG
 ‘I am boiling some tea here.’ (Akhustonba and the horseman, 19)

The verb *tə/də* is used for all other cases, in particular abstract concepts, but also appears with concrete objects which cannot be put into any of the preceding categories.

- (5) *kʰakʰa kʰarma rdzi eatsa tə-rə zəgu ŋayji*
 other animals footprints many exist-SENS however 1PL:GEN
xə-j rdzi de, ŋayji azə-w
 hybrid.yak-GEN footprints DEM 1PL:GEN mother’s.brother-ERG
taŋtar χa kɔ-rə
 clearly understand understand-SENS
 ‘Although there were footprints of many other animals, our uncle could distinguish the footprints of our hybrid yak.’ (The hybrid yak, 22–3).

4 NOUN PHRASE

4.1 Structure of the noun phrase

Noun phrases in Stau follow the template given in (6), as illustrated by example (7).

- (6) NOUN-ADJ-NUM+CL-DEM

Stau has a set of classifiers with numeral prefixes, which present morphophonological alternations and differ from free numerals. For instance, the numeral ‘one’ is *ru*, but the prefix ‘one’ on classifiers is *a-*, *e-*, or *ə-* (the distribution of the three allomorphs is not completely predictable synchronically).

- (7) *mbrovčə ke cʰe ə-lə ci-rə*
 nomad.dog very big one-CL exist:ANIM-SENS
 ‘There was a huge nomad dog.’ (The hybrid yak, 70)

4.2 Case

There are six case postpositions in Stau: the ergative *-w*, the genitive *-j*, the instrumental *-kʰa*, the dative *-gi*, the locative/superessive *-təʰa*, the allative *-xa* and the comitative *-pʰa*.

The ergative and the genitive merge with the last word of the preceding noun phrase, and the regular vowel fusion rules seen in Table 30.3 apply. Table 30.12 presents examples of ergative and genitive forms of some common nouns.

The first and second person pronouns (*ŋa* 1SG, *ŋə* 2SG) do not have ergative forms (except in hybrid indirect speech, see later), as illustrated by example (8), where the noun

TABLE 30.12 VOWEL FUSION IN STAU NOUNS

Base form	Meaning	Ergative	Genitive
<i>kəta</i>	dog	<i>kəta-w</i> → <i>kətow</i>	<i>kəta-j</i> → <i>kətej</i>
<i>vdzi</i>	man	<i>vdzi-w</i> → <i>vdzu</i>	<i>vdzi-j</i> → <i>vdzi</i>
<i>xə</i>	hybrid of yak and cow	<i>xə-w</i> → <i>xu</i>	<i>xə-j</i> → <i>xi</i>

waqi-w ‘the rabbit’ takes the ergative while the pronoun *ŋa* remains invariable in exactly the same context.

- (8) *ŋa zəŋəs qe-w tə-jə-sə ŋə-rə.*
1SG first shoot-1SG→3 PFV-say-IFR be-SENS
teʰəge, waqi-w zəŋətə-f-qe-sə ŋə-rə.
then rabbit-ERG first PFV-INV-shoot-IFR be-SENS
'He said "I will (you) shoot". Then, the rabbit shot him first.' (The rabbit and the tiger, 21–2)

Some morphologically intransitive verbs, such as *rga* ‘like’ take an ergative argument, as in (9), where the experiencer is in the ergative, and the stimulus in the dative case, but the verb only agrees with the argument marked in the ergative. Thus, in example (9), the verb is zero-marked, because the ergatively marked argument *tə-w* is third person (see Table 30.4).

Morphological and syntactic transitivity in Stau should thus be treated separately as they do not necessarily match for all verbs.

- (9) *tə-w ŋa-gi rga-rə*
he-ERG I-DAT like-SENS
'(S)he likes me.'

All postpositions above can occur in the argumental structure of some verbs, including the instrumental *-kʰa* and the comitative *-pʰa*, which are selected by the verbs *mkʰə* ‘need, want?’ (see (10)) and *təə/dzə* ‘meet’ respectively.

- (10) *ŋa rji-kʰa mi-mkʰə-ã*
1SG horse-INSTR NEG:N.PST-want-1
'I don't want a horse.'

5 NOMINALIZATION

All productive nominalization markers in Stau are suffixes. They are commonly used to build relative and complement clauses.

We find the agentive nominalizer *-ŋkʰə*, that can be used with intransitive (*čə* ‘go’ → *čə-ŋkʰə* ‘the one who goes’) as well as transitive verbs (*rə* ‘buy’ → *rə-ŋkʰə* ‘buyer’). The verb loses all person morphology (including the inverse prefix *v-* – note that the finite third person form of *rə* ‘buy’ is *v-rə* ‘he buys’) and only the negative prefixes can be added.

The nominalizer *-lə* can be used with transitive verbs to designate the patient of the action, as in *fee* ‘tell’ → *fee-lə* ‘things that have been told’ or *ŋgə* ‘eat’ → *ŋgə-lə* ‘food’. Alternatively, it can build action nominals with either transitive or intransitive verbs, as in example (11).

- (11) *nə-nə nə-vi-lə de tə-mpʰjə-sə ŋə-rə*
2-PL DOWN-go-NMLZ:P/ACTION DEM PFV-be.late-IFR be-SENS
'You arrived there too late' (=your arrival there was late; Akhustonba and the walnut tree, 34).

The suffix *-re* is used for locative nominalization, as in *rkʰə* ‘put in’ → *pʰjodzə* *rkʰə-re* (money put.in-NMLZ:LOC) ‘wallet, place one puts money in’. It also appears with stative existential verbs such as *tə* ‘be there, exist’ → *tə-re* (be.there-NMLZ:LOC) ‘the place where (it) is’.

Finally, *-sce* is the marker of instrumental nominalization, as in *ra* ‘write’ → *ra-sce* ‘pen, the thing one uses to write’.

Relatives in Stau are always built using one of these four suffixes. The S/A argument is relativized with *-ŋk'ø*, the P argument with *-lø*, and oblique arguments or adjuncts with the other two suffixes.

In relative clauses with relativized P, the agent of the relative can be marked either with the ergative or the genitive. In example (12), it is possible to say *akəstəmba-w* with the ergative *-w* instead.

- (12) [akəstəmba-*j* *fee-lø*] *de* *vdē-ndzø?*
Akhustonba-GEN tell-NMLZ:P/ACTION DEM be.true-SENS
'Does Akhustonba tell the truth?' ('It is true what Akhustonba says', Akhustonba and the walnut tree, 25)

6 COMPLEMENTATION

Most complement clauses have a non-finite verb, suffixed with the *-lø* or *-re* nominalizer.

The modal verb *rø* 'be able' is an example of the first type. As shown by example (13), where the verb of the complement clause *scəqi-lø* is in non-finite form, suffixed with *-lø*, while *rø* takes person and TAM marking.

- (13) *tcʰəge ya le røsa scəqi-lø*
then 1SG TOP immediately look.at-NMLZ:P/ACTION
ma-rø-w-sə
NEG:PFV-be.able-1SG→3-IRR
'I was not able to notice it immediately.' (The hybrid yak, 52)

This type of complement is also found with the phasal complex predicate *ŋgə ftsu* 'start', borrowed from Tibetan *ngo btsugs* 'start' (see example (14)).

- (14) *ŋa zl̥də-lø ŋgə kə-ftsu-w*
1SG boil-NMLZ:P/action start PFV-start-1SG→3
I started boiling it.

The locative nominalizer *-re* appears for instance in combination with *l̥di* 'come' to express the meaning 'be ready to, have almost ...', as in (15).

- (15) *ste-re l̥di-ã*
finish-NMLZ:LOC come-1
'I have almost finished.'

Other types of non-finite complement clauses are also attested. The verb *ŋgə* 'help' takes a complement clause whose main verb is marked with the suffix *-ŋgə* (homophonous with the bare stem of the verb, and probably derived from it), as in (16). No other verb can occur with this type of complement clause.

- (16) *tø-w ya-gi tcədə de zl̥da-ŋgə kə-ŋgə-ã*
3-ERG 1SG-DAT book DEM read-CONV PFV-INV:help-1
'He helped me read this book.'

In the synthetic causative construction, the verb *xts'ø* 'let' takes a complement verb in bare stem form, without any person or TAM marker, and no nominalization or converbial suffix. Example (17) illustrates this construction (in this example, the inverse *-f-/v-* cannot surface due to the initial cluster). If the verb in the complement clause is intransitive, its S is coreferent with the P of *xts'ø* 'let'; if it is transitive, its A is coreferent with the P.

- (17) [e-ze nə] tə-xtə^hə-ã
 one-moment rest IMP-INV:CAUS-1
 'Let me rest a moment.' (The hybrid yak, 48)

Finite complement clauses are found with both phasal auxiliary verbs such as ste 'finish' (as in (18); note that in this case, only the verb in the complement clause receives person indexation), and in reported speech.

- (18) [nəq^hej xtse ña tə-t^hi-w] ste-sə
 2SG:GEN soup 1SG PFV-drink-1SG→3 finish-IFR
 'I drank up your soup.' (The louse and the flea, 9)

Stau presents an extreme case of hybrid indirect (or semi-direct) speech (Aikhenvald 2008; Tournadre 2008).

When the original speaker is different from the current speaker, the reported utterance will only be identical to the original utterance if the current and original speaker and addressee are not referred to in it. Thus, in example (19), the reported utterance is identical to the original one (example 20).

- (19) tʂaci-w_i ña-gi jə-rə ge [tə-w_j dzəma de
 Bkrashis-ERG 1SG-DAT say-SENS LNK 3SG-ERG Sgrolma DEM
 nə-f-se-sə] jə-rə
 PRF-INV-kill-EVID say-SENS
 'Bkrashis, told me (that) he, (= Bkrashis) had killed Sgrolma.'

- (20) tə-w dzəma de nə-f-se(-sə)
 3SG-ERG Sgrolma DEM PRF-INV-kill(-IFR)
 'He killed Sgrolma.'

A different situation is observed when the original speaker refers to himself (in the first person) in the original utterance, as illustrated by (21) and (23). The reported utterance keeps the verb form of the original utterance (example (22)), but the first person pronoun is replaced by the logophoric pronoun ñə. Like the first person pronoun ña, the logophoric ñə does not take the ergative marker.

- (21) tʂaci-w_i ña-gi jə-rə ge /ədə_i dzəma de
 Bkrashis-ERG 1SG-DAT say-SENS LNK LOGOPHORIC Sgrolma DEM
 nə-se-w] jə-rə
 PRF-kill-1SG→3 say-SENS
 'Bkrashis_i told me that he_i had killed Sgrolma.'

- (22) ña dzəma de nə-se-w
 1SG Sgrolma DEM PFV-kill-1SG→3
 'I killed Sgrolma.'

- (23) [ədə-gi mbjɔ̄mbjɔ̄' tea gə tə-f-k^hə-ã, tea gə
 LOGOPHORIC-DAT quickly tea INDEF IMP-INV-give-1 tea INDEF
 tə-f-k^hə-ã,] jə-rə
 IMP-INV-give-1 say-SENS
 'He said "Give me some tea, give me some tea.'" (The hybrid yak, 39)

When the current speaker was referred to as a third person in the original utterance, a different type of mismatch occurs, illustrated by example (24). The verb preserves the form of the original utterance (20), but the third person pronoun is replaced by the first

person *ya*. In addition, in this case the pronoun takes the ergative flagging *-w* of the original utterance, although SAP pronouns normally do not take ergative suffixes.

- (24) *tṣaci-w jə-rə ge [ya-w dʐəma de*
 Bkrashis-ERG say-SENS LNK 1SG-ERG Sgrolma DEM
nə-f-se-sə] jə-rə
 PFV-INV-kill-IFR say-SENS
 ‘Bkrashis said that I had killed Sgrolma.’

We see that in Stau hybrid indirect speech verb forms and case marking represent the point of view of the original speaker, but everything else, including pronouns, represents that of the current speaker.

7 EVIDENTIALITY

Like most languages of the Tibetan cultural area, the Stau verbal system has evidential markers, in particular the sensory evidential *-rə* and the past inferential *-sə*.

The sensory *-rə* is generally used to express a non-past state or an action that the speaker is directly witnessing, be it by vision or by other senses. Thus, sentence (25) can be uttered by someone who sees (or feels, in a car) the state of the road.

- (25) *tce ke rce gə yə-rə*
 road very bad INDEF be-SENS
 ‘It is a bad road.’

The sensory *-rə* is not used in objective (i.e. based upon external observation) statements about the speaker himself: in (26), the verb *yə-ã* cannot take the suffix *-rə*. It is also not used with third person referents when the speaker considers the statement to belong to generally accepted ‘encyclopaedic’ knowledge.

- (26) *ya ke mbjɔv yjəra-ŋkʰə gə yə-ã*
 1SG very be.fast run-NMLZ:S/A INDEF be-1
 ‘I am a fast runner.’

On the other hand, as with the Tibetan sensory *ndug* (see Tournadre and LaPolla 2014), *-rə* can be used to express endopathetic sensations, knowledge (27) or desire (28) with the first person.

- (27) *ya sə rdzi-ka tceʰu xa*
 1SG TOP trace-ALL anything understand(1)
mi-gə-ã-rə
 NEG-understand(2)-1-SENS
 ‘Me, on the other hand, I do not understand anything about tracking (animals)’
 (The hybrid yak, 17)

- (28) *ya tea t'i-snə pre-rə*
 1SG tea drink-want want-SENS
 ‘I want to drink tea.’

The inferential *-sə* indicates that the speaker learnt of the facts in question second-hand (hearsay) or guessed them from indirect evidence, as in sentence (29).

- (29) *tṣaci-w dʐəma de nə-f-se-sə*
 Bkrashis-ERG Sgrolma DEM PFV-INV-kill-IFR
 ‘Bkrashis killed Sgrolma.’

A perfective sentence without the suffix *-sə* is used if the speaker has first-hand authoritative knowledge of the events described. For instance, (30) can only be said by someone who witnessed the crime.

- (30) *tṣaci-w džəma de nə-f-se*
 Bkrashis-ERG Sgrolma DEM PFV-INV-kill
 ‘Bkrashis killed Sgrolma.’

A verb in the inferential can be combined with the sensory form of the verb *ŋə* ‘be’ to form the narrative compound tense (see (31)), which is specifically used to tell stories. It is the most common verb form in narratives and other traditional stories.

- (31) *tṣaci-w džəma de nə-f-se-sə ŋə-rə*
 Bkrashis-ERG Sgrolma DEM PFV-INV-kill-IFR be-SENS
 ‘Bkrashis killed Sgrolma.’

8 CLASSIFICATION

There is clear evidence from morphology and lexicon that Stau and Khroskyabs languages constitute a subgroup within Rgyalrongic (evidence for the Rgyalrongic subgroup itself is presented in Sun 2000b).

First, both Stau and Khroskyabs languages have generalized the inverse forms in the non-local scenario and completely lost the direct 3→3' forms (Lai 2013b; Jacques and Antonov 2014; Lai 2015), a puzzling feature that is unlikely to be an independent innovation.

Second, Stau and Khroskyabs lost the nominalization prefixes found in Core Rgyalrong languages. Only indirect traces of the prefixes remain.

An example of such a trace is the noun *yji* ‘hole, orifice’ in Khanggsar Stau, *ŋjō* in Wobzi Khroskyabs, both cognate to Japhug *-yŋyu* ‘(its) hole, opening’, which derives from *nju* ‘open (intr)’, the anticausative of *cw* ‘open (a door)’.

In Japhug, *y-* is an irregular allomorph of the nominalization prefix *kui-*, found in a handful of examples (Jacques 2014: 4–6). Although the *y-* element in Stau and *ŋ-* in Khroskyabs are not analysable anymore (the forms *yji* and *ŋjō* are synchronically unmotivated), this example and others (see Jacques 2012b: 1228–9) suggest that nominalization prefixes used to exist in Stau and Khroskyabs, and were later replaced by innovative suffixes.

These suffixes, probably originally generic relator nouns, are partially shared between Stau and Khroskyabs (see Table 30.13), in particular the S/A nominalizer *-ŋkʰə* and the oblique nominalizer *-re*. Although contact cannot be excluded as a factor in the development of these suffixal systems, no discussion of Rgyalrongic subgrouping can neglect these data.

**TABLE 30.13 NOMINALIZATION SUFFIXES
 IN STAU AND KHROSKYABS**

	Stau	Khroskyabs
S/A	<i>-ŋkʰə</i>	<i>-pa, -ŋkʰə</i>
P, action	<i>-lə</i>	<i>-spi</i>
locative	<i>-re</i>	<i>-ri</i>
instrument	<i>-sce</i>	<i>-ri</i>

The suffix *-ŋkʰə* in Stau and Khroskyabs is reminiscent of Tibetan *-mkʰan*, which has the same use in some modern languages, but this is likely to be a coincidence (words with the rhyme *-an* in Classical Tibetan never correspond to words *-ə* in Tibetan loanwords in Stau).

Third, both Stau and Khroskyabs present a type of verbal reduplication unattested elsewhere, whereby the replicated syllable appears after the base, and its rhyme is replaced by *-a* (Lai 2013a). This process is productive in Stau and Khroskyabs (for instance *ŋga* ‘eat’ → *ŋgəŋga* ‘eat all kinds of things’).

Fourth, Stau and Khroskyabs languages share several lexical isoglosses which distinguish them from Core Rgyalrong languages, as illustrated by the data in Table 30.14. In these examples Khroskyabs and Stau differ from Japhug (and other Sino-Tibetan languages) and in some cases are clearly innovating. In the case of nouns like *mkʰə* ‘smoke’ for instance, the innovation is the replacement of the simple root by a compound comprising the original root and another element (fire+smoke → smoke, a well-attested unidirectional semantic change; see Urban 2011).

Stau and Khroskyabs are also characterized by a series of retentions not shared by Core Rgyalrong languages. Although these do not provide evidence for the Stau-Khroskyabs branch, they are nevertheless worth mentioning.

First, Stau and Khroskyabs have two distinct roots for ‘year’, one with the numeral prefixes (*-fku*) and the other in year ordinals (*-və*) (Table 30.15). Wobzi Khroskyabs *-dju* corresponds to Thugsrjuchenbo *-dyu* and is related to Stau *-fku* (the velar stop underwent lenition in Khroskyabs languages). Japhug and other Core Rgyalrong languages, on the other hand, have the root *-xpa* / *-pa* everywhere.

TABLE 30.14 POTENTIAL LEXICAL INNOVATIONS

	Stau	Khroskyabs	Japhug
heart	<i>zjar</i>	<i>sjâr</i>	<i>tuu-sni</i>
smoke	<i>mkʰə</i>	<i>mkʰó</i>	<i>tr-k'uu</i>
be big	<i>cʰe</i>	<i>cʰæ</i>	<i>wxti</i>
bread	<i>ləkʰi</i>	<i>laekʰi</i>	<i>qajyi</i>
writing	<i>teəðə</i>	<i>dəðdá</i>	<i>trscoz</i>
wind	<i>χpərju</i>	<i>χpərju</i>	<i>qale</i>
skin	<i>teəðza</i>	<i>dəðdzá</i>	<i>tuu-ndži</i>
water	<i>yṛə</i>	<i>jd̪ə</i>	<i>tuu-ci</i>
experience	<i>zdar</i>	<i>zdâr</i>	<i>rpo</i>
general classifier	<i>ə-lə</i>	<i>ə-lo</i>	<i>tuu-rdo</i>
human classifier	<i>a-ki</i>	<i>ə-kei</i>	<i>tuu-rdo</i>
exist (animate)	<i>ci/ji</i>	<i>jē</i>	<i>tu</i>
exist (be put on)	<i>stʰə</i>	<i>stî</i>	<i>tu</i>

TABLE 30.15 YEAR ORDINALS IN RGYALRONGIC LANGUAGES

	Stau	Khroskyabs	Japhug
one year	<i>e-fku</i>	<i>ə-dju</i>	<i>tuu-xpa</i>
two years	<i>yṇə-fku</i>	<i>jnə-dju</i>	<i>χnuu-xpa</i>
three years	<i>xsə-fku</i>	<i>çsə-dju</i>	<i>χsui-xpa</i>
last year	<i>javə</i>	<i>aχpî</i>	<i>japa</i>
this year	<i>pəvə</i>	<i>pîvi</i>	<i>yuijpa</i>

TABLE 30.16 CASE MARKERS IN STAU, KHROSKYABS AND TANGUT

Stau	Khroskyabs		Tangut
-w	ERG		𢃠 <i>ywu</i> ²
-j	GEN	-ji	𢃡 <i>jii</i>
-xa	ALL	-xa	𢃢 <i>ya</i> ²
-te ^h a	LOC		𢃣 <i>te^hjaa'</i>
-k ^h a	INSTR		𢃤 <i>k^ha'</i> in the middle of

A root suppletion in the noun ‘year’ cognate with the one in Khroskyabs and Stau is found in Tangut, Lolo-Burmese and Naish (see Jacques and Michaud 2011), showing that Core Rgyalrong languages are innovative here.

Second, Khroskyabs and Stau preserve the prohibitive dental prefix (*tɔ-* in Wobzi, Lai 2013b: 130–1, *di-* in Stau) that does not exist in Rgyalrong languages (which instead have a prohibitive form *ma-*).

Third, orientation prefixes are placed *before* the negation in Stau and Khroskyabs, while they appear after it in Rgyalrong languages. Since the order negation–orientation is also found in related languages such as Tangut and Pumi (Jacques 2011b), Rgyalrong languages are most probably innovative here.

Although Stau clearly belongs to the Rgyalrongic branch of Sino-Tibetan, it presents interesting commonalities with Tangut, in particular the case marking system, as shown in Table 30.16.

Among these markers, the only potential cognate in Core Rgyalrong languages is the locative *-i* found in Situ (which appears in the predicative possessive construction; see Lin 1993: 328). This is one area where Stau is more conservative than Core Rgyalrong languages, which have borrowed their ergative and genitive markers from Tibetan (see Jacques 2016).

NOTE

- Glosses follow the Leipzig glossing rules. Other abbreviations used here include: ALL allative, FACT factual/assumptive, IFR inferential evidential, INSTR instrumental, INV inverse, SENS sensory evidential. We would like to thank Randy LaPolla, Graham Thurgood and Nicolas Tournadre for useful comments on previous versions of this work.

REFERENCES

- Aikhengvald, Alexandra. 2008. Semi-direct speech: manambu and beyond. *Language Sciences* 30. 383–422.
- Duorji. 1998. *Daofuyu Geshizhahua yanjiu* 道孚语格什扎话研究 (A Study of Dgebshe Stau). Beijing: Zhongguo Zangxue chubanshe.
- Huang, Bufan. 1991. Daofuyu 道孚语 (The Rtau language). In Qingxia Dai, Bufan Huang, Ailan Fu, Renzeng Wangmu and Ju Liu (eds), *Zangmianyu shiwuzhong* 藏缅语十五种 (Fifteen Tibeto-Burman Languages), 1–45. Beijing: Yanshan chubanshe 北京:燕山出版社., 戴庆夏、黄布凡、傅爱兰、仁增旺姆、刘菊.
- Huang, Chenglong. 2013. Zangmianyu cunzai leidongci gainian jiegou 藏缅语存在类动词的概念结构 (The conceptual structure of the existential verbs in Tibeto-Burman). *Minzu Yuwen* 2. 31–48.

- Jacques, Guillaume. 2011a. A panchronic study of aspirated fricatives, with new evidence from Pumi. *Lingua* 121.9. 1518–1538.
- Jacques, Guillaume. 2011b. The structure of the Tangut verb. *Journal of Chinese Linguistics* 39.2. 419–441.
- Jacques, Guillaume. 2012a. Argument demotion in Japhug Rgyalrong. In Katharina Haude and Gilles Authier (eds.), *Ergativity, Valency and Voice*, 199–226. Berlin: Mouton de Gruyter.
- Jacques, Guillaume. 2012b. From denominal derivation to incorporation. *Lingua* 122.11. 1207–1231.
- Jacques, Guillaume. 2013. Applicative and tropative derivations in Japhug Rgyalrong. *Linguistics of the Tibeto-Burman Area* 36.2. 1–13.
- Jacques, Guillaume. 2014. Denominal affixes as sources of antipassive markers in Japhug Rgyalrong. *Lingua* 138. 1–22.
- Jacques, Guillaume. 2016. From ergative to comparee marker: multiple reanalyses and polyfunctionality. *Diachronica* 33.1. 1–30.
- Jacques, Guillaume and Alexis Michaud. 2011. Approaching the historical phonology of three highly eroded Sino-Tibetan languages: Naxi, Na and Laze. *Diachronica* 28.4. 468–498.
- Jacques, Guillaume and Anton Antonov. 2014. Direct/inverse systems. *Language and Linguistics Compass* 8/7. 301–318.
- Jacques, Guillaume, Anton Antonov, Yunfan Lai and Lobsang Nima. 2014. Person marking in Stau. *Himalayan Linguistics* 13.1. 82–92.
- Lai, Yunfan. 2013a. Erehua de fuyin chongdie 俄热话的辅音重叠 (Consonant reduplication in Wobzi). *Minzu Yuwen* 6. 12–18.
- Lai, Yunfan. 2013b. *La morphologie affixale du lavrungwobzi*. Université Paris III MA thesis.
- Lai, Yunfan. 2014. Causative constructions in Wobzi. In *14th International Symposium on Chinese Languages and Linguistics*, 319–336. Taipei: Academia Sinica. Available at: iscll-14.ling.sinica.edu.tw/files-pdf/Papers/Session7/Lai.pdf, accessed 13 August 2016.
- Lai, Yunfan. 2015. The person agreement system of Wobzi Lavrung (Rgyal-rongic, Tibeto-Burman). *Transactions of the Philological Society* 113.3. 271–285.
- Lin, Xiangrong. 1993. *Jiarongyu yanjiu 嘉戎語研究 (A Study on the Rgyal-rong Language)*. Chengdu: Sichuan minzu chubanshe.
- Sun, Hongkai. 1983. Liujiangliuyu de minzu yuyan jiqi xishu fenlei 六江流域的民族语言及其系属分类 (Minority languages of the Six River Valley and their genetic classification). *民族学报 Minzu Xuebao* 3. 99–273.
- Sun, Jackson T.-S. 2000a. Parallelisms in the verb morphology of Sidaba rGyalrong and Lavrung in rGyalrongic. *Language and Linguistics* 1.1. 161–190.
- Sun, Jackson T.-S. 2000b. Stem alternations in Puxi verb inflection: toward validating the rGyalrongic subgroup in Qiangic. *Language and Linguistics* 1.2. 211–232.
- Sun, Jackson T.-S. 2007. Morphological causative formation in Shanghai Horpa. *Bulletin of Chinese Linguistics* 1.2. 207–227.
- Sun, Jackson T.-S. and Qianzi Tian. 2013. Chuanxi Huoeryu Gexihua dongci duixie chutan 川西霍爾語格西話動詞對協初探 (Verb agreement in Gexi Horpa). *Bulletin of Chinese Linguistics* 7.2. 221–241.
- Tournadre, Nicolas. 2008. Arguments against the concept of ‘conjunct’/‘disjunct’ in Tibetan. In Brigitte Huber, Marianne Volkart and Paul Widmer (eds.), *Festschrift für*

- Roland Bielmeier zu seinem 65. Geburtstag, 281–308. Halle: International Institute for Tibetan and Buddhist Studies.
- Tournadre, Nicolas and Randy LaPolla. 2014. Towards a new approach to evidentiality. *Linguistics of the Tibeto-Burman Area* 37.2. 240–262.
- Urban, Matthias. 2011. Asymmetries in overt marking and directionality in semantic change. *Journal of Historical Linguistics* 1.1. 3–47.
- Wang, Stephen S. 1970–1. Consonantal clusters of Tibetan loanwords in Stau. *Monumenta Serica* 29. 631–658.

CHAPTER THIRTY-ONE

JAPHUG¹

Guillaume Jacques

1 INTRODUCTION

Japhug (*kuruuskṛt*) is one of the four Rgyalrong languages. It is spoken in Gdong-brgyad (*βduŋjṛt*), Gsar-rdzong (*sarndzu*), and Da-tshang (*tats'hi*) areas in Mbarkhams (*mbarkʰom*) county, Sichuan province, China, by around 10,000 speakers, and presents some dialectal diversity, especially in the Gdong-brgyad area.

Documentation on this language includes various articles, a short grammar with glossary (Jacques 2008), a text collection (Jacques and Chen 2010) and a dictionary (Jacques 2015a).

Like all Rgyalrongic languages, Japhug has undergone considerable lexical influence from Tibetan (Jacques 2004), but this influence is less conspicuous in the domain of morphosyntax.

2 PHONOLOGY

Japhug has complex onsets with 50 consonant phonemes and at least 414 consonant clusters (314 biconsonantal and 100 triconsonantal clusters), of which 63 are only attested in ideophones (Jacques 2013c) and 30 in Tibetan loanwords. Among the most unusual clusters in Japhug, we find palatal stop + *y* (*lyar* ‘hang on sth’), clusters with *j*– as first element (*jpum* ‘thick’) and fricative + prenasalized stop (*zmbru* ‘boat’).

Only 12 consonants appear in coda position, and no clusters are allowed: *-p*, *-β*, *-m*, *-t*, *-z*, *-n*, *l*, *-r*, *-j*, *-y*, *-η*, *-χ*. The stop *-p* is only attested in a few ideophones, and the only final stop in non-ideophonic vocabulary is *-t*. Final voiced segments, whether sonorant or obstruents, are devoiced before a pause but voicing surfaces in sandhi.

Japhug differs from all other Rgyalrong languages in that it has lost tonal contrasts. The only suprasegmental feature is the word stress, which is always word final except in the case of a few stress-attracting prefixes (the inverse *wyú-*, the negative sensory evidential *múj-*, and the comitative *kṛ-*). Verb suffixes are always unstressed (and can be realized with voiceless vowels), and stress in verb forms is either on the stress-attracting prefixes or on the last stem syllable.

Japhug dialects vary as to their exact number of vowels. The Kamnyu dialect has eight vowel phonemes *a*, *e*, *i*, *o*, *u*, *r*, *w*, and *y*, the latter only occurring in the word *qay* ‘fish’ and verbs derived from it and in Chinese loanwords.

Partial reduplication in Japhug provides important insight into syllabic structure (Jacques 2007). When partial reduplication is applied to a syllable, the rhyme of the replicated syllable is changed to *u* in the replicant.

Some clusters are affected by partial reduplication: when the last consonant of a cluster is one of the non-nasal sonorants (*r*, *l*, *j*, *w*, *y* or *χ*), and the preceding consonant is neither a sonorant nor a sibilant fricative, the sonorant is deleted, as in example (1).

(1) *prax* ‘cut, break’ → *pui-prax*

When the penultimate consonant of the cluster is a sonorant and the last consonant is one of *r*, {*l*, *y*, *v*}, this last consonant is not deleted.

- (2) *βraꝝ* ‘attach’ → *βru-βraꝝ*

The contrast between (1) and (2) shows that the *r* in these clusters does not have the same phonological status: in the first case, it forms a phonological constituent with the rhyme.

3 VERBAL MORPHOLOGY

3.1 Verb stems

Japhug verbs present stem alternations determined by TAM and person/number marking.

Verbs have at most three distinct stems, which we designate as stem 1, stem 2 and stem 3 following Sun (2000). Stem 1 is the default stem. Stem 2 appears only in the perfective and past imperfective forms. It is very residual in Japhug, attested only in a limited number of verbs indicated in Table 31.1.

Stem 3 in Japhug on the other hand is fully productive, and always appears in the forms 1SG→3, 2SG→3 and 3SG→3' of non-past TAM categories (factual, imperfective, sensory, egophoric present) for transitive verbs. Stem 3 does not appear in verb forms with the inverse marker (see later). Intransitive verbs lack stem 3 alternation.

Table 31.2 presents all stem 3 alternations attested in various Japhug dialects. The vowel alternation applies to the last vowel (in open syllables only) of the verb stem. In the case of verb stems ending in *-u* and *-uu*, some Japhug dialects allow two possible alternations, some verbs displaying vowel fronting and other addition of the *-m* suffixal element (as in Datshang, see Lín and Luóerwǔ 2003), while other dialects have generalized vowel fronting.

**TABLE 31.1 STEM 2 ALTERNATIONS IN
JAPHUG RGYALRONG**

stem 1	meaning	stem 2
<i>ee</i>	to go (vi)	<i>ari</i>
<i>suxee</i>	to send (vt)	<i>sryri</i>
<i>yi</i>	to come (vi)	<i>ye</i>
<i>ti</i>	to say (vt)	<i>tut</i>

**TABLE 31.2 STEM 3 ALTERNATIONS IN
JAPHUG RGYALRONG**

stem 1	stem 3	type
<i>-a</i>	<i>-e</i>	vowel fronting
<i>-u</i>	<i>-e</i>	vowel fronting
<i>-u</i>	<i>-rm</i>	<i>-m</i>
<i>-uu</i>	<i>-i</i>	vowel fronting
<i>-uu</i>	<i>-um</i>	<i>-m</i>
<i>-o</i>	<i>-rm</i>	<i>-m</i>

3.2 Orientational prefixes

All finite verb forms except the FACTUAL have orientational prefixes belonging to one of the four series in Table 31.3 (some Japhug dialects have co- instead of *pjr-* in series D for the ‘down’ prefix, see Lín and Luóérwǔ 2003). The non-periphrastic TAM categories are indicated in Table 31.4. Some of these categories require one particular orientational prefix (sensory, egophoric present, past imperfective, evidential imperfective). The other TAM categories only require a prefix of either set A, B, C or D, but the direction depends on the verb.

Most verbs have one intrinsic direction which is lexically determined, and used in the perfective, imperfective, evidential, irrealis and imperative. For instance, the verb *ndza* ‘eat’ selects the ‘up’ series of orientational prefixes for all its forms: **perfective** 1SG→3SG *tr-ndza-t-a* ‘I ate it,’ **imperfective** (with stem 3) *tu-ndze* ‘He eats it,’ **perfective** 3SG→3’ *ta-ndza* ‘He ate it’ and **evidential** *to-ndza* ‘He ate it.’

In addition to the orientational prefixes, two additional TAM markers appear in specific forms. First, the past –t suffix (–z in some dialects of Japhug) is found in the 1SG→3 and 2SG→3 forms of transitive verbs whose last stem syllable is an open syllable as *tr-ndza-t-a* PFV-eat-PST-1SG ‘I ate it.’ Second, the evidential circumfix *ku-....ci* is required whenever set D orientational prefixes are used with a verb whose stem begins in *a-*, including verb with the progressive prefix *asu-* (see later). For instance, the evidential form of the verb *anba* ‘hide’ is *ko-k-^rnba^k-ci* IFR-EVD-hide-EVD ‘he hid.’

Verbs of motion and some verbs of concrete action can be associated with all seven series of prefixes to indicate the direction of the motion. The ‘no direction’ series of prefixes only occurs with motion verbs.

TABLE 31.3 ORIENTATIONAL PREFIXES IN JAPHUG RGYALRONG

	perfective (A)	imperfective (B)	perfective 3→3’ (C)	evidential (D)
up	<i>tr-</i>	<i>tu-</i>	<i>ta-</i>	<i>to-</i>
down	<i>p^uu-</i>	<i>pjuu-</i>	<i>pa-</i>	<i>pjr-</i>
upstream	<i>ly-</i>	<i>lu-</i>	<i>la-</i>	<i>lo-</i>
downstream	<i>t^hu-</i>	<i>c^huu-</i>	<i>t^ha-</i>	<i>c^hy-</i>
east	<i>kr-</i>	<i>ku-</i>	<i>ka-</i>	<i>ko-</i>
west	<i>nua-</i>	<i>jnu-</i>	<i>na-</i>	<i>nr-</i>
no direction	<i>j^r-</i>	<i>ju-</i>	<i>ja-</i>	<i>jo-</i>

TABLE 31.4 FINITE VERB CATEGORIES IN JAPHUG RGYALRONG

		stem	prefixes
factual	FACT	1 or 3	no prefix
imperfective	IPFV	1 or 3	B
perfective	PFV	2	A or C
past imperfective	PST.PFV	2	<i>p^uu-</i>
inferential	IFR	1	D
inferential imperfective	IFR.PFV	1	<i>pjr-</i>
sensory	SENS	1 or 3	<i>jnu-</i>
egophoric present	PRES	1 or 3	<i>ku-</i>
irrealis	IRR	1 or 3	<i>a-</i> + A
imperative	IMP	1 or 3	A

The simple TAM categories in Table 31.4 can be combined with other TAM markers such as the conative *ji-* and the progressive *asu-*/*azz-*. The prefix *asu-* can only be used with transitive verbs. Verbs prefixed with *asu-* lose some transitivity features, such as stem alternation and past *-t*-suffix, but can still take inverse and 1→2 / 2→1 markers. The inverse prefix is *infixed* within the progressive prefix, as in the form *pjr-k-í<wy>z-nrjo-ci IPFV.IFR-EVD-PROG<INV>-wait.for-EVD* ‘he was waiting for him.’

In addition, periphrastic TAM categories, combining simple verb forms with the copula *yu* ‘be,’ *mas* ‘not be’ or *eti* ‘be (affirmative)’ are also attested. The most common periphrastic tenses are the periphrastic past imperfective and periphrastic inferential imperfective, built by combining the imperfective form of the verb with the past imperfective *puu-yu* and inferential imperfective *pjr-yu* of the copula. Most dynamic verbs cannot be used in the simple past and evidential imperfective in Japhug (see Lin 2011) and periphrastic tenses are required as in example (3): directly combining the prefix *pjr-* with the stem of the verb ‘eat’ results in an incorrect form. The past and inferential imperfective of dynamic verbs is attested only in counterfactuals and in combination with the progressive *asu-* (see Jacques 2014a: 297–301).²

- (3) *tu-ndze pjr-yu*
IPFV-eat[III] IPFV.IFR-be
‘He was eating it. / He used to eat it.’
- (4) *pjr-k-rsuu-ndza-ci*
IPFV.IFR-EVD-PROG-eat-EVD
‘He was eating it.’

3.3 Person marking

Japhug person marking is encoded by a series of prefixes and suffixes, in addition to stem 1 / stem 3 alternations in the case of transitive verbs.

Table 31.5 presents the regular transitive and intransitive verbal paradigms in Japhug in the FACTUAL. The symbols Σ_1 and Σ_3 represent stem 1 and stem 3 respectively. The affixes found in the intransitive paradigm (the second person *tu-* and the suffixes) also appear in the transitive paradigm, in addition to portmanteau prefixes *kui-* and *ta-* for the local scenarios 2→1 and 1→2) and the inverse prefix *wyu-* which is obligatory in 3→1 and 3→2 mixed scenario and cannot occur in local and 1→3 and 2→3 scenarios. When both arguments are third person, the use of the inverse is determined by semantic and pragmatic factors (see Jacques 2010).

The sensory existential copulas *yrtzu* ‘exist’ and *maye* ‘not exist’ are extremely irregular. They cannot be used with orientational prefixes, unlike nearly all other verbs, cannot be nominalized, and have infixed forms for the second person *yrtzru* ‘you are there’ and *matayre* ‘you are not there.’³

3.4 Evidentiality

Japhug has a very rich evidential system, which is typologically very similar to that of Tibetic languages.⁴ For reasons of space, we only discuss the evidential contrasts in the present contexts in this chapter (in addition, Japhug has distinct markers for past perfective and past imperfective evidential categories).

As in Tibetan, the evidential forms present a rule of anticipation (see Tournadre 2008 and Tournadre and LaPolla 2014), whereby the speaker uses in questions the forms that

TABLE 31.5 JAPHUG TRANSITIVE AND INTRANSITIVE PARADIGMS

	1SG	1DU	1PL	2SG	2DU	2PL	3SG	3DU	3PL	3'
1SG				<i>ta-</i> Σ_1 -nd <i>z̄i</i>	<i>ta-</i> Σ_1 -nd <i>z̄i</i>	<i>ta-</i> Σ_1 -mu	Σ_3 - <i>a</i>	Σ_3 - <i>a</i> -nd <i>z̄i</i>	Σ_3 - <i>a</i>	
1DU										Σ_1 - <i>ci</i>
1PL										Σ_1 - <i>ji</i>
2SG	<i>kau-</i> Σ_1 - <i>a</i>	<i>kau-</i> Σ_1 - <i>tei</i>	<i>kau-</i> Σ_1 - <i>ji</i>							Σ_1 - <i>ndz̄i</i>
2DU	<i>kau-</i> Σ_1 - <i>a</i> -nd <i>z̄i</i>									Σ_1 - <i>ndz̄i</i>
2PL	<i>kau-</i> Σ_1 - <i>a</i> -nuu									Σ_1 - <i>ndz̄i</i>
3SG	<i>wyūu-</i> Σ_1 - <i>a</i>	<i>wyūu-</i> Σ_1 - <i>tei</i>	<i>wyūu-</i> Σ_1 - <i>ji</i>	<i>wyūu-</i> Σ_1 - <i>ndz̄i</i>	<i>wyūu-</i> Σ_1 - <i>ndz̄i</i>	<i>wyūu-</i> Σ_1 -mu	Σ_3	Σ_3	Σ_3	
3DU	<i>wyūu-</i> Σ_1 - <i>a</i> -nd <i>z̄i</i>									Σ_1 - <i>ndz̄i</i>
3PL	<i>wyūu-</i> Σ_1 - <i>a</i> -nuu									Σ_1 - <i>ndz̄i</i>
3'										
INTR	Σ_1 - <i>a</i>	Σ_1 - <i>tei</i>	Σ_1 - <i>ji</i>	<i>tu-</i> Σ_1	<i>tu-</i> Σ_1 -nd <i>z̄i</i>	<i>tu-</i> Σ_1 -nuu	Σ_1	Σ_1 - <i>ndz̄i</i>	Σ_1 - <i>ndz̄i</i>	

he expects his addressee will employ in his answer. For this reason, I use the notation ‘1/2’ to describe forms appearing with the first person in affirmative sentences and with the second person in interrogative sentences.

In present tense contexts, we find three evidential categories: the factual (using the prefix-less verb form in stem III, without any auxiliary verb), the sensory evidential (marked by the prefix *nii-* or by suppletion in the case of existential verbs), and the egophoric present (*ku-*).

The factual is used in the present to describe facts considered by the speaker to belong to commonly accepted knowledge, as in (5) or present (and future) situations about which the speaker is fairly certain (example (6)). The factual can be used with all persons without strong constraints.

- (5) *uu-ku kuu-mpuu nii jnú-wy-p'ut tce, nuuya*
 3SG.POSS-head NMLZ:S/A-be.soft DEM IPFV-INV-pluck LNK cow
ra kuu ndza-nii, paꝝ kuu mꝝ-ndze
 PL ERG eat:FACT-PL pig ERG NEG-eat:FACT
 ‘One plucks the (leaves) on the extremities, the soft ones, the cows eat it, the pigs don't.’ (06 qaZmbri, 20)
- (6) *a-ye yuu uu-pyro ci <qiche>*
 1SG.POSS-grandson GEN 3SG.POSS-present INDEF car
kuu-xtei ci to-xtuu. tee andi ra a-ta
 NMLZ:S/A-be.small INDEF IFR-buy LNK west PL PASS-put:FACT
 ‘He bought a present, a small car for my grandson, it is there (at home).’
 (14 gongxun, 2–3)

The egophoric *ku-* is only used for 1/2 arguments (7), or third persons possessed by 1/2 (8). It is used to express intimate knowledge of an event or state on the part of the 1/2, not resulting from guess or recent information mediated through the senses. It cannot express a general or gnomic state of affairs, it is only used to refer to an ongoing state or action. It is virtually absent from procedural texts and narratives (expect in conversations quoted in the stories), but very common in conversations.

- (7) <kuabao> *uu-spa ci ku-taꝝ-a*
 bag 3SG.POSS-material INDEF EGOPH-weave-1SG
 ‘I am weaving a bag’ (conversation, 2014-10)
- (8) *nr-kr-nrma uu-kú-dxñ?*
 2SG.POSS-NMLZ:P-work QU-EGOPH-be.many
 ‘Do you have a lot of work (right now)?’ (2014.10 conversation, Chenzhen)

The sensory evidential⁵ is primarily used when the speaker's knowledge of a situation has been obtained through his senses (whether vision, audition, touch, etc.) and is not yet fully assimilated, as in (9).

- (9) *nrzo nr-skrt kuu-fse zo kuu-sna*
 2SG 2SG.POSS-voice NMLZ:S/A-be.like EMPH NMLZ:S/A-be.nice
nii piu-mts'am-a ndre, azo tce myzui a-zuuf
 DEM PFV-hear-1SG LNK 1SG LNK even.more 2SG.POSS-slumber
múj-yi ma jnu-mperr
 NEG:SENS-come because SENS-be.beautiful
 ‘After hearing your wonderful voice, I am feeling even less sleepy, as it is (so) beautiful.’ (140427 zhambeng he maotouying, 22)

Unlike the egophoric, the sensory evidential is not incompatible with general statements. In these contexts using the sensory rather than the factual expresses the speaker's lack of authority and confidence, for instance when describing animals that do not exist in Rgyalrong areas (as in (10)).

- (10) <banma> *nuu* *nuu-mperr*
 zebra DEM SENS-be.beautiful
 'The zebra is (a) beautiful (animal).' (20 RmbroN, 128)

The sensory is rarely used with the first person in affirmative sentences, except in the case of endopathic predicates (expressing pain, hunger, cold, etc.) and some cognitive verbs (such as *suso* 'think'). The sensory is also usable with endopathic predicates with third person referents (Japhug differs in this regard from Tibetan, see Tournadre and LaPolla 2014: 244).

In addition to evidential markers encoded in verbal morphology, the hearsay sentence final particle *k^hi* is used to indicate that the speaker's knowledge is based on a second-hand source. This particle most commonly appears with the sensory evidential, but examples with the factual also exist.

3.5 Derivation

Japhug is a very rich system of derivations, including voice markers, denominal, and deideophonic prefixes. Only the former will be discussed here. The system of voice derivation in Japhug is very similar to that described in Tshobdun (Sun 2014).

3.5.1 Argument demotion

In Japhug, non-overt arguments in a sentence are always construed as being definite. To express indefinite agent or patient, several strategies are possible, including indefinite pronouns such as *tuci* 'something,' generic person, and argument-demoting voice markers.

Japhug has a wide array of argument-demoting prefixes, including passive *a-*, anticausative, antipassive *sr-* and *rr-* and deexperiencer *sr-* (see Jacques 2012a, and also Sun 2014 on cognate markers in Tshobdun). The antipassive prefixes have been grammaticalized from denominal prefixes (Jacques 2014b).

The *a-* prefix is an agentless passive, that converts a transitive verb into an intransitive one whose S corresponds to the P of the base verb. The original A cannot be expressed, but is semantically recoverable, as in (11) with the light verb construction *srcu*–*lxt* 'lock.'

- (11) *u-ŋgu* *l̥-yi* *jry* *ma* *srcu*
 3sg-inside IMP:UPSTREAM-come be.possible:FACT because key
mr-a-lxt
 NEG-PASS-throw:FACT
 'Come in, (the door) is not locked.'

The anticausative is distinct from the agentless passive in that the action is construed as having occurred spontaneously without any agent. Only 24 pairs of verbs with anticausative derivation have been found in Japhug (see some examples in Table 31.6. Note the presence of the verb *χtr* 'spill' borrowed from Tibetan *gtor* (same meaning): this verb shows that the direction of derivation is from transitive to intransitive, not the other way round, since the corresponding intransitive verb *uṇḍr* 'be spilled' cannot be a Tibetan borrowing).

TABLE 31.6 EXAMPLES OF ANTICAUSATIVE IN JAPHUG

basic verb	derived verb		
<i>fisi</i>	to melt (vt)	<i>ndzi</i>	to melt (vi)
<i>kio</i>	to cause to drop	<i>ngio</i>	to slip
<i>pryt</i>	to break	<i>mbryt</i>	to be broken
<i>χt̪sr</i>	to spill	<i>χndxr</i>	to be spilled
<i>t̪saβ</i>	to cause to roll	<i>ndzab</i>	to roll (vi)
<i>cui</i>	to open	<i>jyuu</i>	to be opened

The semantic difference with the passive can be illustrated with *pj̪r-mbryt* IFR-ANTICAUS:break ‘it broke (of a rope)’ vs *pj̪r-k-χ-pryt-ci* IFR-EVD-PASS-BREAK-EVD ‘(some-one) broke it.’

There are two antipassive prefixes *rr-* and *sr-*, which convert a transitive verb into an intransitive one whose S corresponds to the A of the base verb. The prefix *rr-* is used in two cases: when the P of the base verb designates a non-human (*rvt* ‘write (tr)’ → *rr-rvt* ‘write (intr.)’, *car* ‘search, look for (tr)’ → *rr-car* ‘look for things (intr.)’) or in the case of ditransitive verbs, even when the P corresponds to the (human) recipient (*mbi* ‘give X to’ → *rrmbi* ‘give X to someone’). The antipassive *sr-* derives either a dynamic verb whose demoted argument is necessarily human (*sr-car* ‘look for people (intr.)’) or a stative verb, whose demoted P can be either human or non-human (*sat* ‘kill’ → *sr-sat* ‘have a killing power, be deadly’).

A historically related derivation is the deexperiencer *sr-*, which derives a stative verb from any stative verb, changing an experiencer S into the stimulus) (*ngio* ‘slip’ → *sr-ngio* ‘be slippery’).

3.5.2 Incorporation

Japhug has an incorporation-like construction in which noun-verb nominal compounds are turned into verbs by means of a denominal prefix (Jacques 2012b). For instance, from the noun *cui* ‘stone’ and the verb *p̪ut* ‘pluck, take out’ one can derive an action nominal *cui-p̪ut* ‘clearing the stones,’ which can in turn be made into an incorporating verb by denominal derivation *yuu-cui-p̪ut* ‘take out stones (out of the field).’

- (12) (i) *cui-p̪ut* *nui-βzu-t-a*
 stone-clearing PFV-do-PST-1SG
 (ii) *nui-yuu-cui-p̪ut-a*
 PFV-DENOMINAL-stone-take.out-11SG
 (iii) *cui* *nui-p̪ut-a*
 stone PFV-take.out-1SG
 ‘I cleared the stones (from the field).’

The three constructions above have slightly different meanings: the compound action nominal with light verb construction (i) implies that the action took considerable time, while the incorporating construction (ii) cannot be used if the action refers to specific stones that have been previously mentioned in discourse, contrary to (iii).

3.5.3 Causative

There are two productive causative prefixes in Japhug, *sui-* and *yr-* (Jacques 2015b).

The causative *sui-* prefix is extremely productive, and can be added to nearly any verb, including recent borrowings from Tibetan and Chinese. It presents considerable

allomorphy, and numerous irregular forms. It has four regular allomorphs *sui-*, *suy-*, *sux-*, and *z-* depending on the following element. The *z-* allomorph appears before all prefixes with sonorant initial (such as the denominal *nui-*, *ru-*, *yr-* prefixes, the antipassive *rr-*, etc.). The allomorphs *suy-/sux-* only appear with intransitive verbs stems with an onset containing neither a velar or a consonant cluster. The *sui-* allomorph appears in all other cases, and is considered here as the default form.

The causative *sui-* expresses very broad causative meanings, including indirect causative (including by one's inaction), permissive ('allow to') and is also (optionally) used with instruments marked in the ergative as in (13).

- (13) *u-χto nui mbruutcuu kuu kú-wy-sui-rk^he*
 3SG.POSS-slit DEM knife ERG IPFV-INV-CAUSE-carve
 'The slit is carved with a knife.' (Colored belts 13)

Causative verbs with irregular allomorphs such as *jts^hi* 'give to drink' from *ts^hi* 'drink' generally have unpredictable semantics. The regular form is always possible at least to express an action with an instrument *sui-ts^hi* 'drink with'.

When the causative *sui-* is applied to a transitive verb, whenever the causee (the S/A of the original verb) or the P (of the original verb) is first or second person, it is always encoded on the verb, resulting in ambiguous forms. As an example, the causative of *qur* 'to help' *sui-qur* can have two meanings in specific cases (14).

- (14) *tr-kuu-sui-qur-a-ndzi*
 PFV-2→1-CAUS-help-1SG-DU
 'You_d caused me to help him.' OR 'You_d caused him to help me.'

The causative *yr-* is more restricted, and can only occur with some stative verbs, such as *wxti* 'be big' → *yrwxti* 'make bigger.' Although Sun (2006, 2014) reports a semantic contrast between the cognate prefixes *sə-* and *wə-* in Tshobdun, this contrast appears to have been lost in the variety of Japhug under study.

Some labile verbs, such as *mto* (which means 'see' as a transitive verb and 'have the ability to see' as an intransitive stative verb) have distinct causative forms depending on the base meaning: *sui-mto* 'cause to see, show' is based on the transitive *mto*, while *yr-mto* 'cause (a blind person) to recover sight' is based on the stative *mto*.

3.5.4 Applicative and tropative

In addition to the causative, Japhug has two valency-increasing voice derivations, the applicative *nui-* / *nuy-* and the tropative *nr-* / *nry-* (Jacques 2013a).

The applicative converts an intransitive verb into a transitive one whose A corresponds to the S of the base verb and whose P can be either a stimulus or a recipient (for instance *mu* 'be afraid' → *nuy-mu* 'be afraid of').

The tropative is an extremely productive derivation that can be applied to all stative verbs describing a quality (i.e. it is a criterion for defining an adjective subclass among stative verbs), deriving a transitive whose P corresponds to the S of the base verb (like a causative) but whose A is an experiencer, not an agent. The tropative verb has the meaning 'to find / consider to be,' thus *xtei* 'be small' → *nr-xtei* 'to find small.'

3.5.5 Other voice markers

In addition to the voice markers described above, Japhug also has a reflexive *zyr-* distinct from the reciprocal (expressed by combining the prefix *a-* with a partially reduplicated verb stem).

We also find three verbal derivations with modal meanings: the abilitative *suu-*/*z-* (expressing the meaning ‘be able to X,’ *nrcqa* ‘bear’ → *z-nrcqa* ‘be able to bear’), the stative facilitative *yr-*(*wxti* ‘be big’ → *yr-wxti* ‘grow big easily’) and the passive facilitative *nuuyuu-* (*ntc^hoz* ‘use’ → *nuuyuu-ntc^hoz* ‘be easy to use’).

3.6. Autobenefactive-spontaneous

The autobenefactive-spontaneous *nuu-* prefix expresses a wide range of meanings. First, it is used to express that the S/A is affected by the action; it appears in particular with transitive verbs when the P is a body part of or belongs to the S/A (15).

- (15) *nr-ku puu-nuu-χtei*
 2SG.POSS-head IMP-AUTO-wash
 ‘Wash your head.’

The prefix *nuu-* can also be used to express involuntary actions, as in (16).

- (16) *u-qom ci pa-nuu-εluw ynu-ŋu*
 3SG.POSS-tear INDEF PFV:3→3'-AUTO-drop SENS-be
 ‘She shed a tear (involuntarily).’ (Kunbzang 228)

The prefix *nuu-* is also used to express an action occurring by itself, without exterior agent/causer, even if the action is voluntary and controllable, as in (17).

- (17) *azō pjuu-kuu-γrrat-a-nuu mr-ra ma azo*
 1SG IPFV:DOWN-2→1-throw-1SG-PL NEG-need:FACT because 1SG
pjuu-nuu-mtsaꝝ-a jiy
 NEG-IPFV:DOWN-AUTO-jump-1SG be.possible:FACT
 ‘You don’t need to throw me in there, I will jump of my own free will.’ (Nyima Wodzer 2002, 124)

Finally, *nuu-* also has a permansive meaning (translatable as ‘still’), as (18).

- (18) *te^heme nuu jnr-nuuk^hrdə ri, muu-pj^hr-^hrn, te^heme nuu*
 girl DEM IFR-convince LNK NEG-INFR-be.efficient girl DEM
pjr-nuu-γrwu eti
 IPFV(IFR-AUTO-cry be.AFFIRM:FACT
 ‘She (tried to) comfort the girl, but it was for nothing, the girl was still crying.’ (Bean and linen, 48)

The autobenefactive-spontaneous prefix has no influence on the verb transitivity.

3.7. Associated motion

Japhug has two associated motion prefixes *eu-* and *yuu-* expressing a motion taking place before the action expressed by the verb to which the prefixes are attached (Jacques 2013b). The translocative *eu-* and the cislocative *yuu-* respectively express the meaning ‘go to ..., go and ...,’ or ‘come to..., come and ...,’ as in example (19).

- (19) *t^hrrku kui-fse yuu-tu-ndze nura*
 crops NMLZ:S-be.like CISLOC-IPFV-eat[III] DEM:PL
mr-ŋgryl.
 NEG-be.usually.the.case:FACT
 ‘It does not come to eat crops and things like that.’

These prefixes are semantically close to the motion verb construction with *ce* ‘go’ and *yi* ‘come’ and a complement with the verb in the S/A nominalization form. A crucial difference, however, can be found in the perfective: motion verb constructions such as (20) imply that the motion event has taken place, but do not specify whether the action of the complement clause has taken place or not. Using the associated motion form *yuu-tr-tuu-nrma-t* instead would mean ‘What have you done? (since you came).’

- (20) *te^{hi}* *uu-kuu-nrma* *jy-tuu-ye?*
 what 3SG-NMLZ:S/A-do PFV-2-come[II]
 ‘What have you come to do?’

With causative verbs, the associated motion can refer either to the motion of the causer or that of the causee as in example (21).

- (21) *tee* *kupa* *chu* *nura* *athi* *pco^z* *nura*,
 LNK Chinese LOC DEM:PL downstream direction DEM:PL
ui-peⁱ *nura* *kuu* *kuureri*
 3SG-outside DEM:PL ERG here
yuu-chuu-sui-xtuu-nuu *ju.*
 CISLOC-IPFV:DOWNSTREAM-CAUS-buy-PL be;FACT
 ‘People from the Chinese areas, people from outside send people to come here to buy (matsutake and sell them in the areas downstream).’ (20 grWBgrWB 58)

3.8. Nominalization and other non-finite forms

Japhug has a rich set of non-finite verb forms: participles, action nominalizer, infinitives, and converbs. These forms have in common the facts that: (i) they can only mark person by means of possessive prefixes like nouns, cannot take the inverse prefix, and cannot encode two arguments in the case of transitive verbs; (ii) they cannot use stem 3; (iii) they cannot serve as the predicate of a main clause.

There are three participles in Japhug: the S/A-participle *ku-*, the P-participle *ky-* and the oblique participle *sr-*. In the case of transitive verbs, the S/A-participle takes a possessive prefix (see Table 31.7) coreferent with the P *ndza* ‘eat’ → *ui-kuu-ndza* ‘the one who eats it_i’), while the P-participle optionally takes a possessive prefix coreferent with the

TABLE 31.7 PRONOUNS AND POSSESSIVE PREFIXES

free pronoun	prefix	person
<i>azo, aj</i>	<i>a-</i>	1SG
<i>nyzo, nyj</i>	<i>nr-</i>	2SG
<i>uzzo</i>	<i>u-</i>	3SG
<i>teizo</i>	<i>tei-</i>	1DU
<i>ndzizo</i>	<i>ndzi-</i>	2DU
<i>zrni</i>	<i>ndzi-</i>	3DU
<i>izo, izora, izra</i>	<i>i-</i>	1PL
<i>nuzo, nuzora, nuzra</i>	<i>nu-</i>	2PL
<i>zara</i>	<i>nu-</i>	3PL
	<i>tua-, tr-</i>	indefinite
<i>tuzo</i>	<i>tua-</i>	generic

A *ui-kr-ndza* ‘what he/it, eats’). The oblique participle *sr-* can refer to the instrument, the place, time, or recipient of an action (see the section on relative clauses).

The participles can be used with negation and orientational prefixes (only set A and B); participles with perfective orientational prefixes take stem 2 if they have a distinct stem.

In addition, Japhug verbs have infinitives in *kr-* (for dynamic verbs allowing an animate argument) or *kua-* (for intransitive stative verbs or verbs only used with inanimate arguments). While the infinitives are superficially similar to the S/A- and P-participles, their morphology and uses are different. They cannot take any possessive prefix, and in the case of intransitive dynamic verbs, the forms in *kr-* cannot be interpreted as P-participles since such verbs only have an S argument and thus only a participle in *kua-*. Moreover, in some complements clauses, transitive verbs have a bare infinitive made of the bare stem 1 with a possessive prefix coreferent with the P (Jacques 2014b).

Finally, we find three converbs which are used in various clause linking constructions (see Jacques 2014a). First, the perfective converb is built by combining a set B orientational prefix (normally restricted to imperfective tenses!), the prefix *-tu-* and the verb stem 1, as in *mto* ‘see’ → *pju-tu-mto* ‘as soon as X saw Y’; it cannot receive any person marking. Second, the gerundive is made of the prefix *sr-* and a reduplicated verb stem, as in *mu* ‘be afraid’ → *sr-mu~mu* ‘while being afraid.’ Third, the purposive converb comprises a possessive prefix coreferent with any core argument (A, P, or S), a negative prefix (optional), a set B orientational prefix, the prefix *sr-* and a partially reduplicated verb stem 1, as in *jmut* ‘forget’ → *u-mr-nu-sr-jmu~jmut* ‘in order not to forget.’

3.9. Generic person

Apart from argument-demoting voice prefixes, another way of expressing indefinite arguments in Japhug is generic person marking. It is available only for generic human referents, not for inanimates or animals.

This type of person marking follows an ergative alignment: generic S and P arguments are marked with the prefix *kua-* (see examples (22) and (23)) while generic A arguments are marked with the inverse prefix *wy-* (24); the verb *ti* ‘say’ is irregular in that its generic A form *tu-kua-ti* takes the *kua-* prefix, see example (25)). Verbs with generic person markers cannot take any other person or number markers.

- (22) *tceri tr-prtso pui-kua-nu tce, nu kr-ndza*
but INDEF.POSS-child PST.IPFV-GENR:S/P-be LNK DEM INF-eat
wuma zo pui-kua-rga.
really EMPH PST.IPFV-GENR:S/P-like
'When (we) were children, (we) liked it a lot.' (12 ndZiNgri, 135)

- (23) *tce li nu turme kuunr ku-kua-nufse jnu-nu,*
LNK again DEM people also IPFV-GENR:S/P-recognize SENS-be
'(The monkey) recognizes people.' (19 GzW2, 17)

- (24) *turme kuu tu-wy-ndza mr-sna.*
people ERG IPFV-INV-eat NEG-be.fine:FACT
'People cannot eat it.' (11 paRzwamWntoR, 39)

3.10. Transitivity

The single most important feature in Japhug verbal morphology is transitivity. As seen earlier, transitive verbs differ from intransitive verbs in several ways: the S/A-participle

can take a possessive prefix coreferent with the P, the perfective third person forms take set C as orientational prefixes (or the inverse prefix), and verbs whose base stem is in an open syllable have the past *-t-* suffix in 1SG→3 and 2SG→3 in past forms and stem 3 alternation in non-past forms in some cases.

There is never any ambiguity as to whether a particular verb is morphologically transitive or intransitive, but some verbs, such as *taš* ‘weave’ or *murkuu* ‘steal’ are labile and can be conjugated either transitively or intransitively (Jacques 2012a). All labile verbs in Japhug are A-preserving; there are no examples of P-preserving lability.

There is a class of semi-transitive verbs, such as *rga* ‘like’ or *aro* ‘have, own’ that are morphologically intransitive but present some syntactic transitivity. While their finite conjugation is identical to that of intransitive verbs, and while their S does not take ergative marking, they do have P-participles in *kr-* and thus have an object-like argument that can be relativized exactly in the same way as the P of a transitive verb. This argument however, even if first or second person, cannot be marked in the verb morphology.

Ditransitive verbs can be either secundative (indexing the recipient like the P of a monotransitive verb, like *mbi* ‘give’), indirective (indexing the theme like the P of a monotransitive verb, and treating the recipient as an oblique not indexed in verb morphology, as *t^hu* ‘ask’) or both (causative verbs deriving from transitive base verbs) from the point of view of person marking on the verb. No more than two arguments can be encoded by verbal morphology. The noun phrases corresponding to recipients of indirective verbs receive dative flagging by means of the relator nouns *u-cki* or *u-p^he*, but cannot be marked on the verb, even if first or second person.

The relativization patterns, however, slightly differ from argument indexation on the verb. In the case of secundative verbs, both the R and the T of secundative verbs are relativizable in the same way as the P of a monotransitive verb, even though only the R is indexed on the verb. As for indirective verbs, the R is relativized by means of the oblique participle (see the section on relativization).

4 NOMINAL MORPHOLOGY

Japhug nouns can be divided into three sub-classes: simple nouns, inalienably possessed nouns, and nouns with numeral prefixes (classifiers). Only the first two classes are discussed in this sketch.

The same set of possessive prefixes (see Table 31.7) is used for all nouns, but inalienably possessed nouns cannot be used on their own without one of these prefixes. When there is no definite possessor, the indefinite possessive prefixes *tr-* or *tu-* are used. It is the citation form of inalienably possessed nouns (*tr-lu* ‘milk,’ *tu-ŋga* ‘clothes,’ *tr-rpuu* ‘uncle,’ *tu-ci* ‘water’). The choice of the prefix *tr-* vs *tu-* is lexically determined. When a specific possessor is present, the indefinite prefix is replaced by the appropriate possessive prefix (*u-lu* ‘her/its milk (from her nipple),’ *a-ŋga* ‘my clothes,’ *nr-rpuu* ‘your uncle,’ *u-ci* ‘its juice’). It is possible to turn an inalienably possessed noun into an alienably possessed one by prefixing a definite possessive prefix to the indefinite one (*u-tr-lu* ‘his milk (to drink),’ *u-tu-ci* ‘its water (of irrigated water, to a plant)'). Simple nouns cannot take indefinite possessive prefixes.

The indefinite possessive prefixes should not be confused with the generic possessive prefix *tu-*, which can be added to any noun, and which is coreferent with the argument marked with generic marking on the verb, as in example (25). Note also that inalienably possessed nouns that select the indefinite possessive prefix *tr-* have *tu-* instead when the possessor is generic (*tr-rpuu* ‘an uncle’ vs *tu-rpuu* ‘one’s uncle’).

- (25) *tui-rpuu* *uu-rjit* *uu-cki* *tee* *tee*
 GENR.POSS-uncle 3SG.POSS-offspring 3SG-DAT LNK LNK
“a-rpuu *a-lax”* *tu-kui-ti* *yu.*
 1SG.POSS-uncle 1SG.POSS-aunt IPFV-GENR-say be:FACT
 ‘One has to say “my maternal uncle, my maternal aunt” to one’s maternal uncle’s sons and daughters.’ (140425 kWmdza01, 69)

Some possessed nouns have restrictions on the interpretation of the possessor. Thus, the possessive prefix of the noun *tr-prro* ‘present’ can only refer to the person giving the present, never the recipient: *a-prro* 1SG.POSS-present can only mean ‘my present (to him, to you, etc.)’ not ‘the present (you, he gave me).’

Japhug has a very rich denominal morphology, and also has a productive comitative adverb derivation with the prefix *kryu-* or *k̥-* and reduplicated noun stem, as in *tr-rtas* ‘branch’ → *kryurturtas* ‘with branches.’ These adverbs probably originate historically from the S-participle of denominal propriety verbs with the prefix *ayu-* (*tr-rtas* ‘branch’ → *ayurturtas* ‘have many branches’).

5 THE NOUN PHRASE

Noun phrases in Japhug follow the general template given in (26), as illustrated by examples (27) and (28).

- (26) DEM-NOUN^{modifier}-NOUN^{head}-ADJ-NUM-DEM

The demonstratives *ki* ‘this’ and *nui* ‘that’ can either appear in the beginning or the end of the noun phrase, or be repeated at both ends.

- (27) [*ki* *a-teui* *χsum* *ki*]_{NP}
 this 1SG.POSS-son three this
 ‘These three sons of mine.’

Adjectival stative verbs can serve as noun modifier only in the S/A-participle (in *kuu-*), and the [noun+adjective] constituent should be analyzed as a head-internal relative clause.

- (28) [[*tc^hemehead* *kui-mpcui~mpcrr*]_{RC} *zo* *χsum*]_{NP}
 girl NMLZ:S/A-EMPH~be.beautiful EMPH three
nui-nui-lo&-nui
 IPFV-AUTO-come.out-PL
 ‘Three beautiful girls come out of it (each day).’

Nouns modifying other nouns normally appear before them as in (29). The head noun has an obligatory possessive prefix coreferent with the modifier noun, and a genitive postposition *yu* can be optionally added.

- (29) *azō* [*ndza&lay* *yū* *nui-rjy&pu*]_{NP} *yū-a* *tee,*
 1SG Jambudvīpa GEN 3PL.POSS-KING be:FACT-1SG LNK
 ‘I am the king of the people of Jambudvīpa.’ (2011-4-smanmi, 243)

There is a restricted subclass of possessed nouns referring to old or ragged objects such as *uu-do* ‘old (of animals),’ *uu-mbe* ‘old (of clothes),’ *uu-nqra* ‘shabby’ which are strictly postnominal. Although in translation they would correspond to adjectives, they are not noun modifiers syntactically—rather, it is a structure comparable to the marginal construction in English exemplified by forms such as ‘an old rag of a gown.’

- (30) [tu-rcu *uu-mbe* *ci*]_{NP} *to-ŋga*
INDEF.POSS-leather.jacket 3SG.POSS-old INDEF IFR-WEAR
'He wore an old leather jacket.'

Japhug is an exclusively postpositional language. Common postpositions include the ergative/instrumental *kuu*, the comitative *cʰo* 'with' and the locatives *zuu* and *tcu*. The locative postpositions have a vague locative meaning, and can indicate either a fixed position within or on the surface of an entity, or motion into or from it (they differ in that *zuu* cannot follow a demonstrative). Relator nouns are also similar to postpositions, but are morphologically possessed nouns with an obligatory possessive prefix, as with the dative *uu-cki* or *uu-pʰe* (usage depends on idiolects) or locative relator nouns such as *uu-ŋguu* 'inside' or *uu-taʂ* 'on,' and can always be optionally followed by the locative postposition *zuu*.

6 SIMPLE CLAUSES

Japhug has strict verb-final word order. Only a few elements can appear after a verb: sentence final particles, the adverb *ntsuu* 'always,' ideophones, and afterthought constituents.

Transitive sentences tend to have at least one covert argument (which is then always interpreted as being definite), but when all arguments are overt the canonical word order is A–P–verb (though P–A order is also attested, depending on information structure), the A being obligatorily marked with the ergative *kuu* unless it is a first or second person pronoun (see (31)).

- (31) *rjrlpu nuu kuu pyrteuu nuu uu-kayuu uu-ŋguu*
king DEM erg bird DEM 3SG.POSS-silver.necklace 3SG-inside
nuu tcu pjy-nuu-rku.
DEM LOC IFR:DOWN-AUTO-put.in
'The king put the bird in his necklace.' (2002qaCpa, 202)

Right-dislocated elements (afterthoughts) keep their postpositions or relator nouns, as in (32).

- (32) *sla tu-ngo yu tu-ti-nuu yu, kuruu ra kuu*
moon IPFV-be.sick be:FACT IPFV-say-PL be:FACT Tibetan PL ERG
'They say that the moon is sick, the Tibetans.' (29 mWBZi, 149)

7 COMPARATIVE

Japhug comparative constructions are illustrated by example (33): the standard is marked by the comparative postpositions *sʂz* or *staʂ*, while the comparee can optionally be followed by a marker *kuu* homophonous with the ergative/instrumental (Jacques 2016a).

- (33) *uu-ki sʂz uu-pi*
3SG.POSS-younger.sibling COMPARATIVE 3SG.POSS-elder.sibling
nuu kuu mpɛʂʂ
DEM ERG? be.beautiful:FACT
'The elder sister is more beautiful than the younger sister.' (Elicited)

There is a degree construction in which adjectival stative verbs are nominalized with the prefix *tu-* and take a possessive prefix coreferent with the S, followed either by a

finite verb indicating degree (most commonly *saxas* ‘be extremely ...’) or by a clause containing a simile describing the degree of the property.

- (34) *mtc^{hi}* *uu-mat* *rca* *uu-tuu-teur*
 sea.buckthorn 3SG.POSS-fruit TOP 3SG-NMLZ:DEGREE-be.sour
saxas. *uu-tuu-teur* *kuu*
 be.extremely:FACT 3SG-NMLZ:DEGREE-be.sour ERG?
[tuu-kur] *uu-ŋguu* *lú-wy-rku* *q^he*
 INDEF:POSS-mouth 3SG-inside IPFV:UPSTREAM-INV-put.in LNK
maka nuu-suu-ŋmuzyut *qhe,* *tuu-p^hoŋbu* *ra*
 at.all IPFV-CAUS-be.evenly.distributed LNK INDEF:POSS-body PL
kuunr nuu-sux-teur *kuu-fse* *etii]*
 also IPFV-CAUS-be.sour NMLZ:S/A-be.like be.AFFIRMATIVE:FACT
 ‘The fruit of the sea-buckthorn is very sour, so sour that when one puts it in one’s mouth, it makes it completely (sour), and it is as if one’s (whole) body became sour.’ (09 mi, 66)

Superlative can be expressed by the adverb *stu* ‘most,’ but the most common construction is a construction combining the negative existential copula with a nominalized predicate and an equative adjunct in *kuu-fse* ‘like ...’ as in (35).

- (35) *k^huna* *kuu-fse* *zo* *uu-ts^huxto_K* *kuu-tu*
 dog NMLZ:S/A-be.like EMPH 3SG.POSS-loyalty NMLZ:S/A-exist
me *k^hi*
 not.exist:FACT HEARSAY
 ‘The dog is the most loyal animal (there is no animal whose loyalty is like that of a dog). (05-khWna, 5)

8 RELATIVIZATION

Japhug relative clauses can be classified according to two main criteria: finiteness of the verb in the clause and place/presence of head noun.

Both head-internal (see example (36)) and prenominal relatives (37), (42) are found in Japhug. Head-internal relatives are possible for the relativization of core arguments and possessors, while prenominal relatives are required for all oblique arguments and adjuncts. Prenominal relatives are also possible for core arguments, except S in the case of stative verbs. In this section, the head nouns are indicated in bold.

Non-finite relative clauses have a verb in one of the three participles. The S/A-participle in *kuu-* is used for relativizing the S, the A, or the possessor of an argument, as in (36). Note that the presence of ergative marking on the A shows that the noun phrase *tr-prtso ci* ‘a boy’ belongs to this head-internal relative, as the main verb *jr-ye* ‘he came’ is intransitive.

- (36) [*tr-prtso* *ci* *kuu* <*yangma*> *u-kuu-numbrypu*]RC
 INDEF.POSS-child INDEF ERG bicycle 3SG-NMLZ:A-ride
ci *jr-ye*
 INDEF PFV-come[II]
 ‘A boy who was riding a bicycle arrived.’ (Pear story, Chenzhen, 5)

The P-participle in *kr-* is used for P (37), for the object of semi-transitive verbs, and either the T or the R of secundative ditransitive verbs.

- (37) [azo a-*mr-kx-suž*]RC *txjmry* *nū* *kx-ndza*
 1SG 1SG-NEG-NMLZ:P-know mushroom DEM INF-eat
mr-naz-a
 NEG-dare;FACT-1SG
 ‘I do not dare to eat mushrooms that I do not know.’ (23 mbrAZim, 103)

The oblique participle can occur in relatives whose relativized element is a time or place adjunct, an instrument (marked with the ergative/instrumental *kuu* in the original sentence), the recipient of an indirective verb (38).

- (38) [*u-sr-fçrt*]RC *pjr-me* *q^he tce*
 3SG-NMLZ:OBLIQUE-tell IPFV.IFR-not.exist LNK LNK
tr-prtso *u-eki* *nū tcu* *nura* *tch^hi*
 INDEF.POSS-child 3SG-DAT DEM LOC DEM:PL what
puu-kuu-fse *nura* *pjr-fçrt.*
 PST-NMLZ:S-be.like DEM:PL IFR-tell
 ‘She had no one (else) to tell it to, so she told the boy everything that had happened.’ (140515 congming de wusui xiaohai, 77)

Some verbs require oblique arguments marked with the comitative postposition *c^ho*, and these arguments are also relativized by a participle in *sr-* (example (39), a headless relative)

- (39) *tce* [*uzo u-sr-*amumi**]RC *nū d^hn*
 LNK it 3SG-NMLZ:OBLIQUE-be.on.good.terms DEM be.many:FACT
ma ca kuu-fse qazo kuu-fse, *ts^hrt*
 because musk.deer NMLZ:S-be.like sheep NMLZ:S-be.like goat
kuu-fse, uzo c^ho kuu-naxteuy *sujno, xcaj*
 NMLZ:S-be.like it with NMLZ:S-be.identical herbs grass
ma mr-kuu-ndza nū ra c^ho nū
 apart.from NEG-NMLZ:A-eat DEM PL with DEM
amumi-nū tce,
 be.on.good.terms:FACT-PL LNK
 ‘The (animals) that are on good terms with the rabbit are many, it is on good terms with those that only eat grass, like musk deer, sheep or goats.’ (04 qala1, 33-4)

Finite relative clauses are limited to the relativization of P, the object of semi-transitive verbs, the theme of indirect verbs (40), the recipient or the theme of secundative verbs, and also locative adjuncts in the case of motion/manipulation verbs (*ce* ‘go’ or *tsum* ‘take away,’ as in example (41)).

- (40) *nū* [*qajuu kuu-pa^h*]RC *tu-ti-a*]RC *nū nū* *kuu-fse*
 DEM worm NMLZ:S-black IPFV-say-1SG DEM DEM NMLZ:S-be.like
nū-βze nū-ŋu
 IPFV-grow SENS-be
 ‘The black worm that I was talking about grows like that.’ (28 kWpAz, 30)

- (41) [*k^ha j^h-wy-tsum-nū*]RC *nūnū, lonba [com kuu*
 house PFV-INV-take.away-PL DEM all iron ERG
nū-kx-suž-βzu]RC *k^ha pjr-ŋu*
 PFV-NMLZ:P-CAUS-make house IPFV.IFR-be
 ‘The house to which he had taken them, it was a house made of iron.’ (140505 liuhahuan zoubian tianxia, 148)

While inverse marking does not occur in non-finite relative clauses, it is possible in finite ones as in (41) and (42). Inverse marking has no effect on the accessibility to relativization of the arguments (in (42), the ditransitive verb *suxer̥t* is secundative, and the relativized element is the T). In particular, it does not make any argument or adjunct other than the P or the T (in particular, the A) accessible to relativization with a finite relative clause.

- (42) *tx-teu* *nū kuu*, [sayrjy_z ra kuu pū-wy-suxer̥t]RC *yū*
 INDEF.POSS-boy DEM ERG Buddha PL ERG PFV-INV-teach GEN
kʰynduun *nū w-rzaβ* χsum *nū* *pj̥r-suxer̥t*
 sūtra DEM 3SG.POSS-wife three DEM IFR-teach
 ‘The boy taught his three wives the sūtra that the Buddhas had taught him.’ (sloX-pWn, 354)

Finite relative clauses present some syntactic features distinguishing them from independent clauses (Jacques 2016b). First, the possessive prefixes in the sentence referring to a core argument can be neutralized to the indefinite possessive prefix. In (43) for instance, the first singular possessive form *a-pyro* ‘my present (to him)’ is the only possible form in the corresponding main clause, but the indefinite possessive form *tr-pyro* ‘a present’ can be used instead in the relative clause.

- (43) [tr-pyro *nū-mbi-t-a*]RC *tr-rjɪt* *nū*
 INDEF.POSS-present PFV-give-PST: TR-1SG INDEF.POSS-child DEM
a-teu *yū*
 1SG.POSS-son be: FACT
 ‘The child to whom I gave a present is my son.’ (Elicited)

Second, the verb of a finite relative clause can undergo totalitative reduplication (of the first syllable of the word), a morphological process normally restricted to nominalized verb forms that have a quantifying meaning ‘all’ (in reference to the relativized element, in this case normally the P).

- (44) *tee* [*nū* *ra* *t̥rrkusna* *nū* *pū~pa-yuut*]RC *nū*
 LNK DEM PL good.crops DEM TOTAL~PFV:3→3:DOWN-bring DEM
lo-ji-ndzi
 IFR-plant-DU
 ‘They^{du} planted all the crops that she had brought (from heaven).’ (07 deluge, 111)

Third, evidential marking is neutralized: the inferential cannot be used in relative clauses, only the perfective.

9 COMPLEMENTATION

In Japhug, complement clauses, like relative clauses, can be either finite or non-finite. Non-finite complement clauses include four types. Infinitive clauses, with the verb pre-fixed in *kr-*, are found with a wide range of verbs, including modal verbs (45), some phasal verbs, and causative verbs.

- (45) [*par* *ra* *kuu* *kr-ndza*] *wuma* *zo* *rga-nū*
 pig PL ERG INF-eat very EMPH like:FACT-PL
 ‘Pigs like to eat it.’ (12 ndZiNgri, 149)

Note that *par ra kuu* ‘pigs’ is marked with ergative *kuu* and thus belongs to the complement clause (whose main verb *ndza* is transitive) and is not an argument of the main clause (whose main verb *rga* is intransitive, and cannot take arguments in *kuu*).

S/A-participial clauses are used for the purposive complements of motion verbs and the complements of the verbs *zyṛpa* and *nuecpuz* ‘pretend’ and of the phasal verb *rṛyngat* ‘prepare for, be about to,’ as in (46).

- (46) *nyzo a-kuu-kndiu tu-tuu-rṛyngat tee,*
 2SG 1SG-NMLZ:S/A-hit IPFV-2-prepare LNK
nx-p^buŋgu rdrstas jnu-tuu-rke jnu-jnu
 2SG.POSS-folds.of.clothes stone IPFV-2-put.in[III] SENS-be
 ‘You have put a stone in your clothes, and are about to hit me.’ (2002qaCpa, 171)

A few verbs, such as *za* ‘start’ and *rño* ‘experience,’ take bare infinitive complements (47) when the complement verb is transitive. Unlike the two previous types of complement clauses, bare infinitival clauses appear to be unattested in Tshobdun and other Rgyalrong languages (Sun 2012).

- (47) *tce prjk^bu pjw-si eωŋgu zo u-u-ca*
 LNK yet IPFV-die before EMPH 3SG.POSS-meat
u-ndza tu-za-nu eti.
 3SG-BARE.INF:eat IPFV-start-PL be.AFFIRMATIVE:FACT
 ‘They start eating it before it has died.’ (20 sWNgi, 47)

Causative verbs derived from adjectival stative verbs can be used with a bare infinitive complement, the complement indicating the action and the causative verb the manner or degree in which the action is performed, as in (48) and in (49), the latter with the complex predicate *u-puu -pa* ‘preserve, keep, take care of’ in the complement clause.

- (48) *[c^ba uu-ts^bi] ko-yr-tc^bom*
 alcohol 3SG.POSS-BARE.INF:drink IFR-CAUS-be.too.much
 ‘He drank too much alcohol.’

- (49) *[u-puu u-pa]*
 3SG.POSS-CP:keep 3SG.POSS-BARE.INF:CP:keep
a-tr-tuu-yr-βdi ma
 IRR-PFV-2-CAUS-be.well LNK
 ‘You will have to keep it well’.

Note that the causative verb takes the orientational prefix normally selected by the complement verb: *ts^bi* ‘drink’ and *u-puu -pa* ‘preserve’ select the ‘east’ *ko-* *kr-* and ‘up’ *to-* *tr-* orientational prefixes, respectively.

No verb requires a bare infinitive complement. The phasal verb *za* ‘start’ can also be used with an action nominal in *tu-* (see Jacques 2014b: 6–9), and *rño* ‘experience’ and the causative verbs can also take a *kr-*infinitival complement.

Finite complement clauses appear with some speech, thought and modal verbs, but most verbs taking finite complements also accept infinitival complements.

Complement clauses do not necessarily occur as P-argument or adjuncts; some are used in S position, as noun complements, or in topical position as in (50).

- (50) *[taqaβ ci c^bū-wy-lrt] nuu tuu-k^byfitsuuy tu-kuu-ti*
 needle once IPFV-INV-throw DEM one-stitch IPFV-GENR-say
jnu
 be:FACT
 ‘Passing the needle once (in the fabric) is called “a stitch”?’ (12 kAtsxWb, 21)

Finite complement clauses, like relative clauses, can appear with inverse marking in Japhug, as shown by example (50) (used here in its value as a generic A marker; see section 3.9).

NOTES

- 1 I wish to thank Randy LaPolla, Laurent Sagart, Graham Thurgood, Nicolas Tournadre, and Theo Yeh for useful comments on this chapter. Glosses follow the Leipzig rules, to which the following are added: CISLOC cislocative, FACT factual, GENR generic, EGOPH egophoric, EMPH emphatic, IFR inferential, INV inverse, LNK linker, SENS sensory, TOTAL totalitative, TRANSLOC translocative. The examples are taken from a corpus that is progressively being made available on the Pangloss archive (Michailovsky *et al.* 2014). This research was funded by the HimalCo project (ANR-12-CORP-0006) and is related to the research strand LR-4.11 “Automatic Paradigm Generation and Language Description” of the Labex EFL (funded by the ANR/CGI).
- 2 The presence of the circumfix k-...-ci in example (4) is explained in the previous section.
- 3 In mataje the element ma- is related to the negative prefixes my- and mu-, but cannot be analyzed as such synchronically in this form, as there is no independent stem -ne.
- 4 Other Rgyalrong languages have similar systems, see in particular Lin (2003) on Situ.
- 5 This refers to the category that has been variously referred to as ‘constative,’ ‘mirative,’ or ‘testimonial’ in the literature (see Hill 2012). I here adopt Tournadre and LaPolla’s (2014) term.

REFERENCES

- Hill, Nathan W. 2012. “Mirativity” does not exist: hdug in “Lhasa” Tibetan and other suspects. *Linguistic Typology* 16.3. 389–433.
- Jacques, Guillaume. 2004. *Phonologie et morphologie du japhug (Rgyalrong)*. Université Paris-VII (Diderot) dissertation.
- Jacques, Guillaume. 2007. La reduplication partielle en japhug, révélatrice des structures syllabiques. *Faits de Langues* 29. 9–21.
- Jacques, Guillaume. 2008. 嘉绒語研究 *Jiāróngyǔ yánjiū (Study on the Rgyalrong Language)*. Beijing: Minzu chubanshe.
- Jacques, Guillaume. 2010. The inverse in Japhug Rgyalrong. *Language and Linguistics* 11.1. 127–157.
- Jacques, Guillaume. 2012a. Argument demotion in Japhug Rgyalrong. In Katharina Haude and Gilles Authier (eds), *Ergativity, Valency and Voice*, 199–226. Berlin: Mouton de Gruyter.
- Jacques, Guillaume. 2012b. From denominal derivation to incorporation. *Lingua* 122.11. 1207–1231.
- Jacques, Guillaume. 2013a. Applicative and tropative derivations in Japhug Rgyalrong. *Linguistics of the Tibeto-Burman Area* 36.2. 1–13.
- Jacques, Guillaume. 2013b. Harmonization and disharmonization of affix ordering and basic word order. *Linguistic Typology* 17.2. 187–217.
- Jacques, Guillaume. 2013c. Ideophones in Japhug Rgyalrong. *Anthropological Linguistics* 55.3. 256–287.
- Jacques, Guillaume. 2014a. Clause linking in Japhug Rgyalrong. *Linguistics of the Tibeto-Burman Area* 37.2. 263–327.
- Jacques, Guillaume. 2014b. Denominal affixes as sources of antipassive markers in Japhug Rgyalrong. *Lingua* 138. 1–22.

- Jacques, Guillaume. 2015a. *Dictionnaire Japhug–Chinois–Français, version 1.0*. Paris: Projet HimalCo. Available at: <http://himalco.huma-num.fr/> (accessed 24 March 2016).
- Jacques, Guillaume. 2015b. The origin of the causative prefix in Rgyalrong languages and its implication for Proto-Sino-Tibetan reconstruction. *Folia Linguistica Historica* 36.1. 165–198.
- Jacques, Guillaume. 2016a. From ergative to comparee marker: multiple reanalyses and polyfunctionality. *Diachronica* 33.1. 1–30.
- Jacques, Guillaume. 2016b. Subjects, objects and relativization in Japhug. *Journal of Chinese Linguistics* 44.1. 1–28.
- Jacques, Guillaume and Zhen Chen. 2010. *Une version rgyalrong de l'épopée de Gesar*. Osaka: National Museum of Ethnology.
- Lin, Youjing. 2003. Tense and aspect morphology in the Zhuokeji rGyalrong verb. *Cahiers de Linguistique—Asie Orientale* 32.2. 245–286.
- Lin, Youjing. 2011. Perfective and imperfective from the same source: directional “down” in rGyalrong. *Diachronica* 28.1. 54–81.
- Lín, Yòujīng and Luóérwū. 2003. 茶堡嘉戎語大藏話的趨向前綴及動詞 詞幹變化 Chábǎo jiāróngyǔ Dàzhànghuà de qūxiàng qiánzhù jí dòngcí cígàn biànhuà (The directional prefixes and verb stem alternations in the Datshang dialect of Japhug). 民族語文 *Minzú Yǔwén* 4. 19–29.
- Michailovsky, Boyd, Martine Mazaudon, Alexis Michaud, Séverine Guillaume, Alexandre François and Evangelia Adamou. 2014. Documenting and researching endangered languages: the Pangloss Collection. *Language Documentation and Conservation* 8. 119–135.
- Sun, Jackson T.-S. 2000. Parallelisms in the verb morphology of Sidaba rGyalrong and Lavrung in rGyalrongic. *Language and Linguistics* 1.1. 161–190.
- Sun, Jackson T.-S. 2006. 嘉戎語動詞的派生形態 Jiāróngyǔ dòngcí de pàishēng xíngtài (Derivational morphology in the Rgyalrong verb). *Minzu Yuwen* 民族語文 4.3. 3–14.
- Sun, Jackson T.-S. 2012. Complementation in Caodeng rGyalrong. *Language and Linguistics* 13.3. 471–498.
- Sun, Jackson T.-S. 2014. Sino-Tibetan: Rgyalrong. In Rochelle Lieber and Pavol Štekauer (eds), *The Oxford Handbook of Derivational Morphology*, 630–650. Oxford: Oxford University Press.
- Tournadre, Nicolas. 2008. Arguments against the concept of “conjunction” / “disjunction” in Tibetan. In Brigitte Huber, Marianne Volkart and Paul Widmer (eds). In *Festschrift für Roland Bielmeier zu seinem 65. Geburtstag*, 281–308. Halle: International Institute for Tibetan and Buddhist Studies.
- Tournadre, Nicolas and Randy LaPolla. 2014. Towards a new approach to evidentiality. *Linguistics of the Tibeto-Burman Area* 37.2. 240–262.

§3.4.2 *Rawang-Dulong-Anong*

CHAPTER THIRTY-TWO

DULONG

Randy J. LaPolla

1 INTRODUCTION

Dulong [tərūŋ] is a Tibeto-Burman variety considered a language in China, but it is actually part of a larger language that includes what is called the Rawang language of Myanmar (Burma) and the Anong variety spoken in Fugong county on the Nu river in China.¹ The Dulong speakers mainly live in Gongshan Dulong and Nu Autonomous County in Yunnan, China, and belong to either what is known as the Dulong nationality (population 5,816 according to the 1990 census, the newest figures available in 2014), or to one part (roughly 6,000 people) of the Nu nationality (those who live along the upper reaches of the Nu River). The exonym “Dulong” (or “Taron,” or “Trung”) was given to this nationality because they mostly live in the valley of the Dulong (Taron/Trung) River (the easternmost source of the Irrawaddy). In the past, the Dulong River was known as the Kiu (Qiu) river, and the Dulong people were known as the Kiu (Qiu), Kiutze (Qiuzi), Kiupa, or Kiao. Dulong is usually talked about as having four dialects, based on areas where it is spoken: First Township (northern Dulongjiang), Third Township (middle Dulongjiang), Fourth Township (southern Dulongjiang), and (northern) Nujiang. In this chapter, we will be using data of the First Township dialect spoken in Dizhengdang, Gongshan county. Other sources on Dulong, Rawang, and Anong are listed at the end of this chapter. See especially Yang Jiangling’s (a native speaker linguist) work on the Third Township Dulong variety. See also the Rawang-Dulong-Anong Language and Culture Website.² The Dulong-Rawang-Anong language forms part of the Rung branch of Tibeto-Burman, established on the basis of morphological innovations in LaPolla (2013; see also LaPolla, Chapter 2, this volume—“Overview of Sino-Tibetan Morphosyntax”).

2 PHONOLOGY

Dulong has 24 initial consonants at six points of articulation (Table 32.1), plus the consonant clusters /pr, br, mr, kr, xr, gr, pl, bl, ml, kl, gl/ in initial position; only the consonants /p, t, ?, k, n, m, ɳ, r, l/ occur in final position. As /-w-/ only appears after velar initials, I am treating these combinations as a labio-velar series. When followed in close juncture by a voiceless segment, the final nasals are often pronounced as voiceless stops, e.g. wən ‘buy’ > wət-çüü ‘to buy for oneself.’ The initial /k-/ is pronounced rather back; the palatal stops are often pronounced with slight affrication, and the voiceless affricates are often aspirated.

The Dulong vowels are /i, ε, ə, a, ɔ, u, u/, and there are three diphthongs, /ɔi, ai, ui/, which only appear in open syllables. The syllable can be cv (*bà* ‘thin’), cvc (*cəm* ‘iron’),

TABLE 32.1 THE DULONG CONSONANTS

	Labial	Dental	Retroflex	Alveopalatal	Palatal	Velar	Labio-velar	Glottal
voiceless stop	p	t			c-	k	kw-	?
voiced stop	b-	d-			ɟ-	g-	gw-	
voiceless affricate		ts-		tɕ-				
voiceless fricative		s-		ɕ-		x-	xw-	
voiced fricative		z-		ʑ-				
nasal	m	n			jn-	ŋ	ŋw-	
liquids		l	r					
approximants	w-				j-			

cvv (where ‘vv’ represents a diphthong; *pài* ‘large bamboo basket’), ccv (where cc represents one of the consonant clusters listed above; *blā* ‘picture,’ ‘drawing’), or CCVC (*mlāŋ* ‘dream’).

The structure of syllables such as these also includes one of two tones, level (usually pronounced as high level or mid level, the latter especially on grammatical particles and less-stressed syllables; e.g. *dā* ‘scarecrow’) or falling (usually pronounced as high falling, e.g. *dà* ‘gaze fixedly’), but the structure of many words is sesquisyllabic, where the first part of the word is an unstressed, toneless, cv syllable (‘half-syllable’), e.g. *džəzɪ* ‘a kind of pheasant.’ This in effect makes for something like a three-tone contrast (and it was discussed as such in Sun 1982). This reduced tone also appears on grammatical particles such as the postpositions and verbal suffixes. It is written here as a breve mark (᠁). Stopped syllables only appear in one tone, generally a high short tone, and so tone marks are not given on stopped syllables. There is no regular tone sandhi, but there is a change of falling tone to level tone on verbs as a marker of first person and also when preceded by certain prefixes or followed by certain suffixes.

3 MORPHOSYNTAX

Words can be formed by prefixation, suffixation, or compounding. Word classes include nouns, defined by the ability to appear with a numeral classifier; verbs, defined by the ability to appear with negation and the person and tense marking; postpositions, which are enclitic to NPs; numerals; and classifiers. Adjectives are a subset of stative verbs for which reduplication means intensification or adverbialization rather than perfective aspect (reduplication with nouns has a distributive meaning, ‘every’). Adjectives can be used as predicates or can appear nominalized in a copula clause, e.g. *dèm gē* [full NOM + cop] ‘(It) is full.’ Noun types include plain nouns, personal pronouns (Table 32.2), demonstrative pronouns (*já* ‘this,’ *᠁já* ‘that,’ *k᠁* ‘that (remote)’; used as free pronouns and demonstrative adjectives), and interrogative pronouns (*t᠁y* (*mā*) ‘what,’ *᠁nī* ‘who,’ *k᠁rwā* ‘how,’ *kā* (*d᠁*) ‘where,’ *t᠁kcā* ‘why,’ *kāpē* ‘which’). The latter can also be used as indefinite pronouns, as can *᠁ts᠁y* ‘person,’ as in *᠁ts᠁y m᠁-z᠁y-ci᠁* [person NEG-enter-R/M] ‘Nobody came in.’ There is an emphatic, but not reflexive, pronoun, *᠁diū*, that can be used to emphasize the agency of the actor (He did it HIMSELF). Many nouns are formed using the prefix *᠁y-* ~ *᠁k-*, e.g. *᠁k-sa?* ‘breath,’ ‘steam,’ *᠁y-j᠁u* ‘seed(s).’

Within the noun phrase, in addition to the head noun, there can be a demonstrative, genitive, or verbal/phrasal modifier, a classifier or a numeral plus classifier, and a plural

TABLE 32.2 THE DULONG PERSONAL PRONOUNS

	Singular	Dual	Plural
1	šgò	šjúŋŋpí	šjúŋŋ
2	nà	núŋŋpí/nžjúmpí/nžpní	nžúŋŋ
3	àŋ	žŋpí	žŋjúŋŋ

(*rì* inanimate) or group (*ma?* human; *rà* animate or inanimate) marker. A numeral plus classifier can occur either before or after the head, but a classifier used without a numeral must follow the head, and be used in conjunction with the demonstrative, which precedes the head. Unlike in Rawang, where a noun plus classifier construction without a demonstrative can be used when the referent is specific, in Dulong the demonstrative must be used. To use a classifier before the noun without a numeral, a special form I have called a ‘numeral substitute’ (*pē*) is used in place of the numeral. A genitive noun or pronoun simply precedes the head noun, and does not take any genitive marking, though there is a set of pronominal prefixes (š-/nž-/žŋ-) for kinship and relational terms (e.g. škày ‘my master’) derived from the free pronouns. A demonstrative can modify a noun alone, or a noun plus classifier combination. Adjectives can precede or follow the noun head, but when they precede, if there is no other modifier (e.g. a demonstrative), the adjective is generally nominalized by the prefix žŋ- ~ šk-. If the adjective follows the head, the prefix is optional. Compare zžjè (šk)-sár and šk-sár zžjè ‘new book.’ In many cases the adjective can be reduplicated as well, e.g. mržy mržy lüŋ ‘long stone.’ Adjectives also frequently appear as pre-modifiers in the form of the relative clause construction, e.g. tə̄t tčzžŋ gú kám [very hard NOM bamboo] ‘very hard bamboo.’ Following are some more noun phrase examples:

- | | |
|--|---|
| (1) a. zžjè ſŋpí pžŋ
book two CL
'two books' | b. šgò zžjè rì
1sg book pl
'my books' |
| c. šgò zžjè cč? pžŋ
1sg book one CL
'one of my books' | d. štsáŋ tži cč? jč?
person big one CL
'a big person' |
| e. kš tži zžjè šsúm pžŋ
that big book three CL
'those three big books' | f. jā pē pžŋ zžjè
that NS CL book
'that book' |

The noun phrase can be followed by a semantic-relation marking postposition to mark the referent as agentive, instrumental, or adverbial (*tč*); anti-agentive (animate patient, recipient, benefactive) or allative (*lč*); locative or temporal (*dč*); terminative (*xrc?*); or comitative (*mžnžŋ*). There is also a topic marker (*núñ*) and a noun conjunction particle (*pní*). The agentive marking is not obligatory on transitives or ditransitives, but is often used when there is a specific identifiable patient referent, particularly if the direction of action is inverse (e.g. 3 > 1) and/or the action is completed. It is never used on intransitives. There is no evidence for the grammaticalization of grammatical relations such as ‘subject’ and ‘direct object’ or grammatical-relation-changing constructions such as passive and antipassive. Aside from the topic marker, all relational marking is semantic in nature.

Morphological marking that appears within the verb complex includes direction marking, person marking, inverse-marking, reflexive/middle marking, tense/aspect marking, valency-changing affixes, and negation.

Transitive verbs can be intransitivized by use of the intransitivizing prefix *š-*, or by use of the reflexive/middle marking suffix *-cū*. The main function of the prefix *š-* is intransitivization (e.g. *təl* ‘roll,’ vt.; *štəl* ‘roll,’ vi.), but if the single direct argument of the derived intransitive is a plural animate argument, then the meaning is reciprocal, as in (2a). There is also an optional reciprocal particle (*ma?*; possibly the same morpheme as the human group-marking noun suffix *ma?*) that can be used after the verb in conjunction with the prefix. Reciprocals can be formed on causativized intransitives as well, and in this case will usually take the reciprocal particle and often an adverbial phrase meaning ‘to each other’ as well, as in (2b).

- (2) a. *nžnī ŋa-š-sət ma? cū*
 2dl N.1-RECIP-hit RECIP dl
 ‘You two are arguing/fighting (with each other).’
- b. *kɔlč jälč sž-zà ma?*
 that.way that.way CAUS-hurt RECIP
 ‘(They) are hurting each other.’

The reflexive/middle marker patterns something like the French reflexive pronouns, covering the semantics of both true reflexives (3a) and the middle voice (3b), but the marking is an invariant verbal suffix. The reflexive verb can take a patient noun if the noun is a body part, as in (3b), or something related to the actor, such as something the actor is buying for him/herself. There are a number of deponent forms, roots that take the reflexive/middle suffix for the ‘unmarked’ form, such as *jet-cū* ‘laugh.’ Removing the suffix would make the verb transitive ‘laugh at someone.’ There is a contrast in meaning between the two intransitivizers: the prefix gives a simple intransitive, such as *štəl* ‘roll,’ with no specification of agentivity, whereas the suffix implies agentivity, as in *təl-cū* ‘roll oneself’ (with change of tone). It is also possible to use both affixes together, giving a stative sense, as in *š-jžn-cū* [PREF-look-R/M] ‘be visible.’ (See LaPolla and Yang 1996, 2004 for more detailed discussion.)

- (3) a. *àŋ gžjūm-cū*
 3sg hit.with.fist-R/M
 ‘He is hitting himself with his fist.’
- b. *àŋ mɔr t̚eu?̚-cū*
 3sg face wash-R/M
 ‘He is washing his face.’

Intransitive verbs can be made transitive or causative by the addition of the causative prefix *sž-* ~ *tš-*, e.g. *dət* ‘broken (of string),’ *sždət* ‘to break (string);’ *št̚eu?* ‘pinched,’ ‘closed up,’ *t̚t̚eu?* ‘to pinch,’ ‘close up’ (see also (2b) and (4)). If the tone of the root is a falling tone, it becomes a level tone with the addition of the prefix, as in (4) (< *jì*).

- (4) *àŋ tč àŋ lč sž-jī*
 3sg AGT 3sg DAT CAUS-go
 ‘He made him go.’

Some verbs can also be made causative by simply lengthening the vowel (and changing to a level tone, if it was originally falling), e.g. *lūm* ‘warm’ (vi.), *lū:m* ‘warm’ (vt.) (= *să-lūm*).

An analytical causative/permissive construction is formed using the verb (*să*)-*zūr*, as in *jī săzūr* ‘let/make (him) go,’ and in (5). Again a falling tone changes to a level tone when followed by the causative verb. This form of causation involves less direct causation than the causative prefix. There is also at least one form that shows a remnant of the PTB *-t transitivity suffix: *yu* ‘cry’ > *yut* ‘mourn (cry for) a dead person.’

- (5) *àŋ tě pūŋ lě wà să-zūr*
 3sg AGT Pung DAT do CAUS-allow/make
 ‘He made/allowed Pung do (it).’

There is only one marker of negation, the verbal prefix *mă-*, pronounced *mā-* when the root already has a prefix, e.g. *mā-rănā* ‘not rest’ (with tone change).

A benefactive construction is formed using the verb *sănăŋ* ‘help’ or the benefactive auxiliary *ă* after the main verb (again with changed tone on main verb):

- (6) a. *ăŋ cūŋ săgò tē rī sănăŋ*
 3sg wood 1sg AGT carry help + 1sg
 ‘I carry wood for him.’ (lit. ‘I help carry his wood.’)
 b. *săgò tē ăŋ cūŋ rī ăŋ*
 1sg agt 3sg wood carry BEN + 1sg
 ‘I carry wood for him.’
 c. *ăŋ gwā sā să-tər ă*
 3sg wear NOM CAUS-big BEN
 ‘(S/he) makes it bigger for him/her to wear.’ (lit. ‘makes it big for his/her wearing’)

In terms of the marking of illocutionary force, the declarative is unmarked. The imperative is marked by the prefix *pă-* (pronounced *pā-* when the verb already has a prefix): *pă-kăi* ‘Eat!’ The prohibitive takes the normal negative, but the main verb is followed by *ăl* ‘have’: *me-kăi ăl* [NEG + N.1-eat have] ‘Don’t eat (it)!’. An indirect directive is marked by the prefix *la?*-, which is followed by the negative particle for indirect prohibitives: *la?-mă-wá* ‘Don’t let him do (it).’ For the hortative the verb *gōl* ‘want’ is used: *kər gōl* ‘Let’s eat!’ (with tone change).

Polarity questions are generally formed using the postverbal question particles (*pū*) à (7a). Wh-questions have the interrogative pronoun *in situ*, and do not require a final particle (7b). Another type of polarity question is formed by juxtaposing positive and negative choices (i.e. an A-not-A question), as in (7c).

- (7) a. *năjūŋ jni-kăi jūŋ (gu)* *guī ē ā*
 2pl N.1-eat plural NOM taro be Q
 ‘Is what you(pl) eat taros?’
 b. *jā sără ănī gu* *ē*
 this thing who NOM be
 ‘Whose is this thing?’
 c. *jni-kăi me-kăi*
 N.1-eat NEG + n.1-eat
 ‘Do you want to eat or not?’

The verb in Dulong inflects for person, but only speech-act participants are marked, with first person marked for person and number, while second person is marked only for number. The form of the first person singular marking depends on the final consonant of the root: if there is no final consonant, then *-ŋ* is added; if the final consonant is *-?* then *-k* is added; if there is some other final with falling tone, then there is a change to level tone. First person dual takes the dual marker *-cū*; first person plural involves a change from short vowel to long vowel.³ Second person dual also takes the dual marker *-cū*, while second person plural takes the plural marker *-jūŋ*. In either person, when the root takes a suffix or is changed to a long vowel, the root, if it has a falling tone, changes to level tone. Where the root vowel is *-ə-* and there is a *-p* or *-t* final, the vowel changes to *-a-* for all but first person singular. Aside from this, a prefix I call the N.1 (non-first person actor) prefix (*nə-* ~ *na-* ~ *ni-*) marks situations where a speech-act participant is mentioned, but the speaker is not the actor (contrast (8a–b)). When the root takes the intransitivizing prefix, the N.1 prefix is pronounced *na-* and either appears before the other prefix or incorporates the other prefix, e.g. *na-ə-sət* and *na-sət* ‘hit yourself’ are both possible. When it appears with consonant-initial prefixes, the N.1 prefix is simply marked by a change of the vowel of the other prefix to *-ɛ*, as in (7c), above. The variant *ni-* appears as a sandhi form when followed by a front vowel.

- (8) a. *əgə tə əŋ lɛ rūŋ (<ri)*
 1sg AGT 3sg DAT carry + 1sg
 ‘I carried him.’
- b. *əŋ tə əgə lɛ nə-ruŋ*
 1sg AGT 1sg DAT N.1-carry + 1sg
 ‘He carried me.’

In the case of two human interactants, particularly first and second person, person marking can be of either one, but usually when the agentive marking is used, the person marking is of the higher ranking participant (1 > 2 > 3).

The verbs for ‘come’ and ‘go’ have grammaticalized into direction markers, as in *lɔʔ-jì* [return-go] ‘go back’ and *lɔʔ-rà* [return-come] ‘come back.’ These and other direction-related verbs have grammaticalized into tense markers, *jǐ* (<*jì* ‘to go’) and *lǔŋ* (<*lùŋ* ‘to ascend’), both used for recent past actions. The difference between the two forms is an evidential distinction: the use of *lǔŋ* after the verb implies the speaker did not see the action (9a–a’), whereas the use of *jǐ* implies the speaker did see the action (9b–b’). A guess is marked by adding *mèl* after *lǔŋ*. For a strong assertion, *muū* is added after *jǐ*. For an action completed some time ago, possibly years ago, *buū* is used in place of *jǐ* or *lǔŋ*. Inchoatives take the particle *pəŋ* after the verb or the adverb *təsā* before the verb.

- | | |
|--|--|
| (9) a. <i>əŋ tə kà:i lǔŋ</i> | a' <i>əŋ jì lǔŋ</i> |
| 3sg AGT eat PAST/EVID | 3sg go PAST/EVID |
| ‘He has just eaten.’ (I didn’t see him eat.) | ‘He just left.’ (We didn’t see him leave.) |
| | |
| b. <i>əŋ tə kà:i jǐ</i> | b' <i>əŋ jì àm jǐ</i> |
| 3sg AGT eat PAST/EVID | 3sg go DIR PAST/EVID |
| ‘He ate.’ (I saw him eat.) | ‘He just left.’ (Possibly still can see him) |

The particles *jǐ* and *lǔŋ* are not used with first person actors; instead the adverb *zūr* ‘already’ can be used to mark a completed action, as in (10a). In asking someone about their past actions, *lǔŋ* is used (10b).

- (10) a. *šg̡ juŋ zuřr*
1sg go + 1sg already
'I went (and came back) already.'
- b. *nà jn̡-jì lúŋ ǎ*
2sg N.1-go ASP Q
'Did you go?'

Another evidential distinction is made with the hearsay particle *teiwař*. This appears most frequently in traditional narrative texts.

- (11) *teāpāř teāpāř nū džg̡ kā guř? sř teiwař*
long.ago long.ago TOP dog words say know.how HEARSAY
'(It is said) long ago dogs knew how to talk.'

The word order in the clause is verb final, while the NPs are ordered with the more topical elements being earlier in the clause; the immediate preverbal position is the unmarked focus position (the unmarked position for introducing 'new' referents/information). Adverbial elements usually appear in preverbal position (12a–b), but they can follow the verb (12c), with some difference in meaning. Some adverbs always follow the verb, such as *mătɔl* in (13), below, while others always precede the verb, e.g. *lēlā* in *lēlā mr̡øŋ* 'very long/tall.' Preverbal adverbial phrases that are not reduplicated adjectives or adverbs usually take the adverbial marker *wā* (12a) (derived from *wā* 'say') or sometimes *guū* (12e) (which means 'also' in Rawang). A resultative complement also comes before the main verb, marked by the terminative postposition *xre?* 'until,' as in (12e).

- (12) a. *àŋ blakp̡ì wā àŋzā k̡ì*
3sg monkey ADV food eat
'He eats like a monkey.'
- b. *măd̡ řbrà-brà ji*
car/truck fast-REDUP go
'The car is going quickly.'
- c. *măd̡ ji řbrà*
car/truck go fast
'The car goes (can go) quickly.'
- d. *nit-buřm mă-gɔl guū řmrā wà lě ji teiwař*
mind-many NEG-need ADV field work DAT go HEARSAY
'(The parents) could go to work the field without having to worry (about the child).'
- e. *’ɔp̡e əm̡e’ guř? sř xre? ð: ji teiwař*
daddy mummy say able.to ADV do go HEARSAY
'(It) did this until (the baby) could say "mummy and daddy".'

The comparative construction has the word order Topic–Standard–Marker–Verb, where the marker is the word for 'above' plus the locative marker. The verb (adjective) can be in the plain form or reduplicated and followed by an adverb used only in the comparative construction.

- (13) *šg̡ nà măd̡am d̡ (mr̡øŋ mătɔl)*
1sg 2sg above LOC long/tall REDUP very much
'I am (much) taller than you.'

There is only one verb of possession/existence, and the possessive construction patterns as a simple transitive clause; the possessor does not take any relational marking:

- (14) *šjuŋ cē tsūl jɔ? lāmbrɔ? àl*
1pl one ten CL friend have
'We have ten friends.'

A cleft construction can be used for narrow focus on a particular NP, either in questions (15a) or statements (15b):

- (15) a. *àŋ jì guū ē*
3sg go NOM be + Q
'Did he go?'
b. *pūŋ cīn tē sət gē* (*<guū-ē*)
Pung Cin AGT hit NOM + be
'It was Cin who hit Pung.'

Clefts are also used for achieving the sense of a root modal:

- (16) *šgɔ sələp-εiū sā gē*
1sg teach-R/M NOM NOM + be
'I ought to study.'

Subordinators include clause-final *bē* 'if,' 'when,' *mōn̥d̥y* 'because,' 'when' (<'follow'), *sānn̥d̥y* 'because of (in order to).' Some of the nominal postpositions are used as clausal subordinators as well, e.g. *lē* (allative/dative) is used for purpose clauses, e.g. *kòi lē jì* 'go to eat.' Non-quote complement clauses do not require a nominalizer or complementizer (17a–b), but quoted complements take the complementizer *wā* (< *wā* 'say') (17c).

- (17) a. *šgɔ àŋ lɔ? jì nit*
1sg 3sg return go remember
'I remember that he went back.'
b. *àŋ jì bē ŋgɔ kuu? juŋ du ē*
3sg go if 1sg also go + 1sg might COP
'If he comes, I might go.'
c. *àŋ šgɔ mōz̥i?māli? la-jì wā gu?*
3sg 1sg must have.to-go COMP say
'He says I must go.'

Predication of actions or attributes that occur at the same time is represented in a serial verb construction with *zīn* optionally appearing between the two verbs:

- (18) a. *mōnzū wā zīn zəŋ-εiū*
song do LNK enter-R/M
'He entered singing.'
b. *mr̥d̥y zīn p̥čeūŋ guū jɔ?*
long LNK blue/green NOM cloth
'long blue cloth'

Nominalization to mark a location where an action occurs or a thing that is involved in an action is by the particle *sā*, e.g. *jip sā* ‘sleeping place,’ *gwā sā* ‘clothing,’ *kài sā* ‘the thing which can be eaten/food.’ Relative clauses, which appear before the head noun, are generally nominalized by the particle *gū*. In some lexicalized expressions involving relative clauses, no nominalizer is used, e.g. *mǎnzū wà štsàj* [song do person] ‘professional singer.’ Relativization can be of the patient (19a), the agent (19b), the recipient (19c), or just about any role. In some cases, the noun head can be omitted (19d). There are no relative pronouns in Dulong.

- (19) a. àŋ tě làŋ gū cūŋwət
 3sg AGT bring NOM flower
 ‘the flower(s) he brought’
- b. cūŋwət làŋ gū štsàj
 flower bring NOM person
 ‘the person who brought the flower(s)’
- c. àŋ lě cūŋwət làŋ gū štsàj
 3sg DAT flower bring NOM person
 ‘the person I brought flower(s) for’
- d. zějè brī sā ð gū (sərā)
 book write NOM do NOM thing
 ‘the thing I use to write books’
- e. i tě ū tsərī
 hemp INST make string
 ‘a string made of hemp’

Reduplication of the verb has something of a perfective sense, and is used to mark the doing (completion) of an action in preparation for another in non-final clauses of a clause chain. In procedural texts, this form is used in a pattern with *v-t̥n mǎnəŋ* ‘[v finish follow] ‘having finished ving’ or ‘after ving’ where one or more clauses with reduplicated verbs will lead up to the completion of some act or stage of a process, and then *v-t̥n mǎnəŋ* will lead into the next series of actions, as in the following segment of an explanation of how to make a crossbow (see LaPolla 2001 for more examples):

- (20) tānā ākpləŋ zū t̥n mǎnəŋ,
 crossbow body make finish afterwards
 ‘After finishing the body part of the crossbow,’
- cēlaī ū t̥lī wā r̥mū-mū
 one side LOC bow ADV draw-REDUP
 ‘on one side, (I) draw the shape of a bow,’
- dəkeī tě puk.
 small.knife INST bore.small.hole + 1sg
 ‘(and then) use a small knife to bore a small hole.’
- pu? t̥n mǎnəŋ, nātsēnātsē t̥lī zāŋ.
 bore.small.hole finish after slowly bow fit.into + 1sg
 ‘After having bored a small hole, (I) slowly fit the bow into the hole.’

NOTES

- 1 The dialect split is not clearly Dulong vs Rawang, but north–south, ignoring the national border. That is, the northern Dulong varieties are closer to the northern Rawang varieties and less close to the southern Dulong varieties and Anong, which are closer to the southern Rawang varieties.
- 2 See tibeto-burman.net/rda/.
- 3 See LaPolla and Yang (2007) for a discussion of the origin of the vowel length distinctions and also of the benefactive marking. See LaPolla 2010b on the hierarchical person marking system.

BIBLIOGRAPHY

- Barnard, J.T.O. (1934) *A Handbook of the Rawang Dialect of the Nung Language*, Rangoon: Superintendent of Government Printing and Stationery.
- Bradley, David (2010) ‘Review of *A Grammar of Anong: Language Death under Intense Contact (Languages of the Greater Himalayan Region 9)* by Hongkai Sun and Guangkun Liu, translated, annotated and supplemented by Fengxiang Li, Ela Thurgood and Graham Thurgood, Leiden and Boston: Brill 2009’, *Himalayan Linguistics* 9.2: 126–131.
- Dai, Qingxia, Huang, Bufan, Fu, Ailan, Renzengwangmu and Liu, Juhuang (1991) *Zàng-Miānyǔ shíwúzhōng (Fifteen Tibeto-Burman Languages)*, Beijing: Yanshan Chubanshe (includes chapter on 4th Township Dulong).
- LaPolla, Randy J. (1987) ‘Dulong and Proto-Tibeto-Burman’, *Linguistics of the Tibeto-Burman Area* 10.1: 1–42.
- LaPolla, Randy J. (1995) ‘On the utility of the concepts of markedness and prototypes in understanding the development of morphological systems’, *Bulletin of the Institute of History and Philology* 66.4: 1149–1185.
- LaPolla, Randy J. (1996) ‘Middle voice marking in Tibeto-Burman languages’. In *Pan-Asian Linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics*, Vol. 5, Thailand: Mahidol University, 1940–1954.
- LaPolla, Randy J. (2000) ‘Valency-changing derivations in Dulong/Rawang’. In R.M.W. Dixon and Alexandra Y. Aikhenvald (eds), *Changing Valency: Case Studies in Transitivity*, Cambridge: Cambridge University Press, 282–311.
- LaPolla, Randy J. (2001) ‘Dulong texts: seven narrative and procedural texts’, *Linguistics of the Tibeto-Burman Area* 24.2: 1–39.
- LaPolla, Randy J. (2006) ‘Copula constructions in Rawang’. In Dah-an Ho, H. Samuel Cheung, Wuyun Pan and Fuxiang Wu (eds), *Linguistic Studies in Chinese and Neighboring Languages: Festschrift in Honor of Professor Pang-hsin Ting on His 70th Birthday*, Taipei: Institute of Linguistics, Academia Sinica, 1059–1076.
- LaPolla, Randy J. (2008a) ‘Nominalization in Rawang’, *Linguistics of the Tibeto-Burman Area* 31.2: 45–66.
- LaPolla, Randy J. (2008b) ‘Relative clause structures in the Rawang language’, *Language and Linguistics* 9.4: 797–812.
- LaPolla, Randy J. (2010a) “Transitivity harmony” in the Rawang language of northern Myanmar’. In L. de Beuzeville and P. Peters (eds), *From the Southern Hemisphere: Parameters of Language Variation—E-Proceedings of the 2008 Conference of the Australian Linguistics Society* (University of Sydney, 2–4 July 2008), 1–9. Available at: hdl.handle.net/2123/7107.

- LaPolla, Randy J. (2010b) ‘Hierarchical person marking in Rawang’. In Dai Zhaoming (ed.), *Forty Years of Sino-Tibetan Language Studies: Proceedings of ICSTLL-40*, Heilongjiang: Heilongjiang University Press, 107–113.
- LaPolla, Randy J. (2011) ‘On transitivity in two Tibeto-Burman languages’, *Studies in Language* 35.3: 639–649.
- LaPolla, Randy J. (2013) ‘Subgrouping in Tibeto-Burman: can an individual-identifying standard be developed? How do we factor in the history of migrations and language contact?’. In Balthasar Bickel, Lenore A. Grenoble, David A. Peterson and Alan Timberlake (eds), *Language Typology and Historical Contingency*, Amsterdam and Philadelphia: John Benjamins, 463–474.
- LaPolla, Randy J. and Poa, Dory (2001) *Rawang Texts*, Berlin: Lincom Europa.
- LaPolla, Randy J. and Yang, Jiangling (1996) Dúlóng/Riwàngyǔ dòngcí de fānshēn hé zhōngjiāntài biāozhì (Reflexive and middle marking in Dulong/Rawang). In Dai Qingxia, Zeng Siqi, Cheng Yanyan, Ding Shiqing and Fu Ailan (eds), *Zhōngguó Mínzú Yǔyán Lùncōng (I) (Collected Essays on Chinese Minority Languages, I)*, Beijing: Central University of Nationalities Press, 13–34.
- LaPolla, Randy J. and Yang, Jiangling (2004) ‘Reflexive and middle marking in Dulong-Rawang’, *Himalayan Linguistics* 2 Available at: randylapolla.net/Papers/LaPolla_and_Yang_2005_Reflexive_and_Middle_Marking_in_Dulong-Rawang.pdf.
- LaPolla, Randy J. and Yang, Jiangling (2007) ‘Cóng yǔyán, fāngyán de bìjiào liǎo jí yǔyán de lǐshí fāzhǎn’ (Understanding the historical development of a language by comparing it with related languages and dialects), *Journal of Sino-Tibetan Linguistics* 1.1: 120–126. (Also published in *Dōngfāng Yǔyánxué* 2 (2007): 11–18.)
- Liu, Juhuang (1988) ‘Dúlóngyǔ dòngcí yánjiū’ (Studies on the Dulong verb), *Yǔyán yánjiū* 1988.1: 176–191.
- Lo, Ch’ang-p’ei (1945) ‘A preliminary study of the Trung language of Kung Shan’, *Harvard Journal of Asiatic Studies* 8: 343–348.
- Morse, Robert H. (1963) ‘Phonology of Rawang’, *Anthropological Linguistics* 5.5: 17–41.
- Morse, Robert H. (1965) ‘Syntactic frames for the Rvwang (Rawang) verb’, *Lingua* 15: 338–369.
- Sun, Hongkai (1982) *Dúlóngyǔ jiǎnzhì (A Sketch of the Dulong Language)*, Beijing: Minzu Chubanshe.
- Sun, Hongkai (1988) ‘Notes on a new language: Anong’, *Linguistics of the Tibeto-Burman Area* 11.1: 27–63.
- Sun, Hongkai and Liu, Guangkun (2009) *A Grammar of Anong: Language Death under Intense Contact (Languages of the Greater Himalayan Region 9)* [edited translation and annotation, with the translation, editing, annotation and expansion by Graham Thurgood, Fengxiang Li and Ela Thurgood], Leiden: Brill.
- Yang, Jiangling (1999) ‘Dúlóngyǔ dòngcí qūxiàng fānchóu yánjiū’ (The directional category of Dulong verbs), *Mínzú Yǔwén* 1999.1: 37–42.
- Yang, Jiangling (2000) ‘Dúlóngyǔ de chángyuányīn’ (The long vowels of Dulong), *Mínzú Yǔwén* 2000.2: 38–44.
- Yang, Jiangling (2004) ‘Dúlóngyǔ de qíngtài fānchóu’ (The category of modality in Dulong), *Mínzú Yǔwén* 2004.4: 2–6.
- Yang, Jiangling (2009) ‘Cóng qīnshǔyǔ kàn Dúlóngyǔ de yiyuáncí’ (Examination of lexical differences between Dulong and related languages), *Mínzú Yǔwén* 2009.4: 38–40.
- Yang, Jiangling (2011) ‘Dúlóngyǔ gètǐ liàngcí de shēngchǎn hé fāzhǎn’ (The production and development of individual classifiers in Dulong), *Mínzú Yǔwén* 2011.6: 55–62.

§3.4.3 *Kiranti*

CHAPTER THIRTY-THREE

KIRANTI LANGUAGES

Boyd Michailovsky

1 INTRODUCTION

1.1 ‘Kiranti’ languages and speakers

The Kiranti group consists of some 30 or more languages spoken in the Eastern Nepal hills, from the districts of Ramechhap and Sindhuli to Nepal’s eastern border, and beyond it into the Indian States of West Bengal (Darjeeling district) and Sikkim (West Sikkim). The name ‘Kiranti’, used here as a linguistic classification, is in origin a Sanskrit term (*kirāta*) apparently referring to non-Hindu Himalayan hill tribes in general; in Nepali, *kirāti* (or *kirātī*) serves as an exonym equivalent to ‘Rai-Limbu’.^{1,2} An indigenous autonym **rakdunij* is widely reported between Bahing *rɔdi* in the west and Limbu *jakthunj*³ in the east.

In the Nepal National Census of 2011, roughly 800,000 respondents declared a Kiranti language as their ‘mother tongue’. The number of speakers is probably less than this. The 2001 Census of India reported 50,000 speakers of Limbu and ‘Rai’ in India (most in Sikkim) (*Linguistic Survey of India* n.d.).

1.2 Basic typology

Kiranti languages share in the general word-order typology of the Indian linguistic area: SOV and modifier-head order, zero-pronominalization (no obligatory pronouns). Functional morphemes (postpositions, subordinating conjunctions, complementizers) follow the elements in their scope.

Case of nominal core arguments (defined by verb agreement) or other participants is marked by postpositions. Case marking of core arguments is ergative: intransitive S[ubject] and transitive O[bject] are generally unmarked (absolutive case); transitive A[gent] is marked by an ergative/instrumental case-marker.

The main originality of the group is the agreement system, which indexes person and number (singular, dual, plural) and clusivity (1st person inclusive vs exclusive) of core arguments (with variable precision). Grammatical attention to location or direction on the vertical axis is another particularity.

1.3 Subgroups

A subgrouping of languages within the group was proposed by Michailovsky (1994) in a study of shifts in initial consonants. Opgenort (2005, 2011) has pursued this research, taking into consideration unpublished work by Sergei Starostin and adding lexical isoglosses (2011; see also Winter 1991). Subgrouping is not discussed in the present

paper. I use provisional cover-terms (1) for subsets of the Kiranti group insofar as they show typological similarities.

- (1) Kiranti: Rough typological grouping of languages cited
Western

Hayu

Core Western: Bahing-Sunwar, Thulung, Koyi, Wambule-Jero

West-Central: Dumi-Khaling

Eastern

East-Central: Kulung, Sangpang

South-Eastern: Chamling, Bantawa, Puma (Rai *et al.* 2009), Chhintang
(Rai *et al.* 2011)

Upper Arun: Mewahang, Lohorong

Far Eastern:

Yakkha group: Athpare, Belhare, Yakkha, Yamphu

Limbu

The number of common innovations that I can cite to justify the status of Kiranti as a clade within Sino-Tibetan remains disappointingly small. I can suggest the following (none of them shared by Hayu): (1) a possible lexical innovation, PK **del* ‘village’,⁴ shared by all members of the Kiranti group and to my knowledge unknown elsewhere except in Lepcha; (2) the transformation of PST **mij* ‘name’ into PK **nij* (except in Limbu); (3) the etymon PK **rakduy* ‘Kiranti’.

1.4 Resources

The earliest work on Kiranti is that of Brian Hodgson (1857a, 1857b, 1857–1858, 1858). The *Linguistic Survey of India* (Grierson 1908: 198–205, 283–398) has very uneven material on some 20 languages of the group. Modern typological surveys of Kiranti languages include Ebert (1994), a detailed study, particularly of morphology and syntax, in six Kiranti languages and Ebert (2003b), a short summary based on a dozen or so Kiranti languages, which appeared in the first edition of the present volume. Articles on specific issues and languages are mentioned below.

On Kiranti or Rai oral tradition, see Allen (2012); Gaenszle (1991, 2002); Ebert (2000); and Ebert and Gaenszle (2008). The mythological ‘Orphan’ cycle is shared by Western Kiranti (except Sunwar and Hayu), South-Eastern and Upper Arun languages, but not by the Far Eastern languages.

Recorded texts with annotation are available in the archive of the Volkswagen Foundation’s Documentation of Endangered Languages (DOBES) programme (large collections in Puma and Chhintang); in the Pangloss collection of the Lacito research group, French National Research Centre (CNRS) (Hayu, Limbu, Bahing, Khaling, Koyi and Thulung); and in the Endangered Language Archive (ELAR), at the School of Oriental and African Studies (SOAS) of the University of London (Koyi, Puma, and Yakkha).

2 PHONOLOGY

2.1 Syllable and word structure

Kiranti phonologies are best understood in the framework of a typical Kiranti inventory and syllable and word structure constraints. Syllables are maximally (C_i)V(C_f),⁵ where C_i represents C_i(C_m), i.e. an initial C_i or group C_i + medial.

2.2 Initials

The core inventories of C_i include between two and four manner series of stops and affricates, one of nasals, and j, w, r, l, s, h. Khaling has a relatively complete consonant inventory:

- (2) Khaling consonant inventories (Jacques *et al.* 2012)

Syllable canon (segmental): (C_i(C_m))V(C_f)

C_i:

p	p ^h	b	b ^h	m
t	t ^h	d	d ^h	n
k	k ^h	g	g ^h	ŋ
ts	ts ^h	dz	dz ^h	

j, w, r, l, s, h.

C_f: p, t, k, m, n, ŋ, r, l, s, h, ç (before C_i t in the following syllable of the word)

C_m: r, l (only after velar or bilabial C_i stop)

The opposition of voice is absent in Limbu and nearly so in the other Far Eastern languages. In these languages, syllable-initial stops are generally voiced in certain phonological contexts, including post-vocalic and post-nasal. In Limbu, syllable-initial occlusives are voiced after a preceding vowel or nasal, even potentially across a word-boundary, unvoiced otherwise. (Some older speakers still apply this rule in speaking Nepali, however fluently.) Yamphu (Rutgers 1998) has the same distribution, except that word-initial b is reported in a few words in opposition to p. Yakkha (Schackow 2015) is also reported to lack phonological voicing. Belhare has an opposition between unaspirated voiceless and voiced initials (Bickel, this volume), but the latter are seen to be infrequent in his dictionary (Bickel 1997a).

South-Eastern languages have lacunary (or nearly lacunary) voiced series (b, d, Ø; bh, dh, Ø). Doornenbal reports that dz, dz^h; g, g^h are present, but rare, in native Bantawa vocabulary. Ebert (1994: 13) reports their rarity generally in Athpare, Bantawa and Chamling; this appears to apply to Puma and Chhintang as well.

Aspiration is phonological in all languages. Among languages that have plain voiced series, voiced aspirates are absent (or absent in native vocabulary) in Hayu, Sunwar and Kulung, and very rare except in loans in Wambule and Jero. Their opposition with plain voiced is described as unstable in Thulung (Allen 1975: 12).

All languages have initial m, n, ŋ, j, w, r, l, s, h. Word-initial ŋ is replaced by n in Chhatare (Tumbahang 2011: 24) and (apparently) Phedappe (van Driem 1987) Limbu. Voiced breathy m^h, n^h, l^h, r^h occur in a few Chamling words; y^h, mh are reported as rare in Sangpang. Sources differ on the value of Puma mh ([m] or [m^h]), nh. Hayu has voiceless l in a very few words.

Sunwar C+j clusters occur only with velar (including nasal) and sibilant initials, suggesting a complete palatal series kj, khj, gj, ŋj, sj, ſ (relatively frequent) and ſj are also reported. Hayu has palatal initial stops (c, j) in opposition with alveolar affricates ts, dz,⁶ and a fricative initial x, realized [ç], [x], [x^w], [χ], depending on the following vowel and the dialect. The aspirated affricate ts^h is lacking in Sunwar, Bahing (almost), Wambule and Jero; in Limbu, ts^h occurs only as an allophone of s after dental finals. Hayu appears to have no opposition between aspirated members of the affricated (ts, etc.) and palatal series.

Preglottalized, implosive initial b was reported in Bahing (ba ‘chicken’, bar ‘wound’) and in Sunwar in the 1970s (Michailovsky 1975a, 1988b); in Sunwar, it has been replaced by b or w (Borchers 2008). Both b and d are reported in Wambule (bari ‘wound’; bisi

‘eye’; *dwabu* ‘ear’) (Opogenort 2004). Preglottalized *ɻl* is reported from a Wambule dialect (Wamdayal *ɻluy* ‘stone’). These initials are in opposition with plain voiced initials.

Retroflex initials are absent in most languages and marginal or limited to Nepali loans in the rest, except in Thulung, where *ɖ* appears regularly in native core vocabulary (*ɖuy* ‘drink’, *ɖi* ‘egg’) in opposition to *d* (*duj* ‘liver’).

Initial C_iC_m groups of velar or bilabial stop + resonant (*l* and/or *r*) are found in the western languages (Hayu, Sunwar, Bahing, Wambule, Jero, Thulung and Khaling), and (with unvoiced C_i only) in western dialects of Chamling. They are absent in Dumi and in the eastern languages generally.

Initial clusters with medial approximants (C_m *j* (or ‘y’), *w*) may be interpreted as indicating consonantal orders – such as palatals (as in Sunwar, above) or possibly labiovelars – or the glides may belong to the nucleus in some western languages (see §2.4). In Limbu and Bantawa they are marginal. Doornenbal (2009: 36–9) shows that their occurrence in Bantawa is strictly limited phonologically, and that they only occur in ‘paralexemic’ or non-core vocabulary; in Limbu, they result from syncope (Limbu *c'wa:t* ‘water’, *t'je* ‘he arrived’ (root |ta|, past stem *t'j-*)).

2.3 Finals

Finals are typically limited in number (six to ten), with a single series of C_f stops. The latter are pronounced unreleased, with simultaneous glottal closure reported in many languages (Hayu, Bahing, Dumi, Wambule, Bantawa, Yamphu, Limbu). Final stops have voiced allophones before voiced stop C_i in Hayu, Thulung, Jero and Wambule (not in Dumi). Phonological final glottal stop is reported in several languages; it may have developed from glottal-stop allophones or variants of oral final stops. Such allophones are reported in Hayu, Wambule, Dumi and Limbu.

Sunwar is described by Bieri and Schulze as having no word-final consonants, except for syllable initials stranded in word-final position by apocope of the syllable vowel in fast speech. Final *n* in their transcription represents nasalization. Bahing and Chamling have no word-final stops.

A single C_i (or a complex C_iC_m) can occur between two vowels (-VC_i(C_m)V-). When a C_f is followed by a vowel-initial syllable there are various outcomes. In Limbu, C_f are geminated in this context, which arises very frequently before vowel-initial suffixes, postpositions or clitics: *kume:t* ‘his wife’; *kume:tti?* ‘his wife?’; *ammu:ttay* (|am-u:t-an|) ‘they call us and ...’. In Yamphu, the C_f may be geminated or not, but it remains unvoiced.

Word internal consonant groups occur at syllable boundaries, and are thus in principle limited to C_iC_i(C_m) combinations, but other combinations may occur as the result of syncope (loss of an internal vowel) leading to anomalous consonant groups, e.g. Limbu *car'ppʰemba* ‘butterfly’. Anomalous word-finals result from apocope (loss of a final vowel) leaving a word-final cluster (C_f)C_i, e.g. Kulung *baks* ‘lightning’, Sunwar *lob* ‘younger brother’. Syllabic nasals occur infrequently in a few languages, sometimes as prefixes, e.g. Kulung *m̩* ‘3OBL’ and Chamling (Chapter 36) and Belhare (Chapter 35) *N̩-*.

2.4 Vowels

Between five and ten vowel qualities are reported. Sunwar and the eastern languages Kulung, Athpare, Belhare and Yakkha have *i*, *e*, *a*, *o*, *u*. Other eastern languages add a sixth vowel (not counting *ɔ* or *ʌ* if found only in Nepali loans, as is often reported): Puma has *ʌ*; Bantawa adds *i* (mainly an allophone of *i* before velars), as does Chhintang; Yamphu

adds *œ*. Limbu, and Hayu in the west, have seven-vowel systems, with three degrees of aperture front and back.

Some western and central languages have richer inventories, although their distribution and alternations point to underlying five-vowel systems enriched by both vowel harmony and the effect of tautosyllabic consonants, especially velar finals (Allen 1975; Michailovsky 1975a, 2012; Jacques 2013). Bahing and Thulung have ten vowels, Khaling nine, Dumi, Wambule and Jero eight: *i, e, ε (or ja), a, o, ɔ (or wa), u (ʌ almost always in loans)*. Central or front rounded vowels and vowel harmony are reported in Bahing, Thulung and Khaling.

Up-gliding diphthongs induced by coronal finals (Michailovsky 1975b) are reported in Sunwar, Bantawa, Chamling (Chapter 36) and Belhare (Chapter 35).

2.4.1 Vowel quantity

Phonological length is reported in the western languages, and in Kulung, Yamphu and Limbu. It is not reported in the South-Eastern languages. In Wambule and Jero long *a, i, u* are found in open syllables, and very marginally in closed ones; the vowels *e, o* are the long partners of *ja/e* and *wo/o*. In Limbu, the mid vowels *e, o* are inherently long. In Hayu, Thulung and Bahing, length is pertinent only on open syllables.

In the western languages length appears to be secondary: this is clear in verb paradigms, where open-syllable length compensates for the loss of consonant finals in some stem alternants (e.g. Bahing *glu:ta* ‘he came/went out’, root |glunj|). See Hayu (Chapter 34, this volume), for a variation on this scenario that could be called ‘compensatory shortening’). In Limbu and Yamphu, however, length is clearly primitive in closed syllables with stop finals. Thus Limbu *ha:p-ma* ‘to weep’ vs *thap-ma* ‘to give birth’. Length in open stem-syllables is pertinent only in stems derived from stop-final roots: *ha:b-e* ‘he wept’ vs *thab-e* ‘it gave birth’, *pa:n* ‘speech’ (cf. |pa:tt| ‘speak’) (Michailovsky 1986, 2002).

2.5 Tone, stress

In the 1970s and 1980s, tone was reported in Sunwar (Bieri and Schulze 1971a, 1971b; Genetti 1988), Thulung (Allen 1975) and Khaling (Toba and Toba 1975), with indications of tonal morphophonology. More recent work on Sunwar (Borchers 2008) and Thulung (Lahaussois 2004) has failed to find tone, probably due to dialect or generational differences. But tone and tonal morphophonology in present-day Khaling have now been thoroughly described by Jacques *et al.* (2012). There is an opposition of level vs falling tone on heavy syllables (i.e. syllables with resonant finals or open syllables with long vowels). In verb stem morphophonology falling tone reflects certain stop finals, and can also reflect a lost syllable (Jacques, forthcoming).

(3) Tonal oppositions in Khaling infinitives (Jacques *et al.* 2015)

tsene (toneless short vowel) (root |tsel|) ‘to go bad (of flour)’

tsē:ne (level tone) (root |tseŋ|) ‘to sift’

tsē:ne (falling tone) (root |tsektl|) ‘to close’; (roots |tsek|) ‘to pluck’, ‘to be hard’

These three tonal languages occupy a contiguous area in Ramechhap, Okhaldhunga and Solu Districts in the north-west of the Kiranti range.

A tonal opposition reflecting an underlying stem postfinal (§4.2) has been reported in verb forms of a Bantawa dialect, for some speakers only (Doornenbal 2009; see also

Winter 1997). A length and/or tonal opposition, attributed to an elided vowel, has also been reported in Puma, in the slow speech of older speakers (Bickel *et al.* 2007a: 5).

Predictable, fixed stress has been described in several languages (e.g. Belhare, Dumi (van Driem 1993), Wambule); a stress opposition is reported on polysyllabic words in Sangpang (Huysmans 2007).

3 NON-VERBAL MORPHOLOGY

3.1 Noun formation

Nouns can be of any length, with disyllables probably the most common; many have taxonomic derivational suffixes with meanings like ‘person’, ‘game (animal)’, ‘bird’, ‘male’, ‘female’, ‘child’, ‘fruit’ (4). (See also Opgeort 2004: 127–38 for Wambule examples).

- (4) Bantawa taxonomic suffixes (Doornenbal 2009: 65–9)

chokkwasi ‘orange’, *naysi* ‘hailstone’ (cf. *si* ‘fruit’)
mukwa ‘partridge’, *chojwa* ‘bird’, *berawa* ‘parrot’ (cf. *wa* ‘chicken’)
mikwa ‘(eye-)tear’, *cakwa* ‘water’ (cf. *wa* ‘rain’)
khissa ‘deer’, *bwasa* ‘pangolin’, *maksa* ‘bear’ (cf. *sa* ‘flesh’)

3.2 Nominal morphology

Number marking of nominals is not obligatory. Plural and (often) dual number of nouns, especially those with human referents, is marked by suffixes, e.g. Sunwar *nisi* ‘DU’, *paki* ‘PL’; Bahing *dɔsi* ‘DU’, *dɔ* ‘PL’; Limbu *ha?* ‘PL (NSG)’, *si* ‘COLLECTIVE’: *a-mba-si* ‘my father’s generation’.

Articles are found in Far Eastern languages. Limbu has a suffixed singular definite article (*e*)*n* (see examples (20), (31), (34), (51), (52)). Yakkha also has suffixed markers implying countable, specific reference *na* ‘SG’ and *ci* ‘NSG’. Unlike Limbu (*e*)*n*, they can appear in adnominal position, e.g. *to-na cuwa* (up-DF beer) ‘the (bowl of) beer up there’ (Schackow 2015: 120–1; see also Bickel in this volume).

Limbu also has a full set (3SG is zero) of personal agreement markers suffixed to nominal predicates in identificational sentences (§4.7).

3.3 Personal pronouns and possessive prefixes

All languages have both independent personal pronouns and prefixed possessive pronouns. 3rd person pronominal forms are generally reserved for human referents. Bahing has complete paradigms, for all persons and numbers (5). The dual and plural forms of the 3rd person independent pronouns are formed with the same number suffixes as dual and plural lexical nouns.

- (5) Bahing: Independent pronouns and possessive prefixes

	1SG	1DUEX	1PLEX	1DUIN	1PLIN	2SG	2DU	2PL	3SG	3DU	3PL
free	<i>gu</i>	<i>gusu</i>	<i>guku</i>	<i>grsi</i>	<i>gri</i>	<i>gr</i>	<i>gsi</i>	<i>gni</i>	<i>am</i>	<i>amdɔsi</i>	<i>amdɔ</i>
prefix	<i>ɔ</i>	<i>wɔsi</i>	<i>ɔke</i>	<i>isi</i>	<i>ike</i>	<i>i</i>	<i>isi</i>	<i>ini</i>	<i>a</i>	<i>asi</i>	<i>ani</i>

Examples: Bahing *ɔ-gy* ‘my hand’, and, with a genitive-marked nominal, *wam-ke a-pija* ‘the bear’s head’. (Bantawa example: see (27).) In Wambule, number is neutralized in the

free forms; in Hayu it is neutralized in the free forms *gu* ‘1’ and *kōmi* ‘3.HUMAN’. Limbu has only singular prefixed forms; non-singular possessors are represented by the full pronoun⁷ placed before the noun: *ke-him* ‘your^{sg} house’, *kheni phu’ phoya* ‘your^{pl} brothers and paternal uncles’. In Limbu, as in many eastern languages, 3rd person independent forms distinguish only singular and plural (non-singular).

Eastern languages insert a nasal between the prefix and certain kin terms with stop initials: Limbu *ku-mphu?* ‘his elder brother’. In addition to the independent and prefixed forms, Yakkha and Bantawa also have independent possessive pronouns based on the possessive prefix followed by the genitive suffix (Yakkha *ka*), e.g. Yakkha *akka* ‘my, mine’.

3.4 Deictics, demonstrative and interrogative pronouns

Proximal and distal demonstratives are based on deictic elements ‘this’ and ‘that’. These, often with an added nominalizer or the article, can be used as 3rd person pronouns, and can refer to inanimates. A third, anaphoric demonstrative is reported in Yakkha (Schackow 2015: 93–7).

3.5 Adjectives

In general, attributes are expressed by verb forms or deverbal derivations, but a small class of adjectives is usually present. In Limbu, Athpare, Belhare and Bantawa, a number of notions, including ‘new’, carry the 3rd person possessive prefixes (suggesting ‘its newness’) (Michailovsky 2007): thus Limbu *kusɔŋ him* ‘new house’, Bantawa *i-niŋwa* ‘new’, *i-rokwa* ‘old (thing)’ (Doornenbal 2009: 100). Colour terms may constitute a form class which requires special morphology as adnominal modifiers (Limbu *ku-het-la liŋ* (3-red-[COLOUR] thatch) ‘reddish thatch-grass’). In many languages they are verbs themselves, or are used with a verb ‘to be like’ (active participles Limbu *het ke-lɔ~ba* (red ACT-be.like-NML) ‘red’, Bahing *gigi mae-ba* (green be.colour-ACT) ‘green’).

In Limbu, a small subclass of intransitive verbs is distinguished by having a special adjectival form, used in adnominal or predicate position, marked only by the nominalizer *pa*, without the prefix *ke* of the S-agreeing active participle: thus *cu:k-pa mɔna* ‘small man’ vs *ke-le-ba jeba* ‘knowledgeable shaman’ (see van Driem 1987: 21).

3.6 Numerals⁸

In many Kiranti languages (e.g. Jero, Wambule, Athpare) Nepali numbers have replaced the original Tibeto-Burman roots above ‘two’ or ‘three’ in adnominal use, if not in counting or reciting. Nevertheless the German Linguistic Survey of Nepal, which was carried out between 1980 and 1984 (see, e.g., Winter 1991), collected several systems still intact up to ‘ten’, and beyond (Limbu, Dumi). The majority pattern is a decimal system (Limbu, Kulung, Dumi, some varieties of Yakkha, some varieties of Bahing).

An interesting feature is the presence of quinary structures, lexically built on recognizable roots for ‘hand’ or ‘hand’ and ‘foot’, which are probably reflections of old systems with new roots in Yakkha, Bantawa of Ranitar village in Panchthar (Gvozdanovic 1985), and Hayu (Hodgson 1857a: 393). This system was believed to have been lost in Hayu, but an older speaker produced an example spontaneously during a recent enquiry in the village of Adhamara: *na’ung got kolu le* [two hands one foot] ‘15’.

A vigesimal structure was reported for nineteenth century Bahing (Hodgson 1857–1858: 497), even making use of the typologically rare fractional count: *kwong ásim kwong*

áphlo (one 20 one its-half, i.e. halfway to the next integral multiple of the base) ‘30’ (Mazaudon 2010). A vigesimal structure is also reported in Panchtar Bantawa varieties in general (Gvozdanovic 1985: 188).

Many languages have sets of two classifier-suffixes, human vs non-human, rarely a third classifier for round things (Chamling; see Chapter 36), occasionally also one for flat things: Bantawa *bop* ‘round’ (*ik-bop littim* ‘one guava’), *bak* ‘flat’ (Doornenbal 2009: 113–14). In counting, the classifier-suffixes are generally omitted (Sunwar) or replaced by a different suffix (Kulung *ci*, Limbu *si* in numbers from two to nine). Classifiers are not used on demonstratives.

3.7 Time ordinals

Kiranti languages have words identifying days and years as far as four steps into the past and future (Michailovsky 2003); see example (16).

3.8 Case marking

Case is marked by suffixes. All languages have ergative/instrumental, genitive, comitative, general locative and ablative markers. The comitative marker may serve to coordinate nouns. Ablative and allative markers may be suffixed to the locative. Object marking is generally absent or limited to the optional use of the Nepali dative-accusative suffix *lāi* on personal objects. Case markers are not used on 1st and 2nd person pronouns in Bahing, Thulung or Chamling, or on any personal pronoun in Limbu.

3.8.1 Direction marking and the vertical dimension

The vertical dimension is important in Kiranti grammar and lexicon (Allen 1972; Ebert 1994, 1999; Bickel and Gaenszle 1999). All of the languages seem to have separate verbs ‘come up’, ‘come down’, ‘come across (i.e. on the same level)’ (Bahing imperatives *kywø!*, *jywø!*, *piwø!*, respectively) and corresponding derived applicatives (Bahing *kytø!* ‘bring it up!’, etc.).

In Thulung and in eastern languages, ‘altitudinal’ suffixes, related to directional adverbs ‘up’, ‘down’, ‘across’, may appear with deictic elements, e.g. Chamling *udhi* ‘up here’, *tjudhi* ‘up there’, or with nouns: *tujma-di* ‘up in the village’, *juya-ji* ‘down at the river’, *lam-ja* ‘over on the road’ (Ebert 1997b; Rai 1985). Only Limbu (Tamar Khola varieties) distinguishes gravitational from topographic ‘up’ and ‘down’: *thaŋ* ‘up’ vs *tho* ‘uphill’, *mu* ‘down’ vs *jo* ‘downhill, downstream’, *təl-thaŋ* ‘up on the upper storey’, *him-ettho* ‘up(hill) at the house’ (Michailovsky 2015).

Belhare (Bickel 1997c) and Yakkha (Schackow 2015: 183–203) distinguish between demonstratives and directionals that are speaker-centred (Yakkha *tu* ‘up’, *ju* ‘across’, *mu* ‘down’) and those whose referential zero-point is displaced (*to*, *khe/jo*, *mo*).

3.9 Phonæsthetic manner-adverbs

Rich inventories of phonæsthetic expressions, which mainly serve as precisely defined manner-adverbs used in highly restricted verbal contexts, are found in all languages of the group, as well as in Nepali and other languages of Nepal. They are often set off intonationally like a direct quotation, no doubt because they are felt to be onomatopoeic (see also Hayu, Chapter 34 §5.6; Michailovsky 1988a). They may have marginal phonological

elements, characteristic of ‘paralexemic’ vocabulary (§2.2), e.g. *kk*, *ə*, not found in Hayu core vocabulary, in (6).

(6) Hayu phonæsthetic adverbs (D23)

<i>hararararara</i>	<i>dokdo-ha</i>	<i>mi</i>	<i>bilu</i>	<i>bətəkkəi</i>	<i>men</i>
[MANNER.of.flying]	fall.CONJ.CONV-ADV	DIST	tiger	[MANNER.stone.dead]	die.3SG.PST

‘Flying, flying through the air, that tiger fell and died stone dead.’

4 THE VERB

Basic verb forms are composed of a stem and affixes indicating person-agreement, tense and polarity. A lexical verb may use more than one stem; it is often convenient to represent the root by an internally reconstructed form from which the stems can be derived. The lexical root can usually be represented morphophonologically as an extended monosyllable, $|C_i V(C_f)(C_{pf})|$, where C_i reflects the inventory of syllable initials of the language in question, C_f the syllable finals, and C_{pf} a postfinal |t| or |s| (Michailovsky 1975a, 1985).

4.1 Noun-verb predicates

‘Noun-verb predicates’ composed of a preverb (often a noun) and a conjugated root, are found in most languages, e.g. in Limbu, *lay khek* ‘he walks’ (cf. *lay* ‘leg’), *ja raptu* ‘he sharpens it’ (*ja* ‘edge’). Although the verb morphology generally attaches to the verbal part of these expressions, in Chhintang the preverb can be integrated into the prefix string. Thus Chhintang *ja-a-cept-e* (PREVERB.call-2-call-PST) ~ *a-ja-cept-e* (2-PREVERB.call-call-PST) ‘you^{sg} called’ (< |ja-cept| ‘call’) (Bickel *et al.* 2007b: 66).

4.2 Word-families of verbs

Word-families of roots are relics of two no longer productive processes that have created derived applicatives and causatives (Hodgson 1857–1858: 437–41; Michailovsky 1985). These are of two types:

1. Pairs of roots in which an intransitive verb with a voiced stop initial has a transitive partner with an unvoiced initial, aspirated or not, e.g. Bahing *bok-ti* ‘I got up’ (root |bok|) vs *pok-ti* ‘he got me up’ (root |pok|). A few such pairs occur in the western languages. In Belhare and Limbu, the opposition is between unaspirated and aspirated initials, the voicing opposition having been transphonologized to aspiration (Michailovsky 1994). Hayu (Chapter 34) also has a few pairs with non-occlusive initials.
2. Families of roots differing in the postfinal element: postfinal |t| generally has an applicative sense, e.g. Bahing *la* ‘he goes’ (root |la|) vs *lad-a* ‘he takes it’ (|lat|). Postfinal |s| has a causative sense, e.g. Limbu *ha:be* ‘he wept’, *ha:ptu* (applicative) ‘he mourned him’, *ha:psu* (causative) ‘he made him weep’. Verbs with the applicative postfinal |t| are found in all languages,⁹ causatives with postfinal |s| only in Bantawa, Limbu, Kulung and possibly Wambule. The majority of roots with postfinals are not members of word-families and are not applicative or causative in sense.

Productive syntactic means of benefactive and causative formation are discussed later (§§5.4, 6.3).

4.3 Agreement morphology

Suffixed agreement markers occur in all languages; prefixed ones in Khaling-Dumi, South-Eastern and Far Eastern languages (except Yamphu).

Almost all affixes of the intransitive paradigm also appear in the transitive one. The transitive paradigm shows more or less precise agreement with two arguments except in many Sunwar dialects, which have agent agreement only (Borchers 2008). ‘Conjugations’ that are identified in many descriptions generally reflect different root postfinals and stem alternations, not differing affix paradigms.

4.3.1 *Prefixed agreement markers*

Prefixed agreement markers are less numerous than suffixed ones. $1\text{SG}>2$, $1\text{SG}>3$ and $3\text{SG}>3$ forms are never prefixed. The most widely prefixed forms are those for a 2nd person argument (including intransitives, but excluding $1>2$). Other prefixes have values like ‘1IN’, ‘3PLAS’ and ‘3SGA’.

Dumi and Khaling have only one agreement prefix, which marks 2nd person forms (except $1\text{SG}>2$), but also $3>1$ forms (see Table 33.1). This has given rise to the hypothesis of a conflation of ‘2nd person’ and ‘INVERSE’ prefixes.¹⁰

Yakkha is exceptional among the Far Eastern languages in having only one agreement prefix, an assimilating syllabic nasal indicating 3rd plural; it also has a homonymous negative prefix. The Yakkha (also Belhare) 2nd person suffix *ka* has the distribution mentioned above for the 2nd person prefix, unlike that of any other agreement suffix. Chhintang has eight prefixes (including the negative).

4.3.2 *Suffixed agreement markers*

Suffixed agreement marking systems are highly diverse, although several individual morphemes are widely identifiable across the group. In transitive paradigms, two partly conflicting principles are often reported as governing suffixal agreement (excluding the Yakkha 2nd person suffix mentioned above): (1) ‘hierarchical’ agreement: the first agreement suffix on a transitive verb indexes the argument which ranks highest on a person hierarchy $1>2>3$, regardless of its case role; and (2) role-marking agreement, which generally includes markers for the configurations ‘ $1\text{SG}>2$ ’ and ‘ 3O ’ and overrides hierarchical agreement; (3) when the higher-ranking argument is singular, the number of the hierarchically lower-ranking argument may be marked by a secondary number marker, which may be different from the number-markers for primary arguments.

Table 33.2 shows selected forms of an intransitive and a transitive verb in Khaling and in Bahing, two western Kiranti languages. The forms show the agreement markers with hierarchically ranking arguments of particular persons and numbers (rows in the tables) according to their syntactic function (S, O or A; columns in the tables). In both Khaling and Bahing, for 1st and 2nd person non-singular arguments (here represented by 1PLEX and 2PL), S and A agreement use identical stems and identical hierarchical agreement suffixes (columns 1 and 3 for each language). This identity is broken for singular arguments (rows 1SG and 2SG) because these forms show 3rd person object agreement (they also show object agreement in $1\text{SG}>2$ forms).

Further, in Khaling, for 1st and 2nd persons, S and O agreement are identically marked (columns 1 and 2) (this is also true in Dumi and in Hayu), so that, for singular arguments, all three columns have identical markers. This is not the case for Bahing (or for Thulung

TABLE 33.1 KHALING: INDICATIVE PARADIGMS OF LOP VT (*LŌMNE* ‘CATCH’ AND SOP VI (*SŌMNÉ* ‘HAVE ENOUGH’)

TRANS.	1SG O	1DUIN O	1DUEX O	1PLEX O	1PLIN O	2SG O	2DU O	2PLEX O	3SG O	3DU O	3PLEX O	3PL O
1sg A						lōm-n-e	lōm-n-u	lōm-tenu	lob-u	lob-usu	lob-unu	lob-utunu
1duin A						lōm-n-i	lōm-tensu	lōm-tenu	lob-uta	lob-utusu	lob-utunu	
1duex A									lep-i	lep-itü		
1plin A									lep-u	lep-utu		
1plex A									loop-ki	loop-tiki		
2sg A	i-lōm-n-γ i-lep-ata								i-lō-b-u	i-lō-p-su	i-lō-p-nu	
2du A	i-lōm-η-asu i-lep-atastu								i-lō-p-te	i-lō-p-tesu	i-lō-p-temu	
2pl A	i-lōm-η-anu i-lep-atastu	i-lep-i i-lep-itü	i-lep-u i-lep-utu	i-loop-ki i-loop-tiki	i-loop-ka i-loop-taka	i-loop-i i-loop-te	i-loop-i i-loop-te	i-loop-i i-loop-te	i-lōm-ni	i-lō-b-u	i-lō-p-su	i-lō-p-nu
3sg A	i-lōm-γ-γ i-lep-ata								i-lō-p-te	i-lō-p-tesu	i-lō-p-temu	
3du A	i-lōm-γ-asu i-lep-atastu											
3pl A	i-lōm-γ-anu i-lep-atastu											
REFL.	1SG S	1DUIN S	1DUEX S	1PLEX S	1PLIN S	1PLEX S	2SG S	2DU S	2PLEX S	3SG S	3DU S	3PLEX S
	lōmstüp lōmstu	lep-sijü lep-sijü	lep-sijü lep-sijü	loop-siki loop-sikitü	loop-siki loop-sikitü	loop-siki loop-sikitü	lōm-msi lōm-tesi	lōm-msi lōm-tesi	lōm-sijü lōm-sijü	lōm-sijü lōm-sijü	lōm-sijü lōm-sijü	lōm-sini lōm-temu
INTR.	1SG S	1DUIN S	1DUEX S	1PLEX S	1PLIN S	1PLEX S	2SG S	2DU S	2PLEX S	3SG S	3DU S	3PLEX S
	sōm-γ-γ sōm-ata	sep-i sep-itü	sep-u sep-utu	soop-ki soop-tiki	soop-ka soop-taka	soop-ka soop-taka	i-soop i-soop-te	i-soop i-soop-te	i-soop-i i-soop-te	sep-i sep-itü	sep-i sep-itü	sōm-nu sōm-temu

or Wambule). In Bahing, the markers for 1st person arguments in S and O functions (columns 1 and 2) are different. This gives an idea of structural differences in the paradigms. Still another structure is found when 3rd person object agreement is marked for agents of all numbers, not singular only, as in many languages.

Where the highest-ranking argument is singular, the number of the other core argument can be indicated by a ‘secondary number marker’ (Khaling: *su* ‘DU’, *nu* ‘PL’: see Table 33.1).

A particularity of paradigms in many languages is that 3>3 forms show number agreement with the argument with the highest number (see Tables 33.1 (Khaling) and 34.1 (Hayu)). The notations ‘3>3 (DU)’ and ‘3>3 (PL)’ indicate such agreement.

4.3.3 ‘Impersonal’ or detransitivized 1st person object forms

Several researchers (e.g. Allen 1975: 49) have remarked that speakers had difficulty producing 2NSG>1NSG and 1NSG>2 forms, often responding with non-finite forms. This difficulty may account for apparent paradigm simplifications and the existence of alternative forms in these and certain other parts of the agreement paradigm.

In Limbu the 2SG>1SG form is *ke-dzok-ya* (2-do-1sgSO.NPST) ‘you^{sg} do to me’, as might be expected (the past is similarly formed), but there is a single 2>1(NSG) form *a-ke-dzok* (1PLIN-2-do), with only the prefix *a-* (normally ‘1INCLUSIVE’, but clusivity is neutralized in this context) to indicate 1st person. (In contrast, Limbu 3>1 forms are fully marked for 1st person number and clusivity, and 3SG>1 forms are identical to 1st person intransitive forms (as in Khaling).) In Yakkha there are similar, unique non-singular forms for both 2>1(NSG) and 3>1(NSG). The only person marker on the 2>1(NSG) form in Yakkha is a 2nd person marker, and there is no explicit person marker on the 3>1(NSG) form. Although both of these Yakkha forms turn out to be identifiable, because they are unique in the paradigm, many distinctions are neutralized.

In parallel with these apparent neutralizations in the paradigm, ‘impersonal’ or ‘detransitivized’ 1st person-object forms are found, sometimes as optional or dialectal variants of the appropriate part of the regular 23>1 agreement paradigm, in the South-Eastern and Far Eastern languages (Ebert 1991; see the chapters on Belhare and Chamling in this volume). An intransitive form agreeing with the A (which retains ergative marking) is accompanied by a (usually) preposed element, often related to a word for ‘person, other’, indicating a 1st person (especially non-singular) object. In (7) (Limbu), a formally intransitive form, *mehakte* (cf. transitive *mehaktu* ‘they sent it’) with a preposed element (*japmi/napmi* ‘person, other’), indicating a 1st person non-singular object, occurs in the same sentence and with the same referents as a regular 3PL>1PLEX transitive form.

TABLE 33.2 SELECTED FORMS OF KHALING AND BAHING

KHALING			BAHING		
X is S	X is O	X is A	X is S	X is O	X is A
X=	X has enough	he catches X	X catches him	X stands	he beats X
1SG	so̥m-ŋʌ	i-lo̥m-ŋʌ	lob-u	ram-ŋa	tym-ji
1PLEX	soɔp-kʌ	i-loɔp-kʌ	loɔp-kʌ	rap-ka	typ-ki
2SG	i-soɔp	i-loop	i-le:b-u	ram-e	tym-e
2PL	i-so̥m-ni	i-lo̥m-ni	i-lo̥m-ni	ram-ni	tym-ni
3SG	soɔp	le:b-u	le:b-u	ram	tyb-a

- (7) Limbu: transitive and detransitivized 1st plural object forms (DANCEs15)
- anige nurik me-mett-ige-ay ciffhi japmi me-hakt-e*
 1PLEX well 3PL.AS-do-1PLEX-CONJ letter 1NSG.O 3PLAS-send-PST
 ‘They [a group of dancing partners] treated us^{plex} well and sent us letters.’

It is not clear that these forms are ‘detransitivized’ since, as we have seen, the regular 3>1 forms in, e.g., Limbu, Khaling and Hayu are already identical to intransitive forms. ‘Depersonalized’ (because personal agreement with the object is lost) would perhaps be more accurate.

4.4 Tense morphology

Two basic tenses, usually roughly labelled ‘past’ and ‘non-past’, are identified in most languages. In the Western languages Khaling (Table 33.1), Sunwar, Bahing and Thulung, past tense is marked in the suffix string by *t* (somewhat disguised in Thulung); in Dumi, the forms are very similar to those of Khaling, its closest relative, except that *t* marks non-past. Tenseless finite forms are reported in some contexts in some languages (e.g. the Belhare ‘zero-marked form’ (Chapter 35)). There are also various composite tenses with auxiliaries. Ebert (1994: 29–46) is a useful review of tenses and their negations.

Verbs in Eastern languages have morphophonologically related past and non-past stems. In Limbu and Bantawa, the present stem has the form of a phonological syllable C_iV(C_f) and can stand alone, while the past stem ends with the initial of an obligatory following syllable C_iV(C_f)C_i, e.g. Limbu *jep-jeb-* ‘weep’, *khem-kheps-* ‘to hear’ (cf. Doornenbal 2009: 126–9 for Bantawa). The two stems are commonly referred to, respectively, as ‘preconsonantal/pre-pause’ and ‘prevocalic’ (e.g. van Driem 1987: 71), but this is only a (not entirely accurate) statement of their distribution (see Michailovsky 1986, 2002); in fact, final *p* never has the allophone *b* in Limbu (see §2.3) and no phonological rule explains the alternation *m~ps*.¹¹ The phonological characterization obscures the fact that these are morphologically distinct stems.¹²

In several languages the non-past with present sense is marked as progressive (e.g. by a compound verb, §4.8) and the simple ‘non-past’ stem is only found in negative forms. The progressive auxiliary appears to have become part of the stem in some languages, and many descriptions identify distinct negative stems as well.

The semantics of tense is not discussed in the present article. Bickel (1996) analyses the semantics and pragmatics of tense in Belhare.

4.5 Negation

All languages except Athpare have a negative prefix, and some a suffix, used with finite verbs. Bahing, Thulung and Khaling only have prefixed negation. In Dumi, finite verbs are negated by the suffix *nə*; preterite negative forms also have the prefix *mə*: *tsikh-a* ‘it happened’, *mə-tsikh-i-nə* ‘it didn’t happen’. Non-finite forms use only the prefix (van Driem 1993: 149).

The negative prefix (*mVn*) is also limited to past forms in Yamphu, Kulung and Bantawa, and no suffix is found on these forms. The non-past has a suffix. In Limbu, all finite forms have both a negative prefix and suffix. Many languages have a special negative prefix for non-finite forms and nominals (Hayu *maay*, Kulung *man*, Bantawa *ma?an*, Limbu *me:n*). Only Hayu has a special prohibitive ‘don’t!’: *tha phi!* ‘don’t come!’.

Negation of past forms is often based on a present or perfect form. Thus in Bahing the negative form in (8) is based on the non-past *ba-ya* ‘I eat it’, which itself resembles an

intransitive form, e.g. *pi-ŋa* ‘I come’ (cf. *ba-tɔŋ* (eat-1SG>3PST) ‘I ate it’). The past marker *t* is replaced by an irrealis *wa*, as also in Thulung and Khaling (Ebert 1994).

- (8) Bahing negative past (RANGADIPs7)

<i>ma!</i>	<i>gu</i>	<i>ma-ba-ŋa-wa</i>	<i>ɛm</i>	<i>bebatsa-ds</i>
NEG	I	NEG-eat-1SGNPST-(NEG)	PROX-NOM	child-PL
‘No! I didn’t eat the children!’				

Athpare negation is marked by a suffix; past forms are compounded with a specialized auxiliary, which carries the negative marking (9).

- (9) Athpare negative of *khad-a-ŋ-e* (go-PST-1SG-PST) ‘I went’ (Ebert 1994)

<i>khat-ni-ŋ-get-ni-ŋ-na</i>
go-NEG-1SG-NEG.AUX-NEG-1SG-NMLZ
‘I didn’t go’

In Bantawa the past negative is marked by a prefix and special suffix morphology reflecting a negative auxiliary (Doornenbal 2009: 161–72). See (10) for a Chhintang example.

Some languages have a specialized negative for the existence predicate: Limbu *hop/hopte* ‘there is/was none’, Bantawa *matdiŋ*, Chamling *paina*. In others, the existence predicate is negated regularly.

4.6 Free prefix ordering

Free ordering of prefixes is reported in Bantawa (Rai 1985) and in Chhintang (Bickel *et al.* 2007b), where it is reported that the three prefixes (*u*, *kha*, *ma*) in (10) can appear in any of the six possible orders without any semantic or stylistic difference. Note that of the three prefixes, only one, *u*, is strictly pronominal: *kha* is the ‘impersonal’ object of a detransitivized 1NS object form (§4.3.3), while *ma* is the negative prefix linked to the negative auxiliary *jokt*. ‘Preverbs’ of noun-verb predicates can also get permuted.

- (10) Chhintang (Bickel *et al.* 2007b: 44)

<i>u-kha-ma-cop-jokt-e</i>
3NSGA-1NSO-NEG-see-NEG-PST
‘They did not see us.’

4.7 Copulas

All Kiranti languages have one or more copulas, which can be omitted in straightforward, 3rd person singular identificational contexts (e.g. the second part of (11) below).

Limbu also has a full set of agreement markers that are suffixed to nominal predicates, e.g. *a* ‘1SG’, *adi* ‘1PLIN’, *adige* ‘1PLEX’, *ne* ‘2SG’; the 3SG form is zero. Yakkha and Belhare have very incomplete sets of such markers.

- (11) Limbu <TRADEs267>

<i>ani-tna</i>	<i>tɔksɔŋ</i>	<i>ke-wa?-ba-adi</i>	<i>ɛtna</i>	<i>kɔ</i>	<i>seŋga</i>
1PLIN-LEVEL.LOC	highland	ACT-be-NOM-1PLIN	here.LEVEL	TOP	plains
‘Over in our ^{plin} place, we ^{plin} are highland dwellers; here it’s lowlands.’					

Limbu uses a variety of posture verbs to indicate canonical location/existence (fields ‘lie’, houses ‘sit’, livestock ‘stands’, limbs are ‘hung on’) in addition to a more general verb of existence or circumstantial presence (van Driem 1987: 61–8).

Dumi (van Driem 1993: 168) and Khaling (Jacques *et al.* 2015) have separate existence/location ‘be’ for animate and inanimate subjects.

The verb ‘be, exist’ is usually negated in the same way as other finite verbs. For special negative forms, see §4.5.

4.8 Periphrastic tenses: perfect and progressive

Perfect tenses are formed of nominalized forms or converbs with auxiliaries ‘to be’, or by compounding with vectors ‘sit/remain’ or ‘put/keep’. The pluperfect uses the past tense of the auxiliary.

In Thulung, both the perfect and the progressive use a nominalized finite verb with an auxiliary ‘be’; the progressive can also use a simultaneous converb and auxiliary (Allen 1975: 85; Lahaussois 2002: 176–7). In perfect forms, the auxiliary is impersonal. In Wambule, the perfect is composed of a past/passive participle (a nominalized finite form) followed by an impersonal form of the verb |bak ~ ba:| ‘be/sit’: *lwa-ŋ-meɪ ba:-Ø-m* (go-1SG-FCT be-23SG-AFF) ‘I have gone’ (Opogenort 2004: 318).

In Dumi (van Driem 1993: 237–43), the perfect is made up of a nominalized finite past form followed by a present auxiliary. The auxiliary with transitive verbs is an impersonal form of *gini* ‘to exist, be (attributional)’; with intransitive verbs, the auxiliary *mini* ‘be (locational)’ agrees with the subject. The progressive is expressed by compounding with a vector ‘to continue’.

In the Limbu perfect (van Driem 1987: 163–9) the verb is a finite past form, with the conjunctive subordinator *-ay* (§8.1) (van Driem’s ‘perfect gerund’) followed by a non-past copula. If the verb is intransitive, the copula has the same (i.e. subject-agreeing) form as the verb; if it is transitive, the copula is impersonal.

- (12) Limbu perfect with impersonal auxiliary (MARIs100)

the-the *ke-thakt-um-ay* *wa?*
what-what 2-bring.up-2PL>3-CONJ be.NPST
‘What-all have you brought up [as bride-price]?’

Limbu perfects and progressives can choose among posture-verb copulas. Thus, for example, the use of the auxiliary *pøt* ‘be suspended’ in (13) suggests that the speaker expects to get back his money.

- (13) Limbu perfect with auxiliary ‘be suspended’ (SOGHAs175)

khonn-aŋ *søi* *jay* *pur-uŋŋ-aŋ* *pøt*
that.one-conj 100 rupee give-1SG>3-CONJ be.suspended

The use of the posture-verb ‘lie’ in *ness-uŋŋ-aŋ ne* (lay-1SG>3SG-CONJ lie) ‘I’ve laid it out [the money]’ suggests the resultant state. Progressives use the conjunctive or simultaneous subordinator with a non-past or past verb form and an auxiliary of the same tense. For example, *me-imm-aŋ me-ne* (3PL-sleep-CONJ 3PL-PROG) ‘they are sleeping’, or, with the simultaneous subordinator, (14).

- (14) Limbu simultaneous subordinator with auxiliary *juy* (< ‘sit’) (NAROs135)

blastin *me-phøksu-’lle* *bela* *inga* *tɔ:k* *ca-ŋ-lɔ* *juy-aŋ*,
explosives 3PL-set.off-DEF-GEN time 1SG meal eat-1SG-SIMUL PROG-1SG,
bara *bødze*
twelve o’clock

‘When they set off the explosives, I was eating my meal, at 12 o’clock.’

Bantawa uses the auxiliaries *juyma* ‘sit/remain’ ~ *jukma* ‘put down’, either compounded with the verb as vectors or with a nominalized complement to form the perfect. Yakkha languages use vectors: Yamphu *mindep-pe:-t-c-w-e* (forget-RESULT-PERF-DU-3O-FACTITIVE) ‘They^{DU} have forgotten it’ (Rutgers 1998: 221).

Negative perfects use a negative perfect converb (§4.9.5), which is also used in ‘adverbial’ subordinate clauses ‘without Ving’, or negative perfect vectors. The negative perfect often appears as the standard past tense negation (8).

4.9 Non-finite forms

All Kiranti languages have infinitives, converbs and present active participles; many also have a past/passive participle.

4.9.1 Infinitive

The infinitive is a monosyllabic stem with a suffix such as *ma*, *mu*, *ca*. Infinitives are used as nouns, e.g. as the heads of propositional sentence arguments (36) and of propositional complements with modal and perception verbs. The infinitive can have arguments, and can have a suffix indicating plurality of an object or reflexivity, e.g. in Limbu: *khən-ha² thak-ma-si poj* (3-PL bring.up-INF-PL must) ‘[we] must bring them up here’; *cum-cum tɔŋ-ma-siŋ poj* (friend-friend agree-INF-REFL must) ‘friends must agree among themselves’.

4.9.2 Purposive

The purposive (or ‘supine’) is also marked by a suffix (such as *si*, *se* in the eastern languages). It occurs mainly as a complement of motion verbs. In Limbu (15), Chamling and Thulung it can have a possessive pronominal prefix, indicating its logical object.

- (15) Limbu purposive (SOGHAs133)

<i>a</i>	<i>himdaym'-el-le</i>	<i>ku-la-se</i>	<i>peg-aŋ-ay...</i>
1.OBL	wife-DEF-GEN	3.OBL-bring.home-PURP	go-1SG.SO-CONJ
'I went to bring home my wife and ...'			

In a parallel Kulung example, the wife appears in the absolute case: *m-to:ma lam-s k^ha-let-a* (3OBL-wife find-PURP go-away-PST) ‘He went off to find his wife’.

4.9.3 (Present)/active participle

All languages have a non-finite active participle, which functions as the head of an adnominal relative expression referring to the actor (A or S) of its clause. In Hayu (suffix *ji*, Chapter 34) and Yamphu (suffix *khu-khuba*) the active participle of an intransitive verb signifies non-past tense: Yamphu *hedajna-be² pen-ghu* (H.-LOC sit-ACT) ‘(the one) who is staying in Hedangna’ as opposed to a past/passive participle (18). Active participles in other languages are not reported to indicate tense.

- (16) Bahing tenseless active participle <BEGINs133>

<i>santho</i>	<i>niware-ŋ</i>	<i>bek-ba</i>	<i>ɔsi</i>	<i>loba</i>
last.year	year.before.last-INTENS	die-ACT	1DUIN	younger.brother
'our ^{DUIN} brother who died years ago' (cf. finite <i>bekta</i> (die-3SG.PST) 'he died')				

In South-Eastern languages and Limbu the active participle has a prefix *kV*: Bantawa *ka-tok* ‘respected (man)’ (< *tok* ‘receive’); *ka-tuk mina* ‘sick man’. Bantawa lexicalized

actor-nominals add the nominalizing suffixes *pa* ‘MASC’/ *ma* ‘FEM’: *ka-d^huk-pa-ci* (ACT-hit-MASC-PL) ‘blacksmiths’. Limbu has the prefix *ke*, always with the suffix *pa/ma*, which, as in most Kiranti languages, is adnominal.¹³ Dumi has suffixed (*k*)*pi*: *ma-tsa:p-pi* (not-able-ACT) ‘incapable’. The participle can be marked for reflexivity: *tsem-si-kpi-mil* (play-REFL-ACT-PL) ‘players’. Kulung has the suffix *pa*.

- (17) Kulung active participle (Tolsma 2006: 161)

<i>bayar</i>	<i>k^hai-p</i>	<i>i:bum</i>	<i>mis-lo</i>	<i>tup-a-t^he</i>
market	go-ACT	one	person-COM	meet-PST-REPORT
'He met a man who was going to the market.'				

4.9.4 Past/passive participle

Non-finite past/passive participles (past with intransitive verbs) are less common; their function may be filled by patient nominals or by a general clause nominalizer. Examples are found in Hayu (suffix *ta*), Yamphu (*praʔa*), Kulung (*mpa*), Limbu (*muna*). Some of these can refer (as in Hayu) to any clause participant except A.

- (18) Yamphu past/passive participle on an intransitive clause (Rutgers 1998: 205).

<i>kani</i>	<i>miju-ra</i>	<i>lem-braʔa-ji-ro</i>
1PLIN	DIST.across.MED	come-PPT-NSG-REPORT
'we ^{PLIN} who came from over there (they say)'		

For the Yamphu past/passive participle of a transitive verb, see (46) ('given').

Limbu *muna* serves as an action noun: *thi pi-muna-si-ba cələn* (beer give-ACTN-PL-NMLZ custom) ‘the custom of giving them beer’ (DANCEs139). Note that this form can take the plural suffix. With the addition of the prefix *men* it becomes the negative of the active participle (van Driem 1987: 207–9).

4.9.5 Converbs

By ‘converb’ is generally meant a non-finite form that serves, usually with a suffixed function-marker, as the head of an ‘adverbial’ subordinate clause. Ebert, in her articles on converbs in general and in Kiranti (2008) cites van Awera: converbs are –argumental, –adnominal, +dependent and –finite, but has some reservations concerning the last of these. Finite subordinates are usually considered as being governed by subordinating conjunctions, but where they are marked by a suffix or postposition attached to the verb, the latter may resemble a converb. But, like Ebert, I will not use the term ‘converb’ (or, a fortiori, ‘gerund’) for such finite forms.

Ebert identifies a negative converb in seven Kiranti languages.

- (19) Limbu conjunctive subordinator and negative converb (van Driem 1987: 182 cited by Ebert 2008: 79)

<i>me-dho:kt-u=aŋ</i>	<i>me-dzɔ-i?</i>	–	<i>men-dho:k-?e</i>	<i>me-dzɔ.</i>
3pl.SA-cook-3O=CONJ	3pl.SA-eat-q?	–	NEG-cook-NEG.CONV	3PL.SA-eat
'Do they cook it and eat it? – They eat it without cooking it.'				

The same converb is found in negated perfects. Thus Limbu *mem-be:k-?e wa:-?e* (NEG-go-NEG.CONV be-1SG.NPST) ‘I have not gone.’ Cf. *pe:g-aŋ=aŋ wa:-?e* (go-1SG.PST-CONJ be-1SG.NPST) ‘I have gone.’ (Ebert 2008: 84). Positive perfects are formed variously: either with a converb and auxiliary, as in the Limbu example just cited, or by a compound verb.

Ebert finds a simultaneous ('while Xing') converb in six of the languages, and an equivalent form with a finite verb and the suffix *rɔ* in Limbu (see also the Limbu example (14)).

- (20) Limbu finite simultaneous subordinate (SOGHAs98)

<i>khɔ-n-etyo-nu</i>	<i>khɔn</i>	<i>ku-samm-en</i>	<i>lem-u-rɔ</i>	<i>thakt-u</i>
DIST-DF-downhill-from DIST-DF 3-soul-DF coax-3SG>3SG-PROG bring-up-3SG>3SG...				

'[The shaman] brought it up, brought his soul up from down below, coaxing it...'

Again, progressive tenses in some languages are based on this form, e.g. Limbu *sɔys-um-be-rɔ jag-ige* (sell-1PLEX>3-PROG be.in-1PLEX) 'we were engaged in selling it'.

Hayu has two unique reduplicated converbs, and a number of others consisting of verb stems with suffixes, most of them ordinary nominal case markers (see Chapter 34).

4.9.6 Derived nouns

Specialized actor, patient, place, etc., nominals, are based on non-finite forms. Hayu has an unusually complete set (Chapter 34). In Yamphu (Rutgers 1998) there is a second actor nominal (suffix *jay*) used for professional rather than conjunctural actors: *na se?-jay-ji* (fish kill-AGT-PL) 'fishermen'; a locative/instrumental nominal (suffix *tha~tham*): *wawa pen-dham jaksa* (brother stay-LOC/INST hut) 'shed where elder brother stays'; *sap-tha* 'pen'; a patient/instrument nominal based on *cira* 'thing' (cf. Bantawa *kha*): *cet-cira* 'plow'. In Sunwar, we find *tīke* 'INSTRUMENT': *b[e]-tīke* 'pen', and *ci* 'TEMPORAL': *dhol p/u/ci-nati* (drum loosen-TEMP day) 'day when the drums are loosened'. These are basically nouns, but may bring along some arguments, and in some cases serve as adnominals (like the modifiers of 'shed' and 'day' above).

4.10 Compound verbs

The term 'compound verb' is used in the Indian linguistic area for a special type of serial verb in which the main verb, V1, a stem form expressing the main semantic content of the compound, is followed by a 'vector' or 'light' verb, V2 (often related to a full verb), which contributes a peripheral meaning in the areas of direction or aspect. In principle, the compound verb is the predicate of a single clause, with a single argument structure, and designates a single action or situation, but in the absence of a clear functional definition I have not specified strict morphological criteria for the identification of 'compound verbs' in Kiranti. The typology of these constructions in Nepal has been studied by Pokharel (1991) (for Nepali), Ebert (1994), Bickel (1996) and Doornenbal (2009).

In Indo-Aryan, non-finite V1s are the rule, but in Kiranti the tendency is for V1 to be a finite form, except in Hayu and some western languages. Vectors in Kiranti usually have both intransitive and transitive forms, which facilitates morphological harmony between V1 and V2. Vectors are also often observed to have somewhat irregular morphology.

In Hayu, compounding is not very prominent: the most plausible candidate vector is *lat* 'go', which functions as a 'motionalizer' (*lon laxtsem* 'he ran off') and in the *go-passive* (see Chapter 34 §6.8).

In Bahing, as in Hayu, V1 is a stem form, here (21) with the 'ponent' vector.

- (21) Bahing (NINAMs152)

<i>cacar-than</i>	<i>bhādqo</i>	<i>tok</i>	<i>dzyl-tase</i>	<i>ne</i>
four.REDUPL-CLF	pot	drop	put-3DU.PST	REP

'They each dropped off four cooking pots.' ['Dropped' in part because the agents are airborne.]

Wambule has a number of ‘auxiliary verbs’ with vector-like semantics. With the ‘dispatching’ and ‘ponent’ auxiliaries, the V1 is marked by a connector: *uj kimsul hwal-ti co-y-m* (1ERG door open-CONN send-1SG-ASS) ‘I opened the door’ (Opgeenort 2004: 420). ‘Motionalizers’ directly follow a root form of the V1.

In Thulung (Allen 1975: 72–7), Dumi, and all the eastern languages, both V1 and V2 have some agreement morphology. In Dumi (van Driem 1993: 197–214), the agreement prefix (when required) appears before the compound, and the full suffix string after it, suffixed to V2. V1 has the regular stem alternant appropriate to the complete form, and an abridged suffix string depending on the shape of the stem. Dumi: *aj dza: dza-y-pət-tə* (I rice eat-1SG-ALLATIVE-NPT-1SG) ‘I’m going to eat’ (‘go’ implies displacement here); in many Kiranti languages ‘to eat’ would be a non-finite purposive complement. Three verbs with phasal *Aktionsart* semantics, ‘to begin’, ‘to finish’ and ‘to be about to’, are identified as ‘pseudo-aspectivizers’ by van Driem because they function as main verbs with an infinitive complement (e.g. *tsyet-ni lo:* (ache-INF begin) ‘it began to ache’) and not as V2 vectors.

In Athpare, agreement prefixes appear only before V1 but the tense-marking suffix can appear only after V2.

- (22) Athpare (Ebert 1997a: 71)

<i>a-mund-um</i>	<i>lett-um-e</i>
2-forget-2PL	V2:TELIC-2PL-PST
‘You ^{PL} completely forgot it.’	

Limbu is the only language that compounds whole forms with prefixes: *kemm-oʔr-i kem-sur-i* (3PL>2-burn.PA-2PL 3PL>2-finish.PA-2PL) ‘They^{PL} already informed on you^{PL}’ (TRADEs247). The two verbs may differ in agreement.

Example (23) has two compound verbs: the vector is aspectual in the first, directional (a ‘motionalizer’) in the second. (Note that *nu:ksi* is abbreviated from *nu:ksige* ‘we^{PLEX} returned’.)

- (23) Limbu intransitive vectors ‘finish’ and ‘come.down’ (MARIs119)

<i>pa:nn-en</i>	<i>ton-ε</i>	<i>cur-ε-aj</i>	<i>jammu anige</i>	<i>etjo</i>	<i>nu:ks-i</i>
speech-DF	agree-PST	finish-PST-CONJ	again	we.PLEX	here.down return-1PL
<i>j'je-ige.</i>					
come.down-1PLEX					

‘Once the affair was agreed, we^{PLEX} came back down here.’

In the same Limbu text we find the conjunctive subordinate (§8.1) syntagm *nu:ks-i-aj j'je-ige* (MARIs77), with the linker *aj* ‘and’, apparently with the same meaning as the compound of (23).

The division of labour between compound verbs, verb-complement constructions and conjunctive subordinate constructions is highly variable between languages.

4.11 Complement-taking verbs

Modals, phasal aspectuals and verbs of expression and perception may take clausal complements, either non-finite or finite.

The transitivity and argument structure (agreement) of a complement-taking verb may be largely determined by the semantics of a non-finite complement. Thus in (24), ‘quit’ shows object agreement with the object of ‘push’, its infinitive complement.

- (24) Wambule: Agreement of complement-taking verb (Opgeort 2004: 418)

uy un njap-si pli-ni-m
1 2 push-INF quit-1SG>2SG-ASS
'I will stop pushing you^{SG}.'

5 TRANSITIVITY FRAMES AND DIATHESES

5.1 Intransitive

Intransitive (including reflexive) subjects are unmarked. Opgeort (2004: 149–52, Chapter 37) reports that the Wambule ‘source’ case-marker, which is clearly analogous to the ergative/instrumental of other Kiranti languages, is used on both intransitive subjects and transitive agents to draw attention to their active role, and may be omitted otherwise.

5.2 Transitive

Ergative case-marking of transitive agents is usual in Kiranti languages, but its Wambule equivalent, the ‘source’ marker, may be omitted if the agentivity is regarded as obvious or unimportant. Case marking of the agent of impersonal 1st person object forms (§4.3.3) and antipassives (§5.3.2) varies between languages.

- (25) Bahing transitive (KOKTISALAs13-14)

...ɔ-wa-dɔsi-m gu sa:jisi ɔ-sami-dɔsi-mi'
...1SG-elder.sibling-DU-ERG 1 kill-23DU>1SG.NPST 1SG-sister-DU-ERG
'My elders will kill me, my two sisters will!'

In this transitive sentence, the A ('siblings') is marked as ergative; the object ('me') is absolute; the A is repeated as an anti-topic ('sisters').

Verbs of transfer and giving can have two absolute objects. The recipient or beneficiary is identified as a core argument, the ‘primary object’ (PO) (Dryer 1986), by the verb agreement morphology.

The basic transitive argument structure admits some variation. Most Sunwar dialects lack object agreement morphology, and human objects are marked by *kali*. Transitive objects in several languages (Thulung, Wambule, Jero, Chamling, Bantawa, Puma, Athpare) may optionally be marked by the Nepali dative/accusative postposition *lāi* (33). The optional use of ergative-marking on subjects and agents in Wambule has been mentioned earlier.

5.3 Diatheses

Van Driem (1987: 271) remarks that Limbu transitive verbs can be ‘labile’, that is, conjugated as intransitive in certain uses and transitive in others. Weidert and Subba (1985: 82–3) recognized the Limbu antipassive. Doornenbal (2009: 221–9) distinguishes three valence-lowering configurations in Bantawa.

5.3.1 Middle or ‘break’ verbs

‘Break’ or ‘middle’ verbs have both intransitive and transitive forms, such that the S in the intransitive use (e.g. *the glass* in *The glass broke*) corresponds to the O of the transitive (*He broke the glass*), which alone can have an A (see Fillmore 1970).

- (26) Bantawa middle or ‘break’ verb (Doornenbal 2009: 228)
- | | |
|---|---|
| <i>samba ker-a</i>
bamboo break-PST (intransitive)
‘the bamboo broke’ | vs <i>kho-sa-ʔa samba ker-u</i>
he-PRN-ERG bamboo break-3>3(SG)(transitive)
‘he broke the bamboo’ |
|---|---|

5.3.2 Antipassive

In the antipassive diathesis, a transitive proposition (27) is related to an intransitive one in which the S corresponds to the transitive A (28). The action is presented as an activity; the object, whether mentioned or not, does not have definite reference. Doornenbal remarks that the case marking of the antipassive agent is ‘doubtful’: either absolutive or ergative (similarly, in Limbu, van Driem 1987: 271–3 note 7 and exx. (20), (27)).

- (27) Bantawa transitive (Doornenbal 2009: 224)
- | | |
|--|---|
| <i>sjam-ʔa i-pa-ʔo</i>
syam-ERG 3OBL-father-GEN | <i>i-jan khis-u</i>
3OBL-money steal-3>3(SG) |
|--|---|
- ‘Syam stole [transitive form] his father’s money.’
- (28) Bantawa antipassive (Doornenbal 2009: 224)
- | | |
|--------------------------------------|--|
| <i>sjam khis-a</i>
Syam steal-PST | |
|--------------------------------------|--|
- ‘Syam steals (?stole) [intransitive form].’

Verbs of the ‘break’ type can only be given an antipassive sense (‘he breaks stuff’) by including an indefinite pronominal object (Bantawa *kha*) immediately before the intransitive form of the verb. The verb then indexes the actor as S.

- (29) Bantawa antipassive of a ‘break’ verb (Doornenbal 2009: 226)
- | | |
|--|--|
| <i>k^ho-ci-ʔa k^ha mi-hit</i>
3-PL-ERG ANTIP 3PL-scorch [intransitive form] | |
|--|--|
- ‘They burn things.’ [Cf. transitive *i-hit* ‘they^{pl} burn it’, *mi-hitt-uci* ‘they^{pl} burn them^{ns}’]

Doornenbal (2009: 431) also cites *nam-ʔa kha hitt-a* (sun-ERG ANTIP burn-3.PST) ‘the sun was burning (yesterday)’. Note that in both examples the verb form is 3rd person intransitive, but the indexed nominal is ergative-marked.

According to Schikowski (2013: 111–19), the antipassive in Kiranti is limited to the South-Eastern and Far Eastern groups.

5.3.3 Passive

Passives are not prominent in Kiranti. The Hayu *go*-passive has been mentioned (Chapter 34 §5.8); Belhare has a similar construction (Chapter 35). There is also apparently a Hayu *beg*-passive, which has an exact Nepali parallel (*kuti māgāyo* (beat begged) ‘got a beating’).

- (30) Hayu *beg*-passive (PALIs7)
- “mima uŋ pat-ji tɔp-ji dip-ji siytɔŋ le t^hoy-tɔ,
thus 2OBL fight-AGT beat-AGT crush-AGT man also bring-IMPER.3SGO
dip bin-ji¹⁴ le t^hoy-tɔ!” paha tahale pəkrai pa p̩i:-kɔ
crush beg-AGT also bring-IMPER.3SGO comp officer arrest do CAUS-3SG>3SG.

“Produce both the man who fought, who beat, who crushed [the other], and the one who was crushed!” So saying, he [the headman] had the officer make the arrests.’

The Limbu invariable passive auxiliary *tet* (van Driem 1987: 215–18) is used mainly with the modal ‘to be able’ or with perception verbs like ‘to see’ and a complement (‘be perceived as’). Note that the A of the passivized clause can be retained.

- (31) Limbu: passive (SOGHAs49)

'set-ma	suk	<i>tet</i>	<i>jeba-'l-le'</i>	<i>lɔʔr-e</i>
kill-INF	be.able	PASS	shaman-DEF-ERG	say-PST

“It can be killed, by the shaman,” he said.’

Allen (1975: 74) mentions a passive construction in Thulung, e.g. *qem-la seosimu* ‘to get a beating’, based on *dəp* ‘strike’, a (nasalizing) suffix *la*, an otherwise unattested verb *seo(m)* and the middle/reflexive vector *si(t)*; his informants translated it by the Nepali *beg-passive*.

5.4 Causative

Besides the unproductive derivational causatives in some languages (§4.2), all Kiranti languages have productive means of causative formation. Many have chosen to grammaticalize the causative or the applicative of a verb ‘do’ as a general causative auxiliary. Thus Bahing *pa-ŋa* ‘I’ll do it’, *gu nu pad-u* (I be.good CAUSE-1>3) ‘I’ll make it well’. (See also the Hayu causative |piŋ| (< ‘send’) (§5.3.3) and Bantawa |mett| (§5.5)).

5.5 Deponent/unaccusative

Formally transitive ‘unaccusative’ (Perlmutter 1978) or ‘deponent’ verbs agree with their sole argument as a transitive object (Michailovsky 1997). An example is ‘to boil’:

- (32) Limbu (Tamar Khola) unaccusative verb ‘boil’ (Michailovsky 1997)

<i>cakwa</i>	<i>lokt-u</i>
water	boil-3O

‘the water boiled’

The cognate verbs in Thulung (*loddiu* ‘it boiled’, Allen 1975: 42), Bantawa (*lokt-u*), and Yakkha (Schackow 2015: 335–6) have the same property. An ergative agent cannot appear with these verbs unless with a causative auxiliary, thus Bantawa: *cakwa lok-mett-uŋ* (water boil-CAUS-1SG>3SG) ‘I boil the water.’

With other verbs of this type, e.g. Limbu ‘to be fat’, ‘to get something in one’s eye’, ‘to be soiled’, a (non-agentive) personal argument can appear, indexed by object-agreement morphology on the verb. An instrumental (sometimes labelled ‘impersonal agent’) may also be present (33), but there is no possibility of an ergative personal agent.¹⁵

- (33) (Limbu) deponent verb with object-agreement and instrumental (Michailovsky 1997)

<i>khamm-el-le</i>	<i>ne:kt-usi</i>
earth-DF-INST	soil-3>3NS

‘they got soiled with earth’

It is significant that verbs of this type in Limbu, Thulung and Bantawa all have lexical roots with the postfinal *t* (§4.2); in Hayu (this volume) they have applicative morphology.

6 OTHER CASE FRAMES

6.1 Experiencer constructions

Emotional or physical experience is coded in various ways in Kiranti. Bickel has reviewed the ‘possessive of experience’ construction in Belhare (1997b) and experiencer constructions in general in the wider Himalayan context (2004); see also Matisoff (1986) for Southeast Asia. We have seen ‘experiencer as goal’ examples among unaccusative verbs (33).

In possessive of experience constructions, the experiencer is coded as the possessor of the supposed site (Matisoff’s ‘arena’) of a physical or psychological feeling, or of the feeling itself; this site or feeling can become the subject of an intransitive verb, e.g. (Limbu) *a niywa t’j-e* (1OBL-mind arrive-PST) ‘I am content/satisfied’; *a luywa tuk* (1OBL-liver ache) ‘I miss it’; *a-sik la:k* (1OBL ?hunger sting) ‘I’m hungry’. In Limbu, in the basic, intransitive construction, the stimulus can appear as an absolutive, and the person of the experiencer is not indexed on the verb (34); in a causative construction (‘make-happy’), agreement is with the site of the experience, not with the experiencer (35).

- (34) Limbu possessive of experience with absolute stimulus (SOGHAs157)

ku-tɔkkatt-enn-ay-mu sarik ku-lujma tuk ness-e
3OBL-stick-DEF-CONJ-EMPH very 3OBL-liver ache continue-PST
‘He still sorely missed his walking-stick, as well.’

- (35) (SOGHAs140)

ke-niywa tas-umbe phe-ay ku-yg’wa-re mettu
2OBL-mind cause.to.arrive-1PLEX>3 COMP-CONJ 3OBL-mat.uncle-ERG say-23SG>3SG
“We will satisfy you,” his uncle said.’

Note that in (35) the uncle’s proposed role is not that of the stimulus that will satisfy his nephew (or his nephew’s mind), but that of the agent who will produce the stimulus (etc.). The stimulus is not mentioned or indexed.

In a few examples in Limbu, experiencer-agreement is found on the verb rather than on the site NP. Compare (36) with the more common construction *a-sira thay* (my-?liking comes.up) ‘I am pleased’.

- (36) Limbu (DANCEs30)

khombhe khan-ha'-nu wa'-ma sira a-dhay pek ləchə ni
thus DIST-PL-COM be-INF liking 1PLIN-come.up go.3NPST EVID y’know
‘So one comes to enjoy being with them.’

In the parallel Bantawa expression, the experiencer is both the site-possessor and the (unexpressed) core argument (S) of the verb: *ij-niija no-ja* (1OBL-mind be.good-1SG) ‘I am pleased.’ (Doornenbal 2009: 219). As in (36), the site-possessor (experiencer) displaces the site as core-argument of the verb. The experiencer can simply be thought of as the subject of the complex expression, PRON-*niija no-ma* ‘PRON be pleased’.

Bantawa can also treat the experiencer as agent and the stimulus (here human) as object: *ijka-?a som tuk-na* (I-ERG lung hurt.for-1SG>2SG) ‘I feel for you’, alternatively *ij-som tuk-na* (1OBL-lung hurt.for-1SG>2SG) (same meaning). A similar construction, with the experiencer as agent (indexed on the verb), is cited in Limbu (van Driem 1987 s.v. *lujma*, cited by Bickel 1997b: 149): *a-ndzum-in sa:rik a-lujma hi:pt-uŋ* (1OBL-friend much 1OBL-liver yearn-1SG>3SG) ‘I miss my friend very much’. Bickel remarks that agreement with the experiencer is pervasive in Belhare, but that in Limbu, agreement is more often with the site, e.g. the 3rd person object of *tasumbe* (35).

Schackow (2015: 274–88) has a thorough discussion of ‘possessive of experience’ constructions in Yakkha and lists some 50 such noun-verb predicates in five classes defined by frames of argument realization. As an example, we may consider what she calls an ‘unexpected pattern’, ‘which leads, oddly enough, to a literal translation “my disgust brings up bee larvae”’ for (37). Her realization frame for the four expressions of this class (omitting the possessive+site element that is common to all ‘possessive of experience’ expressions) is: {P[stim]-nom V-a[3].p[P]}.

- (37) Yakkha

<i>thaysu=ga</i>	<i>u-cya=ci</i>	<i>a-chippa</i>	<i>ket-wa-ci=ha</i>
bee=GEN	3SG.OBL-child=NSG	1SG.OBL-disgust	bring.up-NPST-3>3NSGP=NMLZ.NSG
'I am disgusted by the bee larvae.'			

Schackow notes: (1) that psych-nouns with this class of verb do not take ergative marking; (2) that there is an alternative formulation in which the verb is *ket-way-ci=ha*, a 1SG>3PL form, showing agent agreement with the experiencer. We can retain that *a-chippa ke?ma* is an complex NP+PREDICATE expression ‘I loathe —’, and that in the alternative formulation the predicate explicitly indexes the experiencer in place of the psych-NP, just as in the Limbu and Bantawa examples earlier.

6.2 Reflexive

Reflexive clauses have a single absolute core argument occupying the functions of both A and O or A and PO (e.g. beneficiary). In the latter case, the Second Object (theme), also absolute, may also be present.

- (38) Khaling reflexive (with second object) (Jacques *et al.* 2015 s.v. *kārnē (kur)* ‘carry’)¹⁶
- | | | |
|----------------------------|---------------|------------------|
| <i>?uy</i> | <i>dzakla</i> | <i>kār-si-ŋa</i> |
| 1[ABS] | food | carry-REFL-1SG |
| 'I carry food for myself.' | | |

In some languages (e.g. Thulung, Wambule, Bantawa), the reflexive is one reading of a general ‘middle’ form.

Reflexive verbs are formed either by a stem-augment (Bahing), by special agreement suffixes or by a vector V2. Reflexive suffixes are found in Limbu, Bantawa, Chamling, Chhintang and the western languages Hayu (Chapter 34), Khaling (Table 33.1) and Wambule. Often there is no single reflexive marker found in the same position in all forms. Reflexive vectors are used in the Yakkha group (e.g. in Belhare, Chapter 35).¹⁷ The Yamphu reflexive vector is |cimmu-s ~ cimmu| (39); Yakkha has a reflexive vector *ca* (< ‘eat’) and a ‘middle’ vector *si?*, clearly related to reflexive suffixes in other languages, used in cases of ‘low intentionality and volitionality’.

- (39) Yamphu reflexive (Rutgers 1998: 169)

<i>maik</i>	<i>toŋ-jimmu-l-lo</i>
black	make-REFL-NPST-REPORT
'She painted herself black.'	

- (40) Sunwar reflexive (Borchers 2008: 140)

<i>go</i>	<i>gup-sa-ŋ</i>
1SG	wear-REFL.PST-1SG
'I put it [a hat] on.'	

6.3 Benefactive

Benefactive constructions are realized in different ways across Kiranti. In Hayu (Chapter 34) the recipient may be marked by *h:si* ‘for’ (and indexed or not on the verb), but more often, in Hayu as elsewhere, the beneficiary is unmarked, but is indexed on the verb as (primary) object.

In Hayu, the ditransitive construction can be used with most verbs. An applicative form of the verb is used if available (Chapter 34).

In Bahing, the verb ‘give’ is used with a converb.

- (41) Bahing benefactive with converb < NINAMs143 >

<i>nani</i>	<i>gyp-ti</i>	<i>gi-ji!</i>
child	pick-CONV	give-2SG>1SG.IMP
‘Child, pick it [a tick] off for me!’		

In both Thulung and Dumi (van Driem 1993: 205–6), verb compounds with benefactive V2 are found.

- (42) Thulung benefactive V2 *sat* (Lahaussois 2002: 212–14)

<i>go</i>	<i>oram</i>	<i>nem</i>	<i>a-lwak-lai</i>	<i>di-sat-pu</i>
1SG	this	house	1OBL-y.brother-DAT	leave-BEN-1SG>3SG
‘I leave this house to my brother.’				

This is also the usual construction in the eastern languages. Yamphu is reported to have three ‘dative auxiliaries’, a general ‘DAT’, ‘help’ and ‘bring for’ (43).

- (43) Yamphu benefactive (Rutgers 1998: 179)

<i>na:ni</i> ,	<i>joya</i>	<i>jay-ghik-khay-m-e!</i>
child	water	carry-bring.for-CONATIVE-1O-IMPER
‘Child, see if you can’t bring me some water!’		

The 1st person beneficiary (‘me’, here not represented by a pronoun) is indicated as primary object by the agreement morphology of the compound verb; the theme (‘water’) is the second object. The benefactive vector ‘bring.for’ is itself an applicative derivative *khi?ma* related to *khi:ma* ‘carry’. The example contains a second vector *khay* ‘conative’, related to the independent verb ‘see’. Taken separately, the three verbs have quite different argument structures.

In Yakkha, benefactive compounds are formed with the vector *pi?* (< ‘give’). Schackow (2015: 371) reports that in this construction V1 stems of certain shapes are regularly augmented by *t* (if they do not already have postfinal *t*), which she considers to reflect the applicative suffix, hitherto believed to have become unproductive in Kiranti. Stems of the shape CV_k(*s*) become CV_k*t*, while CV_p(*s*) remains unaffected.

- (44) Yakkha benefactive (Schackow 2015: 373)

<i>eko</i>	<i>japmi=ga</i>	<i>o-ken</i>	<i>en-d-hak-t-u-bi-na</i>
one	person=GEN	3OBL-tooth	uproot-BEN-V2.SEND-BEN-3O-V2.GIVE[3O,PST]=NMLZ.SG
‘He pulled out someone’s tooth.’			

Limbu is exceptional in that the two verbs may diverge in object-agreement (do-it benefit-you), together reflecting the ditransitive argument structure of the compound.¹⁸

- (45) Limbu < SOGHAs115 >

<i>nu-rik</i>	<i>cog-uy</i>	<i>pi-niy</i>
good-MANNER	do-1SG>3SG	give-1SG>2PL
‘I’ll do a good job for you ^{pl.} ’		

7 NOMINALIZATION

Kiranti languages make frequent use of ‘nominalization’ of verbs, locative or temporal adverbs and clauses, creating nominals or adnominal modifiers which, in the absence of a head, can function as nominals. Clausal nominalizations with finite or non-finite predicates can serve as relative clauses, sentence complements or subordinate clauses. Examples can be found in the articles on Hayu, Chamling and Belhare in this volume. Hayu (this volume) has a particularly rich set of deverbal derivations nominalizing participant roles like agent/actor, patient, manner, place. Watters (2008) discusses nominalization in a dozen Kiranti languages, Kham and a few other Nepal languages.

Free-standing nominalization of complete sentences, which Matisoff (1972) described in Lahu, is common in Kiranti and has been viewed as related to information structure, e.g. by Ebert (1994) and Bickel (1999). Rutgers (1998) shows that one use of the Yamphu ‘factive’ nominalizer in narrative is to focus on a whole situation as relevant background to a main sequence of events (cf. Lambrecht 1994 on ‘sentence focus’).

- (46) Yamphu (Rutgers 1998: 246)

ikko toyara cess-uj. ikko toyara jay-ma-nuj khad-iŋ. ikko ja:ket wa:wa-æ pi:-?a-braʔa, akko caks-uj-æ.
one stick cut-1sg>3SG one stick carry-INF-SOC go-1SGSO one jacket
e.brother-ERG give-PURP-PP that put.on-1sg>3SG-FACT
'I cut a stick [finite]. I went with a stick in my hand [finite]. I had put on a jacket, the one elder brother had given me [nominalized sentence].'

For a Belhare example in conversation see Chapter 35, ex. (10).

Van Driem (1987, 1993) considers that in Dumi and Limbu nominalizing suffixes on independent sentence verbs are in fact the mark of imperfective aspect, unmarked verbs being perfective. (He also discusses the problem of distinguishing nominalization from aspect-marking.) This analysis has not been widely adopted (see Ebert 1994: 35; Doornenbal 2009: 205).

Ebert notes that questions in Athpare are nominalized, and this appears to be the case in other languages of the Yakkha group, and to be common in Bantawa (Doornenbal 2009: 195) and Panchthare Limbu (47).

- (47) Limbu (Panchthare): nominalized question (Weidert and Subba 1985: 315)¹⁹

ta:ndik ke-be:k-p'-i?
tomorrow 2-go-NML-Q
'Will you go tomorrow?'

(A reading with argument focus, ‘Is it tomorrow that...?’ is probably also possible.) Tamar Khola Limbu prefers a variant without nominalization (*ta:ndik ke-be:kk-i?*).

Nominalization of total questions suggests a higher predicate ‘Is it that...?’ with the whole proposition as complement. This depersonalizes the question, making it less intrusive (see Watters 2008: 36–7 on nominalized questions and imperatives).

7.1 Relatives

Relative clauses are generally marked by a nominalizer, which in fact most often marks them as adnominal modifiers. The nominalizer is suffixed to the verb of the relative, which may be a finite or a gerundive form. Examples of relative clauses with non-finite active or passive participles have been presented in §§3.5, 4.9.3, 4.9.4.

In Hayu (this volume) and Bahing all relatives have non-finite verb forms: an actor-nominal/active participle (Bahing suffix *ba* ‘ACT’) (16), or a general passive participle (Bahing *na* ‘PPT’).

- (48) Bahing relative <BEGINs139>

<i>mem</i>	<i>sa:-na</i>	<i>tselpo-kha-di</i>	<i>dzhɔk-tase</i>
that	kill-PPT	kinsman-place-in	arrive-PST.3DU
‘They ^{du} arrived at the place of the kinsman who had been killed.’ (cf. <i>sa:ta</i> ‘he killed him’)			

Wambule similarly has active and passive participle markers. Note that the passive participle in the example refers to the location, not the object.

- (49) Wambule (Opgenort 2004: 370)

<i>nana-m</i>	<i>si-co</i>	<i>mujo</i>	<i>le-bumco</i>	<i>lam-no</i>
before-NMLZ	die-ACT	person	take-PPT	road-LOC
‘on the road on which the dead person was carried’				

Thulung relatives can be formed similarly (markers *pa* ‘ACT’, *ma* ‘PPT’). But it also has relative clauses with finite verbs and the general nominalizer *m/mi* (Allen 1975; Lahaussois 2004).

A widespread pattern in Dumi, Khaling and the eastern languages is a non-finite active participle, used when the head has the function of S or A in the relative, and a general nominalizer with a finite verb when the head is an O or a non-core argument. Compare the use of the Limbu non-finite active participle (prefix *ke*) in ((50), (51)) with that of the unprefixed finite form (52); both have the nominalizing suffix *pa*.

- (50) Limbu <SOGHAs206>

<i>ke-si-ba</i>	<i>ku-lamsa</i>
ACT-die-NMLZ	3-nephew
‘his dead nephew’	

When the relative itself functions as head, it can carry the Limbu definite article (51).

- (51) Limbu <SOGHAs126>

<i>ke-si-b'-en</i>	<i>kɔ</i>	<i>ha:p</i>	<i>he:kt-ε</i>
ACT-die-NOM-DEF	TOP	weep	begin-23SG.PST
‘The dead man began to weep.’			

The nominalized finite verb is used to form a non-actor relative.

- (52) Limbu <ROADs42>

<i>anige</i>	<i>kam</i>	<i>cog-umbe-ba</i>	<i>pəisa-n</i>	<i>a-ber-ε</i>
1PLEX	work	do-1PLEX>3-NMLZ	money-DEF	2>1(NSG)-give-IMP
‘Give us the money we worked for!’				

In Yakkha, the active participle is non-finite, marked by the suffix *khuba*, while the verb of non-S/A relatives is a finite form, followed by one of the general nominal markers *na* ‘SG’ or *ha* ‘PL’. (These resemble the Limbu definite article and plural marker, but the Limbu morphemes can only mark NP heads.)

- (53) Yakkha (Schackow 2015: 404)

<i>paip</i>	<i>pek-khuba</i>	<i>babu</i>
pipe	break-ACT	boy
‘the boy who breaks/broke pipe’		

- (54) Yakkha (Schackow 2015: 408)
- | | | |
|------------------------------------|------------------|---------------|
| <i>beula-ha</i> | <i>khut-u-ha</i> | <i>tephen</i> |
| groom-ERG | bring-3O-NMLZ.PL | clothes |
| ‘the clothes brought by the groom’ | | |

In Bantawa, both active and passive participles are reported. The active participle is a non-finite prefixed form, as in Limbu, but with the suffix *ba* it usually indicates a profession or other lexicalized meaning. The passive participle is also a non-finite form, composed of the infinitive with the added suffix *yin* ‘word, thing’ and specialized to indicate a patient. The most versatile adnominal modifier is a genitive-marked finite form; the two participles seem more like actor and patient nominals. This is close to the situation in Thulung.

8 SUBORDINATE CLAUSES

In a review of Kiranti converbs (i.e. non-finite, non adnominal forms), Ebert (2008) found that the majority of subordinate clauses used finite verbs, except in Hayu (Chapter 34 §6.9).

8.1 The conjunctive subordinate

The equivalent of the ubiquitous Nepali conjunctive converb in *-era* is non-finite only in Hayu.

Clauses which are parallel in frequency and distribution to the Nepali conjunctive clause are identified in several languages (Ebert 2003a). In Hayu, this is a reduplicated converb, i.e. non-finite as in Nepali (Chapter 34 §6.9). In Thulung there are several candidates: Lahaussois (2002) distinguishes ‘sequencers’ (*ma* ‘anterior’,²⁰ *lo* ‘simultaneous’), which govern finite verbs, from converbs (*saka* ‘anterior’, *to* ‘simultaneous’), which mark stem gerunds; both link their clause to a matrix clause. Example (55) is a Thulung version of the well-known ‘— having said, he said’ conjunctive complementizer (cf. Nepali ‘—’*bhənera bhənyo*).

- (55) Thulung anterior sequencer used as a complementizer (Allen 1975: 79)
- | | |
|--------------------------|--------------------------------|
| ‘—’ <i>rɔk-ta-ma</i> | <i>rɔk-sod-dy</i> |
| — speak-3SG.PST-ANTERIOR | — speak-DEFINITIVE-3SG>3SG.PST |
| ‘He said “—”.’ | |

8.2 Converbs

Negative converbs (‘without —ing’) are non-finite forms in all languages; simultaneous converbs (‘while —ing’) are non-finite in all but Limbu (e.g. Yamphu ‘carrying a stick’ (36)).

8.3 Adverbial subordinates

Subordinated clauses are generally nominalized and marked as non-core arguments of an embedding clause. In Limbu, the most common subordinating suffix is (*e*)*lre*, a combination of the definite article, *en*, used as a nominalizer, and the ergative/instrumental/adverbial marker *re*. In Hayu, the verb of the subordinate is a non-finite root gerund; in the other languages it is finite.

- (56) Limbu <DANCEs48>

ja² ra:kt-i peg-ige-’lle anige nu-ba nis-umsimbe menchia-ha?
 paddy dance-1PL go-1PLEX-SUB 1PLEX good-NOM see-1PLEX>3PL woman-PL
 ‘As we went on dancing, we saw them as beautiful, the women.’

9 INFORMATION STRUCTURE AND FOCUS

All of the languages have sets of frequently used topic and focus particles ('INTENSIVE', 'only', 'CONTRARY TO EXPECTATION') and discourse particles ('REPORTED', 'EVIDENTIAL/INFERENTIAL', 'REASSURING', 'ANIMATING'). Topical elements may appear with lowered intonation after the verb in an anti-topic position ((25), (31)).

WEB RESOURCES

DOBES portal: dobes.mpi.nl (accessed 15 August 2016).

ELAR: elar.soas.ac.uk (accessed 1 January 2016).

Pangloss: lacito.vjf.cnrs.fr/pangloss/index_en.ht (accessed 1 January 2016).

NOTES

- 1 Shafer (1966: 142–57) placed Hayu apart from his 'East Himalayish' subgroup (otherwise equivalent to Kiranti as defined here), considering it to be closer to Chepang and Magar. I have included it in the Kiranti group.
- 2 The label 'Kiranti' is also claimed by some other groups in Nepal and India.
- 3 *Yakthung/jakthuy*: Language transcriptions (in italics in this article), have been regularized and converted to a notation based on the International Phonetic Alphabet (IPA) in a few important respects. Inconvenience to readers familiar with the area is regretted. In particular, *y*, used (as in traditional Indian romanizations) for *jod* in most (but not all) of the sources, has been changed to IPA *j*. Where 'j' appeared in the original sources for voiced palatal [j], it has been changed to *y*. (Original 'dz' has been preserved.) The remaining instances of *y* indicate a front rounded vowel. Non-italicized language and place names are in the modern roman transcription most current in Nepal, in which vowels have pre-Great-Vowel-Shift values ('as in Italian'), the phonemes represented by Devanagari *a* and *ā* are both written 'a', 'y' represents *jod* (IPA [j]), and 'ch', 'chh', 'j' and 'jh' represent lamoinal-alveolar affricates.
- 4 Neither Proto-Kiranti (PK) nor Proto-Sino-Tibetan (PST) can be regarded as worked-out reconstructions.
- 5 Bickel *et al.* (2007b: 56) state that Kiranti languages disallow vowel-initial phonological words: 'If there is no underlying consonant, a glottal stop is added'. It is not clear to me to whether such an element is a phonological segment. Cf. also the 'consonant position' posited by Doornenbal (2009: 28–31) in Bantawa.
- 6 In languages without palatals, the affricates *ts*, *dz* are often transcribed *c*, *j*.
- 7 Limbu pronouns do not take the ergative/instrumental/adverbial/genitive marker *re~le*.
- 8 Thanks to Martine Mazaudon for this section.
- 9 Hayu (this volume) has a functionally equivalent series of applicative suffixes in *-t-*.
- 10 The 'inverse' part of a double-agreement paradigm corresponds to situations where the A is outranked by the O in the person hierarchy 1>2>3 (for example); the other half of the paradigm is 'direct'. Ebert (1991, 1994: 26–8, Chapter 36 this volume) makes the case for attributing 'direct' or 'inverse' semantics to a number of agreement markers in Kiranti languages; her interpretation of Bantawa is contested by Doornenbal (2009: 150n).

- 11 The non-past stem appears before a vowel in the 1SG.NPST forms *jepp-e* ‘I stand’, *khemm-e* ‘I hear’; van Driem (1987: 96) transcribes the suffix *e* (and many other suffixes or postpositions) with an initial glottal stop (ʔ*e*) (e.g. in (19)) to account for the regular phonological gemination of pre-vocalic C_f (see §2.3).
- 12 The stems are identified as ‘S1’ (non-past) and ‘S2’ (past) in the glosses of Limbu texts on the Pangloss website, but not in the present chapter. In Limbu, the labels ‘past’ and ‘non-past’ are roughly justified where this opposition is not neutralized.
- 13 The exceptions are Chamling and Bantawa, in which syntagms marked by *pa/ma* cannot function as adnominal modifiers, only as nominal heads. In the other languages, they can function as nominal heads, like all adnominal modifiers.
- 14 This form is unattested elsewhere: the active participle of ‘beg’ is *bi-ji*.
- 15 Van Driem (1987: 275–6) and Kainla (2002) list several of these Limbu verbs as having either intransitive or transitive forms. This may reflect a dialect difference with Tamar Khola Limbu (Michailovsky 2002).
- 16 This Khaling reflexive is also related to an applicative *kārne* (*kurt*) ‘bring for someone’.
- 17 A reflexive suffix or stem-extension followed by intransitive affixes could resemble a vector. I have taken regularity of distribution and placement of the reflexive marker as a criterion for vector-hood in this case.
- 18 As suggested earlier, I have held off limiting the definition of compound verb *a priori*.
- 19 These authors simply gloss *pi* ‘interrogative particle’. Another apparently nominalizer-based question particle, *pe*, is glossed as a politeness marker.
- 20 Thulung *ma*, like Limbu *ay*, also functions as a postclitic conjunction ‘too, also’. Note that the Nepali conjunctive suffix *era* combines a perfect stem-marker *e* with *ra* ‘and’.

REFERENCES

- Allen, Nicholas J. (1972) The vertical dimension in Thulung classification. *Journal of the Anthropological Society of Oxford* 3:81–94.
- (1975) *Sketch of Thulung Grammar*. Ithaca: China–Japan program, Cornell University.
- (2012) *Miyapma: Traditional Narratives of the Thulung Rai*. Kathmandu: Vajra.
- Bickel, Balthasar (1996) *Aspect, Mood, and Time in Belhare. Studies in the Semantics–Pragmatics Interface of a Himalayan Language*. Zürich: ASAS.
- (1997a) *Dictionary of the Belhare Language*. Electronic database at the Sino-Tibetan Etymological Dictionary and Thesaurus Project, University of California at Berkeley.
- (1997b) The possessive of experience in Belhare. In: David Bradley (ed.) *Tibeto-Burman Languages of the Himalayas*, Canberra: Pacific Linguistics, Australia National University, 135–155.
- (1997c) Spatial operations in deixis, cognition, and culture: where to orient oneself in Belhare. In: Jan Nuyts and Eric Pederson (eds) *Language and Conceptualization*, Cambridge: Cambridge University Press, 46–83.
- (1999) Nominalization and focus constructions in some Kiranti languages. In: Yogendra P. Yadava and Warren W. Glover (eds) *Topics in Nepalese Linguistics*, Kathmandu: Royal Nepal Academy, 271–296.
- (2004) The syntax of experiencers in the Himalayas. In: Peri Bhaskararao and Karumuri Venkata Subbarao (eds) *Non-Nominative Subjects*, Amsterdam: Benjamins, 77–112.
- Bickel, Balthasar and Martin, Gaenszle (eds) (1999) *Himalayan Space: Cultural Horizons and Practices*. Zürich: Völkerkundemuseum.
- Bickel, Balthasar, Martin, Gaenszle, Arjun Rai, Prem Dhoj Rai, Shree Kumar Rai, Vishnu S. Rai, and Narayan P. Sharma (2007a) Two ways of suspending object agreement in

- Puma: between incorporation, antipassivization, and optional agreement. *Himalayan Linguistics* 7:1–19.
- Bickel, Balthasar, Goma Banjade, Martin Gaenzle, Elena Lieven, Netra Paudyal, Ichchha Rai, Manoj Rai, Novel Kishor Rai and Sabine Stoll (2007b) Free prefix ordering in Chhintang. *Language* 83:43–73.
- Bieri, Dora and Marlene Schulze (1971a) *Sunwar Phonemic Summary* (revised version). Mimeograph. Kirtipur, Nepal: Summer Institute of Linguistics, Tribhuvan University.
- (1971b) *A Guide to Sunwar Tone*. Kirtipur, Nepal: SIL. Mimeograph.
- Borchers, Dörte (2008) *A Grammar of Sunwar*. Leiden and Boston: Brill.
- Doornenbal, Marius (2009) *A Grammar of Bantawa*. Utrecht: LOT.
- Dryer, Matthew S. (1986) Primary objects, secondary objects, and anti-dative. *Language* 62.4:808–845.
- Ebert, Karen H. (1991) Inverse and pseudo-inverse prefixes in Kiranti languages: evidence from Belhare, Athpare, and Dungmali. *LTBA* 14.1:73–92.
- (1994) *The Structure of Kiranti Languages*. Zürich: ASAS.
- (1997a) *A Grammar of Athpare*. Munich: Lincom Europa.
- (1997b) *Camling (Chamling)*. Munich: Lincom.
- (1999) The UP-DOWN dimension in Rai grammar and mythology. In: Balthasar Bickel and Martin Gaenzle (eds) *Himalayan Space: Cultural Horizons and Practices*, Zürich: Völkerkundemuseum, 107–134.
- (2000) *Camling Texts and Glossary*. Munich: Lincom.
- (2003a) Equivalents of ‘conjunctive participles’ in Kiranti languages. In: Tej Ratna Kansakar and Mark Turin (eds) *Themes in Himalayan Languages and Linguistics*, Heidelberg: South Asia Institute and Kathmandu: Tribhuvan University, 27–48.
- (2003b) Kiranti languages: an overview. In: Graham Thurgood and Randy J. LaPolla (eds) *The Sino-Tibetan Languages*, London/New York: Routledge, 505–517.
- (2008) Converbs in Kiranti languages. In: Karen H. Ebert, Johanna Mattiesen and Rafael Suter (eds) *From Siberia to Ethiopia: Converbs in a Cross-linguistic Perspective*, ASAS no. 20, Zurich: Zurich University, 63–89.
- Ebert, Karen and Martin Gaenzle (2008) *Rai Mythology: Kiranti Oral Texts*. Harvard Oriental Series 69. Cambridge: Harvard University Press.
- Fillmore, Charles (1970) The grammar of hitting and breaking. In: Roderick A. Jacobs and Peter S. Rosenbaum (eds) *Readings in English Transformational Grammar*, Waltham, MA: Ginn, 120–133.
- Gaenzle, Martin (1991) *Verwandtschaft und Mythologie bei den Mewahang Rai in Ostnepal: Eine ethnographische Studie zum Problem der ‘ethnischen Identität’*. Stuttgart: Franz Steiner.
- (2002) *Ancestral Voices. Oral Ritual Texts and their Social Contexts Among the Mehawahang Rai of East Nepal*. Münster: LIT.
- Genetti, Caroll (1988) Notes on the structure of the Sunwari transitive verb. *LTBA* 11.2:62–92.
- Grierson, George A. (1908) *Linguistic Survey of India III, Tibeto-Burman Family, Part I* [prepared by Sten Konow]. Calcutta: Office of the Superintendent of Government Printing, India.
- Gvozdanovic, Jadranka (1985) *Language System and its Change: On Theory and Testability*. Trends in Linguistics, Studies and Monographs 30. Berlin, New York: Mouton de Gruyter.
- Hodgson, Brian Houghton (1857a) Grammatical analysis of the Váyu language. *JASB* 26.5:372–485 (errata in *JASB* 27.6; the errors are corrected in the reprint: Hodgson

- (1880) *Miscellaneous Essays Relating to Indian Subjects* (two volumes), London: Trübner, (1)216–319.
- (1857b) Comparative vocabulary of the several languages (dialects) of the celebrated people called Kirântis... *JASB* 26:333–371 (errata in *JASB* 27.6; the errors are corrected in the reprint: Hodgson (1880) *Miscellaneous Essays Relating to Indian Subjects* (two volumes), London: Trübner, (1)176–215).
- (1857–1858) Grammatical analysis of the Bâhing dialect of the Kirânti language. *JASB* 26.6:486–522 and 27.5:393–442 (errata in *JASB* 27.6; the errors are corrected in the reprint: Hodgson (1880) *Miscellaneous Essays Relating to Indian Subjects* (two volumes), London: Trübner, (1)320–392).
- (1858) On the Kiranti tribe of the central Himalaya. *JASB* 27.5:446–456.
- Huysmans, René (2007) The Sangpang verbal agreement system. In: Mark Turin and Bettina Zeisler (eds) *Himalayan Languages and Linguistics*, Leiden: Brill, 183–218.
- Jacques, Guillaume (forthcoming) Tonogenesis and tonal alternations in Khaling. See: www.academia.edu/10516416, accessed 1 January 2016.
- Jacques, Guillaume, Aimée Lahaussois, Boyd Michailovsky and Dhan Bahadur Rai (2012) An overview of Khaling verbal morphology. *Language and Linguistics* 13.6:1095–1170.
- (2013) Reflexive paradigms in Khaling Rai. *LTBA* 39.1:33–48.
- Jacques, Guillaume, Aimée Lahaussois, Dhan Bahadur Rai and Yadav Kumar (2015) *Khaling–Nepali–English Verb Dictionary* v. 1.0. Available at: hal-inalco.archives-ouvertes.fr/halshs-01242022v1, accessed 1 January 2016.
- Kainla, Bairagi (ed.) (2002) *Limbu–Nepali–English Dictionary*. Kathmandu: Nepal Academy.
- Lahaussois, Aimée (2002) *Aspects of the grammar of Thulung Rai*. Berkeley: University of California dissertation .
- (2004) Thulung Rai. *Himalayan Linguistics* 0.1:1–25.
- Lambrecht, Knut (1994) *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.
- Linguistic Survey of India* (n.d. [2009?]) *Sikkim Part-1*. Office of the Registrar General, Government of India. Prenominalized Himalayan Languages: Limbu–Rai: 204–327; Non-prenominalized Himalayan Languages: Sunwar: 614–674. (Web page of the Census of India: www.censusindia.gov.in/2011-documents/lst/ling_sikkim.html, accessed 1 January 2016).
- Matisoff, James A. (1972) *The Grammar of Lahu*. Berkeley: University of California.
- (1986) Hearts and minds in South-East Asian languages and English: an essay in the comparative lexical semantics of psycho-collocations. *Cahiers de Linguistique Asie Orientale* 15.1:5–57.
- Mazaudon, Martine (2010) Number-building in Tibeto-Burman languages. *North East Indian Linguistics* 2:117–148.
- Michailovsky, Boyd (1975a) Notes on the Kiranti verb (East Nepal): 1. The Bahing verb. 2. The origin of tone in Khaling. *LTBA* 2.2:183–218.
- (1975b) Some Tibeto-Burman sound changes. In: Cathy Cogen, Henry Thompson, Graham Thurgood, Kenneth Whistler and James Wright (eds) *Proceedings of the First Annual Meeting of the Berkley Linguistics Society*, 15–17 February 1975, Berkeley: Berkeley Linguistics Society, University of California, 322–332.
- (1985) Tibeto-Burman dental suffixes: evidence from Limbu (Nepal). In: Graham Thurgood, James A. Matisoff and David Bradley (eds) *Linguistics of the Sino-Tibetan*

- Area: The State of the Art*, Pacific Linguistics C87, Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University, 363–375. (Reprinted in Yogendra P. Yadava and Warren W. Glover (eds) (1999) *Topics in Nepalese Linguistics*. Kathmandu: Royal Nepal Academy, 478–494).
- (1986) Structure syllabique et variation combinatoire: voisement et gémination en limbu. *Cahiers de Linguistique Asie Orientale* 15.2:193–204.
- (1988a) *La Langue Hayu*. Paris: Éditions du CNRS.
- (1988b) Phonological typology of Nepal languages. *LTBA* 11.2:25–50.
- (1994) Manner vs. place of articulation in the Kiranti initial stops. In: Hajime Kitamura, Tatsuo Nishida and Yasuhiko Nagano (eds) *Current Issues in Sino-Tibetan linguistics*, Osaka: National Museum of Ethnology, 766–772, 781.
- (1997) Catégories verbales et intransitivité duale en limbu. *Studi italiani di linguistica teorica e applicata* 26.2:307–325.
- (2002) *Limbu–English Dictionary of the Mewa Khola Dialect with English–Limbu Index*. Kathmandu: Mandala Book Point.
- (2003) Suffix-runs and counters in Kiranti time-ordinals. In: David Bradley, Randy J. LaPolla, Boyd Michailovsky and Graham Thurgood (eds) (2003) *Language Variation: Papers on Variation and Change in the Sinosphere and in the Indosphere in Honour of James A. Matisoff*, Pacific Linguistics 555, Canberra: Australian National University, 237–251.
- (2007) Pronominally marked noun determiners in Limbu. In: Roland Bielmeier and Felix Haller (eds) *Linguistics of the Himalayas and Beyond*, Berlin: Mouton de Gruyter, 189–202.
- (2012) Internal reconstruction of the Dumi verb: lexical bases and stem formation. *LTBA* 35.2: 49–87.
- (2015) On Limbu directionals and locative expressions. In: Mark W. Post, Stephen Morey and Scott DeLancey (eds) *Language and Culture in Northeast India and Beyond: In Honour of Robbins Burling*, Asia-Pacific Linguistics 023, Canberra: ANU, 114–125.
- Opgenort, Jean Robert (2004) *A Grammar of Wambule*. Leiden: Brill.
- (2005) *A Grammar of Jero: With a Historical Comparative Study of the Kiranti Languages*. Leiden: Brill.
- (2011) A note on Tilung and its position within Kiranti. *Himalayan Linguistics* 10.1: 253–271.
- Perlmutter, David M. (1978) Impersonal passives and the unaccusative hypothesis. *Berkeley Linguistics Society* 4:157–189.
- Pokharel, Madhav P. (1991) Compound verbs in Nepali. *Contributions to Nepalese Studies* 18.2:149–173. (Reprinted in Yogendra P. Yadava and Warren W. Glover (eds) (1999) *Topics in Nepalese Linguistics*. Kathmandu: Royal Nepal Academy, 185–208).
- Rai, Novel Kishore (1985) *A descriptive study of Bantawa*. Pune: Deccan College dissertation.
- Rai, Novel Kishor, Manoj Rai, Netra Prasad Paudyal, Robert Schikowski, Balthasar Bickel, Sabine Stoll, Martin Gaenzle, Goma Banjade, Ichchha Purna Rai, Toya Nath Bhattacharya, Sebastian Sauppe, Rikhi Maya Rai, Janak Kumari Rai, Las Kumari Rai, Durga Bahadur Rai, Ganesh Rai, Dayaram Rai, Durga Kumari Rai, Anita Rai, Chandra Kumari Rai, Shanti Maya Rai, Ravindra Kumar Rai, Judith Pettigrew and Tyko Dirksmeyer (2011). *Chhintang–Nepali–English Dictionary and Grammar*. Kathmandu: Chhintang Language Research Program, Tribhuvan University and University of Leipzig [in Nepali; English glosses and index].

- Rai, Vishnu S., Srikumar Puma Rai, Narayan Prasad Gautam, Diana Schackow, Sabine Günther, Arjun Rai, Ganesh Puma Cintan, Martin Gaenszle, Judith Pettigrew and Balthasar Bickel (2009) *Puma–Nepali–English Dictionary and Grammar*. Lalitpur: Kirat Puma Rai Tupkhabangkhala [in Nepali; English glosses and index].
- Rutgers, Roland (1998) *Yamphu: Grammar, Texts, and Lexicon*. Leiden: Research School CNWS.
- Schackow, Diana (2015) *A Grammar of Yakkha (Studies in Diversity Linguistics 7)*. Berlin: Language Science Press.
- Schikowski, Robert (2013) *Object-conditioned differential marking in Chintang and Nepali*. Zurich: University of Zurich dissertation.
- Shafer, Robert (1966) *Introduction to Sino-Tibetan*. vol. I. Wiesbaden: Harrassowitz.
- Toba, Sueyoshi and Ingrid Toba (1975) *A Khaling–English English–Khaling Glossary*. Kirtipur, Nepal: Tribhuvan University.
- Tolsma, Gerard Jacobus (2006) *A Grammar of Kulung*. Tibetan Studies Library. Leiden: Brill.
- Tumbahang, Govinda Bahadur (2011) *A Study of Chhatthare Limbu*. Kirtipur: Tribhuvan University.
- van Driem, George (1987) *A Grammar of Limbu*. Berlin: Mouton de Gruyter.
- (1993) *A Grammar of Dumi*. Berlin: Mouton de Gruyter.
- Watters, David E. (2008) Nominalization in the Kiranti and central Himalayish languages of Nepal. *LTBA* 31.2:1–43.
- Weidert, Alfons and Bikram Subba (1985) *Concise Limbu Grammar and Dictionary*. Amsterdam: Lobster Publications.
- Winter, Werner (1991) Diversity in Rai languages: an inspection of verb stems in selected idioms. *Lingua Posnaniensis* 34: 141–156.
- (1997) A sound change in progress? In: Raymond Hickey and Stanislaw Puppel (eds) *Language History and Linguistic Modelling: A Festschrift for Jacek Fisiak on his 60th Birthday*, Berlin/New York: Mouton de Gruyter, 1113–1123.

CHAPTER THIRTY-FOUR

HAYU

Boyd Michailovsky

1 INTRODUCTION

The Hayu (Nepali *hāyu*, Hayu *wa:ju*) inhabit an area between 50 and 100 kilometres south-east of Kathmandu. The Nepal *National Population and Housing Census 2011* (Government of Nepal 2012) reports a population of 2,925 Hayus, and 2,920 Hayu speakers, but the language is reliably known to be spoken in only three villages: Mudhajor (Sukajor VDC) in the valley of the Sun Kosi in Ramechhap District, where the data presented here was gathered, and Adhamara and Manedihi (Dadigurane VDC) in the valley of the Marin Khola, in Sindhuli District. All speakers are bilingual in Nepali, the Indo-Aryan national language.

Hayu was first studied by Hodgson (1857) during his retirement in Darjeeling, and was rediscovered and studied in 1972 and 1973 by Michailovsky and Mazaudon (1973; Michailovsky 1974, 1975, 1988) in the village of Mudhajor. It has not been learned by children in Mudhajor since the late 1970s, but it is still being learned in the Marin Khola, 30km to the west across the Mahabharat Range. There is little contact between the two areas, but dialect differences are slight. The present sketch is based on the dialect of Mudhajor as spoken in the 1970s, and by older speakers since then.

Hayu is considered here to be a member of the Kiranti subgroup (see Kiranti languages, this volume).

2 PHONOLOGY

2.1 Syllable structure

The syllable canon is $(C_i)V(C_f)$, where C_i represents either an initial consonant C_i or a cluster.

2.2 Vowels

There are seven vowel timbres, *i*, *ɪ*, *ɛ* (front); *a* (IPA [a]); *ɔ*, *ʊ*, *u* (back). The vowels *i* and *ʊ* are not lax; especially in open syllables they may be pronounced as close [e] and [o].

The Nepali *ə* vowel occurs in loanwords, where it also may also be realized as *a* or *ɔ*. Thus Nepali *əbə* ‘now’ is adopted as *abɔ* or *abə*.

Vowel quantity (length) is distinctive in open first syllables of polysyllabic words, in particular on the stem syllable of verb forms. Closed syllable vowels are phonetically short.

Nasality occurs only on open, non-final syllables. Almost all nasal vowels are long and precede an unvoiced stop or *s* initial of the following syllable. A very few words have short nasal vowels before voiced initial stops.

2.3 Consonants

There are three series of initial stops: plain (unvoiced), aspirated and voiced. The syllable initial consonants (C_i) are as follows:

dorsal-velar: k, k^h, g, η
lamino-palatal: $c ([t\epsilon]), \chi ([d\zeta])$
apico-alveolar: ts, ts^h, dz
dental: t, t^h, d, n
bilabial: p, p^h, b, m
apico-alveolar tap: r
apico-dental laterals: l (voiced), hl (voiceless)
approximants: j (palatal), w (labiovelar)

fricatives: apico-alveolar: s ; (labio-)velar-palatal: $x ([x], [x^w]$ before a, ω ; $[ç]$ before front vowels)

aspirate: h .

The opposition between the apico-alveolar and lamino-palatal stop series is neutralized before front vowels, the realization being apico-alveolar. There is no aspirated lamino-palatal affricated stop.

In Mudhajor, j appears before all vowels (thus ji ‘blood’ is distinct from i ‘this’), and w appears only before the vowels a and ω . The Marin dialect has w in place of j before front vowels: wi (or vi) ‘blood’.

Initial clusters are relatively rare, except in phonæsthetic adverbs, where they are somewhat favoured. The following initial clusters occur: $kl, k^hl, gl; kr, k^hr, gr; pl, p^hl, bl$. All words with initial bilabial+ $/$ clusters have doublets with simple bilabial initials.

Retroflex and voiced aspirated initials may be noted in Nepali loan words.

The inventory of syllable finals (C_f) is $p, t, k, m, n, \eta, r, l, x, ?$; of these, x and (with marginal exceptions) $?$ do not occur as word-finals. The stop finals are pronounced unreleased in word-final position, with simultaneous glottal closure. Before a voiced initial stop, either within the word or in close juncture, they have voiced allophones, realized with glottalization or laryngealization [b], \ddot{d} , \underline{g} , as in *bepdzere* [be β bdzere] ‘bedbug’, *it bi:tɔ* [i β bi:tɔ] ‘allow him to say it’.

Consonant combinations of the form $C_i C_i$ occur in polysyllabic words. In these, and in some close junctures, *homorganic* combinations are almost totally absent. Homorganicity is defined with respect to the three place-categories bilabial, velar and *coronal* (lamino-palatal, apico-alveolar or dental). It is clear from the morphophonology (below) that where a homorganic $C_i C_i$ combination would be expected, an oral stop or nasal C_f has been replaced by $x, ?$ or a prosody (nasality, distinctive short quantity) with no place feature. It will be seen that these function as allophones of the final stops and nasals. (The resonant finals r, l do occur before coronal initials, but coronal nasal and stop finals are replaced before initial r, l .)

3 MORPHOPHONOLOGY

A lexical verb root is a monosyllable ($C_i V(C_f)$), where C_f represents one of the absolute finals $p, t, k, m, n, \eta, r, l$. There is no lexical opposition between roots with long vs short or nasal vs oral vowels. Open roots have long quantity in polysyllabic forms. This is important for rule 4 below.

Verbal morphology is suffixal. Wherever a combination between a stop or nasal root-final and a homorganic suffix initial would be expected, the root final is replaced by an alternant as follows:

1. In combinations of stop+sonorant, the alternant is ? : *bʊʔŋɔ* ‘I get up’, root *bok*; *sɪʔnɔ* ‘I kill you’, root *sit*; *pɔʔluy* ‘tying-place’, root *pst*. Note that although the finals *k*, *t* and *p* share the same alternant, there is no neutralization, as the place feature is recoverable from the following initial. Thus ? functions as an allophone of *k*, *t* or *p*, depending on the context. The alternants 2–4 below function similarly.
2. The C_f alternant in homorganic stop+unvoiced stop combinations is *x*: *poxkɔ* ‘raise it’, root *pok*; *sixta* ‘killed’.
3. In nasal+unvoiced stop (also *n+s*) combinations the alternant is (long) nasality: (*pɔ̃:tse* ‘he becomes’, root *pɔn*; *p̩i:kɔ* ‘send it’, root *piŋ*).
4. In combinations of nasal+nasal, *t+s* and *t+ts^h* the homorganic C_f is lost, leaving a short vowel (redundantly transcribed here): *p̩iŋɔ* ‘he sends me’ (*|piŋ+yɔ|*); *p̩atsʰem* ‘they (dual) fought’ (*|pat+tsʰe+m|*); *s̩isɔŋ* ‘kill me!’ (*|sit+sɔŋ|*) (Michailovsky 2000–: Bs51¹). Compare *ma si:sɔŋ* ‘didn’t you recognize me?’, root *si* (Gs40).

Finally, the opposition between velar and bilabial suffix initials is neutralized after bilabial stem finals. The archiphonemes are represented by bilabials: *txxpɔ* ‘strike it’, *|tɔp+kɔ|; dūmɔ* ‘I become’ (*|dum+yɔ|*). The resulting homorganic combinations are subject to the rules 1–4 described above.

The combinations produced by rules 1–4 above (*x* or long nasality+unvoiced stop, $\text{?}+\text{nasal}$ or resonant) occur frequently in the general vocabulary (e.g. *pixpi* ‘grandmother’, *naʔnum* ‘rain’), while homorganic combinations and bilabial+velar groups are marginal, limited to phonaesthetic words (*pempere* ‘disk shaped’). Thus it seems likely that the rules described have affected the general vocabulary. Although evidence from alternations is limited by the absence of suffixes with voiced stop initials, the fact that neither *x* nor long nasality is found before voiced stops in the general vocabulary suggests that the dissimilation rules did not operate in this context.

Junctures like those between verb and postposition, verb and modal, and noun and postposition are weaker than that between verb and suffix, and dissimilation rules apply only in part.

4 VERB MORPHOLOGY

Verbs are classified as belonging to one or more of the lexical categories ‘intransitive’ (vi), ‘transitive’ (vt, vtt or both, see below), ‘deponent’ (vd) or ‘stative’ (vs) according to their agreement morphology and core arguments. Intransitive verbs, and reflexive forms of transitives, agree with an *s* (‘subject’); transitive verbs may agree with both an *A* (loosely ‘agent’) and an *O* (‘object’). Deponent verbs have the morphology of transitives with invariable 3sg *A*; their sole argument is indexed as a transitive *O*. Stative verbs have no finite forms and no agreement. The *tense* distinction is roughly past/completed vs non-past.

Morphological root alternation occurs in (1) four open roots *C,a* (e.g. *ta* (*ta~tɔ*) vt, vtt ‘put’), and (2) verbs having both open and *t*-final roots (e.g. *bu(t)* (*bu~but*) vt, vtt ‘carry’). The choice of root is part of the paradigm of these verbs (see §4.2 and Michailovsky 1988: 99–104).

TABLE 34.1 HAYU VERB MORPHOLOGY (SUFFIXES)

<i>Transitive</i>									
A↓O: Is	Idi	Idε	Ipi	Ipe	2s	2d	2p	3s	(app)
1s					Ncu	Ntci	Ntci	Ntci	/tɔŋ/
1di					Ncu	Ntci	Ntci	Ntci	/tɔŋtsʰɛ/
KEY:									
NPST									
PST									
Ipi									
Ipe									
2s	cl	tsʰk̚l̚							
2d	cl	tsʰk̚l̚							
2p	cl	tsʰk̚l̚							
3s	cl	tsʰk̚l̚							
3d	cl	tsʰk̚l̚							
3p	cl	tsʰk̚l̚							
<i>Intransitive</i>									
s:	Is	Idi	Idε	Ipi	Ipe	2s	2d	2p	3s
NPST	cl	tsʰk̚l̚							
PST	cl	tsʰk̚l̚							
<i>Reflective</i>									
s:	Is	Idi	Idε	Ipi	Ipe	2s	2d	2p	3s
NPST	ñsou	natsʰik	natsʰik	ñsik̚e	ñsik̚e	natsʰik	natsʰik	natsʰik	natsʰik
PST	ñsou	natsʰij	natsʰij	ñsik̚e	ñsik̚e	natsʰe	natsʰe	natsʰe	natsʰe

The suffixes of the verbal morphology are presented in Table 34.1 (see key in table). In the transitive paradigm of the table, each row corresponds to a person and number of A, and each column to a person and number of O. Intransitive and reflexive paradigms are presented below the transitive section.

Certain suffixes in Table 34.1 are transcribed as morphophonemes or alternations:

1. The suffix element transcribed *N* ('PST'; also see next item) transforms a root-final stop into a nasal.
2. The suffix $\eta \sim N \sim sv\eta$ ('1s→3:NPST') has the realization η after vowel, *N* after stop, *svη* after sonorant.
3. The reflexive suffix element $\tilde{N}ts$ nasalizes open roots.
4. *mi~m*: the assertive suffix *mi* may have the form *m* after a vowel.

4.1 Structure of the paradigm

Although Hayu case-marking is ergative, the verbal paradigm does not follow this principle. Most agreement markers mark person (including ‘clusivity’) and number without regard to function. In the transitive paradigm, the superior argument (whether A or O) according to the hierarchy 1>2>3 is given priority (allowing that 2s and 3s markers may be zero); if this argument is singular, then the number of the second argument is indicated by addition of a *secondary number marker*: *ts^{he}* ‘DU’ (except in 2s↔3d forms), *ne* ‘2p’, *me* ‘3p’. 3→3 agreement is with the argument of the greatest number: the combinations are glossed ‘3s→3s’, ‘3→3(d)’, ‘3→3(p)’.

Hierarchical marking is replaced in parts of the transitive paradigm by two sets of function-marking suffixes: (1) *transitive k-initial* and *applicative t-initial suffixes* (*koy*, *kɔ*, *ki*; *tov*, *tɔ*, *ti*), which indicate 3rd person O in some forms, and (2) 1sg→2 suffixes in *nɔ*.

4.2 Applicative forms

Applicative t-suffixes in *tɔ*, *tov*, (*ti*) are found in Table 34.1 in the same cells as the transitive *k*-suffixes. Transitive verbs which use the *k*-suffixes are identified as *vt*; those which use the *t*-suffixes as *vtt*. For verbs which use both (*vt*, *vtt*), the *t*-suffixes are used in ditransitive clauses (§6.5). Deponent verbs use *t*-suffixes.

Transitive verbs which have alternations between open and *t*-final roots may have a full paradigm (not limited to forms with *t*-suffixes) of distinctively applicative forms based on the *t*-final root. Thus, from *bu(t)* *vt*, *vtt* ‘carry’: *bu:nɔm* ‘I’ll carry you [non-applicative]’ (Us11) vs *bu?nɔm* ‘I’ll carry it for you’.

4.3 Negative indicative

The negative of a finite, indicative verb is indicated by *ma* placed before the positive form (see exx. (8), (25), (29) later). Note that *ma* is intonationally independent and may even be followed by a focus particle: *ma na dzo:nɔm* (not INTS eat:1s→2s:ASS) ‘I will absolutely not eat you!’ (C9). The negative *mak^{hi}* ‘not yet, not’ is used with non-past forms only, with negative perfect sense; the corresponding affirmative is a past form (17).

4.4 Imperative, prohibitive, exhortative, optative

Imperatives are identical to 2nd person past indicative forms: *dza:kɔ* ‘eat it!’ (= ‘you ate it’); *sixtɔmɛ* ‘kill them!'; *tɛsvɔŋ* ([tet+son]) ‘let me go!’ (Bs24); *bojne* ‘get up (PL)!’.

Exception: the intransitive 2s imperative is simply the root, like the non-past: *bok* ‘get up!’ (= ‘you (SG) get up’) (Bs58).

Negative imperatives are identical to the non-past indicative preceded by the prohibitive marker *t^ha* ‘don’t!’: *t^ha bok* ‘don’t (SG) get up!?’; *t^ha bokne* ‘don’t (PL) get up!?’; *t^ha dzɔ* ‘don’t eat it!?’; *t^ha sitme* ‘don’t (SG) kill them (PL)!?’; *t^ha tetŋɔ* ‘don’t (SG) let me go!?’.

Optatives add the suffix *ju* to non-past forms: *dzɔ:-ju* ‘let him eat!?’; *gɔ:ts^hik-ju* ‘[long] may we (1di) live?’ (*gɔ:(t)*) (AAs134). Exhortatives are simply non-past 1st person forms (16).

4.5 Non-finite forms and verbal nouns

The infinitive takes the genitive postposition: *mitsi nt-mo* ‘what to say?’ (AAs11)

Among nominalizers, the present active participle (root+*ji*) and the past/passive participle (root+*ta*) are the most generally used in relatives (§8.1): *metji* ‘dying’, *tɔpjji* ‘striking, who strikes/struck’; *mexta* ‘dead, carrion’, *tɔpta* ‘struck’. They are negated by *maaj*, e.g. *maaj nɔxta da:bɔ* (NNEG be:PPT thing) ‘nonsense [“non-existent things”]’ (DDs13); *maaj gɔtji* (NNEG have:ACTP) ‘poor’ (DDs13).

Verbal nouns are formed by adding suffixes to the root: agent nominal *ji*; patient nominal *tay*; instrument nominal *cay*; locative nominal *luy*; time nominal *siy*; manner nominal *si*, *sina*; descriptive nominal (of a person) *tɔ*. Examples: *tɔpjji* ‘blacksmith’ (*tɔp* ‘strike’); *tɔpcay* ‘hammer’; *piptay* cigarette (*pip* ‘suck’) (ex. (7)); *imluy* ‘bed’ (*im* ‘sleep’); *ɔsiy* ‘time of meeting’ (*ot* ‘meet’); *imsiy* ‘bedtime’; *hɔsina* ‘voice, manner of speaking’ (*hot* ‘speak’); *imtɔ* ‘sleepy-head’ (6). All can also nominalize clauses (§8.1).

Verbal heads of subordinate clauses appear as non-finite gerundive forms: (1) the simple root, followed by a postposition; (2) the progressive participle root+*ni*+root (e.g. *tɔ?nitɔ* ‘(while) pursuing?’); (3) the conjunctive participle, formed from the reduplicated root+*ha* ‘INST’: e.g. *tɔtɔt.ha* ‘after/by pursuing’. In this form, CV roots are reduplicated as CV:CV (*dza:dzaha* ‘after eating’); C_iVC_f roots are fully or partially reduplicated: *tɔptɔp.ha* ~ *tɔtɔp.ha* ~ *tɔptɔha* ‘having struck, by striking’ ((6), (12), (22), (23), (25)).

4.6 Non-productive derivational morphology

Hayu preserves traces of the well-known Tibeto-Burman causative morphology in the form of some 35 pairs of related verb roots. This morphology is not productive. In most pairs, the (roughly) non-causative member has a voiced stop initial and the causative member a plain or aspirated unvoiced one. Examples: *gik* vtt ‘tie’ vs *kik* vt, vtt ‘tie on’; *dzɔl* vi ‘live’ vs *tsɔl* vt ‘save the life of’; *dɔk* vi ‘fall’ vs *tok* vt, vtt ‘drop’, ‘cause to fall’; *bok* vi ‘rise, get up’ vs *pok* vt, vtt ‘rouse, raise’; *dat* ‘run out’ vs *t^hat* vtt ‘use up’; *bek* vi ‘enter’ vs *p^hek* vtt ‘bring/take in’. A few pairs do not involve stops: *im* vi ‘sleep’ vs *him* vt, vtt ‘put to sleep’; *ot* vtt ‘meet’, ‘find’ vs *ho(t)* vt, vtt ‘look for’; *re* ‘be broken in pieces’ vs *xe* vtt ‘break’; *ram* vi ‘be afraid’ vs *x^wam* vtt ‘startle’.

5 NON-VERBAL MORPHOLOGY AND WORD CLASSES

5.1 Noun morphology and derivation

Nominal case is marked by postpositions. There is one nominal suffix, *k^hata* ‘plural/collective’ (20).

There are a number of derivational or taxonomic formants: *tɔ* ‘human’, ‘collective human’, *mi* ‘female human’, *wɔ* ‘male human’, *si* ‘fruit’. Examples: *t^ha:tɔ* ‘grandson’,

ka:tsə ‘friend’, *nɔnɔtsə* ‘pair/group of sisters’ (cf. *nɔnɔ* ‘younger sister’), *ta:wɔ* ‘son’, *ta:mi* ‘daughter’, *tʰa:mi* ‘granddaughter’, *kɔksi* ‘*Ficus semicordata*’.

5.2 Pronouns

The independent absolute pronouns are: first person: *gu* ~ *guu*; second person: *gɔn* (2s), *gɔnts'ε* (2d), *gɔne* (2p); third person (human) *kɔmti*. The first and third person pronouns, like nouns (29), can be followed by *nakpu* ‘two’ or *kʰata* ‘PL’ to show number. The demonstrative *mi* is often used as a third person pronoun.

Ergative forms are based on the absolute with the postposition *ha*: *ga* ‘1ERG’, *gɔna* ~ *gɔn-ha* ‘2S.ERG’.

The oblique/possessive pronouns are: 1s *ay*; 1de *aytsʰε*; 1di *uŋtsʰε*; 1pe *ã:ki*; 1pi *û:ki*; 2s *uŋ*; 2d *uŋtsʰε*; 2p *uni*; 3s *a*, 3d *atsʰε*, 3p *ami*. These appear before nouns ((1), (11), (21), etc.) and before postpositions other than *ha* ‘ERG’ ((8), (9)).

5.3 Demonstratives

The demonstratives are *i~ii* ‘this’, *mi~mii* ‘that’. Only the shorter variants are used before suffixes or postpositions: *mikʰata* ‘those, they’, *mi-ha* ‘he-ERG’, ‘that-INST’. Also: *itha* ‘this much’ ‘(towards) here’, *mitha* ‘that much’; *ithara* ‘to this extent (with an adjective)’, *ibe* ‘here’. A connecting element is found in some forms: *mi-khen~min-khen* ‘then’, ‘from there’; *ine* ‘here’ (<i+n+he>); *ima* ‘in this manner’ (|i+m[?] + ha|).

5.4 Indefinite-interrogative words

The indefinite-interrogative words are *su* ‘who, someone’, *mitsi* ‘what, something’, and a series of words based on the indefinite-interrogative morpheme *ha* (only interrogative glosses are cited): *hanɔj* ‘which?’, *hatha* ‘how much?’, *hani* ‘where?’, *hakhi* ‘when?’, *haya* ‘how?’. In indefinite uses these words may carry the indefinite suffix *dum* (cf. *dum vi* ‘become’): *hatha-dum* ‘however much’.

5.5 Numbers

Nepali numbers are used for numbers over four, and often for lesser numbers as well. The numbers from one to four show a rudimentary classifier system: *pu* ‘human’ vs *uŋ* ‘non-human’. The numbers are: *kɔlu* ‘one’ (general); *kɔŋpu* ‘one:HUM’, *nakpu*, *nauŋ* ‘two’; *tsʰukpu*, *tsʰuuŋ* ‘three’; *bliuŋ* ‘four’ (rarely used) (9).

5.6 Adjectives, adverbs

There is no lexical class of adjectives, i.e. words which serve to modify a noun without the genitive postposition *mo*, apart from demonstratives, although the genitive mark is occasionally omitted. Most descriptive terms are participles or phonæsthetic words (below).

Colour terms are often cited with the suffix *mi* (?< genitive *mo*) e.g. *dawaymi* ‘white’. As modifiers, they usually appear with the suffix *ba~baha~bay* ‘like’ and the genitive postposition: *jitsʰij-baha-mo* *kʰursani* ‘red chili-pepper’ <Is52>. As predicate attributes they appear with the same suffix and a copula: *ay xom jitshij-bay pɔ:tse* ‘my tomb will be red [i.e. freshly excavated]’ <Ys3>.

There is a very large class of phonæsthetic words. Some function as sound-effects, but most function as very specific manner adverbs (often marked by the postposition *ha*) or as adjectives (marked by the postposition *mo*), e.g. *kulkul* ‘round’: *kulkul-mo mo:li* ‘round belly’; *kulkul-ha xʷā:tse?* ‘are you fully [“roundly”] sated?’ Some phonæsthetic manner adverbs have a formative *sit*, e.g. *krimsit* ‘in/to the right spot’: *krimsit tɔ:ta* ‘safely put away’.

Among the basic locative adverbs are *wani~ani* ‘up’, *huti~uti* ‘down’, *wati~ati* ‘across’; the corresponding directionals are *jō:kʰa* ‘down’, *lɔ:kʰa* ‘up’, *dɔkʰa* ‘across’; *hɔ:ku* ‘before’, *nɔŋna* ‘behind, after’.

5.7 Postpositions, particles

Postpositions serve to mark the function of an NP or a nominalized clause in a larger unit, NP or clause. Examples: *ha* ‘ERGATIVE/INSTRUMENTAL/ADVERBIAL’; *kʰen* ‘from/after’; *mo* ‘GENITIVE’; *tʰik* ‘like’; *nɔŋ* ‘at/with/after’; (*mo-*)*li:si* ‘for’; *he* ‘in/while’; *bɔŋ* ‘as far/long/much as’; *tiliq* ‘because of’; *bʰənda* ‘than’ (5); *sa* ‘SUBORDINATOR’ (26), *nana* ‘along, PROGRESSIVE’ (18).

For topic, focus and sentence particles, see §7.

6 SYNTAX

A sentence consists essentially of a finite clause, that is, a clause with a conjugated indicative or imperative verb or compound verb and, optionally, core and non-core arguments consisting of noun phrases, adverbs or subordinate clauses, whose function is marked by postpositions. The finite verb is last in its clause, except when a topical argument is right-dislocated after the verb ((10), (20)). There is no requirement that core arguments be present in the clause in lexical or pronominal form. The informational status of arguments may be marked by topic or focus particles. Functors – postpositions, discourse particles, complementizers, etc. – except the negative markers – follow what is in their scope.

6.1 Assertion, interrogation, negation

The main verb of a declarative sentence, positive or negative, may carry the assertive suffix *m~mi* ((2), (3), (4), (7), (10), (11), (17), (18), (20), (21), (23), (27)). This mark is not used in interrogative, imperative or conditional clauses. Questions are marked by intonation, or by tags like *ki ma* ‘or not’ ((17), (29)).

6.2 Intransitive clause structure

The verb in an intransitive clause shows agreement with an S, in the absolute case.

- (1) *ay abə xʷaptsɔ met* (BBs11)
1S:OBL now spouse die:23S:NPST
'Now my wife will die.'
- (2) *nukun le gu itʰa bela na pʰi:yɔmɪ* (AAs72)
tomorrow too 1 this.much time INTS come:1s:NPST:ASS
'Tomorrow I will come again at this same time.'

Arguments marked by the postposition *ha* in intransitive clauses are instrumental, not ergative in case (but see (12), (13)). The verb in (3) has unambiguously intransitive morphology.

- (3) *mi-k^hata buti-ha benmem*
that-PL meal-INST satisfy:3p:PST:ASS
'They had enough food'. *lit.* 'They were satisfied by the food.'

Meteorological verbs are intransitive, with 3s morphology: *lɔ:gaj kak* (sun shine:3s:NPST) 'the sun is shining'; *naðnum toy̥mi* (rain rain:3s:PST:ASS) 'it rained' or 'it is raining' (the past is used with inchoative sense).

6.3 Copular sentences

Identification or attributive predication requires a verb, typically the copula *nɔ(t)* vi (irregular) 'be'. All such propositions use the nominal negator *maaŋ* 'not', before either the predicate or the copula (6). Inchoative and causative senses are expressed using verbs like *dum* vi 'become', *pa* vt, vtt 'make', *pɔn* vr 'become, pretend to be', etc., in place of the copula ((5), (6), (9)).

- (4) *ɛ gɔn su nɔn ibɛ? -gu nɔ:ŋɔm* (BBs26) (BBs28)
hey you who be:2s here - 1 be:1S:NPST:ASS
[In the dark.] 'Who are you in here? – It's me!'
- (5) *pok^holi-b^hənda mi xɔ:ta dum ixtse* (AAs5)
king-than that big:PPT become:23s say:REFL:23s
'He became richer than the king.'
- (6) *sisit.ha gɔn maaŋ sixtɔɔ pɔ:tse*
know:CJP you NNEG know:HUM make:23S:REFL
'Although you know, you play dumb.'

Existence (8), situation ((7), (27)), and possession (9) are also expressed using the copula, which can be negated by *ma* 'not'.

- (7) *mi nɔkts^huij-he siktaŋ mi hɔ:cay-he nɔ:mi* (AAs107)
that ear-in wear:VN:PATIENT that open:VN:INST-in be:23S:ASS
'The earring is in the box.'
- (8) *ã:ki-t^hik-mo wa:ju han iɛ ma nɔ:me* (DDs18)
1pe:OBL-like-GEN Hayu where also NEG be:3p
'There are no Hayu like us anywhere.'
- (9) *a-mɔ kɔ lɔm-nɔŋ kɔ nakpu x^waptsɔ dum -*
3s-GEN TOP road-on TOP two:HUM spouse become:23s
kem-he kɔlu x^waptsɔ nɔ [...] dzəmmə kɔ ts^hukpu dūme (AAs35)
house-in one spouse be:23s all TOP three:HUM become:3p
'Now he had two wives on the road – he had one at home – altogether they were three.'

6.4 The transitive clause

The verb in a transitive clause shows agreement with two core arguments, A, roughly speaking the agent, marked by *ha* 'ERGATIVE', and O, a patient or dative object, unmarked (absolutive).

- (10) *ga t^hɔŋnɔmi gɔn, mine* (AAs57)
1s:ERG convey:1s→2s:ASS 2s there
'I will take you there.'

-
- (11) *kɔlu bud^ha mants^he-ha a bari-nɔŋ lalat.ha rã:pi*
one old man-ERG 3S:OBL field-in go:CJP taro
lɔxtɔm are (Bs1)
plant:23s→3s:PST:ASS HEARSAY
'An old man went to his field and planted taro, they say.' (Beginning of a story.)

Transitive verbs may have inanimate A (12).

- (12) *t̥ri nu?ma-k^heri na?num-ha t̥ɔpsɔŋ, gu dzǔ:sa-ha ŋsɔŋ*
today day-time rain-ERG throw:23s→1s:PST 1 fever-ERG meet:2s→1s:PST
'Today, during the day, I was soaked by the rain (*lit.* "rain threw me to the ground") and I caught a fever' (*lit.* 'fever caught me').

Although the suffix *sɔŋ* can also index a 1sg intransitive s, the verbs *t̥ɔp* and *ot* are transitive, as can be seen in other contexts, where they take personal A. But the deponent verb in (13), whose morphology also shows agreement with a personal o (often conveniently translated as a subject), could never take a personal A. Still, the impersonal A could be analysed as an instrument (3).

- (13) *kɔmi ti-ha t̥imtɔme*
3 water-ERG splash:3→3(p)
'They got sprinkled with water' (*t̥im* vd 'be struck from above (by a falling object)').

Other verbs of this type, of which there are about a dozen, are *kɔt* 'be pricked', *dzi* 'be affected by a bad odour'. All such verbs use the applicative *t*-suffixes rather than *k*-suffixes.

6.5 The ditransitive clause

Many transitive verbs can appear in a ditransitive clause, with an A and two absolute arguments, po (*primary object*) and so (*secondary object*) (Dryer 1986). The verb shows agreement with A and po. Semantically, so generally corresponds to a theme or patient, and po to a beneficiary, goal, or (rarely) a causee. In a ditransitive main clause, the applicative form of the main verb must be used if such a form is available for the particular verb and for the appropriate persons and numbers of A and po. The lack of such a form, however, is no obstacle to the use of the construction.

The verbs *ha* vtt 'give', *mɔm* vtt 'give (food to eat)', *t̥un* vtt 'give (to drink)' are inherently ditransitive, always agreeing with the recipient as po.

Among verbs that appear in both transitive and ditransitive clauses, transitive verbs which take an inanimate o use the ditransitive construction freely whenever a beneficiary is to be indicated. Examples are *rok* vt, vtt 'plough' and *ke(t)* vt, vtt 'peel': *ga roxkɔŋmi* (–APP) 'I ploughed it', vs *ga roktoŋmem* (+APP) 'I ploughed it for them'; *ke:kɔ* (–APP) 'I peeled it' vs *kextɔ* (+APP) 'I peeled it for him'. The verb *rok* can also take the bullock as po: *gu roksɔŋ* 'you used me to plow' (Cs13) (elsewhere: 'he/you plowed it for me').

In general, the ditransitive construction may be used depending on the verb and on referential and semantic properties of the arguments. It is rare to find a ditransitive construction in which the so is human, or otherwise particularly salient: first or second person sos are never encountered. In many situations, speakers can choose between the ditransitive construction or a simple transitive with a beneficiary marked by *l:si* 'for'. In (14) the speaker first uses a non-applicative verb with a beneficiary marked by *l:si*, but then changes to an applicative form (making the postposition redundant).

Note also the condition marked as a contrastive topic (*tɔ*) and the conclusion marked as contrary-to-fact (*pʰen*).

- (14) *apʰnɔ rɔ:mi-l̥i:s̥ iṭʰa-bɔŋ̥ ta:kɔ l̥e tɔ— ta:tɔ l̥e tɔ*
own wife-for this-EXTENT put:23s→3s:PST also CTR put:23s→3s:APP also CTR
a rɔ:mi-mɔ a iṭʰum-nɔŋ̥ hatʰa blosek nɔ pʰen hɔla (Zs23)
3:OBL wife-GEN 3:OBL heart-LOC how:much joyful be:3s CFCT perhaps
'Had he left even this much [gesture] for his wife – left her even this much – how much joy might have been in her heart!'

In a ditransitive clause, PO and so may have the same referent. Thus, related to *poxkɔŋ* (–APP) 'I raised it', 'I roused him', etc., we find *poktɔŋ* (+APP) 'I raised it for him', 'I roused him for him' but also 'I helped him up' (i=j). That the latter is ditransitive is clear from the use of a distinctively applicative form.

6.6 Reflexive clauses

Reflexive forms of transitive verbs (Table 34.1) are very common, most with a 'middle' sense. The single core argument appears in the absolute case. Non-singular reflexive forms may have either reflexive or reciprocal sense (but note the specialized reciprocal form in (19)).

- (15) *mi wɔlta caĩ u watʰe muxtsem ixtse* (AAs45)
that wither:PPT TOP there over.there sit:23s:REFL:ASS say:23s:REFL
'That withered one sat way over there, they say.' (*mut vtt* 'seat so').
- (16) *ka:tɔ cʊʔnatsʰik*
friend look:2di:REFL:NPST
'Friend, let us look at ourselves (reflected in the water)!'

Reflexive clauses with object NPs are clearly related to ditransitive clauses. For example, in *buti dzā:tse* 'eat your meal!' (*dza*-*dzɔ* vt 'eat') the reflexive indicates identity of reference between the eater and the beneficiary; *buti* 'cooked grain' has the function o2. On the other hand, when a particular food item is in question (17) we find transitive forms.

- (17) *ɛ, gɔna dza:kɔ ki hajə? makʰi dzaymi bvol-i* (Ls9)
hey 2s:ERG eat:23s→3s:PST or how? not eat:1s→3:NPST:ASS e.brother-voc
[Seeing that a quantity of edible flowers is diminished (in fact, the flowers have dried up):] 'Hey, did you eat them or what? – I didn't eat (/haven't eaten) any, brother.'

Similarly, in threats – *ga gɔn dzɔ:nɔm* 'I'll eat you' – or when a child eats something other than food or medicine – *hakʰilehak kɔ dzɔ:mi* 'he always eats earth' – the reflexive is not used.

6.7 Compound tenses, auxiliaries

Compound predicates use the verbs *nɔ(t)* 'be' and *pa* 'do' as auxiliaries; this is also the usual way of integrating Nepali verbal loans (cf. *kʰuwai pɔ:sɔym̥i* 'he fed me' in §7 < Nep. *kʰuwānu* 'to feed'). The bare-root gerundive with the progressive marker may appear as the head of a subordinate clause, or as a main verb with a finite auxiliary.

- (18) *gɔna ka:tso lɔ tu la, ga kʰɔʔ-nana paymi pa*
 2:ERG friend leaf pick go 1:ERG cook-PROG do:1s→3s:ASS COMP
ixtɔm ixtse sjal-ha (Is32)
 say:1s→3s:ASS say:23s:REFL jackal-ERG
 “You, friend, go pick leaves [to make plates]. I’ll go on cooking,” Jackal told him
 (they say).’

The conjunctive participle followed by the auxiliary *nɔ(t)* can have a perfect (*mumut.ha nɔ* ‘he is/was sitting’) or, with a transitive verb, a passive sense (23).

The reciprocal is a transitive construction, formed from the fully reduplicated gerundive and the auxiliary *pa*.

- (19) “*gɔna le dʰerai jiu gɔ:ju*” *pa nakpu-ja “gɔ:tsʰik-ju”*
 2-ERG also much life live:23:NPST:OPT COMP two-ERG live:1pi:NPST:OPT
pa ha-ha pɔ:tsʰɛ, asik ha-ha pɔ:tsʰɛ (AAs134–5)
 COMP give-give do:3d:PST blessing GIVE-GIVE do:3d:PST
 ‘Saying “May you, too, have long life! May we (DU.INCL) have long life!” they
 blessed each other.’

6.8 Compound predicates; modals

A number of operators such as aspectuals, directionals, etc., appear as vectors following the bare-root gerundive of the main clausal verb or vector. In such vv concatenations, only the vector has a finite form. The vector often agrees in transitivity with the governed verb and, if the latter is transitive, may agree with its logical object. Thus, *cut* ‘finish’ appears as a reflexive with the intransitive *pʰi(t)* ‘come’: *pʰi cuxtse* ‘he has already come’, but as a transitive in (20), showing agreement with the object of ‘eat’.

- (20) *ã:ki gaũ-mo tsɔ-kʰata dzəmmɔi dza cuxtɔmem bumi pixpi-ha* (Us2)
 1pe:OBL village-of child-PL all:INTS eat finish:3→3(p):ASS bumi g’mother-ERG
 ‘Grandma Bumi had eaten up all the children of our village.’

The verb *la(t)* vi irr. ‘go’ has a number of uses as a vector: centrifugal: *lon laʔnatsʰem* (run go:3d:REFL:ASS) ‘they (dual) ran away’ (Ds29); change of state with stative verbs (which themselves have no finite forms): *ima wɔl lăsɔŋ* (thus withered go:1s:PST) ‘I have become withered up like this’ (AAs52); agentless ‘go-passive’: *ma jer la* (not see go:23s:NPST) ‘it will not be seen’. The latter construction is also used with intransitive and stative verbs: a *tʰum dzik laxtse* (3s:OBL heart break go:23s:REFL) ‘her heart broke’ (Zs43).

Modals and some other complement-taking predicates also govern the root-gerundive. Examples include *dak* ‘be obliged/necessary to’, *pʰat* vi and *ɔn* vtt ‘be able’, *kn* vtt ‘try’, *wat* vtt ‘stop’, *piŋ* vt ‘CAUSATIVE’ (= ‘send so.’), *bi* vtt ‘permit’, *dzok* vi ‘know how to’, *daj* (invariable) ‘be about to’ (also used with the copula as a main verb). The verbs *ha* vtt ‘give/permit’, *bi* vt ‘request permission to’, *liŋ* vt ‘get to’ take infinitive complements: *gu tɔp-mo ha:sɔŋ* ‘let me play it [a drum]’. In the causative (21) and permissive constructions, the causee appears in the absolute case.

- (21) *gɔna ay ixta tʰaǔ-he tʰɔŋɔ tɔ, ga gɔn umbe*
 2s:ERG 1:OBL say:PPT place-LOC escort:IMP:2s→1s TOP 1:ERG 2 now
dza... piŋnɔm. xʷan-xʷan dzā:tse gɔn (AAs54, 55)
 eat... CAUS:1s→2s:ASS satisfy-satisfy eat:2s:REFL:IMP 2
 ‘You just take me to the place I say, and I’ll let you eat right now. Eat your fill!’

6.9 Subordinate clauses, sentential complements

Subordinate clauses with adverbial function are non-finite. The verb is a gerundive, either the bare root followed by a postposition indicating its function (as for an NP adjunct), or a progressive or reduplicated conjunctive ‘participle’ (§4.5). The latter types tend to share a core argument with the main clause.

When, as often, a core argument of the subordinate is coreferent with the actor (s or a) of the higher clause, its case is determined by its role in the higher clause. Thus, in (11), the intransitive s of the subordinate is the a of the main clause, which appears in the ergative case. Much less frequently, the o of the subordinate may be coreferent with the actor of the main clause (22), in which case the conjunctive participle has a passive interpretation.

- (22) *sjal hū:pukumi-ha tsitsik.ha men* (Xs15)
 jackal wasp-ERG bite:CJP die:23s:PST
 ‘The jackal was stung by the wasps and died.’

The same construction with the copula can be used as a kind of passive.

- (23) *gu kɔ xō:ta pok'oli.ha ima itt.ha nɔ:ŋyɔ* (AAs29)
 I TOP big:PPT chief-ERG thus say:CJP be:1s:NPST:ASS
 ‘I am commanded thus by the great king.’

Subordinate clauses marked by other postpositions have no particular tendency to share an argument with the higher clause (24).

- (24) *a umu kʰɔ(k)-kʰen dayday le dum* (AAs112)
 3s:OBL mother walk-after bright also become:23s
 ‘After her mother left, it became light.’

The subordinator *sa* governs a gerundive negated by *ma*, as in *kɔkɔri ma ɔk-sa* (cock NEG cry-SUB) ‘before cock-crow’.

- (25) *gon kem-he lalat.ha ma dza:-sa ma teʔnɔ* (Ds12)
 2s house-in take:CJP NEG eat-SUB NEG release:1s→2s:NPST
 ‘I’ll take you home and I won’t let you go without eating you!’

Conditions are finite clauses usually marked by *nɔm* ‘if’ (< *nɔm* ‘be:23s:ASS’) (but cf. (14) and (21), where conditions are marked as contrastive topics).

- (26) *vt nɔm dzo:mi ixtsime* (Ds61)
 catch:23s→3s:NPST if eat:23s→3s:NPST say:3p:REFL
 ‘If he catches her he’ll eat her, they say.’

Conjunctive participles (with rare exceptions) and conditionals are negated using *ma* (*dza ma līglīgha* ‘not getting to eat’ (BBs14)), while clauses with the simple gerundive and a postposition are negated with *maay*, e.g. *dza maay bit-kʰen* ‘since they don’t let [me] eat’ (AAs53).

Complements of verbs implying citation usually appear as direct citations marked by the complementizer *paha~pa*, composed of the verb *pa* vt, vtt ‘do’ with *ha* ‘MANNER’ ((18), (19)).

- (27) *dzəmmə si:kɔ inε inε nɔm paha* (AAs109)
 all know:23→3s:PST here here be:3s:ASS comp
 ‘He knew where everything was.’ [lit. ‘He knew everything, that it was here and here.’]

7 INFORMATION STRUCTURE AND DISCOURSE PARTICLES

Information structure is marked by word order (including the omission of NP arguments, on which there is no syntactic restriction), intonation, and discourse particles, which follow the element in their scope.

An active topic carried over from the immediately preceding context may be omitted, or it may be placed in antitopic position after the verb. The topicalizer *kɔ* (or Nepali *cañ*) may serve to reactivate a topic ((9), (15)).

There are numerous focus particles: *na* ‘INTENSIFIER’ (2); *le* ‘also, even’ ((2), (8), (14), (24)); *dzr* ‘only’; *pai*, *pi* ‘counter-expectancy’: *expi pi kʰuwai pɔ:sɔy̩mi* (shit FOC feed do:23s→1s:PST:ASS) ‘he fed me shit!’ (Is43).

Sentence particles: *rɔ* ‘[rhetorical question]’, *re*, *are* ‘[hearsay]’, *de~te* ‘[encouragement]’, *tɔ~tɔ* ‘[contrast/implied condition/challenge]’ ((14), (21)).

8 THE NOUN PHRASE

A noun phrase may be followed by a postposition to mark its function, and by discourse particles. Modifiers precede the head; these include possessive pronouns, demonstratives, NPs marked by the genitive postposition *mo* ((8), (20)), participles, relative clauses, lexical adjectives, etc. Thus: *rā:pi-mo lɔ* (taro-GEN leaf) ‘taro leaf’; *haya-mo siy̩tɔŋ* (how-GEN man) ‘what manner of man?’; *xɔ:ta pukʰvɔli* (big:PPT chief) ‘big chief’; *olawɔ pʰum* ‘huge tree’. Quantifiers may precede or follow; in the latter case any postposition is repeated, e.g. *nɔnɔtsɔ-ha nakpu-ha* (sister:COLLECTIVE-ERG two:HUM-ERG) (Ds19). Adverbial modifiers precede adjectives: *itʰara xɔ:ta le gɔtji* (thus big:PPT foot have:ACTP) ‘having feet big like this’ (AAs1).

Any potential nominal modifier may serve as head. That is, *xɔ:ta* ‘big’ (above) may mean ‘big one’ (e.g. *wɔlta* in (15)). Possessive pronouns require reinforcement by *mo* ‘GEN’ when used alone: *aymo nɔm* ‘it is mine’ (cf. *ay kem* ‘my house’).

Coordination is usually realized by simple juxtaposition: *ga gɔt le tsiknɔ* (1s:ERG hand foot break:1s→2s:NPT) ‘I’ll break your arms and legs for you!’ (ditransitive). The postposition *nɔj* ‘after, with’ may also be used: *bʰalu-nɔj sjal tɔxtɔtsʰem* (bear-with jackal chase:3→3(d):PST) ‘he chased the bear and the jackal’ (Hs6) (note the dual agreement).

8.1 Relative clause; deverbal and participial modifiers

Relative clauses (i.e. clauses which modify a noun) have their verb in a participial or deverbal nominal form; the element corresponding to the head is omitted from the relative. When the head corresponds to the intransitive s of the relative, the verb form used depends on the tense, the past/passive participle in *ta* being used for realized states or completed events, and the agentive/active participle in *ji* for non-past or ongoing ones.

- (28) *āki layka-kʰen dzɔkta deuta* (Qs9)
1pe:OBL Sri.Lanka-from come.up:PPT god
'our goddess who came up from Sri Lanka'

- (29) *inɔŋ jɔ:kʰa latji siy̩tɔŋ nakpu gɔna jɛ:kɔ ki ma?* (Ds36)
here below go:ACTP man two 2s:ERG see:23s→3s:PST or not
– *ma jɛ:kɔŋ, tʰa:tsɔ, ga kɔ* (Ds37)
– not see:1s→3s:PST grandson 1s:ERG TOP
'Did you see two people going down here?' – 'I didn't see them, grandson.'

With transitive verbs, the active participle (*in ji*) is used when the antecedent corresponds to the A of the relative: *jesɔ tjesɔ pa:paha dzɔ:ji siŋtɔŋ* (this.way that.way do:CJP eat:ACTP man) ‘a man who lives by chicanery’ (*lit.* ‘eats by doing thus and so’) (Ns43); the passive participle (*in ta*) is used if it corresponds to the o (or other non-A): *tami-ha cupta cu?wa-kʰata* (daughter-ERG wear:PPT clothing-PL) ‘the clothes and ornaments which the daughter wore’ (AAs83). When the antecedent is the impersonal A of a deponent verb, the active participle is also possible: *mi kɔ̡tji tsu* ‘that pricking thorn’ (arguing against interpretation as an instrumental).

A pronominal s or A in the relative appears in the oblique/possessive form: *ay ixta tʰaū* ‘the place I say’ (21).

When the antecedent corresponds to a non-core argument in the relative, the past/pассив participle in *ta* can generally be used: *ima-mo loxta tʰaū-nɔŋnɔ* (thus-GEN plant:PPT place-from) ‘from the place where it had been planted like this’ (AAs113). Often, however, the appropriate nominal (here the locative nominal) is used: *bumi imluŋ kem-nɔŋ* (bumi sleep:NV:LOC house-at) ‘in the house where Bumi slept’ (Us33). The absolute (not genitive) case of ‘Bumi’ shows that it is the whole clause ‘Bumi sleeps’ that is nominalized. But a pronominal argument would appear in the oblique form: *uy khɔk-siŋ-he* ‘when it’s time for you to leave’ vs *khɔkɔri ɔk-siŋ -nɔŋ* ‘at cock-crow’.

9 ADDITIONAL ABBREVIATIONS

1s, etc.	1, 2, 3 represent persons, s singular, d dual, p plural, i inclusive, e exclusive
3s → 1de, etc.	3rd person singular A, 1st dual exclusive O.
ACTP	present/active participle
APP	applicative
ASS	assertive
CFCT	counterfactual
CJP	conjunctive participle
CTR	contrast
HUM	human classifier
INTS	intensifier (‘— itself’)
NNEG	nominal negator
NP	noun phrase
NPST	non-past
OBL	oblique/possessive (pronoun)
PPT	past/passive participle
PST	past
vd	deponent verb
vs	stative verb
vt	transitive verb using <i>k</i> -suffixes
vtt	transitive verb using <i>t</i> -suffixes

NOTE

- 1 Online-accessible recorded texts (Michailovsky 2000–) are cited by capital-letter text identifier and a sentence number.

REFERENCES

- Dryer, Matthew S. (1986) ‘Primary objects, secondary objects, and anti-dative’, *Language* 62.4: 808–845.
- Government of Nepal (2012) *National Population and Housing Census 2011*. Volume 1. National Report. Kathmandu: Central Bureau of Statistics.
- Hodgson, Brian Houghton (1857) ‘Grammatical analysis of the Hayu language’, *JASB* 26: 373–475 [errata in *JASB* 27].
- Michailovsky, Boyd (1974) ‘Hayu typology and verbal morphology’, *LTBA* 1.1: 1–26 [errata in Michailovsky 1975: 215n1].
- (1975) ‘Notes on the Kiranti verb’, *LTBA* 2.2: 183–218.
- (1988) *La langue hayu*. Paris: CNRS.
- (2000–) Hayu recorded texts, with synchronized transcription and translation. Available at: lacito.vjf.cnrs.fr/pangloss/languages/Hayu_en.htm, accessed 1 January 2016. Cited (by text code and sentence) in this chapter: AA: Merchant; B: Taro; BB: Rich and poor; C: Ingratitude; D: Two sisters (v1); DD: Hayu identity; HBIS: Jackal and Bear (v1); I: Jackal and Bear (v2); L: Fratricide; M: Daughter-in-law (v1); N: Complaint; Q: Origins; U: Cannibal grandmother; X: Hornets’ nest; Z: Daughter-in-law (v2).
- Michailovsky, Boyd and Mazaudon, Martine (1973) ‘Notes on the Hayu language’, *Kailash* 1.2: 135–152.

CHAPTER THIRTY-FIVE

BELHARE

Balthasar Bickel

1 INTRODUCTION

Belhare is a Kiranti language spoken by about 2,000 people living on the Belhara (written *Belahārā*) hill, one of the southern foothills of the Himalayas situated in Eastern Nepal (Dhankutā district; Kośī zone; 87° 18' E and 26° 57' N; c. 1,150 m above sea level). The Belhare are culturally closely related to the neighbouring Athpare community of Dhan-kuta. This ethnic affiliation leads speakers to refer to their language also as ‘Athpare’, although differences in morphology make the two languages mutually unintelligible for practical purposes. In terms of religion and mythology, the Belhare and Athpare are distinct from both the Rai and the Limbu traditions, but in most other respects they share the general Kiranti patterns of shamanist ancestor worship and a high degree of social compartmentalization. The Belhare are virtually all farmers, and a series of food taboos brings about a high degree of self-subsistence and a strong reluctance to travel. Partly as a result of this, language maintenance is relatively high, and most children still learn Belhare as their first language. Nevertheless, speakers are all bilingual in Nepali, and Belhare discourse is rife with code-switching, borrowings and stylistic calques.

Apart from occasional variation in morphophonology (e.g. *ŋka-chi* vs *ŋke-chi* for ‘we two’), Belhare is internally homogeneous; its closest relatives are Athpare, Yakkha, Chintang and probably also Chiling (which is so far undescribed).

2 PHONOLOGY

Table 35.1 summarizes the phoneme inventory of Belhare. *N* is a nasality feature that associates with the preceding syllable, but is realized as an assimilating syllabic nasal when initial in the phonological phrase (examples in section 2.1). Segments in brackets occur in loanwords only, but the breathy voiced consonants also appear as regular allophones of aspirated voiceless consonants between sonorants, e.g. /taŋkhe:/ [t̪an̪.g̪e:k] ‘head’ or /laphe:/ [lab̪e] ‘caught’, and are represented as such in the practical orthography used here. Coronals are postalveolar; /t/ has a secondary glottal release before laterals and dissimilates to [?] before palatal glides. Apart from this, there is not much allophonic variation beyond low-level assimilation of vowels to consonantal points of articulation.

There is no length contrast, but open-syllable vowels optionally lengthen under main stress. Diphthongs arise only as the result of stem alternation in verbs, on which see section 2.3. Belhare has no phonemic tones.

2.1 Phonological domains

The major syllable canon of Belhare is (c)v(c). While onsets can be any consonant, codas are limited to nasals and unreleased versions of /p, t, k, ʔ/. There is in addition a minor

TABLE 35.1 BELHARE PHONEMES

k	kh	g	(gh)	ŋ	i	ĩ	ũ	u
t	th	d	(dh)	n	e	o		
c	ch	(j)	(jh)			ʌ		
p	ph	b	(bh)	m		a		
?	h	r	(rh)	s	y	w	n	

Note: <c> stands for [ts] and <j> for [dʒ].

syllable type consisting of the syllabic nasal that arises from *N* when it is initial in a phonological phrase, e.g. *nka*: [n̩.ka] ‘I’ or *n-ta-he* ‘3NSG-come-PAST’: [n̩.ta.he]. When located inside phonological phrases, *n* docks onto the preceding syllable, e.g. *un-chik n-ta-he* ‘3-NSG 3NSG-come-PAST’ is realized as [*un.chiŋ ta.he*].

Apart from a few lexical exceptions, main stress is initial in the phonological word. Secondary stress follows a trochaic rhythm of bimoraic feet (with closed syllables counting as heavy), but is blocked from final open syllables. Phonetically, unstressed open syllables can be reduced in non-final position, e.g. (ϕ 'phagi)(ϕ , det)(ϕ , lem) ‘butterfly’: ['phagi, det[?], lem] ~ ['phag, det[?], lem].

The phonological word starts at the first stem boundary, e.g. [?uhopchi] from *u-(_ωhop-chi)* ‘3SG.POSS-calebush- NSG’, unless this results in end stress, e.g. ['uhop], and includes enclitics. The prefix-excluding domain also defines the application of inter-sonorant voicing (e.g. [Pababu] from (_ω ap-a=phu) ‘come.over-IMP=HS’ or [*inday*] from (_ω in-tay) ‘egg-LOC.UP’ as opposed to [minta] rather than *[minda] from *min-(_ωta)* ‘NEG.CVB-come’), glottal prothesis in onset-free syllables (e.g. [?uukma] from (_ω u-uk-ma) ‘roast-bring.down-INF’) and various vowel contraction rules (see below). But some phonological rules are defined by a larger domain that includes prefixes, e.g. velar codas are dropped at the end of prosodic words. For this purpose the prefix-stem boundary counts as word-internal, e.g. [tayu] from (_ω ta-yuk) ‘come-NONPAST’ as opposed to [tayukcha] from (_ω ta-yuk=cha) ‘come-NONPAST=ADD’ as well as [aklu], not *[alu], from (_ω ak-lu) ‘OPT-tell’.

2.2 General morphophonology

The single most important morphophonological regularity is due to consonant prothesis to avoid morpheme boundaries within feet (Bickel 2003b), e.g. (ϕ 'lap-)(ϕ , buk-)ma from *lap-uk-ma* ‘to catch and bring down’ instead of *(ϕ 'la)(ϕ , b-uk-)ma. If prothesis is unable to rescue feet from being heteromorphemic, the relevant string is underparsed and no stress is assigned, e.g. (ϕ 'la)b-u-k=cha ‘catch-3P-2A=ADD’ instead of *(ϕ 'lap)(ϕ , b-u-k)=cha or *(ϕ 'la)(ϕ , b-uk)=cha. However, lexical roots always receive stress, and as a result of this, some heteromorphemic feet are left unrepaired, cf. e.g. (ϕ 'lab-he) ‘catch-PAST’. Prothetic consonants assimilate in place of articulation to preceding segments but are always voiced.

Apart from this rule, voicing is also a general phenomenon applying within phonological words (cf. section 2.1). Stops generally voice between sonorants, with a few exceptions of invariably unvoiced markers (most notably -t ‘non-past’). Intervocalic /t/ not only becomes voiced but also continuant, cf. [kar^be] from *kat-he* ‘come.up-PAST’ vs [k^bhemdahe] from *khem-ta-he* ‘hear-come-PAST’. (The aspiration feature in the past tense marker is realized as breathy voice, following the breathy voice rule that affects all aspirated voiceless consonants.)

There are two deletion rules, one affecting vowels, the other consonants. First, /a/ and /i/ regularly delete before /u/ inside phonological words (e.g. [luisuha] from *lui-sa-u-hak* ‘tell-TR.PERF-3P-PERF’) and /u/ deletes after /e/ (e.g. [lur^he] from *lut-he-u* ‘tell-PAST-3P’). Second, the rule of velar-coda drop noted in section 2.1 applies obligatorily in grammatical morphemes and optionally in lexical morphemes, cf. e.g. *maʔi-lok* = to ‘man-COMIT = ID’ vs *maʔi-lo* ‘man-COMIT’ or *dabhek* ~ *dabhe* ‘sickle’.

Nasals assimilate at initial boundaries of suffixes, e.g. /unna/ from *un-ŋa* ‘3sg-ERG’.

2.3 Verb stem alternation

Virtually all verb roots fit into the (c)v(c) syllable canon (often augmented by the reflexes of Tibeto-Burman *-s and *-t suffixes), but a series of inflectional endings requires a cvv shape. cv roots are fit into this shape by epenthesis of /i/ or, after /i/, /u/ (e.g. *so-* ~ *soi-* ‘wait’, *khi-* ~ *khiu-* ‘quarrel’, etc.). Root codas are vocalized, while retaining their tongue and velum states, e.g. *yaj-* ~ *yaū-* ‘carry by hand’, *yak-* ~ *yau-* ‘stay overnight’, etc. Bilabials are exempted from this and remain unchanged. Final /t/ not only turns into /i/, but it also effects a concomitant glottal closure (e.g. *kat-* ~ *kai?* ‘come up’), which is usually incomplete and results in laryngealization (creakiness) of /i/.

3 INFLECTIONAL MORPHOLOGY

Belhare is a double-marking language on both the clause and the NP level: the dependent of a verb or a noun is marked by case affixes on the dependent itself, and also on the head by agreement affixes:

- (1) a. *ŋka-ha a-tak*
1SG-GEN 1SG.POSS-friend
'my friend'
- b. *un-chik-ŋa ŋka ma-ŋ-ni-at-ni*
3-NSG-ERG 1SG 1SGP-3NSGA-see-PAST-NEG
'They didn't see me.'

Grammatical morphemes are phonologically concatenative and divide into prefixes, suffixes, circumfixes (e.g. *mi-n-lui-t-u-n* (3NSGA-NEG-tell-NPAST-3[sg]P-NEG) ‘they won’t tell him/her’) and simulfixes (e.g. *tai-ŋa-chi-ha* (come-INTR.PERF-[3]du[s]-PERF) ‘s/he has come’). Inflectional morphology sometimes splits lexical words into two parts which have no syntactic or semantic status of their own (at least not synchronically): cf. *la ŋŋ-umm-at-ni* (NEG-walk-PAST-NEG) ‘they didn’t walk around’ from *la um-ma* ‘walk-INF’.

3.1 Nominals

The general template for nominal inflection is (POSSESSOR-)Σ-NUMBER-CASE(-AGREEMENT). Apart from some exceptions noted below, the number and case morphology is the same for nouns, demonstratives and pronouns. The set of demonstratives includes proximate (*na* ‘this’) and remote (*i-na* ‘that’) forms as well as roots differentiating UP-, DOWN- and ACROSS-locations (Bickel 1997, 2000a). Pronouns are either personal (roots *ŋka* ~ *ŋke* ‘1’, *han* ‘2’, *un* ‘3’) or interrogative (roots *sa* ‘who’, *he* ‘which’, *yeti* ‘what’); indefinites are expressed by the generic noun *maʔi* ‘person’, the numeral *ibay* ‘one’, interrogatives (*sati* = *cha* ‘who = ADD’, i.e. ‘whoever’) or by zero (on the latter, see Bickel 2003a; Stoll and Bickel 2009).

TABLE 35.2 PRONOUNS AND DEMONSTRATIVES

	Singular	Plural	Dual
1st inclusive		<i>ŋke</i>	<i>ŋkechi</i>
1st exclusive	<i>ŋka</i>	<i>ŋkey</i>	<i>ŋkenχiŋ</i>
2nd	<i>han</i>	<i>hanik</i>	<i>hanchik</i>
3rd	<i>un</i>		<i>unchik</i>
demonstrative	<i>na</i>	<i>nakha</i>	<i>nakhachi</i>
remote demonstrative	<i>ina</i>	<i>inkha</i>	<i>inkhachi</i>

3.1.1 Number

Number is optional and rare with inanimates; it distinguishes singular (zero-marked) from non-singular (-chi or, with pronouns, -chik). Demonstratives as well as first and second person pronouns make an additional distinction between dual and plural (which is absent from the third person non-singular pronoun unchik); the first person differentiates exclusive from inclusive forms (Table 35.2).

Note that first person plural inclusive pronouns (and the corresponding agreement desinences) also cover generic reference (in the sense of French *on* or German *man*).

3.1.2 Case

The case forms are summarized in Table 35.3, with a brief description of their use and distribution.

The absolute is unmarked except for the interrogative pronoun *sa-ti*, where the absolute is marked by *-ti* in the singular (contrasting with *sa-a* ‘who-ERG’ and *sa-chi* ‘who-nsg [ABS]’). Occasionally, the Nepali dative *-lāi* is used on high-empathy objects, but for most speakers, this is not very common. The ergative is confined to transitive actor arguments, but, on inanimates, it also functions as an instrument, cause and force marker (see (3c) for an example). Ergative marking is split insofar as the first person singular pronoun never takes an ergative desinence. C-final stems are restricted to the *-ya* allomorph of the ergative. Except for the first person inclusive pronouns (*ŋke-a*, *ŋkechi-a*), V-final stems, by contrast, show free variation between *-a* and *-ya* (e.g. *ma?i-a ~ ma?i-ya* ‘person-ERG’ or *sa-a ~ sa-ya* ‘who-ERG’); there is a slight preference for *-ya* in repetitions. Free allomorphy is also characteristic of the mediative (*-lam* vs *-lamma*) and the allative (*-sam* vs *-samma*, from Nep. *samma*). The ablative form *-huj* is restricted to temporal and spatial concepts, e.g. *hambahuj* ‘from today on’, while the form *-etnahuj* is distributionally free. The underlying C in the locative triggers gemination of a subsequent C in the same phonological word, but deletes in all other environments (cf. *mi-et=to* ‘fire-LOC=ID’ vs *mi-e* ‘fire-LOC’). The locative allomorph *-pak* (which is subject to the velar coda drop rule, cf. section 2) is limited to environmental space demonstratives (e.g. *tu-ba* ‘up there’), some temporal roots (e.g. *hale-ba* ‘earlier’) and converses (*cok-sa-ba* ‘do-CVB-LOC’). Zero-marking of locatives is possible with inherently spatial nouns such as place names or *khim* ‘house’, ‘home’ in conjunction with directed motion verbs (see (10b) for an example). The interrogative root *he-* has a defective and irregular case paradigm: loc. *hene*, med. *hellam*, dir. *helley*. The form *hena* ‘which’ derives from *he-* suffixed by the enclitic attributive article *=na* (on which cf. sections 3.6 and 5.2). The demonstrative *na* has an irregular locative *ne-e* (but cf. the regular directive *na-lej*); dir. **ina-lej* and med. **ina-lam* contract to *illej* and *illam*, respectively (also cf. abs. pl. *inkha ~ ikha*). Case combination is possible only with the environmental

case locatives: *na-pmu-ley* (DEM-DOWN-DIR) ‘towards down there’. However, it is likely that the ablative allomorph in *-etnahuy* derives historically from the locative *-eC* suffixed by an article (=na) and the older ablative in *-huy*. In complex NPs, there is no case stacking (*Sufixaufnahme*) and neither is there case agreement. Belhare case-marking has phrasal scope.

3.1.3 Possession

Nouns can inflect for possessors (Table 35.4). In the non-singular, the relevant prefixes are closely related and sometimes identical with free pronouns (cf. Table 35.2). Nouns denoting kinship relations (e.g. *-ni* ‘paternal aunt’), body parts (*-mik* ‘eye’), psychological states (*-chom* ‘desire’) and topological notions (*-tem* ‘outer surface’) always take possessive prefixes (except for disyllabic kin terms used for address, e.g. *eba!* ‘uncle! (father’s elder brother)'). Note that topological nouns make up for the complete lack of adpositions in Belhare.

3.1.4 Agreement

When used as predicates, nouns are marked by person-indicating agreement markers. Compared to verb agreement (see Table 35.5), the possibilities are severely limited, however, and only involve *-ya* ‘[1]EX’ and *-kak* ‘2’; third person is zero-marked:

- (1) a. *ŋka n-tak-ŋya*
 1SG[ABS] 2SG.POSS-friend-[1]EX
 ‘I am your friend.’

TABLE 35.3 CASE MARKERS AND THEIR USE

absolutive	<i>-Ø ~ -ti</i>	S, all kinds of objects, experiencers in experiencer-as-goal constructions, predicate nominals
erg./instr.	<i>-ŋa ~ -a</i>	A, instruments, causes, forces
genitive	<i>-ŋahak ~ -hak</i> (after V)	possessors, attributive nouns in NPs
comitative	<i>-lok</i>	accompanying referents (NPs) or situations (clauses)
ablative	<i>-etnahuy ~ -huy</i>	from, after
mediative	<i>-lam(ma)</i>	via, through, from, in (language X)
allative	<i>-sam(ma)</i>	until, up to
directional	<i>-ley</i>	towards, in the direction of
locative	<i>-eC ~ -pak ~ -i</i>	at, in, on, to (on NPs and converb clauses)
UP-locative	<i>-ttay</i>	up at, in, on, to
DOWN-loc.	<i>-pmu</i>	down at, in, on, to
ACROSS-loc.	<i>-ŋya</i>	across at, in, on, to

TABLE 35.4 POSSESSIVE PREFIXES

	Singular	Plural	Dual
1st inclusive		<i>ŋke-</i>	<i>ŋkechi-</i>
1st exclusive	<i>a-</i>	<i>ŋkey-</i>	<i>ŋkeychinj-</i>
2nd	<i>n-</i>	<i>hani-</i>	<i>hanchi-</i>
3rd	<i>u-</i>		<i>unchi-</i>

- b. *un n-tak*
3[SG.ABS] 2SG.POSS-friend[-3SG]
'S/he is your friend.'

3.2 Adjectives

Belhare has a small set of adjectives which are marked either by the nominalizer *-khak* (cf. section 3.5) or by the enclitic article, the latter agreeing in number and indicating specific reference of the head noun (=na 'SG', =*kha* 'PL', =*khachi* 'DU'). One small subset of adjectives denotes concepts like 'big', 'small' or spatial configurations like 'bent' and represents for the most part frozen verb forms (whence the nominalizer). Another subset of five adjectives denotes colours ('white', 'black', 'red', 'green', 'yellow') and takes a different article: =*ma* (=ha in the non-singular), cf. *phabeley=ma khim* 'the red house' vs *ei?=na khim* 'the big house'.

Adjectival concepts are inherently comparative and an expression like *eikha* 'big' can equally be translated as 'bigger'. The standard of comparison can be made explicit by *bhanda*, which is borrowed from Nepali (*lit.* 'while saying'): *na bhanda eikha* 'bigger than this'.

3.3 Verbs

Inflectional affixes are listed in templatic form in Table 35.5 (from Bickel 1996, with revisions); Table 35.6 is a sample paradigm. An important feature of verb paradigms is the copying of nasals. In the imperative, negative *-n* copies iteratively around any syllable (e.g. neg. imp. 2du > 3sg: $\Sigma\text{-}n\text{-}a\text{-}n\text{-}ch\text{-}u\text{-}n$ from $\Sigma\text{-}a\text{-}ch\text{-}u\text{-}n$, 2sg > 3nsg: $\Sigma\text{-}u\text{-}n\text{-}chi\text{-}n$ from $\Sigma\text{-}u\text{-}n\text{-}chi$). In non-imperative forms, this is limited to copying around the sf8 filler *-chi* 'NSG' (e.g. neg. 3sg > 3nsg: $\Sigma\text{-}u\text{-}n\text{-}chi\text{-}n$ from $\Sigma\text{-}u\text{-}n\text{-}chi$). The first person markers of sf 7 copy around *-chi* 'NSG' in all forms.

3.3.1 Agreement

The verb agrees with S, A and P arguments and in most cases distinguishes three numbers and persons as well as, with first person referents, exclusion and inclusion of addressees. While agreement markers are often underspecified (e.g. *-kak* '2') or ambiguous (e.g. *-m* '1PLA or 2PLA'), their combination always resolves this (e.g. *-mgak* < *-m-kak* '2PLA'). The absence of a prefix entails reference to '3SGS/A'; '3SGP' is expressed by *-u*, but note that /u/ regularly deletes after /e/ (cf. section 2.2), and that it also deletes where its retention would violate the syllable canon (e.g. /nlui?nij/ rather than */nluitunj/ from *lui-t-u-n(i)-η* 'NEG-tell-NPAST-3P-NEG-1SGA'; neg. *-ni* is *-n* after vowels).

3.3.2 Negation

Negation is marked by the circumfix *n-* . . . *-n(i)* (pf3 and sf6), but in inverse scenarios (3 > 1, 3 > 2, 2 > 1) only the suffixal part is used. In negative past forms, the subjunctive vs indicative distinction is neutralized in *-att*. Similarly, the distinction between a simple and a resultative perfect (cf. below), as well as between an intransitive and a transitive version of these forms, is neutralized in the form of the suffix *-yatt* in sf1/2. (The distinction between the two perfects is still maintained, however, by the additional suffix *-khak* ~ *-hak* in the simple perfect.) The infinitive is negated by the prefix *min-* and the converb *-sa*

TABLE 35.5 FINITE VERB INFLECTION

pf1	<i>ma?</i> -i	first person non-singular exclusive P		
	<i>ma-</i>	first person singular P		
	~ <i>mai-</i>	if directly followed by the stem		
	~ <i>ka-</i>	in forms with a second person A		
	<i>ka-</i>	first person non-singular inclusive P		
pf2	<i>ak-</i>	optative		
pf3	<i>N-</i>	third person A (used only if acting on a second person)		
	<i>N-</i>	third person non-singular S or A		
	~ <i>mi-</i> (in pf1 position)	in negative forms with third person S or		
	<i>N-</i>	negative (only in direct forms and only in combination with sf6 filler <i>-ni</i>)		
	~ <i>min-</i>	in infinitive forms		
sf1	<i>-yuk</i>	definitive non-past	collapsing sf1 and sf2:	
	<i>-yakt</i>	imperfective	<i>-kett</i>	inceptive
	~ <i>-yau</i> before <i>-t</i> ‘Non-Past’		<i>-hett</i>	temporary
	~ <i>-ya</i> before <i>-att</i> ‘Neg. Past’		<i>-kon</i>	spatially distributed
sf2	<i>-t</i>	‘non-past indicative’	<i>-kone</i>	temporary
	~ <i>-Ø</i> before C (but blocking intervocalic voicing of C)		<i>-ye</i>	inconsequential
	~ <i>-?</i> before nasals		~ <i>-se</i> in transitives	resultative perfect
			<i>-ŋa</i>	
	~ <i>-yuk</i> in [_o σ _—]		~ <i>-sa</i> in transitives	
	<i>-he</i>	past indicative	<i>-ŋatt</i>	perfect negative
	~ <i>-e</i> after /s/ and in C ^ø Vr _—			
	<i>-a</i>	past subjunctive		
	<i>-att</i>	past negative		
	<i>-a</i>	imperative		
	~ <i>-an</i> in plural affirmative forms			
sf3	<i>-na</i>	first person A and second person P		
	~ <i>-nan</i> before <i>-i</i> first or second person plural			
sf4	<i>-chi</i>	dual		
	<i>-i</i>	first or second person plural		
	~ <i>-ik</i> in imperatives (and pronouns)			
sf5	<i>-u</i>	third person P		
sf6	<i>-ni</i>	negative		
	~ <i>-n</i> after V			
sf7	<i>-ŋ</i>	first person singular A (used only if acting on a third person)		
	<i>-m</i>	first or second person plural or non-singular A (used only if acting on a third person)		
sf8	<i>-chi</i>	non-singular A or P		
sf9	<i>-ŋa</i>	first person exclusive S or A		
	~ <i>-ŋ</i> after <i>-ni</i> negation marker and in the environment V_C			
	<i>-kak</i>	second person		
	~ <i>-ka</i>	word-finally or before vowel		
	~ <i>-k</i> in the environment V_C in perfect and past subjunctive forms			
sf10	<i>-khak</i>	perfect (only in combination with sf1/2 filler <i>-ŋa ~ -sa</i>) or ‘nominalizer’		
	~ <i>hak</i> after V or nasals			

has a negative counterpart marked by *min-* (without a suffix). There is no negative supine, a paradigm gap which is compensated for by syntactically forcing main clause negation to target subordinate supine propositions ('NEG-transport'); cf. (38) in section 5.6.

3.3.3 Tense and mood

The past vs non-past opposition also covers a mood distinction between counterfactual and other situations. This use is often supported by the irrealis clitic = *phe ~ = be* (after sonorants), especially in conditionals:

- (2) a. *ŋka n-cha lis-e-ŋa*
 1SG[ABS] 2SG.POSS-child[-SG.ABS] be-PAST-[1SG]EX
 '[Suppose] I were your child.'
- b. *ŋka-na har-e-ŋ = be, kochu*
 1SG[ABS]=TOP bite-PAST[-3SGP]-1SGA=IRR dog[-SGABS]
lis-a-ŋ=naa.
 be-SUBJ.PAST-[1SG]EX=TOP
 'I'd have bitten him if I were a dog.'

The NON-PAST allomorph *-yuk* is used if suffixing *-t* would result in a monosyllabic prosodic word, e.g. *mai(₉ lu-yu)*, not **mai(₉ lu-t)* 's/he tells me'. The definitive non-past in *-yuk* always combines with *-t* and asserts a situation with certainty, e.g. *tup-yuk-t-u* 'he definitely understands it!' Like all non-past categories, reference can be present or future.

The other major mood distinction is between subjunctive and indicative. The subjunctive is marked only in the past. It is limited to nominalized and subordinate clauses (see section 3.5). There is no non-past subjunctive, but this gap is pragmatically filled by a zero-marked form, which is temporally unspecific. In independent sentences, the zero-form is mainly used for deontic requests (3a), adhortatives (3b) and warnings (3c):

- (3) a. *yot-pir-u-ŋ?*
 add-P.BEN-3P-1SGA
 'Shall I pour him some [more tea]?'
- b. *iŋa uŋ-ch-u!*
 beer[ABS]drink-[1]DU[INC S]-3P
 'Let's have some beer!'
- c. *cuj-ŋa si-chi-ga!*
 cold-ERG die-DU-2[S]
 'You may die of cold!'

For a dependent clause example, where the zero-form has a past tense value, see (41).

Another modal category of Belhare is the optative, which is marked by the prefix *ak-*. The marker is paradigmatically and functionally part of the agreement rather than the tense/mood system. It appears in pf2 position (Table 35.5) and signals that the realization of a situation is under someone else's rather than the subject's (S/As) control. This can (4a), but need not (4b) be the speaker (see Bickel 2000a for further discussion):

- (4) a. *ay-khat-ni!*
 OPT:NEG[-3SGS]-go-NEG
 'She shouldn't go!' (i.e. 'I don't want her to go.')

TABLE 35.6 AFFIRMATIVE (UPPER FORM IN EACH CELL) AND NEGATIVE (LOWER FORM) NON-PAST PARADIGM (*LUMA* 'TO TELL' AND *KHATMA* 'TO GO')

Intransitive						
	1SG	1NSG.EX	1NSG.INC	P	2DU	3SG
	1SG	1NSG.EX	1NSG.INC	2SG	2PL	3NSG
1SG				<i>luiñha</i>	<i>luiñnaci</i>	<i>luituj</i>
				<i>nluiñhan</i>	<i>nluiñnanchin</i>	<i>nluiñhij</i>
1PL.EX				<i>luiñnari</i>	<i>luitumma</i>	<i>nluiñhij</i>
				<i>nluiñnachina</i>	<i>nluiñnimma</i>	<i>nluiñnimma</i>
1DU.EX						
1PL.INCL						
1DU.INC						
A	2SG	<i>kahuika</i>	<i>mañlukika</i>			
		<i>kahuñiga</i>	<i>mañluiñiga</i>			
	2PL	<i>kahuítiga</i>	<i>mañluiñiga</i>			
		<i>kahuítiga</i>	<i>mañluitinga</i>			
		<i>kahuichiga</i>	<i>mañluitinga</i>			
		<i>kahuichiga</i>	<i>mañluitinga</i>			
	2DU					
	3SG	<i>mailayu</i>	<i>mañluyu</i>	<i>kaluyu</i>	<i>nluika nluiñiga</i>	<i>nluitiga nluitinga</i>
		<i>mailuñi</i>	<i>mañluñi</i>	<i>kaluñi</i>		<i>nluichiga</i>
		<i>mailuñi</i>	<i>mañluñi</i>	<i>kaluñi</i>		<i>nluichinga</i>
	3PL	<i>manluyu</i>	<i>mañluyu</i>	<i>kanhayu</i>		
		<i>manluñi</i>	<i>mañluñi</i>	<i>kanhuñi</i>		
		<i>mapluñi</i>	<i>mañluñi</i>	<i>kanhuichi</i>		
		<i>mapluñi</i>	<i>mañluñi</i>	<i>kanhuichi</i>		
	3DU					

- b. *ak-tupt-u-ga!*

OPT-understand-3P-2[SGA]

‘You are to understand it.’ (i.e. ‘they want you to understand it’, not in the sense of ‘please understand!’, for which the more specific imperative is used.)

In non-third person use, the optative is typically supported by the hearsay or reportative clitic =*phu* ~ =*bu*.

The imperative is marked by *-a* and, in plural affirmative forms, by *-an* (e.g. *lur-an-u-m* ‘you^p tell him!’). Negative imperatives are marked by the regular negation circumfix *n- . . . -n(i)*, but are additionally characterized, as noted above, by a special rule of nasal copying.

3.3.4 Aspect

There are five aspectual forms defining a scale of increasing markedness and specificity: (zero-marked) simple < imperfective < temporary/spatially distributed temporary < inceptive forms. The simple form often has a perfective value, but since it is maximally unmarked, it is also used for simple statements of events and states. The temporary differs from the imperfective in two respects: it entails a non-past tense value, and it denotes an ongoing situation together with its boundaries, e.g. *cokg-hett-u* ([3SGA]-do-TEMP-3P) ‘s/he is doing it right now and only now’. The spatially distributed temporary has the same aspectual value as the simple temporary, but it additionally indicates that a situation is spatially distributed. This can, but need not, imply motion: e.g. *thali khore wat chi-gon* ‘she is cleaning plates (*thali*) and cups (*khore*)’ can refer to a situation where someone is cleaning dishes that are scattered around him or her at a washing place. The missing past forms are compensated for by the *aktionsart* modifier *kond-* (which is cognate with the spatially distributed temporary aspect marker) (cf. section 4.2.4) combined with the imperfective, e.g. *wat chi-goī-yakt-he* ‘s/he was cleaning here and there’. From its competition with the two temporary aspects, the non-past imperfective generally implicates unlimited situations, which in turn suggests exaggeration; the form is therefore most common in reproachful and ironical statements. The inceptive aspect occurs in affirmative forms only with verbs denoting paths. The marker indicates that the motion has started and is still going on (e.g. *khat-ke* ‘s/he has set off’, whence ‘is going’). In the negative, the inceptive is compatible with all verbs and denotes that a situation has not yet come about (e.g. *n-cok-ket-ni* ‘s/he is not yet doing it’ or ‘s/he has not done it yet’).

3.3.5 Perfect and resultative perfect

Both types of perfect signal a noteworthy relationship between a past situation and the time of reference, but while the simple perfect focuses on the past as, say, the explanation for the present, the resultative present focuses on the results of the past. In line with this, the simple perfect only allows specification of the time of the event (5a), whereas the resultative perfect also optionally allows specification of the time of reference, i.e. of the resulting situation (5b):

- (5) a. *namniyy-etnahuŋ misen niu-s-u-ŋ-ŋ-ha.*
last.year-ABL know-TR.PERF-3[SG]P-1SGA-PERF
‘I have known him since last year.’

- b. *asen* *tai-ye.*
 yesterday [3SGS-]come-INTR.RES.PERF
 ‘She was here yesterday (having arrived).’
 or ‘S/he came yesterday and is now here.’

Both forms have transitive and intransitive variants. Unlike similar forms in other languages (cf. Nedjalkov 1988), the resultative perfect does not require any detransitivization: e.g. *nau-se-na* (ask.for-TR.RES.PERF-1[SG] > 2[SG]) ‘I have asked money from you (which I still owe you)’ has a stative meaning similar to what would correspond to *I have (the) money (that I) asked from you* in English, but is fully transitive.

3.3.6 Inconsequential

The inconsequential denies whatever one might expect as the consequence of a situation:

- (6) *piñ-sa* *un-gone-ya* *khat-ca-he.*
 run-CVB come.DOWN-INCONS-[1SG]EX [3SGS-]go-TEL-PAST
 ‘I ran downhill, but [the bus] already went off.’

Combined with the exclamatory question word *keko* (from Nep., lit. ‘of what’), the inconsequential often suggests that some action is pointless, e.g. *keko pi-gone-chi?* ([3SGA-]give-INCONS-[3]NSGP) ‘What’s the point in her/him giving it to them?’

3.4 Diathesis

Belhare has two types of (finite) passives, one with adversative and one with perfect value, as well as an object-downgrading construction that partially fulfils an antipassive function. The adversative passive is marked by the suffix *-khaca* and implies that the derived S-argument is negatively affected by the event (7a). The perfect passive is realized by the intransitive form of the perfect (7b). Neither form allows the overt expression of an agent:

- (7) a. (**Kancha*/**Kancha-ŋa*/**Kancha-lam*, etc.) *lu-khaca-he.*
 last.born[ABS]/last.born-ERG/last.born-MED [3SGS-]tell-ADVERSATIVE.PASS-PAST
 ‘S/he was scolded.’
- b. (**Kancha* /**Kancha-ŋa*, etc.) *Maiti-pa* *lui-ŋa-ha.*
 last.born[ABS]/last.born-ERG M.-father[ABS] [3SGS-]tell-INTR.PERF-PERF
 ‘Māiti’s father has been told.’

The object-downgrading construction is realized through intransitive agreement morphology and absolutive instead of ergative case, cf. (8a). The active counterpart of this is (8b):

- (8) a. (*i-na*) *wa* *khu²-yu.*
 REM-DEM[-SG.ABS] chicken[-ABS] [3SGS-]steal-NPAST
 ‘This [guy] steals chicken.’
- b. (*i-na-ŋa*) *wa* *khui²-t-u.*
 REM-DEM[SG]-ERG chicken[-SG.ABS] [3SGA-]steal-NPAST-3[SG]P
 ‘This [guy] will steal a/the chicken.’

While the object-downgrading construction is similar to an antipassive in that it promotes A to S, it is different in that the construction does not alter the grammatical relation of the P-argument (the object): P can be relativized on in internally headed relative constructions in the same way as the promoted S-argument, leading to ambiguity (relativization on A is impossible in this construction; cf. (27b) in section 5.4):

- (9) *tombhira wa sei²-sa-ha chitt-he-m.*
lynx[-SG.ABS] chicken [-ABS] [3sgS-]kill-TR.PERF-NOMZR find-PAST[-3SGP]-1PLA
'We found the lynx that had killed chicken.' or
'We found chicken killed by a/the lynx.'

However, the object-downgrading construction alters the categorial status of objects from the usual NP level to a bare noun. As such, the object cannot contain attributes, demonstratives or any marking that could imply a specific determiner value such as possessive or number: the constituent must have a generic kind reading. The categorial downgrading and the associated semantics suggest similarity to noun incorporation, but a downgraded object is not limited to the immediately preverbal position, and it can be modified by focus (e.g. in (8a), *wa=ro* 'chicken=ID') and topic (*wa=na* 'chicken=top') clitics (cf. Bickel 2004). Note that unlike with the perfect passive, the object-downgraded version of perfect forms does not substitute the transitive marker *-sa* by the intransitive marker *-ŋa* (cf. Table 35.5 and section 3.3); the object-downgrading construction affects only agreement and case morphology.

3.5 Nominalization, participles and non-finite forms

3.5.1 Nominalization

The suffix *-khak* (~ *-hak* after sonorants) nominalizes finite verb forms which head relative and complement clauses (cf. section 5.6). The nominalizer can also occur on independently used verb forms where it serves as a focus marker (Bickel 1999c). Scope can be on a constituent (10a) or on the entire proposition (10b):

- (10) a. *hale mand-u-ŋŋ-ha.*
before finish-3P-1sGA-NOMZR
'It's earlier that I finished!'
- b. *hamba Dhankuta khar-a-ŋŋ-ha,*
today Dh.[-LOC] go-SUBJ.PAST-[1SG]EX-NOMZR
rak-khar-e-ŋa, *ŋka!*
tired-TEL-PAST-[1SG]EX 1SG[ABS]
'It's that I went to Dhankuta today. [That's why] I am tired!'
(as an answer to an inquisitive look)

The nominalizer is also used to mark the attributive value of adjectives and demonstratives. In this environment, the nominalizer is always *-khak*, even after sonorants (e.g. *itii-kha* 'small-NOMZR' or *yu-kha* 'DEM:ACROSS-NOMZR'). An alternative means of nominalization is the enclitic article, which marks clauses, adjectives and demonstratives in attributive function (also cf. section 5.2).

3.5.2 Participles

Active participles are formed from verb roots by the prefix *ka-* and the natural gender-indicating suffix *-pa* (unmarked value) or *-ma* ('female'). *Ka*-participles refer to the S/A-argument of the verb, and allow clausal modification:

- (11) *asenle ka-pikg-a-ba* *ma?i-ja* *sappe*
 earlier ACTIVE.PART-fall-go.DOWN-PART person[SG]-ERG all[ABS]
mai-mat-pir-he.
 1SGP-[3SGA-]narrate-P.BEN-PAST
 'The guy who fell down recently, told me everything.'

Passive participles are rare and use one of the suffixes *-pilat*, *-ilan*, *-palak*, *-palat* or *-palak* with no known semantic distinction, e.g. *ten-bilat* 'one who got hit'. Passive participles have a perfect value and are typically used as NPs on their own, e.g. *inbilat* 'the things that one bought' or *thukpilat* 'cooked stuff'. When suffixed by the focalizer/nominalizer *-khak*, passive participles are typically used predicatively:

- (12) *Maiti-pa* *lu-bilat-kha,* *helo!* *iti* *bela*
 M.-father[SG.ABS] tell-pass.PART-NOMZR hey! this.much time
yy-hon-dai-?-ni!
 NEG[-3SGS-]appear-come-NPAST-NEG
 'It's that Maiti's father was told [to come], but what's that? He doesn't show up even now!'

3.5.3 Non-finite forms

The infinitive is non-finite with regard to tense/mood, but inflects for imperfective aspect (e.g. *khoj-yak-ma* 'to keep playing' vs *khoj-ma* 'to play') and number of the P-argument (e.g. *hitma-chi* 'to hit them' vs *hitma* 'to hit it/him/her', but not S-arguments, e.g. **imma-chi* 'their sleeping'). The form is used in non-finite complementation (section 5.6), but can also constitute independent sentences:

- (13) *ikaa=bu semba khat-ma kina mun dhup-ma?*
 why=HS night go-INF SEQ chat-INF
 'Why [should one] go and chat in the night? (they asked)'

Another (quite frequent) use of the infinitive involves topicalized verb repetition of the kind that is common throughout Tibeto-Burman and adjacent languages (e.g. Nepali; on Lahu cf. Matisoff 1973: 423). An example is *cama=na cayautu* 'as for eating, it is eating' in (41) later.

The three other non-finite forms, the supine (*-si*) and the affirmative (*-sa*) and negative (*miN-*) converb suffixes mark subordinate clauses and are discussed in section 5.6.

3.6 Clitics and clause-final particles

Belhare has various clitics modulating information structure (=*na* 'specific article', =*kha* 'specific article plural', =*khachi* 'specific article dual', =*na* 'topic', =(*e*)*tlo* 'restrictive focus', =*cha* 'additive focus', =*tok* 'identificational focus', =*ndo* 'counter-presuppositional focus', =(*k*)*olo* 'contrastive focus', =(*e*)*? wa* 'like') and evidentiality/status (e.g. =*phu*

$\sim =bu$ ‘hearsay’, ‘reportative’, $=phe \sim =be$ ‘irrealis’, $=nno$ ‘confirmative’). Clitics are unrestricted as to the part of speech they attach to (cf. e.g. $=bu$ on an adverb in (13) and $=phu$ on a noun in (44)), with two exceptions: (i) the article is limited to spatial demonstratives, adjectives and verbs; (ii) after finite verb forms, the topic marker $=na$ is replaced by the subordinator $=naa \sim =naya$ (cf. section 5.6; for the historical reasons for this, see Bickel 1999b).

Clause-final particles are the interrogative particle *i*, the ‘obvious’ marker *mu* (signaling clear and obvious evidence), the mirative marker *raicha* (from Nep. *rahecha*) and also include the conjunctions *ki* ~ *kina* ~ *kinahuy(go)* ‘and’, *huycha* ‘whereas’, ‘but’, and *muhuygo* ~ *huydo* ‘since’, ‘because’. Unlike clitics, clause-final particles bear stress and are unaffected by the voicing and prothesis rules that apply inside the phonological word, cf. *taheya ki* (not **tahey=ggi*) ‘I came and’ or *kharega i* (**kharegak=ggi*) ‘did you go?’

4 DERIVATIONAL MORPHOLOGY AND COMPOUNDING

This section discusses lexically restricted but productive morphology. Category-changing morphology that is lexically unrestricted is treated in section 3.5.

4.1 Nominals and adverbs

The most important derivational device in nominals is teknonymy. The Belhare refer to each other by the name of the first-born child, suffixed by *-pa* for men and elaborated by *u-ma* ‘3SG.POSS-mother’ for women. An example appeared in (7) earlier. Natural gender distinction like this is otherwise limited to *ka*-participles (cf. section 3.5) and a few kinship terms such as *eba* ‘father’s elder brother’ and *ema* ‘father’s elder brother’s wife’. Apart from this, the only productive segmental suffix is the diminutive *-cilet*, as in, e.g. *phak-cilet* ‘piglet’ or *khim-cilet* ‘small houses’. Reduplication is used for indicating diminutives with a few demonstrative and adverbial roots (e.g. *mopmo* ‘a bit further down’); with colour adjectives, it has an intensifying function, cf. e.g. *makkhorok-makkhorok=ma* ‘pitch black’.

Nominal compounding involves a small set of generic nouns such as *khim* ‘house’ (e.g. *ma?i khim* ‘a foreigner’s house’, *may khim* ‘god-house’), *sa* ‘meat, flesh’ (e.g. *phak sa* ‘pork’, *ya sa* ‘fish’) or *tay* ‘tree’, ‘plant’ (e.g. *siŋ tay* ‘firewood tree’, but *ambiu u-tay* ‘mango tree’, lit. ‘mango its-tree’). Note that phonologically, nominal compounds (unlike verbal compounds; cf. below) are separate words: there is no voicing or prothesis at compound boundaries.

Adverbs can be derived from adjectives by the suffix *-mu* (*ukurik-mu* ‘in a bent manner’, *phabeley-mu* ‘currently red’), but this formation is rare and often completely lexicalized (e.g. *rommu* ‘together’, but no **rom*, **rom=na*, etc.). Other adverbs primarily include temporal expressions, in particular a ‘year’ and a ‘day’ series counting up to four times away from the present, e.g. *onumba* ‘in three days’, *khonumba* ‘in four days’, *añbu* ‘three days ago’.

4.2 Verbs

The inherited derivational machinery of Tibeto-Burman coronal transitivizers is no longer productive in Belhare, although there are many pairs like *pok-* ‘rise’ vs *phok-* ‘raise’ (from **s-pok*), *ta-* ‘come’ vs *tat-* ‘bring’, *hon-d* ‘appear’ vs *hot-t* ‘drive out’, or *lik-* ‘enter’

vs *lijs-* ‘insert’. However, the coronal suffixes are still treated as transparent entities by the morphology: they are deleted before consonants, while homophonous root components are retained, e.g. *hi-t-* ‘be able; finish’, where /t/ is a suffix, surfaces as *hima* in the infinitive, whereas *hit-* ‘watch’ surfaces as *hitma* because here /t/ is part of the root and not an etymological suffix.

Apart from one suffix (*-ap*), which serves to integrate loanwords (e.g. *har-ap lima* ‘to lose’, ‘be defeated’ from Nep. *hārānu* ‘id.’ and the auxiliary *lima* ‘to be (come)’), the verbal lexicon can be productively expanded only through compounding, which is often lexicalized or grammaticalized.

Compounds often reflect stereotyped ways of performing activities such as *e?wa lu-gat-ma* ‘to bathe and come up (from the river)’, i.e. ‘to go bathing’. Note that the sequence of verbs is not necessarily iconic, e.g. *chandahe* ‘it (the snake) came and curled up’ in (44) later. Compound verbs agree in transitivity with regard to the roots chosen:

- (14) *cama ca-apt-he-i-ga i?* (**ca-ab-he-i-ga*)
 food[-ABS] eat-bring.ACROSS (TR)-PAST-PL-2 Q eat-come.ACROSS (INTR)-PAST-2PL-2
 ‘Did you come over here after you had your meal?’ (Nep. *khānā khāera āyau?*)

While compound verbs usually form a single morphological word, they are occasionally split, e.g. *rut m-pir-he* ‘collect 3NSGA-give-PAST[-3SGP]’~*n-rut-pir-he* ‘3NSGA-collect-give-PAST [-3SGP]’, both ‘they collected it and gave it to him’. Compounds with *metma* ‘to cause’ and *tetma* ‘to [verb] acceptably’ are always split, e.g. *hit mai-mett-he* (look.at 1SGP [-3SGA]-cause-PAST) ‘s/he showed it to me’. Syntactically, however, these units are inseparable verb complexes, cf. the ungrammaticality of, say, positioning *hit* in a post-clausal afterthought position:movement: **maimette, hit*. The most important grammaticalized compounds include those highlighted in the following sections.

4.2.1 Benefactives

When added to another verb, the root *pir-* ‘give’ expresses benefactive or malefactive effects on the P-argument. Similar notions are also encoded by ditransitive versions of environmental space markers, e.g. by *phett-* ‘to do something on an across trajectory for the (dis)advantage of P’ (cf. below). Benefactive or malefactive affection of the S or A argument is expressed by *ca-* (intransitive) and *cand-* (transitive). This typically implies a middle voice or restrictive value:

- (15) a. *a-lan tok-cand-he-y.*
 1SG.POSS-leg[-SG.ABS] hit.against-S/A.BEN -PAST[-3SGP]-1sgA
 ‘I hit my leg against something.’
- b. *u-nucha khoy-ca-he.*
 3SG.POSS-younger.sibling[-SG.ABS] [3SGS-]play-S/A.BEN-PAST
 ‘His younger sister played by herself/alone.’

4.2.2 Reflexives and reciprocals

Whereas other Kiranti languages mark reflexives by an inflectional affix in the agreement system, Belhare has reanalysed the cognate of this affix as a verb stem that behaves morphologically like an auxiliary in compound verbs: *-chind* (cf. the reflexive agreement

suffixes in Bahing *-si*, Dumi *-(n)si*, Hayu *-ntsi* ~ *-ntse* ~ *-na*, Khaling *-si*, Limbu *-siŋ* ~ *-ne*, etc.); e.g. present *ten-chiʔ-yu* ‘s/he hits himself/herself’ vs past *ten-chind-he* ‘s/he hit himself/herself’. Not only transitives but also other multivalent predicates can be reflexivized, e.g. *khi-chind-he* ‘s/he was angry with himself/herself’ from *X-lok* (‘comit’) *khima* ‘to be angry with X, to quarrel with X’. The reflexive also competes with the S/A-benefactive in covering middle voice:

- (16) *i-n-et=to kin-chind-he*
DIST-DEM-LOC=ID [3SGA]-place-REFL-PAST[-3SGP]
'He put [the basket] right there for himself (so he could use it afterwards).'

Occasionally, the loan pronoun *appi* (< Nep. *aphno* ‘refl.’) is added in reflexive constructions.

Reciprocity is expressed by V-*ka*-V or V-*kabila* with the auxiliary *cama*. The construction with *-kabila* incorporates a benefactive notion, cf. *hankahan* or *hankabila cama* ‘to distribute among each other’ (from *hanma* ‘to distribute’), but *tha tokkatok* vs **tha tokkabila cama* ‘to know each other’ (from *tha tokma* ‘to know’).

4.2.3 Telicity

Both atelic and inherently telic verbs can be marked as telic. The regular compounds involve *-khat-* (~ *-hat-* after sonorants, e.g. *tuk-khar-e* hurt-TEL-PAST ‘it hurt’, *chomm-haiʔ-ŋaha* be.crazy-TEL-PERF ‘he has become crazy’) with intransitives, *-tend-* with transitives (e.g. *pum-deʔ-se* clinch-TEL-RES.PERF ‘he has clenched [his fist]’) and exceptional forms after *khat-* ‘go’ and *khatt-* ‘take’ (*khat-ca-* ‘go-TEL-’ and *khat-lott-* ‘take-TEL-’). Telicity is also entailed by the environmental space markers. These auxiliaries, e.g. *-phett* as described above, denote UP, DOWN and ACROSS- trajectories in the same way as the corresponding cases and demonstrative roots (section 3.1).

4.2.4 Other *aktionsarten*

Ambulative (-*kond-* ‘at various places, on the go’; cf. section 3.3), relinquitive (-*ye-* ‘to leave P where the action takes place’), perseverative (-*yuy-* ‘keep doing’), accelerative (-*itt-* ~ *is-* ‘quickly’, ‘just’), diminutive (-*khey-* ~ *-khat-* ‘only’, ‘just’, ‘a little’, ‘quickly’), errative (-*khey-* ~ *-khat-* ‘mistakenly’), restrictive (-*tis-* ‘do nothing else but VERB’), potential (-*sit-* ~ *-si-*).

5 SYNTAX

5.1 Clause structure

Clausal phrase structure is flat and, in line with other South Asian languages, word order generally reflects an increase of informational newsworthiness and focality, ending with the predicate. The prepredicate position often attracts question words and focus-marked constituents. There is an intonationally marked post-clausal afterthought position (see (19), (25b) and (44) later), which bans new or focal referents, and so does not accommodate, for instance, true question words (Bickel 2004). NPs are all optional in the clause and indeed typically absent (Bickel 2003a; Stoll and Bickel 2009).

The predicate usually consists of a verb (or verb complex). Nominal predicates are either marked by agreement morphology (section 3.1), a copula (*yuyya-* in the non-past, *yuyyase-* in the past; or *lis-* ‘be(come)’) or the identificational focus marker *=tok* (e.g. *un=do* ‘it is/was him/her’; cf. the use of *=tok* as a focus clitic in (17) below).

Clause structure is non-configurational insofar as both word order and information structure are syntactically independent of grammatical relations (on which see section 5.4). Therefore, as long as they are focal enough, pronouns can be co-referent with clause-mate antecedents that bear a lower grammatical relation. Focality can be increased by word order and/or focus particles:

- (17) *Mohan-naha_i u-kitab un-na=ro_{i,j} chitt-he.*
 M.-GEN 3SG.POSS-book[ABS] 3[SG]-ERG=ID [3SGA-]find-PAST[3SGP]
 ‘HE*_{i,j} found Mohan_i’s book.’

Apart from the newsworthiness and focality cline, information structure is largely unconstrained. In particular, the ‘potential focus domain’ (Van Valin and LaPolla 1997) is virtually unlimited in the clause and includes even NP sub-constituents and subordinate clauses. Thus, question words are possible even in relative (18a) or adverbial (18b) clauses. Note that these are genuine questions in Belhare, not echo questions:

- (18) a. *sa-a thuu-s-u = na ija chept-he-ga?*
 who[SG]-ERG [3SGA-]cook-TR.PERF-3P=ART beer[ABS] taste-PAST[-3P]-2[SGA]
 ‘*Who did you try the beer that ___ made?’
- b. *sa-a lab-yukt-u-lo m-phig-he?*
 who[SG]-ERG [3SGA-]grip-keep.stable[-SUBJ.PT]-3[SG]P-COM 3NSGA-pick-[3SGP]-PAST
 ‘Who kept down [the branch] when they picked [fruits]? (literally: ‘*Who did they pick [fruits] when ___ held down [the branch]?’)

Instead of *saa* ‘who’, a focused NP like *a-na-ŋa = ro* ‘my-elder.sister-ERG=ID’ would be equally grammatical in these cases.

This contrasts with right-dislocation (afterthought) constructions, which cannot be co-indexed with arguments inside embedded or otherwise subordinated clauses:

- (19) a. *a-ni-a thuu-s-u=na ija chept-he-ŋ.*
 1SG.POSS-father.sister-[SG]ERG [3SGA-]cook-TR.PERF-3[SG]P=ART beer[ABS] taste-PAST[-3SGP]-1SGA
 ‘I tried the beer that my aunt brewed.’
- b. **thuu-s-u-na ija chept-he-ŋ, a-ni-ŋa.*
 [3SGA-]cook-TR.PERF-3[SG]P=ART beer[ABS] taste-past[-3SGP]-1SGA 1SG.POSS-father.sister-[SG]ERG
 Intended: ‘I tried the beer she has prepared, my aunt that is.’

5.2 The noun phrase

NPs have a similar structure as clauses: the head is final, there is a post-phrasal adjunct position, and pre-head order is syntactically free:

- (20) a. [NP *sik-kira phabeley=ma phuŋ] tar-he-ŋ.*
two-NHUM red=COLOUR.ART flower[ABS] bring-PAST[-3P]-1SGA
- b. [NP *phabeley = ma sik-kira phuŋ] tar-he-ŋ.*
red = COLOUR.ART two-NHUM flower[ABS] bring-PAST[-3p]-1SGA
- c. [NP [NP *sik-kira phuŋ] phabeley = ma]*
two-NHUM flower[ABS] red = COLOUR.ART
tar-he-ŋ.
bring-PAST[-3P]-1SGA
'I brought two red flowers.'

The post-NP position accommodates all kinds of attributes, but it is limited to a single element (unlike the post-clausal position, which allows more elements).

Attributes inside an NP are marked as such by the genitive case, an article, a nominalizer, or, in the case of numerals, a classifier. Genitival attribution can combine with possessive marking on the head, in which case the possessor has specific reference. Note that possessive marking requires genitive marking (except in compound expressions), but not vice versa:

- (21) a. *maʔi-ha khim*
person[SG]-GEN house[ABS]
'someone (foreign)'s house'
- b. *maʔi-ha u-khim*
person[SG]-GEN 3SG.POSS-house[ABS]
'the (known) person's house'
- c. **maʔi u-khim*
person[SG.ABS] 3SG.POSS-house[ABS]

The semantic effect of agreement markers is that the possessor has specific reference. Without agreement, as in (21a), the possessor can only be interpreted as generic. Therefore, it would be impossible to add a demonstrative that enforces specific reference (**na maʔiha khim*, with the demonstrative *na*)

Articles and the nominalizer are complementary means of attribution (and some other Kiranti languages merge them formally; cf. Bickel 1999c): the use of articles indicates specific reference of the head noun, while the nominalizer does not entail specificity. Possessed nouns have inherently specific reference and obligatorily require article-marked adjectives:

- (22) *un-naha eiʔ=na /*ei-kha khim*
3[SG]-GEN big=ART big-NOMZR house[ABS]
'his/her big house'

Articles and the nominalizer combine not only with adjectives and verbs (as for example in (18a) earlier), but also with environmental space demonstratives, e.g. *yu=na khim* 'the house over there' vs nominalizer-marked *yu-kha khim* 'a /the house over there'.

One adjective, *uchōñat* ‘new’, can be used as an attribute inside an NP without further marking (*uchōñat khim* or *khim uchōñat* ‘a /the new house’) and appears to go back to a possessive construction meaning ‘its-newness’. Such constructions are found with a few other concepts: e.g. *yatlابu u-soy* [banana 3POSS-ripe] ‘the/a ripe banana’.

Classifiers are obligatory with numerals and draw a basic distinction between humans (-*pay*) and non-humans (-*kira*); more specific classifiers are borrowed from Nepali. The three surviving native numerals (*i-* ‘one’, *siC-* ‘two’, *sum-* ‘three’) require classifiers even in counting. With human head nouns, numerals require number agreement:

- (23) *sip-pay mañi-chi / *mañi*
 two-HUM person-NSG[ABS] person[SG.ABS]
 ‘two people’

With other nouns, number agreement is optional; cf. the examples in (20) earlier.

The NP head is always optional, whence demonstratives and attributes can function as NPs on their own. Indeed, demonstratives often substitute for third person pronouns in discourse.

5.3 Predicate agreement

Like in other Tibeto-Burman languages (Bickel 2000b), Belhare verb agreement is not necessarily identificational (24a) but also allows appositional (24b) and partitional (24c) interpretations of the relationship between the agreement trigger and its target:

- (24) a. *kaepma-chi ñ-khar-e.*
 girl-NSG[ABS] 3NSGS-go-PAST
 ‘The girls went.’ (*ñ-* ‘they’ = *kaepmachi* ‘girls’)
- b. *masiy=cha siy tayy-e thañ-2-ña.*
 old.woman=ADD fire.wood tree-LOC go.UP-NPAST-[1SG]EX
 ‘Even as an old woman I climb trees.’ (-*ña* ‘I’ as *masiy* ‘old woman’)
- c. *sa-ti khar-e-i-ga?*
 who-SG.ABS go-PAST-2PL[S]-2
 ‘Who of you went?’ (*sati* ‘who’ **of** *-iga* ‘you (pl.)’)

This also applies to agreement markers suffixed to predicate nominals (cf. section 3.1).

The semantic flexibility of agreement and the optionality of NPs notwithstanding, it is important to realize that the Belhare agreement system reflects grammatical agreement rather than incorporated or cliticized pronouns: agreement-triggering NPs are true arguments. This is evidenced by the fact that agreement-triggering NPs differ in syntactic possibilities from non-argument NPs: it is possible to extract sub-constituents from agreement-triggering NPs and put them into the post-clausal afterthought position, as shown in (25) for an A argument. By contrast, such extraction is not possible from NPs which are not in argument function (26):

- (25) a. *[a-tak-ñaha u-phu-ña] mai-lur-he.*
 1SG.POSS-friend[SG]-GEN 3SG.POSS-elder.brother[SG]-ERG 1SGP-[3SGA-]tell-PAST
- b. *[u-phu-ña] mai-lur-he, a-tak-ñaha*
 3SG.POSS-elder.brother[SG]-ERG 1SGP-[3SGA-]tell-PAST 1SG.POSS-friend[SG]-GEN
 ‘My friend’s elder brother told me.’

-
- (26) a. [a-tak-*yaha* *u-khimm-e*] *n-tupth-he*.
 1SG.POSS-friend[SG]-GEN 3SG.POSS-house-LOC 3NSGS-meet-PAST
- b. *[*u-khimm-e*] *n-tupth-he*, *a-tak-*yaha**
 3SG.POSS-house-LOC 3NSGS-meet-PAST 1SG.POSS -friend[SG]-GEN
 ‘They met at my friend’s place.’

5.4 Grammatical relations

5.4.1 Primary grammatical relations

While many constructions, most notably coordination and chaining, are syntactically unconstrained as to argument coreference, some constructions are subject to rigid grammatical pivot constraints. Active participles, supines and conversbs (see section 5.6) are subject to accusative-style syntax and require sharing of their S/A-argument with the matrix clause. Two constructions show syntactic ergativity. First, head-internal relatives allow relativization on S/P-arguments, but not on A-arguments (pre-nominal attributes relativize on virtually all arguments; see Bickel 2004):

- (27) a. *ma?* *khiu-2=na* *misen niu-t-u-ga* *i?*
 person[SG.ABS]_s [3SGS-]quarrel-NPAST=ART know-NPAST-3[SG]P-2[sgA] Q
 ‘Do you know the person who is quarrelling?’
- b. *tombhira-*ya** *wa* *sei?-s-u=na* *chitt-he-m.*
 lynx[SG]-ERG_A chicken[SG.ABS]_p [3SGA-]kill-TR.PERF-3[SG]P=ART find-PAST
 [-3SGP]-1PLA
 ‘We found the chicken that the lynx had killed.’
Impossible: ‘We found the lynx that had killed the chicken.’

Second, non-finite complement constructions (section 5.6) involve matrix control of the lower S or P:

- (28) a. [$\emptyset_{i,*}$ *khoj-ma*] *nui-ka_i*.
 [ABS]_s play-INF may-NPAST:2[sg]s
 ‘You may play.’
- b. [$\emptyset_{*,ij}$ $\emptyset_{i,*}$ *lu-ma*] *nui-ka_i*.
 [ERG]_A [ABS]_p tell-INF may-NPAST:2[sg]s
 ‘You may be told.’ or ‘S/he may tell you.’ *Impossible*: ‘You may tell him/her.’

Some matrix verbs are inflected transitively, in which case the pivot is controlled by the P-inflection. Note that the matrix A-inflection does not necessarily control the reference of the embedded A-argument, although there is often coreference for pragmatic reasons:

- (29) a. *khali* [$\emptyset_{*,i,j,k}$ $\emptyset_{i,*j,*k}$ *set-ma*] *ka_i-*y_j*-koy_u*.
 only [ERG]_A [ABS]_p kill-INF 1INCP-3NSGA-want-NPAST
 ‘They just want us to get drunk.’ (*lit.* ‘They want us to get killed by [the beer].’)
- b. [$\emptyset_{*,i,j,k}$ $\emptyset_{i,*j,*k}$ *ten-ma*] *ma-*y_j*-na-rend-he*.
 [ERG]_A [ABS]_p beat-INF 1SGP-3NSGA-stop-TEL-PAST
 ‘They stopped beating me.’ or ‘They stopped x from beating me.’

5.4.2 Objects

With ditransitives, it is mostly the goal rather than the theme argument that triggers P-agreement, see, e.g. (33) later. The same pattern is found in the pivot of non-finite complement constructions. Unlike ditransitive goals, themes cannot be controlled by the matrix inflection:

- (30) [Ø_{i,*j} *kitap-chi_{i,j}* *pi-ma*] *ŋ_i-khe-yu.*
 [ABS]Goal book-NSG[ABS]Theme give-INF 3NSGS-must-NPAST
 'They (*s/he) must be given books.'

With verbs like *pima* ‘to give’, themes and goals are not differentiated by case-marking, where all objects are in the absolute, or by the pivot of internal-head relatives, where all objects are part of the pivot.

Ditransitives with oblique goals (e.g. locations) or oblique themes (e.g. instrument) show different behaviour in these regards, with considerable variation across individual lexical items.

5.5 Experiencer constructions

Belhare features both experiencer-as-goal constructions, corresponding to South Asian-style dative subjects (but of course using the Belhare goal marker, which is an absolute), and experiencer-as-possessor constructions, corresponding to South-East Asian-style psycho-collocations (Matisoff 1986):

- (31) a. *ŋka*caleppa *khikt-he.*
 1SG[ABS] bread[ABS] [3s-]taste.bitter-PAST
 'The bread tastes bitter to me.' (lit. 'To-me the-bread tasted-bitter.')
- b. *a-niūa*ta-he.
 1SG.POSS-spirits[ABS] [3S-]come-PAST
 'I am happy.' (lit. 'My-spirits came.')

Syntactically, it is always the experiencer that counts as s/a-argument; it can therefore be the pivot of active participles or converbs, but not of non-finite complement constructions (see Bickel 2004, 2006 for detailed analyses), e.g.:

- (32) a. *caleppa* *ka-khik-pa*
 bread[ABS]_P ACTIVE.PART-taste.bitter-PART
 'one to whom the bread tastes bitter'
- b. [(*ŋka_{i,j}*) *caleppa_{i,j}* *khik-ma*] *ŋna-rend-he*
 1SG[ABS]_A bread[ABS]_P taste.bitter-INF [3SGS-]stop-TEL-PAST
 / **mai_i-ŋna-rend-he.*
 1SGP- [3SGA-]stop-PAST
 'The bread stopped being bitter to me.'

5.6 Clause linkage

5.6.1 Complement and relative clauses

Both these clause types are marked by the nominalizer *-khak* ~ *-hak* and, in the past tense, both are limited to subjunctive mood. Compare the following relative construction with a complement clause as in (35):

-
- (33) *dahi* *ka-khui?*-*sa-ha* *ma?*
 yoghurt[-SG.ABS] 1INC.P-[3SGA]-carry-TR.PERF-NOMZR person[-SG.ABS]
he-llay *khar-e?*
 wh-DIR [3SGS]-go-PAST
 ‘Where did the guy go who brought us the yoghurt?’

Relative constructions are prenominal, as in (33), or, less frequently, internally headed, as in (9) earlier. Instead of the nominalizer, relative clauses (but not complement clauses) can also be marked by the article; see (18a) for a prenominal and (27) for internally headed examples. Participial constructions, as in (11) and (32a), are a non-finite alternative to relative constructions, but they are limited to prenominal and headless positions.

Unlike other languages of Nepal (Ebert 1986), Belhare makes a rigid distinction between direct and indirect speech. Direct quotation is often unmarked, but is sometimes indicated by the converb *ceksa* ‘saying’ or the Nep. loanword *bhanera* ‘id.’:

- (34) *'hena bela tai-s-u-k-kha?'* (*cek-sa*) *kisi mai-lapt-he.*
 which time reach-TR.PERF-3P-2[SGA]-PERF say-CVB 1SGP-[3SGA]-ask-PAST
 ‘S/he asked me, “When did you arrive ?”’

Indirect quotation involves complementation by means of *-khak* ~ *-hak* and triggers deictic shift:

- (35) *tas khɔŋs-a-ŋŋi-ha* **i,j,ma?*-*yakt-he.*
 card play-SUBJ.PAST-[1SG]EX-NOMZR [3SGS]-Inarrate-IMPFV-PAST
 ‘S/he was saying that I played cards.’ (not: “I, played cards”, s/he, was saying.)

Non-finite complementation involves infinitives governed by various modal, aspectual and other semantically light matrix verbs. As noted in section 5.4, the matrix registers the S or P argument of the lower clause, but three verbs can also be used without matrix-coding (viz. *khes-* ‘must’, *kond-* ‘want’, *mitt-* ‘intend’, ‘think’). Compare the matrix-coded examples in (36a) with the referentially uncontrolled construction in (36b):

- (36) a. $[\emptyset_i \quad \emptyset_j \quad lu\text{-}ma]$ $i,j,k koī\text{-}t\text{-}u *_{i,j,*k}$.
 [ERG]_A [ABS]_P tell-INF [3SGA]-want-NPAST-3[SG]P
 ‘S/he wants x to tell him/her.’
- b. $[\emptyset_i \quad \emptyset_j \quad lu\text{-}ma]$ $i,j,k koī\text{-}yu.$
 [ERG]_A [ABS]_P tell-INF [3SGS]-want-NPAST
 ‘S/he wants that someone tells someone.’

Despite this tight interlacing of argument coding, the infinitive is a syntactically separate constituent and can therefore be extracted into the postverbal afterthought position, e.g. *koītu*, *luma* as an alternative to (36a). One matrix verb, *hima* ‘to be able to V’, ‘to know how to V’ requires a form V-*a* (not attested elsewhere) instead of the regular infinitive.

5.6.2 Converbs, supines and comitative clauses

Subordinate clauses headed by conversbs, supines or comitative-marked verbs are all embedded within the relational periphery of the matrix clause:

- (37) *a-tak-ya* *[ne-e yuy-sa] sappai*
 1SG.POSS-friend[SG]-ERG DEM-LOC sit-CVB all[ABS]

mai-mat-pir-he.

1SGP-[3SGA]-narrate-P.BEN-PAST

‘My friend told me everything while sitting here.’

Embedded clauses attract the scope of matrix negation (‘NEG-transport’), i.e. they cannot have an affirmative value if the matrix is marked as negative, e.g.:

- (38) *wa-si y-khatd-att-i-n-na*
 stroll-SUP NEG-go-PAST-1PL[S]-NEG-ex
 ‘We didn’t go for a stroll.’ (i.e. ‘We went but not for strolling.’ or ‘We neither went nor strolled.’)

In both these regards, embedded clauses contrast with ad-sententially subordinated clauses, on which see below.

Converb clauses typically encode manner or accompanying circumstances (cf. (37)). In repeated form, they are also used for alternating activities, often supported by a generic verb in auxiliary function (e.g. *metma* ‘to do’):

- (39) *kop-sa yuj-sa kop-sa yuj-sa mes-sa khatt-he.*
 pick.up-CVB put-CVB pick.up-CVB put-CVB do-CVB [3sgA-]take-PAST[3P]
 ‘He took them away by picking them up and putting them one by one [into a basket].’

The negative counterpart of *-sa* is marked by *min-*, which is in most cases supported by the locative case ending *-pak*, e.g. *min-cek(-pa)* ‘without saying anything’.

Supines (as illustrated by (38)) occur only with matrix predicates that allow the construal of a purpose notion, typically, but not exclusively, verbs of motion. The choice of controller depends on the semantics of the matrix verb (e.g. A with *tatma* ‘to bring’ but P with *luma* ‘to tell’).

Comitative clauses are finite and are marked by the comitative case suffix, as illustrated for example by (18b). They are restricted to subjunctive mood, and are subject to *consecutio temporum*. Depending on aspect and tense choice, comitative clauses encode a variety of relations, from ‘until’ to ‘when’.

5.6.3 Ad-sentential subordination

Ad-sententially subordinated clauses are marked by the clausal topic marker *=naa* ~ *=naya* (which is etymologically related to the ergative case; see Bickel 1999b) or by one of the phonologically independent, clause-final conjunctions *muhunyo* ~ *huydo* ‘since’ and *huycha* ‘whereas’, ‘although’, ‘but’ (which are all based on the ablative *-huy*). Topic clauses typically present a situational framework for a stretch of discourse and mostly translate as ‘if’ or ‘when’, depending on aspect and mood choice (see (2b) for an example). Occasionally *=naa* ~ *=naya* is replaced by the Nep. loanword *bhane* ‘topic’. Concessives and indefinite conditionals can also be signalled by the additive focus marker *=cha* ‘also’, ‘even’, cliticized directly to a finite verb in the subjunctive or zero-form (e.g. *khat-na=cha* ‘go-[1SG]EX[S]=ADD’, i.e. ‘even if I go’).

5.6.4 Chaining and coordination

Symmetrical coordination of clauses is different from nominal coordination (marked by comitatives: N-*lok* N-*lok*) and involves the additive focus clitic *=cha*:

- (40) *ŋka=na ten-he-ŋ=cha pheyd-he-ŋ=cha.*
 1SG[ABS] = TOP hit-PAST[-3SGP]-1SGA=ADD drive.off-PAST [-3SGP]-1SGA=ADD
 'I hit him and drove him off.'

More common, however, is asymmetrical chaining marked by the clause-final conjunction *ki*, which can be expanded to *kina*, *kinahuy* or *kinahungo*. So far, no semantic correlate of these different forms could be established, but it is likely to involve conceptual distance between events. Unlike subordinate clauses and like independent sentences, chaining allows indicative mood, but unlike independent sentences, chaining also allows the zero-form subjunctive with a modally neutral value (in contrast to examples like (3) in section 3.3); the past subjunctive is banned, however. *Consecutio temporum* applies, but note that the zero-form has no intrinsic tense value:

- (41) *khimm-e n-ta-ch-u (*ntas-e-chu) ki mun n-dhup-chi.*
 house-LOC 3NSGA-reach-DU[A]-3P (-PAST-) SEQ 3NSGS-chat[NPAST]-DU
 'They will reach home and chat.' or 'When they reach home, they'll chat.'
 or 'They reached home and now they will chat.'

There is no requirement of order iconicity, and *ki*-marked clauses can be postponed:

- (42) *ca-ma=na ca-yau-t-u tarə he-lley ley kina?*
 eat-INF=TOP [3SGA]-eat-IMPFV-NPAST-3[SG]P but where-DIR [3SGS]-direct SEQ
 'It (the cow) is eating, but after having turned towards which direction?'

Unlike with subordinate clauses, illocutionary scope may (but need not) extend through both conjuncts:

- (43) *khar-e ki jutta yy-in-ghutt-he-ga i?*
 [3SGS-]go-PAST SEQ shoes[ABS] 3[SG]A-buy-bring.for-PAST-2[SGA] Q
 'Did she go and buy you shoes?' or 'Did she buy you shoes when she went?'

Note that chaining is not subject to any coreference constraint. Zero anaphora can be controlled by any NP in any grammatical relation or, indeed, by referents from the wider context:

- (44) *chokt-he ki, n-celi-ŋa,*
 [3sgA-]point.with.finger-PAST[3sgP] seq 2SG.POSS-unmarried.agnatic.girl-ERG
doko-ep=phu chay-da-he.
 basket-LOC=HS [3SGS-]curl.up-come-PAST
 'When your *celi* pointed with her finger [at him_i], [Rainbow_i] came and curled up in [her] basket.'

Independent sentences can be linked by demonstratives in the ablative (*inetnahuy* 'after that') or the Nepali loans *ani* 'and then' and *pheri* 'again, and then'. For contrastive linkage, Nep. *tarə* 'but' is often used; the native strategy involves either the inconsequential (section 3.3) or subordination by *huycha* 'although', 'whereas', 'but'.

ADDITIONAL ABBREVIATIONS

- | | |
|-----|------------------------|
| ADD | additive focus |
| HUM | human |
| ID | identificational focus |
| IRR | irrealis |

MED	mediative case
NHUM	non-human
pf	prefix slot
SEQ	sequential
sf	suffix slot
SUP	supine
TEL	telic
Σ	stem
'>	designates a transitive relationship between an A (left side of >) and a P (right side).

Elements in square brackets are zero-marked, i.e. entailed by obligatory paradigmatic opposition (e.g. singular vs non-singular with animate referents).

FURTHER READING

Phonology and especially the prosody–morphology interaction is treated in Bickel (1998, 2003b). Agreement morphology is analysed in Bickel (1995a). The tense/aspect/mood system is the topic of a monograph (Bickel 1996), which also includes a full set of verb paradigms. For clausal syntax, see especially Bickel (1999c, 2004, 2006), Bickel *et al.* (2010) and Bickel and Gaenszle (2015), and for clause linkage, Bickel (1993, 1995b, 1999b, 1999c). Spatial deixis is discussed in detail in Bickel (1997, 2000a). Ethnographic information is given in Bickel (1999a, 2000c) and Bickel and Gaenszle (2015) and studies of discourse in Bickel (2003a) and Stoll and Bickel (2009). A dictionary is available in electronic form at www.ivs.uzh.ch/bickel-files/belhare_dictionary.html (accessed 18 August 2016).

REFERENCES

- Bickel, B. 1993. Belhare subordination and the theory of topic, in K.H. Ebert (ed.) *Studies in Clause Linkage*, 23–55. Zürich: ASAS Press.
- 1995a. In the vestibule of meaning: transitivity inversion as a morphological phenomenon, *Studies in Language* 19. 73–127.
- 1995b. Relatives à antécédent interne, nominalisation et focalisation: entre syntaxe et morphologie en bélharien, *Bulletin de la Société de Linguistique de Paris* 90. 391–427.
- 1996. *Aspect, Mood, and Time in Belhare*, Zürich: ASAS Press.
- 1997. Spatial operations in deixis, cognition, and culture: where to orient oneself in Belhare, in J. Nuysts and E. Pederson (eds) *Language and Conceptualization*, 46–83. Cambridge: Cambridge University Press.
- 1998. Rhythm and feet in Belhare morphology, Working Paper No. 287, Rutgers Optimality Archive. Available at: www.ruccs.rutgers.edu/roa.html (accessed 18 August 2016).
- 1999a. Cultural formalism and spatial language in Belhara, in B. Bickel and M. Gaenszle (eds) *Himalayan Space: Cultural Horizons and Practices*, 75–104. Zürich: Museum of Ethnography Press.
- 1999b. From *ergativus absolutus* to topic marking in Kiranti: a typological perspective, *BLS* 25. 38–49.

- 1999c. Nominalization and focus constructions in some Kiranti languages, in Y.P. Yadava and W.W. Glover (eds) *Topics in Nepalese Linguistics*, 271–296. Kathmandu: Royal Nepal Academy.
- 2000a. Deictic transposition and referential practice in Belhare, *Journal of Linguistic Anthropology* 10. 224–247.
- 2000b. On the syntax of agreement in Tibeto-Burman, *Studies in Language* 24. 583–610.
- 2000c. Space, territory, and a stupa in Eastern Nepal: exploring Himalayan themes and traces of Bon, in Y. Nagano (ed.) *New Horizons in Bon Studies*, 685–702. Osaka: National Museum of Ethnology.
- 2003a. Referential density in discourse and syntactic typology, *Language* 79. 708–736.
- 2003b. Prosodic tautomorphemicity in Sino-Tibetan, in D. Bradley, R.J. LaPolla, B. Michailovsky and G. Thurgood (eds) *Variation in Sino-Tibetan and South East Asian Languages*, 89–99. Canberra: Pacific Linguistics.
- 2004. Hidden syntax in Belhare, in A. Saxena (ed.) *Himalayan Linguistics*, 141–190. Berlin: Mouton de Gruyter.
- 2006. Clause-level vs. predicate-level linking, in I. Bornkessel, M. Schlesewsky, B. Comrie and A.D. Friederici (eds) *Semantic Role Universals and Argument Linking*, 155–190. Berlin: Mouton de Gruyter.
- Bickel, B. and M. Gaenzle. 2015. First person objects, antipassives, and the political history of the Southern Kirant, *Journal of South Asian Languages and Linguistics* 2. 63–86.
- Bickel, B., M. Rai, N. Paudyal, G. Banjade, T.N. Bhatta, M. Gaenzle, E. Lieven, I.P. Rai, N.K. Rai and S. Stoll. 2010. The syntax of three-argument verbs in Chintang and Belhare (Southeastern Kiranti), in A. Malchukov, M. Haspelmath and B. Comrie (eds), *Studies in Ditransitive Constructions: A Comparative Handbook*, 382–408. Berlin: De Gruyter Mouton.
- Ebert, K.H. 1986. Reported speech in some languages of Nepal, in F. Coulmas (ed.) *Direct and Indirect Speech*, 146–149. Berlin: Mouton de Gruyter.
- Matisoff, J.A. 1973. *The Grammar of Lahu*, Berkeley: University of California Press.
- 1986. Hearts and minds in South-East Asian languages and English: an essay in the comparative lexical semantics of psycho-collocations, *Cahiers de linguistique asie-orientale* 15. 5–57.
- Nedjalkov, V.P. (ed.) 1988. *Typology of Resultative Constructions*, Amsterdam: Benjamins.
- Stoll, S. and B. Bickel. 2009. How deep are differences in referential density?, in E. Lieven, J. Guo, N. Budwig, S. Ervin-Tripp, K. Nakamura and Ş. Özçalışkan (eds) *Crosslinguistic Approaches to the Psychology of Language*, 543–555. London: Psychology Press.
- Van Valin, R.D. and R.J. LaPolla. 1997. *Syntax: Structure, Meaning, and Function*, Cambridge: Cambridge University Press.

CHAPTER THIRTY-SIX

CAMLING

Karen H. Ebert

1 INTRODUCTION

Camling is one of the many endangered languages of the Kiranti cluster (see Chapter 33). It is spoken mainly in Khotang district and parts of Udaypur in eastern Nepal. The number of speakers is difficult to estimate, as there are many semi-speakers. In the areas I visited children do not learn the language any more. Balamta, a remote village in Udaypur, seems to be the only place where one can still find children speaking Camling. As a consequence, Camling has been heavily influenced by Nepali, especially in syntax and lexicon and on the pragmatic level.

There is a great deal of dialectal variation. A major isogloss bundle runs along the rivers Sapsu and Sun Kosi, dividing northwestern (NW) from southeastern (SE) dialects. Some characteristics of the two dialect groups are:

- NW has initial consonant clusters, SE does not;
- NW has relics of an inverse system, SE marks first person patient by *kha-* with subject agreement;
- agentive nouns are formed with *-pa* in NW, with *ka-* in SE;
- the sequential suffix is *-nʌ* in NW, *-nʌ* or *-ki* in SE.

2 PHONOLOGY

The consonants of Camling are given in Table 36.1.

Voice and aspiration/breathiness are relevant in initial and in medial position.

<i>phuima</i>	pluck	<i>bhuima</i>	pound
		<i>buima</i>	call
<i>thujma</i>	cough	<i>tujma</i>	village
		<i>dujma</i>	drink
<i>ityu</i>	brought from above	<i>idyu</i>	gave him
<i>lhamma</i>	catch	<i>lammha</i>	dump
		<i>lamma</i>	look for

The voiced consonants *g, gh, j, jh* occur only initially and mainly in loans from Nepali. There is no native Camling word beginning with *j*, and the only candidate for *jh, jhara* ‘all’, could be borrowed from some southern source (cf. Dhimal *jhara*). The only lengthened consonants in Camling that occur with some frequency are *p* and *m*. This is partly due to the suffixes *-ma* and *-pa* for female and male, and to grammatical suffixes like *-ma* (infinitive): *chamma, chappa* ‘great-grandmother’, ‘great-grandfather’, *khrumma* (*khrup + ma*) ‘get up’. Five vowel phonemes and five diphthongs can be distinguished on the basis of minimal pairs:

<i>khama</i> (<i>khas-</i>)	be satisfied	<i>khema</i> (<i>khet-</i>)	break (SE)
<i>khima</i> (<i>khi-</i>)	quarrel	<i>khuma</i> (<i>khus-</i>)	steal, hide

TABLE 36.1 CONSONANT PHONEMES

	Labial	Dental	Alveolar	Velar	Glottal
stop, -voice	p	t	c [ts]	k	
	ph	th	ch [ts ^h]	kh	
+voice	b	d	(j) [dz]	(g)	
	bh	dh	(jh) [dʒ ^h]	(gh)	
fricative	f	s			h
nasal	m	n		ŋ	
	mh	nh			
continuants	l, lh, r, rh				
glides	w, y				

<i>khōma (kharj-)</i>	see			
<i>khaima (khat-)</i>	go	<i>kheima (khet-)</i>	split, cut up (SE)	
<i>maima (maid-)</i>	make	<i>muima (muit-)</i>	ripen	
<i>məima (məid-)</i>	forget			

However, there is a great deal of variation in the realization of vowels and diphthongs. Thus we find free or dialectal variation between:

- *i* and *u* after the central consonants (dentals and alveolars): *sumra* ~ *simra* ‘three’, *-yu* ~ *-yi* ‘3rd patient’;
- *o* and *u* in some words: *lodyu* ~ *ludyu* ‘he told him’, *oko* ~ *uko* ‘this’, *tyoko* ~ *tyuko* ~ *tyiko* ‘that’;
- *e*, *o* and *yo* in some words: *de* ~ *do* ~ *dyo* ‘what’, *themma* ~ *thomma* ~ *thyomma* ‘perform’, ‘dance’.

Vowels tend to be centralized in diphthongization. In the most southern dialect all diphthongs are reduced to *e*.

The status of the unrounded back vowel *ʌ* is unclear. As there is no [a] before *r* in verb stems, a pair like *chorma* ‘pay’ vs *charma* ‘urinate’ probably represents the opposition /o/ : /a/, although some speakers have collapsed both to [o]. The only opposition I found with /a/ is between the topic marker *-na* and the linker *-nʌ*. but both are often pronounced the same.

The canonical syllable structure is cv(c). The NW dialect has initial consonant clusters, restricted to *p(h)*, *k(h) + r*, *l*: NW *krupma* ‘get up’, *praima* ‘shout’, *phloma* ‘help’, but SE *khumma*, *paima*, *phoma*. In word-final position only non-aspirated sonorants occur.

3 NOMINALS

3.1 Pronouns and possessives

The various pronouns of Camling are given in Table 36.2.

The Camling pronouns referring to speech-act participants distinguish dual and plural and first person inclusive and exclusive. The second person plural pronoun is used as an honorific in addressing older people (possibly due to Nepali influence). The suffix *-ci* is a non-singular marker with third person pronouns, demonstratives and nouns, a dual marker with first and second person pronouns. The pronoun *khu* is restricted to human referents.

TABLE 36.2 PERSONAL PRONOUNS, POSSESSIVES AND DEMONSTRATIVES

	pers. pronouns	poss. pronouns	poss. prefixes
1s	<i>kaya ~ kā</i>	<i>aya</i>	<i>a-</i>
1di	<i>kaici</i>	<i>icmo ~ acmo</i>	<i>ic- ~ ac-</i>
1de	<i>k^cka</i>		
1pi	<i>kai(ni)</i>	<i>imo</i>	<i>i-</i>
1pe	<i>kaika</i>	<i>amka</i>	
2s	<i>khana</i>	<i>khamo</i>	<i>kap-</i>
2d	<i>khaici</i>		
2p	<i>khaini</i>	<i>khaimo</i>	<i>kai-</i>
HUM 3s	<i>khu</i>	<i>khumo</i>	<i>m-</i>
ns	<i>khuci</i>	<i>khucimo</i>	<i>kic-</i>
PROX s	<i>uko</i>	<i>ukomo</i>	
ns	<i>ukoci</i>	<i>ukocimo</i>	
DIST s	<i>tyuko, tyu(ko)pa</i>	<i>tyukomo, tyu(ko)pamo</i>	
ns	<i>tyukoci</i>	<i>tyukocimo</i>	

The demonstratives, formed from deictic roots *u-* and *tyu-* plus nominalizer, serve as pronouns for both humans and non-humans. When referring to a person, the forms *tyupa* and *tyukopa* (see (8a)) are also possible.

The system of possessive prefixes and possessive pronouns is defective. Sometimes *i-* is used as a general marker for first person plural, e.g. *amka i-tujma* ‘our_e village’, *i-la* ‘our language’ (also when speaking to a non-Camling). Only *aya* and *amka* are genuine possessive pronouns; the other forms are made up of prefix + genitive (*icmo, imo*) or they are genitives based on the personal pronouns (*khamo, khumo*). The possessive pronouns are used as independent nominals or in prenominal position, where they can be combined with a possessive prefix.

<i>uko aya</i>	this is mine
<i>aya a-khim = a-khim = aya khim</i>	my house
<i>khamo kap-khim = kap-khim = khamo khim</i>	your house

The possessive prefixes are obligatory with inherently relational nouns, which include – besides the usual kinship terms, names for parts, etc. – terms for emotions: *a-kurma* ‘my fear’, *m-sikha* ‘his liking’, and for order: *m-dōsi* ‘behind’, ‘later’, *m-selam* ‘the next day’ (cf. *selam* ‘tomorrow’).

3.2 Composite and derived nouns

Nouns are compounded by simple juxtaposition, e.g. *wa-hui* ‘river’ (water + down), *wa-dupkha* ‘water-drinking place’. Often elements of a compound cannot be used alone. Of the following nouns only those with a bisyllabic stem occur without *sa* (‘meat’): *wasa* ‘chicken’, *yasa* ‘fish’, *sosa* ‘porcupine’, *lusa* (frog species), *bose* (**bosa*) ‘pig’, *rapa(sa)* ‘fish otter’, *pira(sa)* ‘deer’, but the head of such compounds can be combined with other nouns, like *khli* ‘excrement’: *wakhli* ‘chicken droppings’, *bokhli* ‘pig droppings’. *wa* itself functions as a class noun both in the meaning ‘bird’ and ‘water’: *tuwa* ‘pheasant’, *khlawu* ‘woodpecker’; *mikwa* ‘tears’, *bouwa* ‘soup’.

Nouns denoting animates often indicate natural gender through the suffixed (seldom prefixed) gender markers *ma* and *pa*, which also mean ‘mother’ and ‘father’: *kokma/*

kokpa ‘grandmother’/‘grandfather’, *masuyma/pasunpa* ‘old woman’/‘old man’, *wama/wapa* ‘hen’/‘cockerel’. A few frozen adjectives with gender endings could be elicited, but in actual discourse they were always replaced by Nepali terms: *cukpa/cukma* ‘small’ (→ Nep. *sano*), *makcukpa/makcukma* ‘black’ (→ Nep. *kalo*). Words corresponding to English adjectives are mostly nominalized verbs.

Derivational processes are very limited in Camling. The most frequent derivational morpheme is the nominalizer *-ko*, which derives nouns and attributes (see section 5.3). Agentive nouns are derived from verbs with *-pa* in NW, *ka-* in SE Camling: *khurpa/kakhur* ‘carrier’, ‘porter’. Patientive, instrumental and locational nouns are formed with *-kha*: *woikha* ‘clothes’ (<*woi-* ‘put on’), *bhuikha* ‘pounder’, *hiykhha* ‘living place’.

3.3 Classifiers

There are three classifiers: *-po* for humans, *-li* for round objects and *-ra* for everything else. They combine only with the numerals ‘one’ to ‘three’ of Tibeto-Burman origin. Higher numerals are borrowed from Nepali. A non-singular number is, as a rule, marked only with human referents.

<i>haka-po</i>	<i>maricha-ci</i>	two women
two-CL	girl-ns	
<i>i-li</i>	<i>suntala</i>	one orange
one-CL	orange	
<i>sim-ra</i>	<i>pyupa/khim/yosi</i>	three cows/houses/bananas
three-CL	cow/house/banana	

The functional load of the classifiers is minimal, and forms with *-ra* are winning general acceptance with all nouns. Quantifying classifiers can be exemplified by *i-ruj* ‘once’, *haka-ruj* ‘twice’, *i-homa rō* ‘one mana (measure) rice’. Nepali measure terms combine freely with Camling numerals, e.g. *i-mana rō* ‘one mana rice’.

3.4 Case markers

Camling has an ergative-instrumental case marker *-wa* and a genitive *-mo*. The dative suffix *-lai* is borrowed from Nepali and optional with human undergoers. All can be demonstrated with the following constructed example:

- (1) *Ram-mo m-nicho-wa a-woini(-lai) bhe-wa ap-u.*
 (name)-GEN 3sPOSS-y.sibling-ERG 1sPOSS-friend-DAT arrow-INST aim-3P
 ‘Ram’s younger brother aimed at my friend with an arrow.’

(For split ergative marking see section 5.1.) Different from other Kiranti languages, Camling has no comitative case marker, which would also serve to coordinate nouns. Nouns are juxtaposed or connected by Nepali *rā* ‘and’: *kusya (rā) puchō* ‘youngest daughter and the snake’. Comitative meaning can sometimes be expressed with the postposition *tōda* ‘near’.

Camling shares with other Kiranti languages a unique system of locative markers indicating altitude in relation to the point of reference.

<i>khim-dhi~ -di</i>	at/to the house (at a higher location)
<i>khim-i</i>	at/to the house (at a lower location)

<i>khim-ya</i>	at/to the house (at the same level)
<i>khim-da</i>	at/to the house (neutral)

The altitudinal case markers can be combined with deictic roots as well: *udhi* ‘up here’, *tyudhi* ‘up there’ (see neutral *uda*, *tyuda*). The ablative marker *-ka* and the optional allative *-ni* are suffixed to a locative: *khim-i-ni* ‘down to the house’, *khim-ya-ka* ‘over from the house’. The opposition up–down, which is equivalent to north–south under the geographical conditions of the Himalayan slopes, is of eminent importance for the Camling spiritual world. The higher regions are associated with the male ancestor gods and purity, the lower regions with the female ancestor gods and fertility. The up–down opposition is a constant theme in Camling mythology and rituals (see Ebert 1999).

4 THE VERB

4.1 Stems

The full verb stem occurs only before vowels. Before a consonant-initial suffix stem-final *t*, *d*, *s* are elided. This results in diphthongization of the stem vowels *a*, *e*:

stem	+ <i>-ma</i> (INF)
<i>lod-</i>	<i>lo-ma</i> tell
<i>khis-</i>	<i>khi-ma</i> comb
<i>caidh-</i>	<i>cai-mha</i> beat
<i>khat-</i>	<i>khai-ma</i> go
<i>set-</i>	<i>sei-ma</i> kill

Stem-final *k* sometimes leaves a trace in the preconsonantal stem. *a + k* yields *o* before a consonant-initial suffix, whereas *i/e + k* sometimes results in a glide + *u/o*.

<i>pak-</i>	<i>po-ma</i>	put
<i>bik-</i>	<i>byu-ma</i>	sweep
<i>hek-</i>	<i>hyo-ma</i>	cut

From some verbs causative or applicative stems can be formed with an augment *t ~ d*, which replaces a stem-final *s*, if there is one.

<i>i-</i>	come down	<i>it-</i>	bring down
<i>si-</i>	die	<i>set-</i>	kill
<i>ri-</i>	laugh	<i>rit-</i>	laugh at
<i>ban-</i>	come over	<i>baid-</i>	bring over
<i>libs-</i>	turn around	<i>libd-</i>	turn sth.
<i>hors-</i>	throw	<i>hord-</i>	throw at

Secondary causatives are possible with a few verbs, e.g. *set- + t > seid-* ‘kill for someone’.

- (2) *kap-ghicro-da a-ma sor bo yyabd-e-yas-e.*
2SPOSS-neck-LOC 1SPOSS-mother louse PART stick-NPST-V 2:stay-NPST
—*seid-uy-na!*
kill:APPL-1S -IMPER
‘Mother, a louse is sticking to your neck.’ – ‘Kill it for me!’

Analytic formations with *maid-* ‘make’ are more frequent as primary causatives: *sei maima* ‘kill for someone’, *khõ maima* ‘show’.

A reflexive-reciprocal stem is formed with the root extension *-umc* ~ *-unc* in NW dialects, *-āic* ~ *-āit* in SE dialects, e.g. *hupd-umc-uja/hupd-āic-uja* ‘I washed myself’, *mi-ras-umca/mi-ras-āica* ‘they separated’ (see *ras-* ‘divide’).

4.2 Person and number affixes

Camling, like the neighbouring Kiranti languages, is characterized by complex verbal paradigms, in which person is marked partly by prefixes, partly by suffixes, sometimes independently of role and number.

Speech-act participants are always marked on the verb. The first person singular marker is independent of semantic role; the second person prefix *ta-* is independent of role and number.

1s	<i>khat-uja</i>	I went
1s → 3s	<i>phlod-uja</i>	I helped him
3s → 1s	<i>pa-phlod-uja</i>	he helped me
	INV-help-1s	
2s	<i>ta-khata</i>	you went
3s → 2s	<i>ta-phloda</i>	he helped you
2p → 3s	<i>ta-phlod-u-m</i>	you _p helped him
	2-help-3p-1/2pA	

The first/second person share the role specific plural markers *-i* and *-m*. But as *-i* has the variant *-im* before exclusive *-ka*, one could postulate the morpheme /im/ with the allo-morphs *m* after a vowel, *i* ~ *im* before *-ka*, and *i* elsewhere.

3 → 2p	<i>ta-lod-i</i>	he/they told you
2p → 3s	<i>ta-lod-u-m</i>	you told him
3s → 1pi	<i>pa-lod-i</i>	he told us
3s → 1pe	<i>pa-lod-i(m)-ka</i>	he told us
1pe	<i>khat-i(m)-ka</i>	we went
1pe → 3s	<i>lod-u-m-ka</i>	we told him

The dual marker *-ci* is independent of person and role.

1di	<i>khata-ci</i>	we went
= 3d		they went
2d → 3	<i>ta-loda-ci</i>	you told him/them
= 3 → 2d		he/they told you

The homophonous third non-singular patient marker *-ci* always follows *-u* and the person-number markers *-m*, *-y*, which together are copied after *-ci*. Outer suffixes like the exclusive and the non-past marker follow the copy.

1pi → 3ns	<i>lod-u-m-c-um-e</i>	we will tell them
	tell-3p-1p-3nsp-[copy]-NPST	
3s → 3ns	<i>set-yu-c-yu</i>	he killed them
	kill-3P-3nsp-[copy]	

The dual marker and the third non-singular patient marker cannot occur together; with a dual actor, a third non-singular patient remains unmarked: *tyoka-ci* ‘they_d saw him/them’; ‘we_{di} saw him/them’.

TABLE 36.3 CAMLING PERSON-NUMBER MARKERS NW DIALECTS

A	P	<i>Is</i>	<i>lde</i>	<i>lpe</i>	<i>ldi</i>	<i>lpi</i>	<i>2s</i>	<i>2d</i>	<i>2p</i>	<i>3s</i>	<i>3ns</i>	<i>itr.</i>
<i>Is</i>							-na	-naci		-nani	-nupa	-unjya
<i>lde</i>											-nuya	
<i>lpe</i>										-eka		
<i>ldi</i>										-umka	-umcumka	-i(m)ka
<i>lpi</i>										-ci		
<i>2s</i>	<i>ta-</i>	<i>ta-</i>	<i>-cka</i>	<i>ta-</i>	<i>-i(m)ka</i>					-um	-umcum	-i
<i>2d</i>										<i>ta-</i>	<i>-u</i>	<i>ta-</i>
<i>2p</i>										<i>-ueyu</i>		<i>ta-</i>
<i>3s</i>	<i>pa-</i>	<i>pa-</i>	<i>-cka</i>	<i>pa-</i>	<i>-i(m)ka</i>	<i>pa-</i>	<i>-ci</i>	<i>pa-</i>	<i>-i</i>	<i>ta-</i>	<i>-um</i>	<i>ta-</i>
<i>3d</i>										<i>-ci</i>	<i>-ueyu</i>	
<i>3p</i>										<i>pa-</i>	<i>-ci</i>	
SE dialects: 1st person patient scenarios												
		<i>Is</i>	<i>lde</i>	<i>lpe</i>	<i>ldi</i>	<i>lpi</i>						
<i>2s</i>							<i>kha-ta-</i>					
<i>2d</i>							<i>kha-ta-</i>	<i>-ci</i>				
<i>2p</i>							<i>kha-ta-</i>	<i>-i</i>				
<i>3s</i>							<i>kha-</i>					
<i>3d</i>							<i>kha-</i>	<i>-ci</i>				
<i>3p</i>							<i>kha-mi-</i>					

4.3 Inverse configurations

The prefix *pa-* is an inverse marker in NW and *-u* (with the variants *-yu* ~ *-yi*) a direct marker. The 3 → 3 configurations are crucial for this interpretation:

3s → 3s	<i>chik-u</i>	he pinched him	3p → 3	<i>pa-chika</i>	they pinched
1pe → 3s	<i>chik-u-m-ka</i>	we pinched him	3s → 1pe	<i>pa-chik-im-ka</i>	he pinched us

The distribution of the affixes is self-explanatory in an empathy hierarchy 1/2 > 3s > 3p. But, as Table 36.3 demonstrates, the system is disturbed. Second person *ta-* does not combine with inverse *pa-* due to a one-prefix-restriction, *-uya* (< u + ya) has been generalized as a first person singular marker, and dual actor forms have no *-u*. That these must be recent developments becomes clear from comparison with closely related languages like Bantawa and Athpare (see ‘Kiranti’ in this volume), which have *-u* in all direct configurations and no *-u* in inverse. Although the analysis of the affixes as direction markers is not totally satisfactory, the alternative of interpreting *pa-* as a third person actor marker is less so, as it does not account for 3s → 3 *-u* vs 3p → 3 *pa-*. I gloss *-u* generally as 3p, as this accounts better for forms in the SE dialect, where the only inverse marker is *pa-* in 3p → 3. Further, it facilitates comparison with Kiranti languages in which *-u* is an unambiguous 3P marker.

The SE dialects have developed different forms to code inverse actions directed towards first person. Instead of the old first person markers the verb takes the clitic *kha-*, which is independent of number and not subject to the one-prefix-restriction. All other forms are as in NW. The new forms have the advantage of unambiguously marking the actor.

NW		SE		
3s → 1di	<i>pa-loda-ci</i>	3s → 1	<i>kha-loda</i>	he told me/us
3d → 1di	<i>pa-loda-ci</i>	3d → 1	<i>kha-loda-ci</i>	they _d told me/us
3p → 1di	<i>pa-loda-ci</i>	3p → 1	<i>kha-mi-loda</i>	they _p told me/us

The *kha-* forms seem to imitate the principle of marking inverse by a prefix, but the pattern of participant marking has been totally restructured, as can be seen from Table 36.3. The inverse forms of NW Camling conform to the principle of marking speech-act participants on the verb; *pa-* indicates only direction. *kha-* in the SE dialect stands for first person patient, and the rest of the verb marks second or third A as in the intransitive paradigm.

2p → 1	<i>kha-ta-lod-i</i>	you told me/us	2p	<i>ta-khat-i</i>	you went
3p → 1	<i>kha-mi-loda</i>	they told me/us	3p	<i>mi-khata</i>	they went

Syntactically the verb remains transitive; third person actors take the ergative marker (see (6b)).

4.4 Tense-aspect forms, imperative

Camling has two basic aspecto-temporal forms: an unmarked aorist/past and a marked imperfective/non-past. Stem + *a* constitutes a finite base. The aorist/past is unmarked, the imperfective/non-past is marked by *-e* following the person-number markers. After *-i* and *-m* it can take the form *-ne*; *-u* + *-e* → *-yo*.

	PST (unmarked)		NPST	
3s	<i>khata</i>	he went	<i>khat-e</i>	he will go
3d	<i>khata-ci</i>	they went	<i>khata-c-e</i>	they will go
1pi	<i>khat-i</i>	we went	<i>khat-i-(n)e</i>	we will go
3s → 2s	<i>ta-tyoka</i>	he saw you	<i>ta-tyok-e</i>	he will see you
1pe → 3s	<i>tyok-u-m-ka</i>	we saw it	<i>tyok-u-m-k-e</i>	we will see it
3s → 3s	<i>lod-yu</i>	he told him	<i>lod-yo</i>	he will tell him
3s → 3s	<i>pak-u</i>	he put it	<i>pak-yo</i>	he will put it

The data suggest that the forms are basically aspects and that the temporal interpretation as past and non-past is secondary. The aorist/past is used to express the sequence of events in narratives or to report past events in conversation. Most narrators change to the imperfective/non-past when they report habitual or iterated actions or describe backgrounding situations. However, the past is regular with purely stative verbs, e.g. *hiya* ‘he was’ (not: ‘he became’, as in most neighbouring languages), and in periphrastic tenses -Ø and -e have purely temporal function. Aktionsart has not been investigated systematically, but verbs of perception for example are momentaneous or ingressive. An actually perceived smell can only be verbalized as *nhama* ‘it stinks’ (i.e. ‘it came to stink’), whereas *nhamē* is a statement based on general knowledge.

A resultative-perfect is expressed by a compound verb formation with the verb *yas*-‘stay’, ‘remain, keep’. The SE dialect has a progressive marker *-uys* ~ *-ōs* (contracted from *yas*-?). Both dialects have a grammaticalized ambulative marked by *-hod*:

<i>mi-khat-e-ŋas-e</i>	they have gone/they are gone
3ps-go-NPST-V2:stay-NPST	
<i>rō m-uys-yo</i> (SE)	she is preparing food
rice make-PROG-3P:NPST	
<i>suy khop-hod-yo</i>	she is moving around cutting firewood
wood cut-AMB-3P	

The imperative has the form of the unmarked 3s/d intransitive and the 3s/3dA transitive (+ P markers); in the plural the second person plural suffix *-ni* is added to stem + *a*.

khata! *khataci!* *khatani!* go!
diju! *dijaci!* *dijani!* drink it!

First person patient forms differ according to the dialectal inverse paradigms: NW *iduya*! SE *kha-ida*! ‘give it to me!’ The first person dual and plural non-PAST forms serve as hortatives: *khatace*, *khatine* ‘let’s go (d/p)’.

4.5 Negation

There is much dialectal and individual variation in the negated paradigms. The following rules generally hold for the basic tense-aspect forms:

- negation is marked by a prefix *pa-* and a suffix *-na*, *-āi*, *-aina* (probably an adaptation to Nepali);
- negative *pa-* is not realized before second person *ta-* due to the one-prefix-restriction;
- *pa-* replaces 3pS *mi-*;
- first person singular has *-n* after the stem;
- there are no direction markers.

	<i>asserted</i>	<i>negated</i>	
1s	<i>khat-u-ja</i>	<i>pa-khai-n-uya</i>	I did not go
1di	<i>ta-khata-ci</i>	<i>pa-khata-ci-na/-c-ãi/-c-aina</i>	we did not go
2d	<i>ta-khata-ci</i>	<i>ta-khata-ci-na/-c-ãi/-c-aina</i>	you did not go
3s → 3s	<i>lod-yu</i>	<i>pa-lod-aina</i>	he did not tell him
1pi	<i>khat-i</i>	<i>pa-khat-imna/-umna</i>	we did not go
1p → 3s	<i>lod-u-m</i>	<i>pa-lod-imna/-umna</i>	we did not tell him
tell-3P-1PA		NEG-tell-1P:NEG	

Some speakers of NW dialects have a different paradigm for the negative non-past, with *-ð* in forms that have no personal suffixes and *-mi* in first and second person forms (again with some variation): *pa-los-ð* ‘he will not sell it’, *pa-los-um-mi* ‘we will not sell it’.

The negative perfect-progressive is formed by *mi-* + stem followed by a finite marked form of *yas-*.

	<i>asserted</i>	<i>negated</i>		
1pi	<i>tip-i-yas-i-e</i>	<i>mi-tim</i>	<i>yas-i-e</i>	we have not met

meet-1ps/p-v2:stay-1ps/p-NPST NEG-meet v2:stay-1ps/p-
NPST NPST

In the negative imperative the stem is negated by *mi-* and followed by the negative auxiliary *da-*:

2s	<i>khata!</i>	<i>mi-khai-da!</i>	don’t go!
2s→1s	<i>setuya!</i>	<i>mi-sei-duya!</i>	don’t kill me!

4.6 Compound verbs

In the Camling compound verb construction both verbs carry finite markers, but prefixes occur exclusively on the first verb, outer suffixes on the second. Only those suffixes remain after v1 which make up a syllable together with the stem-final consonant. The list of Camling second verbs is similar to that found in many other South Asian languages. Second verbs are desemanticized to some degree, most often they telicize the main verb.

<i>ca-m-pak-u-m-ka</i>	we ate it up
<i>eat-1/2PA-v2:put-3P-/2PA-e</i>	
<i>mi-pera-khata</i>	they flew away
<i>3pS-fly-v2:go</i>	
<i>ta-mobdh-yu-kas-yu</i>	you spilt it
<i>2-dump-3P-v2:throw-3P</i>	

pid- occurs only as v2 with benefactive function, but it is easily recognizable as an older form of the verb ‘give’, which became *id-* in Camling. Optative is expressed with the help of *ni-* ‘be good’ as v2: *phlodyo-nyo* ‘let him help’.

4.7 Non-finite forms

Camling has three non-finite verb forms. The infinitive in *-ma* is used as a citation form and as a noun. In questions it expresses the modal meaning ‘shall’: *de muma?* ‘What shall we do?’ Infinitive clauses are complements to modal, evaluative or phasal verbs: *cama tire* ‘let’s eat’ (*lit. eat-INF must*), *kh(r)amma puisyu* ‘he began to cry’.

The purposive is mainly used as a complement of motion verbs. The undergoer is indicated by a possessive prefix.

- (3) *kai-ka kap-la-si ta-i -k-e.*
 we_p-e 2SPOSS-fetch-PURP come-1/2pS/P-e-NPST
 ‘We will come to pick you up.’

The simultaneous converb (-sa, typically reduplicated) describes an activity accompanying the action expressed by the main verb, which is most often a verb of motion or rest. There is necessarily subject identity (see (8d)).

It is remarkable – even in the context of the preferably finite-marking SE Kiranti languages – that the language has only one converb and no participles. The forms that morphologically correspond to participles in other Kiranti languages, i.e. agentive and patientive nouns (see section 3.2), cannot be attributed.

5 SYNTAX

5.1 Basic sentence patterns

The basic word order is generally verb final. In unmarked order the actor or intransitive subject takes the initial position; patient follows a goal. Modifiers, including subordinate clauses, precede the head. There is much freedom to rearrange the elements according to communicative needs. Any major constituent can stand after the verb (see (6b)).

Camling is morphologically a split ergative language. Third person actors are marked ergatively, whereas first and second person actors have no case marker. The subject of an intransitive clause and the undergoer remain unmarked, though a human undergoer may take the Nepali dative suffix *-lai*.

- (4) a. *kaya khana(-lai) lo-na.*
 I you(-DAT) tell-1 → 2
 ‘I told you.’
- b. *khu-wa kaini pa-phlod-i.*
 s/he-ERG we_{pi} INV-help-1/2ps/P
 ‘She helped us.’

‘Dative subjects’ are conspicuously absent in Camling. Emotional states are expressed by an abstract noun, which refers to the emotion, followed by an impersonal form of the verb *la-*, that does not occur independently (possibly from Nepali *lāg-* ‘become perceptible’, ‘be felt’, ‘occur’). The experiencer is indicated by a possessive prefix: *m-bulma lae* ‘he is angry’, *a-sikha lae* ‘I like it’.

Participants are mainly coded by the person-number affixes on the verb, which contain more information than the optional pronouns. Role of first or second person is understood only from the verb.

- (5) a. *(kaika) khat-im-ka* we went
we_{pe} [=itr. s] go-1ps/p-e
- b. *(kaika) phlod -u-m-ka* we helped him
we_{pe} [=Actor] help-3p-1/2pa-e
- c. *(kaika) pa-phlod-IM-ka* he/they helped us
we_{pe} [=Patient] inv-help-1ps/p-e

That participant marking on the verb is not grammatical agreement is evident from examples like the following:

- (6) a. *uko-lai phakai mu-ma sopa-wa ta-cap-u-m-ne?*
 this-DAT court do-INF who-ERG 2-can-3P-1/2P^A -NPST
 ‘Who of you can [go and] court him?’
- b. *khim-da mi-hij-e-ko-ci-lai rairewa-kukuwa kha-maid-e*
 house-LOC 3pS-be-NPST-NOMZR-p-DAT (ritual) 1P-make -NPST
nochuj-wa
 shaman-ERG
 ‘For those in the house, the shaman performs the *rairewa-kukuwa* ritual for us.’

Camling has a broad topic marker *-na*, which can occur several times in a clause to mark the theme or a contextually given element. Constituents can be focused by *cāi* or *pāni* (both from Nepali), or by the nominalizer *-ko* (see section 5.3). Topic and focus markers need further investigation.

Yes–no questions are marked only by intonation. The answer ‘yes’ is usually expressed by repeating the questioned verb. The preferred position of the question word is before the verb. The verb *hijma* ‘sit’, ‘live’, ‘exist’, ‘be’ is used in locational and existential sentences. The negative counterparts have no verb, but the invariable particle *paina*. Identity statements have no copula (see (9), (11a)) and are negated by *aina* ‘no’, ‘not’.

- (7) a. *kap-khim khoda hij-e?*
 2SPOSS-house where be-NPST
 ‘Where is your house?’ i.e. ‘Where do you live?’
- b. *kic-nicho-ci mi -hij-e? – mihije./paina.*
 3nsposs-y.sibling-ns 3ps-exist-NPST
 ‘Do they have younger siblings?’ – ‘Yes’./‘No’.

5.2 Clause linkage

Subordinate clauses are preposed in Camling and the subordinator takes the clause-final position. What is unusual in the South Asian context (and for verb final languages in general) is the fact that – with the exception of the three non-finite forms mentioned – the verb has the full range of finite markers. Event sequences are linked by *-na* following a finite verb or an infinitive; i.e. Camling has no sequential converb (‘conjunctive participle’). Speakers of SE dialects also use *-ki*, which is a connector in some Kiranti languages further east.

The following short excerpt from a mythological story in the SE dialect demonstrates linkage with *-ki*, a temporal clause marked with the nominalizer + topic marker (*-ko-na*), and another one with the grammaticalized temporal subordinator *-pana* (old nominalizer *-pa* + topic). It moreover exemplifies the typical method of connecting discourse by repeating the predicate of the preceding clause.

- (8) a. *khawa-wa dig-u-pana 'lo, u-ko-wa dig-u!*
 woodpecker-ERG drink-3p-TEMP INTJ PROX-NMZL-ERG drink-3P
mi-riya-ki tyu-ko-pa-mo m-lem
 3ps-say-SEQ DIST-NMZL-NMZL-GEN 3SPOSS-TONGUE
pa-bhurda-ko raicha.
 inv-pull.out-NMZL REP

- b. *m-lem pa-bhurda-ki kamalapa pa-maida-ko raicha.*
[repet.] -SEQ upside.down INV-make-NMZL REP
- c. *kamalapa pamaida tyudaka arko thañ mi-khata-ko-na tyu-da*
[repet.] then other place 3ps-go-nmzl-top dist-LOC
pani hija-ko wa-na chirkucipa-wa diy-u-ko raicha. [...] FOC be-NMZL water-TOP grasshopper-ERG drink-3P- NMZL REP
- d. *tyudaka lam-sa lam-sa mi-khata-pana cikalemma*
then search-SIM search-SIM 3pS-go-TEMP (name)
rujma-ko wasa-wa 'tyuda hij-e, *tyuda hij-e!*'
QUOTE-NMZL bird -ERG there be-NPST [repet.]
lod-yi -yas-yi-c-yi-ko raicha-ki tyu-ko-ci-lai
tell-3P-V2:keep-3P-3nsp-[copy]-NMZL REP-SEQ DIST-NMZL-NS-DAT
khõ maid-yi-c-yi-ko . . .
see make-3P-3nsp-[copy]-NMZL

(During a drought the ancestor-god Nayima sends out animals to search for water: . . .)

- a. When the woodpecker drank, they said: ‘Look, he drank!’ and they pulled out his tongue.
- b. They pulled out his tongue and turned it the wrong way round.
- c. Then, when they went to some other place, the grasshopper drank some water that was there. [...]
- d. As they went, searching, searching, the bird called cikalemma kept telling them: ‘There is some, there is some!’ and he showed them . . .

Complement clauses to cognitive verbs are embedded with the quote particle *rujma(pa) ~ rujmanə* (< *rujma* ‘say’), which is optional before a speech-act verb, e.g. SPEECH (*rujmanə*) *lod-yu* ‘he told him’. Thoughts are rendered in the form of direct speech, i.e. with the deictic elements unshifted.

- (9) *a-kuruŋpa rujma pa-chaid-aina, Saphopte-wa pəni*
1sposs-mat.uncle quote NEG-know-NEG (name) also
a-cyodum-ci rujma pa-chaid-aina.
1sposs-niece-ns QUOTE NEG-know-NEG
‘She did not know that he was her maternal uncle, and Saphopte also did not know that they were his nieces.’

5.3 Nominalizer

The nominalizer *-ko* is one of the most frequent morphemes in Camling, as the text passage (8) demonstrates. Any verb, sentence, adverbial or deictic root combines with it, and any nominalized form can stand as a noun or as an attribute, e.g.:

- (10) *lhise-ko siri* a heavy basket
cane-ko rō tasty food
ase-ko dum yesterday’s story
u-ko this
khimda mihiye-ko-ci those who live in the house (6b)

Nominalized sentences can constitute headless relative clauses, or they can be attributed as relative clauses; see *tyuda pəni hijako wa* in (8c). They are complements to perception

verbs, or temporal clauses (8c). In narratives every sentence is nominalized before the report particle *raicha* (< Nepali). Moreover, the nominalizer is used to focus a noun or the whole utterance.

- (11) a. *pa-i-na-n-e, u-ko a-rō-ko.*
NEG-give-1→2-NEG-NPST prox-NMZL 1SPOSS-rice/food-NMLZ
'I won't give it to you; this is my food!'
- b. *kho-ni ta-khat-e-ko.*
which-ALL 2-go-NPST-NMZL
'Where are you going?'

The multifunctional nominalizer has many parallels in other Sino-Tibetan languages.

ADDITIONAL ABBREVIATIONS

INV	inverse
NW	northwestern
SE	southeastern
SEQ	sequential
TEL	telicizer
v2	second verb (explicator, aspectivizer)

BIBLIOGRAPHY

- Ebert, Karen H. (1997) *Camling*, Munich: Lincom.
— (1999) 'The UP–DOWN dimension in Rai grammar and mythology', in Balthasar Bickel and Martin Gängzle (eds) *Himalayan Space: Cultural Horizons and Practices*, Zürich: Völkerkunde Museum, 107–134.
— (2000) *Camling Texts and Glossary*, Munich: Lincom.

CHAPTER THIRTY-SEVEN

WAMBULE RAI

Jean Robert Opgenort

1 INTRODUCTION

Wambule (Nep. *Vāmbule*) is the name of a Western Kiranti language spoken in Okhal- $\ddot{\text{d}}$ hungā, Khotān, Udaypur and Sindhlū districts in eastern Nepal, around the confluence of the Sunkosi and Dūdhkosī rivers. Some other names by which this language is known from the literature are ‘Umbule’ (Hanßon 1991) and ‘Chouras’ya’ (Hodgson 1857), ‘Chourasya’ (Grierson 1909) and ‘Chourase’ (Hanßon 1991). The latter names are anglicised forms of the historical Nepali name *Caurāsiyā* or ‘Eighty-Four’ counties. The name ‘Wambule’ is an endonym, which may be derived from the native word *wam* ‘bear’. There are some 5,000 speakers of Wambule, mostly bilingual in Nepali. The language most closely related is Jero (Nep. *Jero, Jerunī*).

2 PHONOLOGY

The canonical syllable is (C_i) (C_2) (C_3) V (C_f), where C_i is the initial consonant, C_2 is the consonant /l, r/, C_3 is the consonant /w, y/ before /a/, V is a nuclear vowel, and C_f is the final consonant.

2.1 Consonants

The Wambule consonants are presented in Table 37.1.

All the consonants occur in syllable-initial position (C_i). However, only the restricted set /p, t, k, m, n, η , s, r, l/ appears in syllable-final position (C_f). There is no opposition in aspiration and voicing for final stops and there is no affricate. Each final /p, t, k/ has a voiced and a voiceless allophone. The voiced allophone occurs before a voiced initial stop. The voiceless allophone occurs elsewhere, and is often realised with a simultaneous glottal stop in word-final position. The glottal stop /ʔ/ has phonemic status in intervocalic position. Nasalised vowels are analysed as plain vowels plus /ŋ/, and they are phonetically long [V̄]. Even though breathy voiced stops can be found in some indigenous words, they are more common in loans from Nepali. In contrast to Nepali, which distinguishes dental and retroflex stop phonemes, Wambule may be analysed as having just one series of coronal stops. Contrastive alveolar and retroflex series are typically found in loans from Nepali. In native words, the use of retroflex stops instead of alveolar stops depends on personal style. The consonants /l, r/ occur as C_2 in a consonant cluster, where C_i are /p, ph, b/ or /k, kh, g/. The consonants /w, y/ can occur as C_i , but also in a consonant cluster after C_i or $C_i C_2$ when followed by /a/. Wambule does not permit sequences of consonants in syllable-final position. For instance, morphemic root ending <pt> is reduced to /p/ before other suffixes, e.g. <hipt-ca-me> → /hipcam/.

2.2 Vowels

The Wambule vowels are presented in Table 37.2.

The open syllable of polysyllabic words and some closed monosyllabic verb forms are characterised by a length contrast between the short vowels /i, a, u/ and the long vowels /i:, a:, u:/. The vowels /e, o/ are analysed as long vowels and the loan vowel /ʌ/ is short.

2.3 Some historical observations on the modern vowel system

There are reasons to assume that the modern sequences /ya, wa/ reflect the former short vowels /*ɛ, *ɔ/, at least partially, and that the contemporary vowel system derives historically from a five-vowel system.

There are some phonological particularities that plead in favour of analysis of /ya, wa/ as the former vowels /*ɛ, *ɔ/. Firstly, there is the phonotactic restriction, common among the languages of Nepal, that /y, w/ may only be followed by /a/, except that in Wambule, /y/ in C_i position may also be followed by back vowels. Secondly, /ya, wa/ may occur in a consonant cluster after C₂ consonants /l, r/. And thirdly, /ya, wa/ may occur after implosives, nasals and other consonants that cannot be followed by /l, r/.

Verb classes that are characterised by a paradigmatically conditioned stem alternation which involves a contrast in length between <i, a, u> and <i:, a:, u:> show a corresponding contrast between <e, o> and <ya, wa>. The long vowels appear to be the result of a loss of root-final consonants, which was accompanied by changes in vowel quantity and quality.

TABLE 37.1 WAMBULE CONSONANTS

	Bilabial	Alveolar	Retroflex	Velar	Glottal
<i>Stops</i>					
Voiceless	p	t	t̪	k	?
Aspirated	ph	th	th̪	kh	
Voiced	b	d	d̪	g	
Breathy	bh	dh	d̪h	gh	
<i>Affricates</i>					
Voiceless		c (IPA ts)			
Aspirated		ch (IPA ts ^h)			
Voiced		j (IPA dz)			
Breathy		jh (IPA dz ^h)			
<i>Glottalized</i>	6	?l	d̪		
<i>Nasal</i>	m	n		ŋ	
<i>Fricative</i>		s			
<i>Trill</i>		r			
<i>Lateral</i>		l			
<i>Approximant</i>	w	y			h

TABLE 37.2 WAMBULE VOWELS

i i:	e	a a:	o	u u:
ʌi			ʌ	ʌu

Stem 1	Stem 2
$\langle i, a, u \rangle + \langle k(t), \eta, n \rangle$	$\sim \langle i:, a:, u: \rangle$
$\langle ya, wa \rangle + \langle k(t), \eta, n \rangle$	$\langle e, o \rangle$

2.4 The origin of the implosive or preglottalised consonants

Like the closely related Sunwar and Bahing languages, modern Wambule has implosive or preglottalised phonemes (Table 37.3). These sounds appear to be unrelated to the preglottalised stops that Michailovsky (1994) reconstructs for Proto-Eastern Kiranti and Proto-Thulung, and the preglottalised stops that Starostin (1994, 1994–2000) reconstructs for Proto-Kiranti.

The Wambule implosive stop /b-/ in e.g. /b̥ari/ ‘wound’, /b̥allu/ ‘fishing net’, /b̥a-/ ‘eat, bite’ and /b̥o/ ‘chicken’ corresponds to various bilabial consonants in other Kiranti languages, and in particular to Sunwar /b ~ ?w/ and Bahing /b ~ ?b/. This systematic correspondence argues in favour of the reconstruction of the consonant /*?w-/ or /*?b-/ at some stage in the phonological development of Western Kiranti. Michailovsky (1988) points out that Western Kiranti /*?b-/ may have developed from an earlier labiovelar /*kʷ-/.

In contrast to Sunwar and Bahing, Wambule has two implosive stops instead of just one, namely /b-/ and /d-/; and there is also a preglottalised lateral /?l-/. Therefore, Wambule suggests a new origin for glottalised consonants in Kiranti, i.e. modified sonorants.

Although the phonological status of /b, d/ was first demonstrated by Toba (VS 2052 (=1995)), it was Konow who was the first to remark on the systematic correspondences between the stops ‘b’ and ‘d’ in Wambule and nasals ‘m’ and ‘n’ in related languages:

B and m, d and n, respectively, are apparently interchangeable; thus *sālā-me*, young woman ; *tā-be*, daughter ; *bisi*, Dūmi *miksi*, eye ; *dōbū*, Kūlung *nōbo*, nose ; *di*, Kūlung *ning*, name ; *dwām*, Dūmi *nām*, sun, etc. It will be seen that d in the last instance corresponds to n in connected forms of speech.

Grierson (1909: 369–70)

In addition to comparative evidence, internal reconstruction of the verb also suggests that /b, d/ can be traced back to clusters of stops and nasals, at least partially. That is to say, person and number agreement suffixes with initial <m-> or <n-> in their basic form have allomorphs with initial /b- ~ b-/ or /d- ~ d-/ after verb stems that end in <-t>.

- (1) <japt-mi-me> /jaθim/ ‘they^p buy’
 <japt-nu-me> /jabdum/ ‘you^s buy’

TABLE 37.3 DEVELOPMENT OF WAMBULE IMPLOSIVE OR PREGLOTTALISED PHONEMES

Proto-Kiranti	Proto-Western Kiranti	Proto-Wambule	Modern Wambule
*kʷ-	*?w-, *?b-	*?b-	b-
*m-	*m-	*?m-	b-
*n-	*n-	*?n-	m-
?l-	?l-	?l-	n-
		*?l-	?l-
		*l-	l-

In verb forms, the glottal stop /ʔ/ can often be traced back to a first person singular subject suffix of which the basic morph has initial <ŋ>. A former preglottalised velar /*ŋ/ may have been completely substituted by /ʔ/. See also section 5.2.

Investigations by Pokharel (2003) and myself in the late 1990s have led to the identification of an ingressive or preglottalised lateral /ɺ/. According to Pokharel, the ‘ingressive lateral’ in Wambule is a newly discovered sound that is not represented in phonetic literature and hence in the International Phonetic Alphabet. On the basis of Wambule sound files that I sent to Peter Ladefoged, the latter wrote:

The files you sent me are I think fairly conclusive evidence that Wambule has what may be called a ‘glottalized lateral’, a lateral approximant that is preceded by a glottal stop or a form of creaky voice. Similar glottalized laterals (and nasals) occur in Athabaskan and Salishan languages . . .

Peter Ladefoged (personal communication,
 12 August 2005)

In conclusion, the main source of the Wambule phonemes /b, d, ɺ, ɻ/ seem to be the pre-glossed sonorants /*m, /*n, /*ŋ, /*l/ (Table 37.4).

3 MORPHOPHONOLOGY

Many morphemes in Wambule have two alternants, of which the full form ends in a consonant plus a vowel, whereas the abbreviated form ends in a consonant. Morpheme-final vowel deletion is common with free morphemes, clitics and suffixes alike and operates if the resulting word is phonologically acceptable.

- | | | |
|--|---|-----------------------|
| (2) <tico> | /tit/ | (plural/collective) |
| <C _i V C _i V> | /C _i V C _f / | |
| (3) <twa-me> | /twa-m/ | ‘that ^{up} ’ |
| <C _i C ₃ V - C _i V> | /C _i C ₃ V - C _f / | |
| (4) <ʔla=kho> | /la=k/ | ‘by hand’ |
| <C _i V = C _i V> | /C _i V = C _f / | |

There is no difference in meaning between the two forms. The abbreviated forms are unstressed morphs, whereas the full forms are stressed morphs that are used in more carefully enunciated speech. The full forms are used as citation forms here.

TABLE 37.4 WORDS CONTRASTING MODIFIED AND UNMODIFIED SONORANTS /M, N, L/

Kiranti	Wambule
/m-/	/b-/
	/m-/
/n-/	/d-/
	/n-/
/l-/	/ɺ-/
	/l-/

/bico/ ‘wife’, /bim-/ ‘remember’, /bisi/ ‘eye’
 /mama/ ‘mother’, /mi/ ‘fire’, /muyo/ ‘person’
 /di/ ‘name’, /du-/ ‘be good’, /dwam/ ‘sun’
 /nakso/ ‘priest’, /hai-/ ‘rest’, /nyam/ ‘brain’
 /ɺla/ ‘hand’, /ɺlutchu/ ‘marten’, /ɺluŋ/ ‘stone’
 /lwa-/ ‘go’, /lwak-/ ‘boil’, /lyam/ ‘tongue’

4 NOMINAL MORPHOLOGY

4.1 Non-productive markers

Wambule has several non-productive markers e.g. <-co> ‘person’ in e.g. *caco* ‘grandson’, <-ja> ‘grain’ in e.g. *garja* ‘rice plant’, <-ku> ‘water’ in e.g. *glwaku* ‘Dūdkosī river’, <-si> ‘tree, wood’ in e.g. *kuksi* ‘*Ficus cunii*’, <-si> ‘fruit’ in e.g. *twaksi* ‘mango’, <-si ~ -ci> ‘small object’ in e.g. *ywaksi* ‘salt’, <-so> ‘day’ in e.g. *saiso* ‘yesterday’, <-so> ‘flesh, meat’ in e.g. *beiso* ‘water buffalo’, <-pa ~ -po ~ -wa> ‘masculine’ in e.g. *papa* ‘father’, and <-ma ~ -mo ~ -me> ‘feminine’ in e.g. *mama* ‘mother’.

4.2 Number markers

Wambule makes a distinction between dual and plural number. There is no zero marker for singular because non-singular referents do not require the use of number markers. In addition to marking pluriformity of a referent, number markers are also used to create countable units of mass nouns and to indicate various types or varieties.

- (5) <=nimpha> dual/couple
 <=tico> plural/collective

4.3 Case markers and postpositions

Different grammatical/semantic roles are distinguished from one another by means of case markers and postpositions. Case markers have abbreviated forms that show a contraction with the word to which they are attached if the preceding word ends in a vowel. The subject of an intransitive verb, the patient and goal arguments are not marked. The agentive marker is mostly used for disambiguation, and not obligatory. The ablative is usually used together with the locative, except in a number of adverbs, where the ablative is attached directly to the adverb, e.g. *bhata=y* [up.there=ABL] ‘from up there’.

- (6) <=kho> ergative, instrumental ‘by’
 <=la> directive ‘towards’
 <=lo ~ =no> locative ‘in, to’
 <=no> comitative ‘and, with’
 <=nɔ> ablative ‘from’
 <=ŋan> genitive ‘of’
 <=se> similaritive ‘like’

Here is an example:

- (7) *Uŋ . . . am Hwalen des=no=y Wambule yor*
 I/we this Holland country=LOC=ABL Wambule word
 saykalən pa-phu ga-co, Jan-Rabet=kho
 collection do-SUP come.up-ACT Jean-Robert=ERG/INSTR
 rikwad po-u-Ø-meya.
 recording do-3npA-23s-FCT
 ‘I . . . this Dutchman who came up in order to make a collection of the Wambule language, Jean Robert made a recording.’

In contrast to case markers, postpositions do not show a contraction with the noun to which they are attached. In addition, postpositions form the basis of attachment of case markers and can be preceded by possessive pronouns.

- (8) <=swaŋkho> ‘without’
<=bi ~ =bu> ‘in the company of’
<=dambi> ‘in front of, before’
<=nʌuθa(ŋ)> ‘behind, after’
<=kha:ci> ‘in the middle of’
<=tola> ‘on top of’
<=pum> ‘beside’

There are also some loan postpositions, like e.g. <=lagi> ‘for the sake of’ and the marker <=lai> (PAT), a loan from Nepali which is used with living beings to mark patients and recipients.

4.4 Discourse markers

Discourse markers have rhetorical functions and mark inter-sentence relationships or dependencies between elements of different sentences.

- (9) <=ya> ‘also’
<=se> ‘exactly’
<=ke> ‘on the contrary’
<=ŋa> ‘only’

4.5 Pronouns

The personal pronouns are only distinguished for person, not for number or inclusivity/exclusivity. Non-singularity may be expressed by adding number markers to the abbreviated pronouns, e.g. <unj=nimpha> ‘we^d’. Case markers are attached to full pronouns, e.g. <unŋgu=kho>, which is the ergative form of ‘I, we’.

- (10) <unŋgu ~ unŋ> ‘I, we’
<unu ~ un> ‘you’
<aŋgu ~ aŋ> ‘he, she, they’

The possessive pronouns are structurally dependent upon the following constituent, which can be a noun, nominal or postposition (Table 37.5). If no overt nominal head is present, the word <blyame> ‘own, belonging to the self’ is used, e.g. <a=blyame> ‘mine, my own’.

4.6 Interrogative and indefinite words

The interrogative and indefinite words consist of a series of words with the morpheme <a-> and a series with the morpheme <tha->. Some interrogative and indefinite words contain lexicalised instances of case markers.

- (11) <acu> ‘who, somebody’
<acucu> ‘who^b, some (persons)’
<akwalo> ‘how many, much’
<ama> ‘what, something’
<amama> ‘what^b, some (things)’
<ase> ‘how, as’
<asyalo> ‘when, ever’
<asyarŋo> ‘why’

<thalo>	'where, somewhere'
<thame>	'which'
<thamsyala>	'where to, whither'

4.7 Demonstratives

The demonstrative system (Table 37.6) is based on five bound morphemes, which indicate a location relative to the point of orientation given by the context. The morphemes <a-> and <i-> also mark temporal settings. Demonstrative pronouns ('this', 'that') are formed by suffixation of the general nominalising morpheme <-me> to the bound morphemes. Demonstrative pronouns form the basis of suffixation of lexicalised instances of the nominalising <-me>, similaritive <-sya> and directive <-la> markers, e.g. <am-sya-me> 'like this' and <am-sya-la> 'hither'. Demonstrative adverbs of location ('here', 'there') are formed by suffixation of the locative marker <-lo> or the vertical markers <-na> 'same level', <-ta> 'up' and <-ya> 'down'. Demonstrative postpositions ('hither', 'thither') point to a place on a particular side of the location given by the context. Demonstrative similatives ('like this', 'like that') consist of the bound morphemes <a-> and <i-> to which the similaritive marker is suffixed.

4.8 Adjectives and numerals

Although the majority of words that denote qualities and attributes are derived from verbs (section 5.6), Wambule has a few lexical adjectives:

- (12) <blyame> 'own'
 <caku> 'other'
 <ywakka> 'small, little'

With the exception of the native word <kwalo> 'one', Nepali numerals are used in everyday conversation.

TABLE 37.5 POSSESSIVE PRONOUNS

	Singular	Dual	Plural
1st person exclusive	<a=>	<ancuk=>	<ak=>
1st person inclusive		<inci=>	<ik=>
2nd person	<i=>	<inci=>	<in=>
3rd person	<an=>	<anci=>	<an=>

TABLE 37.6 DEMONSTRATIVES

	Near	Distal	Yonder		
			Horizontal	Up	Down
Bound morphs	<a->	<i->	<hwa->	<twa->	<ywa->
Dem. pronouns	<ame>	<ime>	<hwame>	<twame>	<ywame>
Dem. adverbs	<alo>	<ilo>	<hwana>	<twata>	<ywaya>
Dem. postpositions	<=amna>	<=imna>	<=bhana>	<=bhata>	<=bhaya>
Dem. similatives	<ase>	<ise>			

4.9 Nominalisation

Nouns and adverbs can be turned into nominals by attachment of the general nominalising marker <=me>. For instance, in the following utterance, locative *hilepan=lo* ‘in Hilepānī’ is turned into a nominal, which can function as the head of a nominal phrase or as a prenominal modifier.

- (13) *hilepan=lo=m muyo*
Hilepānī=LOC=RES person
‘a person from Hilepānī’

The various types of verb nominalisation are discussed in section 5.7 later.

5 VERBAL MORPHOLOGY

The general structure of the verb is as follows:

$$(\text{NEG}) - \Sigma_{\text{main}} - (\Sigma_{\text{bound}}) - (\text{DETR}) - (\text{ARG}) - (\text{TAM/NOM}) - (\text{CASE})$$

At its minimal, a verb consists of the main root (Σ_{main}), which expresses the basic meaning of the verb. Only the negative marker <a-> (NEG) can precede the main root. The main root is the basis of suffixation for bound roots (Σ_{bound}), which mostly express that the action denoted by the main verb root involves a motion. The detransitivising morpheme <-s> (DETR) is a derivational marker of non-transitive scenarios in second person singular subject imperative forms. The core arguments of the main root are cross-referenced by means of various inflectional suffixes (ARG) that provide limited information about the person and number of the participants involved in the verbal action. A verb that consists of the main verb root and person and number agreement suffixes will be called a ‘simplex verb’. A simplex verb forms the basis of suffixation for tense/aspect/mood markers (TAM), nominalisers (NOM) and case markers (CASE).

5.1 Conjugations and verb classes

Verbs can be classified into three types: intransitive, middle and transitive. Intransitive and middle verbs show agreement with one argument: SUBJECT (S). Transitive verbs show agreement with two arguments: AGENT (A) and PATIENT (P).

Verbs may have one or more (typically two) different roots and can be classified in various conjugations on the basis of paradigmatically conditioned root alternation. Verbs that share a given conjugation can be classified into classes on the basis of shared allomorphy that affects the verb roots and the morphemes that are bound to them.

The reason to analyse the patterns of verbal morphology in terms of a combination of paradigmatically and morphophonologically conditioned root forms is related to the fact that the different forms of a verb cannot be accounted for in terms of the modification of a single verb root under the presence or absence of affixes. This is exemplified by the conjugation of the transitive verb *himcam* ‘to see’ in section 5.1.3.

5.1.1 Intransitive verbs

Intransitive verbs show agreement with one argument, which functions as the intransitive SUBJECT. Intransitive verbs can be divided into two conjugations. The verbs of the first intransitive conjugation have invariable roots that end in <a, ya, i, wa, u>, <a:, e, i:, o, u:>,

<m>, <y> or <r, l>. The verbs of the second conjugation have two roots that end in <k ~ V:, η ~ V:>; n ~ V:>. Table 37.7 gives the conjugation of the verb *lwacam* ‘to go’.

5.1.2 Middle verbs

Middle verbs show agreement with one argument, the SUBJECT, which initiates the action and is affected by it. Self-benefactive, reflexive and passive readings can be ascribed to the meaning of middle verbs. Middle verbs are characterised by a single paradigmatic root that ends <wa, u>, <a:, e, i:, o, u:>, <m>, <y> or <r, l>.

Table 37.8 gives the conjugation of the verb *ca:cam* ‘to climb’. Note that middle verbs that are inflected for person and number have a suffix <-si> (MID) which directly attaches to the main root and which only occurs in third person singular forms, where the 23s agreement morpheme is phonologically zero.

Apart from those middle verbs that do not have attested intransitive or transitive counterparts, e.g. <bal-> ‘be or get tired’, middle verb roots seem to be derived from intransitive or transitive roots by a number of processes. In some cases, the middle root is formally identical to the second root of the corresponding transitive verbs, e.g. <him-> ‘be seen’. In other cases, which seems to be the more productive way, the middle root is derived by suffixation of <-y> to an intransitive or transitive root, e.g. <pay-> ‘get dressed’. All middle verbs require the suffix <-si> (MID) in third person singular finite verb forms.

5.1.3 Transitive verbs

Transitive verbs show agreement with two arguments. The AGENT initiates the action and the PATIENT is affected by the action. Transitive verbs can be divided into three conjugations. The first transitive conjugation is characterised by one root throughout the entire paradigm which ends in <i, u>, <a:, e, i:, o, u:>, <m>, <y> or in <i, u> followed by <r, l>. The second transitive conjugations show a paradigmatically conditioned alternation between <a ~ u>, <ya ~ i>, <wa ~ u>, <ŋ ~ V:> or <a ~ u, ya ~ i, wa ~ u> followed by <l, r>. The third transitive conjugation has an alternation between <pt ~ m>, <t ~ n>, <kt ~ V:>, <m̩t ~ m>, <yt ~ y> or <rt ~ r, lt ~ l>.

TABLE 37.7 FIRST INTRANSITIVE CONJUGATION (NO ROOT ALTERNATION)

	s	d	p
1e	<i>lwa-ŋu</i>	<i>lwa-ŋ-cu-ku</i>	<i>lwa-Ø-ku</i>
1i	<i>lwa-ŋu</i>	<i>lwa-ŋ-ci</i>	<i>lwa-ki</i>
2	<i>lwa-nu</i>	<i>lwa-ŋ-ci</i>	<i>lwa-ni</i>
3	<i>lwa-Ø</i>	<i>lwa-ŋ-ci</i>	<i>lwa-mi</i>

TABLE 37.8 MIDDLE CONJUGATION (NO ROOT ALTERNATION)

	s	d	p
1e	<i>ca:-ŋu</i>	<i>ca-ŋ-cu-ku</i>	<i>ca:-Ø-ku</i>
1i	<i>ca:-ŋu</i>	<i>ca-ŋ-ci</i>	<i>ca:-ki</i>
2	<i>ca:-nu</i>	<i>ca-ŋ-ci</i>	<i>ca:-ni</i>
3	<i>ca:-si-Ø</i>	<i>ca-ŋ-ci</i>	<i>ca:-mi</i>

Table 37.9 shows the conjugation of the verb *hipcam* ‘to see’. This verb is analysed as having two different roots: a first root (Σ_1) <hipt-> and a second root (Σ_2) <him->. The table shows the paradigmatically conditioned distribution of the two verb roots and the 19 distinct agreement suffixes among the 75 transitive forms that are theoretically possible.

The first root can be seen as ‘agentive’ or ‘ergative’, at least historically, and the agreement suffixes <-ŋju> (1s) and <-nu> (2s) following the root mark an AGENT. By contrast, the second root can be seen as ‘patientive’ or ‘absolutive’, and the suffixes <-ŋju> (1s) and <-nu> (2s) mark a PATIENT. It is also worth noting that verb forms that mark a first person singular AGENT acting on a second person PATIENT have a second root followed by the agreement morpheme <-ni> (1s→2), which has an initial nasal /n/ that is normally associated with the second person. Likewise, verb forms that mark a second person singular AGENT acting on a first person PATIENT have a second root followed by the agreement morpheme <-ŋi> (2s→1), which has an initial nasal /ŋ/ that is normally associated with the first person.

The first root of verbs of the third transitive conjugation ends in a final stop <t> or in a final stop <t> that follows the consonants <p>, <k>, <m>, <y>, <r> and <l>. From a historical viewpoint, these roots seem to have been derived from other (mostly intransitive) roots by means of <-*t> suffixation. The following series of Wambule verbs consist of two related etyma. The first series contrasts an intransitive verb on the left side with a transitive verb on the right side. The initial consonants of the intransitive and transitive verbs are identical.

(14) Intransitive	Transitive
<bi-> ‘agree’	<bit- ~ bin-> ‘approve’
<blak- ~ bla:-> ‘arrive’	<blat- ~ blan-> ‘bring’
<si-> ‘die’	<syat- ~ syan-> ‘kill’
<sam->‘die’	<samt- ~ sam-> ‘lose’

In the second series, by contrast, the suffixation of <-t> is accompanied by a change in initial consonant. The transitive directive form has an aspirated/plain voiceless initial consonant, whereas the intransitive form has a voiced initial consonant at the same place of articulation.

(15) Intransitive	Transitive
<blyam-> ‘be folded’	<phlyamt- ~ phlyam-> ‘fold’
<buk-~ bu:-> ‘get up’	<phukt- ~ phu-> ‘erect, lift’
<gar-> ‘be fried’	<khart- ~ khar-> ‘fry’
<glwam-> ‘go to sleep’	<khlwamt- ~ khlwam-> ‘put to sleep’

The third series contrasts transitive verbs, of which the forms on the right side can be called ‘applicative’, since it promotes a benefactive to the PATIENT argument of a verb. The verb agrees with the semantic benefactive and not with the semantic patient, e.g. 1s→3s *kimsul hwalsum* ‘he opened the door’ (‘door’ is marked as PATIENT on the verb) vs 1s→3s *kimsul hwalyatim* ‘he opened the door for me’ (‘me’ is marked as PATIENT on the verb).

(16) Oblique transitive	Patient transitive
<pa- ~ pu-> ‘do’	<pat- ~ pan-> ‘cause to do’
<bar(s)- ~ bur-> ‘throw out’	<bart- ~ bar-> ‘throw out’
<hwal(s)-> ‘open’	<hwalt- ~ hwal-> ‘open for someone’
<phway(s)-> ‘ask for something’	<phwayt- ~ phway-> ‘ask someone’

TABLE 37.9 THIRD TRANSITIVE CONJUGATION: DISTRIBUTION OF PARADIGMATICALLY CONDITIONED ROOT FORMS

1s	1di	1de	1pi	1pe	2s	2d	2p	3s	3d	3p
1s					1 <i>him-ni</i>	2 <i>him-ni-ci</i>	3 <i>him-ni-ni</i>	4 <i>hi-bu</i>	5 <i>hi-bu-ci-m</i>	6 <i>hi-bu-m</i>
1di					10	11 <i>hip-sa</i>	12	13	14 <i>him-eu-k</i>	15
1de								16	17 <i>hip-i</i>	18
1pi					19	20 <i>hip-sa</i>	21	22	23 <i>him-O-ku</i>	24
1pe								34 <i>hib-du</i>	35 <i>hip-du-ci-m</i>	36 <i>hib-du-m</i>
2s	25 <i>him-yi</i>							37	38 <i>him-ci</i>	39
2d	26 <i>him-yi-ci</i>							40	41 <i>him-ni</i>	42
2p	27 <i>him-yi-ni</i>									
3s	43 <i>him-yati</i>	44 <i>hipt-u-O</i>	45	46 <i>hipt-u-O</i>	47	48 <i>him-nati</i>	49	50	51 <i>hipt-u-O</i>	52
3d	54 <i>him-yu-cl-m</i>	55 <i>hipt-u-ci</i>	56 <i>him-ki-ci-m</i>	57 <i>hipt-u-ci</i>	58 <i>him-ki-m</i>	59 <i>him-nu-ci-m</i>	60 <i>him-ci-m</i>	61 <i>him-ni-m</i>	62 <i>hipt-u-ci</i>	63 <i>him-ci-m</i>
3p	65 <i>him-yu-m</i>	66 <i>hi-bi</i>	67	68 <i>hi-bi</i>	69	70 <i>him-nu-m</i>	71	72	73 <i>hi-bi</i>	74

5.2 Person and number agreement morphemes

Table 37.10 gives the inventory of the basic morphs of the person and number agreement suffixes and their positions in the suffixal string.

5.3 Temporal, aspectual and modal inflections

5.3.1 Simplex verb

A verb that just consists of the main root and the agreement suffixes listed in Table 37.10 is called a ‘simplex verb’. A simplex verb can be used as an exhortative to express a proposal for action to be taken by both the speaker and the hearer.

- (17) *di-ki!*

go-1NSPS

‘Let us^{di} go!'

A simplex verb is also used as a deliberative to elicit the hearer’s advice or assent, or in rhetorical questions, in which an answer is not expected.

- (18) *lwa-ŋu?*

go-1S

‘Am I to go?’

- (19) *acu=k khwakt-u-Ø?*

who=ERG/INSTR cook-3NPA-23S

‘Who will cook?’

5.3.2 Affirmative

The affirmative (AFF) is formed by suffixation of <-me> to a simplex verb. The affirmative is used in positive statements and yes–no questions. The affirmative marks dynamic

TABLE 37.10 BASIC MORPHS OF THE SIMPLEX PERSON AND NUMBER AGREEMENT SUFFIXES

	Sf1	Sf2	Sf3
Σ	<-i> (1pi→3)		
	<-ki> (1NSPS)		
	<-mu> (3→3p)	<-Ø ¹ > (23s)	
	<-nati> (3s→2s)	<-Ø ² > (1peAS)	<-ku> (1nseAS)
	<-ni ¹ > (1s→2)	<-ci> (d)w	<-mi> (3/ns)
	<-nu> (2s)	<-cu> (1deAS)	
	<-ŋ> (ds)		
	<-ŋati> (3s→1s)	<-ni ² > (2p)	
	<-ŋi> (2→1s)		
	<-ŋu> (1s)		
	<-u> (3NPA)		

Note that, in Table 37.9, the morphemes <-nu> (2s), <-ŋu> (1s) and <-mi> (3/ns) are realised as /-du/, /-bu/ and /-bi/ after the first verb root <pt>. In addition, the morpheme <-ku> (1nseAS) is generally realised as /-k/ after <-cu>, and the morpheme <-mi> (3/ns) is /-m/ after other agreement suffixes.

actions regardless of the situation in time and omnitemporal statements. The suffix <-me> is absent in negative verb forms, which take the negative prefix <a-> instead.

- (20) *lwa-ŋ-ci-m*
 go-ds-d-AFF
 ‘You^d go/went/will go’
- (21) *a-lwa-ŋ-ci*
 NEG-go-ds-d
 ‘You^d did/do/will not go’

5.3.3 Indefinitive

The indefinite (IND) is formed by suffixation of <-ya> to a simplex verb. The indefinite occurs in clauses with an interrogative and indefinite word. In main clauses, the indefinite is used in questions and exclamations. Subordinate correlative clauses translate as ‘whomever’, ‘whatever’, ‘wherever’, etc.

- (22) *thal lwa-Ø-i?*
 where go-23s-IND
 ‘Where did he go to?’, ‘Where will he go to?’
- (23) *ama de-m-ya!*
 what say-3/ns-IND
 ‘What do they^p say!?’
- (24) *thal ba:-nu-ya, il . . .*
 where be/sit-2s-IND there
 ‘Wherever you^s are, there . . .’

5.3.4 Optative

The optative (OPT) is formed by suffixation of <-kano> to a simplex verb. The optative is used to express wishes and hopes.

- (25) *dum-nu-kan!*
 become-2s-OPT
 ‘May you^s become!’

5.3.5 Volitive

The volitive (VOL) is formed by suffixation of <-ma> to a simplex verb. The volitive expresses the speaker’s will or determination to do things. Volitive verbs function as first person imperatives.

- (26) *uy wak-ti gwa-ʔu-ma*
 I/we scream-CNN give-1s-VOL
 ‘I will shout at him.’

5.3.6 Factitive, or factual participle

Suffixation of <-meya> to a simplex verb yields a factitive (FCT). The factitive, or factual participle, is a verb nominalisation that refers both to the state and object brought about

by the situation expressed by the verb. The factitive can be used as a nominal argument (complement clause), a prenominal modifier (relative clause) or as the main predicator of an entire clause. Predicatively used factitives usually mark past time, completed action or events that the speaker asserts to have happened and hold true. In the case of prenominal modifiers, the head nouns are co-referential with the SUBJECT or PATIENT of the modified verb. Example (7) earlier illustrates a predicatively used factitive.

- (27) *un bla:-nu-meи akwal din dum-Ø-ya?*
 you arrive-2s-FCT how.much day be/become-23s-IND
 ‘How many days ago did you^s arrive?’
- (28) *hwal-si-Ø-meи kimsul thyak-Ø-ta!*
 open-MID-23s-FCT door close-IMP-SAS
 ‘Close^s the door that opened itself!’
- (29) *uy . . . saiso ywam bhai*
 I/we yesterday that_{down} below
dojar phi-b-beи thiyo.
 bulldozer bring^{hrz}-3/ns-FCT was
 ‘I . . . yesterday they^p fetched a bulldozer down there.’

5.3.7 Imperative

The imperative is formed by suffixation of a distinct set of agreement morphemes to verb roots, which are distributed according to a fixed pattern of paradigmatically conditioned root alternation. See Table 37.11.

The imperative is used to express requests and orders. Negative imperatives are used to express prohibition and denial of permission. Some intransitive verbs take the suffix <-s> (DETR) after the verb stem, which is perhaps the same entity as the middle marker <-s> (MID).

- (30) *ilo=η=qa, 'lu, ca-s-Ø-ta!*
 there=ABL=THM all.right climb-DETR-IMP-SAS
ca:-ka-ne! *way-ki-m-me.*
 climb-IMP-23p say.to-1NSPS-3/ns-AFF
 ‘And then he said to us^{pe}, “Well, get^s in! Get^p in!”’
- (31) *unu=k is a-pu-no!*
 you=ERG/INSTR in.that.way NEG-do.self-NEG.SAS
 ‘Do^s not behave like that!’

TABLE 37.11 BASIC MORPHS OF THE IMPERATIVE PERSON AND NUMBER AGREEMENT SUFFIXES

	Sf1	Sf2	Sf3	Sf3
	<-mi> (s→3ns)			
Σ	<-no> (NEG.SAS)	<-ka> (IMP)	<-ta> (SAS)	<-ce> (23d)
	<-ŋ> (ds)		<-ŋ> (1sp)	<-ne> (23p)
	<-sa> (IMP+1ep)			

5.4 Compound verbs

The main root forms a basis of suffixation for several bound roots that express that the action denoted by the main verb root involves a vertical motion or the transportation to or away from the speaker. These bound roots are specified for the vertical level of the place or direction given by the context.

- (32) <-phi-> ‘come or bring (across a horizontal plane)’
 <-kha-> ‘come or bring up’
 <-ywa-> ‘come or bring down’
 <-di- -du-> ‘go or take (and come back)’
 <-lwa-> ‘go or take’

In addition to these bound roots which are unmarked for transitivity and which can be used in relation to intransitive, middle and transitive forms alike, Wambule possesses a special group of ‘middle’ forms, which contain the middle marker <-y>, e.g. <-khay-> ‘come or bring oneself up’.

5.5 Causatives

The causative is an analytical construction that involves a modified form of the main verb called ‘preverb’ plus the causative auxiliary verb *paccam* ‘to cause to do’, which is the directive or applicative form of *pacam* ‘to do’. The causer is marked as the AGENT of the verb and the causee as its PATIENT. The causee functions as the SUBJECT of an associated intransitive verb or as the AGENT of a transitive verb.

- (33) *pro-ki-m*
 run-1NSPS-AFF
 ‘We^{pi} ran’
- (34) *pro pay-ki-m-me*
 run cause.to.do-1NSPS-3/ns-AFF
 ‘He made us^{pe} run’

5.6 Reciprocals

Reciprocals are analytic constructions which are composed of a preverb followed by the reflexive auxiliary *pucam*, which is conjugated as a middle verb. Reciprocals are conceptually similar to reflexives in the sense that both indicate that the AGENT and PATIENT are co-referential, but here the two participants act equally upon each other rather than one participant acts upon himself.

- (35) *phapu ga pu-lwa-m-mwa-l* . . .
 bamboo give act-go/take-3/ns-CTT-PRG
 ‘While they^p were handing over the bamboo to each other . . .’

5.7 Verb nominalisation

One of the characteristics of Wambule as a Tibeto-Burman language is its widespread use of nominalisations; Wambule has various nominalisers with distinctive functions. With the exception of the factitive marker <-meya>, discussed in section 5.3.6 earlier, all the nominalising markers attach to verb roots to which no agreement markers are suffixed.

5.7.1 Infinitive

The infinitive is formed by suffixation of <-si> to the verb root. The infinitive refers to the action expressed by the verb root in a more general way. The infinitive is also used as the main verb with auxiliaries such as e.g. *jyakcam* ‘to start’, *plyacam* ‘to quit’, *tumcam* ‘to finish’, which denote the beginning, stopping and ending phases of an event.

- (36) *un ja-si tum-ci-m.*
you eat-INF finish-d-AFF
'you^d finished eating'

5.7.2 Nomen actionis

The *nomen actionis* is formed by suffixation of <-wa> (NAC) to an infinitive. This nominal refers to the entire action expressed by the verb root in a more manner-like way.

- (37) *Aj=cwak-si-wa ran-co.*
his/her=carve-INF-NAC seem.well-ACT
'His (way of) carving is good.'

5.7.3 Supine

The supine is formed by suffixation of <-phu> to the verb root. The supine refers to the purpose of the action expressed by the verb root. See example (7).

5.7.4 Nominaliser of loan verbs

The nominaliser of loan verbs is formed by suffixation of <-ai> (LN) to a Nepali verb root. These loan nominals are used with the auxiliary *pacam* ‘to do, make’ to form transitive verbs and with *dumcam* ‘to be, become’ to form intransitive verbs.

- (38) *am khwala tar-ai dum-si capt-u-Ø-m.*
this river cross-LN become-INF can-3npa-23s-AFF
'He can cross this river.'

5.7.5 Active participle of specific events

The active participle of specific events is formed by suffixation of <-co> (ACT) to the verb root. This form marks the SUBJECT or AGENT of the modified verb, and refers to a specific event or a single action on a particular occasion. See examples (7) and (37).

- (39) *Im dukso=lo=ŋ ay=dî Ghwaqa*
that day=LOC=ABL his/her=name horse
Yoc-co o-b-bei, dyam.
throw-ACT say.to-3/ns-FCT HRS
'And from that day onward they^p called him The Horse Thrower.'

In this example, the word *dyam* indicates that the information known to the speaker is second-hand, conveying the meaning ‘the people say’ or ‘it is said’. It is repeated continually throughout myths and stories, where it indicates that what is told is generally known in the society.

5.7.6 Active participle of characteristic events

Suffixation of <-ce> (ACT) to a verb root yields the active participle of characteristic events. This nominalisation marks the SUBJECT or AGENT of the modified verb, with reference to characteristic activities.

- (40) *Kwal mufhi sera gwak-Ø-ta,*
one handful husked.grain give-IMP-SAS
im phwai-ce=lai!
that ask.for-^{ACT}=PAT
'Give^s one handful of husked grain to that beggar!'

5.7.7 Passive participle

Suffixing <-munco> (PAS) to the verb root yields a passive participle. This form refers to a non-agentive argument, often a PATIENT, of the action described by the modified verb. The passive participle can be derived from transitive, intransitive and middle verbs. The intransitive passive is a non-agentive form which seems to correspond to the Dutch intransitive passive in e.g. *er wordt gedanst* 'there is dancing'. In the following example, the relativised noun functions as the PATIENT of the modified verb.

- (41) *hwal-mumco kimsul thyak-Ø-ta!*
open-PAS door close-IMP-SAS
'Close^s the door which (someone) opened!'

5.7.8 Attributive participle

The attributive participle is formed by adding <-ji> (AP) and the gender markers <-wa> (MAS) or <-mo> (FEM) to the verb root. The attributive participle refers to the SUBJECT or AGENT of the modified verb and describes a trait or characteristic.

- (42) *jwa-ji-mo* 'female glutton'
de-ji-wa 'talkative male'

5.7.9 Stative participle

The stative participle is formed by suffixation of <-bo> to the root of stative verbs. This morphological process appears to be non-productive. There are a few stative participles and only about half of them have a well-attested verbal base.

- (43) *khwalbo* 'big'
robo 'high, tall'

5.7.10 Purposive participle

Adding <-ca> (PUR) to a verb root forms the purposive participle. This participle refers to the intended application or purpose, including an instrument that is used to accomplish the action represented by the verb. The participle is usually extended by means of the nominaliser <-me> (RES). These extended forms constitute the usual citation form of verbs.

- (44) *lwa-ca-me* 'to go'
glwam-ca-m 'to sleep'

6 SYNTAX AND CLAUSE COMBINATIONS

6.1 Constituent order

The grammatical marking strategy known as ‘constituent order’ is not very important, since Wambule has no rigid order. Yet, there seems to be some kind of canonical order as the verb typically occurs in clause-final position. The most frequent pattern is with the AGENT or SUBJECT as topic.

	SUBJECT	verb
AGENT	PATIENT	verb

6.2 Relative and complement clauses

Relative clauses are clauses which depend on a deverbal form and are placed before the nominal head. Complement clauses are clauses which depend on a factitive and function as an argument of another verb. See sections 5.3.6 and 5.7 earlier.

6.3 Adverbial clauses

Many adverbial clauses are formed on the basis of verbs. In some cases, these verbs are marked with suffixes that are at least historically related to the case markers discussed in section 4.3. I am unsure whether they represent the same markers synchronically, since there are morphosyntactic differences: the case markers are clitics that attach to nouns whereas the verbal markers are suffixes.

- (45) <-kho> perfect ‘after . . .’ (PFG) (<ergative)
<-lo ~ -no> present ‘when . . .’ (PRG) (<locative)
<-ŋo> conditional ‘if . . .’ (CND) (<abative)
<-se> ‘like . . .’ (SIM) (<similaritive)

Here is an example:

- (46) *aygu* *kathmanqu* *di-ca-m* *dum-Ø-mwa-l*
he/they Kathmandu go/come-PUR-RES become-23s-CTT-PRG
hal-sa-k *lwa-ca-m* *thiyo,* *dyam.*
walk-MAN-PFG go-PUR-RES was HRS

‘When it was time to go to Kathmandu, he had to go on foot, it is said.’

In addition, there are some suffixes that have no related case marker:

- (47) <-thu> negative state ‘without . . .’
<-ti> connective ‘and . . .’
<-sʌi> simultaneous ‘whilst . . .’
<-sa> manner ‘while . . .’
<-li> motion ‘moving here and there’

Finally, Wambule also has a few conjunctive particles which are loans from Nepali, e.g. *athaba* ‘or, or else’, *ki* ‘or’, *jaba* ‘when’, *ya* ‘or’.

ABBREVIATIONS

- transitive relationship
- affix boundary

=	clitic boundary
1	first person
2	second person
3	third person
A	AGENT
ABL	ablative
ACT	active participle
AFF	affirmative
AP	attributive participle
ARG	agreement markers
C _i	initial consonant
C ₂	second consonant in cluster
C ₃	third consonant in cluster
C _f	final consonant
CASE	case markers
CND	conditional gerund
CNN	connective gerund
CTT	contemporaneous time
d, ^d	dual
dem.	demonstrative
DETR	detransitive
e, ^e	exclusive
ERG/INSTR	ergative/instrumental
FCT	factive
FEM	feminine
HRS	hearsay
hrz	horizontal
i, ⁱ	inclusive
IMP	imperative
IND	indefinitive
INF	infinitive
IPA	International Phonetic Alphabet
LN	nominaliser of loan verbs
LOC	locative
MAN	gerund of manner
MAS	masculine
MID	middle
NAC	<i>nomen actionis</i>
NEG	negative
Nep.	Nepali
NOM	nominaliser
np	non-plural
ns	non-singular
OPT	optative
P	PATIENT
p, ^p	plural
PAS	passive participle
PAT	patient marker
PFG	perfect gerund
PRG	present gerund

PUR	purposive participle
RES	reifier (general nominaliser)
S	SUBJECT
S, ^s	singular
sf	suffixal slot
SIM	similaritive gerund
SUP	supine
TAM	tense/aspect/mood markers
THM	theme
V	vowel
VOL	volitive
Σ ₁	first verb root
Σ ₂	second verb root
Σ _{main}	main verb root
Σ _{bound}	bound verb root

BIBLIOGRAPHY

- Grierson, George A. 1909. *Linguistic Survey of India*. Vol. III. Tibeto-Burman family. Part I. General introduction, Specimens of the Tibetan dialects, The Himalayan dialects, and the North Assam group. Calcutta: Government of India [Reprinted: Delhi, Varanasi, Patna. Motilal Banarsi das 1967. xxiii. 641 pp].
- Hanbon, Gerd. 1991. *The Rai of Eastern Nepal: Ethnic and Linguistic Grouping. Findings of the Linguistic Survey of Nepal*. Kirtipur/Kathmandu: Linguistic Survey of Nepal and Centre for Nepal and Asian Studies, Tribhuvan University [Edited and provided with an introduction by Werner Winter].
- Hodgson, Brian Houghton. 1857. Comparative vocabulary of the languages of the broken tribes of Népal. *Journal of the Asiatic Society of Bengal* 26: 333–371.
- Michailovsky, Boyd. 1988. Phonological typology of Nepal languages. *Linguistics of the Tibeto-Burman Area* 11.2: 25–50.
- Michailovsky, Boyd. 1994. Manner vs. place of articulation in the Kiranti initial stops. In: Hajime Kitamura, Tatsuo Nishida and Yasuhiko Nagano (eds) *Current Issues in Sino-Tibetan Linguistics*. Osaka: Organizing Committee of the 26th International Conference on Sino-Tibetan Languages and Linguistics, 59–78.
- Opgenort, Jean Robert. 2002. The Wāmbule language. Doctoral dissertation, Leiden University.
- Opgenort, Jean Robert. 2004a. Implosive and preglottalized stops in Kiranti. *Linguistics of the Tibeto-Burman Area* 27.1: 1–27.
- Opgenort, Jean Robert. 2004b. *A Grammar of Wambule: Grammar, Lexicon, Texts and Cultural Survey of a Kiranti Tribe of Eastern Nepal*. Leiden: Brill.
- Pokharel, Madhav. 2003. Ingressive lateral in Umbule Rai. Paper presented at the 36th International Conference on Sino-Tibetan Languages and Linguistics at La Trobe University Melbourne, Australia.
- Starostin, Sergei. 1994. The reconstruction of Proto-Kiranti. Paper presented at the 27th International Conference of Sino-Tibetan Languages and Linguistics, Paris, 11–15 October 1994.
- Starostin, Sergei. 1994–2000. Databases. Available at: starling.rinet.ru/cgi-bin/main.cgi?flags=eygttnl (accessed 9 October 2005).
- Toba, Sueyoshi. VS 2052 (=1995). Implosive stops in Umbule Rai. *Libju-Bhumju* 3: 7–9.

§3.4.4 *Western Himalayan*

CHAPTER THIRTY-EIGHT

SANGLA KINNAURI¹

Anju Saxena

1 INTRODUCTION

Sangla Kinnauri is a Tibeto-Burman language spoken in the Sangla tahsil (Kinnaur district, Himachal Pradesh state, India), with the exception of Rakcham and Chitkul villages. It is subsumed under what is usually referred to as Kinnauri in the literature, although it has been called variously in different sources, e.g.: “Milchan” (Gerard 1841), “Kunawar” (Gerard 1842), “Kanaawari” (Konow 1905), “Kanawari” (Joshi 1909), “(Lower) Kanauri” (Bailey 1908, 1909, 1910, 1911, 1920, 1938), “Kanooring skad”/“Kanooreeanu skad” (Bailey 1909), and “Kinnauri” (Sharma 1988; Saxena 1992, 1995).²

2 PHONOLOGY

2.1 Consonants

There are 31 consonant phonemes in Sangla Kinnauri (Table 38.1).

h occurs only word-initially (e.g. *haray* ‘bone,’ *hon* ‘now’). Recognizably Indic words with non-initial *h* are regularly pronounced without *h* in Sangla Kinnauri, e.g. *mel* ‘palace,’ *braman* ‘priest.’ *r* is occasionally heard immediately after *t* and *tʰ* when there is a vowel (except *e*) following. This is indicated in the examples as *t(r)/tʰ(r)*.

b, *d* and *g* are consistently realized as voiced stops word-initially. In other positions, they are often realized as voiced fricatives (*sub* ~ *suf* ‘foam’; *tʰ(r)og* ~ *tʰ(r)oy* ‘white’; *baymɔd* ~ *baymɔð* ‘footprint’; *t(r)əgmu* ~ *t(r)əymu* ‘to break’). In all these cases the language consultants perceived these consonants as stops. In fast speech the word-final stops were realized, at times, as voiceless (e.g. *tag* ~ *tak* ‘pus’; *motag* ~ *motak* ‘healthy’). Finally, in some cases the duration of the word-final stop is very short, although the hearer can still identify the consonant. This is indicated in the transcription used here as unreleased stops, e.g. *jomed* ‘mother-in-law.’

There are some minimal pairs involving a contrast in consonant length, (e.g. *rɔn* ‘iron(N),’ *rɔnn* ‘a kind of grey-colored lentil’; *gɔl* ‘wooden roof,’ *gɔll* ‘summer’), although they are extremely rare. Long consonants occur only at the end of lexical morphemes.

2.2 Vowels

The vowel phonemes are provided in Table 38.2. Vowel length is phonemic (examples of minimal pairs: *ray* ‘horse,’ *ra:y* ‘mountain top’; *ri* ‘day before yesterday,’ *ri:* ‘pine nut’³). Although vowels preceding a nasal consonant are normally nasalized, nasalization is not phonemic. In some cases, there is a variant pronunciation without the nasal consonant but

TABLE 38.1 CONSONANT PHONEMES¹

	Bilabial	Alveolar	Retroflex	Palatal	Velar	Glottal
Plosive	p b	t d	t̪ d̪		k g	
Aspirated	p ^h	t ^h	t̪ ^h		k ^h	
Nasal	m	n		j̪	ŋ	
Fricative	s z		ʂ	ʃ ʒ		
Affricate	tʂ tʂ ^h			fʂ fʂ ^h dʐ		
Lateral	l					
Trill	r					
Approximant	v			j		

¹ The following consonants do not appear in word-final position in my material: h, p^h, t^h, k^h, j̪, tʂ^h, fʂ^h and v.

TABLE 38.2 VOWEL PHONEMES

i, i:	u, u:
e, e:	o, o:
a, a:	

where the vowel retains nasalization (e.g. *rāŋdəless* ~ *rāŋdəless* ‘widower’), and in a small number of words (e.g. *ãdʒ* ‘male goat’; *ajāraj* ‘dark’), I have not been able to find a variant with a nasal consonant.

Some linguistic variation is found in Sangla Kinnauri. Thus some lexical items as well as affixes vary with respect to their vowels (e.g. *kʰjamu* ~ *kʰimu* ‘to see’; *bjɔmu* ~ *bimu* ‘to go’; *-is* ~ *-es* [-ERG]). More generally, the phonetic realization of vowel phonemes varies both within the speech of an individual and across speakers: *i* is realized along the entire spectrum of [i]–[ɪ]. Similarly, *u*: [u]–[ʊ], *e* [e]–[ɛ], *o* [o]–[ɔ] and *a* [a]–[ɑ]–[ə]–[a].

2.3 Morphophonological stem alternations

There are some stem alternations recurring in several places in the nominal inflectional system, triggered by particular suffixes.

Disyllabic stem truncation: As we will see, when certain inflectional suffixes are added to a disyllabic noun stem ending in *-ay*, *-iŋ* or *-es*, this final part of the stem is replaced by the inflectional suffix.

Final vowel elision: When certain vowel-initial inflectional suffixes are added to a disyllabic or polysyllabic stem ending in *-a* or *-e*, the stem-final vowel is deleted (stems in *-a*) or replaced by a high glide (stems in *-e*). This is normally accompanied by a lowering of the suffix vowel (*-u* > *-o*).

3 NOUN PHRASE

3.1 Noun phrase structure

The noun phrase has the following basic structure:

(DEM) (NUMERAL) ((ADV) ADJ) N-(DIM)-(PL)-(CASE) (EMPH)

- (1) *ds tif val(-i)* *gato-ts* *f^hay-ɔ*
 dem seven much(-EMPH.PRECISELY) small-DIM child-PL
 ‘Those seven very small children’

In some discourse contexts the emphatic marker may precede the locative case marker. The N can consist of a title plus a name. In such cases both orders are possible (e.g. *dafrat^h raza* [PN king] ~ *raza dafrat^h* [king PN] ‘king Dashrath’).

3.2 The components of the noun phrase

3.2.1 Nouns

A large part of the core vocabulary is made up of Indo-Aryan lexical items. Most nouns are monosyllabic or disyllabic. Disyllabic nouns often end in *ay*, *ŋ* or *ts* (e.g.. *dəjəŋ* ‘body’; *kətŋ* ‘basket’; *pələts* ‘blood’). Nouns are also formed by compounding (e.g. [N+N]: *gətəŋhɔŋ* ‘snail(N)’ [*gətəŋ* ‘grindstone’ + *hɔŋ* ‘worm’]; *pʰɔŋ/a* [*pʰo* ‘deer’ + */a* ‘meat’] ‘deer meat’; [ADJ+N]: *rɔkmig* [*rɔk* ‘black’ + *mig* ‘eye’] ‘pupil’). Formal word structure does not distinguish count and mass nouns, or inanimate, animate and human nouns.

Marking for inflectional features on the noun is exclusively suffixing.

3.2.1.1 Number

Generally, a two-way number distinction—singular vs plural—is made in nouns (but see 3.2.2 for some instances of duality). The distribution of the regular plural markers on nouns (-(*j*)*a* and -(*g*)*ɔ*) is not completely systematic, but some tendencies are observable. -(*j*)*a* triggers disyllabic truncation (e.g. *boʈʰaŋ* ‘tree,’ *boʈʰ-a* [tree-PL]; *sapəs* ‘snake,’ *sap-a* [snake-PL]). -*ja* normally occurs after a vowel and -*a* after a consonant. In other cases, the most frequent plural marker is -(*g*)*ɔ*, where -*gɔ* occurs after a stem-final vowel and -*ɔ* after a stem-final consonant (e.g. *gone* ‘wife,’ *gone-gɔ* [wife-PL]; *tʰar* ‘leopard,’ *tʰar-ɔ* [leopard-PL]). In a restricted set of nouns the plural marker is -*e* (e.g. *rɔŋ-e* [bread-PL]; *ha:p-e* [wolf-PL]; *gambu:t-e* [boot-PL]). The plural marker -*e* also occurs with the numeral *id* ‘one,’ forming a generic pronoun (2).

- (2) *id-e-s aŋ f^hay-ts du*
 one-PL-ERG 1sg.NNOM boy-DIM tell-PROG be.PRES
 ‘Some are saying: “(You are) my son.”.’

In a few nouns, the stem-final vowel is lengthened to mark plurality. In many such instances the plural marker -*gɔ* is also acceptable (e.g. *ate:* ~ *ate-gɔ* [o.brother-PL]). The plural marker may also optionally appear in a noun phrase which contains a numeral (3).

- (3) *tif sapəs f^hay-ts-ɔ*
 seven snake child-DIM-PL
 ‘Seven little snakes’

3.2.1.2 Gender

Gender is not a grammatical category in Sangla Kinnauri. However, there are word-formation devices for creating nouns denoting female and male humans and animals. These are generally suffixes (e.g. *zo* ‘mountain bull,’ *zɔ-mo* ‘mountain

cow'; *larɔ* ‘bridegroom,’ *lari* ‘bride’), but there are also two prefixes (or first members of compounds: (*s*)*kjo-ray* [male-horse] ‘stallion’; *mənt(r)-ray* [female-horse] ‘mare’). None of these processes are productive. The following two almost-grammatical processes are, however, productive. The forms of the agentive nominalizer (-*tja/-tse*) and the contrastive specifier (-*sja/-se*)⁴ encode a gender distinction: *-tse* and *-se* occur with female referents and *-tja* and *-sja* with male referents, e.g.: *gas-ɔ ſi-tse* [clothing-PL wash-NOMZR.F] ‘clothes washer (female),’ *gas-ɔ ſi-tja* [clothing-PL wash-NOMZR.M] ‘clothes washer (male); *ʃʰad-sja* [son-in-law-CONT.SPECIFIER.M], *ama-ts-se* [mother-DIM-CONT.SPECIFIER.F].

3.2.1.3 Case markers

All case markers except the nominative are suffixes (or clitics). They are generally agglutinated to the stem, with some exceptions described in section 2.3 and noted below.

3.2.1.3.1 NOMINATIVE

The nominative form is the stem of a noun or pronoun without any other case suffixes. This form can be used for subjects (intransitive and transitive)—i.e. the NP triggering subject agreement in the verb—and direct objects.

3.2.1.3.2 ERGATIVE: -(i)s

-(*i*)*s* functions as the ergative and instrumental marker. *-s* occurs with stems ending in a vowel and *-is* with consonant-final stems. The ergative marker occurs only on the subject of transitive verbs, but its occurrence is not obligatory. In narratives, the ergative marker occurs almost obligatorily on the subject of the matrix clause of direct-speech statements (“he said”). The occurrence of the ergative marker here can be seen as a deictic marker which draws the listener’s attention to the change in the mode of narration—from the descriptive to the expressive mode. Similarly, the ergative marker in the non-*he said* constructions in narratives occurs regularly in situations where the clause describes something which runs counter to expected behavior (including social norms). The ergative marker functions as a linguistic tool to describe a shift in perspective (Saxena forthcoming).

3.2.1.3.3 DATIVE: -*u/-o*, (-*pəŋ* AND -*n(u)*

With nouns in the singular, *-u/-o* occurs⁵ predominantly with stems ending in a consonant and (-)*pəŋ* with stems ending in a vowel, while *-n(u)* occurs only in the plural. The dative suffixes never trigger disyllabic stem truncation. With personal pronouns their distribution is as follows: *-u/-o* occurs with 1SG, 1DU, 2DU.HON, 2DU.NHON. *-n(u)* occurs with 1PL.INCL, 1 PL.EXCL, 2SG.NHON, 2SG.HON, 2PL.HON, 2PL.NHON and 3PL. With 3SG, both (-)*pəŋ* and *-u/-o* are found.

As is the case with many South Asian languages, Sangla Kinnauri, too, has the so-called dative experiencer construction.⁶ The dative marked argument occurs with non-volitional verbs, such as *pərennu* ‘to find’; *gjamu* ‘to like, to want’; *tsalmu* ‘to feel’; *tʰasmu* ‘to hear.’ It occurs, for example, in constructions which describe bodily conditions and emotional states (4) and in the obligative construction (5).

- (4) *ay-u* *əkʰa* *tɔ*
 1SG.NNOM-DAT pain be.PRES
 ‘I have pain.’

- (5) *dɔ-pəŋ dʒəŋ ma-bə-n gja-mig du-ge*
 3SG-DAT here NEG-come-NOMZR want-NOMZR be-PST
 ‘He should not have come here.’

3.2.1.3.4 GENITIVE: *-u/-o* AND *-(n)u*

-u/-o occurs⁷ in the singular and *-n(u)* in the plural (*-nu* and *-n* are interchangeable). The genitive singular suffix *-u/-o* optionally triggers disyllabic stem truncation, being realized as *-o* in this case (e.g. *botʰaj* ‘tree,’ *botʰ-o* ‘tree-GEN’). It also triggers final vowel elision (e.g. *ate* ‘older brother,’ *atj-o* [o.brother-GEN]; *rikʰa* ‘bear,’ *rikʰ-o* [bear-GEN]).

3.2.1.3.5 LOCATIVE: *-o*, *-n(o)*, AND *-r*

-o and *-n(o)* are productive, *-o* occurring in the singular and *-n(o)* in the plural (*-no* and *-n* are interchangeable). *-r* occurs only with demonstrative pronouns. The locative singular marker *-o* optionally triggers disyllabic stem truncation (e.g. *botʰaj* ‘tree,’ *botʰaj-o* ~ *botʰ-ɔ* [tree-LOC]). It also triggers final vowel elision (e.g. *bagitsa* ‘garden,’ *bagits-ɔ* [garden-LOC]). With other vowel-final stems, the stem vowel is retained (e.g. *tsaku-o* [knife-LOC]). In fast speech, in noun stems ending in *-o*, one does not always hear both the stem-final vowel and the locative marker (e.g. *dörko-o* [skeleton-LOC]). Nouns in the locative are sometimes lexicalized into adverbs. For example, *djar-o* [day-LOC] ‘daily.’

3.2.1.3.6 COMITATIVE: *(-)ray*

(-)ray expresses comitative and associative. It occurs with nouns representing human, animate, and inanimate referents. *(-)ray* also forms a coordinate construction with the structure: N(*-ray*) N((*-*)case marker) (6).

- (6) *june-ray gɔlsay-o dəŋ krab-o krab-o*
 sun-COMIT moon-GEN near cry-PROG cry-PROG
 ‘To Sun and Moon, (she) is crying (=complaining), (she is) crying (=complaining)’

Finally, *(-)ray* also follows the verb in non-final clauses, either with a nominalized verb form or where the verb is immediately followed by the manner marker *-(j)e* (7). Such non-final clauses have an adverbial interpretation.

3.2.1.3.7 ABLATIVE: *-(ə)yf*

The normal form of the ablative suffix is *-yf*, with *-əyf* sometimes appearing after a consonant. *-(ə)yf* occurs in the following structures: N-ABL, N-LOC-ABL and N-GEN *dɔk*-ABL. N-ABL and N-LOC-ABL occur only with nouns representing non-human referents, where N-LOC-ABL occurs with nouns whose referents are physically or conceptually viewed as delimited, with clearly defined boundaries; N-ABL occurs elsewhere. N-GEN *dɔk*-ABL occurs only with nouns representing human referents.

3.2.1.3.8 MANNER: *-(j)e*

-(j)e expresses manner or circumstance (e.g. *bal-e* [head-manner] ‘first in the queue’). *-e* occurs after consonants and *-je* after vowels. It is also suffixed to verbs, yielding non-final adverbial clauses (7). In many—though not all—such constructions, the comitative marker *(-)ray* follows the non-final verb with *-(j)e*.

- (7) *gas-ɔ yfi-je ray id-es tsʰetsats-u lɔd-o*
 clothing-PL wash-MANNER COMIT one-ERG girl-DAT say-PROG
 ‘At the time of washing (their) clothes, one (woman) is saying to the girl.’

3.2.2 Personal and demonstrative pronouns

Distinct from nouns, personal pronouns (see Table 38.3) also indicate dual number: *kifaj* functions as the first person dual pronoun. *-suj* marks the dual on third person pronouns (e.g. *nok-suj* ‘those two (who are in sight)’) and also emphasizes togetherness.⁸ The numeral *nif* ‘two’ occurs occasionally after the second and third person pronouns to indicate duality (e.g. *kanif* ‘you two’). With first person reference, *nif-i* [two-EMPH.PRECISELY] occurs on its own, without the first person pronoun, and it is used when one of the two persons or both of them are talking about themselves to a third person.

The 1SG person pronoun has two forms, referred to here as nominative and non-nominative. *gə* [1SG.NOM] is used as subject and also to form the ergative: *gə-s*. *an* [1SG.NNOM] is used as object, as genitive, and as the stem to which other case suffixes are added (including those for dative and genitive). In the third person, *an*, *anegsuj* and *anegə* function as anaphoric pronouns. In the reflexive construction, the dative case marker is affixed to the object pronoun (8).

- (8) *maj-o niŋɔ-s niŋɔ-nu sa-sa*
 dream-LOC 1PL.EXCL-ERG 1PL.EXCL-DAT kill~PFV
 ‘In the dream we killed ourselves.’

The third person pronouns are short forms of the demonstratives (*hɔ)dɔ*, (*hɔ)nɔ* and (*hɔ)dɔ/(*hɔ)jɔ*, (*hɔ)dɔgɔ*, (*hɔ)nɔgɔ* and (*hɔ)dɔgɔ/(*hɔ)jɔgɔ*.**

Plural personal and demonstrative pronouns (e.g. *dɔgɔ*, *nɔgɔ*) are also used to refer to a singular referent, as a marker of respect (e.g. *dɔ-gɔ lama* [DEM-PL lama(SG)] ‘that lama’). In non-honorific situations, the singular pronouns occur with plural head nouns, e.g. *dɔ kim-ɔ* [DEM.SG house-PL] ‘those houses’; *dɔ tɔ'ɛsmi-gɔ* [DEM.SG woman-PL] ‘those women.’

The anaphoric pronoun frequently occurs appositionally, after its antecedent noun. In this construction, the case marker may occur on both the noun and the pronoun. In normal anaphoric usage, *an* often takes an additional emphatic marker (9).

- (9) *ardžun-is kuaj-o la:j se~se an-i fi~fi*
 a.name-ERG well-LOC jump(N) send~PFV 3SG.COREF-EMPH.PRECISELY die~PFV
 ‘Arjun jumped in the well and he died.’

Some interrogative pronouns (and adverbs) are: *hat* ‘who, which’; *tʰu*, *tʰu* ‘why,’ *ham* ‘where’; *tetra* ‘how much, many’; *tʰɔd*, *tʰa* ‘what’; *teraj* ‘when.’ In WH-questions the word order and the verb inflectional endings remain the same as in declarative sentences (10).

- (10) *dʒu pja-ts aŋ jnum's tʰu bəd-o du*
 DEM.PROX bird-DIM 1SG.NNOM after why come-PROG BE.PRES
 ‘Why is this bird coming after (=following) me?’

TABLE 38.3 PERSONAL PRONOUNS (NOMINATIVE)

	SG	DU/PL
1	<i>gə</i>	<i>kifaj</i> (DU), <i>kifa</i> (PL.INCL), <i>niŋɔ</i> (PL.EXCL)
2NHON	<i>kɔ</i>	<i>kanif</i> (DU) <i>kanɔ</i> (PL), <i>kanegə</i> (PL)
2HON	<i>ki</i>	<i>kifi</i> / <i>kisi</i> (DU), <i>kinɔ</i> (PL)
3	<i>dɔ</i> (DISTANT, nonvisible) <i>nɔ</i> (DISTANT, visible) <i>dʒɔ</i> (PROX)	<i>dɔksuj</i> (DU), <i>dɔgɔ</i> (PL.DISTANT, nonvisible) <i>noksuj</i> (DU), <i>nɔgɔ</i> (PL.DISTANT, visible) <i>dʒɔksuj</i> (DU), <i>dʒɔgɔ</i> (PL.PROX)

3.2.3 Adjectives

Adjectives precede their head nouns (e.g. *dam batəŋ* [good news]). Modifying adverbials, such as *val* ‘much (non-countable)’; *bɔdi* ‘more, much (countable)’; *gɔma* ‘very’; *sants* ‘some’; *kjalekʰa* ‘enough, sufficient,’ precede adjectives. One set of adjectives does not inflect for gender of their head noun (e.g. *raŋk ʃʰay* ‘tall boy’; *raŋk tsʰetsats* ‘tall girl’; *raŋk ʃʰay-ɔ* [tall boy-PL]; *raŋk tsʰetsats-ɔ* [tall girl-PL]). There are other adjectives which inflect for the gender of their head noun (e.g. *rɔkalo mi* ‘black man’; *rɔkale tsʰesmi* ‘black woman’). The head noun may be left out, and the inflectional endings suffixed to the adjective instead (e.g. *raŋk-ɔ* [tall-PL]; *rɔkalo-gɔ* [black.M-PL]; *rɔkale-gɔ* [black.F-PL]). Some adjectives (e.g. *ušk* ‘old (inanimate objects)’ take *-e* as the plural marker when they occur without their head nouns (e.g. *ušk-e* [old-PL]). Finally, adjectives such as *mari* ‘weak’; *tsisʰa* ‘rotten’ do not allow any plural marker.

In addition to their attributive use, adjectives also occur as predicatives in copular constructions.

Adjectives have no comparative forms. Comparison is expressed by affixing *-o-ʃ'* [-LOC-ABL] to the standard of comparison (11).

- (11) *sjo dakʰay-o-ʃ'* *səsta du*
 apple grape-LOC-ABL cheap be.PRES
 ‘Apples are cheaper than grapes.’

The superlative is expressed by putting either *tseik-o-ʃ'* [all-LOC-ABL] or *zɔ* ‘maximum’ before the adjective ((12)–(13)).

- (12) *dʒɔ tse(i)k-o-ʃ'* *teg garay du*
 3sg.PROX all-LOC-ABL big river be.PRES
 ‘This is the longest river.’

- (13) *i fare-ts pja-ts zɔ gato-ts*
 one beautiful.F-DIM bird-DIM maximum small-DIM
atj-o nioms bəd-o du-gjɔ
 brother-GEN after come-PROG BE-PST
 ‘One beautiful bird was coming after (=following) the youngest brother.’

3.2.4 Numerals

The numerals one to ten are as follows: *id* ‘one’; *nif* ‘two’; *ʃom* ‘three’; *pə* ‘four’; *ya* ‘five’; *tog* ‘six’; *(s)tʃ* ‘seven’; *re* ‘eight’; *gv̥i* ‘nine’; *se* ‘ten.’ The numeral system is vigesimal (e.g. *nif'niza* [*nif* ‘two’ × *niza* ‘20’] ‘40’; *pə nizɔ se* [*pə* ‘four’ × *niza* ‘20’ + *ɔ* + *se* ‘10’] ‘90’). The terms for ‘hundred’ and ‘thousand’ are *ra* and *həzar*, respectively. Numerals between ‘100’ and ‘1,000’ are compounds with *ra* ‘hundred’ (e.g. *yara* ‘500’). There are two ways of forming numerals beyond ‘1,000’: *sigid ra* [1,100] and *həzaris ira* [thousand-INST *i* ‘one’ + *ra* ‘hundred’], both ‘1,100.’

4 STRUCTURE OF THE CLAUSE

The most frequent word order in Sangla Kinnauri is SOV.¹⁰

Transitivity is determined only by means of formal criteria—transitive verbs can take objects. Objects do not need to be explicitly present in order for a verb to be considered transitive. Subjects of transitive clauses can be either in the ergative or in the nominative.

Objects can be in the dative or nominative. In ditransitive clauses where both DO and IO occur, IO gets the dative marker, and DO remains in the nominative.

4.1 The verb complex

4.1.1 Verb lexeme structure

Simple verbs, like nouns and adjectives, are mostly monosyllabic or disyllabic. There are some verbs (e.g. *bənnu* ‘to come’; *sannu* ‘to kill’; *lənnu* ‘to say’; *vannu* ‘to laugh’) where a different stem-final consonant appears in some contexts, for example, in the progressive aspect (e.g. *bəd-o* [come-PROG]; *sad-o* [kill-PROG]; *ləd-o* [say-PROG]; *vad-o* [laugh-PROG]) and when the manner marker -(j)e is suffixed to the verb stem (e.g. *bəd-e* [come-MANNER]).¹¹

Complex verbs are frequently encountered. One of the two main types consists of a noun followed by a light or support verb (e.g. *bək lan-nu* [warmth make-INF] ‘to warm (TR)’; *pudʒa ma-lan-nu* [prayer NEG-make-INF] ‘not to pray.’ No dative marker occurs on the noun, while the negative marker (including the prohibitive marker) is prefixed to the verb and not to the whole N-V combination. These support verb constructions are treated as verb lexemes here both because of their non-compositional semantics and because the N+V complex often behaves as a simple transitive verb, taking a direct object (14).

- (14) *ama nif-u tseik-əs-i ase ta-tʃ-o du*
 mother 1DU-DAT all-ERG-EMPH.PRECISELY torture keep-OBJ.AGR-PROG BE.PRES
 “Mother, everyone is torturing us two.”

In the other complex verb construction, the compound verb construction, a main verb (in the perfective verb form) is followed by a light or vector verb such as *nimu* ‘to stay’; *rannu/kemu* ‘to give’; *bjɔmu* ‘to go’; *tamu* ‘to keep’; *fennu* ‘to send’ (15).

- (15) *lag-u kom-o fe~fe ta-u du*
 sleeve-GEN inside-LOC send~PFV keep-PROG BE.PRES
 ‘(She) is sending (putting) it inside her sleeve.’

Some valency changing mechanisms are as follows.

4.1.1.1 Reflexive/middle -(j)i

The reflexive/middle in Sangla Kinnauri has parallels in several other Tibeto-Burman languages (LaPolla 1996). -(j)i is realized as -ʃ when the suffix following it starts with a vowel. With some verbs, it is realized as -ʃi (e.g. *legmu* ‘to burn,’ *legfimu* ‘to get burned’). While some verbs with -(j)i describe a resulting state (e.g. *g̊ifimu* ‘to get washed (e.g. in the rain)’), others describe intransitive actions (e.g. *qabfimu* ‘to crawl’). Further, -(j)i occurs in constructions where the subject and the object refer to the same argument (e.g. *sumu* ‘to bathe (someone else),’ *suʃimu* ‘to bathe (oneself)'). It also occurs in constructions where it highlights that more than one person is involved in an activity and that the action is done jointly (16), and in reciprocal constructions, which can optionally also contain the reciprocal pronoun (17).¹² The ergative and the dative marker are not permitted on the core arguments of transitive clauses with -(j)i.

- (16) *sum-ki ləʃ-o du*
 three-EMPH.PRECISELY say-R/M-PROG BE.PRES
 ‘All three are saying (at the same time).’

TABLE 38.4 TRANSITIVITY CONTRASTS OF DISYLLABIC VERBS WITH *-en* VS *-ja*

INTR	TR		
<i>pɔlənnu</i>	‘to turn around, to roll’	<i>pɔljamu</i>	‘to turn (something) around, to roll’
<i>ba:sennu</i>	‘to smell’	<i>ba:sjamu</i>	‘to smell (something)’
<i>bɔjennu</i>	‘to float, to blow’	<i>bɔjamu</i>	‘to float, to blow (something)’
<i>rukennu</i>	‘to stop’	<i>rɔkjamu</i>	‘to stop (something)’
<i>sɔmzennu</i>	‘to understand’	<i>sɔmzjamu</i>	‘to understand (something)’

TABLE 38.5 TRANSITIVITY CONTRASTS OF DISYLLABIC VERBS WITH *-ja* VS *-en/-ʃ(i)*

v (TR)	(cf. Table 38.4)	v (INTR)	
<i>rɔkjamu</i>	‘to stop (something)’	<i>rukennu</i>	‘to stop (on its own)’
<i>dʒɔnljamu</i>	‘to swing (something)’	<i>dʒɔnlennu</i>	‘to swing (on its own)’
<i>metjamu</i>	‘to gather (something)’	<i>metʃasimu</i>	‘to gather (on its own)’
<i>kusjamu</i>	‘to wipe, to sweep’	<i>kusʃasimu</i>	‘to wipe (oneself)’
<i>tujamu</i>	‘to make someone ready’	<i>tujasimu</i>	‘to get ready’

- (17) *tsʰay-ɔ* (*ane-gɔ*) *kul-f-o* *du*
 child-PL self-PL hit-R/M-PROG BE.PRES
 ‘Children are fighting among themselves.’

4.1.1.2 The intransitivizer *-en*

Disyllabic verb stems in *-en* are intransitive. As the examples in Table 38.4 illustrate, their corresponding transitive verb forms have the transitivizing morpheme *-ja*.

As we saw earlier, some underived verb stems exhibit a *-d/-n* alternation. This is also the case with verbs in *-en*. (e.g. *bɔjen-nu* [float-INF], *bɔjed-o* [float-PROG]; *baten-nu* [talk-INF], *bated-o* [talk-PROG]).

4.1.1.3 The transitivizer *-ja*

Disyllabic verb stems in *-ja* are transitive. They often have an intransitive Indo-Aryan base. As we can see from the examples in Table 38.5, the corresponding intransitive verb stems have a valency decreasing marker (*-ʃ(i)* or *-en*).

4.1.1.4 Stem-initial voicing contrasts

With a small number of verbs, when the intransitive verb stem begins with a voiced obstruent, the corresponding transitive verb stem begins with a voiceless consonant, as illustrated in Table 38.6. However, this is not a productive process in Sangla Kinnauri.

4.1.2 Structure of the verb complex

The verb complex exhibits one of the following structures:

Copula construction: Vcop-(TNS)-AGR

Non-copula constructions: V-(OBJ.AGR)-TNS-AGR

TABLE 38.6 STEM INITIAL VOICING CONTRASTS RELATED TO TRANSITIVITY

Intransitive	Transitive	Intransitive	Transitive
<i>bəŋmu</i>	‘to fill’	<i>pəŋmu</i>	‘to fill’
<i>grumu</i>	‘to burn (food items)’	<i>kruₘu</i>	‘to burn (food items)’
<i>bjogmu</i>	‘to blow off fire (on its own)’	<i>pjogmu</i>	‘to blow off fire’
<i>gjulmu</i>	‘to scrape’	<i>k'julmu</i>	‘to scrape’
<i>dʒɔgmu</i>	‘to drip’	<i>ʃɔgmu</i>	‘to drip’
<i>bralmu</i>	‘to fall’	<i>pʰralmu</i>	‘to fall’

TABLE 38.7 SUBJECT AGREEMENT MARKERS

Person	SG	PL/DU
1	- <i>k</i>	- <i>f</i> (DU; PL.EXCL) / - <i>me</i> (PL.INCL)
2NHON	- <i>n</i>	- <i>n(-o)</i> (DU; PL)
2HON	- <i>n</i>	- <i>f'</i> (DU; PL)
3NHON	0	0 (DU; PL), - <i>suy</i> (DU)
3HON	- <i>f</i>	- <i>f(-o)</i> (DU; PL)

N V_{light}-TNS-AGR
 V/PFV V_{light}-(OBJ.AGR)-TNS-AGR
 V-(OBJ.AGR)-ASP (AUX-(TNS)-AGR)
 N V_{light}-ASP (AUX-(TNS)-AGR)
 V/PFV V_{light}-(OBJ.AGR)-ASP (AUX-(TNS)-AGR)

4.1.2.1 Subject and object agreement and tense markers

The subject agreement markers occur in both copula and non-copula constructions. Table 38.7 presents the subject agreement markers. -*o* functions as the plural agreement marker with 2NHON and 3HON, and -*suy* functions as the dual subject agreement marker. In natural discourse, however, the plural marker rarely occurs for plural subjects. Similarly, with dual subjects, the plural marker -*o* occurs more frequently than the dual agreement marker -*suy*.

The object agreement marker is -*f* (except with the verbs ‘to give,’ ‘to tell’¹³). In some cases it appears in free variation with -*dʒ*. It occurs when the first and second person argument is the most affected argument in a clause ((14), (18)–(19)).

- (18) *ay-u birmafṣtən rakses-o dor ſe-f'-is*
 1NNOM-DAT PN demon-GEN along send-OBJ.AGR-PFV
 ‘He sent me with the demon Birma Chosten.’
- (19) *ram-is ay-u zali bat-en-nu se-f'-e*
 a.name-ERG 1NNOM-DAT lie(N) talk-INTR-INF send-OBJ.AGR-PST
 ‘Ram made me tell (a) lie.’

The dative-marked argument in the dative experiencer construction does not control subject agreement. If it is first or second person, it triggers object agreement (20).

- (20) *ki-nu akʰa ker-o du-ge*
 2SG.HON-DAT pain give.OBJ-PROG BE-PST
 ‘You were having pain.’

4.1.2.2 Copula construction

tɔ, *du* and *ni* function both as equational and existential copulas. The copula is, however, not obligatory. Tables 38.8–38.10 present the copula paradigms, where one can also see the distribution of the copulas and which inflectional endings they take.

The distribution of the copulas (*tɔ*, *du* and *ni*) with third person honorific and non-honorific subjects is semantically conditioned. With non-honorific subjects, *tɔ* indicates that the subject is closer to the speaker (kinship or physical proximity). *ni* occurs where the hearer has some doubts either about the very existence of the subject, or in identifying the subject as either A or B, while the speaker definitely knows the answer (either having seen it personally or because s/he has some other way of knowing the truth). *du* occurs elsewhere. With honorific subjects the selection of the copulas (*tɔ* or *du*) is determined by the animacy of the subject: *tɔ-f* occurs with animate subjects and *du-f* occurs with inanimate subjects. The semantic interpretation of *ni* with honorific subjects remains the same as described above for non-honorific subjects.

The copulas *tɔ* and *du* also function as auxiliaries (see section 4.1.2.4). As an auxiliary, *du* occurs with first and second person subjects too, without any apparent change in meaning. With third person subjects, the copula auxiliary *tɔ* occurs when the subject referent is in the close vicinity of the speaker and/or is closely related to the speaker.

TABLE 38.8 COPULA PARADIGM: PAST TENSE

Person	SG	DU/PL
1	<i>tɔ-ke-k</i>	<i>tɔ-ke-f</i> (DU; PL.EXCL), <i>tɔ-ke-me</i> (PL.INCL)
2NHON	<i>tɔ-ke-n</i>	<i>tɔ-ke-n(-ɔ)</i> (DU; PL)
2HON	<i>tɔ-ke-n</i>	<i>tɔ-ke-f</i> (DU; PL)
3NHON	<i>tɔ-ke / du-ge / tɔ-kjɔ / du-gjɔ</i> / <i>tɔ-ke / du-ge / tɔ-kjɔ / du-gjɔ / ni</i> (DU; PL) <i>ni</i>	
3HON	<i>tɔ-ke-f / du-ge-f / ni</i>	<i>tɔ-ke-f(-ɔ) / du-ge-f(-ɔ) / ni</i> (DU; PL)
3DU.HON		<i>tɔ-ke-suj / du-ge-suj / ni</i> (DU)

TABLE 38.9 COPULA PARADIGM: PRESENT TENSE

Person	SG	DU/PL
1	<i>tɔ-k</i>	<i>tɔ-f</i> (DU; PL.EXCL), <i>tɔnne</i> (PL.INCL)
2NHON	<i>tɔ-n</i>	<i>tɔ-n(-ɔ)</i> (DU; PL)
2HON	<i>tɔ-n</i>	<i>tɔ-f</i> (DU; PL)
3NHON	<i>tɔ / du / ni</i>	<i>tɔ / du / ni</i> (DU; PL)
3HON	<i>du-f / tɔ-f / ni</i>	<i>tɔf(-ɔ) / du-f(-ɔ) / n</i> (DU; PL)
3DU.HON		<i>tɔ-suj / du-suj / ni</i> (DU)

TABLE 38.10 COPULA PARADIGM: FUTURE TENSE

Person	SG	DU/PL
1	<i>ni-ta-k</i>	<i>ni-ti-f</i> (DU; PL.EXCL), <i>ni-te</i> (PL.INCL)
2NHON	<i>ni-ta-n</i>	<i>ni-ta-n(-ɔ)</i> (DU; PL)
2HON	<i>ni-tr-n</i>	<i>ni-tt-f</i> (DU; PL)
3NHON	<i>ni-tɔ</i>	<i>ni-tz-(gɔ)</i> (DU; PL)
3HON	<i>ni-tt-f</i>	<i>ni-tt-f(-ɔ)</i> (DU; PL)
3DU.HON		<i>ni-ti-suj, ni</i> (DU)

4.1.2.3 Non-copula constructions

In the finite structure v-(OBJ.AGR)-TNS-AGR, a future–past tense distinction is made. The future tense markers (-ta, -ti and -to) and their distribution are the same here as presented earlier in the copula paradigms. The past tense markers which occur in this structure are: -ke, -gjɔ, - (j)a, -ɔ and 0. They are grouped here in two sets: Set 1: -ke, -gjɔ and Set 2: - (j)a, -ɔ, 0. Set 1 occurs in both copula and non-copula constructions: -ke occurs with 3HON subjects and -gjɔ occurs with 3NHON subjects. -ke is also realized as -ge or as -fe, although their appearance is not phonologically conditioned. The vowel -e in -ke is mostly realized as -i (i.e. -ki, -gr or -fi) when it is followed by the 3HON agreement marker (e.g. *tofi-gr-f* [sit-PST-3HON]). Similarly, when the final syllable in a disyllabic verb stem has an -e (e.g. *tofennu* ‘to sit’), the -e of -ke is realized as -i, resulting in finite verb forms such as *tofi- gi- f* [sit-PST-3HON]. Similarly, when the verb stem ends in a nasal, the consonant of the past tense marker is sometimes dropped in fast speech (e.g. *ray-i-f* [say-PST-3HON]; *pa:ŋ-i-f* [build-PST-3HON]).

The Set 2 markers (- (j)a, -ɔ, 0) occur only in the non-copula constructions, with all persons and numbers. Their distribution is complementary: Some verbs (e.g. *fimū* ‘to wash’) only take the marker -ɔ whereas other verbs (e.g. *zamu* ‘to eat’) only permit 0 (Table 38.11).

However, the Set 2 marker is always realized as -e when the finite verb either has the object agreement marker (21) or the reflexive/middle marker (22).

- (21) *ʃ⁰ay-ɪs ay-u kul-ʃ⁰-e*
 boy-ERG 1SG.NNOM-DAT beat-OBJ.AGR-PST
 ‘The boy hit me.’
- (22) *sjanɔ-gɔ sjan-ʃ⁰-e-f*
 old.person-PL meet-R/M-PST-HON
 ‘(Those) old people met (each other).’

As mentioned earlier, both Set 1 and Set 2 markers are permitted with third person honorific and non-honorific subjects. Their distribution is semantically determined: the Set 2 markers occur when the speaker has direct knowledge of that which is being described (typically by having seen it), while the Set 1 markers are used in other cases.

4.1.2.4 Aspect morphology

A three-way aspectual distinction is made: perfective, imperfective, and progressive.

The perfective aspect markers are -is and reduplication of the final syllable of the verb stem:¹⁴ -is occurs when the verb ends in /f/ or /s/ (e.g. *bɔf-is* [forget-PFV]; *haʃf-is* [become-PFV])

TABLE 38.11 SET 1 AND SET 2 PAST TENSE MARKERS

Infinitive form	Set 1: -ke, -gjɔ	Set 2: 0	Set 2: -ɔ	Set 2: -a
	(3SG.NHON, 3SG.HON)	(1SG, 3SG.NHON, 3SG.HON)		
<i>bjomu</i> ‘to go’	<i>bjɔ-gjɔ, bjɔ-gr-f</i>	<i>bjɔ-k, bjɔ, bjɔ-f</i>		
<i>vamu</i> ‘to laugh’	<i>va-kjɔ, va-ki-f</i>	<i>va-k, va, va-f</i>		
<i>zamu</i> ‘to eat’	<i>za-gjɔ, za-gr-f</i>	<i>za-k, za, za-f</i>		
<i>fimū</i> ‘to wash’	<i>f̩i-gjɔ, f̩i-gr-f</i>		<i>f̩i-ɔ-k, f̩i-ɔ, f̩i-ɔ-f</i>	
<i>fēmu</i> ‘to write’	<i>f̩e-gjɔ, f̩e-gr-f</i>		<i>f̩e-ɔ-k, f̩e-ɔ, f̩e-ɔ-f</i>	
<i>taymu</i> ‘to look’	<i>tay-gjɔ, tay-(g)r-f</i>		<i>tay-ɔ-k, tay-ɔ, tay-ɔ-f</i>	
<i>kərmu</i> ‘to bring’	<i>kər-gjɔ, kər-gr-f</i>			<i>kər-a-k, kər- a,</i> <i>kər- a-f</i>

and reduplication in all other cases (e.g. *t^{has}-t^{has}* [hear~PFV], *sontse-sontse* [hear~PFV]). Perfective verbs also occur as the non-final verb forms in the clause chain construction, in the compound verb construction (V/PFV V_{light}), as the past participle verb forms and in non-final clauses with a temporal adverbial interpretation.

The imperfective aspect marker *-id* occurs after verb stems (including with the object agreement marker) ending in *f* or *f*; *-ts* occurs in all other cases. It also functions as the present participle marker and as the agentive nominalizer.

The progressive aspect marker *-o*, sometimes realized as *-u* when the verb stem ends with an *-o* (e.g. *zo-u* [eat-PROG]). It also occurs in some non-final positions: as a noun modifier indicating an ongoing action, and as a non-final verb in the clause chain construction, where it describes temporal overlap.

4.2 Negation

There are two negative copulas in the present tense: *ma-ni*, which has a contrastive interpretation, and NEG-AGR (e.g. *ma-k* [1SG]; *ma-yf* [2HON.PL]; *ma-du* [3NHON (animate, inanimate)]; *ma-f* [3SG.HON]), which has a neutral negative interpretation.

In the past tense, *mats^h*; *ma-ke-AGR*; *ma-du-ge*; *ma-du-gjɔ* function as negative copulas (existential, equational). *ma-ke-AGR* occurs with third person honorific subjects and *ma-du-gjɔ* occurs with third person non-honorific subjects. While *mats^h* has a neutral interpretation, *ma-ke-AGR*, like *ma-ni*, occurs in situations where the speaker wants to correct the listener's claim or assumption. In the future tense the negative copula is *ma-ni-AGR(-PL)*.

The negative copulas also occur in the following 'or not'-constructions ((23)–(24)).

- (23) *kisi babu to-yf-a ma-yf*
2DU administrator be-2PL.HON-Q NEG-2PL.HON
'Are you two administrators or not.'

- (24) *ds kim-o du-a ma-du*
3SG house-LOC BE.PRES-Q NEG-BE.PRES
'Is he in the house or not?'

In negated non-prohibitive, non-copula clauses *ma-* is prefixed to the finite or non-final verb (e.g. *ma-sa-k* '(I) did not kill'). In finite verb constructions with an auxiliary, *ma-* may be prefixed either to the main verb or to the auxiliary, depending on the scope of the negation.

4.3 Polar questions

Polar questions are formed by affixing *-a* to the finite verb (25). The question suffix *-a* does not occur in WH-constructions (see example (10) earlier).

- (25) *zap-u dejay gja-ti-n-a*
gold-GEN body want-FUT-2HON-Q
'Do (you) want a body of gold?'

4.4 Imperative

The following four sets of inflectional endings are used to form the imperative. Their distribution reflects a complex interplay of semantic and pragmatic factors, where *-(i)ri-n/-yf* is the most polite and (*Ø?*) the most direct alternative (Saxena 2002, 2011).

(PROH-)V(-OBJ.AGR)	- <i>(i)rr-n</i> /- <i>tf</i> <i>-i)n</i> /- <i>tf</i> <i>-i)ra</i> <i>(Ø?)</i> /- <i>ju</i> /- <i>jo</i>
--------------------	--

The various inflectional endings are illustrated below.

	- <i>(i)rr-n</i>	- <i>(i)n</i>	- <i>(i)tf</i>	- <i>(i)ra</i>	(<i>Ø?</i>) ¹⁵
<i>rannu</i> ‘to give’	<i>ran-ri-n</i>	<i>ran-iy</i>	<i>ran-rif</i>	<i>ran-ra</i>	<i>ran</i>
<i>kemu</i> ‘to give (OBJ.AGR)’	<i>ke-ri-n</i>	<i>ke-p</i>	<i>ke-tf</i>	<i>ke-ra</i>	<i>kjɔ</i>
<i>tsemu</i> ‘to write’	<i>tfe-ri-n</i>	<i>tfe-p</i>	<i>tfe-tf</i>	<i>tfe-ra</i>	<i>tʃo</i>
<i>raymu</i> ‘to tell’	<i>ray-ri-n</i>	<i>ray-iy</i>	<i>ray-tf</i>	<i>ray-ra</i>	<i>ray</i>
<i>dʒi-</i> [come.IMP]	<i>dʒi-ri-n</i>	<i>dʒi-p</i>	<i>dʒi-tf</i>	<i>dʒi-ra</i>	—

The occurrence of *-i* in the imperative suffix is phonologically conditioned. It occurs sometimes when the verb stem ends with a consonant. *-n* is the 2SG.HON subject agreement marker and *-tf* is the 2PL.HON subject agreement marker (see section 4.1.2.1). In the imperative constructions *-tf* also occurs, at times, with singular subjects and in situations where the referent is a member of a group.

While most verbs permit the zero imperative marker¹⁶ as one of the possible imperative forms (e.g. *bjɔ* [go.IMP]; *riŋ* [tell.OBJ.AGR.IMP]; *haled* [(take.a).walk.IMP]; *sad* [kill.IMP]), a restricted set of verbs takes *-ju* or *-jo* instead (e.g. *ni-ju* [stay-IMP]; *kʃj-ɔ* [see-IMP]). Exceptionally, the verb *bənnu* ‘to come’ exhibits suppletion—*dʒi-* [come.IMP—but lacks the (*Ø?*)-imperative.

Prohibitives have the same structure as the imperatives, except for the additional prohibitive morpheme *tʰa-* which is prefixed to the verb.

ADDITIONAL ABBREVIATIONS

AGR	subject agreement	NIMP	non-imperative
CONN	connective	NNOM	non-nominative
COP	copula	OBJ	object
COREF	coreferential	OBJ.AGR	object agreement
COUNT	countable	PROX	proximate
DU	dual	PST	past
EXCL	exclusive	TNS	tense
F	feminine	V	verb
INCL	inclusive		
IO	indirect object		
N	noun		
NHON	non-honorific		

NOTES

1 I would like to express my deep gratitude to my language consultants, and in particular Mrs Santosh Negi for her generosity and for her help in numerous ways over the decades that I have studied Sangla Kinnauri. I would also like to thank Graham Thurgood and Randy LaPolla for their comments.

- 2 I refer to this language as *Sangla* Kinnauri to reflect the fact that Kinnaur is the home of a number of Tibeto-Burman (TB) languages, a fact which the label “Kinnauri” tends to obscure. Sangla Kinnauri encompasses a number of closely related language varieties spoken in the Sangla tahsil (Saxena and Borin 2013; Saxena in preparation). These varieties all belong to a traditionally recognized lower-level branch of TB. This is the classification given by the *Ethnologue*, with some variations depending on the edition: The 16th edition (Lewis 2009) classifies Kinnauri (ISO 639-3: ksf) together with 11 other languages as “Sino-Tibetan>TB>Himalayish>Tibeto-Kanauri>Western Himalayish>Kanauri>Kinnauri,” while the 17th edition (Lewis *et al.* 2013) groups it with 11 other languages (although not the same 11 languages as in the 16th edition) as “Sino-Tibetan>TB>Western TB>Bodish>West Himalayish>Kinnauri>Kinnauri,” in both cases placing them fairly close to Tibetic. In contrast to this, e.g. LaPolla (2003) places Tibetic and West Himalayish rather far apart in the TB family tree (under the higher-level units Bodish and Rung, respectively). Matisoff (2000) also does not recognize a Tibeto-Kanauri branch, but simply places Tibetic and “the westernmost TB languages of Himachal Pradesh” together under the higher-order grouping Himalayish (Matisoff 2000: 351).
- 3 There is no noticeable word-final glottal closure in the latter pair of examples.
- 4 The contrastive specifier *-sja/-se* is distinct from the agentive nominalizer *-tsja/-ts'e*. The agentive nominalizer is affixed to a verb, while the contrastive specifier is affixed either to a noun (animate male/female), a pronoun, or to an adjective not followed by a head noun.
- 5 *-u* and *-o* appear interchangeably.
- 6 In the South Asian context, this is often referred to as the “dative subject” or “experiencer subject” construction (e.g. Verma and Mohanan 1990), but in Sangla Kinnauri the dative-marked argument in this construction behaves as an object, not a subject (see section 4.1.2.1).
- 7 *-u* and *-o* appear interchangeably.
- 8 In direct-elicitation language consultants accepted *-suj* with a few [+human] referents (e.g. *tsʰetsats-suj* [girl-DU]; *dəkʰrats-suj* [young.man-DU], but not with, e.g. **tsʰetsmi-suj* [woman-DU]; **mi-suj* [man-DU]).
- 9 *jums* has a range of functions, e.g. as a subordinator where it has an ‘after’ interpretation. It also occurs in the expressions for ‘afternoon,’ ‘leftover food,’ ‘last’ and ‘pregnant’ (Saxena 2008).
- 10 A systematic exception to this is the dative experiencer construction, where the dative-marked argument precedes the subject; see (4) and (21).
- 11 Unlike most verbs, this set of verbs has *-nu* as the infinitive marker, and not the normal *-mu*. The infinitive form is best seen as the result of a mutual assimilation process (e.g. *sad-mu* > *sannu*).
- 12 In the reflexive constructions (section 3.2.2, ex. (8)), the reflexive/middle marker *-si* is not permitted.
- 13 The verbs ‘to give’ and ‘to tell’ display suppletion: *kemu* ‘to give,’ *rəŋmu* ‘to tell’ occur with 1st and 2nd person objects, and *rannu* ‘to give’; *lənnu* ‘to tell’ with 3rd person objects. *kemu* ‘to give (OBJ.AGR)’ has the past-tense form *ker-ɔ* (see (20)). This is not a regular phenomenon (morphophonological alternation or phonetic variation); it seems to be unique to this verb.
- 14 In the perfective form of verb stems with the transitivizer *-ja*, the last consonant of the first syllable together with the morpheme *-ja* are reduplicated (e.g. *tsʰinjamu* ‘to cut (vegetables),’ *tsʰinjə-nja* [cut~PFV]).
- 15 In some cases there is a change in the stem vowel.
- 16 In the case of a small number of verbs, the (Ø?)-imperative verb form is the same as one of the past verb forms with third person non-honorific subjects (*dɔ bɔ* [3SG go.PST] ‘he went,’ *bɔ* [go.IMP] ‘go!’). The distinction in function is indicated here by means of prosody and the context in which it occurs. Further, while the verb form is often repeated while giving orders, this rarely happens with declaratives.

REFERENCES

- Bailey, Thomas Grahame. 1908. *The Languages of the Northern Himalayas, Being Studied in the Grammar of Twenty-Six Himalayan Dialects*. Asiatic Society monographs vol. XII. London: Royal Asiatic Society.
- Bailey, Thomas Grahame. 1909. A brief grammar of the Kanauri language. *Zeitschrift der deutschen morgenländischen Gesellschaft* 63: 661–687.
- Bailey, Thomas Grahame. 1910. Kanauri vocabulary in two parts, English–Kanauri and Kanauri–English. *Journal of the Royal Asiatic Society* 42.3: 659–705.
- Bailey, Thomas Grahame. 1911. Kanauri vocabulary in two parts, English–Kanauri and Kanauri–English. (Continued from 1910, p. 705.) *Journal of the Royal Asiatic Society* 43.2: 315–365.
- Bailey, Thomas Grahame. 1920. Lower Kanauri. In: Thomas Grahame Bailey (ed.), *Linguistic Studies from the Himalayas Being Studied in the Grammar of Fifteen Himalayan Dialects*. Asiatic Society Monographs vol. XVIII. London: Royal Asiatic Society, 47–86.
- Bailey, Thomas Grahame. 1938. A brief grammar of the Kanauri language. In: Thomas Grahame Bailey (ed.), *Studies in North Indian Languages*. London: Lund Humphries, 79–105.
- Gerard, Capt. Alexander. 1841. *Account of Koonawur, in the Himalaya*, edited by George Lloyd. London: James Madden.
- Gerard, Capt. Alexander. 1842. A vocabulary of the Kunawar languages. *Journal of the Asiatic Society of Bengal* 11: 478–551.
- Joshi, Pandit Tika Ram. 1909. *A Grammar and Dictionary of Kanawari, the Language of Kanawar, the Bushahr State, Punjab*. Edited by H. O. Rose. *Journal and Proceedings of the Asiatic Society of Bengal, V (Extra number)*. Calcutta: Baptist Mission Press. [Reprint 1989. Shimla: Himachal Academy of Arts, Culture and Languages.]
- Konow, Sten. 1905. Some facts connected with the Tibeto-Burman dialect spoken in Kanawar. *Zeitschrift der deutschen morgenländischen Gesellschaft* 59: 117–125.
- LaPolla, Randy J. 1996. Middle voice marking in Tibeto-Burman. *Proceedings of the Fourth International Symposium on Languages and Linguistics: Pan-Asiatic Linguistics* (8–10 January 1996), Volume 5: 1940–1954.
- LaPolla, Randy J. 2003. Overview of Sino-Tibetan morphosyntax. In: Graham Thurgood and Randy J. LaPolla (eds), *The Sino-Tibetan Languages*. London: Routledge, 22–42.
- Lewis, M. Paul (ed.). 2009. *Ethnologue: Languages of the World* (16th edition). Dallas: SIL International.
- Lewis, M. Paul, Gary F. Simons and Charles D. Fennig (eds). 2013. *Ethnologue: Languages of the World* (17th edition). Dallas: SIL International.
- Matisoff, James. 2000. On the uselessness of glottochronology for the subgrouping of Tibeto-Burman. In: Colin Renfrew, April McMahon and Larry Trask (eds), *Time Depth in Historical Linguistics*. Cambridge: McDonald Institute for Archaeological Research, 333–371.
- Saxena, Anju. 1992. Finite verb morphology in Tibeto-Kinnauri. PhD dissertation, University of Oregon.
- Saxena, Anju. 1995. Finite verb morphology in Kinnauri. *Cahiers de Linguistique, Asie Orientale* 24.2: 257–282.
- Saxena, Anju. 2002. Speech reporting strategies in Kinnauri narratives. *Linguistics of the Tibeto-Burman Area* 25.1: 165–190.

- Saxena, Anju. 2008. On *ñum* and *Om* in Kinnauri. In: Birgitte Huber, Marianne Volkart and Paul Widmer (eds), *Chomolangma, Demawend und Kasbek: Beiträge zur Zentralasienforschung*. Bonn: International Institute for Tibetan and Buddhist Studies, 153–163.
- Saxena, Anju. 2011. Context shift and linguistic coding in Kinnauri narratives. In: Roland Bielmeier and Felix Haller (eds) *Linguistics of the Himalayas and Beyond*. Berlin: De Gruyter Mouton, 247–264.
- Saxena, Anju. forthcoming. Charting the linguistic landscape of Kinnaur—a part of the Indian Himalayas.
- Saxena, Anju and Lars Borin. 2013. Carving Tibeto-Kanauri by its joints: using basic vocabulary lists for genetic grouping of languages. In: Lars Borin and Anju Saxena (eds), *Approaches to Measuring Linguistic Differences*. Berlin: De Gruyter Mouton, 175–198.
- Sharma, Devidatta D. 1988. *A Descriptive Grammar of Kinnauri*. Delhi: Mittal Publications.
- Verma, Manindra K. and Tara Mohanan (eds). 1990. *Experiencer Subjects in South Asian Languages*. Stanford, CA: CSLI Publications.

§3.5 Qiangic

CHAPTER THIRTY-NINE

QIANG¹

Randy J. LaPolla

Qiang is spoken in Aba Tibetan and Qiang Autonomous Prefecture in northwest Sichuan Province, China, by people classified as either Tibetan (Heishui county) or Qiang (other areas), though all use the autonym *զմե* or some variant of this form. The Qiang language has been said to belong to the Qiangic branch of Tibeto-Burman (H. Sun 1982; but see Chirkova 2012 questioning the validity of this grouping, and Thurgood, this volume for current thinking on this question). Hongkai Sun (1981a: 177–8) has suggested there are two major Qiang dialects: Northern Qiang (spoken in Heishui County and the Chibusu district of Mao County; roughly 70,000 speakers) and Southern Qiang (spoken in Li County, Wenchuan County, Mao County, and Songpan County; about 60,000 speakers). Recent, as yet unpublished, work by Jackson T.-S. Sun and Jonathan Evans, and by Nathaniel Sims, has called this division into question, but a clear division of the varieties will have to await further fieldwork and comparison (see Chang 1967 and work by Evans 2001a, 2001b for comparative studies). The dialect presented here is the Northern Qiang variety spoken in Ronghong Village, Yadu Township, Chibusu District, Mao County. See LaPolla with Huang (2003, including texts and glossary) and other work by LaPolla and Chenglong Huang (a native linguist) for a more complete description of this variety. See Liu (1998b) and other work by Guangkun Liu and Hongkai Sun and also Jackson T.-S. Sun and Jonathan Evans on the Mawo variety; C. Huang (2004) and C. Huang and Wang (2007) on the Puxi variety of Qiang; Huiqun Liu (2011) on the Muka variety; and work by Bufan Huang and Facheng Zhou and also LaPolla (2003a) and LaPolla and Poa (2003) on the Qugu variety. Other works on the history, culture and literature of the Qiang are given in the references.

1 THE PHONOLOGICAL SYSTEM

Qiang has 39 consonants at seven points of articulation (Table 39.1), plus complex consonant clusters, both in initial and final position.

Items in parentheses are not phonemic: [v] is an allophone of /w/ when it appears before front vowels; [z] and [ɣ] are allophones of /č/ and /χ/, respectively, when followed by a voiced consonant. There is no phonemic contrast between a glottal stop onset and a pure vocalic onset or between /u/ and /wu/. Almost all of these consonants, except the aspirated stops and aspirated affricates, can be finals. All of the original Proto-Tibeto-Burman finals were lost (see G. Liu 1984), but new ones were created from the merging of two syllables where the de-stressing of the second syllable led to the loss of the final vowel (and often reduction of the original initial, e.g. [səf] ‘tree’ < /sə/ ‘wood’ + /phə/ ‘forest’).

TABLE 39.1 THE QIANG CONSONANTS

	Labial	Dental	Retroflex	Palatal	Velar	Uvular	Glottal
Voiceless stop	p	t			k	q	
Aspirated stop	ph-	th-			kh-	qh-	
Voiced stop	b	d			g		
Voiceless affricate		ts	tʂ	tɕ			
Asp. affricate		tsh	tʂh	tɕh			
Voiced affricate		dz	dʐ	dʑ			
Voiceless fricative	ɸ(f)	s	ʂ	ɕ	x	χ	h
Voiced fricative	(v)	z	ʐ	ʑ	(ɣ)	ʁ	ɦ
Nasal	m	n		ɳ	ŋ		
Voiceless lateral	ɿ						
Voiced lateral	ɿ						
Approximant	w			j			

Phonemically, consonant clusters are formed by /ʂ/ and one of the following initials: /p, t, k, q, tɕ, b, d, g, m, dz/, /x/ followed by /k, s, tɕ, ʂ, tʂ, l, l, z, dz, z, dz/. or /χ/ followed by /q, s, ʂ, tʂ, ʂ, l, d, z, n, dz, ɳ, z, dz/. Phonetically /ʂ/ becomes [s] before /t/ and /d/, and becomes [ʂ] before /pi/, /pe/, /bi/, /te/ and /dʐ/, and the pre-initials all become voiced before voiced initials (e.g. /mi:-xkam/ ‘eyebrow’; /χsu/ ‘living’, ‘to be alive’; /sta/ ‘entrust to’; /ʐbu/ ‘drum’; /zdu/ ‘deer’; /zdzi/ ‘disease’; /udzəs/ ‘toenail’). Some examples of clusters in final position: /tʂhexl/ ‘sip (vt.)’; /dʐætɕe/ ‘laugh (v)’; /wəχʂ/ ‘horse dung’; /laxʂ/ ‘palm’; /əxtʂ/ ‘shade (vt.)’.

The Qiang vowels are given in (1):

- (1) i, i: y, y: u,u:
 e, e: o,o:
 ə
 a, a: ə, ə:

There are 14 native diphthongs (/ia, ia, ie, ye, eu, əu, ei, əi, oi, ua, ua, uə, ue, ui/) and one native triphthong (/uəi/). The diphthongs [ya] and [ya:] occur when the first person suffix /-a/ or the prospective aspect marker /-a:/ is added to a root such as /təye/ ‘carry’: [teya] ‘I carry’, [teya:] ‘going to carry’. Two diphthongs (/ai/, /au/) and two triphthongs (/uai/, /iau/) appear only in Chinese loan words.

Four of the basic vowels (i, e, ə, a) show a lexical contrast in r-colouring, a retroflexion of the tongue at the end of the vowel, and all vowels can take r-colouring when they are the final vowel of a verb with 1pl marking (which is /-i/). The r-colouring participates in the vowel harmony scheme (see below) and so is treated as a vowel feature rather than a consonant. We do not find the sort of tongue backing (uvularization) found in the Mawo dialect (Evans *et al.* 2016).

The syllable canon is given in (2):

- (2) (c) (c_i) (v) v (v) (c) (c_f)
 [fric] [glide] [glide] [fric]

The minimum syllable is a single vowel, e.g. /a/ ‘one’, the maximum is ccVVCC, e.g. /epiexl/ ‘scar’. Any of the consonants listed in Table 39.1 can be the initial consonant of a syllable, but only fricatives can be the first consonant of a cluster. The same restriction applies to final consonant clusters.

When certain consonants appear in non-word-initial position due to affixation or compounding, they undergo lenition, e.g. /ph/ > [ɸ] /phə/ 'blow' > [ɸ] 'blow (imperative)'; /kh/ > [x] /khə'te/ 'hit (people)' > [nəxte] 'hit (perfective)'; /dz/ > [l] /ma/ NEG + /dʒə/ 'able' > [mə-lə] 'not able'; /dz/ > [z] /dzə/ 'sit' > [əzə] 'sit (imperative)'; /k/ > [k] /kə/ 'go' > [dəz] 'go out'; /b/ > [w] /bə/ 'pile' > [təw] 'piled'. Comparing Ronghong and Mawo dialect (H. Sun 1981a) forms, we can see that a similar type of weakening has occurred historically to pre-initial consonants in Ronghong (e.g. Mawo /khs/, Ronghong /xsə/ 'god').

In general, stress is trochaic, which leads to the loss of second syllables in bisyllabic words, particularly if the final is /ə/, e.g. /sə/ prefix + /tchə/ 'drink' > [sətə] 'drink!' (imperative).

There is a pattern of vowel harmony where the vowel of the first syllable of a compound or prefix + root combination harmonizes wholly or partially (e.g. becomes fronted) with the vowel of the second syllable or root (e.g. /wə/ 'bird' + /ʃpu/ 'flock' > [wuʃpu] '(wild) pigeon'; /ha/ 'ten' + /tʂi/ 'one' > [hatʂi] '11'). If the second syllable of a compound or prefix + root form has r-colouring, in many cases the first syllable also takes on r-colouring (e.g. /me/ 'not' + /we/ 'reduce' > [me'-we] 'unceasingly').

When a collocation of consonants due to derivation or compounding results in an unacceptable cluster of consonants, an epenthetic schwa is inserted to break up the cluster (e.g. [zdzi-tʂop-əm] [illness-heal-NOM (<-m)] 'doctor').

These phonological processes (the stress pattern, harmony, epenthesis) occur within a unit that can be defined as the phonological word.

Many lexical items in Qiang allow free variation of the pre-initial, initial or final consonant (e.g. *phiʂ ~ phiχ* 'white'; *ɳiq ~ ɳix* 'black'; *muxü ~ muɸü* 'smoke'; *squ ~ χqu* 'mouth'; *xupa ~ fupa* 'fur'; *mutu ~ mutup* 'sky'; *qha ~ qhaq* 'bitter').

2 THE NOUN PHRASE

The order of the constituents in an NP is given in (3):

- (3) GEN + REL + Head + ADJ + DEM/DEF + (NUM + CL)/PL

Any combination of the elements in (3) is possible, though a numeral must be followed by a classifier. Classifiers also occur with demonstratives. Adjective modifiers can appear either as non-nominalized post-head adjectives (generally simple adjectives) or nominalized adjectives in pre-head relative clause structures (generally complex modifiers). When more than one adjective appears in an NP, the order of the adjectives in terms of type of adjective is the mirror image of that in English. Example (4) contains two NPs (bracketed).

- (4) [the:-teə̯-teiwsa ba-the-zı] piena-la [zawa ba-xsə-zı] sə.
3sg-GEN-house old-that-CL beside-LOC rock big-three-CL have/exist
'There are three big rocks beside that old house of his.'

NPs and pronouns can be omitted if they are recoverable from the context.

A noun in Qiang is an element that can take definite marking and case marking. A deverbal noun may be formed from a plain verb or a noun + verb combination using one of two nominalizers: /-s/ for inanimate nouns, e.g. /nə/ 'sleep' + /-s/ > /nəs/ 'bed'; /-m/ (</mi/ 'person') for animate nouns, e.g. /wu/ 'help' + /-m/ > /wuam/ 'servant'. Nouns can also be formed from adjectives by simply adding one of the definite markers, e.g. /ɳiq-le/ [black-DEF] 'the black one'. As in all Sino-Tibetan languages, in compound nouns where one noun modifies another, the modifying noun always precedes the modified noun, e.g. /sə-kuatṣa/ [wood-bowl] 'wooden bowl'.

Many Qiang speakers also speak Chinese (and/or Tibetan), and education is generally in Chinese. Qiang has absorbed a large number of Chinese and Tibetan loanwords (see G. Liu 1981 on Tibetan loans). Verbs borrowed into Qiang are treated as nouns, and take the verbalizing suffix /-tha/, if monosyllabic, or the verb /-pa/ ‘to do’, if polysyllabic, e.g. /tuen-tha/ ‘squat’ < dūn (蹲); /eunqian-pa/ ‘train’ < xùnlìan (訓練). A few intransitive stative verbs, older Chinese loans, have /-ti/ (< Chinese nominalizer *de* (的)) instead of /-tha/, e.g. /lan-ti/ ‘blue’ (< Chinese *lán de* (藍的)).

In Qiang only the natural gender of animals is marked: /-mi/ or /miahə/ for females; /zdu/, /zø/, /xø/, /ci/, and /pi/ for males, depending on the animal. The diminutive has the concrete sense of ‘child’; it is not used for hypocoristic or other abstract uses: /-tṣuə ~ tṣü/ (< /tṣuə/ ‘child’ (general)); /-kł/ (dogs), /-zdue/ (sheep).

Most Qiang kinship terms are comprised of a vocalic prefix plus a root, where the prefix harmonizes with the vowel of the root, e.g. *a-pa* ‘grandfather’, *u-tuma* ‘grandmother’.

It is obligatory to have either definite or indefinite marking on all referential count nouns. Of the two definite markers, /le/ and /te/, /le/ is used more frequently for animate referents, while /te/ is used more frequently for inanimate referents. The definite markers are occasionally used with proper names. Newly introduced referents and predicate NPs generally take the non-referential/indefinite marker /ke/. Newly introduced referents can also be marked with just a number and a classifier.

Number marking on nouns is singular (zero) or plural. There are two plural markers: /ha/, used for the vast majority of referent types, and /le/ (> /yłe/ ‘few’), used only on words referring to people, e.g. /ipi-le/ ‘uncles’. Following a proper name, the plural means that person ‘and others’, e.g. /upu-bəs-la-ha/ [uncle-snake-DEF-PL] ‘Uncle Snake and others’. The number ‘one’ can be added to /ha/ to form /aha/ ‘a few’. Plural marking is not used when a numeral + classifier phrase is used.

The personal pronouns are given in Table 39.2.

The third person pronoun /qupu/ is used to refer to a third person who has a close relationship to the speaker, such as a spouse, and as a logophoric pronoun, that is, in indirect quotes when the person quoted and the one being talked about are the same. The form [the:] is a reduced form of /the ze/ (‘that’ + classifier) ‘that one’. Reflexive pronouns for first and second person are formed by reduplication of the regular pronouns. The reflexive pronouns are also used as emphatic pronouns. There are no possessive/genitive pronouns or prefixes.

The demonstrative pronouns mark only proximate /tse/ (plural [tsaha]) and distal /the/ (plural [thaha]). They must take a classifier, or the vowel can be lengthened to represent a classifier (e.g. /tse-ze/ or [tse:]). The same form of the demonstrative pronoun is used for both free pronoun and adjectival uses.

The main interrogative pronouns are given in (5):

- | | | | | |
|-----|-----------------------|------------------|--------------------|----------------|
| (5) | <i>sə-(le)</i> | who? / whoever | <i>ɳawu / ɳawe</i> | how much/many? |
| | <i>tca -la ~ tea:</i> | where? | <i>ɳi:ke</i> | how? |
| | <i>ɳiyi</i> | what? / whatever | <i>təho:</i> | when? |
| | <i>ɳiyi-χuanqi</i> | why? | | |

The numeral system is a simple decimal system, with ‘one’ to ‘ten’ being unique forms, ‘11’ to ‘19’ being ‘ten’ + ‘one’ etc., ‘20’ to ‘90’ being ‘two’ + ‘ten’, etc., and the numerals in between being ‘two’ + ‘ten’ + ‘one’, etc. There are no ordinal numbers in Qiang; /tei-qə'-le/ (most-front-DEF) ‘the first one’, /tsə-steke-le/ (this-back-DEF) ‘the next one’, /thə-steke-le/ (that-back-DEF) ‘the one after that’ are used for the ‘first’ to ‘third’, but after that the cardinal numbers plus classifiers are used as ordinal numbers.

TABLE 39.2 THE QIANG PERSONAL PRONOUNS

	Singular	Dual	Plural	Reflexive (sg/pl)
1	<i>qa</i>	<i>teizzi</i> [tei-zi]	<i>tei-le</i>	<i>qa-qəi/teiil-teile</i>
2	<i>ʔü</i>	<i>ʔizzi</i> [ʔi-zi]	<i>ʔi-le</i>	<i>ʔi-ʔi:/il-ilie</i>
3	<i>the:/qupu</i>	<i>thizzzi</i> [the-zi]	<i>them-le</i>	(<i>the:</i>) <i>ŋiqi/ŋil-ŋile</i>

Classifiers or measure words are necessary whenever a number or demonstrative pronoun is used. A number of both types are clearly related to nouns, e.g. /qu/ ‘mouthful’ < /ʂqu/ ‘mouth’, /sa/ ‘classifier for sections’ < /ʂaq/ ‘joint’. Many others are loans from Chinese. Some common classifiers: /ze/ general classifier, used for people and many other objects; /la/ for stick-like objects; /xʂe/ for one item of a pair.

3 NOMINAL RELATIONAL MORPHOLOGY

The semantic and pragmatic roles of the major arguments of a sentence are mainly expressed by word order and the following enclitics:

Topic marker	<i>ŋuəŋi</i>
Agent, instrumental, ablative, perative ('through', 'along') marker	<i>wu</i>
Genitive, recipient marker	<i>teš</i>
Locative, allative marker	<i>la</i>
Locative, allative, temporal, goal marker	<i>ta</i>
Locative, temporal marker	<i>ka</i>
Comitative, conjunction marker	<i>ŋa</i>
Comparative marker	<i>sɔ/ŋiki</i>

These markers are generally used alone, but there are some instances where a locative and ablative marker, or a semantic marker and the topic marker, are used together.

In both transitive and ditransitive sentences, if the actor is the topic (initial NP), then the NP representing the actor need not take any agentive marking. Generally only when there is marked word order, or when there is a need to emphasize the agentivity of the actor, is the agentive marker /-wu/ used after the NP representing the actor, as in (6).

- (6) *the:-teš pi:-xsə-la sum-wu de-l-ji ŋuə.*
 3sg-GEN pen-three-CL teacher-AGT DIR-give-CSM COP
 ‘THE TEACHER gave him three pens.’

In some cases, even when the word order is actor-undergoer, if the flow of action is marked (e.g. a third person referent is acting on a first person referent), or if the actor is inanimate, then agent marking is necessary for clarity, as in (7):

- (7) a. *mi-wu qa zə-dzi.*
 person-AGT 1sg DIR-hit
 ‘Somebody hit me.’
- b. *moŋu-wu qa da-tuə-ʔɔ-ŋa.*
 wind-AGT 1sg DIR-fall.over-CAUS-1sgU²
 ‘The wind knocked me over.’

An instrumental NP is marked by the postposition /wu/, the same form as the agentive and ablative markers.³

- (8) a. *qa* *wduaq-te-wu* *the:* *dzeta.*
1sg hammer-DEF-INST 3sg hit:1sg
'I use the hammer to hit it.'

A genitive NP appears before the noun it modifies, and can be followed by the genitive marker /-teø/, e.g. /eiautṣaq-teø puṇu/ [Little.Zhang-GEN cat] 'Little Zhang's cat'. When the relationship between two nouns in a genitive relationship is clear, as in most cases of inalienable possession, the genitive particle is not needed.

In general, the NP representing the undergoer of a transitive verb does not take any marking of its undergoer status, though if the undergoer is animate and the NP representing the actor does not have agentive marking the locative marker /-ta/ can be used after the NP representing the undergoer for clarity.

- (9) a. *the:* *qa-ta* *dzel!*
3sg 1sg-LOC hit
'He is hitting me!'

In ditransitive clauses Qiang formally distinguishes between goal, the referent at which an action is directed, and recipient, the referent who receives some object as a result of the action. The postposition used after an NP which represents a goal argument is the locative /-ta/.

- (10) *tshetsə-żə-ŋuəŋi* *qa* *the:-ta* *kəja.*
car-affair-top 1sg 3sg-LOC tell:1sg
'I told him about the car.'

The postposition used to mark an NP representing a recipient argument is /-teø/, the genitive postposition:

- (11) *sum* *the:-teø* *pi:-xsə-la* *de-l.*
teacher 3sg-GEN pen-three-CL DIR-give
'The teacher gave him three pens.'

The NP representing a benefactive, the referent for whose benefit an action is performed, is also generally marked with the postposition /-teø/. As the form of the clause involving a benefactive argument with this marking is the same as the genitive construction, the addressee must depend on the context for proper interpretation.

- (12) *ma:* *qa-teø* *fa* *spetc.*
Mom 1sg-GEN(BEN) clothing sew
'Mama sews my clothes.' / 'Mama sews clothing for me.'

It is also possible to mark a benefactive with the postposition /χuanqi/ 'because', 'in order to', 'in place of', 'for':

- (13) *the:* *qa-χuanqi* *sə* *ge-kə.*
3sg 1sg-because firewood chop-go
'He went to chop wood for me.'

A comitative relation can be represented by two NPs conjoined together in a single large NP with the comitative/conjunction particle /-nq/ between them, in which case the meaning is that the two referents are doing something together, or one NP can be made the topic

and the second NP is then followed by /-ɳa/. In the latter case the meaning is that the referent represented by the topic NP does something ‘with’, or ‘follows’ the second referent in doing something.

- (14) a. *mutsitsu-ɳa-tugantsu zəpəq-ta ho-lua-ji.*
Mutsitsu-com-Tugantsu earth-LOC dir-come-CSM
'Mutsitsu and Tugantsu came to earth.'
- b. *qa khumtsi-ɳa tıantsə-ka ka:*
1sg Khumtsi-COM store-LOC go:1sg:PROSP
'I'm going to go to the store with Khumtsi.'

The particle /ɳa/ is also used for the arguments of certain verbs, such as in (15):

- (15) a. *?i?i-ɳa tce-khue.*
2sgREFL-COM NEG.IMP-upset
'Don't be angry at/with yourself.'

In comparative clauses, the positive comparative marker is /-sə/, while the negative comparative is /-ɳiki/.

- (16) a. *qa the:-sə tce-fia.*
1sg 3sg-COMPAR still-white:1sg
'I am lighter (in colour) than him.'
- b. *qa ?ū-ɳiki ma-wa.*
1sg 2sg-COMPAR NEG-big:1sg
'I am not as big as you are.'

The unmarked locative/allative postpositions are /-ta/ and /-la/, e.g. /petein-la kə/ [Beijing-LOC go] ‘go to Beijing’. It is also possible in some cases for the locative postposition to appear as a lengthened vowel on the noun representing the location, e.g. /zdzyta-la ~ zdzyta:/ ‘in/to Chengdu’.

The locative /-ka/ generally marks containment in some sort of vessel or movement in or out of one, e.g. /pankonjə-ka/ [office-LOC] ‘in the office’.

The ablative postposition is the same form as the agentive and instrumental postposition, /-wu/. It can be used alone (/petein-wu/ ‘from Beijing’), or with one of the other locative post-positions introduced above (/məq-ta-wu la/ [above-LOC-ABL fly] ‘fly from above (from on high)'). It can also be used with a perlocative meaning, ‘along’.

In a possessive construction, if the object is owned by the person, or is physically part of the person, then there is no marking on the possessor (17a), but if the situation is simply one of temporary possession and not ownership, then the possessor takes the locative/dative marker /-ta/ (17b). If the situation involves ownership of an object or relationship (brother, sister, etc.) (17c), then no marking is required on the nouns, but the causative suffix must be used with the appropriate verb of possession, which differs with the type of noun possessed or the nature of the possession (see below).

- (17) a. *qa dzoqu-ji-tua wa.*
1sg leg-two-CL have/exist:1sg
'I have two legs.'
- b. *?ū-dzobu-le qa-ta sə.*
2sg-key-DEF 1sg-LOC have/exist
'I have your key.'

- c. *khumtsi tutʂə-yzə-z-i* *z-i-zə.*
Khumtsi younger.brother-four-CL have/exist-CAUS
'Khumtsi has four younger brothers.'

4 THE VERB COMPLEX

The verb complex is defined as the predicating part of the clause, not including the sentence-final mood particles. In its most expanded form, the verb complex has a manner adverbial, an orientation prefix, a negative prefix, an aspectual prefix, the verb, causative marking, aspect marking and person marking, in that order. A particle which means 'again' can also follow the verb, though does not occur with the negative or the aspectual prefix. There is an adverb of degree that follows some intransitive stative verbs. If there is an auxiliary verb, then it follows the main verb.

The verb in Qiang can be defined as an element that can take the directional prefixes, the negative prefix and/or the causative suffix. Many verbs in Qiang can be used either intransitively or transitively. There is no applicative construction for adding an undergoer or benefactive argument. It is possible to derive verbs from nouns by putting the verb /pə/ 'to do' after the noun.

Intransitives can be formed by reduplicating the verb to make a reciprocal, e.g. /ku/ 'curse' > [ku ku] 'curse each other'. The verb in this construction can either take one plural argument or two arguments, one of which is marked as an indirect argument by the comitative postposition /-ŋa/. While there is marking of the reciprocal on the verb, there is no marking of reflexives on the verb, and also no middle voice or passive constructions.

Transitive verbs can be formed from intransitives using the causative suffix /-zə/, which increases the valency of intransitive (18a), transitive (18b) and ditransitive (18c) verbs. Causatives derived using this suffix can be permissive or causative, and either direct or indirect causatives. The NP representing the causee can take agentive/instrumental marking if the basic clause from which it is formed is transitive. The person marking on the verb reflects the person and number of the causer.

- (18) a. *qa tsə tu-χsu-zə-ja.*
1sg water DIR-boil-CAUS-CSM:1sg
'I brought the water to a boil.'
- b. *qa the:-wu pəitsə-e-ze zə-pə-za.*
1sg 3sg-AGT cup-one-CL DIR-buy-CAUS:1sg
'I made him buy a cup.'
- c. *qa χumtsi-wu ləyʂ-te-pen khumtsi-ta tə-χua-zə:*
1sg Xumtsi-AGT book-DEF-CL Khumtsi-LOC DIR-buy-CAUS:1sg:PROSP
'I'm going to make Xumtsi buy the book for Khumtsi.'

Three sets of verbs reflect an old voiced–voiceless (aspirated) contrast in simplex-causative pairs (see LaPolla, this volume, Chapter 2).⁴ This type has a sense of direct causation, and not permission; they cannot take a further productive causative suffix.

- (19) *simplex causative*
- | | | |
|---------------|---------------|--------------------------------------|
| <i>de-pe</i> | <i>he-phe</i> | tear (of clothes) |
| <i>da-ve</i> | <i>ha-qhe</i> | break (of bowls, etc.) |
| <i>da-vli</i> | <i>ha-χli</i> | break (in two) (of tree limbs, etc.) |

There is also one clear example of a reflex of the old Proto-Sino-Tibetan *s- causative prefix (see Chapter 2): *etcə* ‘feed’ < *tchə* ‘eat’ (with assimilation of the prefix to point of articulation of the initial). While this is the only such example found in Qiang, it matches well with forms found in closely related languages and dialects; the lack of examples may simply be due to lack of data or the obscuring of the prefix by assimilatory processes.

Intransitive stative verbs form a separate class from transitive and activity intransitive verbs, so can be called ‘adjectives’. They can be predicates without the use of the copula, and take the same person marking forms as other intransitive verbs, but unlike verbs, they can be nominalized using the definite and indefinite markers, and those representing gradient concepts can take the postpositive adverb /-wa/ ‘very’. The meaning of reduplication for most verbs is reciprocity, while the meaning of reduplication for adjectives is intensification or plurality. Intransitive stative verbs are comparative even without overt marking of comparison.

There are only three types of reduplication of adjectives: AA (marks plurality; *patspats* ‘some round things’), Au:A (intensification; *patsú:patq* ‘very round’), AAu: (plurality plus intensification; *patspatṣú*: ‘some very round things’; /u:/ is a stressed syllable added to the reduplicated form).

There are four main existential/locative verbs: /ʂə/, for inanimate referents that are not in containers or immovable or inalienably connected to some larger entity; /le/, for a referent located in a containment of some type; /zj/, for animate referents; and /we/, for possession of qualities and for immovable referents or referents inalienably connected to a larger entity.

Following are some auxiliary verbs and their meanings/uses: /yʐə/ learnt ability; /dʐə/ or /qe/ natural (physical) ability; /gu/ ability to fit into something else; /kʊ/ willingness to perform an action or to allow others to perform an action, or in some cases the possibility of some situation; /xʂu/ ‘to dare’; /ʐze/ ‘ought to’; /se/ permission or lack of it; /etcəq-lu/ ‘want’; /zulu/ ‘wait’ (used in optative constructions); /dze/ experiential aspect. These verbs take a complement clause that is not nominalized, but does not take person marking.

Person marking suffixes on the verb generally reflect the person and number of the actor of a transitive clause, the single direct argument of an intransitive clause, and the causer of a causative construction.

All verbs can take person marking, but only animate arguments are marked. In some contexts, such as nominalizations and some complement clauses, no person marking is used, while in other contexts, such as with some third person plural actors, the person marking is optional. Table 39.3 gives the forms of the suffixes.

Another set of suffixes can be used for marking empathy with an affected salient non-actor human referent.⁵ These forms are given in Table 39.4.

A set of eight verbal prefixes marks the orientation of the action vis-à-vis the speaker. Not all verbs can take all eight prefixes; e.g. /tse/ ‘watch, read’ only takes one prefix. The form of the prefix follows the rules of vowel harmony. An example of the prefixed forms for the verb /lə/ ‘look’ is given in (20) below.

TABLE 39.3 THE QIANG PERSON MARKING SUFFIXES FOR INTRANSITIVE VERBS

	1	2	3
Singular	-ə	-n	zero
Plural	-i	-i	-tei

TABLE 39.4 THE QIANG NON-ACTOR PERSON MARKING SUFFIXES

	1	2	3
Singular	-ʂɑ	-san	-wə~u
Plural	-ʂɑ ^j	-sai	-wə~u

- (20) *təl* look upwards *hal* look downwards
zəl look towards centre *dal* look outwards from centre
nəl look upstream *səl* look downstream
əl look in *hal* look out

Aside from marking the actual direction of the action, the orientation prefixes are also used to mark a change in the *Aktionsart* of the verb, from state or activity to achievement or accomplishment, e.g. /ba/ ‘big’, [təwa] ‘become big’; /təhə/ ‘eat’ (activity), [sətəhə] ‘eat’ (active accomplishment); /ylu/ ‘roll’ (activity), [doylu] ‘roll’ (active accomplishment). This change is often used to affect something like a perfective sense. For this usage usually only one of the eight prefixes is regularly used, but which prefix is used differs between verbs.

Use of a different orientation prefix can also affect the interpretation of the agentivity or volitionality of the argument of some intransitive verbs, or can be part of the causativization or transitivization of some verbs, e.g. [da-la] ‘slip’, [fia-la] ‘slide’; [do-tshu] ‘fall (e.g. of fruit from tree)’, [fio-tshu] ‘pick (fruit from tree)’ (see also (19); cf. C. Huang 1997: 73).

There are several suffixes for marking the different types of aspect. Prospective aspect is marked by the suffix /-a:/, which replaces the root vowel of the verb if it is /ə/ or /a/, as in (21), where /məpa/ becomes [məpa:] in the second clause.

- (21) *pəs məpa wa, təp-ŋi tsə-s tca-məpa: lu.*
 today cold very tomorrow this-COMPAR still-cold:PROSP will
 ‘Today is very cold, and tomorrow is going to be even colder than this.’

The most common aspect marking is /-ji/, which marks a recent change of state or situation.

- (22) a. *me': ci-ji.* b. *me': de-ci-ji.*
 rain fall-CSM rain DIR-fall-CSM
 ‘It’s started raining.’ ‘It has already rained (and stopped).’

As /-ji/ expresses a recent change of state, it can have the sense of an inchoative aspect marker. This suffix can also be used together with prospective aspect marking to express the idea ‘about to v’.

- (23) *me': ea:-ji.*
 rain fall:PROSP-CSM
 ‘It’s about to rain.’

The form /-jy/ marks an action that has ‘already’ begun or been carried out, like a perfect:

- (24) *qa a-tian u-zəlu-jya.*
 1sg one-hour DIR-wait-PERF:1sg
 ‘I have already waited for one hour.’

The prefix [t̥ee- ~ t̥ei- ~ t̥ea- ~ t̥eo-] ‘still’, ‘yet’ is used to express present progressive actions. In some cases this prefix has the same form as the prohibitive prefix, but as the two prefixes appear in different types of contexts (and the prohibitive is not used with prospective marking) there usually is no problem of ambiguity.

- (25) *me': t̥ee-ci.*
rain still-fall
'It's still raining (has been raining all along).'

The marking of an imperative sentence involves the same prefixes used for direction marking, plus the optional polite imperative particle /-na/. The prefix, which may be any one of the directionals, is stressed, unlike non-imperative directional prefixes.⁶ In an imperative clause the person marking is optional, though the imperative sense is stronger if person marking is used (*ə-zə-n-a!* [DIR-eat-IMP] ‘Eat!’ vs *ə-zə-n-na* [DIR-eat-2sg-IMP] ‘You eat!’).

The prohibitive is expressed by the prefix /tea-/ ([*tea* ~ *tee* ~ *t̥ea* ~ *t̥eo*]) (< PTB *ta-), which appears in the same position as the negative prefix. For example: *ha-t̥ea-kə!* [DIR-NEG.IMP-go] ‘Don’t go out!’.

Polarity questions are marked by rising intonation and by the addition of the clause-final particle /-na/ (2sg /-n/ plus question particle) for 2sg actors/topics, or /-ŋua/ (often pronounced [wa]) for all other persons or numbers.

- (26) a. *?ū z̥me ŋua-n-a?* b. *the: z̥me ŋua-ŋua?*
2sg Qiang COP-2sg-Q 3sg Qiang COP-Q
'Are you a Qiang (person)?' 'Is s/he a Qiang?'

Polarity questions can also be formed by repeating the entire verb complex, with the first token in the positive and the second token in the negative, and the question particle on both tokens:

- (27) *?ū zdzyta: ha-qə-n-a ha-mə-qa-n-a?*
2sg Chengdu DIR-go-2sg-Q DIR-NEG-go-2sg-Q
'Did you go to Chengdu?'

Question particles are used even if interrogative pronouns are used in the sentence.

Epistemic and root modals are expressed using the same structure, a nominalized clause followed by the copula, or the auxiliary verb /kze/ ‘ought to’. Person marking on the verb is optional in this construction, but if it appears, it is the non-actor marking that is used.

- (28) *the: t̥ea-u-la lu-s ŋua-ŋuə.* (< wə)
3sg home-LOC come-NOM COP-3sgu
'S/he must come home!'

The potential to perform an action is also expressed by the use of auxiliary verbs, with the choice of auxiliary verb depending on the type of potentiality (see the discussion of verb types earlier).

In Qiang the unmarked clause is assumed to represent knowledge that the speaker is sure of, like a direct evidential. To express the fact that what the speaker is reporting is hearsay, /-i/ (< /jə/ ‘to say’) is added to the end of the verb complex.

- (29) *the: zdzytā ha-qə-i.*
3sg Chengdu DIR-go-HS
'He went to Chengdu.' (indirect evidential, hearsay)

If rather than hearing about an action, one sees the result of the action (but not the action itself) and infers that the action took place based on that evidence, this lack of direct evidence is expressed by adding the particle /-k/ to the verb, after the change of state marker and any other aspect markers, but before the person marking.

- (30) *dzy de-zge-ji-k.*
door DIR-open-CSM-INFR
'The door is open!' (guess)

If the situation is such that one has just discovered the evidence of the action (mirative), then this can be expressed by adding the particle /wa/ after /-k/.

5 ADVERBIALS

The relative degree marking adverb /teə/ ([teə ~ tei ~ tea ~ t eo]) (also used to mean 'still', 'yet') is generally used in the comparative construction (*tea-wa* [still-big] 'relatively big'). The superlative of adjectives and some stative verbs is marked by the prefix /tei/: *tei-wa-la-la* [most-big-that-CL] 'the biggest (stick-like object)'. The form of the superlative is similar to one of the harmony forms of the preverbal adverb for marking a relative degree, but the superlative does not undergo vowel harmony.

The negative adverbial prefix /mə/ ([ma ~ me ~ mi ~ mə ~ mo]) appears after the orientation prefix, e.g. /ha-ma-qa/ (orientational prefix + negative + 'go') 'didn't go out'. The same negative adverbial prefix is used for all types of negation except the prohibitive.

Generally, manner adverbs take /-ŋi/, though if reduplicated, then /-ŋi/ is not used:

- (31) a. *akha-kha ə-tchə* b. *akha-ŋi ə-tchə*
slow DIR-eat slow-ADV DIR-eat
'eat slowly' 'eat slowly'

Some adjectives, when acting as manner adverbs, take /-ji/ or /tei/ rather than /-ŋi/, e.g. /na/ 'good' > /na-ji/ 'well'.

6 THE CLAUSE

The order of the NPs in the clause is affected by pragmatic factors such as topicality, but the verb always appears in final position. The only exception to this is the occasional afterthought clarification of an NP that was omitted or expressed as a pronoun in the clause. The most unmarked word order in the clause is given in (35).

- (32) (TEMP)-(LOC)-(actor)-(goal/recipient)-(undergoer)-VC-(PRT).

The main type of relative clause is a pre-head nominalized clause. Which nominalizer a relative clause takes depends on the semantics of the head noun. If the head noun is an undergoer or other non-instrument, including an inanimate actor, then the genitive marker /-te/ is used:

- (33) *pətsa-ŋi tə-bəl-jy-te tʂuatsə*
[just.now-ADV DIR-make-ASP-NOM]RC table
'the table just made'

If the head noun is an instrument (even if it is animate), then the nominalizer /-s/ is used:

- (34) *doqu-żə-s kħuə-le*
afraid-CAUS-NOM dog-DEF
'The dog used to frighten people.'

If the head noun is an animate actor, then the nominalizer /-m/ is used. This form derives from the word /mi/ 'person', but it has fully grammaticalized, to the point that it can be used together with /mi/ as the head noun. (Contrast (34) and (35).)

- (35) *qa ha-żdże-m kħuə-le*
1sg DIR-bite-NOM dog-DEF
'the dog which just bit me'

Complement clauses of most secondary verbs are not nominalized (e.g. (36)), but complements of the copula are generally nominalized, mostly by /-s/ (e.g. (37)) but sometimes by /-m/.

- (36) *the: e-ze stu nə mo-xṣu.*
3sg one-CL alone sleep NEG-dare
'S/he doesn't dare sleep alone.'

- (37) *teile-γuəŋi pə-s pies γuə.*
1pl-TOP buy-NOM meat COP
'What we need is meat.'

Generally hypotactic clause juncture involves nominalization of the subordinate clause, with various particles used to express the relationships between the actions expressed by the two clauses. If the action expressed by the second clause preceded the action expressed by the first clause, then the predicate can take the form /ma-tei-Verb/ 'had not yet Verb' in the first clause and be nominalized by /-te/, as in (38).

- (38) *nəs, qa ma-te-kə-tc, the: qa səimi de-l.*
yesterday 1sg NEG-yet-go-GEN 3sg 1sg fruit DIR-give
'Yesterday before I left, s/he gave me a package of fruit.'

Another option is to have the initial clause nominalized by /-s/ and followed by /qe^l/ 'before'. If instead the action expressed by the second clause is said to follow the action of the first clause (whether or not the first action was completed), then the particle /ŋike/ or /ŋiantei/ 'following' is used at the end of the first clause:

- (39) *qa stuaha sə-tchħə-ŋike, suə xuəla.*
1sg food/rice DIR-eat-following teeth wash
'After I eat a meal, I brush my teeth.'

To make explicit the idea that an action immediately followed another, the particle /ŋiaufu/ is used instead of /ŋike/ or /ŋiantei/. To mark the purpose of an action, the postposition /χuani/, can be used, either after a noun, a clause or a nominalized clause. The cause of an action or situation can also be marked by /χuani/.

A cause-effect relation can also be marked by adding the instrumental postposition /wu/ or the manner adverbial marker /ni/ to the end of the first clause:

- (40) *the:-dzoqu-le dagħ-wu, pitc sei ma-lə-jy.*
3sg-foot-DEF break-INST now walk NEG-able-ASP
'His/her foot is broken so he/she cannot walk now.'

- (41) *the: dziq ha-qə-ŋi die-ße.*
 3sg cliff DIR-go-ADV DIR-die
 ‘S/he died from falling off the cliff.’

To express the concessive, the phrase /ha-ŋuə̥-lu/ (DIR-COP-‘come’) is added to the end of the first clause:

- (42) *qa quaha ha-χuəla-haŋuə̥lu, xo:ts he-me-qhua.*
 1sg face DIR-wash-although beard DIR-NEG-shave:1sg
 ‘Although I washed my face, I didn’t shave.’

In quoting another’s speech, the quoted speech generally follows the NPS representing the speaker and addressee, and is followed by the verb /jə/ ([jə ~ ji]) ‘to say’. Both direct and indirect quotation are possible. If there is a more specific verb of asking or replying, then this verb may precede the quoted speech, though the verb /jə/ ‘to say’ still follows the quote, as in the two tokens of this structure in (43):

- (43) *χutsqha-le-wu tə-ke'l-kui, 'ha! ɻi ɳi-χuanji kə-zei-n-a?'*
 stomach-DEF-AGT DIR-ask-HS EXCL 2sg what-because DIR-cry-2sg-Q
jə-kui, ' ɳiχuanji kə-zei-n-a?' jə-kui-tu, kapətʂ-tou
 say-HS why DIR-cry-2sg-Q say-HS-LINK orphan-DEF:one:CL
he-ʐgue-kəi, 'qa-ŋuə̥ni ep lə me-ʐə-i
 DIR-answer-HS 1sg-TOP father also NEG-have/exist-CSM
əw lə me-ʐə-i, gəs-ɳi ha-xtʂəp ɳiaufu
 mother also NEG-have/exist-CSM night-ADV DIR-dark as.soon.as
ʂkup-le qa dzə-la:, jə-kəpə-ʂə, ...
 orangutan-DEF 1sg eat-come:PROSP say-HAB-LINK
 ‘The stomach asked him, “Why are you crying?”, he said, “Why are you crying?”
 The orphan answered, “I have neither father nor mother. As soon as it gets dark, the orangutan (a boogeyman in such stories) is going to come eat me.”’

In terms of cross-clause coreference, there are neither accusative nor ergative syntactic restrictions on control of the zero anaphor of the second clause.

NOTES

- 1 I would like to thank R.M.W. Dixon and Alexandra Aikhenvald for valuable comments on a draft of this chapter.
- 2 U is the gloss for a bound non-actor marker, e.g. ‘1sgU’ means ‘1sg non-actor’.
- 3 See LaPolla 1995b on the development of such isomorphic patterns in Tibeto-Burman languages.
- 4 See H. Sun 1981a: 192–3 for more examples of this phenomenon in the Mawo dialect of Qiang.
- 5 This distinction of actor vs non-actor parallels the use of agentive and non-agentive marking on NPS (as in exx. (6)–(9) above; see LaPolla 1992a, 1994, 1995a), and makes the Qiang system quite different from many of the other Tibeto-Burman person marking systems, which are hierarchical, that is, mark person primarily, and not semantic role (see LaPolla 1992b, 1994 for discussion of person marking).
- 6 In some cases the prefix usually used for the imperative is different from the prefix usually used for the other uses of the directional prefixes: *sə-z* ‘(S/he) ate.’ vs *ə-z* ‘Eat!'; *sə-te* ‘(S/he) drank.’ vs *ə-te* ‘Drink!’

BIBLIOGRAPHY

- Chang, Kun (1967) ‘A comparative study of the Southern Ch’iang dialects’, *Monumenta Serica* 26: 422–443.
- Chirkova, Katia (2012) ‘The Qiangic subgroup from an areal perspective: a case study of languages of Muli’, *Language and Linguistics* 13.1: 133–170.
- Editorial Board (1985) *Aba Zangzu Zizhi Zhou Gaikuang (Overview of Aba Tibetan Autonomous Prefecture)*. Chengdu, Sichuan: Sichuan Nationalities Press.
- Evans, Jonathan P. (2001a) *Introduction to Qiang Lexicon and Phonology: Synchrony and Diachrony*. Tokyo: ILCAA, Tokyo University of Foreign Studies.
- Evans, Jonathan P. (2001b) ‘Contact-induced tonogenesis in Southern Qiang’, *Language and Linguistics* 2.2: 63–110.
- Evans, Jonathan P. and Jackson T.-S. Sun (2016) ‘Qiang’, in *Encyclopedia of Chinese Language and Linguistics*, Leiden: Brill.
- Evans, Jonathan P., Jackson T.-S. Sun, Chenhao Chiu and Michelle Liou (2016) ‘Uvular approximation as an articulatory vowel feature’, *Journal of the International Phonetic Association* 46.1: 1–31.
- Graham, David Crockett (1958) *The Customs and Religion of the Ch’iang (Smithsonian Miscellaneous Collection Vol. 135, No. 1)*. Washington, DC: Smithsonian Institution.
- Huang, Bufan (1987) ‘Qiangyu yuyin yanbian zhong paichi biyin de qushi’ (The tendency to lose nasals in the development of Qiang phonology), *Minzu Yuwen* 1987.5: 19–26.
- Huang, Bufan and Faching Zhou (2006) *Qiangyu Yanjiu (Study of the Qiang Language)*. Chengdu: Sichuan Renmin Chubanshe.
- Huang, Chenglong (1992a) ‘Qiangyu fufuyin de yanbian’ (The changes in the consonant clusters of Qiang), *Qiangzu Yanjiu* 1992.2: 152–157.
- Huang, Chenglong (1992b) ‘Qiang zu’ (The Qiang), in *Zhongguo ren de xingming (The Names of the Chinese People)*, Beijing: Chinese Academy of Social Sciences Press, 492–508.
- Huang, Chenglong (1993) *Zhongguo Shaoshu Minzu Yuyan Dang'an: Qiangyu Rong-honghua (Phonetic Files on China’s Minority Languages: The Ronghong Variety of Qiang)*. Beijing: Nationalities Institute, Chinese Academy of Social Sciences.
- Huang, Chenglong (1994) ‘Qiangyu xingrongci yanjiu’ (Study on the adjectives of Qiang), *Yuyan Yanjiu* 1994.2: 181–189.
- Huang, Chenglong (1997) ‘Qiangyu dongci de qianzhu’ (Verbal prefixes in Qiang), *Minzu Yuwen* 1997.2: 68–77.
- Huang, Chenglong (1998) ‘Qiangyu yinjie ruohua xianxiang’ (Syllable weakening in Qiang), *Minzu Yuwen* 1998.3: 59–67.
- Huang, Chenglong (2004) ‘A reference grammar of the Puxi variety of Qiang’, PhD dissertation, City University of Hong Kong. (Published in Chinese as *Puxi Qiangyu Yanjiu*, Beijing: Minzu Chubanshe, 2007).
- Huang, Chenglong (2008) ‘Relativization in Qiang’, *Language and Linguistics* 9.4: 735–768.
- Huang, Chenglong (2010a) ‘Qiangyu de shishizhe ji qi xianguan biaoji’ (Agent in Qiang and its related marking), *Language and Linguistics* 10.2: 249–295.
- Huang, Chenglong (2010b) ‘Qiangyu de feishishizhe ji qi xianguan biaoji’ (Non-agent arguments and their related marking), *Yuyanxue Luncong* 41: 339–367.
- Huang, Chenglong (2011) ‘Qiangzu de niangjiu gongyi ji qi wenhua’ (The art and culture of making liquor among the Qiang), *Kexue Zhongguoren* 2011.10: 22–25.

- Huang, Chenglong (2013) ‘Qiangyu zhong de shengmingdu dengji xulie’ (The animacy hierarchy in Qiang), *Hanzangyu Xuebao (Journal of Sino-Tibetan Linguistics)* 2013.7: 25–43.
- Huang, Chenglong and Shude Wang (2007) ‘Puxi Qiangyu de huati-pingshu jiegou’ (Topic–comment structure in Puxi Qiang), *Language and Linguistics* 8.2: 519–551.
- LaPolla, Randy J. (1992a) ‘Anti-ergative marking in Tibeto-Burman’, *LTBA* 15.1: 1–9.
- LaPolla, Randy J. (1992b) ‘On the dating and nature of verb agreement in Tibeto-Burman’, *BSOAS* 55.2: 298–315.
- LaPolla, Randy J. (1994) ‘Parallel grammaticalizations in Tibeto-Burman: evidence of Sapir’s drift’, *LTBA* 17.1: 61–80.
- LaPolla, Randy J. (1995a) ‘Ergative marking in Tibeto-Burman’, in Yoshio Nishi, James A. Matisoff and Yasuhiko Nagano (eds), *New Horizons in Tibeto-Burman Morpho-Syntax (Senri Ethnological Studies 41)*, Osaka: National Museum of Ethnology, 189–228.
- LaPolla, Randy J. (1995b) ‘On the utility of the concepts of markedness and prototypes in understanding the development of morphological systems’, *Bulletin of the Institute of History and Philology* 66.4: 1149–1185.
- LaPolla, Randy J. (2003a) ‘English–Qiang glossary by semantic field’, in Ikeda Takumi (ed.), *Basic Materials on Minority Languages in East and Southeast Asia (Endangered Languages of the Pacific Rim series, A03–004)*, Suita: Faculty of Informatics, Osaka Gakuin University, 153–181. (Qugu variety.).
- LaPolla, Randy J. (2003b) ‘Evidentiality in Qiang’, in Alexandra Y. Aikhenvald and R. M. W. Dixon (eds), *Studies in Evidentiality (Typological Studies in Language 54)*, Amsterdam and Philadelphia: John Benjamins, 63–78.
- LaPolla, Randy J. (2005) ‘Sino-Tibetan languages’, in Keith Brown (ed.), *Encyclopedia of Language and Linguistics* (2nd edition), London: Elsevier, 393–397.
- LaPolla, Randy J. (2011) ‘On transitivity in two Tibeto-Burman languages’, *Studies in Language* 35.3: 639–649.
- LaPolla, Randy J. (2013) ‘Subgrouping in Tibeto-Burman: can an individual-identifying standard be developed? How do we factor in the history of migrations and language contact?’, in Balthasar Bickel, Lenore A. Grenoble, David A. Peterson and Alan Timberlake (eds), *Language Typology and Historical Contingency*, Amsterdam and Philadelphia: John Benjamins, 463–474.
- LaPolla, Randy J. with Chenglong Huang (2003) *A Grammar of Qiang, with Annotated Texts and Glossary*. Berlin: Mouton de Gruyter.
- LaPolla, Randy J. and Chenglong Huang (2004) ‘Adjectives in Qiang’, in R. M. W. Dixon and Alexandra Y. Aikhenvald (eds), *Adjective Classes: A Cross-linguistic Typology (Explorations in Linguistic Typology 1)*, Oxford: Oxford University Press, 306–322.
- LaPolla, Randy J. and Chenglong Huang (2007) ‘The copula and existential verbs in Qiang’, *Bulletin of Chinese Linguistics* 2.1: 233–248.
- LaPolla, Randy J. and Dory Poa (2003) ‘Texts in the Qugu variety of Northern Qiang’, in Kitano Hiroaki (ed.), *Descriptive and Theoretical Studies in Minority Languages of East and Southeast Asia, Endangered Languages of the Pacific Rim*, series, Suita: Faculty of Informatics, Osaka Gakuin University, 77–94.
- Li, Ming, Lin Zhongliang and Wang Kang (1994) *Qiangzu Wenxue Shi (The History of Qiang Literature)*. Chengdu: Sichuan Minzu Chubanshe.
- Liu, Guangkun (1981) ‘Qiangyu zhong de Zangyu jieci’ (Tibetan loanwords in Qiang), *Minzu Yuwen* 1981.3: 19–28.
- Liu, Guangkun (1984) ‘Qiangyu fuyin yunwei yanjiu’ (A study on the consonant finals of Qiang), *Minzu Yuwen* 1984.4: 39–47, 63.

- Liu, Guangkun (1987) ‘Lun qiangyu daici de “ge”’(On the ‘cases’ of Qiang pronouns), *Minzu Yuwen* 1987.4: 50–58.
- Liu, Guangkun (1997) ‘Qiangyu fufuyin yanjiu’ (A study on the consonant clusters of Qiang), *Minzu Yuwen* 1997.4: 25–32.
- Liu, Guangkun (1998a) ‘Lun Qiangyu shengdiao de chansheng he fazhan’ (On the origin and development of tones in Qiang), *Minzu Yuwen* 1988.2: 1–8.
- Liu, Guangkun (1998b) *Mawo Qiangyu Yanjiu (Studies on the Mawo Dialect of Qiang)*. Chengdu: Sichuan Minzu Chubanshe.
- Liu, Guangkun (1999) ‘Lun Qiangyu dongci de rencheng fanchou’ (On the person category of Qiang verbs), *Minzu Yuwen* 1999.1: 30–36.
- Liu, Huiqun (2011) ‘Muka Qiangyu yanjiu’ (Study of the Muka Qiang language), MA thesis, Xianan Jiaotong Daxue.
- Ran, Guangrong, Li Shaoming and Zhou Xiyin (1984) *Qiangzu Shi (The History of the Qiang People)*. Chengdu: Sichuan Minzu Chubanshe.
- Sun, Hongkai (1981a) *Qiangyu Jianzhi (A Brief Description of the Qiang Language)*. Beijing: Nationalities Press.
- Sun, Hongkai (1981b) ‘Qiangyu dongci de quxiang fanchou’ (The category of directionality in the Qiang verb), *Minzu Yuwen* 1981.1: 34–42.
- Sun, Hongkai (1982) ‘Qiangyu zhishu wenti chutan’ (A first look at the genetic position of the Qiang language), in Minzu Yuwen (ed.), *Minzu Yuwen Yanjiu Wenji*, Qinghai: Qinghai Minzu Chubanshe, 189–224.
- Sun, Jackson T.-S. (2003) ‘Issues in Mawo Qiang phonology’, *Journal of Taiwanese Linguistics* 1.1: 227–242.
- Sun, Jackson T.-S. and Jonathan Evans (2013) ‘Mawo Qiangyu yuanyin yinxì zaitan’ (Again on the vowel system of Mawo Qiang), in Shi Feng and Peng Gang (eds), *Eastward Flows the Great River: Festschrift in Honor of Professor William S.-Y. WANG on his 80th Birthday*, Hong Kong: City University of Hong Kong Press, 135–151.
- Wang, Ming-ke (2001b) *Barbarians, the Han and the Rma*. Taipei: Sanmin Press.
- Zhou, Facheng (2010) *Han-Qiang Cidian (Chinese–Qiang Dictionary)*. Beijing: Zhongguo Wenlian Chubanshe.
- Zhou, Xiyin and Zhirong Liu (1993) *Qiang Zu (The Qiang Nationality)*. Beijing: Minzu Chubanshe.

CHAPTER FORTY

PRINMI¹

Picus Sizhi Ding

1 INTRODUCTION

Prior to the mid-twentieth century, Prinmi speakers held the extonym of Hsifan/Sifan (西蕃, or *Xīfān* in Mandarin romanization). Due to different opinions in self-identity after the establishment of the People's Republic of China, this group of people has been divided into two ethnic nationalities under the current classification scheme which recognizes a total of 55 minority nationalities. Speakers of Prinmi identify themselves as *Pǔmí* /普米 in Yunnan, whereas those living in the following three counties in Sichuan are classified as *Zàng*/藏 (i.e. Tibetans): Mùlì Tibetan Autonomous County, Yányuán County (both in Liángshān Yí Autonomous Prefecture) and Jiǔlóng County (in Garzê Tibetan Autonomous Prefecture); see Harrell (2001: 193–215) for a detailed ethnic study between these two clusters of Prinmi speakers in southwest China.

Prinmi belongs to the Qiangic branch of the Tibeto-Burman family (Sun 1982; Bradley 1997: 35–7). With widespread settlements in the mountainous area, Prinmi speakers live close to many communities of distinct Tibeto-Burman languages. Consequently, language shift, to varying degrees, to such locally dominant languages as Bai (in Lán-píng Pǔmí and Bai Autonomous County), Na (in Yúlóng Nàxī Autonomous County) and Nosu (in Nínglàng Yí Autonomous County) were reported among Prinmi communities in Yunnan. Under the successful promulgation of a nine-year compulsory education programme, aided with the use of Mandarin in television broadcasts, Mandarin has posed the greatest threat to the future of Prinmi in recent decades. With little resistance from parents, children (with or without language shift in infancy) are shifting their daily life language to Mandarin across most parts of China (see the case study on Southern Min in Ding 2016).

Lu (1983, 2001) proposed two major dialectal groups for Prinmi: Northern Prinmi and Southern Prinmi. Ding (2014: 8–10) further divides the latter into two, resulting in a tripartite grouping: Northern Prinmi, Central Prinmi and Western Prinmi. Mutual unintelligibility of varying degrees exists among these dialectal groups. Niuwozi Prinmi, the variety presented in this chapter, belongs to Central Prinmi. It is spoken in central Nínglàng Yí Autonomous County in Yunnan, which adjoins both Mùlì County and Yányuán County in Sichuan.

2 PHONOLOGY

2.1 Syllable structure

Unless treated as an independent segment, the glide is considered as part of the nucleus of the syllable. This avoids overloading the consonant inventory (but see Matisoff 1997 for a different approach). An analysis of the syllable structure is illustrated in Figure 40.1.

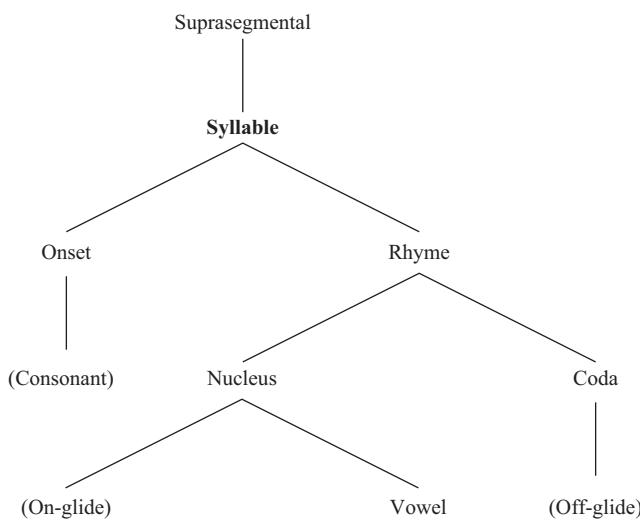


FIGURE 40.1 THE STRUCTURE OF THE SYLLABLE IN PRINMI

2.2 Vowels, glides and consonants

Niuwozi Prinmi has 15 monophthongs. These are given in Table 40.1 according to their height, frontness, roundedness and nasality.

After a fricative or an affricate, /i/ is realized as a syllabic fricative homorganic with the preceding consonant, e.g. *s^{t^R}*[sz] ‘to die’, *t^{f^H}*i^R [tʃ^Hɪ^R] ‘dog’ and *z^{t^F}*[zz] ‘four’ (super-script letters indicate surface tones of different tonal categories: R = rising, F = falling and H = high). The vowel /i/ also contrasts, in non-fricative form, with its rounded counterpart, e.g. *b^{t^H}* ‘sun’ versus *b^{t^H}*u^H ‘(of grain) heap’.

Phonetically /y/ is realized as [qi], but /ŷ/ is a true monophthong. Discounting /y/, /ŷ/, and /i/, high and mid high vowels are prone to lowering, hence /i/: [i]~[i̥], /o/: [o]~[ø], and so forth. Preliminary cross-dialectal comparison suggests that /ɜ/ is historically related to *ɛ. This central vowel contrasts with schwa, e.g. *m^{z^F}* ‘bamboo’ versus *m^{z^F}*ə ‘sky’.

Niuwozi Prinmi has three glides: /j/, /w/ and /w/. The latter two are almost in complementary distribution but contrast in such minimal pairs of native words as *kya^H* ‘melon’ versus *kwa^H* ‘(of eye) to open’. Treating a glide as part of a rhyme leads to a number of rising diphthongs. All three glides can function as the on-glide in diphthongs, but only /j/ and /w/ may serve as the off-glide:

ju,	jʒ,	ja,	jɛ̃,	jɔ;	ɥe,	ɥɛ,	ɥa,	ɥɛ̃;
wʒ,	wa,	wa,	wã	ʒj,	aw;	wɛj		

Table 40.2 displays all consonants in Niuwozi Prinmi. All of them, including sonorants, contrast in voicing. Further, the various types of oral stop show distinctive aspiration. The rhoticized stops involve a special secondary articulation (Ding 2009).

Major allophones of consonants in Niuwozi Prinmi include a set of uvular plosives which are allophones of the velar set before /a/; the palatalized fricative [j], an allophone

TABLE 40.1 MONOPHTHONGS OF NIUWOZI PRINMI

		Front		Central		Back	
High	oral	i	y	i		u	
	nasal	ĩ	ÿ				
Mid high	oral	e					o
	nasal						õ
Mid					ə		
Mid low					ɔ		
Low	oral	a				ɑ	
	nasal	ã					ð

TABLE 40.2 CONSONANTS OF NIUWOZI PRINMI

	Bilabial	Dental	Post-alveolar	Retroflex	Velar
Plosive	b p p ^h	d t t ^h			g k k ^h
Fricative		z s	ʒ ʃ	ʐ ʂ	ɣ χ
Nasal	m n̩	n ŋ			
Lateral		l ɿ			
Rhotic				r ɻ	
Affricate		dz ts ts ^h	dʒ tʃ tʃ ^h	ɖʐ ʈʂ ʈʂ ^h	
Rhoticized stop	b ^r p ^r p ^h				g ^r k ^r k ^h

of /ʒ/ before /i/ or /j/ (not applicable to /ʃ/); the bilabial fricative [ɸ], an allophone of /x/ before /u/; and the optional palatalization of /n/ to [ɲ] before /i/ or /j/.

2.3 Suprasegmentals

Prinmi morphemes are predominantly monosyllabic. Monosyllabic words may bear a high tone, a falling tone or a rising tone; e.g. *bje^H* ‘busy’, *bje^R* ‘urine’, *bje^R* ‘to fly’. The suprasegmental contrast between the high tone and the falling tone is difficult to perceive in monosyllables (Ding 2001, 2007a) due to the fact that a monosyllabic domain is too short to manifest the underlying contrast between them. However, when words are followed by clitics in connected speech, these three tones effectively contrast two pitches: high versus low. The falling and rising tones are often split into a H-L and a L-H sequence respectively in longer domains, whereas the high tone gives rise to H-H (see Table 40.3 for details).

The suprasegmental patterns observed with longer domains indicate that Prinmi tones are differentiated along two parameters: (a) the location of the H tone on the underlying quadrisyllabic prosodic domain (which covers the basic scope of a tonal category); and (b) whether the H tone spreads to the successive syllable. The combination of these two settings yields a total of seven categories.

Tone spreading is optional for Tone VII, which is indicated with the spreading feature unspecified in Table 40.3. Tonal patterns found in verb morphology point to both Tone III and Tone V as the sources for the rising tone in monosyllabic verbs, e.g. *bje^R* ‘to fly’, *gə^L-bjə^H* ‘fly away’, *gə^L-bjə^H=si^L* ‘flew away’ versus *si^R* ‘to die’, *nə^L-si^R* ‘die’, *nə^L-si^L=si^H* ‘died’. The basic tonal pattern of each category, as specified by the two parameters in the table, remains unchanged on each row. The marked tone H appearing on the syllable

TABLE 40.3 SUPRASEGMENTAL CATEGORIES IN NIUWOZI PRINMI

Category	Parameters	Quadrисyllabic	Trisyllabic	Disyllabic	Monosyllabic
I	[1] [-spread]	H-L-L-L	H-L-L	H-L	\widehat{H} (Falling)
II	[1] [+spread]	H-H-L-L	H-H-L	H-H	H (High)
III	[2] [-spread]	L-H-L-L	L-H-L	L-H	$\widehat{L}H$ (Rising)
IV	[2] [+spread]	L-H-H-L	L-H-H	L-H	
V	[3] [-spread]	L-L-H-L	L-L-H	$L-\widehat{L}H$	$\widehat{L}H$ (Rising)
VI	[3] [+spread]	L-L-H-H	L-L-H	$L-\widehat{L}H$	
VII	[4] [spread]	L-L-L-H	L-L-LH		

indicated by the first parameter is essential to a tonal category, but other tones following it are subject to truncation, as the length of the prosodic domain becomes shorter. These tonal categories form a phonological system similar to that in Japanese (Ding 2006), which is termed the ‘melody-tone system’ in Ding (2007a) to avoid problematic application of the notion ‘pitch-accent’; see Hyman’s (2009) discussion on this issue. The tonal system of Prinmi should be distinguished, on the one hand, from the better known syllable-tone system (as found in Mandarin, Burmese and Thai, etc.), and, on the other hand, from the word-tone system discussed in Mazaudon (1977: 76–84) for languages such as Central Tibetan and Tamang.

2.4 Phonological alternations

Phonologically conditioned segmental variation is relatively scarce, but ‘tone sandhi’ is pervasive in Prinmi. Most tone changes ensuing from differences in the prosodic domains where a morpheme is situated are predictable. However, extra-prosodicality may occur in prefixed verbs and numeral-classifier compounds (see Ding 2014: 60). Phonological processes which affect lexical tones include: (a) shifting of the H tone to an adjacent syllable (usually to the right); (b) reversal of the H-spreading feature; (c) merger of two prosodic domains; and (d) split of a prosodic domain into two.

2.5 Intonation

Intonation exerts a great influence on the surface tones of lexemes. This is achieved mainly by merging two prosodic domains in an intonational phrase, as in (1), where the word having lost its lexical tone in the merger process is underlined. It is also possible to insert a H boundary tone (symbolized in examples as H%) to the right edge of an intonation phrase, after the original tone is deleted and replaced by the default low tone; see (2).

- (1a) | zə̝^H to^L=bo^L | tʃ̝^H=ge^H lʒj^L |
 sheep on=FRM water=TOP heavy
 ‘On the sheep, water is heavy.’
- (1b) | tʒ^Hŋi^L=bo^L | ko^Lji^L=ge^L gə^L-tʃ̝^Hə̝^F |
 one:day=FRM crow=TOP CLC-appear
 ‘Then one day Crow came over.’
- (2) | tʃ̝^R H% | gə^L-zə̝^H | (tʃ̝^H → tʃ̝^L+H% → tʃ̝^R)
 water out-carry
 ‘the water washed (him) away.’

3 MORPHOLOGY

3.1 Lexical categories

Strictly speaking, Prinmi lacks derivational morphology that simply changes lexical categories without introducing additional meaning. There is considerable overlap between nouns, verbs and adjectives. Monofunctional adjectives such as *bu^Lla^H* ‘many’ and *gu^Hʒ^L* ‘(adjective) middle’ are scarce. Incidentally, the latter has a noun form *gu^Lʒ^H* ‘(noun) middle’; this is the only minimal pair with an identical meaning contrast in tones for parts of speech. As a rule of thumb, when a simplex non-grammaticalized word modifies a noun postnominally, it functions as an adjective, often in a noun-adjective compound. Many adjectives can also be used as nouns, verbs or both.

Eleven minor lexical categories have been identified, including the following (no suprasegmental marking for toneless morphemes):

- existentials – */f^R* ‘(abstract things) to exist’, *ʒe^F* ‘(animate) to exist’, *dʒi^F* ‘(animate) to exist among a group’, *dʒo^F* ‘(inanimate) to exist’, *kye^F* ‘(inanimate) to exist within something’, and *tʒi^F* ‘(inanimate) to exist on something’;
- auxiliary verbs – e.g. *kʰu^R* ‘must; need’, *ʒiɛ^H* ‘can’;
- demonstratives – e.g. *ta^H* ‘this’, *ə^Hta^H* ‘that’;
- pronouns – e.g. *z^H* ‘I’, *ne^R* ‘you (singular)’, *ne^Ldza^H* ‘you (dual)’, *e^L.a^H* ‘we (inclusive)’, *a^L.a^H* ‘we (exclusive)’, *ni^F* ‘she/he/it (present with interlocutors)’, *tsɔ̃^H* ‘she/he/it (absent with interlocutors)’;
- numerals – e.g. *ti^R* ‘one’, *ni^R* ‘two’, *sɔ̃^R* ‘three’;
- classifiers – e.g. *tsi* ‘(for humans)’, *bɔ̃^R* ‘(for trees)’;
- ideophones (which are image-oriented and always appear in reduplication) – e.g. *tʃatʃa* ‘an intensive state of whiteness’, *gɔ̃gɔ̃* ‘an excessive state of abundance’;
- adverbs – e.g. *ti^Lti^H* ‘slowly’, *lɔ̃ljɛ^R* ‘very’;
- onomatopes (which are sound-oriented and tolerant of paralinguistic sounds) – e.g. *kʰɔ̃-tʰɔ̃* ‘sound of weighty objects falling into deep water’;
- interjections – e.g. *wi^F* ‘expressing an unpleasant surprise’; and
- postpositions – *po* ‘under’, *to* ‘on’, *kʰu* ‘on top of’, *wu* ‘inside’, *lo* ‘outside’, *dʒe* ‘at one’s place’, *be* ‘at; to; from’ (also marking the semantic roles of locative, goal and source), *ð* ‘with (instrumental)’, *ni* ‘with (comitative)’, *bo* ‘for (benefactive)’, and *ki* ‘(beneficiary marker)’.

Many postpositions are grammaticalized from nouns. The ongoing process is changing these words into clitics; the fully cliticized instrumental can form composite variants with other clitics, e.g. *.ð* < *.a* (plural clitic) + *ð*.

3.2 Affixes

Suffixes are the most common type of affix in Prinmi. They impart additional meanings to a word, e.g. the diminutive *-tsi*: *pɔ̃^L-tsi^H* ‘flower’; the nominalizer *-ji*: *tʰjɛ̃^L-ji^H* ‘beverage’; *-di*, a nominalizer for deriving an instrument from a verb: *tsaw^H-di^L* ‘beating implement’; and *-zaw*, which signifies that an uncontrollable state related to the human body is going to start: *ja^L-zaw^H* ‘to be prone to laugh’.

Of the three negation morphemes (see Table 40.4), *ma-* is the only one capable of the additional prefix function to derive antonyms of some descriptive verbs/adjectives, e.g. *ma^L-gɔ̃^H* ‘unclean’.

TABLE 40.4 ATTACHMENT PATTERNS OF THE INTERROGATIVE AND THE NEGATIVE CLITICS

Clitic group	Attaching position	Critic type
Bare monosyllabic verb	<i>a/ma/me/tja</i> + Verb	Proclitic
Bare disyllabic verb	Verb + <i>a/ma/me/tja</i> + Verb	Endoclitic
Prefixed, no other clitics	Prefix + <i>a/ma/me/tja</i> + Verb	Mesoclitic
With other verbal clitics	(Prefix) + Verb + <i>a/ma</i> + Aspect/Modal	Enclitic

TABLE 40.5 THE DIRECTIONAL PREFIXES

Category	Prefix	Meaning	Prefix	Meaning
Person	dʒ-/də-	cislocative (towards the speaker)	tʰʒ-	translocative (away from the speaker)
Space	gə-/kʰə-	outwards	(x)ʒ-	inwards
Vertical	nʒ-	downwards	tʒ-	upwards

Most verbs can be affixed with at least one of the six directional prefixes given in Table 40.5 (see Fu 1998: 27–72 for verb prefixation in Lánpíng Prinmi).

There are two unproductive nominal affixes: the vocative prefix *a-* and the conjunctive infix *-mə-*, e.g. *a^L-pu^H* ‘grandfather’; *ŋi^H-mə^L-ŋi^L* ‘day after day’.

3.3 Inflection

A small number of verbs and auxiliary verbs in Niuwozi Prinmi are inflected for agent agreement in terms of person and number, as exemplified in Table 40.6.² The inflections are often associated with the perfective aspect but their occurrence is variable.

Reciprocal verbs are derived through reduplication, e.g. *tsʒ^Htsaw^L* ‘to fight’ < *tsaw^H* ‘to beat’; *gʒʒ^Lge^R* ‘to help each other’ < *ge^R* ‘to help’. Notice the modification of the vowel in the first syllable.

A few verbs manifest remnants of an archaic causative prefix of Proto-Tibeto-Burman. They contrast with the base form in terms of controllability and voicing, e.g. *k^Hjɛ̃^H* ‘to break (something)’ versus *g^Hjɛ̃^H* ‘(something) to break’; *p^He^F* ‘to destroy’ versus *be^F* ‘to collapse’.

3.4 Compounding

Compounding is extremely productive in Prinmi. Except for those permitted in specific patterns, affixes and clitics do not partake in compounding. Most compounds are bisyllabic or quadrasyllabic; trisyllabic ones typically involve ideophones, e.g. *bt^Hlja^Llja^L* ‘paper-thin’ and *p^Hi^Htʃa^Htʃa^L* ‘snow-white’. Iterative compounding results in complex structure. For instance, ‘fibula’ consists of three levels of compounding:

{[(k^Hə^H lju^H) ɻʒ^Lka^L] [(lo^L tʃ^Hye^H) ɻʒ^Lka^L]}
 foot stem bone out side bone

TABLE 40.6 CONJUGATION OF THE COPULA AND VERBS

Root	First person singular	Second person singular	1/2 person plural	Third person	Meaning
<i>dzi^F</i>	<i>dje^F</i>	<i>daw^F</i>	<i>di^F</i>	<i>dzi^F</i>	be
<i>pa^F</i>	<i>pjɛ^F</i>	<i>pu^F</i>	<i>pi^F</i>	<i>pa^F</i>	do
<i>t^bjɛ^R</i>	<i>t^bð^R</i>	<i>t^bjð^R</i>	<i>t^bi^R</i>	<i>t^byɛ^R</i>	drink
<i>ma^Lsi^F</i>	<i>ma^Lsjɛ^F</i>	<i>ma^Lsaw^F</i>	<i>ma^Lsi^F</i>	<i>ma^Lst^F</i>	know

3.5 Clitics

Clitics are attached to phrases rather than words. Except for negative and interrogative clitics, all others are enclitics. Major nominal clitics include the plural *.ia* and the dual *dzã*, e.g. *tʃɛ^L.ia^H* ‘children’ and *k^bð^Hdzã^H* ‘two feet’. Other nominal clitics mark topics (see section 4.9) and modifiers (see section 4.2). They can cling to a VP occasionally. Discourse clitics such as *bo* and *ne* can be considered a minor type of nominal clitic. They function much like gap fillers and can be readily omitted (as witnessed in editing recorded spontaneous speech).

Verbal clitics are richer than nominal ones in number and form. The experiential *tu* signifies experience of doing something. The durative *nð* expresses a continual state; see (21) and (28). The hortative *gi* conveys the wish to engage in some activity together. The quotative *tʃi^Lju^H* is used for marking hearsay information whose truth is not committed to by the speaker, e.g. (3). Its precise function in (3b) is to mark mirative, a psychological sense of surprise (see DeLancey 1997).

- (3a) *mja^Hbu^F=ge^L=bo^L k^bð^Hgjð^L to^L n^bz^L-t^bswa^L=si^H ***tʃi^L=i^Hju^H***
 eyelid=TOP=FRM knee on down-contact=PFV QUOT
 ‘It is said that their eyelids touch down to the knees.’*
- (3b) *pu^Hsi^L 3^H 3i^Hme^L ti^L ga^L-t^bð^R*
 last.night 1sg dream one out-appear
*3^H xã^L ti^H be^H t^b3^L-p^Law^F=si^L ****(tʃi^L=i^Hju^H)***
 1SG parrot one to TLC-transform=PFV MRT
 ‘Last night I dreamt of myself transforming into a parrot.’*

Aspect, modality and evidentiality can be expressed with sets/pairs of clitics sensitive to the factor of speaker’s controllability. Of the two perfective markers *si* and *sjɛ*, the basic one is *si*; see (9) and (15). When an activity is controllable by the speaker, *sjɛ* is chosen (for first person singular). Likewise, the pair of modal clitics, *so* and *kɔj*, differ in whether or not the wish to do something is under the speaker’s direct control; see (25). Controllability also plays an important role for the set of imperfective clitics *.ð*, *.u* and *.i*, versus the default one *.ju*; see (10) and (12). The former set conveys controllability of a situation by interlocutors, whereas the latter indicates a lack of such control. The variants of the set for controllable situations are selected according to the person and number of the agent: *.ð* (first person singular), *.u* (second person singular) and *.i* (first/second person plural).

Occurring at the end of the utterance, attitudinal clitics usually bring in some extra information about the speaker’s attitude or emotional state at the time of speaking. Commonly seen attitudinal clitics are: *gja* (expressing surprise), *ma* (making a suggestion), *pa* (speculation) and *mo* (making an assumption). Of these, *ma* and *pa* are likely to be loans from Southwestern Mandarin.

Prinmi has three specialized negative clitics: the general negator *ma*, the perfective negator *me* and the desiderative negator *tja* (often found, but not exclusively, in imperative sentences). These negators and the interrogative clitic *a* are the only ones that may precede the verb host, as they typically occupy the penultimate position in the clitic group. Depending on the number of syllables in the clitic group, the negator and the interrogative clitic may appear as a proclitic, an endoclitic, a mesoclitic or an enclitic, as summarized in Table 40.4 above.

4 SYNTAX

4.1 Grammatical system

Prinmi does not have clearly defined subjects and objects. Although unstable agent agreement is found on a small number of verbs, no other grammatical properties suggest the existence of subject in the language.

Prinmi tends to mark the agent with the instrumental clitic *ð*, especially after a third person singular noun in the form of *gð*; see (14) and (23). Sometimes this optional marking is also found with the agent of an intransitive verb. The inconsistency suggests that the development of ergativity in Prinmi is at an early stage (see LaPolla 1995).

4.2 Structure of the noun phrase

Six kinds of modifying elements may appear in a Prinmi noun phrase. According to their closeness to the head noun, the noun phrase can be analysed as having four layers (see Figure 40.2). The ‘modificator’ is a special type of modifier, typically signified by the modificatory clitic *a* or its complex variants such as *gja*. Genitive expressions, e.g. *ʒ^H-gja^H k^Hə^H* ‘my foot’, and the relative clause, exemplified in (4), are effected through the modificatory relation. The modificatory clitic introduces an attributive element to the head noun without nominalizing it, and it is often omitted; see the genitive expression in (6). Regarding its range of functions and optionality, the modificatory clitic *a* is comparable, but not identical, to *de* in Mandarin.

The following contains all the possible elements in a noun phrase, except for the adjective and postposition/clitic:

- (4) [mja^Hbu^Lt₃j^L] [tʃi^Ld^H] di^H-mi=a^L] t₃^H
 eyelid:big flood cast-NMLZ=M this
ʃi^Hba^H *ŋə^Lgjɛ^H* *f^Hba^H* *pu^Ldi^L* *n³tsi^L*
 dragon:family old.man dragon:family old.woman two:CLF
 ‘these two big-eyelidded old man and old woman of the Dragon family who have started the flood’

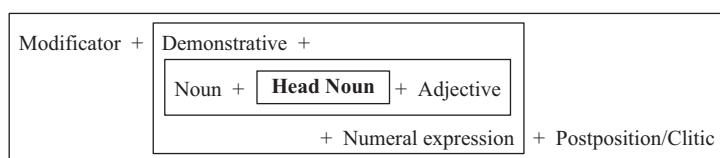


FIGURE 40.2 A LAYERED ANALYSIS OF THE NOUN PHRASE

The pair of coordinate heads (in boldface) in (4), each taking an attributive noun, is qualified by two modicators (placed within brackets), a demonstrative, and a numeral expression.

4.3 Structure of the clause

The structure of Prinmi clauses can be described elegantly with the layered analysis advanced in Role and Reference Grammar (see Van Valin and LaPolla 1997 and references therein). The smallest layer is the Nucleus, containing the predicate. The Core layer consists of the Nucleus and the core argument(s) of the predicate. Clausal modifiers such as temporals and locatives are situated in the Periphery. These three layers together form a clause expandable to a simplex sentence, as shown in Figure 40.3.

Within the scope of a clause, word order is quite rigid, starting from the adjunct/modifier in the Periphery, followed by the core argument(s) on the Core layer and then the verb on the Nucleus layer. The default order of arguments is Agent–Beneficiary–Theme, but core arguments can also appear outside the clause when they serve as sentential topics. Note that the extra clausal positions are reserved for pragmatic topics; simultaneous topics at both ends are prohibited. Sentence-initial topic is dominant while the sentence-final one is marginal. For instance (bracketing indicates the layers within a clause; N = Nucleus, C = Core and P = Periphery):

- (5) *ne^R P[*zə^Lbe^H*]P C[*tʂi^Ltʂu^L=ge^H*] N[*nʒ^L-di^L*] nʒ^L-ʃi^H ke^H.]N]C
 2sg first pestle=TOP down-cast down-go CAUS
 ‘You throw the pestle down first.’*

The presence of the temporal expression *zə^Lbe^H* ‘first’ in (5) is helpful for recognizing a left-detached argument. For the core argument *ne^R* ‘you’ to precede the temporal at the periphery, it must be situated at the left-detached position outside the clause.

4.4 Major sentence types

Exemplified in (6)–(8) are the major sentence types: declarative, negative and interrogative.

- (6) *pʒ^Hdi^H ʒ^H dʐu^L dzɪ^H.*
 frog 1SG friend COP
 ‘Frog is my friend.’
- (7) *mʒ^Htʃe^H xi^R me^{L=t^Hʃo^H}*.
 seek get PF.N=succeed
 ‘(He) searched (but) couldn’t get (one).’

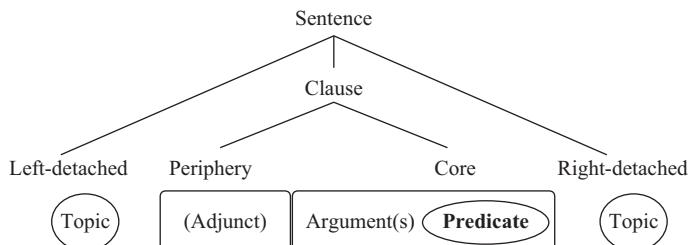


FIGURE 40.3 THE STRUCTURE OF SIMPLEX SENTENCES

- (8) *pō^Hpō^L a^H=3e^L?*
 uncle Q=EXT.AN
 ‘Is uncle (here)?’

Prinmi also has tag questions which imply a pragmatic presupposition. These are constructed by rendering the interrogative clitic as the tag, with its own intonation phrase, as shown in (9). An information-seeking question arises when an interrogative pronoun is employed *in situ*. Compared with the statement given in (10c), the formation of questions made in (10a) and (10b) apparently does not cause any changes in word order; see also (16).

- (9) *ne^R dʒ^L-ji^H=si^L a^L?*
 2SG CLC-come:2SG=PFV Q
 ‘You’ve come over, haven’t you?’
- (10a) *ne^R hʒ^Hki^L fi^L=.u^L?*
 2SG where go=IPFV:2SG
 ‘Where are you going?’
- (10b) *hʒ^He^L ba^Hdʒe^L fi^H=.ju^H?*
 who Ninglang go=IPFV:3
 ‘Who goes to the county seat of Ninglang?’
- (10c) *ʒ^H ba^Hdʒe^L fi^H=.i^Lō^H.*
 1SG Ninglang go=IPFV:1SG
 ‘I am going to the county seat of Ninglang.’

Verbs in imperative sentences generally show consistent agreement with the agent, even though it is often expressed implicitly. Negation of imperative sentences requires the use of the desiderative negator, as in (11b).

- (11a) *(ne^L) dzaw^H.*
 2SG eat:2SG
 ‘(You) eat.’
- (11b) *(ne^L.i^Lō^H) tja^L=dzit^H.*
 2PL DS.N=eat:2PL
 ‘(You) don’t eat.’

4.5 Clause compounding

When two Prinmi verbs are adjacent in a sentence, they may conjoin into a unit: the ‘double-verb predicate’ (the term ‘double’ is not to be taken literally). On account of recursion it is possible to have more than two verbs in this type of complex predicate. As with morphological compounds, the conjoining is binary in nature and it occurs on the Nucleus layer. The verbs in the complex predicate are syntactically interdependent on each other, but without subordination. A prime example of double-verb predicate (boldface) is provided in (12). Note that *fi^H ‘go’* functions as an auxiliary verb and takes the clause headed by the double-verb predicate as its complement.

- (12) *tsō^Lgy^H=ge^H **ni^f be^L** gə^L-j₃^R t₃^L-k^bj^L fi^H a^L=.ju^L?*
 attire=TOP 3SG to out-fetch TLC-give go Q=IPFV
 ‘Will (Golden Pheasant) go bring and give the clothes back to him?’

- (13) *ts $\tilde{o}^L gy^H = ge^H g\mathcal{O}^L \cdot j\mathcal{Z}^R ni^F be^L t^h\mathcal{Z}^L \cdot k^h j\tilde{\mathcal{E}}^L \mathfrak{f}t^H a^L = ju^L?$*
attire=TOP out-fetch 3SG DAT TLC-give go Q=IPFV
'Will (Golden Pheasant) go bring the clothes and give (them) back to him?'

Given in (13) is another type of clause compounding. The erstwhile shared argument (single-underlined), intervening between the verbs, has prevented formation of a double-verb predicate. Notwithstanding its identical meaning to (12), (13) takes a different syntactic structure for the complement clause, which is similar to the clause-chaining sentence in the following:

- (14) *ma $\mathcal{F} = go^L = ne^L m\mathcal{O}^F to^L bi^H \mathcal{Z}^H \cdot ty^L li^H \mathcal{Z}^H \cdot ty^L$* .
sky=INS=DSC sky on sun in-put:3 moon in-put:3
'The heavenly god placed a sun and a moon in the sky.'

Clause-chaining sentences involve conjoining of clauses at a level higher than the nucleus. They are looser in structure and laxer in prerequisites, and thus occur more commonly than the double-verb predicate in Prinmi.

4.6 Subordinate clauses

Clause subordination is found in relative clauses and periphrastic constructions where an embedded clause serves as a complement to the auxiliary verb. Sometimes verbs of cognition may also take a complement clause. Prinmi uses the nominal clause construction to convey a sense of discovery based on cursory observation of the surroundings. This construction consists of a relative clause and the marker *ti*, which is probably grammaticalized from the numeral *ti^R* 'one'; this marker serves as the structural head. For instance:

- (15) *ti $\mathcal{H} mi^F = ge^L n\mathcal{Z}^L \cdot si^L = si = a^H ti^L$* .
this person=TOP down-die=PFV=M NMLC
'This person (is) dead.' (occurring in a direct speech)

As with other relative clauses, the modificatory clitic in (15) is ommissible.

When the copula is used in periphrastic constructions, its complement clause is nominalized by the suffix *-ji* for the obligational construction, e.g. (16), or by the suffix *-mi* for the focus–presupposition construction,³ e.g. (17).

- (16) *me $\mathcal{H} t^h\mathcal{Z}^L \cdot t^h i^L \cdot ji^L dz\mathcal{Z}^L ?$*
what TLC-do-NMLZ COP
'What should (we) do?'
- (17) *fj $\tilde{z}^H = g\tilde{o}^H g\mathcal{O}^L \cdot tsaw^H \cdot mi^H dz\mathcal{Z}^L$* .
iron=INS out-hit-NMLZ COP
'What I'm saying is: (the tool) is forged from iron.'

Clausal complements to auxiliary verbs are not marked by any grammatical means, and this frequently leads to juxtaposition of two verbs, e.g.:

- (18) *kw $\mathcal{Z}^H = bo^H tja^L = kw\mathcal{Z}^H ku^H$* .
cry=FRM DS.N=cry must
'As for crying, (you) mustn't cry (any more).'
- (19) *[[[e $\mathcal{Z}^L \cdot \mathcal{O}^H = bo^L d\mathcal{Z}\mathcal{O}^L dz\mathcal{U}^L t^h i^H] 3 \cdot j\tilde{o}^H \cdot i^L = a^L = ju^L] 2 \mathcal{Z}^L \cdot fja^L fja^H] 1 tt^L pu^L$*
1PL.IN=FRM friend do want=Q=IPFV in-try one do:2SG
'(You) give it a try and see if (she) wants to make friends with us.'

That the head verbs of these complements do not constitute a double-verb predicate with the auxiliary verb is shown in the negation pattern in (18) and subordination in (19). When a double-verb predicate is negated, the negative clitic appears on the final unit, as in (20). However, the negator occurs within the complement clause in (18).

- (20) *ne^R t^H-p^he^H t^hi^L g^o^L-z^l^R m^a^R=ke^L=bo^L*
 2sg up-cough.out do out-come N=CAUS=FRM
 ‘If you don’t cough (him) out ...’ (predicate: cough-do + come-not.let).

Three orders of subordination are observed in (19). Of these, only the head of the innermost complement is adjacent to an auxiliary verb. The other instances show the intervention of clitics between two verbs of cognition and separation of the auxiliary verb from its complement by the numeral *ti* ‘one’.

Non-embedded subordination is rare. Adverbial clauses of manner may involve such subordination, e.g.:

- (21) *ni^F.a^o^L t^hz^L-f^u^Lf^u^H=n^o^ð^L t^hw^a^Lt^zj^Hp^a^L=n^o^ð^L k^o^Lkⁱ^H.*
 3PL TLC-gather=DUR voice big do=DUR sing
 ‘They sing aloud together.’

Both adverbial clauses in (21) happen to contain the durative clitic *nð*, but it does not signal subordination.

4.7 The causative construction

Prinmi has various means for expressing causation. The most important one is with the causative verb, built on the following structure:

CAUSER + SOCIATIVE / GOAL / AFFECTEE + Complement clause + Causative verb *ke*

According to the degree of volitionality and/or affectedness, the causee can be expressed as a SOCIATIVE, a GOAL or an AFFECTEE (see Shibatani and Pardeshi 2001). As shown in (22), the causee (rendered in boldface in the examples) is marked by *ki* to signify a willing participant in the sociative causation; in a situation where the causee could have avoided the event, it is marked by *be* in the indirect causation, as in (23); finally, when the causee is helplessly involved in a direct causation, it is encoded as affectee without any postposition marking, as exemplified in (24). Irrespective of how the causee is encoded, it is always co-referenced with the agent in the complement clause.

- (22) *z^H ne^L ki^H go^R po^L n^z^L-f^h^H ke^F. (SOCIATIVE)*
 1sg 2sg BNY hill below down-go CAUS
 ‘I let you go down the mountain.’
- (23) *f^hð^F=g^ð^L ki^Hp^u^L be^L n^z^L-gw^zj^H ke^L. (GOAL)*
 golden.pheasant=INS cuckoo DAT down-dress CAUS
 ‘Golden Pheasant got Cuckoo to put (the clothes) on.’
- (24) *ni^H.a^o^L z^H z^L-f^h^H me^H=kye^H. (AFFECTEE)*
 3PL:INS 1SG in-go PF.N=CAUS:3
 ‘[I wanted to go in;] they wouldn’t let me in.’

In (24) the causative verb is inflected for the perspective of third person to indicate that the causation is viewed without implying the causer’s volition and control of the situation

(see Caughley 1982: 18–21). The use of this variant form is not conditioned by morphological agreement.

4.8 The comparative construction

The comparative construction has the following structure:

Theme + Noun phrase / Clause + *to* + Descriptive verb

Strictly speaking, the comparative construction lacks morphosyntactic marking. The intended comparison is achieved only when the postposition *to* ‘than (literally: above, on)’ is construed as a sort of complementizer which introduces the standard for comparison. This is often a noun phrase, but it can also be a short clause.

In transitive sentences, the comparative construction may be embedded in a relative clause to modify a core argument. An instance of this is provided in (25), where the comparative construction is underscored.

- (25) *3^H n3^Lge^H t2^H djɛ̄lba^H to^L tʃ^Hya^H tt^L=bo^L m3^Htʃe^H p^Hð^R ma^L=k3j^H.*
 1SG again this place than good:M one=dsc seek meet N=VLT
 ‘I won’t be able to find a place better than this one again.’ (*lit.* ‘seek not meet one better than this place’).

4.9 The topic–comment construction

Prinmi has two topic markers: the frame-setting clitic *bo* for establishing discourse scenes, typically regarding a locative frame (see (1a) earlier) or a temporal frame, e.g. (26), and the general topic clitic *ge* for indicating a topic on which a comment is made. Note that *ge* is often used to mark specificity rather than a topic; see (1b) earlier, where *ko^Lji^R* ‘crow’ is newly introduced to the discourse. On the other hand, a topic can also occur without any morphosyntactic marking; see (6).

Major types of topic–comment construction include the double-topic construction, the chained comment construction, and the embedded topic–comment construction. The first type may be simplex in structure, but the other two are always complex. The double-topic construction consists of a frame-setting topic (in boldface) and a topic (single-underlined) in an aboutness relation with a comment (placed inside brackets), e.g.:

- (26) ***māl****to^H k^be^H=bo^L bo^R,* *13^Lpu^F=ge^L* [n3^L-di^L].
 final time=FRM DSC rooster=TOP down-cast
 ‘Finally, the rooster (you) throw (it) down.’

In the chained comment construction a topic is shared by two or more comments in chained clauses, e.g.:

- (27) *l1^Hbō^L=ge^L* [sjɛ̄l^Hbō^L] [t3^Hts^Lðō^L], [bɪ^H], [l3^Lʃe^R gyɛ̄^H].
 China.fir=TOP tree one:kind radially.thick very tall
 ‘The China fir is a kind of tree, radially thick, and very tall.’

The topic–comment relation may be recursive, giving rise to the embedded topic–comment construction (see Ding 2007b: ch. 7). The outer topic and the embedded topic in the construction must hold the semantic relation of set–member. For instance, in (28) the sentence initial topic represents a whole and the other topics body parts. In (29), the set–member relation is realized as agent-and-activity, with the set denoting activities

done by the agent. This semantic relation is necessary, but not sufficient, for topic–comment embedding. Note that the complex structures of (28) and (29) involve not only topic–comment embedding but also chained comments.

(28)	<i>jō^Ldzi^Hpa^Lpa^L</i>	<i>[zut^L=ge^H</i>	<i>[dʒɔ^L=nɔ̄^L]</i>	<i>[tsɔ^L=ge^H</i>	<i>[dzi^L].</i>
	bat	face=TOP	bad=DUR	organ=TOP	wonderful

‘The bat, (its) face is ugly; (yet its) organs are perfect.’

(29)	<i>ts^L3^H</i>	<i>to=a^L</i>	<i>dʒɔ^Ldʒɔ^H</i>	<i>tʃ^Hi^H</i>	<i>mi^L=bo^L</i>
	earth	on=M	bad	do	person=FRM

[tsi^Hlɔ^H=bo^L *[wu^Ldʒa^H* *k^Hɔ^L-tʃ^Hjɔ̄^L* *jɔ̄^L],*
 seed:sow=FRM harvest:poor out-appear ASR

[gi^Lfaw^H=bo^L *[dʒɔ̄^R* *ma^R=jɔ̄^L],*
 livestock:raise=FRM smooth N=ASR

[ts^Hɔ̄^Ltʃ^Hi^H *la^H* *[ku^Hdzi^H* *xi^R* *ma^R=jɔ̄^L].*
 business:do also profit get N=ASR

‘Persons who behave badly on earth, for cultivation, a poor harvest will certainly come; for livestock-raising, (it) certainly won’t be smooth; for doing business, profit certainly won’t be made.’

NOTES

- 1 The description of Prinmi presented here is based on fieldwork done in Ninglāng County, Yunnan. I am indebted to many of the *Pumi* people for their invaluable help. Except for the suprasegmental notation, Prinmi data are presented in IPA according to the latest revision of 2005. Uncommonly used abbreviations – ASR: assertive; BNY: beneficiary; CLC: cislocative; DSC: discourse clitic; DS.N: desiderative negator; EXT.AN: (animate) existential; FRM: frame-setting marker; M: modificatory; MRT: mirative; NMLC: nominal clause marker; PF.N: perfective negator; PL.IN: inclusive plural; TLC: translocative; and VLT: volitive.
- 2 A rich verb inflection is also found in Western Prinmi spoken in Lánpíng County, see Qinghua in Lu (1983: 42–5) and Dayang in Fu (1998: 104–14). Such verb inflection, however, was lost in Northern Prinmi (see Lu 2001).
- 3 The focus–presupposition construction may be similar to cleft constructions in function, but not in structure; no syntactic division is involved in the Prinmi focus–presupposition construction.

REFERENCES

- Bradley, David (1997) ‘Tibeto-Burman languages and classification’, in David Bradley (ed.) *Papers in Southeast Asian Linguistics: Tibeto-Burman Languages of the Himalayas*, Canberra: Australian National University, 1–72.
- Caughley, Ross (1982) *The Syntax and Morphology of the Verb in Chepang*, Canberra: Australian National University.
- DeLancey, Scott (1997) ‘Mirativity: the grammatical marking of unexpected information’, *Linguistic Typology* 1: 33–52.
- Ding, Picus Sizhi (2001) ‘The pitch-accent system of Niuwozi Prinmi’, *LTBA* 24.2: 57–83.
- Ding, Picus Sizhi (2006) ‘A typological study of tonal systems of Japanese and Prinmi: towards a definition of pitch-accent languages’, *Journal of Universal Language* 7.2: 1–35.

- Ding, Picus Sizhi (2007a) ‘The use of perception tests in studying the tonal system of Prinmi dialects: a speaker-centered approach to descriptive linguistics’, *Language Documentation and Conservation* 1.2: 154–181.
- Ding, Picus Sizhi (2007b) *Studies on Bā Resultative Construction: A Comprehensive Approach to Mandarin Bā Sentences*, Munich: Lincom.
- Ding, Picus Sizhi (2009) ‘Rhoticization as a secondary articulation in stops: evidence from Prinmi’, *Chinese Journal of Phonetics* 2: 74–81.
- Ding, Picus Sizhi (2014) *A Grammar of Prinmi: Based on the Central Dialect of Northwest Yunnan, China*, Leiden: Brill.
- Ding, Picus Sizhi (2016) *Southern Min (Hokkien) as a Migrating Language: A Comparative Study of Language Shift and Maintenance across National Borders*, Singapore: Springer.
- Fu, Ailan (1998) *Pumiyu Dongci de Yufa Fanchou (Grammatical Categories of Pumi Verbs)*, Beijing: Chinese Cultural and Historical Publisher.
- Harrell, Stevan (2001) *Ways of Being Ethnic in Southwest China*, Seattle: University of Washington Press.
- Hyman, Larry (2009) ‘How (not) to do phonological typology: the case of pitch-accent’, *Language Sciences* 31: 213–238.
- LaPolla, Randy J. (1995) ‘Ergative marking in Tibeto-Burman’, in Yoshio Nishi, James Matisoff and Yasuhiko Nagano (eds) *New Horizons in Tibeto-Burman Morphosyntax*, Osaka: National Museum of Ethnology, 189–228.
- Lu, Shaozun (1983) *Pumiyu Jianzhi (An Outline of Pumi)*, Beijing: Nationalities Press.
- Lu, Shaozun (2001) *Pumiyu Fangyan Yanjiu (A Dialectal Study of the Pumi Language)*, Beijing: Nationalities Press.
- Matisoff, James A. (1997) ‘Dayang Pumi phonology and adumbrations of comparative Qiangic’, *Mon-Khmer Studies (Memorial Volume for Paul K. Benedict)* 27: 171–213.
- Mazaudon, Martine (1977) ‘Tibeto-Burman tonogenetics’, *LTBA* 3.2: 1–123.
- Shibatani, Masayoshi and Pardeshi, Prashant (2001) ‘The causative continuum’, in Masayoshi Shibatani (ed.) *The Grammar of Causation and Interpersonal Manipulation*, Amsterdam: John Benjamins, 85–126.
- Sun, Hongkai (1982) ‘On the subgrouping of the Qiangic languages’ (in Chinese), in *Minzu Yuwen Yanjiu Wenji*, Xining: Qinghai Nationalities Press, 189–224.
- Van Valin, Robert and LaPolla, Randy J. (1997) *Syntax: Structure, Meaning, and Function*, Cambridge: Cambridge University Press.

CHAPTER FORTY-ONE

TANGUT

Gong Hwang-cherng

1 INTRODUCTION

Tangut (also known as the Xixia language) is an extinct Tibeto-Burman language that was spoken in the Xixia empire that existed from 1038 to 1227 in northwestern China. The language was buried in oblivion till 1908 when the Russian geographer P.K. Kozlov discovered the ruins of a Tangut city at Khara Khoto. A large number of books and manuscripts written in the Tangut script were brought to St Petersburg. Among them were a Tangut-Chinese glossary, the Fan-Han Ho-shi Zhang-zhong-zhu ('The Tangut-Chinese timely pearl in the palm'), various kinds of Tangut rhyme books, such as the Tong-yin ('The homophones'), the Wen-hai ('The ocean of characters'), the Wen-hai Bao-yun ('The precious rhymes of the ocean of characters'), and the Wu-yin qie-yun ('The rhyme tables of five sound categories'), and Tangut literary works, including Tangut translations of Buddhist sutras and Chinese classics as well as secular texts. There are also original Tangut texts. These materials have made it possible for us to reconstruct Tangut phonology and study its grammar.

1.1 Linguistic affiliations

There is no general agreement on the precise position of Tangut within the Tibeto-Burman languages. It has been claimed to have special affinities with Lolo-Burmese. Recent studies, however, reveal that it is most closely related to the Qiangic branch, with which Tangut shares some important cognates and grammatical features.

2 PHONOLOGY

The reconstruction of Tangut phonology is based on the *fan-qie* (use of known words to represent initial and rhyme of a word) given in Tangut rhyme books on the one hand and Chinese and Tibetan transcriptions of the Tangut characters on the other. Several attempts have been made to reconstruct Tangut sounds. The most important contributions have been made by Nishida (1964/1966, 1986) and Sofronov (1968). The system presented here is an improved version based on their studies.

2.1 Syllable structure

The Tangut syllable is (C)(G)V(G) (G = glide). The G slot can be filled by /j/ or /w/. Tangut has no other syllable-final consonants. The V can be either long or short, tense or lax, retroflexed or non-retroflexed. The tense vowels are always short. There are several nasalized vowels which occur mainly in Chinese loanwords.

2.1.1 Reconstruction of Tangut finals

The reconstruction of Tangut finals is based on Tangut rhyme books and foreign transcriptions. Tangut distinguishes two tones: the level tone and the rising tone. The level tone comprises 97 rhymes, whereas the rising tone comprises 86 rhymes. In the following table as well as throughout this chapter the level tone is indicated with 1 and the rising tone with 2, followed by the order of rhymes in each tone. All together there are 105 overall rhymes, represented by *r* in what follows. Tangut rhymes are divided into 12 rhyme groups, abbreviated to RG in Table 41.1.

2.2 Vowels

Tangut has a seven-vowel system consisting of /i i u e ə o a/. The vowels can be either long or short, tense or lax, retroflexed or unretroflexed. The tense vowels are always short. In the reconstruction the long vowels are represented by geminating the symbols, as /ii ii uu ee əə oo aa/. The tense vowels are represented by putting a dot under the vowels, as /i̥ ɿ̥ ɿ̥ ə̥ o̥ ḁ/ and the retroflex vowels by adding *r* after the vowels, as /ir ir ur er ər or ar/. The long retroflex vowels are written as iir iir uuur eer əor oor aar. In addition to these vowels there are four nasalized vowels /ĩ û õ ã/, which are found mainly in Chinese loanwords.

2.3 Consonants

There are 28 consonants and two semivowels. Stops (except the glottal stop) and affricates contrast in voicing and aspiration. Spirants manifest a contrast between voiced and voiceless. The Tangut consonants are presented in Table 41.2.

2.4 Tones

There are two tones in Tangut: the level tone and the rising tone. The level tone comprises 97 rhymes, whereas the rising tone comprises 86 rhymes.

2.5 Phonological alternations

2.5.1 Alternations of initial consonants

The alternation between voiced and voiceless aspirated consonants is relevant to the change in syntactic category. In the following examples the voiced initial consonants represent intransitive verbs, whereas the voiceless aspirated initial consonants represent transitive and/or causative verbs.

- (1) 繩 *bie²* to release, to open (vi)
縛 *phie²* to release, to open, to untie (vt)
- (2) 削 *bja²* end, cut off (vi)
纏 *phja¹* cut off (vt)
- (3) 燭 *dwar²* to burn, to blaze (vi)
𦗷 *thwər¹* to burn, to cause to blaze (CAUS)
- (4) 洗 *džjwi¹* to clean, be pure, melt (vi)
靚 *tšhwji¹* make clean, to purify, melt (vt)

TABLE 41.1 RECONSTRUCTION OF TANGUT FINALS

		Short vowels			Long vowels		
RG	Lax vowels	Tense vowels	Retroflex vowels	Lax vowels	Long vowels	Retroflex vowels	
I	1 [1.1] u	61 [1.58] ɿ	80 [1.75] ur	5 [1.5]	uu		
	[2.1] [2.51]	[2.51] [2.69]		[2.5]			
	2 [1.2] ju			6 [1.6]	juu		
RG	3 [1.3] ju	62 [1.59] jɿ	81 [1.76] jur	7 [1.7]	juu		
	[2.3] [2.52]	[2.52] [2.70]		[2.6]			
	4 [1.4] u						
II	8 [1.8] e	68 [1.65] ɛ	82 [1.77] er	12 [1.12]	ee	99 [2.84] eer	
	[2.7] [2.58]	[2.58] [2.71]		[2.11]			
	9 [1.9] ie	69 [1.66] iɛ	83 [1.78] ier	13 [1.13]	iee		
RG	10 [1.10] ji	70 [1.67] ji	84 [1.79] jir	14 [1.14]	jii	101 [1.93] jiir	
	[2.9] [2.60]	[2.60] [2.72]		[2.12]		[2.86]	
	11 [1.11] ji						
III	15 [1.15] ē						
	[2.13]						
	16 [1.16] jĩ						
IV	17 [1.17] a	66 [1.63] ɑ	85 [1.80] ar	22 [1.22]	aa	88 [1.83] aar	
	[2.14] [2.56]	[2.56] [2.73]		[2.19]			
	18 [1.18] ia		86 [1.81] iar	23 [2.20]	iaa		
RG	19 [1.19] ja	67 [1.64] ja	87 [1.82] jar	24 [1.23]	jaa	89 [2.75] jaar	
	[2.16] [2.57]	[2.57] [2.74]		[2.21]			
	20 [1.20] ja			21 [1.21]	jaa	[2.18]	
V	25 [1.24] ā						
	[2.22]						
	26 [1.25] iā						
RG	27 [1.26] jā						
	[2.23]						
	[2.24]						
VI	28 [1.27] ə	71 [1.68] ɿ	90 [1.84] ər	32 [1.31]	əə		
	[2.25] [2.76]	[2.53] [2.76]					
	29 [1.28] iə		91 [1.85] iər				
RG	30 [1.29] ji	72 [1.69] ji	92 [1.86] jir	33 [1.32]	jii	100 [1.92] jiir	
	[2.27] [2.61]	[2.61] [2.77]		[2.29]		[2.85]	
	31 [1.30] ji						
VII	34 [1.33] ej	63 [1.60] ej	77 [1.73] ejr	38 [1.37]	eej		
	[2.30] [2.53]	[2.53] [2.66]		[2.34]			
	35 [1.34] iej	63 [1.60] iɛj	78 [2.67] iejr	39 [1.38]	ieej		
RG	36 [1.35] ijj	64 [1.61] ijj	79 [1.74] ijir	40 [1.39]	ijjj		
	[2.32] [2.54]	[2.54] [2.68]		[2.35]			
	37 [1.36] ijj						
	[2.33]						

(Continued)

TABLE 41.1 (CONTINUED)

		Short vowels			Long vowels		
RG	Lax vowels	Tense vowels	Retroflex vowels	Lax vowels	Retroflex vowels		
VIII	41 [1.40] eɪj						
	42 [1.41] iɪj [2.36]	76 [2.65] iɪj					
RG	43 [1.42] jɪj [2.37]	65 [1.62] jɪj [2.55]					
	44 [1.43] ew [2.38]		93 [1.87] ewr [2.78]	48 [2.41] eew			
IX	45 [1.44] iew [2.39]						
	46 [1.45] jiw [2.40]		94 [1.88] jiwr [2.79]	49 [1.47] jiiw			
RG	47 [1.46] jiw						
	51 [1.49] o [2.42]	73 [1.70] ɔ [2.62]	95 [1.89] or [2.80]	54 [1.52] oo [2.45]	102 [1.94] oor		
X	52 [1.50] io [2.43]	74 [1.71] iɔ [2.63]	96 [1.90] ior [2.81]	55 [1.53] ioo [2.46]			
	53 [1.51] jo [2.44]	75 [1.72] jɔ [2.64]	96 [1.90] jor [2.81]	55 [1.53] joo [2.46]	103 [1.95] joor		
RG	50 [1.48] jwo						
	56 [1.54] ow [2.47]		97 [1.91] owr [2.82]				
XI	57 [1.55] iow [2.48]			59 [1.57] ioow			
	58 [1.56] jow [2.48]		98 [2.83] jowr	60 [2.50] joow			
XII	104 [1.96] ū						
	105 [1.97] jwa						
rg							

TABLE 41.2 INVENTORY OF CONSONANTS

	Bilabial	Dental	Palato-alveolar	Velar	Glottal
Voiceless stop	p	t		k	? (·)
Aspirated stop	ph	th		kh	
Voiced stop	b	d		g	
Voiceless affricate		ts	tʃ (tś)		
Aspirated affricate		tsh	tʃh (tsh)		
Voiced affricate		dz	dʒ (dź)		
Nasal	m	n		ŋ	
Voiceless spirant		s	f (ś)		h (x)
Voiced spirant		z	ʒ (ż)		ħ (y)
Tap		r			
Lateral		l			
Lateral spirant			l(lh)		
Glide	w		j		

Note: Symbols given in the parentheses are those used in the reconstruction.

2.5.2 Alternation of medials

The alternation of medials includes the alternation between the presence or absence of medial *-j-* as well as *-w-*.

(5) 緣 <i>dzuu²</i>	to plant, to erect	緘 <i>dzjuu²</i>	idem
(6) 繢 <i>thu¹</i>	to draw (a bow)	緥 <i>thju¹</i>	idem
(7) 爰 <i>dzu¹</i>	to love, to like	緝 <i>dzuu¹</i>	idem
(8) 繸 <i>lwu¹</i>	to mix	緒 <i>ljuu¹</i>	idem
(9) 繸 <i>yie¹</i>	strength, influence, power	緝 <i>yie¹</i>	idem
(10) 繸 <i>sji¹</i>	poor, exhausted; to come to an end	緝 <i>sjuu¹</i>	strength, force, power
(11) 繸 <i>šji¹</i>	go, reach, enter	緝 <i>šjuu¹</i>	idem

2.5.3 Alternation of tones

As mentioned earlier, there are two tones in Tangut: a level tone and a rising tone. The alternation of tones can be classified into three groups in connection with the change in the syntactic categories.

a. Level tone representing nouns and rising tone representing verbs

(12) 繸 <i>ljii¹</i>	trousers (n)	繸 <i>ljii²</i>	to put on trousers (v)
(13) 繸 <i>t'siej¹</i>	cap, hat (n)	繸 <i>t'siej²</i>	wear cap (v)
(14) 繸 <i>zji¹</i>	leather shoes, boots (n)	繸 <i>zji²</i>	to put on shoes (v)
(15) 繸 <i>wa¹</i>	the shoulder (n)	繸 <i>wa²</i>	both characters mean ‘to carry on the shoulder’ (v)

b. Level tone representing verbs and rising tone representing nouns

(16) 繸 <i>tji¹</i>	to put, to place (v)	繸 <i>tji²</i>	a place (n)
(17) 繸 <i>yewr¹</i>	to count (v)	繸 <i>yewr²</i>	a number (n)
(18) 纓 <i>džjwo¹</i>	to pierce, to bore a hole (v)	縫 <i>džjwo²</i>	hole (n)
(19) 繸 <i>sju¹</i>	to hoard, to conceal (v)	縫 <i>sju²</i>	a cupboard (n)

c. Tone alternation not related to syntactic categories

(20) 繸 <i>sjwi¹</i>	who	縫 <i>sjwi²</i>	idem
(21) 繸 <i>wiji¹</i>	there is/are	縫 <i>wiji²</i>	idem
(22) 繸 <i>mjjij¹</i>	tail, end, the lower part	縫 <i>mjjij²</i>	idem
(23) 纇 <i>do¹</i>	poison	縫 <i>do²</i>	idem
(24) 纇 <i>nər¹</i>	yellow	縫 <i>nər²</i>	idem
(25) 纇 <i>ywər¹</i>	blue	縫 <i>ywər²</i>	idem

In addition to the above alternations of tone, there are also extensive vocalic alternations which are related with causative formation (see section 4.1), compounding (4.2.2), and verb inflection (4.2.3).

3 MORPHOLOGY

3.1 Substantives

3.1.1 Nouns

Tangut lacks inflectional morphology. Grammatical relations are expressed by particles. Genitive, dative, and accusative are expressed by the same particle 篤 *jij¹*, instrumental

by 故 *ywu²*, locative by 屏 *u²* ‘in (the garden)’, 漢 *kha¹* ‘in (the book, the water)’, 鮑 *gu²* ‘in (the water)’, 蔡 *ya²* ‘in (the sky, the heart)’, 飛 *tshjaa¹* ‘on’, 蔽 *khju¹* ‘under’, allative by 鮑 *do²* ‘to’, comitative by 蔽 *rjir²* ‘with’, and comparative by 蔽 *su¹*. In addition to this, there is an agentive marker, 蔽 修 *džji² wji¹*, that is sometimes used to mark the actor of transitive verbs, and a topic marker, *tja¹*.

3.1.2 Pronouns

There is a rich system of pronouns. Plurality can be expressed by the suffix 蔽 *nji²*. However, there are also special words expressing first person plural, that have a contrast between inclusive vs exclusive forms (Kepping 1994: 339). The pronominal system distinguishes between familiar and honorific forms. There is one reflexive pronoun 蔽 *jij¹*, three demonstrative pronouns, and several interrogative pronouns.

3.1.2.1 Personal pronouns

<i>person</i>	<i>singular</i>	<i>familiar</i>	<i>honorific</i>	<i>plural</i>				
1st	她 <i>ya²</i>	她 <i>mjo²</i>	她 <i>ya²</i>	她 <i>nji²</i>	她 <i>gja²</i>	她 <i>mji²</i>	we	
2nd	她 <i>nja²</i>	她 <i>nji²</i>	她 <i>nji²</i>	她 <i>nji² nji²</i>	她 <i>gji²</i>	她 <i>mji²</i>	we	[inc]

								[exc]

3.1.2.2 Reflexive pronouns

蔽 *jij¹* oneself, himself, herself

3.1.2.3 Demonstrative pronouns

她 <i>thji²</i>	this	她 <i>thja²</i>	(third person, irrespective of animate, inanimate, male, or female)
她 <i>thju²</i>	here	她 <i>thja¹</i>	he, his, him, she, her, it, that
她 <i>tshju²</i>	this, that	她 <i>thja²</i>	there
她 <i>tshju¹</i>	that	她 <i>tjij¹</i>	there
		她 <i>tshji¹</i>	that
		她 <i>tshjiw¹</i>	that

3.1.3 Adjectives

Adjectives occur after the nouns they modify. Many adjectives appear in reduplicated forms.

Examples:

她 *ηq²* 她 *ηq²* 故 *ljuu²* 故 *ljuu²* 故 *lji¹* 懈 *mji¹* 懈 *džij¹* 懈 *lju¹* 懈 *məə¹* 懈 *rjijr¹* 懈 *lji¹*
 good good pretty pretty though not only flute blow skilled AFF
 ‘(She) is not only very pretty, but also good at playing a flute.’

Comparison is expressed by the postposition 蔽 *su¹* ‘than’ followed by an adjective.

𠂇 **𢃥** **龍** **𠂇** **𠂇** **𠂇** **𠂇** **𠂇** **𠂇** **𠂇**
dzjwo² *we²* *ljw4¹* *džjwow¹* *su¹* *bjuu¹* *tja¹* *tij²* *džij²* *nioow¹* *lji¹*
 man be beast bird than noble TOP propriety have therefore AFF
 ‘The reason why humans are nobler than beasts and birds is that humans know
 decencies.’

3.2 Verbs

3.2.1 Direction marking

3.2.1.1 Directional prefixes

There is a special set of prefixes for verbs: seven directional prefixes (which have developed into prefixes of perfective aspect) and six prefixes of optative aspect (Kepping 1982: 77–8). The latter are derived from the former by changing the finals into *-ij*.

Directional markers

	Perfective markers	Optative markers
(1)	𢃥 <i>ja²</i>	𠂇 <i>·jij¹</i> upward
(2)	𢃥 <i>nja¹</i>	龍 <i>njjij²</i> downward
(3)	𠂇 <i>kji¹</i>	𠂇 <i>kjjij¹</i> here, inside
(4)	𠂇 <i>wji²</i>	𠂇 <i>wjjij²</i> there, outside
(5)	𢃥 <i>dji²</i>	𢃥 <i>djjij²</i> towards the speaker
(6)	𢃥 <i>dja²</i>	𢃥 <i>djjij²</i> away from the speaker
(7)	𢃥 <i>rjir²</i>	𢃥 <i>rjijr²</i> (direction not found)

Examples:

𠂇 **𠂇** **散** **𠂇** **𠂇**
mə¹ *dzju²* *lji²* *nja¹* *lji²*
 sky rain big DIR come
 ‘It has rained hard.’

𢃥 **𠂇** **𢃥** **𠂇** **𠂇**
ŋa² *·jij¹* *dji²* *khjjij¹* *nja²*
 me DAT DIR give -2sg
 ‘Give (them) to me!’

𢃥 **𠂇** **備** **縫** **𢃥** **𢃥**
sjjij¹ *·jt²* *zji¹* *mjaat¹* *dji²* *we²*
 now again husband wife OPT become
 ‘Let’s be husband and wife again.’

3.2.2 Agreement

The Tangut verb shows hierarchical agreement only with first and second person. Third person is left unmarked (Kepping 1994: 339).

The verb-agreement system also marks actor agreement in the case of first and second person actors (Kepping 1976). The actor agreement causes phonological alternations of

the verbs, which in certain classes of verbs change the vowels *-i-* and *-u-* of the verbs into *-o-* (Nishida 1986):

	<i>Basic form</i>	<i>Derived form</i>	
(26)	藏 <i>džijj²</i>	藏 <i>džjo²</i>	have
(27)	醜 <i>tji¹</i>	醜 <i>tjo¹</i>	let drink
(28)	醜 <i>phji¹</i>	醜 <i>phjo²</i>	make, let
(29)	替 <i>wji¹</i>	替 <i>wjo²</i>	to do, to make
(30)	藏 <i>rjir¹</i>	藏 <i>rjor¹</i>	to get
	藏 <i>rjir²</i>	藏 <i>rjor²</i>	to get
(31)	葬 <i>pju²</i>	葬 <i>pjo¹</i>	to burn, to bake
(32)	龜 <i>lhjut²</i>	龜 <i>lhjo²</i>	to get, to acquire, to obtain

When the actor is in the third person, the basic forms of verbs such as 藏 *džijj²* ‘have’, 醜 *tji¹* ‘let drink’, and 藏 *rjir¹* ‘get’ are used. For example:

藏 葬 麻 聽 藏
jow¹ *swē¹* *mijj¹* *gji²* *džijj²*
Wang Sun daughter a have

‘Wang Sun has a daughter.’

藏 替 駕 醜 罷 聽 藏 醜
buu² *kow¹* *tšhji²* *rjar²* *·o²* *khjow¹* *dja²* *tji¹*
Mu Gong immediately wine bestow PERF let drink
‘Mu Gong immediately entertained (them) with wine.’

藏 席 漢 麻 藏
džijj¹ *mjjjr²* *dzjw¹* *rjir¹* *sju²*
ferry people boat get like
‘As if when the people who want to cross a river get a boat.’【如渡得船】

When the actor is the first or second person, the derived forms of verbs such as 藏 *džjo²*, 替 *tjo¹*, and 藏 *rjor²* are used.

藏 造 藏 鏊
ŋa² *lji¹* *džjo²* *ŋa²*
1sg treasure have -1sg
‘I have treasure.’

藏 人 藏 替 席 藏 仁 藏 驅
khji² *džwo²* *dja²* *tjo¹* *ŋa²* *njj¹* *·o¹* *dja²* *lhjij²*
ten thousand person PERF let drink -1sg near owner PERF lack
‘I have fed 10,000 people but have neglected my neighbours.’

藏 𠵼 藏 麻 藏 漢 仁 藏 鏊
sq¹ *jir²* *kji¹* *lji²* *·o²* *dijj²* *rjor²* *ŋa²*
three hundred bottle jar wine OPT get -1sg
‘Let me get 300 bottles of wine.’

Only when the personal suffixes of the verbs refer to the actor of the clause are the derived forms of the verbs used. In the following examples 藏 *ŋa²* refers to the actor of the clause, therefore the derived forms of the verbs such as 駕 *phjo²* and 替 *wjo¹* are used.

緋 緒 簡 繢 祕 種 縱 綴 繢 繢 繢
ljo² *tsəj¹* *jjij¹* *dʒju²* *phji¹* *wo²* *kha¹* *gji²* *jjij¹* *kji¹* *dʒju²* *phjo²* *nja²*
 brother little ACC rule CAUS should while son ACC PERF rule CAUS -2sg
 ‘You have made your son, instead of your younger brother as you should have, a ruler.’

櫛 賽 羅 斧 暈 緯 繢 繢 繢 繢
nji² *śja¹* *bji²* *·iɔ¹* *·u²* *dʒjo²* *·ja²* *dzwə²* *rjir²* *wjo¹* *nja²*
 2sg[HON] seven step scope in poem one CL PERF make -2sg
 ‘Write a poem in the time it takes you to make seven steps.’

In the following examples 錢 *ŋa²* and 貝 *nja²* do not refer to the actors of the clauses, accordingly the basic forms of the verbs such as 祕 *phji¹* and 種 *wji¹* are used.

蔬 瓜 緯 盡 祕 種
ko¹ *tshjaa¹* *wji²* *dzuu²* *phji¹* *ŋa²*
 coach on PERF sit CAUS -1sg
 ‘Let me sit on the coach.’

錢 暈 瓜 緯 經 繢
kji¹ *dji²* *lji²* *ljii¹* *ljii¹* *wji¹* *nja²*
 certainly grace heavy return make -2sg
 ‘I will certainly give you a large reward.’

4 WORD FORMATION

4.1 Derivational morphology

The causative forms of Tangut verbs are formed by changing the lax vowels into the corresponding tense vowels. In the following examples the causative forms are derived from transitive and/or intransitive verbs.

- | | |
|--|---|
| (33) 頭 <i>nju²</i> to suck a breast (vt) | 纏 <i>nju²</i> to suckle (CAUS) |
| (34) 鮐 <i>thji¹</i> to drink (vt) | 餌 <i>tji¹</i> to give to drink, to feed (CAUS) |
| (35) 疊 <i>gjwi²</i> to wear clothes, to put on clothes (vt) | 縫 <i>gjwi¹</i> to make to wear clothes, to clothe (CAUS) |
| (36) 戰 <i>ywej¹</i> to fight (vi) | 織 <i>ywej¹</i> to set to fight, to cause war (CAUS) |

Transitive verbs are also derived from intransitive verbs by changing the lax vowels into the corresponding tense vowels.

- | | |
|---|--------------------------------------|
| (37) 翻 <i>lwu¹</i> to get mixed (vi) | 縛 <i>lwu¹</i> to mix (vt) |
| (38) 烹 繩 <i>le²</i> to boil (vi) | 縛 <i>le¹</i> to boil (vt) |

Nouns and adjectives are changed into verbs or causative verbs by the same mechanism.

- | | |
|--|--|
| (39) 繁 <i>nju²</i> tame, mild (adj) | 繆 <i>nju²</i> to tame (vt) |
| (40) 繼 <i>bji¹</i> low, below, down (adj) | 縛 <i>bji¹</i> to lower, to bend |
| 緒 <i>bji²</i> low, below, down (adj) | 縛 <i>bji¹</i> to heighten, to elevate, to promote (v) |
| (41) 尾 <i>bjjij²</i> high (n, adj) | |

- | | | | |
|---------------------------------------|-----------------------------|----------------------------------|---------------------------------------|
| (42) 𢃠 <i>khwa¹</i> | far, distant (adj, adv) | 𢃠 <i>khwa¹</i> | keep away from, keep at distance (vt) |
| (43) 𢃡 <i>lə¹</i> | a pit, a gully, a ditch (n) | 𢃡 <i>lə¹</i> | to throw into a pit, to bury |
| (44) 𢃢 <i>mji²</i> | traces, footprints (n) | 𢃢 <i>mji²</i> | to trace, to trace source (v) |

Nouns are derived from verbs, as in:

- | | | | |
|---------------------------------------|------------------------------|----------------------------------|---------------------------|
| (45) 𢃣 <i>thji¹</i> | to drink (vt) | 𢃣 <i>tji¹</i> | food (n) |
| (46) 𢃤 <i>džju¹</i> | to deceive, to cheat (vi,vt) | 𢃤 <i>džju²</i> | a bait, an allurement (n) |
| (47) 𢃥 <i>sji²</i> | to know (v), knowledge (n) | 𢃥 <i>sji²</i> | wisdom (n) |

4.2 Compounding

4.2.1 Reduplicative compound

Reduplication is an important device for Tangut word-formation.

效 <i>ljuu²</i>	效 <i>ljuu²</i>	(beautiful beautiful) beautiful
發 <i>lji¹</i>	發 <i>lji¹</i>	(one one) one by one, in every detail
福 <i>yowr²</i>	福 <i>yowr²</i>	(all all) all, every
處 <i>rjur¹</i>	處 <i>rjur¹</i>	(every every) here and there, everywhere

4.2.2 The first variant of reduplicative compounds

The reduplicative compounds are formed with a phonological change in the preceding syllable, where the front vowels are changed into the corresponding central vowels.

In the following examples the succeeding syllables represent the original morphemes, whereas the first syllables are derived from the second by dropping the syllable coda *-j*, if there is any, and changing the front vowel to the corresponding central vowel.

打仗	<i>ywə¹</i>	打仗 <i>ywej¹</i>	(fight fight) to fight
清潔	<i>sə²</i>	清潔 <i>sej¹</i>	(clean clean) clean
旋轉	<i>džiə²</i>	旋轉 <i>džiej²</i>	(revolve revolve) to revolve; transmigration
報應	<i>šiə¹</i>	報應 <i>šiej¹</i>	(lead lead) to lead
歸宿	<i>bji¹</i>	歸宿 <i>bji²</i>	(low low) the lower part
醜陋	<i>mji²</i>	醜陋 <i>mji²</i>	(silent silent) silent(ly)
既往	<i>bji¹</i>	既往 <i>bjj²</i>	(high high) the upper part
該當	<i>sji²</i>	該當 <i>sji²</i>	(know know) acquaintance
逃離	<i>lji¹</i>	逃離 <i>lji²</i>	(midday midday) midday
難堪	<i>šji¹</i>	難堪 <i>šji¹</i>	(doubt doubt) to doubt; irresolute

4.2.3 The second variation of reduplicative compounds

Either the preceding or the following syllable changes the front vowels into the corresponding back round vowels.

愚蠢	<i>giee¹</i>	愚蠢 <i>gioo¹</i>	silly, foolish
請求	<i>šiee¹</i>	請求 <i>šioo¹</i>	to collect
交換	<i>wji¹</i>	交換 <i>wjoo¹</i>	to exchange, to trade
肢解	<i>džii¹</i>	肢解 <i>džoo¹</i>	to divide, to share
糞糲	<i>gjiit¹</i>	糞糲 <i>gjoo¹</i>	to bite, to peck

毳	毳	<i>no²-nej²</i>	tranquillity, stability, safety, peace
纖	纖	<i>ror²-rejr²</i>	to surround
穢	穢	<i>kio²-kie²</i>	hate, dislike
穢	穢	<i>sjwo²-sjwij¹</i>	to grind, to whet
祚	祚	<i>djo-djij</i>	to grieve, grievance

5 SYNTAX

5.1 Structure of the noun phrase

Modifying nouns and pronouns occur before, and attributive adjectives and numeral phrases after, the nominal heads.

a. Noun + adjective

𠙴	𦵃	<i>tji¹ŋq²</i>	(food good)	good food
𩶻	𩶻	<i>lji¹ lji²</i>	(wind big)	strong wind
遜	遜	<i>dji² rjijr¹</i>	(doctor good)	good (medical) doctor
鷺	鷺	<i>we¹ nər²</i>	(bird yellow)	yellow bird

b. Noun + noun

蓆	蓆	<i>məə¹ joow²</i>	(fire stove)	brazier
𦵃	𦵃	<i>tsur¹ t'siej¹</i>	(winter hat)	winter hat
𦵃	𦵃	<i>sjt² dzjwo²</i>	(female person)	woman
𧈧	𧈧	<i>džit¹ gjwi¹</i>	(leather clothing)	leather jacket

c. Noun + adjective/noun + numeral

𩶻	𩶻	𩶻	<i>rjijr¹ ya² gjit²</i>	(horse good one)	a good horse
𩶻	𩶻	𩶻	<i>sji² dzjwo² gjit²</i>	(female person one)	a woman

d. Noun + adjective + noun

𩶻	𩶻	𩶻	<i>rjijr¹ ya² tshji¹</i>	(horse good meat)	meat of a good horse
---	---	---	---	-------------------	----------------------

e. Noun + adjective + adverb

遜	遜	遜	<i>dji² rjijr¹ yiej¹</i>	(doctor good really)	really good doctor
---	---	---	---	----------------------	--------------------

f. Verb + auxiliary (Here the auxiliary verb 繼 *lew²* functions as a nominalizer.)

驛	驛	<i>thwər¹ lew²</i>	(burn should)	fuel
𦵃	𦵃	<i>gjwi² lew²</i>	(wear should)	clothing
𧈧	𧈧	<i>zji² lew²</i>	(wear-shoes should)	shoes
𦵃	𦵃	<i>phjii¹ lew²</i>	(send should)	servant, slave

g. Noun + numeral + classifier

Classifiers occur after numerals.

綱	綱	綱	<i>dzjo² ·ja² dzwə²</i>	(poem one CL)	a poem
𦵃	𦵃	𦵃	<i>sow¹ ·ja² phu²</i>	(mulberry-tree one CL)	a mulberry-tree
𦵃	𦵃	𦵃	<i>da² ·ja² gjwi²</i>	(word one CL)	a word
𦵃	𦵃	𦵃	<i>kja² ·ja² kho¹</i>	(song one CL)	a song

h. Pronoun + noun

龜	龜	$\eta a^2 wə'$	(my husband) my husband
龜	龜	$mjo^2 \eta wuu^l$	(my word) my words
龜	龜	$thji^2 dəq^2$	(this word) these words
龜	龜	$nja^2 wjij^2$	(2sg uncle) your (paternal) uncle

i. Noun/pronoun + genitive marker + noun

龜	龜	龜	龜	$thej^l\text{-}pie^l\text{-}jij^l\text{-}gji^2\text{-}lhji^l$	(proper noun poss sons grandsons)
龜	龜	龜	龜	$\eta a^2\text{-}jij^l\text{-}la^l$	‘the descendants of Thej-pie’, (1sg poss hand) ‘my hand’

5.2 Structure of the clause

5.2.1 Word order

The basic pattern of the Tangut clause is verb final. The main verb can be followed by an auxiliary verb.

龜 務 駕 犬 駕 祀 脣 殺 犁 懈 驗 慢 駕
 $xu^l\text{-}jow^l\text{-}jir^2\text{-}da^2\text{-}phja^l\text{-}dzijj^2\text{-}rjijr^l\text{-}mji^l\text{-}džijj^l\text{-}nioow^l\text{-}mjjij^l\text{-}seew^2\text{-}rjijr^l$
Fu Rong learn thing cut judge good NEG only also dream foretell good
‘Fu Rong is not only good at deciding criminal cases, but also good at interpreting dreams.’

5.2.2 Serial verb constructions

In Tangut two or more verb phrases are juxtaposed to refer to events which are related as parts of one overall event.

翁	散	龜	龜	猶	駕	數	脩	鞍
$dzjwi^l$	$ljij^2$	$tshja^l$	$žji^l$	$tsier^l$	$dzjwo^2$	$zeew^2$	pju^2	$rjir^2$
emperor	greatly	angry	left	right	person	send	palace	hall

駕	龜	禡	禡	屏	龜	禡	駕
$khju^l$	$tshjij^l$	$phji^l$	$rjir^2$	$·u^2$	$njjij^2$	yie^l	$·jt^2$
under	bring	CAUS	caldron	in	OPT	boil	COMP

‘The emperor got very angry. He ordered the people surrounding him to bring (Wen Zhi) down to the palace court and boil (him) in the cauldron.’

5.3 Major sentence types

5.3.1 Negated sentences

There are six negative forms in Tangut: 懈 mji^l , 犹 mji^l , 禿 $mjjij^l$, 犁 $mjjij^2$, 犭 tji^l (expressing prohibition), and 惟 nja^2 .

The negative markers generally precede the verb they negate. The exceptions are 犻 $mjjij^l$ and 惟 nja^2 , which occur in the sentence-final position.

The most general form of negation is 諦 *mji*^l, as in:

形 緣 諦 誓
*dq*² *mjor*^l *mji*^l *tshijij*^l

word true NEG say

‘(S/He) does not tell the truth.’

般 蔑 諦 誓 蔑 諦 誓
*dzjij*² *rjur*^l *zji*^l *lji*² *zji*² *džijij*^l *mji*^l *dwewr*²

other all children all coldness not feel

‘All the other children did not feel cold.’

疑 *mji*^l is used to negate auxiliary verbs, such as 詢 *kjir*² ‘dare’, 諦 *njwi*² ‘can’, 諦 *lji*² ‘possible’, and 諦 *džijij* ‘willing’, etc. For example:

謐 蔑 諦 誓 蔑
*tjij*² *phji*^l *mji*^l *kjir*² *ŋa*²

decency discard NEG dare -1sg

‘I did not dare to offend decency.’

蔽 蔑 蔑 蔑 蔑 蔑 蔑 誓
*thjij*² *sjo*² *thji*² *dzjwo*² *tjij*^l *jij*^l *rjur*^l *·ar*² *mji*^l *njwi*² *·jt*²

why this person alone ACC restrain NEG can COMP

‘Why can’t you restrain this person alone?’

磼 諦 諦
*dza*² *mji*^l *lji*^l

estimate not possible

‘It is not possible to estimate it.’

薦 諦 諦
šji^l *mji*^l *džijij*

go NEG willing

‘S/He is not willing to go.’

Finely differentiated existential verbs such as 蔑 *dju*^l, 諦 *tshju*^l, 倒 *wijij*², and 諦 *džijij*² are all negated by 諦 *mji*^l, which normally occurs after nouns.

薦 *dju*^l ‘there is’ and 諦 *mji*^l ‘not exist’ are put in contrast as in:

眞 諦 惇 織 蔑 蔑 蔑 蔑 薦 誓
*thjt*² *we*² *·u*² *tja*^l *yu*^l *khjwi*^l *mji*^l *le*² *bjuu*² *dju*^l *mji*^l *dijij*²

this city in TOP head cut NEG fear general there is but

穀 織 織 諦
yio^l *bjuu*² *tja*^l *mji*^l

surrender general TOP not exist

‘In this city there are generals who are not afraid of being beheaded, but no generals who will surrender.’

藏 **藏** ‘have’ and 藏 **藏** ‘not have’ are put in contrast in:

藏 藏 藏 藏
sij² *tshju¹* *sij²* *mij¹*
wisdom have wisdom not have

‘No matter whether (he) has wisdom or does not have wisdom.’

The contrast between 有 **有** ‘there is’ and 藏 **藏** ‘not exist’ is observed in the following pair of sentences.

有 有 有 有
khjow¹ *tji²* *ljø²* *wjij²*
give place where there is
‘How is it possible to give (her) away.’

有 有 有 有 有
lo¹ *we²* *phji¹* *tji²* *mij¹*
rich become make place not exist
‘It is not possible to make (him) rich.’

藏 **藏** ‘have’ and 藏 **藏** ‘not have’ are put in contrast in the following sentences.

藏 藏 藏 藏 藏
zji¹ *mij¹* *so¹* *mij¹* *džij²*
son not have three daughter have
(He) has three daughters but no sons.’

藏 藏 藏 藏 藏 藏 藏
njwo² *lew¹* *lji¹* *mij¹* *sij¹* *gwə¹* *lhwu¹* *džij²*
ancient one trousers not have now five clothes have
‘In the old days (I) had not a single pair of trousers, but now (I) have five suits of clothes.’

However, it also happens that the existential verb 藏 **藏** *džij²* is negated by 藏 **藏** *mji¹*.

藏 藏 藏 藏 藏 藏 藏
yjwā¹ *xia¹* *lu²* *nioow¹* *war²* *gjwi²* *lew²* *mji¹* *džij²*
Ruan Xian poor therefore property clothes not have
‘Ruan Xian is so poor that he has no property or clothes.’

藏 **藏** can also be used as a negative answer to a yes–no question.

藏 藏 藏 藏 藏 藏
thjij² *sjo²* *mji¹* *lj¹* *rjur²* *tshjii²* *nja²*
why not exist AFF PERF say -2sg
‘Why did you say “no”?’

藏 **藏** *mji¹* is used in a negative sentence in reference to future.

藏 藏 藏 藏 藏
thja¹ *džjwo²* *rjur¹* *kha¹* *mij²* *wee¹*
that person world in NEG born
‘That person has not been born yet in the world.’

The equational verb 故 *ywu*² is negated by 惣 *nja*². These two words are put in contrast in the following sentence.

故	龍	蛇	猶	情	諺	終	竚	故	貳	故
<i>sji</i> ¹	<i>nja</i> ²	<i>na</i> ²	<i>džia</i> ¹	<i>nja</i> ²	<i>ku</i> ¹	<i>síwə</i> ¹	<i>wə</i> ¹	<i>ywu</i> ²	<i>nja</i> ²	<i>lji</i> ¹
now	2sg		fox	NEG	then		rat	be	-2sg	AFF
'You are either a fox or a rat.'										

情 *nja*² can occur alone without a verb as a negative answer to a question. In that case it is equivalent to English ‘no!’

應	行	能	情	猶	諺	緘	能	貳	勢
<i>xjwi</i> ¹	<i>tshjow</i> ¹	<i>dq</i> ²	<i>nja</i> ²	<i>ηq</i> ²	<i>ηq</i> ²	<i>wji</i> ²	<i>sjij</i> ²	<i>nja</i> ²	<i>jt</i> ²
Fei	Zhong	say	no	good	good	PERF	think	-2sg	COMP
'Fei Zhong said, “No! Please think about it”.'									

懼	龍	緘	情
<i>jiw</i> ²	<i>šjwo</i> ¹	<i>lew</i> ²	<i>nja</i> ²
doubt	raise	should	NEG

'Don't have a doubt!'

莫 *tji*¹ ‘don’t’ is the negative marker used in imperatives.

莫	莫	貳
<i>tji</i> ¹	<i>khjij</i> ¹	<i>nja</i> ²
NEG	give	-2sg

'Don't give (it to them)!'?

莫	緘	能	貳
<i>tji</i> ¹	<i>sjwii</i> ¹	<i>sjij</i> ²	<i>nja</i> ²
NEG	worry	-2sg	

'Don't worry!'

懼	麗	莫	麗
<i>jiw</i> ²	<i>ljij</i> ¹	<i>tji</i> ¹	<i>šjwo</i> ¹
doubt	NEG	create	

'Don't suspect!'

5.3.2 Questions

There are two grammatical devices that explicitly mark a sentence as a question. The first device is the question particle, in which the presence of the question particle 故 *mo*² in sentence-final position signals the construction is a question, as in:

穢	耗	故	緘	所	故
<i>nji</i> ²	<i>kă</i> ¹	<i>tšja</i> ¹	<i>wji</i> ²	<i>dzjo</i> ¹	<i>mo</i> ²
2sg[HON]	sugar	cane	PERF	eat	Q

'Do you eat sugar cane?'

The second device is the question word, in which the presence of a question word signals the construction to be a question. The following is a list of the most commonly used question words in Tangut.

覩	<i>sjwi¹</i>	who
覩	<i>sjwi²</i>	who
覩	<i>wa²</i>	what
覩	<i>ljo²</i>	where
覩	<i>lji¹</i>	which
覩	<i>thjij² sjo²</i>	why, how
覩	<i>wa² zjij¹</i>	how many, how much
覩	<i>zjij¹ m²</i>	how many kinds

In general, question words occur in the immediately preverbal focus position or in predicate position. Examples:

覩 蘭 覩 覩 覩
sj¹i *w²a* *gji²* *tja¹* *sjwi¹*
 former husband TOP who
 ‘Who is your former husband?’

覩 蘭 覩 犀 覩
thj²i *lhwu¹* *sjwi¹* *:jij¹* *ηwu²*
 this clothes who POSS be
 ‘Whose clothes are these?’

覩 紗 犀 貝
thjij² *sjo²* *nw²a* *nja²*
 how know -2sg
 ‘How do you know it?’

僕 賈 紗 紗 覩
ηw²a¹ *bji¹* *nja¹* *nji²* *ljo²* *dju¹*
 five brightness divine pearl where be
 ‘Where is the divine pearl of five planets?’

微 蘭 數 楠 犀 薦 覩
sju² *po¹* *tji¹* *nji²* *phji¹* *ya²* *thjij²* *sjo²*
 Su bhū ti 2sg[HON] mind in how
 ‘How does it appear to thee, Subhūti?’

姪 犀 紗 覩 紗 紗 覩
sji² *:jij¹* *ljir¹* *tshja²* *dju¹* *nji²* *zjij¹* *m²* *džjo²* *nja²*
 woman GEN four virtue there.are 2sg[HON] how.many kind have -2sg
 ‘There are four virtues of women. How many do you have of them?’

5.3.3 Subordinate clauses

A modifying clause is placed before the noun it modifies, without any marker, nominalizer or relative pronoun.

緘 軋 懈 憚 罷
yu¹ khjwi¹ mji¹ le² bjuu²
head cut NEG fear general

‘The generals who are not afraid of being beheaded.’

5.3.3.1 Adverbial clause

The conditional meaning is expressed by means of the conditional conjunction 疱 *tjj¹*.

榦 蔽 懈 蔭 肥 錄 網 犬 稜 肥
nji² tjj¹ mji¹ sjt¹ nja² ku¹ kq¹ phja¹ wji¹ nja²
2sg[HON] if NEG go away -2sg then life cut do -2sg
‘If you don’t go away, I’ll kill you.’

The causal meaning is expressed by means of the causal conjunction 疠 *bju¹*.

姦 犬 師 術 繼 網 犬 稜 純 畏
sjt² dzjwo² ku² da² sjij¹ thjt² gji² kju¹ kjiir² u²
woman person answer say now this child carriage room in
縕 羣 總 肥 繢 網 犬 稜 純 畏
wji² wee¹ bju¹ tshjwo¹ mjiij² kju¹ tsə¹ ·ji²
PERF born because therefore name Carriage Son COMP
‘The woman said, “Because this child was born in a carriage stable, he is named Son of Carriage”.

The concessive meaning is expressed by means of the concessive conjunction 故 *lji¹*.

榦 犬 犬 稜 網 犀 犬 犬 犬 犬
nji² nji² tṣjiw¹ ·ji¹ sjt² mji¹ dzji¹ lji¹ ·ji² dji² tṣjiw¹
you [pl,HON] Zhou POSS grain NEG eat although COMP but Zhou
矣 通 遂 犬 犀 犀 犀 犀 犀 犀
ŋər¹ lwu² dzjiij¹ tṣjiw¹ naa² dzjij¹ dzji¹ tja¹ thji² yie¹ ·ji²
Mountain hidden dwell Zhou vegetable only eat TOP why true COMP
‘Why is it so that although you refuse to eat the Zhou grain, yet you dwell in seclusion on Zhou Mountain and eat the Zhou vegetables?’

FURTHER READING

- Gong, Hwang-cherng (1988) ‘Phonological alternations in Tangut’, *Bulletin of the Institute of History and Philology (BIHP)* 59.3: 783–834.
— (1989) ‘The phonological reconstruction of Tangut through examination of phonological alternations’, *Bulletin of the Institute of History and Philology (BIHP)* 60.1: 1–45.
— (1993) ‘Phonological alternations and derivational morphology in Tangut’, *Bulletin of the Institute of History and Philology (BIHP)* 64.4: 935–968.
— (1994) ‘A hypothesis of three grades and vowel length distinction in Tangut’, *Journal of Asian and African Studies* 46–47: 305–314.
— (1997) ‘A possible origin of certain Tangut syllable-final phonological alternations – reduplicated compounds’, *Chinese Languages and Linguistics IV (Typological Studies of Languages in China)* 4: 265–289.

- (2001) ‘Personal agreements and phonological alternations in the Tangut verb’, *Language and Linguistics* 2.1: 21–67.
- Kepping, Ksenia B. (1983) «Les Kategorij» – *Utrazennaja Kitajskaja Lejšu v tangutskom Perevode (The Grove of Classifications: A Lost Chinese Encyclopedia in Tangut Translation)*, facsimile of the xylograph, publication of text, introductory essay, translation, and notes, Moscow: Izdatel’stvo ‘Nauka’.
- (1985) *Tangutskij jazyk, morfologija (The Tangut Language: Morphology)*, Moscow: Izdatel’stvo ‘Nauka’.
- LaPolla, Randy J. (1992) ‘On the dating and nature of verb agreement in Tibeto-Burman’, *Bulletin of the School of Oriental and African Studies* 55.2: 298–315.
- Laufer, Berthold (1916) ‘The Si-hia language, a study in Indo-Chinese philology’, *T’oung-pao* 17.2: 1–126.
- Nevskij, Nikolaj Aleksandrovic (1960) *Tangutskaja filologija: Issledovanija i slovar’* (in two vols), Moscow: Izdatel’stvo Vostočnoj Literatury.
- Shi, Jinbo, Huang, Zhenhua and Nie, Hongyin (1993) *Leilin Yanjiu (A Study in the Forest of Categories)*, Yinchuan: Ningxia Renmin Chubanshe.

REFERENCES

- Kepping, Ksenia B. (1976) ‘Subject and object agreement in the Tangut verb’, *LTBA* 2.2: 219–231.
- (1982) ‘Deictic motion verbs in Tangut’, *LTBA* 6.2: 77–82.
- (1994) ‘The conjugation of the Tangut verb’, *Bulletin of the School of Oriental and African Studies* 57.2: 339–346.
- Nishida, Tatsuo (1964/1966) *Seikago no Kenkyū, Seikago no Saikōsei to Seikamoji no Kaidoku (A Study of the Hsi-Hsia Language: Reconstruction of the Hsi-Hsia Language and Decipherment of the Hsi-Hsia Script)*, Vols 1 and 2, Tokyo: Zauhō Kankōkai.
- (1986) ‘Seikago dōshi kukōzō no kōsatsu’ (A study of the structure of Hsi-Hsia verb phrases), *Memoirs of the Faculty of Letters Kyoto University* 25: 84–116.
- Sofronov, Mikhail Viktorovich (1968) *Grammatika tangutskogo jazyka* (in two vols), Moscow: Izdatel’stvo ‘Nauka’.

CHAPTER FORTY-TWO

LIZU (ERSU)

Katia Chirkova

1 INTRODUCTION

Lizu is one of three closely related Tibeto-Burman languages spoken in Sichuān 四川省 Province in the People’s Republic of China: Ersu, Lizu and Duoxu (see Map 42.1). The three languages are currently classified as dialects of one Ersu language (ISO-639 code *ers*). The Ersu language itself is classified as a member of the Qiangic subgroup of the Tibeto-Burman language family (e.g. Bradley 1997: 36–7; Sun 2001).

The Lizu people refer to themselves as /lî-zû/ or /lŷ-zû/ ‘white people’. In Chinese, the group is variously known as *Lírǔ* 里汝, *Lǚsū* 吕苏 or *Lísū* 僕苏 (Sun 1982; Huang and Renzeng 1991; Wang 2010, 2012); in English, as Lizu or Lyuzu (Chirkova 2008; Ikeda 2009; Yu 2012). The total number of speakers is estimated at 7,000 people (Wang 2010: 3).

Lizu is part of a continuum of closely related linguistic varieties which are spoken in Jiǔlóng County (九龙县, Written Tibetan, hereafter WT, *brygad zur*), Mùlì Tibetan Autonomous County (木里藏族自治县, WT *smi li rang skyong rdzong*) and Miǎnníng County (冕宁县) (see Map 42.2). Of these three locations, the group has the longest history of residence in Jiǔlóng. Migration to Miǎnníng and Mùlì started from the second half of the Qing 清 dynasty (1644–1911) and was triggered by an influx of Chinese and Yí migrants to Jiǔlóng around that time (Wang 2012: 35).

Lizu is spoken in a multiethnic area and has been influenced throughout its history by many languages, most importantly, Tibetan, Southwest Mandarin (hereafter Mandarin), Pǔmǐ 普米, Námuyī 纳木依 and Northern Yí 彝.

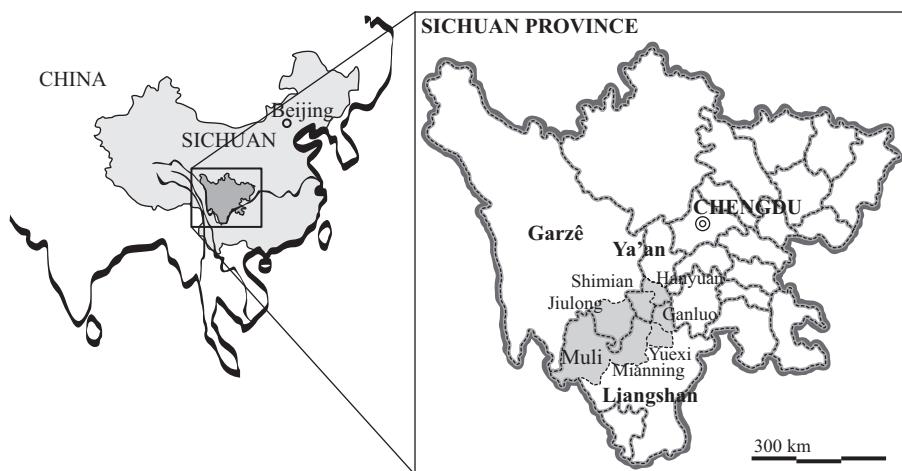
Lizu is essentially used as the primary language of oral communication in family and community events. Most speakers are bilingual in Mandarin, and the current trend for the school-going generation is to become monolingual in Mandarin.

Lizu is but little documented and described, with only two grammatical sketches to date, both of which focus on the Lizu varieties spoken in Mùlì County (Huang and Renzeng 1991; Chirkova 2008). In addition, there are phonological sketches of the Lizu varieties of Miǎnníng, Mùlì and Jiǔlóng (Ikeda 2009; Yu 2012; Chirkova and Chen 2013), an analysis of auxiliary verbs in the Lizu variety of Miǎnníng (Lin *et al.* 2014), and basic vocabulary lists for Mùlì and Jiǔlóng varieties of Lizu (*Zang-Mianyu Yuyin he Cihui Bianxie zu* 1991; Huang *et al.* 1992; Ikeda 2009). The present overview is based on first-hand fieldwork data on /ʃæt̥əhō pâ/ ‘eastern dialect’, as spoken in Kǎlā 卡拉 Township of Mùlì County.

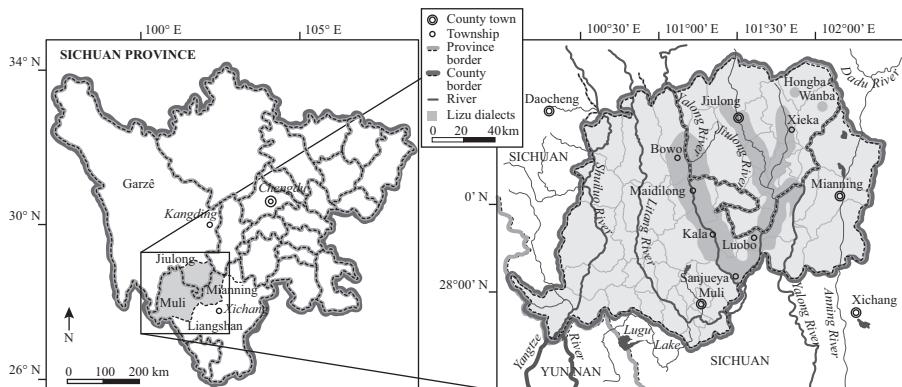
2 PHONOLOGY

2.1 Initial consonants

The Lizu consonant inventory consists of 39 phonemes, listed in Table 42.1. Velar and uvular stops contrast before /o/, e.g. /kǒ/ ‘beg’ vs /qô/ ‘hole, pit’. Elsewhere, they are in



MAP 42.1 DISTRIBUTION OF ERSU, LIZU, AND DUOXU (MAP BY FRANZ HUBER)



MAP 42.2 DISTRIBUTION OF LIZU DIALECTS (MAP BY FRANZ HUBER)

TABLE 42.1 LIZU INITIAL CONSONANTS

	Bilabial	Alveolar	Postalveolar	Alveolopalatal	Velar	UVular	Glottal
Plosive	p p ^h b	t t ^h d			k k ^h g	q q ^h	
Affricate		ts ts ^h dz	tʃ tʃ ^h dʒ	tɕ tɕ ^h dz			
Nasal	m	n		jn		ŋ	
Fricative		s z	f ʒ	ç z	x ɣ		h
Approximant	w	r		j			
Lateral	l						
Lateral fricative	ɿ						

complementary distribution. /g/ only occurs in the prenasalized cluster [Ng] and only before /o/ and /e/. It shows varying degrees of frication at the beginning of the following vowel, after the release, e.g. /ne-Ngê/ [ne-Ng̥ê] ‘stab’. The voiced velar fricative has a uvular allophone occurring before /o/, /e/ and /we/, e.g. /χô/ [χ̩ô] ‘wine’.

/h/ is a nasal voiceless approximant (e.g. /hē/ ‘bamboo’). Before /e/, /ɪæ/ and /ɪə/, it is realized as voiced (e.g. /hè/ [ɦ̩è] ‘five’, /hɪɔ̄/ [ɦ̩ɪɔ̄] ‘mushroom’). Comparative Ersu-Lizu-Duoxu evidence suggests that words with the /h/ initial derive from voiceless nasals (*m *ŋ *j) at the Proto-Ersu-Lizu-Duoxu level. The set of words manifesting the voiceless nasal correspondence pattern is one shared innovation that can set Ersu, Lizu and Duoxu apart from other languages in the region (Chirkova and Handel 2013).

/ŋ/ is the only syllabic consonant. It only occurs after /k/ and in very few words (e.g. /kŋ̩/ ‘seven’).

2.2 Initial consonant clusters

Lizu has complex consonant onsets of three types:

- (a) prenasalized clusters with a homorganic nasal ('N'), which may occur before voiced and voiceless aspirated stops and affricates (e.g. /Nbō/ ‘hat’, /Npʰō/ ‘be diligent’);
- (b) clusters with approximants (j w ɪ) (e.g. /mjæ/ ‘face’, /xwæ/ ‘bird’, /mɪæ/ ‘be tasty’);
- (c) clusters with fricatives (z z ɛ ſ) (e.g. /bz̥e/ ‘fly (v.)’, /ʃibz̥e/ ‘green’, /pēi/ ‘throw’, /pʃ̥/ ‘Tibetan’). In addition, /ptsʰ/, is observed before low vowels (e.g. /kʰe-ptsʰæ/ ‘taste’).

Elements in the second position in consonant clusters have restricted distribution: (i) /j/ only occurs after bilabials and laterals; (ii) /w/ occurs after postalveolars, velars, uvulars, /h/ and /ɪ/; (iii) /ɪ/ is restricted to bilabials and /h/; (iv) /z z ɛ ſ/ are restricted to /b p/.

2.3 Vowels

Lizu has eight vowel phonemes, listed in Table 42.2. Nasalization is not contrastive in vowels. It is observed before prenasalized clusters (e.g. /mēNt̥hō/ [mēnt̥hō] ‘tail’), in the environment of the /h/ initial (e.g. /hû/ [ɦ̩û] ‘want’), and in recent loanwords from Tibetan and Mandarin, where the donor language has a nasal coda (-m, -n, -ng) (e.g. /s̥ə mû/ ‘make smoke offering’, WT *bsang*). (In old loanwords, the original nasal element is in most cases lost without compensation, as in /jēy/ ‘potato’, Mandarin /iaŋ²¹y²¹³/ 洋芋.) Nasalized vowels are also exceptionally found in two native words: /ōw̥e/ ‘goose’, /ɦ̩jæ/ ‘duck’. Diphthongs are observed in recent loanwords from Mandarin (e.g. /xai/ ‘still’, Mandarin /xai²¹/ 还).

2.4 Syllable structure

The syllable structure is (N)(C1)(C2)VT, where N is nasal; C1 can be any initial consonant; C2 can be one of the following set: j, w, ɪ, z, z̥, ɛ, ſ; V is vowel; T is tone; and parentheses indicate optional consonants. The minimum syllable is VT, e.g. /æ/ ‘I’.

TABLE 42.2 LIZU VOWELS

i	y	u
e	ɛ	o
æ		a

The maximum syllable is NC1C2VT, e.g. /Nb̥j̥/ ‘mountain’. Most syllables have a simple structure C1VT, e.g. /n̥e/ ‘you, thou’.

Lizu morphemes are generally monosyllabic but words are generally disyllabic (over 80 per cent of the collected vocabulary). Monosyllables are of two types: (i) roots (free and bound), and (ii) affixes. Monomorphemic words are for the most part monosyllabic, e.g. /dʒ̥e/ ‘water’. Polysyllabic words are mostly composite, e.g. /dʒ̥e-kʰwæ/ ‘flood’ (from /dʒ̥e/ ‘water’, /kʰwæ/ ‘be large’), /tōNbu-mu/ ‘nose hair’ (from /tōNbu/ ‘nose’, /mū/ ‘animal hair, fur’). There are also a handful of disyllabic and trisyllabic monomorphemic words, e.g. /m̥idz̥ə/ ‘hare, rabbit’, /ʃ̥ēNb̥elj̥e/ ‘buttocks’.

2.5 Prosodic system

Lizu has a hybrid prosodic system which combines lexical tone on monosyllabic words and prominence patterns with stress-like and tonal characteristics on polysyllabic words (both monomorphemic and polymorphemic) and compounds. Roots are lexically specified for tone, whereas affixes are toneless.

The two contrastive lexical tones are: rising and falling, e.g. /l̥e/ ‘be old’ vs /l̥e/ ‘be heavy’. The three prominence patterns are:

- (a) The Equally-Prominent Pattern: both syllables sound equally prominent, each syllable has a mid-level pitch contour, though the first is slightly higher than the second, and the second syllable has a falling pattern, e.g. /m̥idz̥ə/ ‘hare, rabbit’.
- (b) The Left-Prominent Pattern: the first syllable is relatively longer and sounds more prominent. The f0 peak is typically realized before the end of the first syllable, where the pitch starts to fall and continues to fall in the second syllable, e.g. /m̥idz̥ə/ ‘pepper’.
- (c) The Right-Prominent Pattern: the second syllable is relatively longer and sounds more prominent. Within the first syllable, there is a mid-level pitch contour with a slight rise. The f0 peak is realized within the second syllable where there is also a clear fall, e.g. /mutsə/ ‘cat’.

In words, compounds, and combinations of words with clitics (nominal and verbal particles), lexical tones undergo sandhi changes, where only the tonal contour of the domain-initial root or word is retained and realized over the entire domain. If the domain-initial monosyllabic root or word has the rising tone, the resulting pattern is generally Right-Prominent. Examples include /iwæ-dz̥e/ ‘chicken pen’ (from /iwæ/ ‘chicken’, /dz̥e/ ‘shed, pen’), /dzə=bō/ ‘be eating’ (from /dzə/ ‘eat’, the egophoric progressive aspect auxiliary /bō/; tone is not indicated for those forms that do not occur in isolation). If the domain-initial monosyllabic root or word has the falling tone, the resulting pattern is generally Equally-Prominent. Examples include /wō-dz̥e/ ‘pig shed’ (from /wō/ ‘pig’), /tō=bō/ ‘be looking’ (from /tō/ ‘look’). In domains that begin with a disyllabic root or word, the resulting contours are determined by the tonal contour of the initial disyllabic item. Examples include /sâNg̥e-mêNt̥hō/ ‘lion’s tail’ (from /sâNg̥e/ WT *seng ge* ‘lion’, /mêNt̥hō/ ‘tail’), /tōNbu wu-li/ ‘tip of the nose’ (from /tōNbu/ ‘nose’, /wū-li/ ‘head’), /mutsə Ndōqō/ ‘cat’s eye’ (from /mutsə/ ‘cat’, /Ndōqō/ ‘eye’).

In words beginning with (toneless) affixes, words with the root with the rising tone generally have the Right-Prominent Pattern (as in /ne-dz̥ə/ ‘eat up’, from /dzə/ ‘eat’), whereas words with the root with the falling tone generally have the Left-Prominent Pattern (as in /dē-dz̥ə/ ‘give birth’, from /dz̥ə/ ‘give birth’).

These general rules are not without exceptions. To use reduplicated forms as examples, in addition to the two regular patterns (that is, Right-Prominent and Equally-Prominent, as in /je-jē/ ‘be small’, /cî-cî/ ‘be fine’), some lexicalized reduplicated forms also have the irregular Left-Prominent Pattern (as in /dzô-dzo/ ‘be flat’). Some exceptions are also found in productive compounds and in combinations of words with clitics. Examples include /Nbjə-meNtʃ̩ô/ ‘horse tail’ (from /Nbjâ/ ‘horse’), the expected pattern is Equally-Prominent, */Nbjə-mêNtʃ̩ô/; /æ=i/ ‘my’, (from /æ/ ‘I’, the genitive particle /i/), the expected form is */æ=i/. For these reasons, prominence patterns on polysyllabic domains are notated phonetically.

3 MORPHOLOGY

Lizu is isolating (weakly agglutinative). Major word-formation processes include compounding, affixation and reduplication. The majority of affixes are derivational.

Nominal prefixes and suffixes are restricted to nouns referring to animate referents. They include: (i) one vocative prefix /æ-/ in kinship terms, e.g. /â-wu/ ‘maternal uncle’; (ii) three gender suffixes: (a) one feminine /-mæ/, e.g. /wô-mâ/ ‘sow’; (b) two male /-Np^hæ/ or /-bu/, e.g. /wô-Np^hâ/ ‘boar’, /yâ-bû/ ‘male yak’; and (iii) one diminutive suffix /-je/, e.g. /wô-jê/ ‘piglet’.

Important verbal prefixes include: (i) three directional prefixes: /de-/ ‘upward’, /ne-/ ‘downward’, /k^he-/ ‘inward’ (as in /dê-ji/ ‘go up’, /nê-ji/ ‘go down’, /k^hê-ji/ ‘go inside, enter’); (ii) one perfectivizing suffix, /t^he-/ (as in /t^he-k^he=ë/ ‘have given’); and (iii) one comparative prefix /jæ-/ ‘more’ (e.g. /(n)i jæ-ljê/ ‘(illness) get better’, from /ljë/ ‘be good’).

Reduplication is mostly attested on verbal roots. It is part productive and part lexically idiosyncratic. It displays cross-linguistically recurrent meanings (e.g. Moravcsik 1978), such as iterativity and repetition (/te^he-te^hê/ ‘chop’), reciprocity (/wô-wô/ ‘help [each other]’) and intensivity (/fë-fë/ ‘be [very] long’).

Inflectional morphological changes are limited. Vocalic alternation is attested in the derivation of the imperative stem of the verb ‘come’: /lî/ (imp) from /læ/. Consonant alternation is observed in a few lexical causative verbs. Such verbs contain voiceless initials, whereas the corresponding non-causative verbs contain voiced initials, e.g. /tsô/ ‘feed’ vs /dzô/ ‘eat’; /fû/ ‘dress’ vs /wû/ ‘wear’. Suppletion is attested in the formation of the verb ‘go’, which has two stems: non-past /jî/ (both egophoric and non-egophoric) vs past /dë/ (non-egophoric).

Lizu has many loanwords from the languages with which it has been in contact, most importantly, Tibetan and Mandarin. Examples include /ŋêmô/ ‘camel’ (WT *rnga mo*), /qômr/ ‘emperor’ (WT *gong ma*), /kûNtš^hê/ ‘coffin’ (Mandarin /kuan⁴⁴ts^hai²¹/ 棺材), /jëy/ ‘potato’ (Mandarin /ian²¹y²¹³/ 洋芋).

Recorded neologisms are coined by extension and meaning shift of native Lizu words, e.g. /zæ-tî pimâ/ ‘cell phone, lit. speech-transmitting frog-shaped device’, /fë-dzû/ ‘plane, lit. iron crane’.

4 THE NOUN PHRASE

Nouns are those forms that can take (in)definite marking, numeral-classifier phrases and nominal particles (analytic case markers). Nouns can modify other nouns directly or in a genitive phrase (both in the pre-head position). There is no agreement with nouns of any kind marked on the verb.

The elements of the noun phrase and the order in which they may occur in relation to the head noun (N) are:

Dem	Gen phrase/Rel clause	N	Adj	Num	Clf	Def	Particles
-----	-----------------------	---	-----	-----	-----	-----	-----------

- (1) [[ŋô-Ngô læmû ne-ɦukæ=î] jâeqâ=bî]
 copper-pot INS downward-cover=GEN child=DEF
 ‘the child who was covered with a copper pot’
- (2) teĕ de-nê eĕ-pû
 cloud upward-black three-item
 ‘three black clouds’

Definiteness is expressed by three definite markers (i) /bi/ (singular, both animate and inanimate), e.g. /jâeqâ=bî/ ‘the child’, /gâemi=bi/ ‘the garment’; (ii) /bo/ (plural, animate), e.g. /jâeqb=bô/ ‘the children’; and (iii) /bæ/ (plural, inanimate), e.g. /gâemi=bæ/ ‘the garments’. Alternatively, definiteness can be expressed by the use of a bare classifier, e.g. /mâ=kâ/ arrow=strip ‘the arrow’. The numeral /tê/ ‘one’ can be optionally used as an indefinite marker, as in /meljô tê/ ‘a lightning’.

Possession may be expressed by a simple juxtaposition of two nouns or noun phrases, e.g. /læpʰâ-Ndʒâ/ ‘tiger skin’. Alternatively, possession, part–whole relationship and other related meanings may be expressed by the genitive particle /i/, e.g. /læpʰâ=i mèNt̥h/o/ ‘tiger’s tail’. /i/ may fuse with the root vowel of its host, also triggering palatalization on the initial of the host. Some frequent forms that result from this type of fusion include: (i) /teʰi/, the genitive form of the third person singular pronoun /tʰê/; and (ii) /dzì/ ‘of the family’ (from /dê/ ‘family’), as in (20).

Other nominal particles, which indicate the semantic and pragmatic roles of the noun phrase in the clause, include:

- (a) Contrastive topic and focus marker /le/, and non-contrastive topic marker /ne/ /le/ indicates the presence of alternatives to the topicalized or focalized element. It typically occurs in parallel predication, in which two referents are contrasted for some properties. In (3), /le/ marks clause-initial topics (see also (18), (22), (27) and (31) for more examples).

- (3) [Nbõ tsu=sû=bi le] [kʰwæ-kʰwâ]
 hat wear-person=DEF CTR father’s.elder.brother-father’s.elder.brother
 [o-tʰê teĕ sə-sə-su=bî le]
 that-that hair be.long-be.long-person=DEF CTR
 [nâ-pâ]
 father’s.married.sister-father’s.married.sister
 ‘The one wearing a hat is my uncle, that one with long hair is my aunt.’

In (4), /le/ marks contrastive focus in immediate pre-verbal position. It signals that the focused content is different from the hearer’s expectation. This sentence is an answer to the suggestion that the speaker, a magic animal, is the one who ate the hearer’s father.

- (4) [ne=î â-bâ dzə=sû] [â le] me=zî
 2sg =GEN VOC-father eat=NMLZ.AGT 1SG CTR NEG=COP
 ‘The one who ate your father is not me.’

/le/ can follow another nominal particle, as in (5):

- (5) dʒəmæ=bi=v le mě tsə=bo=tô
wife=DEF=N-AGT CTR butter feed=PROG.EGO=MIR
'As for [Gesar's] wife, [the demon] was apparently feeding her butter [to fatten her in order to eat her].'

The non-contrastive topic marker /ne/ refers to the given information which is shared by the speaker and the hearer. Consider sentence (6), which occurs in a traditional story after sentence (5), and in which 'butter', already introduced in the story, is marked by /ne/.

- (6) mě=bi ne zō ne-dzâ
butter=DEF TOP 3SG.N-VSB downward-eat
'As for butter, he ate it [himself].'

(b) Non-agentive marker /v/ (or /wə/, in free variation)

/v/ signals primarily human arguments of the verb (except for agent). It indexes the (definite and animate) patient in monotransitive clauses (as in (7)), and the goal or beneficiary in ditransitive clauses (as in (5)).

- (7) nê=v mə=sâ sê=v sə=qê?
2SG=N-AGT NEG=kill who=N-AGT kill=PROG.N-EGO
'If I don't kill you whom do I kill?'

(c) Coordinative /læ/ 'and', as in /təjê læ jê-pê/ 'today and yesterday'

(d) Instrumental /læmû/ (possibly, a combination of /læ/ 'and' and the verb /mû/ 'make'), e.g. /ŋô-Ngô læmû ne-ɦukâ/ 'cover with a copper pot' (as in (1))

(e) Locative /ke/ 'at, on', e.g. /gôpu=ke/ 'on the plane'

(f) Comparative /pə/, as in /t̪ê=pê/ 'like that'. /pə/ is also the marker of the basis of comparison in the comparative construction:

- (8) jê-pê tteje=pê jæ-de-ts^bâ
previous?-day today=CMPR more-upward-be.hot
'It was hotter yesterday than it is today.'

4.1 Pronouns

Personal pronouns in Lizu distinguish singular, dual and plural number in all persons. Singular forms are /â/ 'I', /nê/ 'you, thou' and /t^bê/ and /zō/, both meaning 'he/she/it'. /t^bê/ refers to something or someone visible in the speech situation. It functions as both pronoun and demonstrative. /zō/ refers to something or someone absent or out of sight.

Dual forms are formed with the root /dze/, e.g. /æ-dzê/ 'the two of us'. They may be optionally followed by the expression /ne-t^bê/ 'two-that', e.g. /æ-dzê ne-t^bê/ 'the two of us'.

Plural forms are formed with the root /iə/, e.g. /â-iə/ 'we', /nê-iə/ 'you', /t^bê-iə/ and /zô-iə/ 'they'.

Demonstrative pronouns are /kû-t^bê/ 'this' and /o-t^bê/ 'that'. Plural forms are formed with the plural definite markers /bæ/ and /bo/, /kû-t^bê=bâ/ 'these', /o-t^be=bô/ 'those (people)'.

Lizu has one anaphoric pronoun, /jô/ ‘self’. It is typically used in logophoric function in indirect speech:

- (9) jô læ tê-Ndzê=sâ ̄hû dži=gê
self and one-ride=INF? want say=PROG,N-EGO
“I would also like to ride [on the tree branch],” said [the bear].”

/jô/ may be reduplicated to express reflexive meaning:

- (10) milô=i kʰû jo-jo=ê kʰe-Ndo=ĕ
mirror=GEN inside self-self=N-AGT inward-see=CS
‘He saw himself in the mirror.’

The main interrogative pronouns include /xâ[-te]/ ‘what, lit. what-one’, /sê/ ‘who’, /kʰâ[-ke]/ ‘where’, /nî-kʰâ/ ‘when’, /têhî-bi/ ‘how many’.

4.2 Numerals

Lizu has a decimal counting system, with numerals for ten, /têh-e-têhâ/; 100, /zâ/, WT *bryga*; 1,000, /tû/, WT *stong*; and 10,000, /Nbô/, WT *bum*. Cardinal numerals one to nine are: /tê/ ‘one’, /ně/ ‘two’, /cê/ ‘three’, /zê/ ‘four’, /hě/ ‘five’, /tjhû/ ‘six’, /kñ/ ‘seven’, /dzê/ ‘eight’, /Ngê/ ‘nine’.

Lizu has a multiplicative-additive number system. For example, the numeral /zê-zâ Ngê/ ‘49’ is formed by multiplying ten by four and then adding nine. Numbers above 100 are formed with the intrusive conjunction /læ/ ‘and’ between the 100 and the adjoined number, e.g. /tê zâ læ tê/ ‘101’. Numeral formation is mostly regular. Numbers for ‘one’, ‘three’ and ‘ten’ distinguish between free forms (as above) and bound forms: (i) /tç/ ‘one’ in /tsʰe-tç/ ‘11’; (ii) /sâ/ ‘three’ in /tsʰâ-sâ/ ‘13’ and /sâ-tsʰâ/ ‘30’; and (iii) three bound roots for ‘ten’: (a) /tsʰe-/ in the teen numbers, (b) /-tsʰə/ in /nê-tsʰə/ ‘20’ and /sâ-tsʰâ/ ‘30’, and (c) /-zâ/ in numbers from 40 to 90.

Ordinal numbers are formed by adding the form /pʰôe/ to cardinal numbers, e.g. /pʰôe-tê/ ‘first’, /pʰôe-ně/ ‘second’, etc.

4.3 Classifiers and measure words

Lizu classifiers and measure words only combine with non-human nouns. They are used together with numerals to count nouns or to substitute for them.

Lizu has two bound classifiers: (i) /pu/ ‘item’ for small objects, e.g. /rwâ tê-pû/ ‘one chicken’; and (ii) /kæ/ ‘strip’ for elongated objects, e.g. /břæ tê-kâ/ ‘one rope’. They may occur bare and trigger a definite interpretation (e.g. /břæ=kâ/ ‘the rope’).

Free forms which lend themselves to classifier use can be subdivided into measure words (e.g. /džê tê kezâ/ ‘one bucket of water’), and repeaters, i.e. classifiers that have the same form as the noun they modify (e.g. /gû tê gû/ ‘one sentence’, /Nbřê zê Nbřâ/ ‘four horses’).

5 THE VERB PHRASE

Verbs are those forms that can take directional or perfectivizing prefixes, the causative marker /su/, and the interrogative and negative marking. Verbs can be preceded by adverbial expressions, followed by markers expressing aspect, evidentiality and modality, and

nominalized by one of the nominalizers (see section 6.3.1). Most verbal markers behave like suffixes. They fall into two groups: (i) those that can be preceded by negative and question markers (hereafter auxiliaries, e.g. the progressive auxiliaries /bo/ and /ge/), and (ii) those that cannot (hereafter particles, e.g. the inferred particle /sæ/).

Lizu does not have grammaticalized tense. Important types of aspects include perfective, imperfective, change of state, experiential, progressive, resultative and continuative. For most auxiliaries and particles, the expression of aspect interacts with that of evidentiality. Lizu has a hybrid evidentiality system, which combines specification of source of information (direct, inferred, reported) and specification of speaker's perspective towards the source of, and access to, information (egophoric vs non-egophoric). The latter type bears close similarity to the 'egophoric' evidential pattern best described for Tibetic languages (e.g. Tournadre 2008; Tournadre and LaPolla 2014). Lizu egophoric markers are prototypically correlated with the overt presence of the first person in an utterance or with its anticipation in direct questions (the answer to which normally uses the first person); non-egophoric markers are used in all other cases. Use of the markers is also correlated with the semantic category of the verb: (i) controllable (verbs that denote actions performed by knowing agents of their own will, e.g. 'drink') vs (ii) non-controllable (verbs that denote actions over which agents have no direct control, e.g. 'cough') (e.g. Tournadre 2008). Controllable verbs may be used with both egophoric and non-egophoric markers, while non-controllable verbs are generally only used with non-egophoric markers. Overall, use of an egophoric marker implies that the speaker is responsible for conceptualizing or observing the reported event or situation and committed to its truthfulness; whereas use of a non-egophoric marker has the opposite effect (see section 5.1 for examples).

Adjectives ('lexeme[s] that denote a descriptive property and that can be used to narrow the reference of a noun', Haspelmath 2010: 670) are formally a subset of verbs (intransitive stative verbs). Like verbs, adjectives function as (intransitive) predicates, take verbal prefixes, and the causative, interrogative and negative marking. In contrast to other verbs, adjectives can modify a noun directly in the post-head position (e.g. /Nbřâ dê-ly/ 'white horse'), whereas other verbs can modify a noun only in the pre-head position (e.g. /Ndzê-Nbřâ/ 'riding horse').

Lizu has a subclass of existential verbs that categorize S/O arguments in terms of their animacy and orientation in space:

- (a) /dʒô/ 'have, exist (animate)', e.g. /jæqê dʒô/ 'have children'
- (b) /dʒwă/ 'exist (inanimate)', e.g. /tʰê=kê læpʰî tê dʒwă/ 'there is a vase there'
- (c) /bô/ 'possess (inanimate)', e.g. /bædžê bô/ 'have money'
- (d) /nê/ 'have (abstract)', e.g. /sâ nê/ 'have things to do'
- (e) /hă/ 'exist (be attached)', e.g. /se-pû tê-pû hă/ 'there is a tree [there]'
- (f) /dʒê/ 'exist (inside a container)', e.g. /somô dʒê/ 'have strength'

Lizu does not have passive voice as a morphological category. To use as examples situations depicted in the 'Cut and break clips' (Bohnemeyer *et al.* 2001), a situation in which a man tears a piece of cloth (clip cb01chand.mpg) and that in which a piece of cloth gets torn by itself (clip cb08cspont.mpg) are both described in Lizu with the same verb form: /tʰe-tsê dě/ PFV-tear N-EGO, respectively /tʰê gämi tʰe-tsê dě/ 3SG.VSB garment PRF-break N-EGO 'he tore the cloth' vs /gämi=bi tʰe-tsê dě/ garment=DEF PRF-break N-EGO 'the garment got torn'. The functional equivalent of a passive construction is hence derived by

omitting the agent and using a transitive verb ambitransitively. If the agent is not omitted, it may be marked by the nominal expression /le-kê/ ‘hand?-LOC?’:

- (11) kû-t^hê ts^h=bi (lêp^h=i le-kê) ne-dzâ d^h
this-that goat=DEF tiger=GEN hand?-LOC? downward-eat N-EGO
‘This goat was eaten (by a tiger).’

5.1 Aspect and evidentiality

The basic distinction between Lizu perfective and imperfective aspect is expressed by lexical-derivational means (verbal stems and prefixes). Bare verb stems tend to be imperfective. Perfective forms are derived by adding (telicity-inducing) prefixes to bare verb stems (e.g. /dz^h/ ‘eat’ vs /ne-dzâ/ ‘eat up; have eaten’, /ts^h/ ‘feed’ vs /t^h-ts^h/ ‘have fed’).

Bare verb stems may be used to refer to generic facts and permanent situations. Prefixed verb stems may be used to refer to iterative bounded events:

- (12) â nê=v xâ-te dzî t^hê ne-mû
1SG 2SG=N-AGT what-one say 3SG.VSB downward-make
‘Whatever I tell him, he does.’

References to concrete bounded events often combine the use of a (prefixed) verb form with the change of state particle /v/, which denotes a change of state or situation (e.g. /k^he-Ndo=v/ inward-see=CS ‘saw’, as in (10)). In reference to past events, /v/ stands in opposition to the inferred particle /sa/ (which denotes that the reported event is not directly witnessed by the speaker, but deduced from sensory observation) and the non-egophoric particle /d^h/ (likely from the non-egophoric form of the verb ‘go’, /d^h/). By contrast with these two particles, /v/ comes to be associated with an egophoric reading. Compare the three particles in the same context:

- (13a) t^hê k^hâ=dz^h jæqw^v=bâ dz^h t^hê-dzy =sæ
(13b) =d^h
(13c) =v
3 SG.VSB cooked.rice-eat when all=DEF.PL eat PFV-finish =INF
N-EGO
=CS

‘At dinner, he ate it all.’ (13a: inferred, for example by empty plates on the table; 13b: non-egophoric; 13c: egophoric, direct source of information)

The experiential auxiliary /m^h/ indicates that an event has taken place at least once in the past and could be repeated:

- (14) Ndžû-dz^h k^hû-p^hô lă m^h
lunch-eat inside-side come EXP
‘[He] came by in the morning.’

Progressive aspect marking distinguishes between the egophoric auxiliary /bo/ and the non-egophoric auxiliary /ge/:

- (15) â t^h=bô z^h t^h=g^h.
1SG look=PROG.EGO 3SG.N-VSB look=PROG.N-EGO
‘I am looking. He is looking.’

/bo/ mostly co-occurs with controllable verbs, whereas /ge/ freely co-occurs with both controllable and non-controllable verbs, albeit with different meanings. Used in combination with a

controllable verb and the second or third person actor, /ge/ denotes an activity in progress (as in (15)). Used in combination with a controllable verb and the first person actor, /ge/ denotes an activity certain to occur in the immediate future (e.g. /â tō=qê/ 1SG look=PROG.EGO ‘I will look’). Used in combination with a non-controllable verb /ge/ always has a progressive reading, regardless of the person of the actor (e.g. /â tsʰê=qê/ 1SG cough=PROG.N-EGO ‘I am coughing’, /zō tsʰê=qê/ 3SG.N-VSB cough=PROG.N-EGO ‘He is coughing’). Conversely, a combination of a non-controllable verb with the first person egophoric marker implies that the speaker has some control over the execution of the action or that she performs it deliberately (e.g. /â tsʰê=bô/ 1SG cough=PROG.EGO ‘I am coughing [on purpose]’).

Adjectives (which are in their majority prefixed verbal forms) generally do not co-occur with either /bo/ and /ge/. Such combinations may have an inchoative reading (e.g. /mjæpʃə dê-nifu=ge/ face upward-be.red=PROG.N-EGO ‘the face is beginning to get red’). Alternatively, they may portray the actor as actively experiencing the reported event (triggering certain pragmatic implications) (e.g. /â de-Npʰje=bô/ 1SG upward-be.cold=PROG.EGO ‘I am feeling cold [please fetch a coat for me].’).

The resultative particle /tæ/ is used to indicate the resulting state of an event.

- (16) zō pûNgê de-tr-êtæ læ ne-zî=tæ
 3SG.N-VSB knee upward-hold-hold and downward-sit=RES
 ‘She sits with her knees pulled up.’

The continuative particle /ei/ indicates that the event is continuous or reiterative.

- (17) â jî de-njê ei
 1SG belly upward-be.hungry CONT
 ‘I am still hungry.’

5.2 Modality

Future events are expressed using modal particles and auxiliary verbs. These include:

(a) The epistemic particle /tʃü/ ‘likely’, which signals likely and hypothetical situations:

- (18) tjejê le mě ne-Ndê=sâ dê-tsʰæ tʃü
 today CTR sky downward-be.good=INF upward-be.hot likely
 ‘Today the weather turned out lovely, it’s going to be a hot day.’

(b) The auxiliary verbs /Ndʒō/ ‘know how to’ (expressing the potential to perform an action, in relation to inevitable events, as in (19)) and /pʰâ/ ‘can’ (expressing possibility, as in (20)). The two auxiliary verbs can be used for both epistemic and deontic readings.

- (19) kû-tʰê hîê=bi ne-dzə=ë le nê-Nbzi
 this-that mushroom=DEF downward-eat=CS CTR downward-be.poisoned
 Ndʒō
 know.how.to
 ‘If you eat this mushroom, you will die.’

- (20) tʰe-dzî wô-tsʰû lŷ mêNtʃö=i eo
 3SG.VSB-family.GEN pig-be.fat year tail=GEN top
 de-tsʰû pʰâ
 upward-be.fat can
 ‘Their pig may fatten by the end of the year.’

- (c) Desire to perform an action is expressed by the use of the auxiliary verb /hû/ ‘want’. Its use is restricted to first person actors in declarative sentences or their anticipation in direct questions. By contrast, used with second and third person actors, /hû/ must be followed by the reported marker /dzi=gê/ (see section 6.3.3):

- (21) zö jî-wu hû dzi=gê
3SG.N-VSB house-buy want say=PROG.N-EGO
'He wants to buy a new house.'

Unexpected information is signalled by the mirative particle /to/ (likely from the verb /tô/ ‘look’), which is added as the last element of the verb complex (as in (5)).

5.3 Imperative

The imperative form is the same as the citation form of the verb, but it is distinguished from it by the rising intonation. Both prefixed and unprefixed forms may be used for imperative, as in the following examples. The difference between the two is that the prefixed form implies a definite limit to the demanded action, whereas the unprefixed form has no implication as to whether the demanded action is to attain a definite limit or not.

- (22) ê-wu lêphâ nê le jæ-k^hwâ wûli dzð
voc-maternal.uncle tiger 2SG CTR more-be.large head eat
'Uncle tiger, you are bigger [than I am], eat the head [of the bear].'

- (23) xû-xû tê-mê ne-dzð
much-much one-bit? downward-eat
'Help yourself to more food.'

An imperative can be made more polite by adding the particle /me/ after the verb, as in /ne-dzð me/ ‘please eat up’.

5.4 Causatives

Lizu has two types of causative: (i) non-productive lexical causative (such as /tsø/ ‘feed’, /fû/ ‘dress’) and (ii) productive causative with /su/. The former type expresses direct causation, as is common cross-linguistically (e.g. Shibatani and Pardeshi 2002), whereas the latter type can express both direct and indirect causation.

Productive causative derivations can be formed from intransitive verbs (both activity and stative) and transitive verbs. Both types of verbs share the same marking: the causative particle /su/ is added to the verb stem, as in /læ=sû/ ‘make come’, /dê-dʒu=su/ ‘make dry’, /t^he-le=sû/ ‘make release’. The causee takes non-agentive marking, if denoting a human referent (as in (24)), or receives no overt case marking, in all other cases (as in (25)).

- (24) â zo=ë ne-zi=su=ë
1SG 3SG.N-VSB=N-AGT downward-sit=CAUS=CS
'I made him sit down.'

- (25) jë ne=t^hë=qo=su
blood downward=PROH=appear=CAUS
'Do not let any blood out.'

In non-past contexts, a combination of the causative particle and the change of state particle expresses permission: /nê jî=su=ë/ 2SG go=CAUS=CS ‘You may go now.’

5.5 Adverbs

High frequency adverbs that modify adjectives include: degree marking adverbs: /the-tei-ê/ ‘very’ (of verbal origin, etymology unclear; e.g. /the-tei-ê Nb̥iə-Nb̥iâ/ ‘very tall’) and /d̥ežu/ ‘particularly’ (e.g. /d̥ežu Nb̥iə-Nb̥iâ/ ‘particularly tall’). Greater intensity in adjectives is also commonly expressed by reduplication (also of disyllabic forms), e.g. /Nb̥iə-Nb̥iâ/ ‘very tall’, /d̥e-ly d̥e-ly/ ‘very white’. The comparative form of adjectives is formed with the comparative prefix /jæ-/ (e.g. /jæ-Nb̥iâ/ more-be.tall ‘taller’). The superlative degree of adjectives is formed by adding the form /jî-lâ/ to the comparative form, e.g. /jî-lâ jæ-Nb̥iâ/ ‘the tallest’. The two degree marking adverbs and the comparative prefix may also be used with some stative verbs, e.g. /ʃ̥Ng̥ê=i eo ſ̥jæ-hâ/ thigh=GEN top meat more-be.attached ‘there is more meat on the thighs [of the bear]’.

Manner adverbs are derived from adjectives by adding /mû/ ‘make’, as in /lje-lje-mû/ ‘carefully’ (from /ljé/ ‘be good’).

6 PHRASE AND CLAUSE STRUCTURE

Lizu is verb-final. Syntax operates predominantly through word order and the use of nominal and verbal particles and auxiliaries. The grammatical relations of subject and object are not grammaticalized. The clause structure is based on the pragmatic relations of topical material (clause-initial) vs focal material (clause-final). The ‘unmarked’ word order is S/A – DirO – IndO – V. Tracking of referents in discourse often involves the use of zero anaphora. Temporal and local phrases are most often clause-initial.

- (26) Nb̥jê-wu de-tê tê jî=i eo d̥ežu
mountain-head upward-take one day=GEN top particularly
mje=mje-mû mû=ly me=jê
many-many-make make=NMLZ.PNT NEG=exist.ABST
‘The entire day when [the body of the deceased] is transported up the mountains, [the ritual priest] does not have much to do.’

The verb complex is the only necessary element for an utterance to be considered a clause, and the verb complex may be simply a predicate noun.

6.1 Nominal predicates

Declarative equational sentences juxtapose a subject with a nominal predicate without an overt copula, as in (3). The copula verb /zî/ is used in negated equational sentences, as in (4). The use of /zî/ in declarative sentences conveys factual or assertive meaning.

- (27) [kû-^tê pimâ le] [cŷzâ] zî
this-that frog CTR deity COP
‘That frog was in fact a deity.’

6.2 Negatives and interrogatives

Negative particles include the negative /me/ (as in (7) or (26)) and the prohibitive /t^be/ (as in (25)). They are prefixed just before the main verb (or auxiliary), or after the directional prefix, if there is one.

The interrogative particle distinguishes between: (i) a free form, /jê/, which occurs as the last element of the sentence, and (ii) a bound form, /e/, which is prefixed just before the main verb (or auxiliary), or after the directional prefix, if there is one. The free form is used in content questions with nominal predicates (as in (28)) and in disjunctive questions (as in (29)). Disjunctive questions are formed by repeating the entire verb complex, first positive and then negative. The free form of the interrogative particle is inserted in between the two tokens.

- (28) jô â-bâ mê=d3o tê jê
self voc-father NEG=exist.ANM one Q
'Am I the one without a father?'

- (29) lă jê mê=læ nê tō
come Q NEG=come 2SG look
'See for yourself if you come or not.'

The bound form is used in polar questions.

- (30) nê jêqê ê=d3o
2SG child Q=exist.ANM
'Do you have children?'

The interrogative word appears in the same position as the noun phrase in the declarative form; compare the bold items in the two clauses in (31):

- (31) sê le dzi=gê? **jô-dzî** **dʒæmæ** dzi=gê
who CTR say=PROG.N-EGO self-family.GEN wife say=PROG.N-EGO
"Who says that?" "My wife says that."

6.3 Complex sentences

Juxtaposition is a typical strategy to combine clauses. It may convey various meanings, including condition (as in (7)), a sequential order of events (as in (12)) or contrastive sense:

- (32) tsêxe de-dzû ɻwẵ mē=dû dʒûmæ
pheasant upward-raise chicken NEG=be.complete fox
de-dzû teʰê mē=dû
upward-raise dog NEG=be.complete
'You may feed a pheasant, but it won't turn into a chicken; you may feed a fox, but it won't turn into a dog (proverb, A leopard cannot change its spots).'

6.3.1 Noun-modifying constructions

Lizu employs multiple strategies for nominalization, including:

- Nominalization with the genitive particle /i/, as in /ŋô-Ngô læmû ne-̃ukæ=i/ 'that covered with a copper pot' (see (1));
- Nominalization with one of the three nominalizers: (i) agentive /su/ 'the one who V' (from the generic noun /sû/ 'person'), e.g. /dza=sû/ 'the one who eats'; and (ii) two patient nominalizers: /ly/ 'the one to V', e.g. /dza=ly/ 'things to eat'; and /teæ/ 'the one being V' (from the resultative particle /teæ/), e.g. /gæ-gæ=teæ/ 'toy, lit. the one being played with';

- (c) Nominalization with definiteness markers (e.g. /dê-su=bi/ ‘the yellow one’, /Nt̪ʰə=bô/ ‘the smart ones’) and classifiers (e.g. /de-ne=kâ/ ‘the black one’).
(d) Nominalization and relativization constructions are essentially the same type of noun-noun constructions; see (1), (3) and (4) for examples.

6.3.2 Adverbial clauses

Adverbial clauses appear in a topical position. Subordinating morphemes combine native Lizu and borrowed Mandarin conjunctions. These include /kʰâ/ ‘when’ (as in (33)), /læ/ ‘and’ (as in (17) or (33)), /t̪əu/jiù 就 ‘just’, /tsɔ/ cái 才 ‘then and only then’.

- (33) \hat{a} læ kʰâ ae-dzî \hat{a} -jê dĕ
1SG come when 1SG-family.GEN VOC-elder.brother go.N-EGO
'When I came, my brother had already left.'

In conditional clauses, the protasis is marked by the contrastive marker /le/, and the hypothetical verb form is built by adding the change of state particle /v/ to the verb, as in (19). In counterfactual conditionals, the verb of the main clause is also marked by the change of state particle.

- (34) nê læ=gê hûsə=v le \hat{a} m̪=ji=v
2SG come=PROG.N-EGO know=CS CTR 1SG NEG=go=CS
'If I had known you would come, I would not have left.'

6.3.3 Complement-taking predicates

Complement clauses introduced by complement-taking predicates (Noonan 1985: 110–33) may be connected to the matrix clause hypotactically (as in (35)) or paratactically (as in (36)):

- (35) ae-dzî \hat{a} -jê Npse=su=bî Npse=gê
1SG-family.GEN VOC-elder.brother steal=NMLZ.AGT=DEF steal=PROG.N-EGO
kʰ-e-Ndo=ĕ
inward-see=CS
'My brother saw the thief stealing.'

- (36) \hat{a} kʰ-e-Ndo=ĕ zō dĕ
1SG inward-see=CS 3.N-VSB go.PST.N-EGO
'I saw him leave.'

Complements of verbs of speaking and thinking are introduced by the complementizer /dzî/ ‘say’, which may be further followed by the egophoric progressive auxiliary /bo/ or the non-egophoric progressive auxiliary /ge/. /dzî/ is not a completely grammaticalized element. It continues to be used as a verb of speaking /dzî/ ‘say’, as in (12). It may: (i) specify the exact author of the information (in combination with either /bo/ or /ge/, as in (9), or (31)); and (ii) be used as a reported marker to signal that the reported situation is based on someone else’s verbal account, so that the speaker is uncommitted to its truthfulness (only in combination with /ge/, as in (21)).

ABBREVIATIONS

Abbreviations follow the Leipzig Glossing Rules (LGR, www.eva.mpg.de/lingua/resources/glossing-rules.php). Non-standard abbreviations (those not included in the

LGR) are: ABST = abstract, AGT = agentive, ANM = animate, CMPR = comparative, CONT = continuative, CS = change of state, CTR = contrastive, EGO = egophoric, INF = inferred, MIR = mirative, PNT = patient, VSB = visible.

REFERENCES

- Bohnemeyer, Jürgen, Melissa Bowerman, and Penelope Brown. 2001. Cut and break clips. In: Stephen C. Levinson and N.J. Enfield (eds), *Manual for the Field Season 2001*, 90–96. Nijmegen: Max Planck Institute for Psycholinguistics. Available at: fieldmanuals.mpi.nl/volumes/2001/cut-and-break-clips/, accessed 12 August 2016.
- Bradley, David. 1997. Tibeto-Burman languages and classification. In: David Bradley (ed.), *Tibeto-Burman Languages of the Himalayas (Papers in Southeast Asian Linguistics no. 14)*, 1–71. Canberra: Department of Linguistics, Research School of Pacific and Asian Studies, Australian National University.
- Chirkova, Katia. 2008. Essential characteristics of Lizu, a Qiangic Language of Western Sichuan. Presented at the Workshop on Tibeto-Burman Languages of Sichuan, Institute of Linguistics, Academia Sinica, 21–22 November 2008, Taipei, Taiwan. Available at: www.katia-chirkova.info/resources/presentations/2008EssentialLizu.pdf, accessed 12 August 2016.
- Chirkova, Katia and Yiya Chen. 2013. Lizu: Illustrations of the IPA. *Journal of the International Phonetic Association* 43.1: 75–86.
- Chirkova, Katia and Zev Handel. 2013. Diachronic developments of voiceless nasals in Ersu, Lizu, and related languages. Presented at the 46th International Conference on Sino-Tibetan Languages and Linguistics, Dartmouth College, New Hampshire, 7–10 August. Available at: www.katia-chirkova.info/resources/presentations/2013DiachronicDevelopments.pdf, accessed 12 August 2016.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in cross-linguistic studies. *Language* 86.3: 663–687.
- Huang, Bufan and Renzeng Wangmu. 1991. Lüsuyu [The Lüsu language]. In: Dai Qingxia, Huang Bufan, Fu Ailan, Renzeng Wangmu and Liu Juhuang (eds), *Zang-Mianyu Shiwu Zhong / Fifteen Tibeto-Burman Languages*, 132–152. Beijing: Beijing Yanshan Chubanshe.
- Huang, Bufan, Xu Shouchun, Chen Jiaying and Wang Huiyin (eds). 1992. *Zang-Mian Yuzu Yuyan Cihui [A Tibeto-Burman Lexicon]*. Beijing: Zhongyang Minzu Daxue Chubanshe.
- Ikeda, Takumi. 2009. 200 basic words of the Lyuzu language (Naiqu dialect). In: Yasuhiko Nagano (ed.) *Progressive Report*, vol. 3, Project No. 16102001, Grant-in-Aid for Scientific Research (S).
- Lin, Youjing, Yin Weibing and Wang Zhi. 2014. Lüsuyu de zhudongci [Auxiliary verbs in Lizu]. *Minzu Yuwen [Minority Languages]* 2: 48–54.
- Moravcsik, Edith A. 1978. Reduplicative constructions. In: Joseph Harold Greenberg, Charles A. Ferguson and Andrew Moravcsik (eds), *Universals of Human Language. Volume 3: Word Structure*, 297–334. Stanford: Stanford University Press.
- Noonan, Michael. 1985. Complementation. In: Timothy Shopen (ed.), *Language Typology and Syntactic Description. Volume II: Complex Constructions*, 42–140. Cambridge: Cambridge University Press.
- Shibatani, Masayoshi and Prashant Pardeshi. 2002. The causative continuum. In: Masayoshi Shibatani (ed.), *The Grammar of Causation and Interpersonal Manipulation*, 85–126. Amsterdam and Philadelphia: John Benjamins.

- Sun, Hongkai. 1982. Ersu (Duoxu) hua jianjie [An outline of Ersu (Duoxu)]. *Yuyan Yanjiu [Linguistic Study]* 2: 241–264.
- Sun, Hongkai. 2001. Lun Zang-Mian yuzu zhong de Qiang yuzhi yuyan [On the Qiangic branch of the Tibeto-Burman language family]. *Language and Linguistics* 2.1: 157–181.
- Tournadre, Nicolas. 2008. Arguments against the concept ‘conjunct’/‘disjunct’ in Tibetan. In: Brigitte Huber, Marianne Volkart, Paul Widmer and Peter Schwieger (eds), *Chomolangma, Demawend und Kasbek: Festschrift für Roland Bielmeier zu seinem 65. Geburtstag*, Bank 1, 281–308. Halle (Saale): International Institute for Tibetan and Buddhist Studies.
- Tournadre, Nicolas and Randy J. LaPolla. 2014. Towards a new approach to evidentiality: Issues and directions for research. *Linguistics of the Tibeto-Burman Area* 37.2: 240–263.
- Wang, Dehe. 2010. *Ersu Zangzu Wenhua Yanjiu [Study of Ersu Tibetan Culture]*. Chengdu: Sichuan University Press.
- Wang, Yuqin. 2012. *Yalongjiang Zhong-Xia-You Diqu de Zuqun Hudong yu Rentong – Yi Jiulong “Liru” Zangzu wei Zhongxin de Kaocha [Ethnic Groups in the Middle and Lower Reaches of the Yalong River: Interaction and Identity; A Case Study of the Lizu Tibetans of Jiulong County]*. Beijing: Chubanshe.
- Yu, Dominic. 2012. *Proto-Ersuic*. PhD Dissertation: University of California, Berkeley.
- Zang-Mianyu Yuyin he Cihui Bianxie zu [Tibeto-Burman Phonology and Lexicon Editorial Group] (eds). 1991. *Zang-Mianyu Yuyin he Cihui [Tibeto-Burman Phonology and Lexicon]*. Beijing: Zhongguo Shehui Kexue Chubanshe.

§3.6 Naic [Naxi]

CHAPTER FORTY-THREE

YONGNING NA (MOSUO)

Liberty A. Lidz

1 INTRODUCTION¹

1.1 Position

Na is a Tibeto-Burman language of the Naic subgroup, which includes Laze, Na, and Naxi (Naish branch), as well as Namuyi and Xumi (Jacques and Michaud 2011; Chirkova and Chen 2013). Naxi has three main varieties, Baoshanzhou, Dayanzen, and Lijiang Naxi, and these are mutually intelligible. Na has three varieties, Ninglang / Beiqu, Yongning, and Guabie Na; these are not fully mutually intelligible (Yang 2009; He and Jiang 1985). Naic is near to Lolo(=Yi)-Burmese, but remains a distinct subgroup, with initials having undergone significant changes and syllable structures much reduced from their Proto-Tibeto-Burman (PTB) origins. On a continuum among Tibeto-Burman languages from phonological conservatism to phonological innovativeness, Naic languages are well on the side of phonological innovativeness.

Na is spoken in southwestern China, in northwestern Yunnan province and adjacent areas of Sichuan province, as well as a small region in southern Tibet, and has an estimated speakership of perhaps 40,000. The Na are adherents of Daba shamanism and Tibetan Buddhism, and some Na know Classical Tibetan. The data presented here are from the Yongning variety of Na, which is also sometimes referred to as Mosuo or Moso.

1.2 Contact

Na shows the influences of the three great linguistic areas which surround it: the mainland Southeast Asian linguistic area, the Sinitic linguistic area, and the Himalayan linguistic area. The influence of the mainland Southeast Asian linguistic area can be seen in shared grammaticalization processes, while Sinitic influence can be seen in compounding processes, tonality, and loanwords. Himalayan influence can be seen in the series of retroflex allophones, agentive marking, the evidential system, egophoricity, and borrowings from Tibetan.

Older Chinese loanwords in Na typically are phonologically reduced and likely were borrowed from other languages which were in closer contact with Chinese, rather than borrowed directly. Until the late 1990s, when new roads were built, the Na areas were quite isolated from the outside world, although the Na region is quite multi-ethnic, with Pumi, Han, Yi, Lisu, Tibetans, Naxi, and Bai also living in a number of the villages. Some villages continue to have relatively little direct contact with outsiders, but others have undergone dramatic changes. Elderly speakers typically are most comfortable speaking Na, middle-aged speakers usually are fluent in Na and Yunnanese or Sichuanese (the local

TABLE 43.1 INITIAL CONSONANTS IN NA

	Bilabial	Labiodental	Alveolar	Alveo-palatal	Retroflex	Palatal	Velar	Uvular	Glottal
Stop	p ^h p b		t ^h t d				k ^h k g	q ^h q G	
Trill							jn		
Nasal	m		n						
Fricative		f	s z	ç z		ç z	y		
Affricate			tʂ ts dz	tç̚ te dz		tʂ̚ tç̚ dz̚			
Lateral fricative			l						
Approximant w							y		
Lateral approximant			l						

varieties of Mandarin), while younger speakers are still fluent in Na but also speak one of the local varieties of Mandarin, as well as Standard Mandarin, which is the language of schooling, television, tourism, the government, and trade.

2 PHONOLOGY

2.1 Sound system

Na syllables contain four elements: initial (C), glide (G), vowel (V), and tone (T). The minimal syllable contains a vowel and tone, although a non-phonemic glottal stop or on-glide will typically appear preceding the vowel. Although the Lijiang and Weixi varieties of Naxi have a prenasalized series of consonants, Yongning Na does not. Syllables with the structure CGVT are not uncommon, but syllables with the structure CVT are more common. Na syllables do not have final consonants, although nasalization on the vowel may appear, particularly through rhinoglottophilia or in borrowed words, and some varieties of Na have, to a greater or lesser degree, a rhotacized vowel. Table 43.1 shows the initial consonants in Na.

Na has a three-way contrast (voiced, voiceless, aspirated) among both stops and affricates. /w/ is the only segment which can fill the glide position in a syllable. The inventory of initials in Yongning Na and Lijiang Naxi differ primarily in that Naxi has a prenasalized series of initials (/mb, ndz, nd, ndz̚, ndz, ñg/) and Yongning Na has a voiceless lateral fricative (/l̚/). The most notable characteristic of Na initials is that the series of bilabials and alveolars, the palatal nasal, and the velar fricative have allophones where the place of articulation has moved further back by one position. (Though note that the allophones of the bilabial stops are trills rather than labio-dentals.) Thus, the bilabial series /p^h, p, b, m, w/ has allophones [ɸ̚, ɸ, β, m̚, v̚]; the alveolar series /t^h, t, d, n, l/ has allophones [t̚, t̚, d̚, n̚, l̚]; and the palatal nasal /j/ has the allophone [ŋ̚]. The uvular phonemes /q^h, q, g/ are, as in Tibeto-Burman more generally, a recent development from the velar series /k^h, k, g/; the uvular phonemes appear in very few phonetic environments, but do contrast with /k^h, k, g/. The velar fricative /y/ has a uvular allophone [χ̚]. The conditioning environments for the allophones vary for the different consonant series, but are some set of low (/æ, a/), mid back (/ɤ/), or high back (/ɯ, u, y/) vowels or a diphthong (/wɤ, wae/). The affricate series /tʂ̚, tç̚, dz̚/ shows a tendency to have a phonetic realization where the stop component is retroflex, but this tendency is not systematic, so no retroflex allophones

TABLE 43.2 ALLOPHONES OF THE NA INITIALS

	Bilabial	Labio-dental	Alveolar	Alveo-palatal	Retroflex	Palatal	Velar	Uvular	Glottal
Stop					t ^h t̪ d̪				
Trill	b ^h , b̪, b̪								
Nasal	m̪				n̪		ŋ̪		
Fricative									k̪
Lateral					l̪				
approximant									

TABLE 43.3 NA VOWELS

	Front	Central	Back
High	i		ɯ, u, ɤ
Mid	ɛ	(ə)	ɔ, ɔ̄
Low	æ		a
Diphthongs		w̄ȳ, w̄ō, w̄ǣ	
Nasalized		ã̄, w̄ã̄	

(i.e. [tʂ^h, tʂ̪, dʐ̪]) are suggested for this series. The allophones of the Na initials are shown in Table 43.2.

Table 43.3 shows the Na vowels. Na has 14 vowels, of which ten are monophthongs and four are diphthongs.

One of the monophthongs and one of the diphthongs are nasalized. Other vowels have nasalized phonetic realizations, but in these cases, nasalization is due to rhinoglottophilia from the glottal initials /h/ and [?] rather than being phonemic. Among its vowels, Na has the high back ultra-closed, unrounded vowel /ȳ/, which is not uncommon among Tibeto-Burman languages of southwestern China. /ȳ/ is distinct from the other high back vowels, /u/ and /ū/, with spectrograms of /ȳ/ showing that the distance between the first and second formants is very compressed. Measurements of /ȳ/ indicate that its place of articulation is further back than that of /u/, much further back than that of /ū/, and generally similar in height to /ɔ̄/. [ə] is a reduced vowel subject to harmonization processes. In some cases, it derives historically from the low back vowel /a/—for example in the negative marker *ma*³³—yet in other cases, it is not clear what [ə] derives from, i.e. in the accomplished marker *la*³³.

Na has four surface tones: a mid tone 33, a high tone 55, a mid falling tone 31, and a low rising tone 13. These four surface tones can be further analyzed into a number of underlying tones based on the patterns in which the tones appear in combination (Michaud 2008). Tonogenesis in Na can be partially explained. The low rising tone very clearly appears in words which descend from PTB forms ending in the stops -p, -t, and -k; it also appears in Na reflexes of PTB forms with an *s- prefix where the prefix has been lost in Na, and in Na forms where a PTB nasal prefix has been lost preceding a velar initial. The mid falling tone appears in words which descend from PTB forms with an /a/ or /u/ vowel and a nasal final; however, other Na words which descend from other PTB forms with the same structure have mid tone, so an as yet unidentified conditioning factor must play a role, too. The mid tone appears in words which descend from PTB forms which have a

nasal, lateral, or glide final, or have no final at all. The conditioning environment which led to the development of the high tone remains somewhat unclear, in part due to the relative paucity of high tones in general and the subsequent small number of high tone Na forms for which a PTB reconstruction exists.

2.2 Phonological processes

Na has vowel harmony, which in Na is a regressive process where the affected vowel harmonizes to the vowel in the following syllable. Vowels primarily harmonize based on frontness / backness, but in a few instances, height is a factor. Vowel harmony occurs with a handful of affixes and function words: the vocative/kinship prefix a^{33} -, the question word prefix a^{31} -, the negative marker ma^{33} -, the accomplished prefix $lə^{33}$ -, the durative prefix t^hu^{33} -, and the conjunction la^{33} . Vowel harmony also occurs fairly productively in compounds, and can operate on free morphemes found in compounds. Vowel harmony does not cross the phonological word boundary, and only occasionally crosses the grammatical word boundary (for example, in expressions with the high frequency function word la^{33} ‘and’). Examples which illustrate harmony in Na include the kinship prefix a^{33} -, which appears as [ɑ] in $a^{33}-pʰv^{33}$ ‘grandfather’ and $a^{33}-da^{33}$ ‘father,’ but appears as [æ] in $a^{33}-mi^{33}$ ‘mother, aunt,’ as [ɛ] in $e^{33}-tei^{31}$ ‘aunt,’ and as [ə] in $a^{33}-su^{33}$ ‘great grandparents,’ $a^{33}-yu^{33}$ ‘uncle,’ and $a^{33}-dy^{33}$ ‘friend’; the accomplished prefix $lə^{33}$ -, which appears as [ə] in $lə^{33}-tɔ^{33}$ ‘see, look’ and $lə^{33}-po^{13}$ ‘take,’ but as [æ] in $ləe^{33}-qæ^{13}$ ‘burn up,’ $ləe^{33}-ba^{33}$ ‘open up,’ and $ləe^{33}-dzi^{33}$ ‘eat up,’ and as [ɛ] in $lə^{33}-su^{33}$ ‘die’ and $lə^{33}-su^{33} du^{33}-h̄i^{33}$ ‘thoughts’; the word na^{13} ‘eye’ which appears as [ɑ] in $na^{31}-na^{33}$ ‘pupil of the eye,’ but as [æ] in $næ^{31}-qʰæ^{33}$ ‘eye sand’ and $næ^{31}-tsi^{31}$ ‘eyes and eyebrows; eyelash’; and the conjunction la^{33} ‘and; etc.’ which appears with the vowel [æ³³] in the expression $yy^{33} la^{33} hæ^{33}$ ‘silver and gold,’ which is pronounced as one intonation unit and thus can be considered one phonological word.

3 WORD CLASSES

Nouns and verbs are the open word classes in Na; the closed word classes are classifiers, adjectives, numerals, and particles, which include postpositions, TAM markers, and evidentials.

3.1 Verb formation processes and verbal morphology

Verbs in Na have the morphological structure (Prefixes) + Root(s). Na does not have inflectional morphology for person, number, or gender. With the exception of the *aktionsart* prefixes $lə^{33}$ - and t^hu^{33} -, TAM is marked with post-verbal particles and/or reduplication processes. By far the most productive process for expanding the verbal lexicon in Na is compounding. Serial verb constructions, particularly those indicating the direction of a movement, are common.

Na verbal prefixes include ma^{33} - NEG, and the two *aktionsarten* $lə^{33}$ - ACCOMP and t^hu^{33} - DUR, which alter the verb’s lexical aspect. The accomplished marker $lə^{33}$ - appears on verbs which express a process, ongoing state, or liminality, to indicate that the process or state has in fact been attained. Examples of $lə^{33}$ - prefixation include:

$tɔ^{31}$	‘see’	$lə^{33}-tɔ^{31}$	‘saw, spotted’
ny^{33}	‘know’	$lə^{33}-ny^{33}$	‘found out’

ba³³ ‘open’ lə³³-ba³³ ‘open up’
pi³³ ‘be full’ lə³³-pi³³ ‘filled up’

In contrast, the durative marker *t^hu³³-* appears on verbs of process or liminality and marks that an achieved state endures. Examples of *t^hu³³-* include:

ku³¹ ‘hide’ t^hu³³-ku³¹ ‘hide (and stay hidden)’
k^hu¹³ ‘light’ t^hu³³-k^hu¹³ ‘light (and stay lit)’
di³³ ‘exist’ t^hu³³-di³³ ‘exist (enduringly)’

Na has the following types of verb compounds: [N v]_v and [v v]_v. Nouns can be incorporated, for instance:

my³³-dzæ³³ ‘is named’
name-call
hɪ³³-q^hγ³³ ‘invite guests’
person-invite
bɔ³¹-lu¹³ ‘raise pigs’
pig-raise
hæ³³-t^hu³³ ‘be windy’
wind-blow

Nouns can also be incorporated with stative verbs, for example:

my³³-dza³³ ‘be cloudy’
sky-bad

An example of a verb-verb compound is:

lɔ³¹-yɪ³³ ‘labor’ + zu³¹-yɪ³³ ‘work’ → lɔ³¹-yɪ³³-zu³¹-yɪ³³ ‘work’

In Na, some verbs can reduplicate to give a reading of reciprocal action (Yang 2009); this process is also common in Qiangic (LaPolla, personal communication). Some examples are: *tʂæ¹³* ‘grab’ → *tʂæ³³-tʂæ³¹* ‘fight over,’ *gwr¹³* ‘to circle’ → *gwr³¹-gwr¹³* ‘roam,’ *wr³³* ‘stack’ → *wr³³-wr³³* ‘stack together,’ and *st³³* ‘know’ → *st³³-st³³* ‘meet each other.’

The copula in Na is *ni³³*. It appears in equative constructions, but does not appear with stative verbs. Na, like some other Tibeto-Burman languages of southwestern China, has multiple existential verbs. The Na existential verbs include *dzo³³*, the generic existential verb, which can be used with animates, either human or animal, and inanimates. *di³³* EXIST.P is used with items which are perpendicular to a surface and typically appear in clusters, such as trees, villages, and feathers. It usually appears with *t^hu³³-* DUR, though this is a tendency based in semantics rather than a grammatical requirement. *ku³³* EXIST.T is an existential verb used to refer to the passage of time, and may be grammaticalized from a lexical verb relatively recently. *zu³³* EXIST.C is used with objects which appear within something else, and can be used with physical objects, as well as metaphorical objects, such as memorized scripture:

nu³¹-mi¹³ kwo³³ t^hu³³-zu³³
heart LOC DUR-EXIST.C
'in the heart'

Na has a sizable number of stative verbs, which appear in the predicate with a single (optional) nominal argument, or adjectivally, following the noun head, though the latter is less common.

3.2 Noun formation processes and nominal morphology

Nouns in Na typically consist of a root, or in the case of compounds, more than one root. There is also the kinship prefix *a³³-*, so common in Sino-Tibetan languages, as well as two nominalizers: *-h̄i³³*, from *h̄i³³* ‘person,’ which can be used for general nominalizations and agentive nominalizations, and *-di³³*, from *di³³* ‘earth, land,’ which can be used for locative nominalizations and purposive nominalizations.

a³³-da³³ ‘father’
a³³-ma³³ / æ³³-mi³³ ‘mother or maternal aunt’

lɔ³¹-yi³³-h̄i³³ ‘laborer’
labor-do-NOM.AGT

z̄wæ³³-dzæ³³-h̄i³³ ‘horse rider’
horse-ride-NOM.AGT

lɔ³³-zwx³³-h̄i³³ ‘words’
ACCOMP-say-NOM

di³³ ‘earth, land’

wy³³-di³³ ‘mountainous areas’
mountain-NOM.LOC

dzi³³-di³³ ‘things to eat’
eat-NOM.PURP

Other suffixes include *-mi³³* ‘mother,’ which appears in words for some females, either human or animal: *a³³-mi³³* ‘mother,’ *tʂʰu³³-mi³³* ‘wife,’ *ku³³-mi³³* ‘younger sister,’ *nu³¹-mi³³* ‘girl,’ *æ³¹-mi³³* ‘hen,’ *z̄wæ³³-mi³³* ‘mare,’ *bɔ³¹-mi³³* ‘sow,’ and names for women in Na mythology, such as *tsʰy³¹ hɔ⁵⁵ tse³³ tse³³ mi³³* ‘Tsuhodzedzemi.’ It is also an augmentative, as in the following:

tcʰɔ³¹-mi³³ ‘large ladle’
ladle-AUG

gi³³-na⁵⁵-mi³³ ‘black bear’
bear-black-AUG

khi³³-mi³³ ‘main entrance’
door-AUG

lɔ³¹-mi³³ ‘thumb’
finger-AUG

The suffix *zo³³* derives from a root which means ‘son; male,’ as in *zo³³-hɔ̡³³* ‘little boy.’ It appears as a suffix in names of males in Na mythology, as in *tsʰɔ̡³¹ du³³ lu³³ yi³³-zo³³* ‘Tsodeluyizo,’ but does not appear in male kinship terms in the way that *-mi³³* appears in a number of female kinship terms. It is also a fairly productive diminutive marker:

tʰɔ̡³³-dzu³³-zo³³ ‘pine tree seedling’
pine-CLS-DIM

yi³³-zo³³ ‘calf’
cow-DIM

teʰɔ̡³¹-zo³³ ‘small ladle’
ladle-DIM

By itself, *mo¹³* means ‘old,’ but it also appears as a suffix meaning ‘dear’ with an indication of respect, as in *hi³³-mo³³* ‘elder’ and *to³³-kʰu³³-mo³¹* ‘Tibetan mastiff.’

The plural enclitic =*æ³¹* pl marks a handful of animates, but is not very productive. Examples of the plural enclitic are:

pʰæ³¹ teʰi³³=æ³¹ ‘men’
man=PL

pʰæ³¹ teʰi³³-my³³ zo³³=æ³¹ ‘people’
man-young woman=PL

wɔ³¹ bo³³=æ³¹ ‘livestock’
animal=PL

su³¹ pʰi³³=æ³¹ ‘kings’
king=PL

a³³ pʰy³³-a³³ su³³=æ³¹ ‘ancestors’
grandfather-great grandparent=PL

zo³³ my³³-zo³³=æ³¹ ‘children’
child-boy=PL

Na has little derivational morphology; compounding is the most important morphological process. Noun compounding processes are: [N N]_N (either endocentric or coordinate), [N ADJ]_N, [N CLS]_N, [V N]_N (endocentric), and [N V]_N (exocentric). Examples of noun-noun compounds are *bɔ̡³¹-gɛ³³* ‘pork’ (pig-meat), *bɔ̡³¹-gu³³* ‘trough for pigs’ (pig-trough), *la³³-yu³¹* ‘tigerskin’ (tiger-skin), *mu³³-ei³¹* ‘flame’ (fire-tongue), and *zo³³-my³³* ‘child’ (son-daughter). Examples of noun-adjective compounds are *dzi³³-tsʰi¹³* ‘boiled water’ (water-hot), *mu³³-po³³* ‘blaze’ (fire-bright), and *na³¹-na³³* ‘pupil of the eye’ (eye-black). Examples of noun-classifier compounds are *su³³-du³³* ‘tree’ (wood-CLS), *mi³³-kʰwɔ³³* ‘scar’ (scar-CLS), *nr³¹-lv³³* ‘soybean’ (soy-CLS), *teʰi³³-lv³³* ‘mulberry’ (mulberry-CLS), *ha³³-lu³¹* ‘rice ball’ (rice-CLS), and *dzi³³-pʰæ¹³* ‘ice’ (water-CLS). Examples of noun-verb compounds are *na³¹-tɔ³³* ‘daylight’ (eye-see) and *zu³³-lu¹³* ‘earthquake’ (building-move).

The Na personal pronouns *na³³* and *nɔ³³* derive from PTB² **ŋa-y* ✕ **ka* ‘1ST PERSON PRONOUN / SELF’ and **na-ŋ* ‘2ND PERSON PRONOUN,’ respectively. The plural pronouns are formed by adding the suffix *-sŋ³³ ku³¹* to the singular forms. For the first person plural, Na has both a first person exclusive, formed by adding *-sŋ³³ ku³¹* to the first person singular *na³³*, and a first person inclusive, formed by adding *-sŋ³³ ku³¹* to the reflexive pronoun *ጀ³³*.

na ³³	1SG.PRO
no ³³	2SG.PRO
th <u>w</u> ³³	3SG.PRO
na ³³ -sy ³³ ku ³¹	1EXC.PRO
z ³¹ -sy ³³ ku ³¹	1INC.PRO
no ³³ -sy ³³ ku ³¹	2PL.PRO
th <u>w</u> ³³ -sy ³³ ku ³¹	3PL.PRO

Deictic pronouns in Na are the proximate/neutral demonstrative *t^hu*³³ (which is also the third person pronoun, a cross-linguistically common development), the distal demonstrative *du*³⁴, and the emphatic demonstrative *gr*⁵⁵. These deictics can be used to mark definiteness.

4 SYNTAX

4.1 Verb phrases

Table 43.4 shows the structure of the verb phrase in Na. The only required element of a verb phrase, and of a sentence more generally, is a verb head. Nominal arguments can, and frequently are, simply understood from the discourse context. Adverbs precede the verb head, and all other non-verb constituents in the verb phrase follow the verb head, including the intensifier *zwee*¹³. The evidential and epistemic particles appear towards the end of the verb phrase as these are relatively recent grammaticalizations, mostly from verbs, which have scope over the entire sentence.

Negation is marked with the prefix *mə³³- NEG*, which descends from PTB *ma NEG. It can prefix to active verbs, stative verbs, modals, auxiliaries, existential verbs, and the copula, and can appear in serial verb constructions (SVCs). Its vowel undergoes vowel harmony in some varieties of Na.

Na has three causative verbs, *k^buu¹³*, *tei³³*, and *yī³³*, which are all clear grammaticalizations, with *k^buu¹³* and *tei³³* also verbs meaning ‘put’ and *yī³³* a verb meaning ‘do.’ Matisoff (1973: 243) and LaPolla (2003: 33) postulate an *s prefix causative for PTB, but this has not yet been attested in Na.

SVCs are common in Na, with asymmetrical SVCs, where one of the verbs must be from a limited set of verbs, being prevalent. This limited set of verbs in Na is comprised

TABLE 43.4 STRUCTURE OF THE VERB PHRASE

ADV + V _{II} + CAUS + INTS +	$\left\{ \begin{array}{l} \text{TENSE} + \text{QUOT} + \text{CERT.STR} \\ \text{ASPECT} \qquad \text{pi}^{33} \qquad \text{ni}^{33} \\ \text{MODAL} \end{array} \right.$	$\left\{ \begin{array}{l} \text{REP} \\ \text{tsi}^{13} \end{array} \right.$	+ CERT.M mae ³³
zvæ ¹³	AUX		INFR p ^h æ ³³ -di ³³

TABLE 43.5 STRUCTURE OF THE NOUN PHRASE

N	+	REL	+	N_h	+	ADJ	+	INTS	+	DEM/NUM	+	CLS	+	CASE
---	---	-----	---	-------	---	-----	---	------	---	---------	---	-----	---	------

of four directional verbs: *yɔ³³* ‘venitive (come, where the speaker is the deictic center),’ *bi³³* ‘andative (go, where the speaker is the deictic center),’ *tsʰu³³* ‘come, where the speaker is not the deictic center,’ and *hu³³* ‘go, where the speaker is not the deictic center.’ The directional verbs usually, but not necessarily, follow the verb head. Na also has symmetrical SVCs, where the verbs are not restricted to a particular set. These tend to be either verb phrases where the nominal arguments are understood from the discourse context and the verbs thus concatenate, or resultative SVCs, which are another common SVC in Na. In resultative SVCs, the verb head comes first and is followed by the resultative verb, which can be either active or stative.

4.2 Noun phrases

Table 43.5 shows the structure of the noun phrase in Na. All elements are optional other than the noun head. Na has a rather elaborate system of classifiers, which classify nouns according to shape (for example, slice, strip, ball, kernel, stick); number (one person, more than one person, group, pair); measure (ladleful, bowlful, length of cloth); time (instance, period of time, day, night, year, generation); quantification (a little, some, many, all); round number (tens, hundreds, thousands); large plants (trees and other large plants); small plants (flowers, saplings, vegetables); domestic animals (oxen, cows, horses, yaks, pigs, dogs); birds and some other small animals (birds, chicken, cats, fish); and echo classifiers (in which the classifier is the same as the noun it modifies, for example: *tsɔ¹³ du³³ tsɔ¹³* ‘one room’ and *na³¹-wr³³ du³³ wɔ³³* ‘one Na village’). There is also a generic classifier *lu³³*, which can be used when a noun does not belong to any of the categories for which there is a classifier.

Quantification is indicated through a number of strategies, such as numerals, quantifiers, measure classifiers, the plural enclitic =*æ³¹* (though this enclitic is not productive), and plural pronouns. Like other languages of the Sinosphere, however, quantity frequently is simply understood through the discourse context. The Na numerals are given here; many of these can be clearly identified as reflexes of PTB reconstructions:

<i>du³³</i>	‘one’	
<i>ni³³</i>	‘two’	from PTB *g/s-ni-s TWO
<i>sɔ³³</i>	‘three’	from PTB *g-sum THREE
<i>zu³³</i>	‘four’	from PTB *b-ləy FOUR
<i>jnw³³</i>	‘five’	from PTB *l/b-ŋa FIVE
<i>qʰv¹³</i>	‘six’	from PTB *d-k-ruk SIX
<i>gur³³</i>	‘seven’	from PTB *s-ni-s SEVEN
<i>ho¹³</i>	‘eight’	from PTB *b-r-gyat ✕ *b-g-ryat EIGHT
<i>gv³³</i>	‘nine’	from PTB *d/s-kəw NINE
<i>tsʰe³³</i>	‘ten’	from PTB *ts(y)i(y) ✕ *tsyay TEN
<i>ci³³</i>	‘hundred’	
<i>tu³³</i>	‘thousand’	from PTB *s-tawŋ THOUSAND

The possession types in Na are alienable, inalienable, and inabsoluble. Alienable possession is marked with the possessive marker *bu³³*, and is used with kinship terms when the possessor is lexical or when the possessum is inanimate.

lu ³³ -ṣu ³¹	la ³³	na ¹³	bu ³³	a ³³ -p ^h y ³³ -a ³³ -su ³³	‘ancestors of Luoshui and the Na’
Luoshui	and	Na	POSS	ancestors	
ya ³³ q ^h a ³³		bu ³³	yæ ¹³		‘buckwheat seed’
buckwheat		POSS	seed		
zui ³³ -mi ³³		bu ³³	tsu ³³ tæ ³³		‘the hearth room’s foundation’
hearthroom		POSS	foundation		

Inalienable possession does not use an overt possessive marker; rather, the possessum simply follows the possessor. Inalienable possession is used with kinship terms when the possessor is pronominal or when referring to animal body parts.

ja ³³	a ³³ -wɔ ¹³	‘my family’
1SG.PRO	family	
t ^h u ³³	a ³³ -mi ³³	‘her mother’
3SG.PRO	mother	
zwæ ³³	wɔ ³³ -ly ³³	‘horse head’
horse	head	

Inabsoluble possession is indicated by just giving the possesum; the possessor is understood from the discourse context. It is used when the possesum is a human body part. In the following examples from narrative texts, the possessor is understood from the discourse context.

lo ³¹ -zui ³³	‘(her) fingers’
fingers	
p ^h æ ³³ -q ^h wæ ³¹	‘(his) face’
face	
ha ³³ py ³³	‘(her) hair’
hair	

As in many Tibeto-Burman languages, there is a distinct relationship among nominalization, relativization, and possession. The locative and purposive nominalizer, *-dt³³* (discussed in §3.2), acts as a relativizer in relative clause constructions. It also marks non-relative attributive constructions. In relative clauses, the relativized clause is marked with *-dt³³* and directly precedes the noun head:

teæ ³³ =æ ³¹	ku ³¹	di ³³	tsɔ ³³ mi ³³	‘wooden barrel for making pickles’
pickled vegetables=PL	make	REL	wooden barrel	
wɔ ³¹ bɔ ³³	ki ³³	ha ³³	ki ³³	di ³³
livestock	BEN	food	give	REL
				tsɔ ³³ -kwɔ ³³ du ³³ lʊ ³³
				kitchen one CLS

‘a kitchen to prepare food for the livestock’

The possessive marker *bu³³* marks associative constructions and non-relative attributive constructions. In associative constructions, the modifying element is marked with *bu³³* and precedes the noun head:

di ³³ -kwɔ ³³	bu ³³	my ³³ -zo ³³	‘mortal young woman’
earth-LOC	ASSOC	young woman	

a³¹-yi³³-ʂɛ³³ bu³³ hĩ³³ ‘people of long ago’
long, long ago ASSOC people

In non-relative attributive constructions, the modifying element again is marked with *bu*³³ and precedes the noun head:

tei³³ thæ³³-kwo³³ bu³³ bæ³¹ bæ¹³ la³³ pʰi³³ li³¹ la³³ wu³¹-dze³³ la³³
dirt under-LOC NRA flower and butterfly and bird etc.
‘Flowers, butterflies, and birds on the ground, and more’

Indefinite reference can be indicated by following the noun head with *du*³³ ‘one, a’ and a classifier, while definite reference is given by following the noun head with a demonstrative and classifier.

Na has a wide range of strategies for coordinating noun phrases. In all cases where a conjunctive coordinator appears, it is *la*³³. For conjunctive coordination, there are: asyndetic conjunction ('A, B, and C' with no formal marking of conjunction), monosyndetic coordination ('A, B, and C' where A and B are marked with the conjunction), representative coordination ('A, B, C, and such' where A, B, and C are all marked), coordinator omission ('A, B, and C' where B is marked), emphatic coordination ('A, B, and C all . . .' where A, B, and C are marked and the quantifier *du*³³-*ta*¹³ 'all' follows or only C is marked and the quantifier *du*³³-*ta*¹³ 'all' follows), and emphatic disjunction ('A and B do not V at all . . .' where A is either marked or not marked, the quantifier *du*³³-*sɔ*³³ 'at all' follows B, and the verb takes a negative marker), or a coordinate compound. The structure employed depends upon what the coordinated elements are—common nouns, nominalizations, enumerative expressions, or noun phrases—as well as what meaning is intended (in representative coordination, emphatic coordination, and emphatic disjunction). Disjunctive coordination ('or') is simpler. There are two main types of disjunctive coordination ('A or B'); in each case, A is marked. In the first type, the disjunctive coordinator is *mɔ*³³ *ni*³¹ (possibly deriving from *mɔ*³³-*ni*³³ NEG-COP), in the second type, the conjunctive coordinator is used; discourse context gives the correct reading.

Asyndetic coordination:

bu³³-dzae³³-hĩ³³ la³³-dzae³³-hĩ³³ zwa³³-dzae³³-hĩ³³ tʰw³³-ni¹³ du³³-wo³¹
yak-ride-NOM_{Agt} tiger-ride-NOM_{Agt} horse-ride-NOM_{Agt} this way a type
dzae³³-hĩ³³ tʰw³³-ta³³ la³³ ku¹³
ride-NOM_{Agt} DUR-come off ABLT
‘the yak rider, the tiger rider, the horse rider, and one who rides this way, able to (ride) coming off’

Monosyndetic coordination:

zɔ³³ la³³ my¹³
son and daughter
‘son and daughter’

Representative coordination:

ei³¹ dzi¹³ la³³ lɔ¹³ la³³
coals and ashes and
‘coals and ashes and (such)’

Coordinator omission:

ga³³ la³³ dæ³³ p^y³³ la³³ t^hui³³-te^hi³³ d^y³³ p^y³³
abbot lama and 3PL.PRO.FAM daba
'abbot, lama, or those dabas'

Emphatic coordination:

t^o³¹ mi¹³ bi³³ la³³ kw^y³³ bi³³ la³³ zu³³-mi³³ k^hi³³-mi³³ wo³³-q^hw^y³¹
pillar on top hearth on top hearthroom door on top

la³³ du³³-ta¹³
and all

'on top of the pillars, on top of the hearth, on top of the doorway to the hearth room, in all of these places'

Emphatic disjunction:

tei³¹ h^u³³ bi³³ mi³³ d^zi³³ ha³³ du³³-s^o³³ m^o³³-d^zo³³
clothes stomach eat food at all NEG-EXIST
'clothes, things to eat, none of these things'

Disjunctive coordination:

dæ³³ p^y³³ ki³³ m^o³³ ni³¹ d^y³³ p^y³³ ki³³ bi³³
lama ALL or daba ALL go
'go to (visit) a lama or a daba'

hw^a³¹ li³³ s^o³³ lu³³ la³³ zu³³ lu³³ t^hæ¹³
cats three CLS and four CLS such
'three or four cats'

4.3 Clauses

In pragmatically neutral contexts, Na has (A)(O)V basic word order. Agent/subject and object are not grammatically required to appear, and often are understood from the discourse context. Additionally, as Na is better analyzed as a topic–comment language than as a subject–predicate one, a topicalized object can appear preceding the agent. The Na topic can be a noun phrase, an independent clause, a dependent clause, or an adverbial, and is marked with *dzo*³³ TOP. The topic comes at the beginning of the sentence, which the comment then follows. In noun phrases, modifying elements such as adjectives, numerals, demonstrative, classifiers, and case markers follow the noun head. In verb phrases, modifying elements, with the exception of adverbs, follow the verb head. Agentive marking appears in clauses with non-basic word order or ellipsed patients, although the structural ambiguity of such clauses typically does not yield semantic ambiguity. Agentive marking in Na is often used to indicate agency, volitional focus, or switch in actor (Lidz 2011).

Declarative mood in Na is not explicitly marked morphosyntactically, though the final word in the sentence generally appears with falling tone. Imperative sentences show a tendency to employ an SVC with a directional verb (i.e. 'come and V' or 'go and V'), and there is a suppletive imperative form of the verb 'go,' *hɔ*³³ go.IMP. Prohibitive constructions are formed with *t'a*³³ PROH, which usually marks either the verb head or *hɔ*³³ go.IMP. *t'a*³³ PROH descends from PTB *da × *ta NEGATIVE IMPERATIVE.

In the following example, the SVC $pɔ^{31}$ - $yɔ^{33}$ ‘bring’ (where $pɔ^{31}$ means ‘take’ and $yɔ^{33}$ means ‘come’) gives an imperative reading.

bæ³³ pɔ³¹-yɔ³³. ‘Bring rope!’
rope bring.IMP

Here, a prohibitive construction is created with $tʰa^{33}$ PROH and $hʒ^{33}$ ‘go.IMP’.

gy³³ k^{hu¹³} ku³³ mə³³-ni³³ lə³³-ba³³ tʰa³³ hʒ³³ pi³³.
nine year EXIST.T unless ACCOMP-open PROH go.IMP QUOT
(He) said, “Unless nine years have passed, don’t open (it).”

Interrogatives are most commonly formed by placing the question marker a^{31} preceding the verb, or using question words such as a^{31} - $ts^hε^{33}$ ‘what, how,’ $q^hα^{33}$ - ni^{13} ‘how,’ ni^{13} ‘who,’ and $q^hα^{33}$ - $ts^hε^{33}$ ‘when.’ Interrogative mood can also be indicated by using rising intonation across the end of a sentence which would otherwise resemble a declarative. A-not-A questions are attested but may be due to contact with Sinitic. When the speaker anticipates a yes answer, the interrogative is formed by placing the question marker a^{31} after, rather than before, the verb.

A constellation of attributes show the effects of significant contact with Himalayan languages on Na: the series of retroflex allophones (discussed in section 2.1 earlier), agentive marking, an evidential system, egophoricity, and lexical borrowings from Tibetan. The evidential system is comprised of five parts: direct/visual evidence, which is not formally marked; the common knowledge evidential = a^{31} dzo^{33} ; the reported evidential tsi^{13} ; quotative evidential pi^{33} ; and the inferential evidential $pʰae^{33}$ - di^{33} (Lidz 2007). The latter three evidentials have clearly grammaticalized from verbs, which fits with LaPolla’s analysis of evidentiality as a recent phenomenon in Tibeto-Burman (LaPolla 2003). The system of egophoricity in Na allows only first person statements to be made categorically, while second person “statements” are phrased as questions, and third person statements are modified depending whether the verb is an internal state, observable state, or volitional verb (Lidz 2007). The motivation behind this system is that the speaker does not have direct access to another person’s thoughts, feelings, or conscious action, and thus must qualify such statements.

4.4 Clause combining

Combined clauses in Na fall along a continuum from a lesser degree of integration (coordination) to a greater degree of integration (subordination). Distinguishing between coordination and subordination in Na is not always straightforward because nominal arguments need not be overt when understood from the discourse context and TAM marking is not obligatory, so it can be difficult to distinguish an independent clause with nominal arguments that are understood from discourse context and a dependent clause.

Na has the following types of coordinative constructions: CONJUNCTIVE (‘and,’ indicated with asyndesis), DISJUNCTIVE (‘or,’ marked with $nɔ^{33}$), ADVERSATIVE (‘but,’ marked with dzo^{31} or asyndesis), CONCESSIVE (‘even though,’ marked with $nɔ^{31}$), THEN (‘then,’ marked with pe^{33} or la^{31}), PURPOSE (‘in order to,’ indicated with asyndesis), NEGATIVE PURPOSE (‘lest,’ indicated with asyndesis and the prohibitive marker $tʰa^{33}$ on the result to be avoided), REASON (‘because,’ indicated by asyndesis or $tʰi^{13}$), SUBSTITUTIVE (‘instead of,’

marked by *mə³³- NEG* on the first clause and asyndesis), and SEQUENTIAL (indicated with perfective marking on the first clause and asyndesis).

The following example shows conjunctive coordination through asyndesis.

t^bw³³ a³¹ wu⁵⁵ z^hwæ¹³ lɔ³¹-y¹³-z^hu³¹-y¹³³³ la³³ a³³-tsɔ³³-mə³³-ni³³ hu¹³ z^hwæ¹³.
3SG.PRO good-looking INTS work and whatever strong INTS
'She was very beautiful; (she) excelled in work and such.'

The subordinative constructions in Na are: JUST/UNTIL (marked with *nɔ³³*), THEN (marked with *pe³³*, and differs in usage from coordinative *pe³³*), and SIMULANEOUS ('while,' marked with *ts³¹* 'when' or *kwɔ³³* LOC).

Na has the following types of conditionals: present, habitual/generic, past, hypothetical, counterfactive, predictive, and negative conditionals. Realis conditionals (present, habitual/generic, and past) are not formally marked and the clauses join through asyndesis, while irrealis conditionals (hypothetical, counterfactive, and predictive) are formally marked. Negative conditionals code following their affirmative counterparts, but with the addition of taking negative marking on the A clause. The irrealis conditionals are marked in the following ways. Hypothetical conditionals are marked with *pi³³* or *pi³³ zɔ³¹* on the A clause, while counterfactive conditionals are marked with *pi³³ dzɔ³³* on the A clause, and predictive conditionals are marked with *pi³³ dzɔ³³* on the A clause and a future marker on the B clause. *pi³³*, which by itself is a hypothetical conditional marker, but also appears as part of the other hypothetical conditional marker *pi³³ zɔ³¹*, the counterfactive conditional marker *pi³³ zɔ³¹ dzɔ³³*, and the predictive conditional marker *pi³³ dzɔ³³*, may be grammaticalized from *pi³³* 'to say'; QUOT.

The following is an example of a predictive conditional.

t^bw³³ k^hwy³³ lu³³ pi³³ dzɔ³³ my³³ la³³ di³³ lə³³-tɔ³¹ pi¹³ ho³³.
this CLS till COND.PRED heaven and earth ACCOMP-turn over FUT
'If this piece (of land) is tilled, heaven and earth will switch places.'

Na has a number of complementation constructions. These include those in which the complement clause is sentence-like (utterance and immediate perception complementation constructions) and those in which the complement clause is not sentence-like (propositional attitude, pretence, commentative/factive, knowledge and acquisition of knowledge, fearing, desiderative, manipulative, modal, achievement, and phasal/aspectual complementation constructions).

The following shows an example of a complementation construction with an utterance complement taking predicate, *pi³³* 'say,' and a direct speech complement. Although most Na complements are not marked with a complementizer, extraposed direct speech and immediate perception complements can, but need not, take the complementizer *dzɔ³³*, which is also the topic marker.

wɔ³³ ta³³ ɔ³¹-sy³³ ku³¹ pi³³ dzɔ³³ lui³¹ dzɔ¹³ k^hu³¹ mə⁵⁵-ku³¹
before 1INC.PRO say COMPL feathers good nest NEG-make
k^hu³¹ z^hu³³ z^hy³¹ mə³³-dɔ³³.
skill disguise NEG-should
'Before, we said, "(With) good feathers, (one) doesn't build a nest; skill shouldn't be disguised.'"

5 CONCLUSION

Na has many features typical of Tibeto-Burman languages; some of these features are genetic, but many are either areal or due to Sapirian drift. These features include the *a³³*-kinship prefix; an augmentative suffix derived from ‘mother’ and a diminutive suffix from ‘son’; compounding as the primary morphological process; a complex system of classifiers; asyndetic coordination; a relationship among nominalization, relativization, and possession; serial verb constructions; topic–comment information structure; and agentive marking.

Phonologically, Na is quite innovative, in strong contrast with rGyalrongic, which is also spoken in Sichuan and is arguably the most conservative of the Tibeto-Burman language groups.

Na has a number of particularly distinctive and interesting features. These include a complex system of tones; a sizable set of allophones which are one place of articulation further back from their phonemes; vowel harmony; evidentiality; and an unusual system of egophoricity.

NOTES

- 1 This chapter is based upon my dissertation (Lidz 2010), the fieldwork for which was supported by the National Science Foundation under Doctoral Dissertation Improvement Grant No. 0345862 (PI Tony Woodbury), as well as a grant from the National Science Foundation Office of International Science and Engineering. I am extremely grateful for this support. I would like to thank Geze Dorje, the primary native speaker consultant for this research. His intelligence, seriousness of purpose, long hours of work, and good humor greatly contributed to this work. I would also like to thank Yang Zhenhong, Lamu Gatusa, Awu Daba, Dibi Daba, Mupha Daba, Yongzhutser Daba, Hua'er Zhao, Tseren Dorje, and Erqing.
- 2 All PTB forms cited in this chapter are from the *Sino-Tibetan Etymological Dictionary and Thesaurus*, online version: corpus.linguistics.berkeley.edu/~stedt-cgi/rootcanal.pl, accessed 11 January 2015.
- 3 Tonal changes between isolation and compounded forms are common in Na. In isolation form, *pɔ¹³* has the low rising tone 13 rather than the mid falling tone 31 that appears in compounds.

REFERENCES

- Chirkova, Katia and Yiya Chen. 2013. Xumi (part 1): Lower Xumi, the variety of the lower and middle reaches of the Shuiluo River. *Journal of the International Phonetic Association* 43.3: 363–379.
- He, Jiren (和即仁) and Jiang Zhuyi (姜竹仪). 1985. *A Brief Account of the Naxi Language*. Beijing: Minzu Publishing.
- Jacques, Guillaume and Alexis Michaud. 2011. Approaching the historical phonology of three highly eroded Sino-Tibetan languages. *Diachronica* 28.4: 468–498 and appendix.
- LaPolla, Randy J. 2003. Overview of Sino-Tibetan morphosyntax. In Graham Thurgood and Randy J. LaPolla (eds), *The Sino-Tibetan Languages*, 22–42. London: Routledge Language Family Series.
- Lidz, Liberty. 2007. Evidentiality in Yongning Na (Mosuo). *Linguistics of the Tibeto-Burman Area* 30.2: 45–87.

- Lidz, Liberty. 2010. A descriptive grammar of Yongning Na (Mosuo). PhD dissertation, University of Texas at Austin.
- Lidz, Liberty. 2011. Agentive marking in Yongning Na (Mosuo). *Linguistics of the Tibeto-Burman Area* 34.2: 49–72.
- Matisoff, James A. 1973. *The Grammar of Lahu*. Berkeley: University of California Press.
- Michaud, Alexis. 2008. Phonemic and tonal analysis of Yongning Na. *Cahiers de Linguistique Asie Orientale* 37.2: 159–196.
- Yang, Zhenhong (杨振洪), transl. Liberty Lidz. 2009. An overview of Mosuo. *Linguistics of the Tibeto-Burman Area* 32.2: 1–44.

§3.7 Lolo-Burmese

CHAPTER FORTY-FOUR

BURMESE¹

Julian K. Wheatley

1 INTRODUCTION

1.1 Affiliations

Burmese is the official language of the country of Myanmar, also called Burma.² It is spoken as a first language by the dominant Burman ethnic community and as a first or second language by the majority of the many minority groups living within or along the borders of the country (about 30 percent of the population). The 2014 census (Ministry of Immigration and Population 2015: iii) puts the country's population at 51.4 million. Standard Burmese evolved from a dialect spoken in the central Irrawaddy River valley, and as the formal means of exchange, it penetrated to most parts of the country. Minor variations in pronunciation, lexical choice, and usage distinguish the language as spoken in Upper Burma, the traditional pre-colonial cultural center, from that spoken in Lower Burma—the traditional homeland of the Mon and other groups—that was only intensively settled by Burmans in the colonial period.

Distinct dialects survive in regions that have historically been insulated from the centralizing forces of the state by mountain ranges or other geographical features. One such dialect is Arakanese spoken in the Arakan (Rakhine) region in the southwest of Burma, formerly home to a littoral state that remained independent until the late eighteenth century. Varieties of Arakanese are also spoken by groups of people across the western border of Burma, notably the Marma, who settled in the Chittagong Hills several centuries ago. Arakanese is known for preserving certain distinctions lost in the standard, but preserved in the (standard) spelling. Thus ‘chicken,’ spelled <krak> (ကြာက်) and pronounced *tee?* in standard Burmese, is *kro?* in Arakanese. Arakanese pronunciation accounts for the pronunciation “Rangoon” for Burma’s largest city, spelled <rankun> (ရန်ကုန်) in Burmese, but pronounced *jàŋgòuŋ* in the standard (and now usually written Yangon in English).

Another dialect grouping, Tavoyan, is spoken in the Tenasserim (Taninthayi) region that forms the southeast tail of Burma, particularly around the towns of Tavoy (Dawei), Palaw, and Mergui (Myeik, pronounced *mjei?* or *bei?* in Burmese). Tavoyan (like Intha, noted below) preserves a medial <-l-> that is attested in Old Burmese spelling, but represented by medial <-j-> or <-r-> in the modern orthography and pronounced as either a medial *-j-* or a palatal consonant in speech, e.g.: OBs <mli> ‘earth,’ spelled <mre> (မြေ) and now pronounced *mjè* in standard speech; OBs <klon(:)> ‘monastery,’ now spelled <kyon:> (ကျော်း) and pronounced *teáun*. Other regional dialect groups include Intha, spoken in the Inle Lake region of the Shan highlands, and Yaw, which is spoken in isolated regions between the confluence of the Irrawaddy and Chindwin rivers and the Chin

Hills. The regional dialects differ from the standard most noticeably in pronunciation, but also in some lexical and grammatical vocabulary (see Okell 1995).

None of the current varieties of Burmese preserves the full set of final consonants attested by the orthography (Written Burmese). In most cases, written stops are reduced to a glottal stop in speech and written nasals, to nasalization on the vowel. The exception is the Yaw dialect, which preserves final velar stops (oral and nasal). Otherwise, contemporary support for historic final stops comes from the layer of more distantly related languages such as Atsi (“Zaiwa” in China) and Achang, spoken in the far western part of Yunnan and neighboring parts of Burma.

1.2 History

The distribution of the Lolo-Burmese languages suggests that people speaking a Burmese prototype spread westwards from what is now southwestern China. Precisely when Burmese-speaking people appeared in the central plains of Burma is very uncertain, but it is likely that they encountered the people known as Pyu, speakers of a Tibeto-Burman (TB) language of uncertain subgrouping, who are associated with walled cities such as Sri Ksetra (Tharehkittara) and Halin (Hanlin, Halingyi) that flourished in what is now central Burma during the middle centuries of the first millennium. Profoundly influenced by Indian political and religious tradition, the Pyu left inscriptions written in a distinctive Indic script. Burmese itself is not attested until the late eleventh and early twelfth centuries, when it begins to appear on stone inscriptions.

One of the earliest Burmese inscriptions appears as one face of the quadrilateral Myinkaba Kubyaungyi inscription (also called the “Rājakumar” after the sponsor, or the “Myazedi” after the current site), dated 1113 CE. The other faces are versions of the same text (recording the donation of a gold Buddha image) in Pali, Mon, and Pyu. Pyu is likely to have had only ceremonial function by that time. But Mon was an important language of the first Burmese kingdom, and remained a presence—at times powerful—in Lower Burma until the Mon “state” was absorbed into the Burmese in the eighteenth century. Mon, genetically related to Cambodian and other languages in the Mon-Khmer family, continues to be spoken in the Moulmein region by perhaps as many as three quarters of a million people.

Stone inscriptions recording acts of merit and captions on Buddhist temple paintings are the main extant genres of Burmese literature up to the end of the Pagan period (1287 CE) and beyond. Manuscripts make their appearance later, with the writing scratched on specially prepared palm leaves bound flat with cord (“peisa”) or on a type of thick paper folded in zigzag fashion (“parabaik”). Much of the post-inscriptional literature is verse, often with complicated rhyming structures and special diction, and mostly dealing with king, court, or religious belief. Traditional prose ranges from religious to secular subjects such as law, history, and Pali grammar. Thai influence following the Burmese conquest of Ayutthaya in the mid-eighteenth century, gave rise to secular drama.

The first significant European presence in Burma was that of the Portuguese in the sixteenth century. They were followed by small numbers of Dutch, British, and French. Conflict with the British in India eventually led to the annexation of Lower Burma (1826, 1853) and finally, the removal of the monarchy and the complete loss of independence (1886), which was not to be regained until 1 January 1948. By the late nineteenth century, the introduction of printing and the influence of Western literary forms saw the rise of new secular genres. Newspapers appeared in the 1870s. The first novel appeared in 1904, a work said to have been inspired by *The Count of Monte Cristo*.

1.3 Transcription

Forms in this chapter are cited in broad phonetic transcription using International Phonetic Alphabet (IPA) conventions, following the precedent of Watkins (2005). The one exception is that in order to improve readability, nasalization on the syllable is indicated with a syllable-final Greek “eng,” [-ŋ].³ Other systems of transcription are also current. Okell, for his textbooks and other writings, uses a system that preserves word form, is easy to type, can be learned quickly and is simple to read. The Myanmar Language Commission, publisher of dictionaries and other reference works in Burma, uses a romanized system slightly different from Okell’s, as well as a phonetic system that gives regular, predictable values to a subset of Burmese letters. Thus, *pjóuη* ‘to smile,’ spelt unpredictably with anusvara (the superscript dot) and medial <-r->, <prum̥:> (ပူမ္မား), can be transcribed with regular spelling as <pyun:> (ပျော်း). Burmese personal names, as well as the names of shops and products are often romanized according to older, or more informal methods of transcription. These are sometimes based on English-language spelling conventions, e.g. “U Pyone” for the name ရှို့ ပြုဗော် (ရှို့ပြုဗော်း). In such transcriptions, the creaky tone is often indicated with a final “t”, as in the first two syllables of the historian Thant Myint-U (သန် မြတ် ရှုံး) (သန်မြတ်းရှုံးး).

In this chapter, Burmese orthography is cited in the transliteration recommended by Okell (1971), which follows Duroiselle (1916). Burmese script is provided in parentheses to speed recognition of forms for those who can read it, as well as to provide backup for the romanized representations.

2 CONTACT

From earliest times, Burmese have been in contact with non-Burmese ethnic groups whose writing and speech have left, and continue to leave, their mark on the language. Prominent among the early written influences were the classical Indian languages of Pali, the language of the Buddhist scriptures and, to a lesser extent, Sanskrit. The former has been the main source of specialized words for religious, philosophical, and other abstract lexical material. Pali (and Sanskrit) loans have given the Burmese lexicon a two-tiered structure, with verbs, most grammatical words, and non-specialized nouns being composed of monosyllabic morphemes from the TB stock, and learned vocabulary—mostly nouns—with polysyllabic roots of Indic stock.

Since Burmese script preserves symbols and conventions that allow it to represent Pali precisely, material can be imported unchanged from the classical language, e.g. <mettā> (မေတ္တာ) ‘kindness, affection’ and <kutī> (ကုတ္တာ) ‘latrine for monks,’ spelled identically in both Pali and Burmese. Similarly, the lower ordinal numbers are borrowed from Pali almost intact, though the first drops the Pali retroflex (cf. Pali <pathama>) from the spelling: <pathama, dutiya, tatiya> (ပုံစု၊ ဒုတိယ၊ တတိယ) ‘1st, 2nd, 3rd.’ In other cases, longer Pali words appear in truncated form in Burmese (a process that might have been abetted by the requirements of poetic scansion). This is especially common for disyllabic and polysyllabic words ending in short “a” in Pali. Burmese <cit> (ခိုစ်) ‘mood; thought’ from Pali <citta>; <tiracchān> (တိရှဇ္ဍာန်) ‘animal’ from Pali <tiracchāna>; and <maṇḍap> (မဏေပ်) ‘pavilion’ from Pali <maṇḍapa>. Pali continues to be a source of word formation, often in combination with native Burmese material, e.g. the formal term <patijīwa-che:> (ပဋိဂါဌာဆေး) ‘antibiotic [medicine]’: Pali <paṭī> ‘anti’ and <jīva> ‘life’ adjoined to Burmese <che:> *s'ē* ‘medicine.’ Other examples diverge farther from the Pali prototype: <paccāññ:> (ပစ္စည်း) ‘things, products, property’ from Pali <paccaya>; <takkasuil>

(တက္ကသိလ်) ‘university’ after Pali *<Taxila>* Taxila, the Gandharan city that served as a center for education. In some cases, spelling shifts may reflect transmission through a third language, such as Mon. Sanskrit loans are far fewer in number, e.g.: *<prāssād>* (ပြောဆုံ) ‘multitiered roof [characteristic of Burmese gateways],’ from Sanskrit *<prāsāda>*; *<samuddarā>* (သမ္မဒြရ) ‘ocean’ after Sanskrit *<samudrā>* (with spelling likely influenced by Pali *<samuddā>*).

Contributions from other languages are likely to have been transmitted primarily through the spoken language. Mon loans are extensive and date from early times: *<tuik> tai?* (တိုက်) ‘brick or stone building,’ from Mon *<tuik> tāk*; *<puwā> pəwā* (ပွဲ) ‘cloth, handkerchief,’ from Mon *<pawā> pəwa*, originally from Pali *<pavāra>* ‘woollen cloth’; *<arak> ရောင်* (အရောင်) ‘alcoholic drinks, liquor,’ from Mon *<arak> ərəak*, originally from Arabic, probably via Malay.⁴ In the colonial period and later, Burmese absorbed—and continues to absorb—a great deal of material from English, through both spoken and written channels: *bàηbà* (ဘန်ပဲ) ‘(car) bumper’; *pʰeʃin* (ဖော်ရင်) ‘fashion’; *dimokərèsi* (ဒီမိုက်ရေးစီ) ‘democracy’; *teiʃin* (ကျေရှင်) ‘tuition classes (after school)’; *teʰənpiʃən lei?* (ချုန်ပိယလိုင်) ‘champions league (football).’

3 PERIODIZATION

The ancient inscriptions reflect a language, Old Burmese, that differs quite extensively from the modern in phonological structure and in lexical content, particularly function words. Word order and grammatical categories have been more stable. Between Old Burmese and the modern language, it is useful to distinguish a Middle Burmese stage, represented loosely by an orthography which stabilized by the eighteenth century or sooner, and which survives today as modern Written Burmese. Pronunciation changed much more between the Middle and Modern periods than between the Old and Middle, with the result that literal values of the orthography differ rather drastically from spoken: *<lakphakre kram>* (လက်ဖက်ရောက်မဲ့) ‘plain tea,’ for example, is read *ləpʰɛ?jɛ.dzáñ*, with shift in vowel qualities and loss of final contrasts. However, sound change has been consistent enough that, with the exception of weak syllables and juncture voicing (the former in the first syllable of the ‘tea’ example, the latter in the last), pronunciation can be read off from spelling fairly predictably.

Major sound changes include the following: The four points of articulation of final consonants have been reduced to nasalization for nasals (represented as *-η*) and to glottal stop for oral stops (*-ʔ*). Vowels have split (sometimes with subsequent merger) according to syllable type, open versus closed, so that the spellings, *<i, in, it>* are read *i, eiŋ, ei?*, and *<u, un, ut>* read *u, ouŋ, ou?*, etc. Another important change, also revealed by comparing spelling to pronunciation, involves an ordered shift of initial consonants: *<s>* is read *θ*; *<c, ch>* are read *s, s^h*; and *<ky/kr, khy/khr, gy/gr>* are read *te, te^h, dz*. Thus, Burma’s best known person, spelled *<’oñ chan: cu kraññ>* (အောင်ဆန်းစုံကြည်), is pronounced *paun.s^háñ.sy.tei* (officially rendered—with title—as Daw Aung San Suu Kyi).

Just as phonological features of Middle Burmese survive in the modern orthography, archaic grammatical and lexical material survive in the various literary styles. While it is possible to represent spoken language in writing, and though some modern authors have made it a practice to write using colloquial diction and usage, most writing, and particularly most printed material, is written in Literary Burmese (LBs). Over the centuries, LBs has taken on more and more colloquial features, but it remains distinct in its use of literary grammatical forms, some of which are cognate with modern forms, some not. Thus the LBs equivalent of *teáuŋ.gò ၣဥwá.dé* (ကျောင်းကိုသွားတယ်) ‘[He]’s going to school’

(‘school-GOAL go-RLS’) would be *teáun.ðøθwá.di* (တော်မြေးသွေးသည့်), with both spoken particles replaced, the first with a non-cognate form, the second with a cognate. LBs forms are not always homologous with colloquial ones. The LBs particle pronounced *jwø* (ဂျိ), for example, subsumes functions of two colloquial particles, the causal *lø* (လွှာ) and the conjunctive *pí* (ပြီ); while the functions of LBs *θø* (သွေး) ‘towards,’ *há* (အား) ‘DAT,’ and *kò* (ကွဲ) ‘OBJ; goal’ are all subsumed by a single colloquial particle, *kò* (ကွဲ), seen with voiced initial in the above example. (For a recent discussion of written style, see U Saw Tun (2005).)

4 PHONOLOGY

It is useful to distinguish two types of syllable in Burmese: full, or major syllables and reduced, or minor ones. In the latter, phonological distinctions are confined to the initial consonant: the vowel is realized as schwa; and there are no tonal contrasts. Words may contain one or, less commonly, two minor syllables prefixed to a major: *løpʰε?* (ဝါၢန္ဂ) ‘tea’; *θàηbajø* ‘lemon’ (သံပျိုး); and *kəla?* (ကလတ်) ‘clutch’ (the last from English). In many cases, minor syllables clearly derive from full ones: *ηəpø* (ငါး) ‘fish paste’ (cf. *ηá* (ငါး) ‘fish’); *tʰəmənɛ* (ထေမန္ဒာ), a glutinous rice dish (cf. *tʰəmín* ‘[cooked] rice’).

4.1 Major syllables

Major syllables contain five syntagmatic positions: initial (C_i), medial (M), vowel (V), final (C_f) and tone (T), with vowel, tone, and initial present in all (major) syllables.

Burmese vowel values vary according to whether the syllable is open (-v) or closed (-η or -?). While it is possible to follow the writing system by identifying closed *au* with open ɔ (writing ɔ, ɔη, ɔ?), the reduction in symbols does not seem worth the loss of transparency. Here, as indicated in italics in Table 44.1, we use a broad transcription rather than a strictly phonemic one (with -η realized as nasalization on the vowel in citation syllables).

The four tonal distinctions of major syllables can be termed “creaky,” “low,” “high” (also called “heavy”), and “checked,” the last symbolized by -?. As the names suggest, tone in Burmese has a complex realization, which may include features of phonation, pitch and contour, length, and vowel quality. In general, creaky and checked can be grouped together as “constricted” and “short,” while high and low group together as “contoured” (in pitch) and “long.” However, the salience of these features varies according to context, with phrase internal position contrasting with phrase final (citation). In many cases, the creaky is distinguished from checked not by features of constriction or by pitch but by juncture features of vowel quality and the consonant assimilation characteristic of the latter (see below).

The creaky tone is less common than the others, which reflects its secondary origin (Thurgood 1981). Some compensation is provided by grammatically induced shifts of

TABLE 44.1 VOWELS OF MAJOR SYLLABLES

Syllable type	<i>i</i>	<i>e</i>	<i>ɛ</i>	<i>a</i>	<i>ɔ</i>	<i>o</i>	<i>u</i>
(open) - v	<i>i</i>	<i>e</i>	<i>ɛ</i>	<i>a</i>	<i>ɔ</i>	<i>o</i>	<i>u</i>
(closed) -η	<i>iη</i> [i]	<i>eiη</i> [ɛi]	<i>aη</i> [ã]	<i>aiη</i> [ãi]	<i>auη</i> [ãʊ]	<i>ouη</i> [õõ]	<i>uη</i> [õ]
(closed) -?	<i>i?</i> [i?]	<i>ei?</i> [ɛi?]	<i>ɛ?</i> [ɛ?]	<i>a?</i> [a?]	<i>ai?</i> [ai?]	<i>au?</i> [au?]	<i>ou?</i> [oo?]

low tone and, less often, high, to creaky. With vocatives, for example, creaky tone can indicate impatience (Okell 1969: 20): *θəmí je* (သမီးရေး) ‘hey, girl’ versus neutral *θəmí je* (သမီးရေး). With pronouns and other personal referents, the shift to creaky can, with or without (redundant) P_o marking, indicate grammatical relationships such as possession (or attribution) or object: *teənɔ̄.go* (အဲနှင့်တော်(ကို)) ‘me’ (‘1p^{OBJ}(.OBJ)’) versus *teənɔ̄* (အဲနှင့်တော်); *ηg t̄ù j̄e* (လို့တူရေး) ‘my boy’ (‘1p^{GEN} nephew voc’). Figure 44.1 gives the typical features of the tones.

Figure 44.2 shows the C_i arranged in three series, based on morphological considerations. The voiced series is less common than the aspirate or plain. Voiced stops are almost completely restricted to nouns, many of them loans from languages like Pali and English. Native Burmese nouns with voiced C_i often pair with verbs with aspirate C_i (and more rarely, with plain): *dzei?* ‘hook’ and *te^hei?* ‘to hook’ (both ချိတ်). The aspirate and plain series are associated with a vestigial morphological process that links, in regular pattern, about 50 pairs of verbs in which the causative member has an initial from the aspirate series, and the simplex, one from the plain: *te^hó* (ချိုး) ‘to break’ (aspirate initial), defined in the dictionary as *teó.zé.ðì* (အဲဗြို့စေသည့်) ‘cause to be broken’ (plain initial); *s^hou?* (ဆတ်) ‘to tear’ (aspirate), defined as *sou?.sé.ðì* (စွတ်စေသည့်) ‘cause to be torn’ (plain). Pairs such as *θei?* (ဆိုံး) ‘sleep’ and *θei?* (ဆိုပ်) ‘put to bed’ reveal a likely origin in s-prefixation (*s > θ); pairs like *j_z* (ခေါ်ပုံ) ‘be slack’ and *f_z* (ခေါ်ဖူး) ‘slacken’ support alignment of /s/ and /j/ in the table of initials, reflected in the convention followed by some transcriptions—including the Burmese phonetic script tradition—of writing /s/ as <hy> (သွေး) to align with <jy> /j/ (သွေး).

Items in parentheses, such as (w) are rare: r is found in Pali and other loanwords (such as *tarei?s^hāñ* (တံရာဏ်) ‘animal,’ from a Pali root, and *bərásijà* (ဘရာစိယာ) ‘bra,’ from English); and but for a very few cases (e.g. phrases involving the conjunction ðø such as *ðø.bèmè* (ထိပေမဲ) ‘but, however’), ð is a result of word or phrase internal sandhi. The aspirate series, it should be noted, includes not only heavily aspirated oral stops, but a typologically rare aspirated fricative s^h- (more consistently distinguished from plain s- in the speech of older speakers), as well as a set of pre-aspirated nasals (m̥ = [mm] etc.) and a voiceless lateral fricative, f̥).

The only medials are -j- and -w-. The subscript <-r-> attested by the orthography (like the subscript <-l-> of Old Burmese orthography) is reflected in the modern language as -j- or as a palatal affricate (*te*, *te^h*, *dz*). Medial -j- patterns with C_l, appearing with the labial stops (*p^hjù*, *mjó*, etc.). A few words contain medial -j- after the laterals: *ljg* (ချဉ်) ‘slight,

creaky (y):	short; tense voicing becoming creaky; high and sharply falling
low (v̄):	long, with sharp drop in intensity; lax voicing; relatively low
high (v̄̄):	long, with sharp drop in intensity; lax voicing; relatively high
checked (-?):	very short; sharply falling

FIGURE 44.1 TONAL CONTRASTS (CITATION)

Series:								
Aspirate	<i>p^h</i>	<i>t^h</i>	<i>te^h</i>	<i>k^h</i>	<i>s^h</i>	<i>m̥</i>	<i>ŋ̥</i>	<i>j̥</i>
Plain	<i>p</i>	<i>t</i>	<i>te</i>	<i>k</i>	<i>s</i>	<i>θ</i>	<i>m</i>	<i>n</i>
Voiced	<i>b</i>	<i>d</i>	<i>dz</i>	<i>g</i>	<i>z</i>	<i>ð</i>		

FIGURE 44.2 INITIAL CONTRASTS

slender,’ *ljaʔsi?* (လျှိုစ်) ‘electricity’; but most of these have alternatives with *j-*, e.g. *jaʔsi?* for ‘electricity.’ Medial *-w-*, on the other hand, patterns with the rhymes, appearing only with non-back vowels in open syllables *-wi*, *-we*, *-we*, *-wa*, and a few others in closed syllables (e.g. *-wiŋ*, *-waŋ* *-waiŋ*, *-we?*). Historically, *w*-rhymes have been the source of the series of *u*-rhymes in closed syllables (<*phwat*>, read *pʰu?* (ပွဲ) ‘monitor lizard’), but there remains some regional variation in the modern reflexes, e.g. <*kwam*> ‘betel nut,’ read *kun* or *kwáŋ* (ကွမ်း).

4.2 Juncture

The phonetic realization of final consonants is conditioned by the degree of juncture between syllables. At least two extremes need to be recognized: open juncture, with minimal assimilation between syllables, and close juncture, with maximal assimilation. In citation form, or before a syllable in open juncture, the two *c_i*s are realized as nasalization on the vowel (-*η*) and final glottal stop (plus associated tonal features) respectively. Before a syllable in close juncture, however, the final nasal is realized at the place of articulation of the following initial: *páŋ* + *pwíŋ* (ပန်းပွဲ) ‘flower’ is pronounced [pámbwín]; and the final oral stop -*?-* is realized as a consonant from the plain series homorganic in both manner (stop, fricative, nasal, resonant, etc.) and place (labial, dental, etc.) with the following *c_i*: *θau?* + *té* (သောက်တယ်) ‘to drink (drink+RLS)’ is realized as [θaut.té], with plain dental stop; *me?* + *si* (မျက်စီ) ‘eye’ is read [mjɛssí], with plain fricative.

Close juncture also affects the form of the following *c_i*. Following a checked syllable (-*?*), the *C_i* retains its original manner, but following a smooth syllable (-*v*, -*η*), if the *c_i* is voiceable (i.e. if it is aspirate or plain and there is a voiced *c_i* corresponding to it), then it is realized voiced (manner varying with speed): contrast *daʔ.ə̥hí* (ဓာတ္ထား) ‘gasoline, petrol’ with *þóuŋ.zí* (အုန်းဆီ) ‘coconut oil,’ both containing *s̥í* ‘oil’; or *þei?kʰáŋ* (အိုင်ခုံး) ‘bedroom’ and *þe.gáŋ* (အိုင်ခန်း) ‘living room (guest room),’ both containing the major syllable of *þəkʰáŋ* (အေခန်း) ‘room.’⁵

Juncture also affects the combination of minor plus major syllables, but in such cases, the voicing process often does not extend to the aspirated *c_i* (Okell 1969: 15–17). Thus, *zəbwé* (ဇားပြီ) ‘table,’ composed of *sá* ‘eat’ + *pwé* ‘festival,’ shows internal voicing; but *θəpʰóuŋ* (ထွေးဖုံး) ‘[teeth] gums,’ composed of *θwá* ‘teeth’ and *pʰóuŋ* ‘cover,’ does not. Sometimes, as the example for ‘table’ shows, where the major syllable has a voiced stop, a process of voicing harmony may induce voicing on the initial of the weak syllable as well: *sá* + *pwé* is read *zəbwé*, with voiced initial.

Close juncture (with voicing), characteristic of nominal compounds, is less common in verbal compounds, and is a regular feature of certain phrasal combinations, such as ‘noun plus (unprefixed) stative verb’ (*seʔkiù.dú*) (စောင့်ခြား) ‘thick paper,’ with *tʰù* ‘thick.’ Most particles also follow in close juncture.

5 THE WRITING SYSTEM

The Burmese writing system has the features common to the Brahmic family of scripts. It is written left to right; spaces separate phrases rather than words; and punctuation is limited to the symbols | and ||, corresponding roughly to comma or semicolon and period (full-stop). It is an alpha-syllabary (or “abugida script”), with vowels written horizontally or vertically around a consonant: ခ <*mi*>, မ <*mu*>, ဧ <*me*>, ဃ <*mā*>, ။ <*mo*>, ဃ <*mui*>, etc. Consonants without vowel sign have an inherent vowel, transliterated as a

short <a> and pronounced on the creaky tone, *g*. Consonants are marked as final by the “killing stroke” ($\partial\theta\alpha$? ‘killer’), which has the effect of suppressing the intrinsic vowel: $\infty\theta g$ ‘keep trim; hone,’ $\infty\omega\theta\omega mg$ ‘female practitioner of an art or trade,’ but $\infty\dot{\xi}\theta\dot{\alpha}\eta$ ‘strong.’

The immediate prototype of the Burmese script was used to write Pali before it was adapted to the writing of both Burmese and Mon. For this reason, the script is usually labeled Burmese-Mon (or Mon-Burmese). In addition to writing Burmese (and Mon), the script is also used in Burma to write Pali (and occasionally, Sanskrit). In more recent times, it has been adapted to the writing of Shan, certain Karen languages, and a number of other languages spoken in Burma.

5.1 The consonant signs

The 33 consonant signs of Burmese are traditionally divided into six groups beginning with the velars, then progressing towards the front of the mouth to the labials and ending with a disparate group of sonorants. These are shown in Table 44.2. Boxed consonants are the regular finals that are typical of native material. In Indic loans, all consonants except for ၁၁ <gh> ၂၂ <jh> ၃၃ <d> ၅၅ <dh> can be final. In Old Burmese, even the “vowel support sign,” ၆၆ <> appears in final position (or, in some cases, as a subscript) to mark creaky tone.

5.2 The vowel signs

In adapting the Indic script to Burmese, a tone language, the early scribes matched the Indian short vowel signs, (inherent) <a>, and short <i> and <u> to the shortest (non-checked) Burmese tone, the creaky (top section of Table 44.3): ဗ <ma> *ma*, ဃ <mi> *mi*, စ <mu> *mu*. Written syllables closed with an oral or nasal stop are also written with the short vowel signs: ဗစ် <mat> *ma?*, ဃစ် <mit> *mei?*, စစ် <mut> *mou?*, ဗန် <man> *màñ*, ဃန် <min> *mèin*, စန် <mun> *mòun*. The long version of the Indic vowels (all of which contain an additional stroke) were used for the high- and low-toned open

TABLE 44.2 CONSONANTS

TABLE 44.3 WRITING OPEN RHYMES

TONE	FRONT VOWELS		CENTRAL VOWELS		BACK VOWELS	
	Initial	Diacritic	(Initial)	Diacritic	Initial	Diacritic
V	V	(V)	V	V	V	V
PRIMARY VOWELS						
creaky	၃	၇	၃၁	<i>	ၫ	
level	၅။	၇	၃၁	<ା>	ၫ	
high	၇။	၃၁။	<ି>	ၫ	၂	
				- ၃၁	<ା>	၂
				၃၁ ၃၁၁	<ାର>	၂၁
				၃၁ ၃၁၁	<ାର>	၂၁
				၃၁ ၃၁၁	<ାର>	၂၁
MID VOWELS						
High level	၇	၆၀	၆၃၁	<ୟ>	၇	
high	၆၀။	၆၃၁။	<ୟ>	၇		
creaky	၆၇	၆၃၇	<ୟ>	၇		
Low level	၅	၆၁၅	<ୟ>	၇	၆၇၁	၆၃၁၇
high	၇	၃၈	<ୟ>	၇	၆၇	၆၃၁၇
creaky	၇	၃၈	<ୟ>	၇	၆၇	၆၃၁၇

syllables: ၁၁ <mā> *mà* and ၁၁: <mā:> *má*; ၇ <mī> *mì* and ၇: <mī:> *mí*; ၂၁ <mū> *mù*, and ၂၁: <mū:> *mú*.

Since in the Indian prototype only the three primary vowels, <a>, <i>, and <u>, had long and short versions, tonal marking for the mid-vowels (mid-high and mid-low) had to use different conventions. In the Burmese system, the mid-high vowels in open syllables are inherently low-toned (lower section of Table 44.3): ၁၁ <mē> *mè*; ၇ <muī> *mò*; but their mid-low counterparts in open syllables are inherently high-toned: ၇ <lway> *lwé*, ၁၁ <mo> *mò*. The latter set are then specially marked for low tone, as follows: ၇၁ <lway> *lwè*, with a ‘killed-y’; and ၁၁ <mō> *mò*, with the additional “killing stroke” (which originated as a ‘killed-w,’ ၁).

In most native words, initial vowels are written as diacritics around the vowel support sign, ၁ (၃ <ା>, ၃ <ି>, ၁၁ <ୟ>, etc.); but in Indic and other loanwords (as well as in a few native words), initial vowels are written with special initial vowel signs, such as: ၇ (= ၁၁၁) <ୟ>, as in ၁၁၁၀၁၍ <erāwati> ‘Irawaddy [River]’; and ၁၁ (= ၁၁၁) <ୟ>, as in ၁၁၁၁ <ogut> ‘August.’ As Table 44.3 shows, these special initial vowel signs do not occur for all possible vowels.

Table 44.4 shows the restricted combinations of vowel and ending associated with closed rhymes in native TB vocabulary. Indic and other loans frequently show other combinations (which are accommodated to regular pronunciation). Tonal contrasts only appear on the smooth rhymes (open or final nasal). For all cases where tones are not inherent, they are (in the modern language, though not consistently so in the inscriptive period) marked as high-tone by *visarga* (<>), written as a final (၁) <ି> before the early thirteenth century); and as creaky tone by a subscript dot (<. >), written as a final or subscript <’> (၁) in pre-sixteenth-century writing). Thus: ၁၁ <se> ၇ୟ ‘die’; ၁၁: <se> ၇ୟ ‘small’; ၁၁၁ <se> ၇ୟ (in *jəθə̯* ‘hermit’); ၇၁ <cim> ၇ୟ ‘soak, steep’; ၇၁: <cim> ၇ୟ ‘ooze’; and ၇၁ <cim> ၇ୟ ‘green.’

TABLE 44.4 WRITING CLOSED RHYMES

	-ñ	-ññ/-ñ	-n	-m	-k	-c	-t	-p
a	အင်	အည်/အဉ်	အန်	အမဲ/အံ	အက်	အစ်	အတ်	အပ်
	ရါ့	ရါ/ရိ	ရာ့	ရာ့	ရေ့	ရါ?	ရာ?	ရာ?
i		အိန်	အိမ်				အိတ်	အိပ်
		ရော့	ရော့				ရော့?	ရော့?
u		အုန်	အုံ				အုတ်	အုပ်
		ရွှေ့	ရွှေ့				ရွှေ့?	ရွှေ့?
o	အောင်				အောက်			
	ရာ့				ရာ့?			
ui	အိင်				အိဂ်			
	ရာ့				ရာ့?			

5.3 Special conventions

In addition to the special initial vowel signs, a number of other signs appear primarily with Indic loanwords. These include the letters for the Indic retroflex and voiced aspirate series (see Table 44.2).

The convention of preserving the original spelling of Pali and other loanwords has also given rise to a number of irregular rhymes. These have combinations of vowel and final consonant that are not usual in Burmese native material. Thus: ခါတ် <dāt> ‘element’ (cf. Pali <dhātu>), pronounced *da?*, as if spelled ဒေဝ <dhat>; ခိုလ် <bui> ‘commander, leader’ (cf. Pali <bala>), pronounced *bò*, as if spelled ဗြို <bui>; ဥယျာဉ် <uyyāñ> ‘garden’ (cf. Pali <uyyāna>), read ဥယျာဉ်, as if spelled ဒ္ထယာဉ် <'uyañ>. Burmese scribes also followed the Indian practice of stacking geminate and homorganic consonants (see Okell 1994 for details): ဗုဒ္ဓ <buddha> ‘Buddha’ (Pali <Buddha>) rather than ဗုဒ္ဓဗ <regular rhyme ဗတ်ဒ <butda>>; မဏေလေး <mantale>: *máñdəlé* ‘Mandalay’ rather than မန်းထေလေး . This last example demonstrates the fact that the top member of a stack is neither “killed” nor marked for tone. In general, the second consonant of a cluster is pushed down to subscript position; but when -ñ is the first consonant, it is pushed up and keeps its killing mark: အိုင်းနေ့ <'angāne> ‘Tuesday’ (rather than အင်းငါးနေ့). In some cases, the lower member of a stack is abbreviated, or reoriented (see Okell 1994, “special stacks”).

The consonants သ ya, ရ ra, ဝ wa, and ဟ ha have medial versions: ယ, ယု, ယံ and ယုံ (most of them showing shifts or variants to accommodate other signs). Thus ချိုက် <khyak> *te^hɛ?*, မြန် <mran> *mjàñ*; မြေး <mwe> *mwé*; မှု <mha> *mq*; ရှု <rha> *ʃq*. Two medials may cluster on a single consonant sign: ယွှု <lhya> *tʃq* or *ʃq*.

Particularly in modern times, the script has developed conventions for the expression of foreign language material. Weak syllables can represent initial clusters, for example: ပလုပ်စတစ် *pəlaðsatí?* ‘plastic.’ Differences in the vowel qualities of open versus closed syllables can be exploited. For example, the combination of diphthong and final in “blouse”—foreign to Burmese—can be represented by choosing the closed rhyme *lau?* <lok> (for the /aʊ/ diphthong) and adding a second final, <s> (to give /aʊ-s/): ဘလောက်စ် <balok-s> *bəlau?*s. Similarly the first syllable of the brand “Top Choice” is written with an open syllable (for [ə]), the creaky tone (for short duration) and a final <p>: တော်ပျိုက် <to.pkhyuik> /tɔp.tə^hai?/. The diphthong in “choice” is regularly represented in Burmese by the *ai* vowel; cf. မီးပိုင် <mi:pwuñ> *mí.pwəin* ‘traffic light,’ a compound of ‘fire’ and ‘point’ (borrowed from English).

Of the more numerous abbreviations that appear in pre-modern texts, four have earned official status, appearing in modern dictionaries. These are, in effect, logograms. They are: ဂျိ *jwé* a literary subordinate marker (sometimes spelled out as ရှယ်. <ruy>); roughly equivalent to colloquial ပြုပါ *p̥i*; ဂု *nai*, a literary locative marker (sometimes spelled out နှိက <nhuik>; equivalent to မှာ *m̥à* in colloquial); ၏ *hi*, a genitive/realis declarative marker (equivalent to colloquial ခဲ့ *je* and သေွ် *-tè*); and င်း *ləgáun*, a coordinate marker and resumptive pronoun. င်း is sometimes written out in full form: လည်းကောင်း <lənñi-kóñ:>⁶ A set of rarely encountered letters used for the writing of Sanskrit are listed and discussed in Okell (1994: §34, 1971: 27). The numbers zero to nine are written as follows: ၁၀၁၃၄၅၆၇၈၉၁၂.

6 WORD CLASSES

6.1 Verbs

The verb repertoire is expanded by compounding rather than by derivation from other word classes or by borrowing. Disyllabic verbs can usually be analyzed as lexical compounds (two roots), even if one is or has become restricted in the modern language: *kù.nì* (ကူည့်) ‘help,’ from *kù* ‘to help’ and restricted *nì*; *squn.sau?* (စောင့်ရွှော၍), with *squn* ‘look after’ and rare *sau?*. (The two compounds appear in coordination in public appeals to “assist and look after” tourists: *kù.nì squn.sau?*.) Many such compounds are clearly pleonastic, in which a more literary stylistic effect is achieved by combining the root with a (usually rarer) synonym: *pjó.sʰò* (ပြောဆုံး) ‘to speak, utter,’ a pleonastic compound made up of roots that both mean ‘speak’; *tʰéin.θéin* (ထိန်းသိမ်း) ‘preserve, observe [precepts],’ with *tʰéin* ‘control, mind’ and *θéin* ‘keep, take possession of’; and *pjáun.láun jè.mó* (ပြောင်လျောင်ရပ်မော) composed of compounds meaning ‘mock’ and ‘laugh heartily (at).’ Verbs may also draw noun complements into lexical configurations of “verbs with tied nouns” (Okell 1969: 36): *þá ná* (အားနာ) ‘be embarrassed’ (‘strength hurt’); *ðábó t̥eg* (သဘောကျို) ‘to like’ (‘disposition fall’); *o̥t̥:wá wún.θá ~ wáŋ.θá* ‘be happy’ (‘stomach/bowels pleasant’).

6.1.1 Adjectives

A class of adjectival verbs can be distinguished from functive verbs. Both can act as predicatives but they behave differently as modifiers. While both may precede a N_h if mediated by a relative marker, adjectives more usually follow their N_h , either as derived ∂ -prefixed nouns, or without prefix and in close juncture: *ləp'εʔ.jè ɻatɛh̥o.p₂* (လက်ဖော်ရည့်အချိပ်) ‘lightly sweetened tea’ (*teh̥o* ‘sweet,’ *p₂* ‘light’); contrast *teh̥o.dɛ ləp'εʔ.jè* (ခြုံတလက်ဖော်ရည်) ‘tea that is sweet.’

6.1.2 Adverbs

While a few adverbs may resist a derivational analysis, most words that appear in the adverbial position in clause structure are transparently derived from verbs (or adjectival verbs) by processes of reduplication, prefixation, rhyme or chime, or combinations thereof: *sózó tʰq-* (စွာစွာဝါ-) ‘to get up early’ (reduplication of *só* ‘be early’); *?əlu? teʔ?-* (အလှုတ်ကျက်-) ‘learn by heart’ (prefixation of *lu?* ‘be free’); *bjáunbján pjó-* (ပြားဌားပြန်ပြား-) ‘to say with syllables switched around’ (*pjáun* ‘change’ [with the initial unpredictably voiced in the compound] and *pján* ‘return’); *bəjóun-bəjín*

tʰg- (ဗရှိုးဗရှင်းသဲ-) ‘get up [out of bed] in a panic’ (*bəjóuη-bəjíŋ* formed by *bə*-prefixation of the pleonastic *jóuη.jíŋ*, based on the verb *jóuη* ‘struggle,’ and *jíŋ* ‘be intimate’); *səni?* *təteq lou?* (စန်တကျလွယ်) ‘do systematically’ (from *səni?* *teq-* ‘be systematic’). The last example, in which the *tə*-prefixed verb keeps its complement noun, is fairly productive when the outcome is, as in the example, two iambic phrases. This is one of many examples of the Burmese predilection for ‘elaborate,’ rhythmically matched phrases (see Wheatley 2014 for details). Derived adverbials may become conventionalized enough to be considered true adverbs, but those that can retain complements, at least, suggest a grammatical process of subordination or embedding. In the following example, the adverb formed by reduplication (based on verb *teà* ‘be long [in time]’) retains its complement *bəlau?* ‘how’:

- (1) *Bəlau?* *teà.dzà nè.ʔóuη.mə.lé* ဘလောက်ကြာကြာနေအုံမလဲ။
How long.long stay.still.IRR-Q
'How long are [you] staying?'

6.2 Nouns

6.2.1 Nouns

Nouns are found as head of most kinds of clausal constituent other than the verb phrase. The class is enlarged through borrowing, and through derivational processes such as prefixation, reduplication, compounding (lexical and pleonastic), and combinations thereof, e.g. the elaborate pleonastic noun *ʔəlou?* *ʔəkàiŋ* (အလုပ်အကိုင်) ‘work,’ with roots *lou?* ‘to do’ and *kàiŋ* ‘to handle.’ Nominal compounds have more possible configurations than verbal: in addition to N+N, V+N, (N+V)+N, one finds a large class of exocentric compounds of the form N+V: *gəbai?* *nai?* (ခါးပိက္ခနိက်) ‘pickpocket’ (‘pouch-pick’); *θə?* *θa?* *lu?* (သက်သတ်လွတ်) ‘vegetarian’ (‘life killing free’).

6.2.2 Pronouns

Except for the informal first person and second person pronouns, *ŋà* (၏) and *nìn* (၏၏), which derive from ST pronominal prototypes, pronouns in Burmese are derived from nouns. The polite first person (male speaker) *teiñ.dɔ* ~ *teənɔ* (ကျွန်တော်), for example, is etymologically a compound of ‘slave+HON.’ (Its more informal version is *teənou?* (ကျွန်းပွဲ).) The second person polite (male speaker) is *k'əmjá* (ခံုံးများ), based on *k'hiŋ* (ခံုံး) ‘master’; the female version is *ʃiŋ* (၏၏)—cf. *ʔəfiŋ* (အရှင်) ‘master, owner.’ Pronominal functions are also frequently performed by kin terms, names, titles, or other nouns, such as *mei?* *s'hwè* (မိတ်ဆွေ) ‘friend’ or *s'əjá* (ဆရာ) ‘teacher.’

Standard pronouns with typical usage are shown in Figure 44.3. (For greater detail, see Cooke 1968; Okell 1969: 100–1; and Bradley 2005: 68.) The pronouns shown are ordered roughly along the dimension of social distance, with deferential (distant) first and familiar (intimate) last. Pairings of first person and second person cannot be safely inferred from the chart. As indicated, in addition to registering social distance, some pronouns also register the sex of the speaker (e.g. *teiñ.dɔ* ~ *teənɔ*, male speaker, versus *teəmg*, female). There are also pronouns for showing deference to people of special status, e.g. *dəbbi.dɔ* (တေပည့်တော်) ‘first person to monks (male speaker)’ (‘disciple.HON’) or *dəbi.dɔ* (female speaker) and *ʔəfiŋ.p'əjá* (အရှင်ဘုရား) ‘second person, to monks’ (‘lordship’).

1p	teənɔ̄ (ကျော်တော်) teəmə (ကျော်မ) teənou? / teou? (ကျော်း / ကျူး) kò (ကိုယ်) ηà (ခါ)	polite, male speaker polite, female speaker neutral colloquial, familiar familiar
2p	k ^b əmjmá (ခင်ဗျား) ʃiŋ (ရှင်) míŋ (မင်း) (ʔə)θiŋ ((အ)သင်) tò (တော်) ní (လည်း) niŋ (နှင့်)	polite, male speaker polite, female speaker slightly superior sounding rather august sounding intimate, among women familiar, among women familiar
3p	θù (စူး) PL + tq (ဝ္ား)	‘he,’ ‘she,’ but not ‘it’ teənɔ̄.dø, θù.dø, etc.

FIGURE 44.3 PRONOUNS

For reflexive and emphatic pronouns, we can note the following forms built on partially repetitive structures such as $\theta_y+[]+\theta_{\bar{y}}$ (သဲ+[]+သဲ) ‘him-/herself’ (‘3p^{GEN}+[]+3p’ or $k_0+[]+k_0$ (ကိုယ့်+[]+ကိုယ့်) ‘oneself’ (‘1p^{GEN}+[]+ACC’). The slot may be filled with a body term, as for example *pá* ‘cheek’ in the following:

- (2) *Kq.pá.kò θwá te^bg.bà.lá* ကိုယ့်ပါးကိုသွားချပါလား။
Self.cheek.self go strike.POL.Q
‘Why don’t you go and slap yourself on the cheek?’

The inserted element may also be *kò* (ကိုယ်) ‘body, self,’ giving a phrase in which all three elements are etymologically related:

- (3) *Kq.kò.kò θj.bà* ကိုယ့်ကိုယ်ကိုသိပါ။
Self-body-self know.POL ‘Know yourself!’

Insertion of *hà* (သာ) ‘thing’ (or *p^bà* or *p^bàdà* (သာသာ) ‘on one’s own’) forms emphatic pronouns:

- (4) *Kq.hà.kò.bé lou?.pà.yg.zè* ကိုယ့်ဟာကိုပဲလုပ်ပါရမေး။
Myself.only do.POL.must.CAUS ‘May I do it on my own?’

The pronominal registration of age, sex, and rank is supported by the use of “final tags,” some of which are similar or even identical to pronominal forms:

- (5) *Hou?.pà.dé ſiŋ* ဟာတပါတယ်ရှင်း။
Be-the-case.POL.RLS TAG ‘That’s right, sir/ma’am.’ [woman speaker]

- (6) *tej.jòuŋ tej.dà.bà kwà* ကြည်ရုံကြည်တာပါကွာ။
Look.only look.RLS^{NOMZR}.POL TAG ‘Just looking, dude.’

A subclass of pronouns are deictics, which include the proximate *dì* (ဒဲ) ‘here; this’ and the distal *hò* (ဒဲ) ‘there; that,’ plus *dà* (ဒဲ) ‘this one, that one.’ The three also occur with the ejaculation *hó* (ဟော) ‘oh, aha’ prefixed (*hódi* ‘this’ etc.) or with the ejaculation *ye* (အဲ) ‘yes, yeah’ (*yédi* ‘this,’ etc.). The former are often exophoric (pointing), the latter anaphoric (textual).

6.2.3 Noun classifiers

Burmese, like so many other languages in the region, makes use of noun classifiers to count nouns or to substitute for them: e.g. *pá* (ပာ) for monks, *jau?* (ယောက်) for (ordinary) people, *káun* (ကောင်) for animals, *t'è* (ဝါ၍၌) for clothes, *s'hù* (ဆူ) for sacred objects (e.g. pagodas), *lóun* (လုံး) for round things. Choice of classifier may also reflect different aspects of the referent: *lóunjái tøt'hé* (လုံချည်တထည်) ‘sarong one-cloth’ or *tøkwín* (တကွင်း) ‘one-circle.’ In a few cases, the classifier echoes the noun: *?ein tø. ?ein* (အိမ်တအိမ်) ‘house one-house.’ Nouns may also select quantifying measures: *yøpj. ði tø p'hí* (ပုဂ္ဂိုလ်ပူသီးတဖိုး) ‘a (small) bunch of bananas.’ Nominal forms of classifiers appear in a special construction for counting items in multiples of ten (10, 20, 100, etc.): *p'hóun.dzí kó.bá* (ဘုန်းဒို့ကိုးပါး) ‘nine monks’ (‘monks 9-CL’) but *p'hóun.dzí ?apá ñø.s'hé* (ဘုန်းဒိုးအပါးဆူ၍၌) ‘20 monks’ (‘monks 20CL 2.10’).

6.2.4 Other noun subclasses

Associated with nouns, there are a number of subclasses of words with grammatical and semantic specialization. These include plural markers. Pronouns are regularly made plural by the addition of *tq* (ଓঁ): *teənou?.tq* (ତେନୁହିଲ୍ଲାଟି) ‘we, us.’ The same word appended to nouns, however, means ‘and the like, and the others’: *sʰi.mi.dq* (ସିମିଟାଟି) ‘oil lamps and so on’; *mg só só.dq* (ମୋହିସିଟି) ‘Ma So So and company.’ The usual plural for countable nouns is *-tè* (ୱେଁ), or in careful speech, *-twè*, as spelled: *mí.bóuŋ.dè* (ମିବୁଙ୍ଗଦେ) ‘lanterns.’ With mass nouns, *tè* signifies ‘a large amount of’: *seʔkù.dè.n̥e* (ସେକୁଦେନ୍ଥେ) ‘with lots of paper.’ In more formal contexts, *mjá* (ଯପି), derived from a verb meaning ‘be many’ (and also occurring as a postposition meaning ‘or something’), may also be used as a plural marker.

Other suffixial elements worth noting are the ‘female’ suffix, *ma* (ω) (cf. *ʔəməg* ‘female; main’), the augmentative suffix, *teí* ‘big,’ and the diminutive, (*kʰə*)*lé*. Examples: *sʰəjá.ma* (ဆရမ) ‘female teacher’; *sʰəjá.dzí*, often pronounced *sʰəjá.ʔi*, (ဆရ္တိ) ‘eminent teacher’; *sʰəjá.gəlé* (ဆရကေလေ) ‘little teacher (addressing a nun).’

6.2.5 Location nouns, subordinate nouns

Another set of nouns specify location: *?ein.sə.n̩à* (အိုင်ရွှေမာ) ‘in front of the house’ (*‘house.front.LOC’*); contrast *?ein ?əsə* ‘the front of the house,’ in which *?əsə* is the *N_h*. One of these location nouns, *tʰe?* ‘above,’ has become specialized as a comparative marker:

- (7) *Lù bəwq̥.hà kʰwé bəwq̥.de?* káuŋ.ðə.lá
Man life.s dog life.than be-good.RLS.Q
‘Is a man’s life better than a dog’s?’

There is also a class of nouns that serves to form subordinate clauses: *ɻək'à* 'time,' in *jauʔ.kʰà* (ເຈົ້າກໍາ) ('arrive.when') 'when [we] arrived'; *ɻəsə* 'start' in *jauʔ.sg* (ເຈົ້າກົດ) 'since arriving.' These nouns may also act as heads for RCs:

- (8) *Mə.θ̥l.ðə haŋ s^hàuη.nè.d̥e* မသိသဟန် အဆင်နေတယ်။
 NEG.know.RLS appearance adopt.stay.RLS ‘[They] are pretending not to know.’

7 SYNTAX

7.1 The verb complex

Minimally, the verb complex consists of a head followed (in most cases) by one of a disparate set of main (see Table 44.5) or subordinate clause particles (§7.5). In most cases,

other elements are also present. Verbs may appear before the v_h to form concatenations (“serialized verbs”):

- (9) $\theta wá kʰa?$ ^{pá} သွားခံပါ။
go ladle.POL ‘Go and draw some [water].’
- (10) $P^h_θjáuηdàiη t^h_ɛ t^h_úη.dè$ ဖေယောင်းတိုင်ထည့်ထွန်းတယ်။
candles insert light.RLS [They] stick candles [in] and light [them].
- (11) $jó sá.jìη . . .$ ရောစားရင် . . .
mix eat.if ‘if [you] eat [them] mixed [with other fruit] . . .’

Concatenations can often be resolved into temporally consecutive clauses by the insertion of the subordinating particle $pí$ (ပါး) ‘and, having v’d.’ For instance, the last example could be broken into two clauses: $jó pí sá.jìη (ရောပြီး စားရင်)$ ‘if you eat [them] after mixing . . .’

Auxiliary verbs may appear in unmediated strings directly after the main verb. They show a range of grammatical properties and semantic specialization along a cline of abstractness. Those whose relative position is close to the v_h show greater independence, often combining in open juncture, for example, and being able to be directly preceded by the negative prefix or a complement marker; while those relatively late in the phrase are closely bound, and inseparable.

- (12) $Mə.pjìn.pé.nàiη.bú$ မပုံင်ပေးစိုင်ဘူး။
NEG.repair give.able.NEG ‘[I] can’t fix [it] for [you].’

In the above phrase only $pé$ allows direct negation, not $nàiη$: $Pjìn mə.pé.nàiη.bú$.

Auxiliaries include verbs like $nè$ ‘stay,’ as in $Là.nè.bì$ (လာနေပုံး) ‘Here [he] comes!’ (‘come stay-PUNC’); or $t'á$ ‘put’ and $pé$ ‘give’ in $K^h_θng tei.t'á.pé.bà$ (ခုက္ခည် ထားပေးပါး) ‘Watch [it] for [me] for a while, please’ (‘awhile watch.put.give.POL’). Some auxiliaries, such as $pé$ ‘give’ (benefactive) in the previous example and $kʰáin$ (ခိုင်း) ‘ask; tell’ (causative) in the next, add arguments to the clause:

- (13) $θapʰwá.dzí.gò$ ဒေါ်မှု ဂုံး ဝါး အသွားကြီးကိုအုန်းသီးခြုံပခိုင်းရတယ်။
old-woman.big.OBJ coconut scrape show ask.get.RLS
‘[I] had to ask the old woman to show [me] how to scrape coconuts.’

At the more abstract end of the phrase are particles such as $pʰú$ (ဖူး) ‘have ever’ (etymologically related to the homophonous verb meaning ‘visit, behold’⁴) as in:

- (14) $Dəpj.i.dzò mə.sá.bú.bú.lá$ ငါးပိုင်းကြုံ မစားဖူးဘူးလား။
Shrimp.paste.fried NEG.eat.ever.NEG.Q
‘Haven’t [you] ever eaten fried shrimp paste?’

Another particle, $kʰɛ$ (ခဲ့), conveys displacement in space or time:

- (15) $Leθsʰáuη.gò jù.lá.ge.mè$ အက်နောင်ကိုယဲလာခဲ့မယ်။
Present.OBJ carry.back.IRR ‘[I]’ll bring [you] back a present [from there].’

$K^h_ɛ$ also appears in subordinate clauses conveying counter-factual conditions: $θàuηmjiη.gɛ.jìη$ (ဆောင်မြင်ခဲ့ရင်) ‘if [she] had succeeded’ (‘succeed- $K^h_ɛ$ -if’). $K^h_ɛ$ is paradigmatically paired with $lai?$ (ချိုက်) by Allott (1965: 296). As a full verb, $lai?$ means ‘follow.’ As a particle its prototypical meaning is ‘action away,’ but it may also suggest abrupt or effective action: $Jé.pé.lai?.pá$ (ရေးပေးလိုက်ပါး) ‘Would [you] mind just writing [it] down [for me]?’ (‘write.give.lai?.POL’). Another paradigmatically related set includes the aspectual

particles, *θé* (၁၁။) and *ʔóun* (၃၂။), both meaning ‘still, not yet (i.e. no change),’ the first used only in realis clauses and the second, only in unrealis; and *t₂* (၁၀၃။) ‘about to, anymore (change of state,’ used in both realis and unrealis clauses. The last is difficult to distinguish in function from the ‘punctative’ main clause marker, *pi* (၅။) (see Table 44.5). The following five short examples illustrate the differences between these aspectual markers:

- (16) *sá.nè.θé.dè* (စားနေသေးတယ်) ‘still eating’; (RLS)
sá.ʔóun.mè (စားအဲးမယ်) ‘will eat’; (IRR)
sá.d₂.dè (စားတော့တယ်) ‘finally ate’; (RLS)
sá.d₂.mè (စားတော့မယ်) ‘about to eat’; (IRR)
sá.bi (စားပို) ‘am eating (now, at last); have eaten’; (PUNC)

7.2 The noun phrase

The following sentence (taken from a recorded text) is cited to illustrate some of the features of the noun phase:

- (17) *ʔédi sà.gò ʔəlúŋ mjá.mjá pjàŋ.nàiŋ.* these texts.OBJ very many recite.able.
d₂e pʰóun.d₃.dzi.mjá.gò RLS_{ATTR} monks.HON.great.PL.OBJ
lu.dè.gá.lé.bé people.PL.S.also.EMPH
ʔin̥matàn tɕʰí.tɕ'ú.dzaŋ.bà.dè very praise.COLL.POL.RLS
အဒီစာကို အပွဲန်း မူးမျှးမျှး ပြန်ဝိုင်တယ့် ဘုန်းတော် ရွှေးများကို လူတွေ့ကလည်းပဲ
အင်မတန် ချိုးကျူးကြပါတယ်။ ‘Monks who can recite back large parts of these texts, the people praise them highly.’

In the NP, modifiers, for the most part, precede the *N_h*. In the example, these include demonstratives (*ʔédi*) and RCS (marked by *-d₂e*); SVs follow (*dzi*), as do classifier phrases, noun suffixes (*mjá, dè*), and classes of grammaticalized nouns. General P_o’s (*lé, bé*) follow case-marking P_o’s, and occupy final position in the NP. Several of the most common case marking P_o’s show “local” and “non-local” functions: *-gg*, marking subject in our example, also marks source and certain kinds of past-time temporal phrases. It also appears as a general P_o associated with “topic.” While one is tempted to try to find a single notion such as “source” that will subsume these functions, a careful study of the distribution of *gg* requires positing homophonous particles with several senses. The P_o *kò*, with

<i>hà</i> (၁၃)	subject; common in stories, often followed by a pause
<i>kq</i> (၁၂)	source, contrastive subject, at (past time), attributive
<i>kò</i> (၁၇)	goal, object, extent, at (future time)
<i>nɛ</i> (၁၄)	manner (with adverbs), instrument, accompaniment (“with”)
<i>mà</i> (၁၅)	locative
<i>tin̥/twin̥</i> (၁၁၃/၁၂၃)	locative (“within”)
<i>jɛ</i> (၁၆)	genitive (<i>kɛ</i> (၁၇) after checked tone)
<i>mg(lq)</i> (၁၅(၁၉))	[colloquial] cause (“as, because of”)
<i>teguŋ</i> (၁၂၁၄)	cause (“for, on account of, thanks to, because of”)
<i>pʰg</i> (၁၅)	purpose (“intended for”)

FIGURE 44.4 CASE AND OTHER CORE POSTPOSITIONS

meanings that cohere around the notion “goal” (object, goal of motion, extent, etc.) requires the same type of analysis. The core P_o ’s—those associated with subject and object—are optional, in the sense that they are expressed only when certain discourse conditions are met (disambiguation, contrast, etc.; see Kashevitch 2005 on “optionality” in this context.) In the long example above, object precedes subject, and both are marked by P_o ’s (*gg* and *kò*) lest there be confusion about the roles of the initial phrases. Figure 44.4 lists the main case marking postpositions and the notional/grammatical relationships that they mark.

7.3 The clause

Burmese is consistently verb-final, though positive categorial sentences such as *?ú mjññ màun tèáun s?əjá* (ဦးမျဉ်မောင် တော်ဆရာတဲ့) ‘U Myint Maung is a school teacher’ tend not to express a copular verb. Among nominal clause elements, though certain orderings are frequent, e.g. discourse and scene setting elements (typically subject, time, and location phrases) before objects and inner locatives, and all those elements before adverbials and interrogative phrases, the order of nominal elements can be manipulated for various rhetorical effects (see Simpson and Watkins 2005 for an investigation of word order permutations in Burmese).

The VP is the only obligatory element of clause structure (other than in noun clauses), and although a sentence such as *Pjññ.pé.lai?.mè* ([ပျော်ပေးလိုက်မယ်] ‘[I]’ll fix [it] for [you]’ (‘repair.give.follow.IRR’) does not express any of the nominal arguments implied by the verb, it is nonetheless a complete sentence. Sentences such as these, with zero pronominalization, which are the norm in Burmese, should be contrasted with elliptical sentences, e.g. *Mənəe?pʰjāñ* (မနောက်ပွဲ) ‘Tomorrow,’ given as a one-word answer to a question ‘When?’ Repetition of the verbal phrase is often more natural than ellipsis.

The verb string is anchored by clause particles, and closed off with various sentence particles that include the interrogatives, and others that express notions such as “certainty” or “doubt”: *θwá.mè n̩* (သွားမယ်နော်) ‘[I]’ll be off then’ (‘go.IRR okay?’). Main clause particles are shown in Table 44.5.

The indicative particles are unstressed in certain contexts. The distinctions, realis, irrealis, and punctative, are neutralized in the negative. “Punctative” (which might also be labeled “change of state”) expresses the realization of a state or initiation of an action: *Jau?.pi* (ရောက်ပါ) ‘[We]’ve arrived’; *Lou?.pi* (လှုပိပါ) ‘[She]’s begun [doing it].’

Burmese exhibits a typical range of clause types. It can be noted, though, that ambient clauses that predicate time of day or state of weather show a subject argument: *Mó jwà.dè* (မြေးရှာတယ်) ‘It’s raining’ (‘sky rain.RLS’); and that “complex transitive” clauses may be built around verbs of “shaping” and “processing,” as well as “naming” and “selecting”:

- (18) *Sʰáŋaŋ.jè.gò sʰá tèʰε?.nè.dè* ဆားငြန်ရောကို ဆား ချက်နေတယ်။
Salt.water.OBJ salt cook. PROG.RLS ‘They are refining salt out of salt water.’

Double-subject (topic–comment) sentences are a particularly common type:

TABLE 44.5 MAIN CLAUSE PARTICLES

	Affirmative			Negative
Imperative	ø			ng
Indicative	Realis tè (~ θə)	Irrealis mè (~ mə)	Punctative pi (~ pə)	pʰú

- (19) *Bəlú.dzí.hà mjε ?lóuη.dè gg.lé ləgwíη.lau?* *ʃi.dè*
Ogre.big.s eye.PL.S.also cymbal.like be.RLS
ဘို့လူးကြီးဟာ မျက်လုံးတွေကလဲ လင်းကွင်းလောက် ရှိတယ်။
'And the ogre's eyes were as large as cymbals.'

The inner subject and verb of such clauses have a tendency to lexicalize (as shown by the position of the adverbial in the following example):

- (20) *Bəlú.dzí.hà ?əlùη ?əjou?* *s'ho.dè* ဘို့လူးကြီးဟာ အလွန် အရှပ်ဆိုတယ်။
Ogre.big.s very appearance bad.RLS 'The ogre was very ugly.'

Existence and possession are expressed by a single clause type containing a locative expression and the verb *ʃi* 'be': *Dì.hmà ʃi.dè* (ဒီမာ ရှိတယ်) 'It's here' ('here.LOC have.RLS'); *θy.ηà ʃi.tè* (သူမှ ရှိတယ်) 'He has [one]' ('3p^{LOC}.LOC have.RLS').

7.4 Processes

Questions in Burmese require no special structural reorganization. Yes–no questions are formed by the addition of the sentence particle *lá* (လား) to the declarative. Content questions are based on the interrogative pronoun *bè* (ဘယ်), usually placed directly before the VP, in conjunction with another final interrogative sentence particle, *lé* (လည်း). *Bè* phrases are also indefinites: *bè.dɔ...lé* (ဘယ်တော့...လည်း) 'when,' but *bè.tɔ.ŋa* + neg (ဘယ်တော့မှာ) 'never' (with general P_o *ŋa*, which in other contexts has the sense of 'only when, only if'). Negation is indicated by a preverbal clitic, *mə*, which in main clauses is supported by the clause markers *p'úi*, *nɛ*, etc.: *Mə.pau?.p'úi.lá* (မပေါက်ဖူးလား) 'Don't [they] grow [here]?' ('NEG.grow.NEG.Q'). Alternative questions pattern like coordinate clauses, with redundant material often omitted.

7.5 Complex and compound sentences

Independent clauses are usually conjoined asyndetically, the only indication being parallel structure and non-final intonation in the first clause, though redundant clause elements, including the verb complex, may be omitted (gapped). Verbs of locution, as well as certain verbs of cognition, take clauses as complements, marked with the "quotative" particle, *lo* (လွှဲ). In some cases, the verb [e.g., *teàn̥tʰá.dè* in the following example] may be 'understood': *θwá.mə.lo* [*teàn̥tʰá.dè*] (သွားမလို့ [ကြွထားတယ်]) '[I] intend to go' ('go.IRR.QUOT[intend.put.RLS]'). A more general "evidential" particle, *tɛ*, appearing in final position as if meaning 'it was said,' reports on any locution: *θwá.mə.lo.dɛ* (သွားမလို့တဲ့) '[She] said [she] intended to go.'

With other verbs, clauses are embedded through the replacement of the main clause particle with an appropriate member of the set of nominalizing nouns, or by addition of the general nominalizers *tà* (စဲ့) or *ŋà* (ခွဲ) (parallel in function to the main clause markers *tè* and *mè*):

- (21) *teənou?* *pəsʰó teu?.təq.ŋà sójèiη.dè* ကျွန်ုပ်ပုဆိုး ကျွတ်ကျုမှာစိုးရိမ်တယ်။
1p sarong undo.fall.IRL^{NOMZR} fear.RLS 'I was afraid my sarong would fall off.'

Embedding is especially common with verbs of elapsed time:

- (22) *Jau?.nè.dà θóuη.lq cà.bi* ကျောက်နေတာသံးလဲကာပီ။
Arrive.state.RLS^{NOMZR} 3.months long.PUNC '[I]'ve been here three months.'

The general nominalizers also appear in matrix clauses in a construction (“sentence nominalization”) that presupposes the verb and focuses on a nominal element (with *tà* replacing *te*, in the following example):

- (23) *Lùdzòuŋ.gga yù.là.pé.dà* လူကြံကယူလာပေးတာ။
courier.s carry.come.give.RLS^{NOMZR} ‘It was brought by courier.’

Cleft sentences perform a similar identifying function by recasting a verb sentence as a noun sentence:

- (24) *Tzəjouʔ zəgá pjó.taʔ.tà.ga kò Nàŋda.bà* တရုပ်စကား ပြောတတ်တာက ကိုနှစ်ပါ။
Chinese language speak able.RLS^{NOMZR}.s [Name].POL
‘Ko Nanda’s the one who speaks Chinese.’

Focus may also be placed on the verb by first citing it in nominal position (marked with one of the set of general *P_o*’s), then reiterating it in verbal position, a process known as “exposure” or “verb focus”: *θè.gò θè.ʔóuŋ.mà* (ဆောကိုသေအုံမှာ) ‘[You]’ll get [yourself] killed’ (*die.extent die.FUT.IRL^{NOMZR}*, e.g. warning a child).

In some cases, clauses may be transformed into derived nominals that retain nominal arguments. This is normal with the verb *kʰàŋ* ‘suffer,’ for example, where it occurs in a construction that is as close to a passive as can be found in Burmese:

- (25) *θù.gga baʔsəká ʔətai? kʰàŋ.ja.de* သူက ဘတ်စကား အတိုက် ခံရတယ်။
3p.s bus hitting suffer.get.RLS ‘He got hit by a bus.’

Clauses are subordinated by one of a set of particles that mark the clause boundary (and can be followed by general *p_o*’s). Included in this set is the weakly subordinating particle, *pí* (ပါး), that connects temporally consecutive as well as simultaneous clauses:

- (26) *tʰí sʰàuŋ.bí seʔbéiŋ.ng lá.dé* ထိုးဆောင်ပြီး စက်သီးနှံ လာတယ်။
umbrella hold.ing bicycle.by come.RLS ‘[She] rode up holding an umbrella.’

The final example, a proverb (for people who wear sarongs), illustrates conditional clauses, and gives us a glimpse of the wit of Burmese gnomic expressions:

- (27) *ʔəpjɔ́ mə.taʔ.jìŋ sʰé.ðə.lo ʔətʰàiŋ mə.taʔ.jìŋ pʰjé.ðə.lo*
အခြောမတတ်ရင်ခဲ့သလို အထိုင်မတတ်ရင်ဖြဲသလို။
speaking NEG.capable.if swear.RLS.like sitting NEG.capable.if gape.RLS.like
‘If you don’t speak well it’s like swearing; if you don’t sit right it’s like “baring”.’

ADDITIONAL ABBREVIATIONS

<i>IRL^{NOMZR}</i>	irrealis nominalizer
<i>POL</i>	polite
<i>PUNC</i>	punctative
<i>RLS</i>	realis
<i>RLS^{NOMZR}</i>	realis nominalizer
<i>3p^{LOC}</i> , etc.	the superscript indicates the function of the “induced creaky tone”

FURTHER READING

- Allott, Anna J. (1967) ‘Grammatical tone in modern spoken Burmese’, *Wissenschaftliche Zeitschrift der Karl-Marx Universität, Leipzig, Gesellschafts- und Sprachwissenschaftliche Reihe* 16, 1/2: 157–162.

- Allott, Anna J. (1985) ‘Language policy and language planning in Burma’, in David Bradley (ed.) *Papers in South-East Asian Linguistics No. 9: Language Policy, Language Planning and Sociolinguistics in South-East Asia (Pacific Linguistics, A-67)*, Canberra: Pacific Linguistics, 131–154.
- Becker, Alton L. (1995) ‘The figure a classifier makes: describing a particular Burmese classifier’, in Alton L. Becker (ed.) *Beyond Translation: Essays toward a Modern Philology*, Ann Arbor, MI: University of Michigan Press, 211–230.
- Jenny, Mathias (2015) ‘The far west of Southeast Asia: ‘give’ and ‘get’ in the languages of Myanmar’, in Nick J. Enfield and Bernard Comrie (eds) *Mainland Southeast Asian Languages: State of the Art and New Directions*, Berlin: de Gruyter, 155–208.
- Sawada, Hideo (1995) ‘On the usages and functions of particles *-kou/-ka*. in colloquial Burmese’, *Senri Ethnological Studies* 41: 153–187.
- Spriegg, R.K. (1957) ‘Junction in spoken Burmese’, in W.S. Allen (ed.) *Studies in Linguistic Analysis*, Oxford: Philological Society, 104–138.
- Vittrant, Alice (2014) ‘Psycho-collocational expressives in Burmese’, in Jeffrey P. Williams (ed.) *The Aesthetics of Grammar: Sounds and Meaning in the Languages of Mainland Southeast Asia*, Cambridge and New York: Cambridge University Press, 255–279.

NOTES

- 1 Special thanks to Christian Lammerts, Mimi Tian, and Myo Thwin Myint for providing invaluable feedback on drafts of this chapter.
- 2 In 1989, the government changed the English name of the country from Burma to Myanmar (with final “r” silent). In Burmese terms, the change from Burma to Myanmar reflects a shift between names that are ultimately cognate: from the more colloquial <Bamā> (pronounced *bəmà*) to the literary (and earlier attested) <Mranmā> (pronounced *mjəmà*). Even though the name Myanmar now has official status, Burmese as the name for the language seems preferable to the awkward “Myanma(r) language.” The names of important cities and geographical features were also given new English versions, some reflecting the Burmese pronunciation more closely: thus Rangoon became Yangon, Pagan, Bagan, Irrawaddy, Ayeyarwadi (with the silent “r” again), and so on. The full name of the country is the Republic of the Union of Myanmar.
- 3 Unfortunately, the standard IPA usage for representing tones clashes with earlier conventions, particularly the popular Okell system. In the current system, the three open tones of Burmese are marked over the main vowel sign as follows: low tone, with a grave accent (à); high, with an acute (á); and creaky with the sign for creaky voice (ä). Checked syllables, which can be regarded as a fourth tone, are marked by final glottal stop. A dot between syllables within a phonological phrase separates words or morphemes.
- 4 See Hla Pe (1973) for a long list of Mon loans that includes these examples.
- 5 One of the advantages of the Okell system of transcription for teaching materials is that phonemic shifts resulting from close juncture (e.g. *k^háñ ~ gáñ*, *té ~ dè*) are represented by underlining (i.e. *k^háñ ~ k^háñ*) rather than substitution, thus preserving word shape.
- 6 The abbreviation (၁၄) takes a vertical version of the numeral “4” (၄) pronounced *lé* to stand in for ၁၄, *lé* ‘also, as well,’ and combines it with the final of ၁၄၏ ‘good’; hence ‘as well ၁၄, also the case; ditto.’

REFERENCES

- Allott, Anna (1965) ‘Categories for the description of the verbal syntagma in Burmese’, in G.B. Milner and Eugénie J.A. Henderson (eds) *Indo-Pacific Linguistic Studies II (=Lingua 15)*, Amsterdam: North-Holland, 283–309.

- Bradley, David (2005) ‘Reflexives in literary and spoken Burmese’ in Justin Watkins (ed.) *Studies in Burmese Linguistics*, Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University, 67–86.
- Cooke, Joseph R. (1968) *Pronominal Reference in Thai, Burmese, and Vietnamese*, University of California Publications in Linguistics 52, Berkeley and Los Angeles: University of California Press.
- Duroiselle, Charles (1916) ‘Literal transliteration of the Burmese alphabet’, *Journal of the Burma Research Society* 6: 81–90.
- Kassevitch, Vadim B. (2005) ‘Syntactic and morphological markers in Burmese: are they really optional?’ In Justin Watkins (ed.) *Studies in Burmese Linguistics*, Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University, 87–96.
- Ministry of Immigration and Population (2015) *The Republic of the Union of Myanmar, Population and Housing Census of Myanmar, 2014, Provisional Results. Census Report Volume (I)*, Nay Pyi Taw, Myanmar: Department of Population, Ministry of Immigration and Population.
- Okell, John (1969) *A Reference Grammar of Colloquial Burmese* (2 Vols), London: Oxford University Press.
- Okell, John (1971) *A Guide to the Romanization of Burmese*, James G. Forlong Fund Vol. 27, London: Royal Asiatic Society of Great Britain and Ireland.
- Okell, John (1994) *Burmese: An Introduction to the Literary Style*, Southeast Asian Language Text Series, DeKalb, IL: Center for Southeast Asian Studies, Northern Illinois University.
- Okell, John (1995) ‘Three Burmese dialects’, in David Bradley (ed.) *Papers in South-East Asian Linguistics No. 13: Studies in Burmese Languages (Pacific Linguistics, A-83)*, Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University, 1–138.
- Pe, Hla (1973) ‘A tentative list of Mon loan words in Burmese’, *Journal of the Burma Research Society* 50.1: 71–94.
- Simpson, Andrew and Justin Watkins (2005) ‘Focus in Burmese: an investigation and experimental study of information structure and prosody’, in Justin Watkins (ed.) *Studies in Burmese Linguistics*, Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University, 27–66.
- Thurgood, Graham (1981) *Notes on the Origins of Burmese Creaky Tone*, Monumenta Serindica 9, Tokyo, Japan: Institute for the Study of Languages and Cultures of Asia and Africa.
- U Saw Tun (2005) ‘Writing Modern Burmese: an examination of the status of colloquial Burmese’, in Justin Watkins (ed.) *Studies in Burmese Linguistics*, Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University, 201–219.
- Watkins, Justin (ed.) (2005) *Studies in Burmese Linguistics*, Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.
- Wheatley, Julian (2014) ‘Delight in sound: Burmese patterns of euphony’, in Jeffrey P. Williams (ed.) *The Aesthetics of Grammar: Sounds and Meaning in the Languages of Mainland Southeast Asia*, Cambridge and New York: Cambridge University Press, 237–254.

CHAPTER FORTY-FIVE

ZAIWA

Yanhua Zhu

1 INTRODUCTION

Zaiwa [exonym Atsi] is spoken in Yunnan, China; Kachin State, Myanmar; Northeast India; and northwest Thailand. In China the Zaiwa speakers are considered a branch of the Jingpo nationality. The number of Zaiwa people in China is about 76,500 (2000 census), and they mainly live in Mangshi, Ruili, Longchuan, Yingjiang counties of Dehong Dai and Jingpo Autonomous Prefecture in Yunnan Province.

Zaiwa belongs to the Burmic branch of Tibeto-Burman, and is similar to Leqi, Bola, Langsu, and Burmese. It includes three dialects: Longzhun, Tingzhu, and Bengwa. The Longzhun dialect is spoken in Xishan Township, Zhefang Town of Mangshi City, Longchuan County and Ruili City; the Tingzhu dialect is spoken in Mangzhang Township, Zhanxi Town, Zhina Township, Xincheng Town, Taiping Town of Yingjiang County; and Bengwa dialect is spoken in Santaishan Township and Wuchalu Township. This paper mainly focuses on the Zaiwa spoken in Zhefang Town, Mangshi City.

2 PHONOLOGY

Zaiwa has 29 initial consonants, which have the following characteristics: stop and affricate consonants show an opposition between aspirated and unaspirated but no opposition between voiced and unvoiced. Fricative consonants show a contrast between voiced and unvoiced. Bilabial and velar consonants show a contrast between palatalized and non-palatalized stops. Affricate consonants only occur in the dental and post-alveolar positions. /f/ only exists in Chinese borrowings, like /fun⁵¹/ ‘flour.’ The initial consonants are listed in Table 45.1.

There are 122 finals, with major characteristics including lax and tense vowels, which exist in orderly correspondence. There are seven monophthongs plus their corresponding tense versions, and four diphthongs, all off-glides, plus their corresponding tense versions. The consonant endings include the stop codas /p/, /t/, /k/, /ʔ/ and the nasal codas /m/, /n/, /ŋ/. They can combine with all the monophthong vowels except /ɿ/ and /ɻ/, but cannot combine with the diphthongs. The pronunciation of the codas /p/, /t/, /k/ is occlusive but not plosive (i.e. they are unreleased). The finals are listed in Table 45.2.

There are four tones: high-level tone 55, as in /mjap⁵⁵/ ‘forever’; low-falling tone 31, as in /mjan³¹/ ‘horse’; high-falling tone 51, as in /mjər⁵¹/ ‘see’; mid-rising tone 35, as in /mjau³⁵/ ‘temple.’ Mid-rising tone 35 is only seen in Chinese borrowings, such as /tsəŋ³⁵/ ‘Tibetan,’ and as a sandhi tone, as in /pəŋ⁵⁵wɔʔ³¹fak³¹/ ‘loom.’

The phonetic processes of Zaiwa include weakening, epenthesis, and tone sandhi. The weakening only exists for vowels. Generally when the first syllable of disyllabic words is a single vowel or vowel with final glottal stop, the vowel will have a weak pronunciation. A non-central vowel will be pronounced as the central vowel [ə], with a slightly short tone. For example, /ʃɔ̃lɪkjɔ̃l/ > [ʃəlɪkjɔ̃l] ‘pitiful’; /sú⁵⁵pə⁵⁵/ > [sə⁵⁵pə⁵⁵] ‘someone else’;

TABLE 45.1 THE INITIAL CONSONANTS OF ZAIWA

		Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Palatal	Velar
Stop	Unaspirated	p pj		t			k kj	
	Aspirated	ph phj		th			kh khj	
Affricate	Unaspirated			tʃ		ʈʃ		
	Aspirated			tsh		ʈh		
Nasal		m mj		n			ɳ ɳj	
Lateral				l				
Fricative	Voiceless	f	s		ʃ		x xj	
	Voiced	v			ʒ			
Semivowel		w				j		

TABLE 45.2 THE FINALS OF ZAIWA

Monophthongs	Diphthongs	Finals with consonant coda								
		With nasal coda				With stop coda				
		-i	-u	-m	-n	-ŋ	-p	-t	-k	-ʔ
i				im	in	ɪŋ	ip	it	ik	i?
ɪ				ɪm	ɪn	ɪŋ	ɪp	ɪt	ɪk	i?
u	ui			um	un	ʊŋ	up	ut	uk	u?
ɯ	ui	ɯm	ɯn	ɯŋ	ɯp	ɯt	ɯk	ɯ?		
ə		əm	ən	əŋ	əp	ət	ək	ə?		
ə̥		ə̥m	ə̥n	ə̥ŋ	ə̥p	ə̥t	ə̥k	ə̥?		
ɛ		ɛm	ɛn	ɛŋ	ɛp	ɛt	ɛk	ɛ?		
ɛ̥		ɛ̥m	ɛ̥n	ɛ̥ŋ	ɛ̥p	ɛ̥t	ɛ̥k	ɛ̥?		
ɔ	ɔi	ɔm	ɔn	ɔŋ	ɔp	ɔt	ɔk	ɔ?		
ɔ̥	ɔi	ɔ̥m	ɔ̥n	ɔ̥ŋ	ɔ̥p	ɔ̥t	ɔ̥k	ɔ̥?		
a	ai au			am	an	aŋ	ap	at	ak	a?
ɑ	ai au	ɑm	ɑn	ɑŋ	ɑp	ɑt	ɑk	ɑ?		
l									l?	
ɿ									ɿ?	

/tǎʔ³¹ wu⁵¹/ > [təʔ³¹ wu⁵¹] ‘cotton.’ The vowel in the three weakened syllables is respectively pronounced as [ɔ], [u], [a] when the syllable is pronounced alone, but as [ə], with the tone pronounced weakened and shorter, when in combination with the other syllable.

3 MORPHOSYNTAX

Zaiwa words are mainly compounds, including root + root combinations and affixed forms (prefix + root, root + suffix). There are 11 word classes, including nouns, pronouns, numerals, quantifiers, verbs, auxiliary verbs, adjectives, adverbs, conjunctions, onomatopoeic forms, and interjections.

The pronouns include personal pronouns, demonstrative pronouns, and interrogative pronouns. The singular personal pronouns feature a case difference: possessive vs other. The difference is in the tone and vowel. For example, the first and second singular non-possessive pronouns are /ŋɔ⁵¹/ and /naŋ⁵¹/, and the possessive pronouns are /ŋa⁵⁵/ and /naŋ⁵⁵/ . The possessive form of the third single personal pronoun only differs from the non-possessive form in terms of tone. Its non-possessive form is /jan³¹/ and the possessive

form is /jan⁵¹/. However, the dual and plural personal pronouns show no variations for case. The first person dual and plural include exclusive and inclusive forms. The exclusive forms only refer to the speaker and one or more others, but not the addressee, while the inclusive form refers to both speaker and addressee and others (in the case of the plural). The exclusive form of the dual pronoun is formed by adding the suffix /-nik⁵⁵/ after the singular possessive form /ŋa⁵⁵/, while the inclusive pronoun is formed by adding /-ník⁵⁵/ after /i⁵⁵/ . The plural exclusive pronoun is formed by adding the suffix /-moʔ³¹/ after /ŋa⁵⁵/ . The plural inclusive pronoun is formed by adding the suffix /-nun⁵⁵/ after /ŋa⁵⁵/ or /i⁵⁵/ . The second and third person dual pronouns are formed by adding the suffix /-nik⁵⁵/ after /naŋ⁵⁵/ and /jan⁵⁵/ respectively, while the plural pronouns are formed by adding the suffix /-moʔ³¹/ after /naŋ⁵⁵/ and /jan⁵⁵/ . See Table 45.3.

There is also a reflexive (1a) and emphatic (1b) pronoun /jum⁵¹səŋ⁵¹/ ‘self.’ If it is used as an emphatic pronoun, it always needs to be followed by /jan³¹/, possibly the third person pronoun but used adverbially here.

- (1) a. jan³¹ jum⁵¹səŋ⁵¹ mǎ³¹tu⁵⁵ tsa³¹ mjit³¹ 3a⁵⁵.
 3sg himself about only care REAL
 ‘He only cares about himself.’
- b. jan³¹ jum⁵¹səŋ⁵¹jan³¹ ŋɔ^{51/31} 3l⁵⁵ ap⁵⁵thau³¹ pji^{31/51} 3a⁵⁵.
 3sg personally 1sg P injection give REAL
 ‘He gave me an injection personally.’
- c. than³¹ kə³¹ jum⁵¹səŋ⁵¹ pum⁵¹ mǎ⁵⁵ tsan^{31/51} θ⁵⁵.
 firewood TOP one’s own hill LOC cut NMLZ
 ‘The firewood was cut on my own hill.’

The demonstrative pronouns differ in terms of distance, elevation, and number. The distance is categorized as proximal (‘this’), medial (‘that’), and distal demonstrative (‘that’). The distal demonstrative is further divided according to elevation: same level as speaker (‘that’), above speaker (‘that above’), below speaker (‘that below’). If the speaker wants to emphasize the remoteness of a referent, it can be expressed by lengthening the tone. The longer the tone, the more distance it represents. All of the demonstrative pronouns are also used for locations by adding /ma⁵⁵/ ‘at’ or /mai³¹/ ‘from.’ For example: /xji^{51/55}ma⁵⁵/ ‘here’; /xau^{51/55}ma⁵⁵/ ‘there’; /xje^{51/55}ma⁵⁵/ ‘there (farther)’; /xu^{51/55}ma⁵⁵/ ‘above’; /mɔ^{51/55}ma⁵⁵/ ‘below’; /xji^{51/55}mai³¹/ ‘from here’; /xau^{51/55}mai³¹/ ‘from there’; /xje^{51/55}mai³¹/ ‘from there (farther)’; /xu^{51/55}mai³¹/ ‘from above’; /mɔ^{51/55}mai³¹/ ‘from below.’ Number includes singular and plural, the latter formed by adding the suffix /-pg⁵⁵/ or /-paŋ³¹/ to the singular. See Table 45.4.

Auxiliary verbs mainly serve to provide further information about the verb. In Zaiwa, there are various types of auxiliary verbs with complex usage. These verbs also have cross-functions and can be used as other categories of words. Normally they derive from

TABLE 45.3 THE PERSONAL PRONOUNS OF ZAIWA

		1st person	2nd person	3rd person
Singular	Non-possessive	ŋɔ ⁵¹	nan ⁵¹	jan ³¹
	Possessive	ŋa ⁵⁵		
Dual		ŋa ⁵⁵ nik ⁵⁵ exclusive; i ⁵⁵ nik ⁵⁵ inclusive	nun ⁵⁵ nik ⁵⁵	jan ⁵⁵ nik ⁵⁵
Plural		ŋa ⁵⁵ moʔ ³¹ exclusive; ŋa ⁵⁵ nun ⁵⁵ inclusive i ⁵⁵ nun ⁵⁵ inclusive	nun ⁵⁵ moʔ ³¹	jan ⁵⁵ moʔ ³¹

TABLE 45.4 THE DEMONSTRATIVE PRONOUNS OF ZAIWA

Proximal demonstratives	Neutral demonstratives	Distal demonstratives	Same level	Above	Below
xji ⁵¹ xji ^{51/55} pə ⁵⁵ xji ^{51/55} panj ³¹	xau ⁵¹ xau ^{51/55} pə ⁵⁵ xau ^{51/55} panj ³¹	xje ⁵¹ xje ^{51/55} pə ⁵⁵ xje ^{51/55} panj ³¹	xu ⁵¹ xu ^{51/55} pə ⁵⁵ xu ^{51/55} panj ³¹	mə ⁵¹ mə ^{51/55} pə ⁵⁵ mə ^{51/55} panj ³¹	

action verbs. For example, the direction auxiliary verbs /lo⁵¹/ ‘come’ and /lo⁵⁵/ ‘go’ are grammatical uses of the action verbs /lɔ⁵¹/ ‘come’ and /lɔ⁵⁵/ ‘go.’ For example:

- (2) a. xau^{51/55} mai³¹ lo⁵⁵ aṛ³¹.
 that ABL go IMP
 ‘Let us go from there.’
- b. jan³¹ sək⁵⁵kam⁵¹ va³¹ lo⁵⁵ pə⁵¹.
 3sg tree carry away CSM
 ‘He carried the tree away.’
- (3) a. jan⁵⁵məṛ³¹ lɔ⁵¹ pə⁵¹.
 3pl come CSM
 ‘They came.’
- b. jan³¹ tum³¹ təu^{31/51} lɔ⁵¹ pə⁵¹.
 3sg again return towards CSM
 ‘He returns again.’

The auxiliary verbs /pjii³¹/ ‘away’ and /ju⁵¹/ ‘toward center’ indicate action away from the center and action toward the center, respectively. These two auxiliary verbs come from the verbs /pjii³¹/ ‘give’ and /ju⁵¹/ ‘take,’ respectively. For example:

- (4) a. jan⁵¹ ḡ⁵⁵ laṛ³¹khaṛ³¹ thuk³¹-pjii³¹ pə⁵¹.
 3sg POSS salary transfer-CENTRIF CSM
 ‘His salary was given to him.’
- b. ḡ⁵⁵ laṛ³¹khaṛ³¹ thuk³¹-ju⁵¹ pə⁵¹.
 1sg:POSS salary transfer-CENTRIP CSM
 ‘I received my salary.’
- (5) a. jum⁵¹ sū⁵⁵pə⁵⁵ ʒl⁵⁵ poi⁵⁵-pjii^{31/51} tɔ³¹ lɛ⁵¹.
 house other.person P loan-CENTRIF CONT IRR
 ‘The house is loaned to someone else.’
- b. ḡ⁵¹ jum⁵¹ xji^{51/55} jum⁵¹ poi⁵⁵-ju^{51/31} tɔ^{31/51} lɛ⁵¹.
 1sg house this CL rent-CENTRIP CONT IRR
 ‘I rent this house.’

The noun phrase can be followed by a postposition marking the referent as agentive or instrumental (əṛ³¹); adverbial, ablative and container instrumental (mai³¹); patient, temporal and material instrumental (ʒl⁵⁵); locative and temporal (ma⁵⁵); comitative (əṛ⁵⁵); comparative (tho⁵⁵). There is also a topic marker (kə³¹). The agentive marking is not obligatory on transitive or ditransitive clauses, but is often used when there is a specific identifiable patient referent, particularly if the direction of action is inverse.

Aspect marking is used to show the different stages of the action or state, including contemplated ($a^{31} \dots \underline{f}_1?^{31}$); imminent ($p\underline{o}^{31}za^{51}, p\underline{o}^{31}k\underline{o}^{31} za^{51}$); prospective ($za^{51}, k\underline{o}^{31} za^{51}$); inchoative ($l\underline{o}^{51}$); change of state ($p\underline{o}^{51}, p\underline{o}^{31}k\underline{o}^{51}$); progressive ($\underline{n}ji^{51}$); durative ($t\underline{o}^{31}$); ongoing ($a^{31}\underline{f}_1^{51}$); perfective ($t\underline{o}m^{31}$); experiential (wu^{55}).

The realis and irrealis markers are used to show whether the action has happened or not. If the action has happened, the realis marker / za^{55} / or / $a^{31}k\underline{o}^{51}$ / should be added to the end of clause, while if the action hasn't happened, the irrealis marker / le^{51} / should be used. / za^{55} / is used when the agent is singular or first person plural, while / $a^{31}k\underline{o}^{51}$ / is used when the agent is a second or third person dual or plural form referring to an animate referent. The composition of / $a^{31}k\underline{o}^{51}$ / is / a^{31} / (a reduced form of / za^{55} /) + plural marker / $-k\underline{o}^{51}$. For example:

- (6) a. $paj^{55}tai^{51} \underline{fa}^{?^{31}} w\underline{o}^{55} tin^{51/31} za^{55}$.
 rabbit very well can run REAL
 ‘The rabbit runs very well.’
- b. $\underline{jum}^{51} m\underline{a}^{55} k\underline{o}^{31} pju^{51} sum^{31} ju?^{31} le^{55} a^{31}k\underline{o}^{51}$.
 home LOC TOP people three CL come REAL
 ‘Three people come to my home.’
- c. $na\underline{j}^{51} k\underline{o}^{31} i^{55}nu^{31}i^{55}va^{31} \underline{zi}^{55} phau^{31}tap^{55} za^{51} tut^{31} le^{51}$.
 2sg TOP parents P respect should AUX IRR
 ‘You should respect your parents.’

Hearsay is expressed by adding the evidential marker / ka^{51} / ‘hear from’ at the end of the sentence. / ka^{51} / can refer to information from other sources as well as folk tales or oral histories. For example:

- (7) a. $\underline{l}a^{31}kam^{55} na\underline{j}^{51/31} \underline{zi}^{55} pat^{31} \underline{\theta}^{55} ka^{51}$.
 PN 2sg P beat NMLZ HEARSAY
 ‘I heard Legam beat you.’
- b. $a^{31}nam^{31}saj^{51}x\underline{\theta}^{31} \underline{zi}^{55}, pak^{55}pan^{55} k\underline{o}^{31} \underline{l}a^{31} kun^{51}tu^{31} a^{31}p\underline{ju}^{55} tsa^{31} \underline{\eta}ut^{55}le^{51}$
 long time ago TEM, pak⁵⁵pan⁵⁵bird TOP one body slim ADV COP IRR
 $\underline{s}^{31}mau^{55} \underline{l}a^{31} khat^{55} \underline{zi}^{31} a^{31} ju?^{31} ka^{51}$.
 feather one CL all NEG grow HEARSAY
 ‘It is said, in the past, pak⁵⁵pan⁵⁵ birds were slim, without any feathers.’

Yes-no questions always take the mood particle / lu^{55} / or / $lu?^{31}$ /. Alternative questions always take the mood particle / $lu?^{31}$ /. Before the disjunction there is always a disjunction marker, / $a^{31}\underline{\eta}ut^{55}zi^{31}/, /a^{31}\underline{\eta}ut^{55}fan^{55}/ or / $m\underline{o}^{55}k\underline{o}^{31}$ /. Positive-negative questions are a unique form of alternative question. They are formed juxtaposing affirmative and negative choices, with a conjunction before the second option: / $a^{31}\underline{\eta}ut^{55}zi^{31}/, /a^{31}\underline{\eta}ut^{55}fan^{55}/ or / $m\underline{o}^{55}k\underline{o}^{31}$ /, is normally added. Content questions have an interrogative pronoun and need no mood particle at the end of the sentence.$$

The imperative sentence should have an imperative marker at the end of the sentence to express imperative mood. The imperative marker has complex variations according to the person that the imperative is directed towards. These are listed in Table 45.5.

The imperative marker for second person has as its singular form a low tone / $a?^{31}$ / or high level tone / $a?^{55}$ /. Its dual or plural form is / $k\underline{o}^{55}$ / or / $k\underline{o}^{?^{55}}$ /. / $k\underline{o}^{?^{55}}/\sim/k\underline{o}^{55}/$ is probably a fusion of the plural marker / $k\underline{o}^{51}$ / and the singular imperative form / $a?^{55}$ /, as the native speakers consider that all / $k\underline{o}^{?^{55}}$ /(or / $k\underline{o}^{55}$ /) can be pronounced / $k\underline{o}^{51/31} za?^{55}$ /. When / $k\underline{o}^{51}$ /

TABLE 45.5 THE IMPERATIVE MARKERS IN ZAIWA

	1st person	2nd person	3rd person
Singular		pa ⁵⁵	a ²⁵⁵ (or a ²³¹)
Dual	inclusive exclusive	k ² ? ³¹ pa ⁵⁵	k ² ? ⁵⁵ (or k ² ? ⁵⁵)
Plural	inclusive exclusive	fan ⁵⁵ pa ⁵⁵	k ² ? ⁵⁵ (or k ² ? ⁵⁵)
			k ² ? ^{51/31} fan ⁵⁵ ka ²³¹

and /a?⁵⁵/ combine together, the vowel of /k²?⁵¹/ becomes the central vowel /ə/. The tone changes from high falling tone /51/ to low falling tone /31/, and as there is a central vowel /ə/ before /a?⁵⁵/, an epenthetic /ʒ/ is added to the start of the syllable, forming [ʒa?⁵⁵].

The imperative marker for third person has as its singular form /fan⁵⁵ka²³¹/ and dual or plural form as /k²?^{51/31}fan⁵⁵ka²³¹. /k²?^{51/31}fan⁵⁵ka²³¹/ is the combination of the plural marker /k²?⁵¹/ and /fan⁵⁵ka²³¹. /ka²³¹/ can be omitted, and the plural form /k²?⁵¹/ is increasingly omitted as well. Therefore, the third person imperative particle can be simply /fan⁵⁵. See the following examples:

- (8) a. ηɔ⁵¹ kut⁵⁵ pa⁵⁵! b. ηa⁵⁵nik⁵⁵ ka⁵⁵ pa⁵⁵!
1sg do IMP 1dl:INCL write IMP
'Let me do it!' 'Let us write!'
c. i⁵⁵nik⁵⁵ ka⁵⁵ k²?³¹! d. ηa⁵⁵mɔ?³¹ ka⁵⁵ pa⁵⁵
1dl:INCL write IMP 1pl:EXCL write IMP
'Let us write!' 'Let us write!'
e. ηa⁵⁵nun⁵⁵ tʃɔŋ³¹ ts?^{31/51} je⁵¹ fan⁵⁵ f. nan⁵¹ i³¹ʃam³¹ fu?⁵⁵ a?³¹
1pl:INCL school go to IMP 2sg water drink IMP
'Let us go to school!' 'You drink water!'
g. nun⁵⁵nik⁵⁵ ka⁵⁵ k²?⁵⁵ h. nun⁵⁵mɔ?³¹ tai³¹ k²?⁵⁵
2dl write IMP 2pl say IMP
'You two write!' 'You say!'
i. jan³¹ pau⁵⁵ kut⁵⁵ fan⁵⁵ka²³¹! j. jan³¹mɔ?³¹ je⁵¹ k²?^{51/31}fan⁵⁵ka²³¹!
3sg alone does IMP 3pl go IMP
'Let him do it alone!' 'Let them go!'

Within the noun phrase, the modifiers include nouns, pronouns and adjectives (adjective phrase), verbs (verb phrase) and quantitative phrases, demonstrative-quantifier phrases and relative clauses. If the attributive is a single-syllable adjective, the word order is "head word + attributive"; if the attributive is a multiple-syllable adjective or adjective phrase, the attributive can appear before or after the head noun, but the nominalizing particle /ə⁵⁵/ must be used. For example:

- (9) a. jan³¹ k²?³¹ [pj⁵¹ k²?⁵¹] ηut⁵⁵ le⁵¹.
3sg TOP person good COP IRR
'He is a good person.'
b. jan⁵⁵mɔ?³¹ [tui³¹ xəŋ⁵¹ xəŋ⁵¹ ku⁵¹ ə⁵⁵] 3a⁵¹ a³¹k²?⁵¹.
3pl rope long REDUP AUX NMLZ need REAL
'They need a long rope.'

- c. pum⁵¹ mā⁵⁵ [ŋjui⁵¹ ŋjui^{51/31} ə⁵⁵ man³¹] ju?³¹ t₂^{31/51} 3a⁵⁵.
 mountain LOC green REDUP NMLZ grass grow CONT REAL
 ‘There is green grass growing on the mountain.’

The quantifier phrase serves as attributive, with two word orders: “head noun + quantifier phrase,” “quantifier phrase + head noun.” This is true also of demonstrative-quantifier phrases: “head noun + demonstrative-quantifier phrase” and “demonstrative quantifier phrase + head noun.” For example:

- (10) a. ŋjɔ³¹tsɔ³¹ lā³¹ kjin⁵⁵ fish one CL ‘One catty of fish’ b. lā³¹ kjin⁵⁵ ŋjɔ³¹tsɔ³¹ one CL fish ‘One catty of fish’
 c. i⁵¹phə³¹ xji^{51/55}tau⁵⁵ wine this CL ‘this wine’ d. xji^{51/55} tau⁵⁵ i⁵¹phə³¹ this CL wine ‘this wine’
 e. khji⁵¹tsunj³¹ xje⁵¹i⁵⁵ tsum⁵⁵ f. xje⁵¹i⁵⁵ tsum⁵⁵ khji⁵¹tsunj³¹ shoes that two CL that two CL shoes ‘Those two pairs of shoes’ ‘Those two pairs of shoes’

In the verb phrase, the modifiers include adjectives, verbs, adverbs, volitive auxiliaries, nouns (generally temporal nouns, locative nouns), pronouns (demonstrative pronouns and interrogative pronouns), and quantifier phrases (generally verbal classifier phrases and timing classifier phrases). Sometimes there are several modifiers. According to the meaning of adverbial modifiers, the word order is generally the following: mood, time, position, attitude, negative, degree, modal, number, trait, head word. For example:

- (11) a. jan³¹ wui⁵¹ wui⁵¹ kut⁵⁵ ja⁵⁵mɔ?³¹ 3l⁵⁵ tai^{31/51}kjɔ³¹ ə⁵⁵.
 3sg laugh REDUP AUX 1pl P tell NMLZ
 ‘She laughs and says to us.’ b. jan³¹ tum³¹ a³¹ kam^{55/51} je⁵¹ lɔ?³¹ ka⁵¹.
 3sg again NEG want go DECL HEARSAY
 ‘He does not want to go. (hearsay)’ c. nuŋ⁵⁵mɔ?³¹ ſəŋ⁵¹ʃl³¹ nap³¹ma⁵⁵ ſəŋ³¹ ma⁵⁵ po³¹luŋ⁵¹ kha⁵⁵ je⁵¹ pat³¹.
 2pl simply tomorrow school LOC ball NEG go play
 ‘You had better not play ball at school tomorrow.’ d. jan⁵¹ va³¹ a³¹ŋji⁵⁵nap³¹ ſəŋ³¹ ma⁵⁵ lā³¹tsup³¹ le⁵⁵ wu⁵⁵ 3a⁵⁵.
 3sg.POSS father yesterday school LOC one CL come look REAL
 ‘His father came to school to have a look yesterday.’

The number of transitive verb arguments can be increased or reduced through two major grammatical means. The first is to add an argument by using a causative form. Generally there are three causative constructions in Zaiwa: (i) inflected forms (verb initial consonants with aspirate non-aspirate alternations, verbs with tense versus lax vowel alternations, and verbs with voiceless fricative initials rather than semivowel initials), e.g. /kjp³¹/ ‘break’ ~ /khjp⁵⁵/ ‘make break’; /tsup³¹/ ‘sit’ ~ /tsun⁵¹/ ‘seat’; /wut³¹/ ‘wear’ ~ /xut⁵⁵/ ‘get dressed’; (ii) analytic form (adding the prefix /l₂?⁵⁵-/ to a simplex verb), e.g. /xa⁵⁵/ ‘loose’ > /l₂?⁵⁵xa⁵⁵/ ‘loosen’; (iii) combination of inflected and analytic form (with variation of consonant or vowel plus the prefix /l₂?⁵⁵-/), e.g. /pun³¹/ ‘wake’ ~ /l₂?⁵⁵pun³¹/ ‘wake up.’

The construction which reduces the number of arguments by adding the adverbial /lum³¹/ ‘mutual’ after the verb: For example, /pat³¹/ ‘fight’ > /pat³¹lum³¹/ ‘fight each other,’ /pji³¹/ ‘give’ > /pji^{31/51}lum³¹/ ‘give to each other.’ Examples of both types of valency-altering constructions are given in (12) and (13):

Most complex sentences rely on conjunctions or adverbs to show the relationship between the joined clauses. The position of the conjunction or adverb fits the principle “linked part in the middle.” Normally it is at the end of the first clause (14a) or different connectives are added at the end of the first clause and the beginning of the second clause (14b). A few of the complex sentences have connectives at the beginning and end of the first clause (14c).

BIBLIOGRAPHY

- Xu Xijian and Xu Guizhen. 1982. *Jingpozu yuyan jianzhi: Zaiwayu* (*Brief Description of the Languages of the Jinghpaw People: Zaiwa Language*). Beijing: Ethnic Publishing House.

Zhu Yanhua. 2012. The category of evidentiality in Zaiwa. *Journal of Yunnan Normal University* 5: 25–31.

Zhu Yanhua. 2012. Aspect in Zaiwa. *Journal of Sino-Tibetan Linguistics* 6: 112–130.

Zhu Yanhua and Lepai Zaozha. 2013. *Reference Grammar of Zhefang Zaiwa*. Beijing: China Social Sciences Press.

CHAPTER FORTY-SIX

AKHA

Inga-Lill Hansson

1 INTRODUCTION

The Akha people are spread over the border areas between Burma, Thailand, China, Laos and Vietnam. Their homeland is presumed to be in southwest China, where most of the Akhas still live. The migration routes have gone from China southwards to Burma and then on to Thailand since around 1900, and also to Laos and Vietnam. The Akha have traditionally lived on semi-high mountains, growing dry rice, but by now many Akhas live in the lowlands and in the cities. Population figures are not reliable. Approximately there are said to be 45,000 in Thailand, 200,000 in Burma, 100,000 in Laos and 7,000 in Vietnam. In China, where Akha is included in the Hani minority consisting of more than a million, Akha make up around 250,000.

Linguistically they belong to the southern part of the Burmese-Yipho (or Lolo-Burmese) subgroup of the Tibeto-Burman branch of the Sino-Tibetan language family. Akha is closely related to Hani (Hansson 1982, 1989) and I have suggested a scheme for southern Yipho (Figure 46.1).

This scheme does not include Gókhý, living mainly in China, where more work needs to be done (Hansson 1991).

Traditionally, Akha has had no writing system. A system based on Roman letters was devised in Burma in the 1950s and used in the *Akha–English Dictionary* by Paul Lewis (1968). Later a writing system based on Thai letters was developed by missionaries in Thailand but has not been widely used. In China, a system based on the *pinyin* transcription system for Chinese has been devised for Hani. This writing system is used, together with Chinese, in the first four years in schools in Hani areas, but as Akha is so different from Hani, it cannot be used for Akha without changes. There is no official system for writing Akha in China, but the Akha community there has made a system according to the same principles, and some Akhas in Thailand use a similar system. Very few are literate in Akha though.

Akha has a rich oral tradition with a huge body of ritual texts transmitted orally (Hansson 2013). The texts belong to the offices of *phíma* ‘priest’, *dz:ma* ‘village leader’ and *njíphà* ‘shaman’. The texts are recited on the occasion of death or sickness. The ritual language is slightly different from the modern language of today. Most disyllabic nouns in the modern language are monosyllabic; the syntax is also changed to fit the metrical pattern, which makes the texts difficult to understand for young Akhas of today (Hansson 1991). Apart from the ritual texts, there is also a vivid tradition of songs, storytelling and

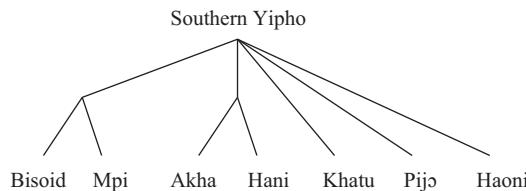


FIGURE 46.1 RELATIONS AMONG THE SOUTHERN YIPHO LANGUAGES

myths. With the rapid economic changes in the whole area affecting the traditional life of the Akha people, these texts are in danger of disappearing. Six years ago, the Hani and Akha of China initiated conferences dealing with all aspects of Akha life and history, and there are now small groups working in China and in Thailand together with some foreign researchers recording those texts, translating them into Modern Akha, and into Chinese, Thai and English. The third conference was held in Jinghong, China, in December–January 1999–2000, where Akha from China, Thailand and Burma, together with Hani from China, discussed all such issues.

Most Akhas are and have probably always been bilingual or even trilingual or more. The older generation in Thailand often speak Lahu, northern Thai, and sometimes also Chinese. All over the area, the Akhas by now speak the local dominant dialect of the majority language and – if they have gone to school – the standard language of the country. Akha is at present not an official school language anywhere. Some Akhas have received further education, some even up to university level.

The following description is based on Akha as spoken in northern Thailand. There are very few differences between Akha in Thailand, Burma and China, and as far as I know also in Laos, and they can communicate easily. As the Akha in the various countries of course relate to different majority languages and also learn that language in school, many new loanwords enter Akha from different directions. This causes more problems in mutual understanding, especially when discussing modern issues requiring new terminology, e.g. in the fields of economics, law and politics. Some new phonemes may also enter the language in loanwords, as e.g. /f/ and /w/ in Akha in China.

2 PHONOLOGY

Akha has the following initials and finals:

There are two phonation types, one breathy, lax (written ‘zero’), one laryngealized, tense (written *-q*). Aspirated initials only occur in non-laryngealized syllables (there are a few exceptions, mainly in names). Laryngealization in Akha is the reflex of Proto-TB final *-p/t/k. Except for the back nasal *-ay*, all finals occur non-laryngealized as well as laryngealized. There are three tones in open syllables: high *-á*, mid *-a*, low (slightly falling) *-à*, and three in laryngealized syllables: high *-áq* (very rare), mid *-aq*, low *-àq*, for example:

<i>má</i>	full
<i>ma</i>	mother
<i>mà</i>	not
<i>maq</i>	dream
<i>màq</i>	group

In Akha, there is both an initial *m*- and a final, syllabic *-m*, which may also be laryngealized:

<i>i-nìmq</i>	brain (cf. WB <i>nok</i>)
<i>nìm-shì</i>	sesame

The syllable structure is (c)v. The back vowels combine with more initials than the front vowels. The top scorers are *-ɔ* and *-a*, which occur with all initials; the least versatile final is *-ii*, which occurs with only a few initials in very few words.

There are some words with the syllabic structure cv + v, which may be analysed as two syllables or as one with a diphthong. Some of them have longer forms, used for emphasis, with a consonant initial in the second syllable:

<i>míε < mí ne</i>	because
<i>xòqi < xòq ni</i>	although
<i>jò yà naqù < jò yà na lùq</i>	everybody (where the last syllable has lost the initial and the laryngealization has moved forward)

In a few other ones, there is no long form to help in the analysis. The second vowel is always *-a*, which could be the topic particle *-a* or something else, for example:

<i>náa</i>	when
<i>mía</i>	sentence particle, non-past
<i>mìa</i>	sentence particle, past
<i>ŋáa</i>	sentence particle, non-past
<i>ŋàa</i>	sentence particle, past

(For discussion of sentence particles, see later.)

For the time being, I analyse these words as consisting of two syllables and not long vowels or diphthongs.

Akha has very few cases of shift between phonation types, vowels or tones to mark a syntactic function or different word class, which is otherwise not uncommon in the language group. A few verbal auxiliaries have a change in phonation type, and also in the vowel:

<i>làq</i>	movement towards first person
<i>lèq</i>	movement towards non-first person
<i>nèq</i>	action towards first person
<i>nè</i>	action towards non-first person

First and second person pronouns change phonation type and/or tone when followed by certain noun particles:

<i>ŋá</i>	1sg	<i>ŋà ne</i>	by me	<i>ŋà áŋ</i>	to me	<i>ŋàq ε</i>	mine
<i>ná</i>	2sg	<i>nò ne</i>	by you	<i>nò áŋ</i>	to you	<i>nòq ε</i>	yours

There are a few simple sandhi rules governing some of the sentence particles which have high tone for non-past and low tone for past. They both have mid tone as sandhi:

High + High > High + Mid	<i>lá mé</i> > <i>lá me</i>	he is coming
Low + Low > Low + Mid	<i>dzà mè</i> > <i>dzà me</i>	he ate

There are a number of sound changes in Akha readily observable. Some mergers have affected the phonemic system and seem to be almost completed in the speech of the younger generation. The most conspicuous change is the merging of the alveolar affricates and fricatives with the palatal series:

<i>Open syllables</i>	<i>Laryngealized syllables</i>
tsh > tjh	ts > tj
dz > dj	dz > dj
z > j	z > j

but the other way round in the following, where there is a loss of palatalization:

sjh > sh	sj > s
----------	--------

Many older speakers still keep the distinctions but you can clearly hear that the change is happening to everybody. Over the last few years, the final *-ay* has changed to *-ɔ* at a surprising speed. With some speakers I have noted the change being completed within only a few years. The same change seems to be going on in Akha in China, and it is interesting to note that the closely related Hani has no final nasals at all. The remaining nasal final *-m* seems to be stable so far.

3 WORD FORMATION

Most verbs in Akha are monosyllabic:

<i>dzà</i>	eat
<i>dó</i>	drink
<i>sjhí</i>	die
<i>jùq</i>	sleep
<i>kha</i>	plant
<i>dù</i>	dig
<i>dì</i>	beat
<i>n:</i>	think
<i>m̥</i>	do

But more often than not the verb phrase consists of a concatenation. In the modern language most nouns are compounds, where either each syllable has a discernible meaning or one of the syllables is an affix. Some very common ones are monosyllabic though, for example:

<i>tjhé</i>	paddy
<i>hò</i>	cooked rice
<i>m̥</i>	heaven

Akha does not abound in affixes. The only widespread ones in the modern language are the prefix *a-* and the suffix *-ma*:

Prefix *a-* (high or low tone)

in many kinship terms and plant names:

<i>àma</i>	mother
<i>àda</i>	father
<i>àbú</i>	daughter
<i>àli</i>	son
<i>àdu</i>	corn
<i>ábjèq</i>	bamboo shoot

The prefix *a-* can also be attached to other free or bound nominal or verbal roots, which in their turn may occur without the prefix in combinations with other morphemes:

ábò (GEN) bugs, insects, cf. *bò-xòq* caterpillar
àbjìm head of a plant, cf. *bjìm-ma* Caladium

The suffix *-ma* has two clear meanings: (1) female, mother, as *àma* ‘mother’, (2) ‘big’, ‘major’. It also seems to have a general noun formative function with no clear meaning:

<i>ja-ma</i>	hen (< <i>ja-tjiq ja-ma</i>)
<i>khy-ma</i>	daughter-in-law
<i>dú-ma</i>	male's sister
<i>phí-ma</i>	priest
<i>dzò-ma</i>	village leader
<i>ja-ma</i>	elephant
<i>de-ma</i>	wet field
<i>gá-ma</i>	road
<i>ny-ma</i>	heart
<i>gy-ma</i>	attic (storage area)

Some prefixes are limited to certain kinds of animals, as *bò-* for many insects, *ho-* for rats and rodents, *yà-* for fishes (<*yà-sjhà* ‘fish’) and *xhà-* for some birds and beasts:

<i>xhà-hm̄</i>	bear
<i>xhà-là</i>	tiger
<i>xhà-xhò</i>	dove
<i>xhà-phà</i>	frog

Many bound morphemes occur in a number of compounds together with other free or bound morphemes. Some of them can be assigned a meaning, but many cannot so far. Some examples:

bjá flat, broad

<i>dəq-hà-bjá</i>	kind of axe with a broad blade
<i>bò-bjá</i>	large kind of turtle
<i>ù-bjá</i>	kind of flat headdress
<i>bjá-yá</i>	side

luq move (free verb)

<i>i-luq</i>	billow (< <i>i-tjùq</i> water)
<i>mí-luq</i>	earthquake (< <i>mí-tshà</i> earth)
<i>dzaj-luq</i>	rolled up topknot

hà – no clear meaning

<i>dəq-hà</i>	axe
<i>tshó-hà</i>	person
<i>tshò-hà</i>	dirt area for fireplace (hearth)
<i>tjí-hà</i>	barking deer
<i>bó-hà</i>	bat

bó – no clear meaning

<i>abó</i>	grandfather
<i>ábó</i>	tree
<i>nà-bó</i>	ear

A common suffix in the ritual language, *-ba*, is in the modern language lexically restricted to a few spatial nouns, as *xhó-ba* ‘over there’. In the ritual language it occurs after many place words and time words:

<i>gý-ba</i>	downhill
<i>nja ba</i>	uphill
<i>mé-ba</i>	upstream
<i>bí-ba</i>	downstream
<i>zá-ba</i>	later
<i>hù-ba</i>	before

The stress in a disyllabic noun is always on the syllable with the highest tone irrespective of whether the syllable is an affix or not. If both syllables have the same tone, the stress is even.

<i>phí-ma</i>	priest
<i>dzò-ma</i>	village leader

In both words, *-ma* is a suffix.

4 SYNTAX

An Akha sentence with expressed agent, indirect and direct object, time word, adverb and a verb phrase ending with a sentence particle and a final particle, may look like this:

àjòq tshó-mò jɔ-m� xhó njì yà áy áshi jɔ-khý jɔ nanjɔ ishàŋ i biq i mòq ́ nja djé
3sg old people good those two CLF fruit tasty every day very give go want FEEL:NPAST
INDIRECT.QUOTE

‘He says he very much wants to go to give those two good old people tasty fruit every day.’

Like most Tibeto-Burman languages, Akha can be said to have the basic word order – agent, object, verb – but you rarely find such sentences. Usually, the clauses or sentences consist of a verb or verb concatenation followed by a verb particle (vp), sentence particle (sp), final particle (fp) or a conjunction. Then in order of frequency, object-verb, agent-verb with or without place and time words. The agent is a very loose member of the sentence, not being grammatically obligatory, and is often topicalized. Any noun or noun phrase can be – and often is – topicalized, placed either first or last in the sentence,

marked with the topic particle *à*. When clarity requires it, the agent and object may be marked for their function.

4.1 The noun phrase

The word order of the noun phrase is: noun, adjective, demonstrative pronoun, numeral, classifier:

tshó.hà jɔ-myò xhó njì yà
person good those two CLF
'Those two good persons.'

There is no marking for number on nouns. Nouns and noun phrases may be marked for function by postpositional noun particles (NP):

Noun particles

- | | |
|-----------|--|
| <i>ne</i> | agent, instrument, ablative (place from, time from) |
| <i>áy</i> | direct object, indirect object, locative/allative (place at or towards, time at) |
| <i>à</i> | topic (subject, object, time, place, clauses) |
| <i>ké</i> | object (often after verbs like 'call, name', enumeration of objects) |
| <i>jà</i> | possessor |

An inanimate subject or object is unmarked – a few verbs require their object marked though. An animate noun introduced can shift its function without being overtly expressed or marked, from object to agent for the following verbs:

*jàq.the.jà hə hy ye ó njèq ó.ne, tjhu hì ó.ne kuq áy
young.wild.boar this big be SASEQ catch SASEQ raise big SASEQ pen LOC
mà láy ni ð bi djó hɔ shá è, tjeq bjoq i má
NEG keep ADVP VP let stay try please run disappear away CONFIRM:NPAST:1*
'Catching wild young boars being this big, having raised (them) to a size, (I) ask you to try to let (them) stay without keeping them in a pen, (they) will run away.'

An animate agent of a transitive verb is marked with the NP *ne* in the past tense (marked with low tone on the sentence particle). It is not marked with an intransitive verb:

*àjòq lá mè
3sg come CONFIRM:NPAST:N1 (sp *mè*: mid tone as sandhi of high)
'He is coming.'*

*àjòq lá mè
3sg come CONFIRM:PAST:N1
'He came.'*

*àjòq àzàq dì sèq mè
3sg pig beat death CONFIRM:NPAST:N1
'He will kill a pig.'*

*àjòq ne àzàq dì sèq mè
3sg ABL pig beat death CONFIRM:PAST:N1
'He has killed a pig.'*

With inanimate nouns the NP *ne* marks an instrument:

àjòq *ne* mì-tjhe *ne* àzàq *dì* sèq *mè*
3sg ABL knife ABL pig beat death CONFIRM:PAST:N1
'He has killed a pig with a knife.'

An animate object may be marked with *áy* if clarity requires it, but usually it is unmarked:

ja-tjq *dì* *sèq* *ya* *djé*
chicken beat kill VISUAL:PAST INDIRECT.QUOTE
'The chicken was killed.'

An indirect object precedes a direct object and is marked with *áy*:

àjòq *áy* áshì *thì* *shì* *bìq* *má*
3SG loc fruit one CL give CONFIRM:NPAST:1
'I will give him one fruit.'

In the pronominal system of Akha there is a tone change from high to low in first and second persons singular when followed by the NPs *ne* and *áy*, and to low laryngealized when followed by the NP *à*:

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
1st	<i>ŋá</i> - <i>ŋà</i> - <i>ŋàq</i>	<i>ŋá</i> <i>ŋjàq</i> (exclusive)	<i>ŋá</i> <i>màq</i> (exclusive)
		<i>àdy</i> <i>ŋjàq</i> (inclusive)	<i>àdy</i> <i>màq</i> (inclusive)
2nd	<i>nó</i> - <i>nò</i> - <i>nòq</i>	<i>nó</i> <i>ŋjàq</i>	<i>nó</i> <i>màq</i>
3rd	<i>àjòq</i>	<i>àjøq</i> <i>ŋjàq</i>	<i>àjøq</i> <i>màq</i>

There is also a third person remote pronoun, *àtjhò*, 'other', 'somebody else' and two forms, *àha njàq* 'those two' and *àha màq* 'those', referring to people not present.

As mentioned before, the agent is very often omitted when the context does not require it. That is also true of the pronouns. When present they are often stressed or contrasted:

ŋá *àjòq* *áy* áshì *thì* *shì* *bìq* *má*
1sg 3sg LOC fruit one CLF give CONFIRM:NPAST:1
'I will give him one fruit.'

In the past tense, the subject pronoun is marked with *ne* and change of tone, and the object pronoun is marked with *áy*, sentence particle on low tone (mid tone as sandhi for low after another low tone):

ŋà *ne* *àjòq* *áy* áshì *thì* *shì* *bìq* *ma*
1sg ABL 3sg LOC fruit one CLF give CONFIRM:PAST:1
'I gave him one fruit.'

Possession is marked with the NP *à* (mid tone as sandhi when preceded by a low tone) after the possessor:

xhà.là *ə* *dò.mi*
tiger POSS tail
'tiger's tail'

ŋàq *ə* *àma*
1sg POSS mother
'my mother'

When first or second person pronouns modify a kinship term and when not contrasted or stressed, the low tone form without laryngealization and NP is used followed by an unprefixed kinship term:

- ŋà ma* my mother (< *àma*)
nò da your father (< *àda*)

The interrogative pronouns are:

- àshú ~ àshú yà* who
àdjè what
àgə + CLF which

Demonstrative pronouns are followed by a classifier and have a three-way distinction for distance:

- hə* here (close to me)
thə there (close to you)
xhó over there (away from both of us)

There is a fourth demonstrative pronoun, only occurring with a suffix for space and usually pronounced with an exaggerated high, prolonged vowel:

- xhü-ba ~ hü-ba* far over there

A relative clause precedes the noun and is marked with the particle *á* for non-past and *à* for past. The noun may be deleted and the classifier acts as the head:

- mí.nay lá á xhó njì yà*
yesterday come POSS those two CLF:HUMAN
'Those two persons who came yesterday.'

- tjhé kha lá á tshó.hà*
rice plant come POSS person (*á* with mid tone as sandhi of high)
'Those persons who are coming to plant rice.'

Classifiers are compulsory in Akha after a demonstrative pronoun and/or a numeral. A number of disyllabic nouns have a classifier as one of its syllables, e.g.:

- shà-áy thi áy* one spindle
ló-góq thi góq one stream

Most classifiers are used with a number of nouns, e.g.:

- yà* for human beings (= strength)
mó for animals (< *mó-do* = body)
shí for small round things, fruits (< fruit)
bjm for heaps

Nouns may be conjoined with the coordinate conjunction *hóə* between the two nouns, or with the comitative conjunction *nèqé* after the second noun:

- àma hóə àda* mother and father
àma àda nèqé mother with father

An adjective may modify a noun in two ways:

- 1 Noun + *jɔ* + adjective
ŋà-sjhà jɔ-né red fish
phi-di jɔ-pjhú white skirt

- 2 Noun + one syllable of the noun repeated + adjective

ŋà-sjhà *ŋà né* redfish
phi-di phi pjhú white skirt

There is a difference in meaning between the two structures. A noun modified with the *jɔ-* prefix denotes a temporary quality, something that happens to be so – a fish that is described as being red, a skirt that happens to be white – while the repeated structure denotes a category of the noun – a red fish as a kind of fish, a white skirt used for ceremonial purposes. If a disyllabic noun contains an affix, it is the other syllable that is repeated. A monosyllabic noun can be modified in both ways, i.e. with or without the prefix but without repetition, with the same semantic difference.

As in many Southeast Asian languages, there are a great number of intensifiers of adjectives in Akha. Some of them have a clear meaning, of the type ‘black as burnt’, but most of them only occur with one specific adjective and have no independent status. There is no obvious pattern of phonetic resemblance between the adjective and its intensifiers. Syntactically, they behave like adverbs and are marked with a final *-i*, but often occur last in a sentence, as an afterthought. Here are some examples with explanation, if any (for more details, see Hansson 1982):

<i>naq pyq i</i>	very black (black + burn)
<i>nàq kha i</i>	very deep
<i>pjhú dà i</i>	very white
<i>phý dù i</i>	very blue
<i>yò lò i</i>	very big
<i>bú tjøq i</i>	very clean

Comparison is structured as adjective (without the *jɔ-* prefix) + *dzè* (marking degree), and in questions as (noun + coordinate conjunction + noun +) interrogative pronoun + adjective + *dzè*:

<i>ŋá hò dzè</i>	I am bigger/the biggest
<i>àshú yà hò dzè</i>	who is bigger
<i>àma hò àda àshú yà hò dzè</i>	who is bigger, mother or father

4.2 The verb phrase

The verb can be defined as a word marked by the vp *ə* in its citation form and able to be preceded by the negation *mà*. When asking ‘How do you say “eat” in Akha?’, you always get as answer *dzà ə*, i.e. the verb + the verb particle. The adjective on the other hand, has in its quotation form the prefix *jɔ-*: *jɔ-sjhý* ‘yellow’, *jɔ-hò* ‘big’. Akha thus has different morphological marking on verbs and adjectives, even though the adjective otherwise functions like a verb, i.e. it may be a predicate and be part of a verb concatenation. When the adjective is negated or part of a concatenation, the prefix disappears.

The concatenation in Akha (Hansson 1985) consists of restricted (can only occur as verb head and in the first post-head position – vr) or versatile (can occur as verb head and in more than one post-head position – vv) verbs and verbal auxiliaries with a clearly defined internal word order. The verbal event must be only one with regard to acting persons, affected object, time and place, and describe one single state or action. The negation can only precede the whole chain and can never be placed inside it. The verbal auxiliaries cannot function as verb head (vh). They can be placed either before the verb head – pre verb head auxiliaries (auxv) – or after – post verb head auxiliaries (vaux) – and most of them

are versatile. The Akha verb concatenation is very post-head oriented. There are only four pre-head auxiliaries – words for ‘cause’, ‘must’ and ‘further’/‘again’. There can be up to three pre-head auxiliaries in a row and – so far – up to five post-head positions:

<i>phá</i>	<i>ja</i>	<i>bi</i>	<i>khú</i>	<i>xòq</i>	<i>nèq</i>	<i>làq</i>	<i>djí</i>
auxv	auxv	auxv	vh	vr	vaux	vaux	vv
further	must	cause	call	back	for/towards	CISLOCATIVE	all
'further must cause (somebody) to call (them) all back towards me'							

<i>mà</i>	<i>phá</i>	<i>bi</i>	<i>khú</i>	<i>xòq</i>	<i>nèq</i>	<i>làq</i>	<i>djí</i>	<i>tjhó</i>
NEG	auxv	auxv	vh	vr	vaux	vaux	vv	vaux
not	further	cause	call	back	for/towards	CISLOCATIVE	all	able
'further not able to cause (somebody) to call (them) all back towards me'.								

The verbs may have a change of meaning when changing position, the post-head position having a more abstract or grammaticalized meaning (for more details see Hansson 1985). Some examples:

			<i>Verb head</i>	<i>Post verb head</i>
	<i>lá</i>		come up	up, start, become, develop, come to vh
	<i>í</i>		go down	down, become, go down to vh
	<i>hɔ</i>		see	have experienced, try to vh
	<i>sèq</i>		kill	vh to death, very
EX	<i>lá</i>	ž		come!
	vh	FP		
	<i>djø</i>	<i>lá</i>		come to rob
	vh	VV		
	<i>tshè</i>	<i>lá</i>		start to shine
	vh	VV		
	<i>já-ŋe</i>	í		go to the fields
	Obj	vh		
	<i>sjhý</i>	í		go yellow, wither
	vh	VV		
	<i>phá</i>	<i>ló</i>	<i>i</i>	further topples over
	auxv	vh	VV	
	<i>mà</i>	<i>hɔ</i>	<i>xhm̩</i>	not allowed to look
	NEG	vh	vaux	
	<i>mà</i>	<i>kha</i>	<i>hɔ</i>	have never planted
	NEG	vh	VV	
	<i>ázàq</i>	<i>sèq</i>	ə	to kill a pig
	o	vh	VP	
	<i>dì</i>	<i>sèq</i>	ə	beat to death
	vh	VV	VP	
	<i>ná-hà</i>	<i>sèq</i>	ə	ask very much
	vh	VV	VP	

When several post-head verbal auxiliaries occur in one concatenation, the one with the most abstract meaning is placed farthest away from the verb head. They can be grouped together according to their semantic contents as follows – with a few examples of each group:

Benefactive	<i>nè</i>	for non-first person
	<i>nèq</i>	for first person

Directional	<i>lāq</i>	towards first person (cislocative)
	<i>lèq</i>	towards non-first person (translocative)
Attitudinal	<i>mòq</i>	want
	<i>náy</i>	dare
Potential	<i>tjhó</i>	able
	<i>xhm̄</i>	allow

The internal order is: benefactive, directional, attitudinal, potential:

<i>mà</i>	<i>djáŋ</i>	<i>nè</i>	<i>mòq</i>	don't want to do (it) for him
NEG	vh	vaux	vaux	
		Ben	Att	

Adjectives may occur unprefixed in concatenations, but usually only as verb head or in the first post-head position

<i>sjhý i</i>	go yellow, wither (< <i>jɔ-sjhý yellow</i>)
vh vv	
<i>de à</i>	beaten wet (by rain) (< <i>jɔ-à wet</i>)
vh vv	

Adverbs, preceding the verb phrase and the negation, are either always marked with *-i* or *-é*, as e.g. *íshàŋ i* ‘very’, *átyq tèqé* ‘a little bit’ or constructed by repeating a *jɔ-* prefixed adjective or a verb phrase marked with *-é*:

<i>jɔ-khɔ.khɔ</i>	<i>é</i>	<i>i</i>
fast	P	go
‘go fast’		
<i>hɔ.hɔ</i>	<i>é</i>	<i>djó</i>
look	ADVP	stay
‘stay looking’		
<i>lɔq thà lɔq thà</i>	<i>é òq</i>	<i>lé</i>
peel leave	peel leave	P return
‘return peeling and leaving (leaves on road)’		

A sequence of events where the agent must be the same but there may be different objects is marked with the verb particle *ž* (longer form: *ž ne*) after each verb or concatenation:

<i>àjɔq já-ŋe i ž, tjhé kha ž, òq lá ž, hɔ dzà dzà ž . . .</i>
sub place vh vp obj vh VP vh vv VP obj vh VP
3sg field go SASEQ rice plant SASEQ return come SASEQ rice eat SASEQ
‘He went to the fields and planted rice and then he returned home and ate and . . .’

A sentence may be connected to another sentence through a conjunction, placed immediately after the verb phrase and the subject and object may of course change. They may follow a verbal predicate as well as a nominal one, (some exceptions) replacing a verb particle or a sentence particle. The most common ones are:

<i>náa</i>	if/when
<i>ŋε</i>	while
<i>lɔ-nní -à</i>	if
<i>míe</i>	as, because
<i>xɔqi</i>	although, even if

5 SENTENCE PARTICLES

In Akha, most of the sentence particles can follow both a nominal and a verbal predicate. They carry quite a lot of information: first person prime mover contra non-first person prime mover, expectation contra non-expectation or statement vs reaction, kind of knowledge, i.e. know for sure (first person: *e*, *ma*, non-first person: *a*, *me*) or infer from seeing (*ya*, *yaa*), hearing (*mia*) or feeling (*nja*), excluding (first: *le*, non-first: *la*) or emphasizing (first: *the*, non-first: *tha*) an assumption. All these occur with high tone for non-past (sandhi mid tone after high tone) low tone for past (sandhi mid tone after low). Following these comes four particles of prediction, only with high tone, expressing fear (*á*) or threat (*njá*), doubt (*mí*) or certainty (*má*) (for more details, see Egerod and Hansson 1974; and Hansson 1996a). Some examples:

Plain statement:

<i>àkhà é</i>	I am an Akha (plain statement)
<i>àkhà á</i>	he is an Akha
<i>dzà é</i>	I will eat
<i>dzà á</i>	he will eat
<i>dzà e</i>	I have eaten (mid tone on sp sandhi of low)
<i>dzà a</i>	he has eaten

Confirmation or positive reaction:

<i>àkhà má</i>	yes, I am an Akha (confirmation, reaction)
<i>àkhà mé</i>	yes, he is an Akha
<i>dzà má</i>	yes, I will eat/let's eat
<i>dzà mé</i>	yes, he will eat
<i>dzà ma</i>	yes, I have eaten
<i>dzà me</i>	yes, he has eaten

Nominal predicates and plain *jɔ-* prefixed adjectival predicates occur for semantic reasons only with non-past sp. In questions the sp is reversed, i.e. the sp that is going to be used in the answer is used in the question + an interrogative final particle: *ló* (asking for information) or *là* (asking for confirmation):

Questions:

<i>àkhà má ló</i>	Are you an Akha?
<i>àkhà má</i>	Yes, I am.
<i>àkhà mé ló</i>	Is he an Akha?
<i>àkhà mé</i>	Yes, he is.
<i>lá mè ló</i>	Has he come?
<i>lá mè</i>	Yes, he has come.

Four sentence particles mark how the speaker got his knowledge and whether it is a plain statement or a reaction (high or low tone for non-past and past): *njá* ‘feeling’ and plain statement, *ya* ‘seeing’ and plain statement, *mia* ‘hearing’/‘feeling and reaction’, *yá* ‘seeing and reaction’. Some contrasting examples:

<i>hò mèq mia</i>	I am hungry (it's dinner time).
<i>hò mèq é nja</i>	I am hungry (although I had dinner a while ago). (the vowel of a preceding verb on low or mid tone must be repeated with a high tone before the sp <i>nja</i>).
<i>akhà ya</i>	He is an Akha (I can see it).

- àkhà ñáa* Yes, he is an Akha (I can see it).
àkhà mía He is an Akha (I can hear him speaking).
àkhà á njá He is an Akha (I feel it).

These four sentence particles may be negated and the implication is that the speaker does not know the reason for what is happening:

<i>àjòq</i>	<i>áy</i>	<i>dì</i>	<i>ə</i>	<i>àshú yà</i>	<i>mà</i>	<i>ñá</i>
3sg	LOC	beat	VP	who	NEG	VISUAL:NPAST

‘I don’t know/can’t see who is beating him.’

This is a very special and infrequent function and does not warrant us to regard those sentence particles as verbs. They cannot take a verb particle or be part of a concatenation as a verb can.

6 NOUN INCORPORATION

In Akha, there are numerous examples (so far around 300) of a three-syllable noun-verb construction where the second and the third syllables are the same. A negation or a classifier may be inserted between the second and the third syllables, proving the structure to be basically noun-verb, mostly object-verb but also instrument-verb, place-verb. The noun cannot be marked with a noun particle but it is rather a kind of noun incorporation. The third syllable, i.e. the verb, may function as the verb head of a concatenation or as a versatile verb following an inserted verb head. (For more details see Hansson 1996b.) This construction occurs also in Lahu (Matisoff 1973) but only rarely.

The first syllable may have its origin in a disyllabic noun or be a free noun. The first syllable may also be the prefix *a-* or only occur together with the second syllable. The verb seems always to be able to occur by itself together with other nouns.

6.1 The first syllable comes from the first or second syllable of a disyllabic noun

This is the most frequent structure:

From *sjhà-djí* meat e.g.

- | | |
|-----------------------|------------------------|
| <i>sjhà bjím bjím</i> | divide meat into piles |
| <i>sjhà bjeq bjeq</i> | chop raw meat |
| <i>sjhà the the</i> | hunt |
| <i>sjhà tjhé tjhé</i> | pickle |

From *ùdù* head

- | | |
|-------------------|--------------------------|
| <i>ùdzàn dzàn</i> | wear a turban |
| <i>ùtjhó tjhó</i> | wear a woman’s headdress |
| <i>ùxòq xòq</i> | wear a hat |

From *àlàq* hand, arm

- | | |
|----------------------|-------------------|
| <i>làq bè bè</i> | wear a ring |
| <i>làq duq duq</i> | wear a bracelet |
| <i>làq tsàq tsàq</i> | hold hands |
| <i>làq tuq tuq</i> | clench one’s fist |

These first syllables can only function as a noun together with the second syllable, or otherwise in their original disyllabic form. The negation and numeral + classifier can only be placed in front of the third syllable.

<i>sjhà bjm̄ ò bjm̄ bjm̄</i>	divide into four piles of meat
<i>ùdzàŋ thi kháj dzàŋ</i>	wear one turban
<i>làq bè thi bè mà bè</i>	not wear one ring

6.2 The first syllable may function as a monosyllabic noun

About 20 per cent of the sample belongs to this category. The most clearly prevalent and most versatile is *tjhé* ‘rice’, ‘paddy’ as, for example:

<i>tjhé boq boq</i>	pop rice
<i>tjhé dì dì</i>	thresh paddy
<i>tjhé dyq dyq</i>	soak rice
<i>tjhé lɔq lɔq</i>	dry paddy in the sun
<i>tjhé ná ná</i>	for rice to become diseased
<i>tjhé kha kha</i>	plant rice
<i>tjhé phi phi</i>	carry rice on back
<i>tjhé thàŋ thàŋ</i>	pound rice
<i>tjhé tsoq tsoq</i>	transplant rice

Also frequent are combinations with *m̄* ‘heaven’; ‘sky’ and *hò* ‘cooked rice’:

<i>m̄ djè djè</i>	to thunder
<i>m̄ mjɔq mjɔq</i>	to lighten
<i>m̄ tjìq tjìq</i>	get towards evening
<i>hò thàŋ thàŋ</i>	make rice cakes
<i>hò tjhu tjhu</i>	wrap rice in leaves
<i>hò tsuq tsuq</i>	make rice balls
<i>hò dzà dzà</i>	eat rice

6.3 The first syllable is the prefix *a-* (in various tones)

The prefix *a-* constitutes about 16 per cent of the sample. The second syllable may be part of another disyllabic noun but does not occur as a free noun, but of course as a verb. Some examples:

<i>ádzyq dzyq</i>	itch	<i>ájseq jeq</i>	to flower
<i>ám lm̄</i>	measure in fathoms	<i>ápjhà pjhà</i>	have a fever
<i>áshì shì</i>	produce fruit	<i>átjhù tjhù</i>	to bud
<i>áðq ðq</i>	to burp		

Sometimes the second syllable occurs both with prefix *a-* and in other combinations:

<i>ábòq bòq</i>	make a pattern
<i>shán bòq bòq</i>	write (wood +)

where *ábòq bòq* has a more general meaning and *shán bòq bòq* a more specific one.

6.4 The first syllable occurs as a free noun only together with the second syllable

<i>ny dag dag</i>	be filthy, dirty
<i>tháŋ djø djø</i>	to rob

<i>bɔ̄ duq duq</i>	plant a bush
<i>gý dəq dəq</i>	make war (<i>dəq</i> ‘cut’)

When the first syllable is a free noun it may occur without the second syllable being repeated. There is then a contrast in whether the stress is on the noun or on the verb:

<i>hò dzà dzà</i>	eat (action of eating)
<i>hò dzà mà dzà</i>	not eat
<i>hò dzà</i>	eat rice (have a meal)
<i>hò mà dzà</i>	not eat rice (but may be something else)

When the first syllable is not a free noun, this contrast can only surface by retrieving the disyllabic noun:

<i>sjhà tjhé tjhé</i>	pickle (meat)
<i>sjhà tjhé mà tjhé</i>	not pickle (meat)
<i>sjhà mà tjhé</i>	(not possible)
<i>sjhà-djí mà tjhé</i>	not pickle meat (but may vegetables)

7 FINAL PARTICLES

The most frequent final particles are:

<i>᳚</i>	imperative
<i>lé</i>	direct quotation
<i>djé</i>	indirect quotation
<i>ló</i>	questions – asking for information
<i>là</i>	questions – asking for confirmation

The negative imperative *thà*, is placed before the verb:

<i>lá ᳚</i>	come!
<i>thà lá</i>	don’t come!

REFERENCES

- Egerod, Søren and Hansson, Inga-Lill (1974) ‘An Akha conversation on death and funeral’, *Acta Orientalia* 36: 225–284.
- Hansson, Inga-Lill (1982) ‘A phonological comparison of Akha and Hani’, *LTBA* 7.1: 63–115.
- Hansson, Inga-Lill (1985) ‘Verb concatenation in Akha’, in Graham Thurgood, James A. Matisoff and David Bradley (eds) *Linguistics of the Sino-Tibetan Area: The State of the Art*, Pacific Linguistics, Series C, No. 87, Canberra: Australian National University, 287–309.
- Hansson, Inga-Lill (1989) ‘A comparison of Akha, Hani, Khatu and Pijé’, *LTBA* 12.1: 6–91.
- Hansson, Inga-Lill (1991) ‘The language of Akha ritual texts’, *LTBA* 14.2: 155–167.
- Hansson, Inga-Lill (1996a) ‘The interplay between the verb particle “᳚” and the sentence particles in Akha’, *LTBA* 19.1: 65–76.
- Hansson, Inga-Lill (1996b) ‘Object-verb in Akha: the ABB structure’, *LTBA* 19.1: 77–95.

- Hansson, Inga-Lill (2013) ‘Grammatical aesthetics of ritual texts in Akha’, in Jeffrey P. Williams (ed.) *The Aesthetics of Grammar: Sound and Meaning in the Languages of Mainland Southeast Asia*, Cambridge: Cambridge University Press, 280–288.
- Lewis, Paul (1968) *Akha–English Dictionary*, Data paper 70, Linguistics series III, Ithaca, NY: Cornell University Southeast Asian Program.
- Matisoff, James A (1973) *The Grammar of Lahu*, Berkeley, CA: University of California Press.

CHAPTER FORTY-SEVEN

LISU¹

David Bradley

1 INTRODUCTION

Lisu has nearly a million speakers: over 600,000 in northwestern Yunnan and southwestern Sichuan provinces, China; over 300,000 in northeastern Burma; 45,000 in northern Thailand and 2,700 in northeastern India, with a few elsewhere overseas. The very closely related Lipo language is spoken by a further 250,000 in north central Yunnan and nearby in Sichuan in China; about a quarter of these are also officially classified as Lisu nationality in China, while the rest are included in the Yi nationality. In some of the older literature the Lipo are referred to as Eastern Lisu. Most of nearly 600,000 Lolopo, members of the Yi nationality to the west of the Lipo area in north central Yunnan, speak another language fairly similar to Lisu and Lipo; another closely related major language is Lalo, widely spoken by members of the Yi nationality west of Lolopo and east of most Lisu in west central Yunnan. Speakers of Lipo suggest some intelligibility between Lipo, Lisu and nearby varieties of Lolopo, but Lisu and Lolopo speakers cannot understand Lipo; there is no contact with Lalo. Some other languages, particularly Lamu (Bradley 2004) and Mique (Bradley and Bradley 2002), are also close to this complex.

The documented migrations of the Lisu took them from western central Yunnan up the upper Salween and upper Lancang or Mekhong rivers from about 1750, thence into extreme northern Burma from about 1880 and further into Tirap District of Arunachal Pradesh, India from the early 1940s; the easternmost Lisu are widely spread north of the Jinsha or upper Yangtse river in northern Yunnan and southwestern Sichuan, north of the Lipo. Southward migration took the Lisu into Shan State of Burma from the early nineteenth century, and into Thailand since 1918. The largest concentration of Lisu is now in Nujiang Lisu Autonomous Prefecture and Weixi Lisu Autonomous County of northwest Yunnan and in Kachin State of Burma.

Lisu is divided by its speakers into three subgroups. The northernmost /lo³⁵ ne⁴⁴/ ‘black Lo’ (called /lo³⁵ wu³⁵/ ‘northern Lo’ by other Lisu, *Hei Lisu* (‘Black Lisu’) in Chinese and Black Lisu in English) are in the northwest of Yunnan and into extreme northern Burma and India. The central /eq⁴⁴ eq⁴⁴/ (sometimes called Hua Lisu in Chinese and Flowery Lisu in English) are in western Yunnan and adjacent areas of northeastern Burma. The Southern Lisu /lo³⁵ si³³/ ‘yellow Lo’ are in parts of Shan State in eastern Burma, the extreme southwest of Yunnan and in Thailand; they are sometimes called Lisaw or Lishaw in Shan, Burmese and Thai. The subgroups correspond to the three main dialects within Lisu proper. These are mutually intelligible, but with some initial difficulty. The fourth cluster of Lisu dialects in southwestern Sichuan and adjacent areas of northern Yunnan speak a somewhat divergent dialect not further discussed here, and are not in contact with the main Lisu concentration further west.

There are several current Lisu and Lipo orthographies. The first was for Lipo, and is a Pollard² script which is still used by the numerous Lipo Christians in north central Yunnan. The second is the Fraser script, named after the Scottish missionary James Outram

Fraser, whose initial work eventually led to the conversion of the majority of the Lisu in China to Christianity.³ This is a composite mainly based on the Flowery Lisu dialect, and is universally used by Lisu Christians of whatever dialect background for writing and in some formal spoken domains. The third is a syllabary using Chinese characters and other symbols devised by Wang Renpo, a Black Lisu traditional religious leader from Weixi County, in the 1920s; it is hardly used nowadays. The fourth is a 1950s romanization representing the Black dialect following the principles of Chinese *pinyin*.⁴ This was used fairly widely in China from the late 1950s, but is now rapidly losing ground to the Fraser script. Hope (1976) prepared a Thai-based script for Southern Lisu, but no one uses this.

2 PHONOLOGY

Lisu has the inventory of consonants given in Table 47.1.

There is one central Flowery Lisu subvariety, the first which Fraser encountered, that distinguishes a complete retroflex series /tʂ tʂʰ dʐ ʂ z/ from the alveopalatals /te teʰ dz e/ before /a/. Retroflexes and alveopalatals are also distinguished in Lipo and in Lolopo, as in many other related languages; this distinction is made in the Fraser orthography, redundantly for most speakers, by writing the alveopalatals with a following *Y* and the retroflexes without the *Y*; so *CY* represents /teə/ and *C* represents /tʂə/. In most varieties, the retroflexes are in complementary distribution with the alveopalatals: the retroflexes occur before the fricative syllabic [ɻ], before back vowels /u/ and /o/, and before medial /w/; the alveopalatals occur before /a/ and elsewhere including before /i/. In Southern Lisu retroflexes before /a/ are replaced by alveolar affricates or fricatives /ts tsʰ dz s z/; in Black Lisu they usually have a following medial /w/ plus /a/. Some varieties which completely lack other retroflex versus alveopalatal affricate and fricative distinctions have /z/ (written *R*) contrasting with /j/ (written *Y*) in a few words before /a/ and /o/. Other dialects have /z/ or /j/ corresponding to these /z/.

The palatal nasal occurs mainly before /a/ and /i/, and in some dialects before /o/. Words which have initial /n/ before /i/ in many varieties of Lisu often palatalize this to initial [ɲ] in some varieties, but not the reverse: many words are /ni/ in all dialects. A non-nasal [h] occurs variably in the final imperative marker [hq²¹]; in /h~/-initial words with other tones or vowels the entire syllable including the initial is nasal. The Fraser orthography provides a separate letter, inverted *G*, to represent oral [h].

In some northwestern Flowery Lisu subdialects there is a contrast between /w/ and /v/ before /u/ and of /j/ and /j/ before /i/; in Black Lisu and Southern Lisu words with /wu/ and /ji/ in such Flowery Lisu subdialects are pronounced as /o/ and /e/ with no initial. In most varieties without these contrasts, where /w/ occurs it is [v] before front vowels and [w] in most other places, but [vʷ] before /o/. The /f/ is also marginal; it occurs exclusively

TABLE 47.1 CONSONANTS OF LISU

p	t	ts	te	(tʂ)	k
p ^h	t ^h	ts ^h	te ^h	(tʂ ^h)	k ^h
b	d	dz	dz	(dʐ)	g
m	n		ɲ		ŋ
f		s	e	ʂ	x
			z	j	ʐ
w	l				y

before /u/ or /y/ in native Lisu words, as discussed later. There are also a few areas where /f/ and most /w/ are replaced by alveolar or velar fricatives.

Table 47.2 shows the vowel system that Lisu has in most areas.

There are also fricative syllabics [h] and [ɺ] which occur after homorganic affricates and fricatives: [tsɻ ts^hɻ dzɻ sɻ zɻ] and [tʂɻ tʂ^hɻ dʐɻ ʂɻ ʐɻ] respectively. In one small central subdialect of Flowery Lisu described in Yu (2007) there are also sequences of [tsji ts^hi dzji sji zji], so these could be treated there as an additional set of initial consonant clusters, or /ɻ/ could be recognized as an additional vowel, with [tsji] and so on are treated as /tsi/, and [tsɻ] and so on then treated as /tsɻ/. The varieties on which the various Lisu orthographies are based do not have sequences like [tsji], so the orthographies treat the [ɻ] as an allophone of another vowel: of /y/ in the Fraser orthography, and of /i/, following *pinyin*, in the Chinese orthography. All dialects have sequences of alveopalatals plus /i/, so here a distinction must be made; either a vowel /ɻ/ is needed, occurring only after retroflexes; or five distinct retroflex initials are needed; the alveopalatal fricatives are needed in any case as vowels other than [ɻ] occur after them in Flowery Lisu. The Fraser orthography recognizes five retroflex initials in any case, since it also distinguishes these initials before /a/; it treats [tsɻ] and so on as sequences of retroflex plus /y/, and of course [tei] and so on as alveopalatal plus /i/. The Chinese orthography also distinguishes retroflexes, partly in order to represent the many Chinese loans, and treats [ɻ] as another allophone of /i/, again following *pinyin*. But Black Lisu, which it represents, does not have retroflexes before /a/ in native words, except allophonically before medial /w/ and back vowels.

The Fraser treatment of fricative syllabics is possible because of the restricted distribution of the front rounded vowels. These are largely absent in some Black Lisu dialects, and are thus not distinguished in the Chinese orthography: all /ø/ are replaced by /e/ written *ei*, and all /y/ are replaced by /u/ written *u*. In other dialects there is some variation between /ø/ and /e/; for example the demonstrative ‘this’ is Flowery Lisu /tʰø³³/, but usually /tʰe³³/ in Southern and Black Lisu; ‘child’ is Flowery, Southern, and sometimes Black Lisu /za²¹ nø³³/, but also sometimes Black Lisu /za²¹ ne³³. Thus the Chinese script is somewhat underdifferentiated in failing to distinguish the front rounded vowels, but this reflects their marginal status in that dialect. In most kinds of Lisu, /y/ occurs only after bilabial, alveolar stop, nasal, lateral, /w/ and /h~/ initials, and never after any affricate, fricative, or velar initials.

Black Lisu fronts /a/ to /e/ after alveopalatals and after bilabials plus medial /j/ and palatalizes the bilabials to the corresponding alveopalatal; thus, for example, *JY*: /dza⁴⁴/ ‘cold’ becomes /dze⁴⁴/ and *BY*: /bjɑ²¹/ ‘bee’ palatalizes to /dze²¹/.

The high and mid-back unrounded vowels /u/ /y/ do not occur syllable-initially; they are preceded by a voiced velar fricative [ɣ] if no other consonant is present. The velar fricative is otherwise only found before /a/, and some dialects including Southern Lisu replace some or all of these [ɣɑ] sequences with [ja], thus eliminating the need for postulating a /ɣ/ at all. The Fraser script writes both [w] after other consonants and the syllable [yu] with inverted *L*, and [ɣ] after a consonant and the syllable [ɣɣ] with inverted *D*. The distinction between /u/ and /y/ is highly marginal, but those literate in the Fraser script

TABLE 47.2 VOWELS OF LISU

i	y	w	u
e	ø	ɣ	o
ɛ			
		a	

are now accustomed to it. The Chinese *pinyin* orthography does not distinguish the two; both are written with *e*.

In Southern Lisu, the vowel /u/ conditions a slight labiodental friction after a consonant, [f] after voiceless initials and [v] after voiced ones; for example, /tu/ is [t^fu] and /du/ is [d^vu]. The vowel /e/ conditions a preceding [j]; for example /de/ is [dje]. All the mid vowels, /e ə ɔ ɔ:/, can variably be raised to [i ɿ ɯ u], most often in Southern Lisu. Often a raised /e/ retains its glide, becoming [ji] rather than [i], thus giving a contrast of /i/ [i] versus /e/ [ɿ] ~ [ji]. In the case of underlying /u/ versus /o/, in Southern Lisu these are still distinguished by the labiodental: /u/ is [f'u] or ['u], while /o/ is [o] or [u]. Velar stops followed by front vowels are variably palatalized to alveopalatals; thus, /ge/ can be [gje], [gji], [gi], [dzje], [dzji] or [dzi]. This palatalization of velars is least frequent in Black and most frequent in Southern Lisu; for example the very frequent verb ‘go’ and postverbal directional marker for motion away is underlyingly /ge³³/, but is consistently written as *ggi* /gi³³/ in the Chinese *pinyin* for Black and as *JE*, /dze³³/ in written Flowery Lisu, reflecting the most frequent pronunciations.

Several initials are in complementary distribution with initial glottal stop. There are no native syllables with initial glottal stop before /u/ or /y/ and [f] occurs in native Lisu words only before /u/ or /y/; so [f] and [ɸ] are in complementary distribution. Similarly, initial [y] before /u/ and /v/ is in near-complementary distribution with [ɸ], since these vowels also do not occur initially, other than in baby talk, onomatopoeia, or loanwords. This could completely eliminate the need for an /f/ phoneme in native words and greatly reduce the distribution of /y/, though /f/ is reinforced by loanwords with other vowels. As we have seen, the Fraser orthography treats [y] as allophonic preceding an initial vowel with /u/ and /v/, but /f/ is written distinctly, with inverted *J*. The Chinese orthography writes /f/ as *f*, but writes initial /y/ before /a/ as *e*; thus ‘rock’ [ya⁵⁵] is written *eal*. As noted above, [yu] and [yv] are both written in this orthography with *e* representing both initial and vowel, as in [yu⁵⁵] ‘left’ written *el*.

The phonetic value of the /y/ vowel differs in different areas. In Southern Lisu, it is made with lip compression; Hope (1973a) therefore describes it as /wi/. Most of Flowery Lisu has front [y] with lip projection rounding, but the degree of rounding differs from area to area. In Black Lisu it is central rounded [ɯ] in some varieties, but completely merged into /u/ in others.

Lisu has six tones. They are high level /⁵⁵/, rising /³⁵/, higher-mid creaky /⁴⁴/, mid non-creaky /³³/, low falling /²¹/, and low falling with final glottal stop /²¹/'. The six tones are numbered 1 to 6 in this order in Fraser (1922), and this is also their alphabetical order in the Flowery Lisu orthography.

The rising tone is the least frequent, and has a restricted distribution: it is rare with voiced stop or affricate initials, as is the /⁵⁵/ tone; and infrequent with nasal initials. Rising tone is however usual in the second syllable of reduplicated baby talk words which have a sequence of /²¹/ + /³⁵/ tones, such as /ma²¹ ma³⁵/ ‘rice’, and also occurs after voiced stop initials in a few very frequent words, especially post-head modal elements such as /dq³⁵/ ‘be at’/‘durative’ and in Chinese loanwords.

The phonetic value of the sixth tone is given as [⁴²] in most Chinese phonetic descriptions, but this is wrong. There is considerable variation between tones 3 and 4; there are dialectical differences in their distribution, and within a variety there is also some fluctuation. In some dialects the creaky tone 3 is higher in pitch than the non-creaky tone 4; in others, such as Southern and some kinds of Flowery, tone 3 has mid pitch, [³³], and the contrast is entirely one of creaky vs non-creaky. The usual Chinese transcription represents it just as [⁴⁴], taking the pitch difference as primary and the phonation difference

as secondary. In addition, there are various combined versions of two tones which occur when two syllables fuse; this is most frequent for postverbal elements in clause-final position with a low or low stopped tone in second position, giving rise to a variety of surface falling tones.

The basic syllable structure of Lisu is c(g)vt. All dialects have clusters of velar plus medial /w/ before /a/, and nearly all have bilabial or /l/ (but not labiodental) plus medial /j/ before /a/; /lja/ also occurs in Chinese loans. Some Black Lisu dialects have retroflex affricate and voiceless fricative plus medial /w/ clusters before /a/; like the retroflex allophones before [tʃ], here also these do not contrast with the alveopalatals which never occur before /w/ in any dialect. Most dialects have clusters of bilabial plus medial /j/ before /ø/, but these are replaced by bilabial plus /i/ or bilabial plus /e/, i.e. [je], in some areas including most of the northern area.

One interesting example of an orthography creating a marginal distinction is in the very frequent postverbal continuous marker of Flowery and Southern Lisu written TY, or more frequently as fused TY⁵ since the first page of the very first 1915 catechism. It is pronounced /teɑ³³/ by non-literate Lisu, but has come to be pronounced /tja³³/ or /tja³⁵/ by Lisu Christians in reading or careful speech. It is the only such /t + j/ cluster, and appears to have arisen due to a spelling pronunciation. Presumably it should have been written CY, as it is pronounced by non-literate speakers.

Speakers of dialects with quite different canonical sequences are happy to follow the now-traditional orthography. For example, though all bilabial plus /ja/ sequences are replaced by palatoalveolar plus /e/ in most kinds of Black Lisu, thus /bjə²¹/ ‘bee’ usually becomes /dze²¹/ and /mja²¹/ ‘many’ becomes /ne²¹/, Black Lisu Christians still write these as BY: /bjə²¹/ and MY: /mja²¹/ . The same convention applies in the Chinese orthography: so ‘bee’ is written *bbiat* and *notjjait* and likewise ‘many’ is *miat* and never *niait*. Similarly, in Fraser Lisu script, Southern Lisu and Black Lisu speakers make every effort to write the retroflex initials before /a/ which they never pronounce as such in their own speech, unless they are trying to speak literary Lisu.

3 MORPHOSYNTAX

3.1 Nominals

Like most other Tibeto-Burman (TB) languages, Lisu is mainly verb-final, though one or more noun phrases (NPs) can be postposed for focus, most frequently in Southern Lisu. The order of NPs before the verb is to some extent pragmatically determined, as Hope (1973b) has pointed out; however the most frequent NP order for transitive verbs is temporal–subject–object–locative–verb, and for ditransitive verbs is temporal–subject–indirect object–direct object–locative–verb. Temporal phrases, which are most often clause-initial, are NPs in origin and in form. Lisu is a zero anaphora language; that is, any NP may simply be omitted if it is clear from the context. Overt pronouns for syntactic core NPs are relatively infrequent in context. Thus some clauses contain no overt NPs at all, and many contain fewer than the number of slots provided by the verb.

Postpositions marking nominal case exist, but are not obligatory for syntactic core cases. They come last in the NP. One very frequent postposition is the marker /te⁵⁵/ in Black and Flowery Lisu and /le⁵⁵/ in Southern Lisu. It is almost obligatory on the causee subject of a causative or causativized verb, quite frequent on a DAT or BEN, and less frequent on an ACC. For most speakers it may only occur on one NP in a sentence, normally on the DAT or BEN if present.

-
- (1) ji⁵⁵ ŋwa³³ te⁵⁵ nu³³ (?te⁵⁵) t^ho²¹yu²¹(*te⁵⁵) gu²¹ ts¹³³
 3sgANIM 1sg-OBJ 2sg-OBJ book-(OBJ) give CAUS
 'He makes me give you a book.'

Much less frequent is the subject marker /le³³/, also pronounced [ne³³] in some places, homophonous with the COMIT and INST. Though it is more frequent on transitive subjects and has thus been called an ergative by Hope (1973b), it also occurs on subjects of agentive intransitive verbs such as motion verbs. Postpositions on non-core NPs are much less optional. These include the locative/allative /kwq⁴⁴/ 'at'. In some varieties these have alternative forms with a voiced initial; there are also shortened forms such as Southern Lisu /wg⁴⁴/ and /q⁴⁴. Various locational nouns can precede the LOC, such as /k^hu²¹/ 'inside' and so on.⁶ The 'from' ABL postposition is LOC /kwq⁴⁴/ plus /teo⁵⁵/ in Southern Lisu, probably a Chinese loan from 从; other dialects use /kwq⁴⁴ be³³/ here. There are also some dialect forms which are completely different, such as Southern /to⁵⁵/ for INST, which is thus distinct from COMIT /ne³³/ and AGENT /le³³/ in that variety.

Possessor NPs precede the NP possessed (see (2a) below). The genitive postposition /yu²¹/ is very infrequent between a possessor pronoun and an NP; in many varieties of Lisu this is ungrammatical. It is somewhat more frequent between an NP possessor and the possessed NP, as in (2b), especially if the possessor has more than one constituent, as in (3). It is obligatory if there is no overt possessed NP. Another way of expressing possession is with /ji⁵⁵/, as in (2c). This cannot occur after a non-third person pronoun; so it can be viewed as noun + 'his/her' + noun; of course this is homophonous with the formative prefix before bound nouns discussed later.

- (2) a. /ŋwa³³ ba³⁵ ba²¹/ b. /a⁵⁵ mo²¹ (ji⁵⁵) mu⁵⁵/ c. /ji⁵⁵ mu⁵⁵/
 1sg father horse PREF tail PREF tail
 'my father' 'horse's tail' 'tail' or 'his/her tail'
 (3) /t^ho³³ mq⁴⁴ ŋwa³³ yu²¹ (ŋq³³)/
 this CL 1sg GEN COP
 'This one is mine.'

There is also a topic marker /ŋq⁴⁴/. It is very frequent on subjects, especially transitive ones, but also occurs on NPs in other case roles. This topic marker can follow another case marker, such as on OBJ, a LOC and so on. Another less widespread topic marker which is also mainly used with objects is /nq²¹/.

Demonstrative and quantifier phrases, if any, follow the head noun. The demonstrative phrase consists of a demonstrative plus either a nominal postposition, or if no nominal postposition is present, then an obligatory nominalizer /mq⁴⁴/. The quantifier phrase consists of numeral plus classifier. When a demonstrative is present, the number 'one' may be omitted. Curiously, and unlike closely related languages, the quantifier phrase is surrounded by the demonstrative phrase: the demonstrative precedes and the nominalization follows the quantifier phrase, as in (4). As in many other TB languages, NPs with only a demonstrative plus nominalizer or numeral plus classifier or both, without a head NP, are frequent.

The demonstrative system (Bradley 2003) marks relative height as well as distance: /t^ho³³/ 'this', /go³³/ 'that (same level)', /no³³/ 'that (higher level)' and /dzo³³/ 'that (lower level)'. Some varieties also have a form /a⁵⁵ t^ho³³/ 'that (near you)'; all five are within visible distance. Southern Lisu has an anaphoric series in which each demonstrative also has an anaphoric form with /a⁵⁵/ prefix; it does not use /a⁵⁵ t^ho³³/ in the meaning 'that (near you)'. There are dialect differences in the demonstratives for greater distance or beyond vision, but they include modified forms /dzo⁵⁵/, /no³³/, /go⁵⁵/ of the three 'that' forms with

high tone, additional forms with voiceless initials /t^o⁵⁵/ and /k^o⁵⁵/, as well as various reduplicated forms. All demonstratives also occur fused with the LOC /kwa⁴⁴): /t^ha³³/ ‘here’, /gwa³³/ ‘there’, /nwa³³/ ‘up there’ and so on; these are used as independent NPs.

The numbers follow the usual TB pattern, with higher round numbers used as classifiers of a lower number.

- (4) /a⁵⁵ na²¹ go³³ t^hi²¹ hja^{~33} li⁵⁵ ts^hl³³ ku⁴⁴ ma³³ ma⁴⁴/
dog that one hundred four ten nine CL NMZR
‘those 149 dogs (on the same level)’

In Lisu, even in counting, numbers are usually given with a classifier; if nothing else, the general classifier /ma³³/. There are tone sandhi forms of the numbers /sa⁴⁴/ ‘three’, /li⁴⁴/ ‘four’ and /ku⁴⁴/ ‘nine’; when the following classifier has tones /³³/ or /⁴⁴/, a sandhi tone [⁵⁵] occurs for ‘four’ and ‘nine’, and usually [²¹] or sometimes [²¹] for ‘three’. However this sandhi tone does not occur before the general classifier /ma³³/. The number ‘one’ has an alternative unaspirated form /ti⁵⁵/ which occurs in ‘11’ /ts^hl³³ ti⁵⁵/ and as a noun suffix ‘sole, only’. The number ‘ten’ has an unaspirated form in ‘20’ /ni²¹ ts^hl³³/, as in many Ngwi languages. For more discussion of the history of numerals in Lisu and related languages, see Bradley (2005).

The general classifier /ma³³/> is nearly homophonous with the nominalizer/relativizer /mq⁴⁴/ discussed later. It classifies many types of nouns including most animals. Humans are classified with /zo⁴⁴/ in Flowery, the corresponding form /jo⁴⁴/ in Black, and /zo³³/ or more frequently /wa²¹/ in Southern Lisu; the latter is also the pronoun plural marker in most varieties of Lisu. There is the usual range of semantically determined classifiers, many of which are also bound nominals like /sl²¹/ ‘fruit’ and /dzl³³/ ‘tree’. Some non-bound nouns like /hl³³/ ‘house’ are also used to classify themselves; in such a case the head noun is often omitted.

- (5) /(hl³³) li⁵⁵ hl³³/
(house) 4 CL
‘four houses’

An interesting and crosslinguistically rare set of dyadic two-syllable classifiers, which is also found in many other Ngwi languages (Bradley 2001), is used for groups of family members. The forms are /md⁵⁵ lq²¹/ for mother and children, /pa⁵⁵ lq²¹/ for father and children, and /pi⁵⁵ ll²¹/ for grandparent or grandfather and grandchildren. A fourth form /mi⁵⁵ ll²¹/ for grandmother and grandchildren is used by some Lisu. These occur without a head noun or can follow a pronoun. Eastern Lisu has a more elaborate system, see Yu (2007: 118).

- (6) /sq⁴⁴ pa⁵⁵ lq²¹/
three CL (father + children)
‘father and two children’ (also ‘father, mother and one child’ or ‘father, child and child’s spouse’)

There are various forms of attributive constructions corresponding to relative clauses in Lisu, the use of which is determined by the category of the verb. For stative verbs, the relative has clause-final /(a⁵⁵) mq⁴⁴/; this is pre-head or more frequently post-head and then precedes the demonstrative plus quantifier in the NP.⁷ The /a⁵⁵/ is often omitted. This also occurs as a general complementizer for other verb types. Unlike the following constructions, an /(a⁵⁵) mq⁴⁴/ construction with a verb usually has an overt nominal head, though such clauses can also occur headless. A sequence of /mq⁴⁴/ plus /md³³/, where the first is

the relativizer/complementizer and the second is the general classifier in a quantifier phrase with ‘one’ omitted, is fairly frequent with stative verbs, whether headed or headless. When the verb is preceded by the intensifier / α^{21} k^hui⁵⁵/ ‘very’, there is no marking of embedding.

- (7) /h̄i³³ p^hu³³ (α^{55})mq⁴⁴ t^hø³³ (t^hi²¹) h̄i³³/
 house white REL this one CL
 ‘this (one) house which is white’
- (8) /p^hu³³ (α^{55}) mq⁴⁴ (h̄i³³) t^hø³³ (t^hi²¹) h̄i³³/
 white REL house this one CL
 ‘this (one) house which is white’
- (9) /p^hu³³ (α^{55}) mq⁴⁴ (t^hi²¹) ma³³/
 white REL one CL
 ‘this white one’
- (10) /h̄i³³ α^{21} k^hui⁵⁵ p^hu³³ t^hø³³ (t^hi²¹) h̄i³³ mq⁴⁴/
 house very white this one house NMZR
 ‘this (one) house which is very white’

Stative extentive verbs of positive dimensional extent (‘many’, ‘long’, ‘far’, ‘big’ and so on) have additional sandhi forms; see Bradley (1995). These include nominal forms; question forms with a prefix /q⁴⁴/ parallel to the substance question words; diminutive forms with a following /e⁵⁵/; and a variety of reduplicated augmentative forms. Six to nine such verbs pattern this way in different varieties of Lisu; Lipo varieties have eight to ten.

- (11) /wu²¹/ /wu³³/ /q⁴⁴ wu³³/ /q⁴⁴ wu³³ e⁵⁵/ /q⁴⁴ wu⁵⁵ wu³³ be³³/ / α^{21} wu⁵⁵ wu⁵⁵ be³³/
 ‘big size how big? small very big extremely big’

Verbs occur in nominalizations with one of a variety of clause-final markers indicating the case function within the nominalized clause. These include /su⁴⁴/ ‘(animate) subject’, /g²¹/ ‘(non-human, usually) object’, /gu³³/ ‘locative’, /du³³/ ‘thing for’ and /t^he²¹/ ‘temporal’ as well as the general /mq⁴⁴/. Such a nominalization must have a case role in the main clause and may be followed by a case-marking postposition, except that the temporal nominalizations are normally just temporal and not further marked other than by a possible topic marker. These nominalizations are frequent without a head NP and thus look like complements; but they sometimes occur with a head NP, which usually follows; if it precedes, it cannot be distinguished from an internally headed nominalization unless there is a pause. What does not occur is a nominalization which both contains (or is preceded by) its head and is also followed by its head. The nominalized clause may contain modal and directional postverbal elements, but not aspect markers.

- (12) /(ŋwa³³) la³³ su⁴⁴ te⁵⁵ t^ho²¹yw²¹ gu²¹ o⁴⁴/
 1sg come NOMZR OBJ book give PFV
 ‘I gave the person who came/comes a book.’

In addition to such nominalizations, there are also compounds consisting of a head noun, sometimes bound, plus a verb, usually stative, without embedding marking. The degree of lexicalization varies.

- (13) /h̄i³³ p^hu³³/
 house white
 ‘white house’

- (14) /mi³³ nɛ⁴⁴/
field black
'field in steeply sloping valley (which is often in the shade)

As in other related languages, there are four-syllable equivalents of many two-syllable nominal compounds and nominalizations. These are used in formal speech and in song language.

A further kind of nominal form is with the prefix /ji⁵⁵/ . This is extremely frequent as a dummy prefix making bound nominal elements into full nouns. This is homophonous with the third person animate pronoun, and may also mean ‘his/her NOUN’. It is also frequently used as an abstract nominalizer of a stative non-extensive verb.

- (16) a. /ji⁵⁵ mjo⁴⁴/ b. /ji⁵⁵ mjo⁴⁴/ c. /ji⁵⁵ nɛ⁴⁴/
 3sgANIM name PREF name PREF black
 'his/her name' 'name' 'blackness; the colour black'

Nouns may be conjoined without overt marking, or this may be marked by /be³³/ in Flowery and Black Lisu. A string of conjoined nouns may be followed by a single /be³³/; this is most likely if there are three or more. In principle this /be³³/ may also occur between the nouns and after them, but this is infrequent and rejected by many speakers. Conjunction of nouns referring to animates can also use COMIT/INST /le³³/ 'with', medially or finally. The corresponding nominal conjunction in Southern Lisu is /tee³³/.

The prounoun forms /*ŋwa*³³/ '1sg prounoun' /*nu*³³/ '2sg prounoun' /*ji*⁵⁵/ '3sg animate prounoun' are used to refer to humans. The system includes a third person remote 'other' category, for which the form is /*su*³³/; it is of course not an accident that this is similar to the animate nominalizer suffix /su⁴⁴/ . For second and third person forms there is an optional plural suffix /*wa*²¹/; these also sometimes fuse to single-syllable forms [na²¹] from /*nu*³³ *wa*²¹/ '2pl prounoun' and [*ja*⁵⁵] from /*ji*⁵⁵ *wa*²¹/ '3pl animate prounoun'. This suffix does not occur after /*su*³³/ or /*ŋwa*³³/ . There is an inclusive collective first plural form /*zo*²¹/ or sometimes /*zo*²¹/ in Flowery, /*jo*²¹/ in Black, and /*a*²¹ *zo*²¹/ in Southern Lisu. The form /*ŋwa*³³ *nu*²¹/ or its presumably more archaic song form equivalent /a⁴⁴ *nu*²¹/ is the first plural exclusive; note the tone difference from the 'you' prounoun.

A final pronoun is the reflexive. There are dialect differences, but the most widespread construction is an invariant reflexive pronoun /tei⁵⁵ te¹¹a²¹/ from Chinese source which can be in any case role other than subject, but is most frequently direct object; it may be followed by a postposition marking case. This form is absent from Lipo, which has a construction more frequently encountered outside Lisu: pronoun + reflexive marker + pronoun. Lisu also has such a construction, but with a non-reflexive emphatic meaning: pronoun + /da³³/ + pronoun ‘by oneself’. Another reflexive possibility is pronoun + /go³³ dø²¹/ ‘body’ (or other body part noun) + pronoun, but only where the action physically affects one’s body. The invariant /tei⁵⁵ te¹¹a²¹/ form may also be preceded, followed or even preceded and followed by a pronoun.

Other than optional marking on pronouns or in a quantifier phrase, Lisu has no number marking on nominals. There is also no requirement for gender marking, except that there are male and female suffixes which can be attached to nouns, and to agentive verbs to form nominals: /ma⁴⁴/ 'female' and /p^hq²¹/ 'male', as in /ma⁵⁵ ma⁴⁴/ 'female teacher' and

/ma⁵⁵ pʰa²¹/ ‘male teacher’, both from the verb /ma⁵⁵/ ‘teach’. These are alternatives to the gender-neutral /su⁴⁴/ nominalizer, but unlike it they can also be added to nominals, even some of those which already contain /su⁴⁴/, as in /li⁴⁴ su⁴⁴ ma⁴⁴/ ‘Lisu woman’, /li⁴⁴ su⁴⁴ pʰa²¹/ ‘Lisu man’; but not in Lipo where these are /li³³ ma⁴⁴/ and /li³³ pʰa²¹. Note that the female suffix is homophonous with the nominalizer, but these have a different tone from the general classifier /mc³³/.

3.2 Verbs

There is no agreement with nominals of any kind marked on Lisu verbals. There is also no switch reference marking, but of course there are a range of markers of non-final clauses: sequential, simultaneous, conditional and so on. Reflexive is also marked entirely by the nominals. A morphosyntactically distinct passive does not exist, though the order of NPS is pragmatically determined. There is a rich evidential and epistemic system which differs between dialects (Bradley 2010).

Lisu has a complex system of pre- and post-head modals of various types, many of which have been grammaticalized from verbs which still exist. In addition, sequences of verbs are quite frequent as nominals are often absent. There can be ambiguity between a sequential verb + verb meaning and a verb + modal or modal + verb meaning.

Pre-head elements are far fewer in number. If there is a negative, it precedes the pre-head element. Some are homophonous both with a main verb and with a post-head element; for example /wq⁴⁴/ is a verb meaning ‘get’, a pre-head meaning ‘have the opportunity to’ and a post-head meaning ‘must’.

Post-head elements are quite numerous. They include a wide range of modal meanings: /da³³/ ‘can’, /wq⁴⁴/ ‘must’, /teo³⁵/ ‘need to’, /teo⁴⁴/ ‘experiential’, /da³⁵/ ‘durative’, /tea³³/ ‘continuous’, /ji⁴⁴/ ‘try to’ and so on. The post-head modal ‘want to’ has two syllables: /ni³⁵ sl³³/; the first syllable is the noun ‘heart’. Some do have full verb uses, such as /wq⁴⁴/ ‘get’, /da³⁵/ ‘be at’, /tea³³/ ‘stay’ and /ji⁴⁴/ ‘look at’; others do not.⁸ Some post-head modals may be directly negated and/or questioned, others may not. The two-syllable ‘want to’ has negation between its elements. There are also various sequences of post-head modal elements which are less frequent; some speakers prefer to avoid these by using complement constructions or breaking them into a sequence repeating the main verb with one post-head in a non-final clause and with another post-head form in a final clause. In context, such as when answering a question, some post-head elements can be used in their modal meaning as the sole verbal element in a sentence.

As in many TB languages, there are pairs of etymologically related intransitive and transitive forms, the latter related to a Proto-TB *s- prefix usually described as a causative prefix. In Lisu, the causative/transitive forms have voiceless initials and the intransitive forms have voiced initials, as in /dzo⁴⁴/ ‘fear’, /teo³⁵/ ‘scare’ or /dza²¹/ ‘eat’, /tsq⁵⁵/ ‘feed (an animal)’. There are often also tonal and semantic differences, as seen here. In addition to these, there is a post-head causative /tsq⁴⁴/, as in (1), (17b) and (18b), with various alternative dialect forms. Furthermore, /gu²¹/ from the verb ‘give’ is also used post-head in a benefactive meaning,⁹ as in (18a), sometimes with causative-like semantics, as in (18c). As a pre-head it is a permissive causative, as in (17a). A lexical causative/transitive can also be followed by a postverbal causative, as in (17d). While it is infrequent, for many speakers it is possible to combine a pre-head benefactive, main verb and causative or a main verb, causative, and a post-head benefactive in these orders, as in (17e).

- (17) a. /gu²¹ dza²¹/ b. /dza²¹ ts₁⁴⁴/ c. /tsa⁵⁵/ d. /tsa⁵⁵ ts₁⁴⁴/ e. /gu²¹ dza²¹ ts₁⁴⁴/
 p.CAUS eat eat CAUS feed feed CAUS p.CAUS eat CAUS
 'give' 'to eat' 'eat' 'cause to eat' 'feed' 'feed' 'cause to feed' 'cause to give to eat'

- (18) a. /ma⁵⁵ gu²¹/ b. /ma⁵⁵ ts₁⁴⁴/ c. /ma⁵⁵ ts₁⁴⁴ gu²¹/
 teach BEN teach CAUS teach CAUS BEN
 'teach' 'cause to teach' 'cause to teach' (for someone's benefit)

Directionals are the last of the postverbal elements before the aspect and epistemic/evidential/imperative markers, and often follow other postverbal elements. The reciprocal, Flowery /lɛ²¹ k^ho³³/, Black and Southern Lisu /lɛ²¹ xo³³/, forms part of the directional system, the rest of which is derived from motion verbs. These include /la³³/ 'motion towards/come', /ge³³/ (with various raised and /or palatalized dialectal forms including Flowery and Southern Lisu [dze³³]) 'motion away'/go' and the slightly less transparent forms /je³³/ 'motion away' and /le³³/ 'change of state' derived from motion verbs with cognates outside Lisu but no longer used as main verbs in Lisu.

Sentence-finally most non-negated sentences contain an aspect marker, either imperfective declarative /q⁴⁴/; perfective /q⁴⁴/; or habitual/generic /lo⁴⁴/. In Southern and some subvarieties of Flowery Lisu, the intransitive perfective is very often preceded by a directional (/le³³ q⁴⁴/, /je³³ q⁴⁴/), but in other varieties such sequences are not preferred.

Some equational copula sentences have no overt copula, just a sequence of two nominals, as in (19a). Alternatively, there can be an overt final equational copula /ja³³/ which becomes obligatory when negated (19b); when the copula is present and not negated, this can be followed by the HAB /lo⁴⁴/, or with the copula plus perfective fused to /jo³³. These copula forms cannot be combined with other pre- or post-head verbal elements. Another copula-like form is /p^hjø²¹/ 'become', but unlike /ja³³/ it is a normal verb and can be combined with modal, directional and non-habitual aspect postverbal elements, as in (20).

- (19) a. /jwa³³ ma⁵⁵ ma⁴⁴ (ja³³)/ b. /jwa³³ ma⁵⁵ ma⁴⁴ ma²¹ ja³³/
 1sg teacher-female COP 1sg teacher-female NEG COP
 'I am a (female) teacher.' 'I am not a (female) teacher.'

- (20) /ji⁵⁵ ma⁵⁵ ma⁴⁴ wa⁴⁴ p^hjø²¹ tea³³ a⁴⁴/
 3sgANIM teacher-female get.to become CONT IMPFV
 'She is getting to become a teacher.'

There are various types of nominalization discussed above and below; these convert a clause ending in a verbal into a nominal or part of a nominal by means of various postpositions. The most abstract of these is /ma⁴⁴/, which is often used as a complementizer without a head noun. Other complement types include non-final clauses embedded by a final /(a⁴⁴) be³³/, and with certain final verbs involving human action a non-final clause with a verb and then /be⁴⁴/ (the verb 'say').

3.3 Other form classes

Temporal expressions are nominals, not adverbs, in Lisu. They contain nominal elements such as the sandhi form of /ji³³/ 'day' seen in /ji⁵⁵ me³³/ 'today' or verbal elements nominalized by a clause-final /t^he²¹/. There is one very frequent intensifier adverb /a²¹ k^hu⁵⁵/ 'very' which immediately precedes a negative if any and then an intransitive verb. This cannot be regarded as a preverbal modal, since the negative precedes those. Another type

of adverb in Lisu is a small but open class of reduplicated two-syllable onomatopoeic forms. There are Southern Lisu manner adverbs with the prefix /teŋ³³/ (varying between [teŋ³³] ~ [tee³³] ~ [teɪŋ³³] ~ [tei³³]) preceding a partially or fully reduplicated two-syllable form with the same vowel and tone in both syllables; many such adverbs have two identical syllables. A moderately productive Northern Lisu adverbial pattern is a reduplicated verb plus the suffix /mu⁴⁴/. Adverbs usually occur immediately before the verb; dialect differences in adverbial forms are extreme. Apart from adverbs and exclamations, all other form classes occur within either a nominal or a verbal element. This includes demonstratives, numerals and classifiers as well as nominalizing suffixes and prefixes and NP-final case and topic post-positions within nominals; and negatives, pre- and post-head modals, directionals, aspects etc. within verbs.

3.4 Clause types

The final clause has the widest range of possibilities for postverbal elements, as discussed earlier. Among final clauses, statements most often end in the imperfective declarative /q⁴⁴/ or the perfective /o⁴⁴/. These aspect markers are usually not found in negatives or negative imperatives, which have the negative /mq²¹/ or negative imperative /t^ha²¹/, respectively, immediately preceding the verb or modal to be negated. An imperative is most frequently just a bare verb without a following marker, but various final imperative markers such as /ha²¹/ exist, and some final particles may occur after a negated verbal, most frequently /se²¹/ 'still'/yet'.

- (21) /dza³³ dza²¹ q⁴⁴/
rice eat PFV
'(I) have eaten.'
- (22) /dza³³ md²¹ dza²¹ (se²¹)/
rice NEG eat (still/yet)
'(I) have not eaten (yet).'
- (23) /dzi³³ p^hwi²¹ do³³ (ha²¹)/
liquor drink (IMP)
'Drink liquor!'
- (24) /dzi³³ p^hwi²¹ t^ha²¹ do³³/
liquor NEGIMP drink
'Don't drink liquor.'

Questions containing a question nominal ('WH word') are like other clauses; fronting of the question nominal is not necessary, though it is not infrequent as this is the most topical element in a question. These question nominals all start with the prefix /a/, with various tones: /a²¹ mq³³/ 'who?', /a⁵⁵ s²¹/ 'what?', /a⁵⁵ t^he²¹/ 'when?', /a⁴⁴ to⁵⁵/ 'where (at)?' and so on. The question word /a⁴⁴ le⁴⁴/ 'which?' /how? can occur alone, as in the frequent greeting given in (25) below, or with a following nominal element, especially a classifier. 'Where?' can be expressed by 'which?' plus LOC /kwq⁴⁴/ or various contracted forms, Southern Lisu /a⁴⁴ lq⁴⁴/ and Black Lisu /a⁴⁴ kwq⁴⁴. It can be noted that some of the second syllables of question words are homophonous with clause nominalizers of similar meaning. These and other questions may have a sentence-final question element /q²¹/ fused onto the last syllable in the clause; this is realized as a falling tone, as in (25) and (26), and final glottal stop if that syllable ends in /a/ with another tone already.

- (25) /q⁴⁴ le⁴⁴ nq³³ (q²¹)/ [nq³¹]
 how? Be Q
 ‘How are (you/things)?’ (a greeting)
- (26) /q⁴⁴ le⁴⁴ (t^{h1}²¹) ma³³ la³³ q⁴⁴ (q²¹)/ [q⁴¹]
 which? One CL come DECL Q
 ‘Which one is coming?’

Yes–no questions have a number of alternative structures which differ in frequency and grammaticality between varieties of Lisu. Alternative questions are fairly frequent: verb + negative + verb, as in (27), or for some modals verb + modal + negative + modal. As for other negatives, these usually lack most final markers such as directional, aspect and so on. Another widespread possibility is verb + /q²¹/, as in (28), the final question marker we will also see as an option for yes–no questions; this is obligatory if it is the sole marker of a yes–no question, and is most frequent in Southern Lisu. Final /la²¹/, as in (29), which is an earlier form of this final with cognates outside Lisu, is also seen in some Flowery Lisu varieties; Northern Lisu has a corresponding form / le²¹/. This can be used finally or combined with the alternative question, in which case it comes after both instances of the verb or modal, as in (30). There are some other sentence-final question markers for specific situations, such as the first person rhetorical /ne³⁵/, as in (32). Some varieties of eastern Black Lisu use a reduplicated verb to form a yes–no question, as in (31), but speakers of most other varieties find this ungrammatical.¹⁰

- (27) /dza³³ dza²¹ ma²¹ dza²¹/
 rice eat NEG eat
- (28) /dza³³ dza²¹ q⁴⁴ q²¹/ [dza³³ dza²¹ q⁴¹]
 rice eat IMPFV Q
- (29) /dza³³ dza²¹ la²¹/
 rice eat Q
- (30) /dza³³ dza²¹ la²¹ ma²¹ dza²¹ la²¹/
 rice eat Q NEG eat Q
- (31) /dza³³ dza²¹ dza²¹ – Eastern Black Lisu only
 rice eat eat
 ‘Have you eaten?’
- (32) /dza³³ dza²¹ ne³⁵/
 rice eat Q
 ‘Have I eaten?’

Comparisons are formed with /s_j⁵⁵ ma²¹ ts_j⁴⁴/ ‘above/more’ following the compared NP; so this is structurally like an NP plus postposition.

- (33) /(j)⁵⁵) nwa³³ s_j⁵⁵ ma²¹ ts_j⁴⁴ mo²¹ q⁴⁴/
 3sgANIM 1sg more old IMPFV
 ‘He/she is older than I.’

Various kinds of non-final clauses have restrictions on the elements following the verb. Non-final complement clauses include the /(q⁴⁴) be³³/ type, whose marker is homophonous with a very frequent nominal conjunction, and the /be⁴⁴/ type, whose marker is the verb ‘speak’. As we have seen, there are also a variety of rather frequent clause

nominalizations including temporals, locatives, instrumentals, objects and subjects functioning as NPs or parts of NPs.

Some constructions which are directly embedded by a conjunction in other languages must be nominalized and followed by an abstract head noun in Lisu. For example, one type of ‘because’ clause is followed by nominalizer /ma⁴⁴/ + noun /pu⁵⁵ do³³/ ‘cause’ and then the main clause with its final verb.¹¹

(34)	/mu ²¹	ha ³³	le ⁴⁴	la ³³	q ⁴⁴	ma ⁴⁴	pu ⁵⁵	do ³³ ,
	rain	fall	DIR	IMPFV		NOMZR		cause
	ne ⁴⁴	ji ³³	dza ²¹	la ³³	q ⁴⁴ /			

river flood DIR PFV
‘Because rain fell, the river flooded.’

One syntactic genre which should be mentioned is songs, which are in seven syllable lines; in most cases the first four syllables are nominal material and the last three are verbal or sometimes nominal. After each line there is another line repeating the content with alternative lexical material; see Bradley *et al.* (2008) for examples. Interestingly, the markers /q⁴⁴/ and /q⁴⁴/ do not count as a syllable in songs. Like most oral literature, songs are very repetitive, with overlap between adjacent lines. Another genre with even more distinct syntax is proverbs, which are two lines having an equal number of three or more syllables and without postverbal markers.

NOTES

- 1 The assistance of many Lisu and other colleagues and friends over the years and the support of the Australian Research Council (DP0772046 and earlier support) and of UNESCO is gratefully acknowledged. Thank you for comments goes to Lisu colleagues, the editors and Takashi Kato; of course all remaining errors are solely the author’s responsibility. Examples are from Flowery Lisu unless otherwise stated.
- 2 Samuel Pollard worked with the Miao, mainly in Guizhou; his script uses a variety of large symbols (some letters, some symbols from Pitman shorthand, and so on) for consonants; the vowel symbols are smaller, and are placed in different positions adjacent to the consonants to represent the tones: above for high tone and so on. For details of the initial development of this script starting from 1904 see Enwall (1994). Its use was extended to Lipo by Nicholls and Metcalf before 1910, with the first printed materials appearing in 1912.
- 3 Fraser devised this script working with a Karen and an American during a visit to Burma in 1914; the final version came into use in 1919. It uses upper-case letters, upright and inverted, to represent consonants and vowels; tones are written with punctuation marks following the syllable. The vowel [a] is inherent in a consonant letter not followed by another vowel. For more details of this script, see Bradley (1979: 55–65), Bradley and Bradley (1999), Bradley *et al.* (2006) and Bradley (2006).
- 4 For details of this orthography see Bradley and Kane (1981) and Bradley (1994). Briefly, consonants and vowels are written using Chinese *pinyin* values where possible. Voiced stops are written double: *bb* represents /b/, because *b* is /p/ and *p* is /p^h. Tones are written with a final consonant, such as *-l* for /⁵⁵/ and *-t* for /²¹/.
- 5 TY, is TY., fused with the following declarative marker /q⁴⁴/.
- 6 The ‘inside’ form may be related to the verbal locative nominalizer /gu³³/. This locative postposition has cognates elsewhere, including Lahu /ki²¹/, Nasu /gu³³/ and so on; in other languages it is usually just a verbal locative nominalizer, not also a nominal locative postposition as in Lisu.

- 7 There is no obvious meaning difference between preceding and following REL; speakers often reject preceding REL out of context, but these do occur, especially if the REL itself is complex.
- 8 For example, the post-head form ‘able to do skilfully’ /ku⁵⁵/, with widespread Ngwi cognates such as Lahu /pi³⁵/, has no corresponding verb form.
- 9 For the benefit of any human of whatever person.
- 10 These are the speakers who are in contact with Nosu, which also forms yes–no questions in this way.
- 11 This is exactly like the structure of various closely related languages including Lahu, which has clause + NMZR /ve³³/ + /pa³³ tɔ³³/ ‘cause’; the noun forms are also cognate.

REFERENCES

- Bradley, David. 1979. *Proto-Loloish*. Scandinavian Institute of Asian Studies Monograph Series 39. London and Malmö: Curzon Press.
- Bradley, David. 1994. *A Dictionary of the Northern Dialect of Lisu (China and Southeast Asia)*. Pacific Linguistics C126. Canberra: Department of Linguistics, Research School of Pacific and Asian Studies, Australian National University.
- Bradley, David. 1995. Grammaticalisation of extent in Mran-Ni. *Linguistics of the Tibeto-Burman Area* 18/1.1–28.
- Bradley, David. 2001. Counting the family: family group classifiers in Yi Branch languages. *Anthropological Linguistics* 43/1.1–17.
- Bradley, David. 2003. Deictic patterns in Lisu and Southeastern Tibeto-Burman. In: *Language Variation: Papers on Variation and Change in the Sinosphere in Honour of James A. Matisoff*, ed. by David Bradley, Randy J. LaPolla, Boyd Michailovsky and Graham Thurgood, 219–236. Pacific Linguistics 555. Canberra: Department of Linguistics, Research School of Pacific and Asian Studies, Australian National University.
- Bradley, David. 2004. Endangered Central Ngwi languages of northwestern Yunnan. Keynote presentation at 37th International Conference on Sino-Tibetan Languages and Linguistics, Lund, Sweden.
- Bradley, David. 2005. Why do numerals show ‘irregular’ correspondence patterns in Tibeto-Burman? Some Southeastern Tibeto-Burman examples. *Cahiers de Linguistique, Asie Orientale* 34/2. 222–238.
- Bradley, David. 2006. Lisu orthographies and email. In: *Lesser-known Languages of South Asia: Status and Policies, Case Studies and Applications of Information Technology*, ed. by Anju Saxena and Lars Borin, 125–135. Berlin: Mouton de Gruyter.
- Bradley, David. 2010. Evidence and certainty in Lisu. *Linguistics of the Tibeto-Burman Area* 33/2. 63–84.
- Bradley, David and Daniel Kane. 1981. Lisu orthographies. Working Papers in Linguistics, University of Melbourne 7. 23–38.
- Bradley, David and Maya Bradley. 1999. Standardisation of transnational minority languages: Lisu and Lahu. *Bulletin Suisse de Linguistique Appliquée* 69/1.75–93.
- Bradley, David and Maya Bradley. 2002. Language policy and language maintenance: Yi in China. In: *Language Endangerment and Language Maintenance*, ed. by David Bradley and Maya Bradley, 77–97. London: RoutledgeCurzon.
- Bradley, David, with Byabe Loto and David Fish. 2008. *Lisu New Year Song*. Chiang Mai: Actsco Publishing.
- Bradley, David, with Edward R. Hope, James Fish and Maya Bradley. 2006. *Southern Lisu Dictionary*. Linguistics of the Tibeto-Burman Area Monograph Series 4. Berkeley: STEDT.

-
- Enwall, Joakim. 1994. *A Myth Become Reality: History and Development of the Miao Written Language*. Stockholm East Asian Monographs 5. Stockholm: Institute of Oriental Languages, Stockholm University.
- Fraser, James Outram. 1922. *Handbook of the Lisu (Yawyin) Language*. Rangoon: Government Printer.
- Hope, Edward Reginald. 1973a. Selected phonological rules for Thailand Lisu. *Pacific Linguistics* A30.19–34.
- Hope, Edward Reginald. 1973b. Non-syntactic constraints on Lisu noun phrase order. *Foundations of Language* 10.79–109.
- Hope, Edward Reginald. 1976. Lisu. In: *Phonemes and Orthography: Language Planning in Ten Minority Languages of Thailand*, ed. by William A. Smalley, 125–148. *Pacific Linguistics* C43. Canberra: Department of Linguistics, Research School of Pacific and Asian Studies, Australian National University.
- Yu, Defen. 2007. *Aspects of Lisu: Phonology and Grammar*. *Pacific Linguistics* 588. Canberra: Department of Linguistics, Research School of Pacific and Asian Studies, Australian National University.

CHAPTER FORTY-EIGHT

LAHU¹

James A. Matisoff

1 INTRODUCTION

Lahu is a member of the Central Loloish branch of the Lolo-Burmese subgroup of Tibeto-Burman (TB). The Lahu people refer to themselves as *Lâhū-yâ* (*yâ* ‘son,’ ‘child’).² The Burmese, Shan, Thai, and Lao call the Lahu by a name romanized variously as Muhsur, Mussur, etc., most likely derived from Burmese *mouʔ-hsôu* ‘hunter.’ In Vietnam, the Lahu are called *Coxǔng* (also spelt Co Sung, Co Xung, or Khu Xung).³

Lahu villages are to be found over a wide area between the Salween River on the west and the Mekong River on the east: the southwest portion of Yunnan; the Kengtung area of Shan State in Burma; the northern Thai provinces of Chiang Mai, Chiang Rai, Mae Hong Son, Tak, and Kamphaeng Phet; Nam Tha Province in northwest Laos; and a few scattered areas in northern Vietnam. Lahu enjoys rather more prestige among other groups of hillfolk than the average minority language, and is often used as a lingua franca by such groups as the Akha and the Mien.

Based on purely linguistic criteria, there are two fundamental branches of the Lahu people: the *Black Lahu* (*Lahu Na*, *Lâhū-nâ?*) and the *Yellow Lahu* (*Lahu Shi*, *Lâhū-ši*). The Black Lahu (BL) are far more numerous than the Yellow Lahu (YL) in China and Burma, and are certainly the more prestigious group.⁴ In Thailand, the biggest Lahu groups are known as Red Lahu (*Musəə Deey*, *Lahu Nyi*, *Lâhū-ni*) and Lahu Sheleh (*Lâhū ſe-lé*), but these can be shown to be merely subvarieties of BL. The relatively few Lahu villages in Laos include at least two varieties of YL (*Lâhū-ši-bân-kew* and *Lâhū-ši-bân-lán*), as well as two interesting-sounding varieties called White Lahu (*Lâhū-phu*) and *Lâhū A-phû-be-le*. Next to nothing is known about the Lahu dialects spoken in northern Vietnam. BL is the dialect described in my grammar, dictionary, and other work (Matisoff 1973/82, 1988, 2006, 2009, 2011, 2014).

The 1990 census of China gives the Lahu population in China alone as 411,800. Estimates of the number of Burmese Lahu have fluctuated wildly, ranging from 40,000 to 230,000. The Lahu in Thailand have been accurately numbered at about 40,000. The small Lahu community in Laos is now thought to comprise from 8,000 to 10,000 persons, while the even smaller Lahu presence in Vietnam seems to number about 1,500. The tiny but growing Lahu population (mostly YL from Laos) of the US is approximately 800, clustered near Visalia, a farming community near Fresno, California. Thus our latest best guess at the world’s total Lahu population is about 600,000.

The oldest Lahu settlements are those of China. In the eighteenth and nineteenth centuries, the Lahu, frequently led by messianic “priest-chiefs,” gained some notoriety as rebels against imperial Chinese rule. Their movement from Yunnan to Burma dates back to the early nineteenth century, motivated partly because of pacification measures in Yunnan, but also because the richly forested, sparsely inhabited hills beckoned them ever southwards.

The Lahu in Thailand and Laos have all immigrated within the last 100 years, and many much more recently. The Lahu in Thailand have been relatively well off, with extensive trade links forged between them and the lowlanders. In present-day China, the Lahu are one of the 55 officially recognized minorities. Since 1953, they have played a major role in the local administration of Lancang Lahu Autonomous County in the far southwest of Yunnan, where Chinese and Lahu are the joint official languages, and in the affairs of adjacent Menglian County (officially the “Tai-Lahu-Lawa Autonomous County”).

2 PHONOLOGY

BL has seven tones. Five of these tones are smooth and open, pronounced without constriction; the other two are checked by a glottal stop.⁵ See Table 48.1.

Lahu syllables have a very simple structure, consisting of only an (optional) initial consonant, a vowel, and a tone. There are no syllables with final consonants, since glottal stop -ʔ is best regarded as a tonal feature.⁶ Even in syllable-initial position, only a single consonant may occur; there are no consonant clusters.⁷ BL has 24 initial consonants and 9 vowels, as shown in Table 48.2.

The five palatals /c ch j š y/ have special dental pronunciations before /i/: /ci chi ji ši yi/ → [tsʃ tshʃ dzʃ sʃ zʃ]. These dental sibilants [ts tsh dz s z] do not occur before any other vowel. Similarly, the four labials /p ph b m/ have special affricated variants before /u/: /pu phu bu mu/ → [pfw̚ pfhuw̚ bv̚ mv̚]. In these situations, /i/ and /u/ have special allophones, [ɪ] and [ʊ], which only occur after the palatals and labials, respectively.

Besides the nine simple vowels, various types of diphthongs also occur, both rising and falling; the most common native Lahu word with a diphthong is *qay* ‘go.’

TABLE 48.1 THE SEVEN TONES OF BLACK LAHU

Name of tone	Pitch	Symbol	Example
Mid	33	unmarked	ca look for
High-rising	35	/'/'	cá boil
High-falling	53	/'/'	câ eat
Low-falling	21	/'/'	cà ferocious
Very low	11	/'/'	cā feed
High-checked	54	/?'/'	câ? rope
Low-checked	21	/?'/'	câ? push; machine

TABLE 48.2 CONSONANTS AND VOWELS

Consonants (24)				Vowels (9)			
p	t	c	k	q	i	i	u
ph	th	ch	kh	qh	e	ə	o
b	d	j	g		ɛ	a	ɔ
m	n		ŋ				
f	h	š	h				
v	l	y	g(y)				

3 LAHU VOCABULARY AND WORD FORMATION

Lahu morphemes are almost always monosyllabic, though there are many polysyllabic words, compounds consisting of two roots, words containing a root plus a prefix and/or suffix, and reduplicated roots. Although grammatical gender is not much developed in Lahu, there do exist suffixes to distinguish the sex of some nouns referring to people or animals: *-pā* ‘masculine’/-*ma* ‘feminine’ (*chɔ-mɔ-pā* ‘old man,’ *chɔ-mɔ-ma* ‘old woman’); *-šē-phā* ‘masculine’/-*šē-ma* ‘feminine’ (*qhâ?-šē-phā* ‘headman,’ *qhâ?-šē-ma* ‘headman’s wife’).

Most Lahu compounds have only two or three syllables, though certain extra-long ones (especially flora and fauna names) may run to five or six, e.g. *á-lâ-mì-ší-jɔ* ‘rainbow’; *pā-pā-qu-ti-ni* ‘dragonfly’; *a-gɔ-a-lí-pē* or *na-gù-na-gá-pē* ‘spider’; *khâ?-pà-mē-cí-câ-kwì* ‘greater racket-tailed drongo’ (*Dicrurus paradiseus*).

Important Lahu prefixes include: *a-* ‘(vocative) kinship prefix,’ e.g. *a-e* ‘mother!'; *á-* ‘noun-forming prefix,’ e.g. *á-qhâ* ‘ragweed,’ *á-qhɔ* ‘home’; *ð-* ‘noun-forming prefix,’ e.g. *ð-cè* ‘a plant,’ *ð-š̄t* ‘a ball; a sphere,’ *ð-u* ‘an egg’ (<*u* ‘to lay an egg’), *ð-qð?* ‘a curved thing’ (<*qð?* ‘to bend’).

Nouns are reduplicated to achieve several semantic effects: (a) “Inclusive” reduplication signifies all the members of the class represented by the noun: *yâ-mî yâ-mi* ‘all the women’; (b) “Sequential” or “distributive” reduplication shows that the things represented by the reduplicated noun are considered one after the other: *ð-ce ð-ce* ‘pair by pair’; (c) “Indefinite” reduplication indicates uncertainty about the exact scope of reference of the noun: *té chi kilô-lô* ‘about ten kilos’; (d) “Emphatic” reduplication heightens or intensifies the meaning of the noun: *ð-le-le* ‘the very last.’ Particularly interesting are the hundreds of four-element compounds I call “elaborate expressions,” where the structure is of the form A-B-A-C or A-B-C-B, e.g.: *khâ-mu-khâ-nè* [word-high-word-low] ‘high and low pitched words; tones (of a tone-language),’ *ha-lè-ha-qâ* [spirit-warm-spirit-qâ] ‘happy and relaxed,’ *dâ-šâ-gâ-šâ* [think-easy-gâ-easy] ‘be serene and easy in one’s mind.’

Besides its basic stock of vocabulary inherited from Proto-Lolo-Burmane (PLB), and ultimately from Proto-Tibeto-Burman (PTB), Lahu contains many words borrowed from the languages with which it has come in contact. Many of the earliest loans are from Chinese (e.g. *ð-li* ‘custom,’ *šš-ci* ‘shuttle’). Most loans in the BL villages of Thailand that I studied in the mid-1960s were from Shan, since these Lahu had just crossed the border from Shan State 25 years before. Many Burmese words have also filtered down into Lahu via Shan, including a number of religious and governmental terms that ultimately derive from Sanskrit/Pali. The Lahu of Yunnan are now undergoing a massive lexical assault from Chinese. There are a few English loans in Lahu from the days of British rule in Burma (e.g. *lɔlî* ‘lorry’; ‘truck’), or because of missionary activity (e.g. *kɔmītî* ‘committee,’ *nò?* ‘note’ (in music)).

Aside from borrowing words outright, Lahu keeps pace with the modern world by creating neologisms out of its own lexical resources, or by creating novel blends of parts of foreign words with native morphemes, as illustrated by several of the compounds with *š̄t* ‘round object’ as their final element. Sometimes these are new combinations of native Lahu morphemes: *ð-mî-š̄t* ‘hand grenade’ (fire ball); *ð-fî-qð-š̄t* ‘stomach ulcer’ (stomach ball); *cí-ðî-š̄t* ‘tonsils’ (spit-spheres). Sometimes they are blends of *š̄t* with Shan or Thai roots, e.g. *mâ?-pâw-š̄t* ‘coconut’ (1st element < Thai *məphráaw*); *šé(n)-š̄t(a)* ‘jewel’; ‘precious stone’ (b) ‘lightbulb’ (1st element < Shan *s’ey* ‘precious’); or blends with Burmese morphemes (via Shan), e.g. *dâ?-š̄t* ‘lightbulb’ (1st element < Bs, ult. < Pali/Skt *dhâtu*

‘element’; ‘primitive matter’; *yā-cū-śī* ‘weight’ (< Bs *ra-ju* ‘scales’). Most recent are blends with English syllables: *bó-śī* ‘ball’ (first syllable < Eng).

4 CLAUSE STRUCTURE

Lahu, like the vast majority of TB languages, is verb-final. In Lahu clauses the VP is king, since a clause need contain no NPs at all to be complete:

<i>mɔ-chwɛ</i>	<i>ch̄e</i>	<i>tù</i>	<i>ve</i>	<i>yō</i>
ADV	V	PV	PU	PU
‘(X) will stay quite a while.’				

In isolation a large number of translations are possible (‘We’ll stay . . .’, ‘They’ll stay . . .’, ‘People will stay . . .’, etc.), though the context will make clear what is meant. If absolutely necessary a noun may be added for contrast or clarity, but there is no particular noun or pronoun which is felt to be “understood.” As this example illustrates, a VP may contain adverbs before the verb, and/or particles after it. The verbal nucleus of the VP may itself consist of a single verb (as here) or may contain as many as five verbs in direct juxtaposition. The following simple sentence contains four NPs before the VP:

<i>yàʔ-ni</i>	<i>ŋà-hi té qhā? 5</i>	<i>šā</i>	<i>qhɔ-qhe-phɔ̄</i>	<i>gà?</i>	<i>cā</i>	<i>e</i>	<i>tù</i>	<i>le</i>
NP1	NP2	NP3	NP4					
TIME	TOPIC	OBJECT	PLACE					
‘Where shall our village go off to hunt today?’								

Verb finality goes along with relatively free order of NPs before the verb; although this sentence is quite natural as it stands, other orderings of these four NPs are possible to convey slightly different emphases. In general, NPs indicating time and place tend to come before those referring to participants in the verbal event. Interrogative NPs tend to occur right before the verb. In the above example NP4 refers to place, but since it is interrogative (‘where?’) it appears right before the verb. The VP of this sentence contains a verbal nucleus of two verbs (*gà?* ‘hunt’ and *cā* ‘eat’), followed by two verb particles, *e* ‘motion away from,’ *tù* ‘future; unrealized action’; the clause ends with the unrestricted particle *le* ‘question marker.’

5 FORM CLASSES

Lahu has only three major form classes: nouns, verbs, and particles. Minor classes include numerals, classifiers, adverbials, conjunctions, and interjections. The numerals are a closed set of morphemes which can simply be listed; classifiers always occur after a numeral. Nouns can then be defined as words which can be modified or counted by a numeral-plus-classifier. By this definition, pronouns and demonstratives are considered subclasses of the nouns. Verbs are defined as those words which can be negated by the adverb *mā*. By this criterion, Lahu “adjectives” are just a subclass of the verbs. Adverbials must precede the verb they modify. Particles are bound morphemes with abstract grammatical functions. Even though they cannot occur alone in a phrase, they are considered to be separate words, not inflectional endings, and are written with spaces before and after them. The dozens of Lahu particles may be conveniently divided into several subtypes: (1) Noun particles (pn) only occur after nouns: e.g. *thà?* ‘object marker’; (2) Verb particles (pv) only occur after verbs: e.g. *tù* ‘irrealis’; (3) Unrestricted particles (pu) may occur after either nouns or verbs. The pns may be further subclassified according to the

types of clause in which they appear: (a) Non-final unrestricted particles (pufn) occur only in non-final clauses: e.g. *qo* ‘if’; *thɔ̄* ‘even,’ ‘also’; *l̥e* ‘because’; (b) Final unrestricted particles (puf) occur only in final clauses: e.g. *lâ* ‘yes–no question marker,’ *hé* ‘possibility.’ Several particles may occur in a row in an NP, VP, or at the end of a clause. Unrestricted particles always follow PNs or PVS.

6 ARGUMENT-VERB RELATIONS

When a noun meets a verb with no intervening particle, the grammatical relationship between them must be deduced from the inherent semantic features of the noun and verb themselves. The most important of these relationships include:

- (a) direct object + verb, e.g. *ጀ cā ve* ‘to eat rice’; *nâ? tâ? ve* ‘to carry a gun.’⁸ Lahu does have an object-marking particle, *thâ?*, but it is not much used after inanimate nouns except for special contrast or emphasis, since it is already clear that they are not the initiators of the action;
- (b) indirect object + verb, e.g. *pɔ-khû (thâ?) phu tân ve* ‘offer money to the priest.’ The particle *thâ?* is especially common after indirect objects, since these are typically human, and might otherwise be interpreted as the initiators of the action: *pɔ-khû phu tân ve* ‘the priest offers money’;
- (c) topic + verb, e.g. *i-kâ? gò ve* ‘water is cold’;
- (d) instrument + verb, e.g. *i-kâ? chî ve* ‘wash with water’;
- (e) location + verb, e.g. *á-qhɔ chê ve* ‘stay at home,’ *i-kâ? pā-tî? ve* ‘sink into water’;
- (f) ‘tied noun’ + verb: a few verbs are so tightly bound to a particular preceding noun that the two words form a sort of compound: *i-kâ? hé ve* ‘bathe’ (‘water-bathe’), *i-kâ? ší ve* ‘be thirsty’ (‘water-thirsty’). The verbs *hé* and *ší* never occur without the preceding noun *i-kâ?* ‘water’;
- (g) purpose + te ‘do/make,’ e.g. *ú-gê te ve* ‘use for a pillow’;⁹
- (h) physical characteristic + te, e.g. *ጀ-qhè? te ve* ‘have ridges’ (‘do ridges’), *ጀ-dì te ve* ‘be lumpy.’

6.1 Final unrestricted particles (puf)

The final unrestricted particles operate on whole sentences. They may be divided into several subclasses:

- (1) Declarative: *yò*, *ጀ*, *l̥o*, *ጀ* and *l̥o* show more emotional involvement than the neutral (and most common) *yò*, e.g. *Lâhū-yâ yò* ‘He’s a Lahu.’ *Lâhū-yâ ጀ* ‘He’s a Lahu, of course.’ *Lâhū-yâ l̥o* ‘Gee, he’s a Lahu!’
- (2) Dubitative: *hé*, *nè-ጀ*, e.g. *Lâhū-yâ hé* ‘I guess he’s a Lahu.’ *câ ò nè-ጀ* ‘I suppose he’s eaten by now.’
- (3) Persuasive: *mē* (sense of urging, often requesting assent): *qha-bû? cā mē* ‘Please eat your fill, won’t you?’
- (4) Interrogative: Lahu uses four interrogative pufs to form various kinds of questions:
 - (a) *lâ* ‘Yes–no questions’: e.g. *ጀ cā ò lâ* ‘Have you eaten already?’ Yes–no questions may be disjunctive: *nò qay lâ mâ qay lâ* ‘Are you going or not?’ (*lit.* ‘You go? You not go?’);
 - (b) *l̥e* ‘request for assent; tag questions’: e.g. *ጀ cā ò l̥e* ‘You’ve eaten, haven’t you?’;

- (c) *nā* ‘rhetorical or indirect questions’: e.g. *yô kâ? qaī tû nā* ‘I wonder if he’ll go too?’;
- (d) *le* ‘substance questions,’ used at the end of sentences that contain a specific interrogative noun or adverbial: *a-šu le* ‘Who is it?’; *à-thò?-ma le* ‘What is it?’¹⁰
- (5) Quotative: *cé* indicates that the speaker is reporting something at second hand, repeating what someone else has told him/her. Storytellers tend to use it in almost every sentence, but it is usually best left untranslated: *à-śwè thâ ð-châ nî ġâ cò ve cé* ‘Once upon a time there were two friends (it is said.)’ *cé* may be ordered differently with respect to other final particles, with change of meaning. When *cé* appears in the same clause as an interrogative particle (*lâ, le*), it marks a quoted question; but according to whether *cé* follows or precedes the interrogative particle the sentence is interpreted as a direct or an indirect quoted question:
- nò ve lâ cé* He said, ‘Is it yours?’ / *nò ve cé lâ* Did he say it was yours?
- (6) Interjectory (to show emphasis or emotion): *è ?, ma, vâ, nê, yâ, lè?, qô?-ma, qô?-lè?*.

In rapid colloquial Lahu these particles appear in practically every sentence, sometimes several in a row in a single clause: *yâ?-tɔ-phâ kâ? mâ šî ò yâ ma nê* ‘She’s absolutely shameless about it now!’ (lit. ‘She does not even know the way of shame any more.’)

6.2 Compound sentences and non-final unrestricted particles (punfs)

These particles appear at the end of non-final clauses to indicate the nature of their semantic relationship to what follows in the sentence, e.g.:

- (a) *le* ‘conjoining’; ‘suspensive’
*yô tê ni qha-gâ yâ?-qɔ jû le ## hâ jâ ve yô*¹¹
he all day road walk punf tired very puf
'He walked on the road all day, and is very tired.'
- (b) *pa-tɔ* ‘causal’
yô tê ni qha-gâ yâ?-qɔ jû ve pa-tɔ ## hâ jâ ve yô
'Since he has been walking all day on the road, he's very tired.'
- (c) *thâ* ‘concessive’
yô tê ni qha-gâ yâ?-qɔ jû thâ ## mâ hâ še
he all day road walk punf NEG tired yet
'Although he's been walking all day on the road, he isn't tired yet.'
- (d) *thâ* ‘temporal’
yô tê ni qha-gâ yâ?-qɔ jû thâ ## hâ jâ ve yô
'When he's been walking all day on the road, he's very tired.'
- (e) *qo* ‘conditional’
yô tê ni qha-gâ yâ?-qɔ jû qo ## hâ jâ ve yô
'If/whenever he walks on the road all day, he's very tired.'

7 THE NOUN PHRASE

7.1 Personal pronouns

Lahu does not distinguish number with common nouns, but pronouns (and proper names) can take the plural suffix *-hi*, or dual suffixes like *-hí-mà* or *-hí-nè*. There is also an impersonal third person pronoun which cannot be pluralized: *šu* ‘remote or contrastive third person’; ‘they’; ‘others.’¹²

<i>ŋà</i>	1sg	<i>ŋà-hi</i>	1pl	<i>ŋà-hí-mà</i>	1dl
<i>nò</i>	2sg	<i>nò-hi</i>	2pl	<i>nò-hí-mà</i>	2dl
<i>y᷑</i>	3sg	<i>y᷑-hi</i>	3pl	<i>y᷑-hí-mà</i>	3dl

The inclusive and exclusive distinction in pronouns is alien to the Lolo-Burmese branch of TB to which Lahu belongs. The egalitarian nature of Lahu society is reflected by the lack of pronouns that make distinctions of politeness, or establish relative rank or social distance.

7.2 Numerals and classifiers

A Lahu NP that is quantified has three parts: the head noun, a numeral, and a classifier, in that order: *chɔ tê gā* ‘a/one person’; *pâ-vî nî khe* ‘two civet cats.’ There are 11 numerals, the numbers from one to nine, the interrogative ‘how many?’, and a Shan-derived word for ‘several’ (10, 100, 1,000 and other ‘round numbers’ are classifiers, not numerals):

tê ‘one’; *nî* ‘two’; *šē* ✕ *šē?* ‘three’; *᷑* ‘four’; *ŋâ* ‘five’; *kh᷑?* ‘six’; *š᷑* ‘seven’; *hí* ‘eight’; *q᷑* ‘nine’; *qhâ-nî* ‘how many?’; *lây* [< Shan; cf. Thai *lāaj*] ‘several.’

A numeral must virtually always be followed by a classifier. Only in counting and doing arithmetic may the numerals occur alone; but even here, the general classifier *mà* is often used. Lahu classifiers may be divided into several types:

- (a) echo-classifiers: some nouns take a copy of themselves as classifier, e.g. *y᷑ tê y᷑* ‘a house’ (‘house one house’); *qhâ? nî qhâ?* ‘two villages’;
- (b) special classifiers: many nouns require special classifiers which place them in a particular semantic category, or mark them as having a particular shape, e.g. *gā* for living human beings; *khe* for animals; *kâ* for places; *cè* for plants; *pê?* for fields; *qô?* for books or papers; *qh᷑* for elongated objects; *câ?* for stringlike objects; *š᷑* for round objects;
- (c) measure classifiers: e.g. *i-kâ? tê lî?* ‘a/one liter of water’; *lâ tê khê* ‘a/one cup of tea’;
- (d) time classifiers: these usually occur without a head noun (they are auto-classifiers), e.g. *tê ha-pa* ‘one month’; *᷑ ni* ‘four days’; *hí qh᷑?* ‘eight years’;
- (e) group classifiers: these appear only with the numeral *tê* ‘one’; ‘whole,’ e.g. *Lâhû tê phâ* ‘all the Lahu’; *᷑-qâ tê ñì* ‘a bunch of buffalo’; *Hê?-pâ tê ca* ‘a (married) couple of Chinese’;
- (f) general classifier *mà*: this may be used instead of most auto- or special classifiers. To use it for humans is rather pejorative. The only other classifier that compares with *mà* in generality is *c᷑* ‘kind’: *y᷑ nî mà* ‘two houses’; *yâ?-q᷑ tê mà* ‘a/one road’; *nâ?-chî tê c᷑* ‘a/one kind of medicine’;
- (g) round-number classifiers: all round numbers are classifiers. The way to say ‘ten’ is *tê chi*, i.e. ‘one ten-unit,’ just as ‘500’ is *ŋâ ha* ‘five hundred-units.’ The higher round numbers are borrowings from Shan: *hé(n)* ‘thousand’; *mâ(n)* ‘ten thousand’; *šé(n)* ‘hundred thousand’; *lâ(n)* ‘million.’¹³

7.3 The object noun particle *thà?*

As an independent noun prefixed by *᳚-*, or a constituent in noun compounds, Lahu *thà?* (< PTB **l-tak* ✕ **g-tak* ‘ascend’; ‘above’) means ‘upper surface’; ‘top part.’ As a noun particle, *thà?* (often reduced to *hà?* or even *à?*) has developed into an object marker. In this function it is used sparingly, only where clarity demands or emphasis is required. When both direct and indirect objects are present, *thà?* will follow the indirect object. This is because indirect objects are typically human, so that an explicit marker is sometimes required to exclude an agentive interpretation. Thus without *thà?*, the following sentence could be interpreted as ‘I have given (someone) that book?’:

li? chi ḷà **thà?** pî tā ve yò
book this 1sg OBJ give PV PU PUF
(Someone) has given me that book.'

In complex sentences, *thà?* may also mark a whole clause which functions as the object of another clause:

{y᳚ q᳚? la ve} **thà?** ḷà d᳚.łɔ ve yò
3sg return PV PU OBJ 1sg hope PU PUF
'I hope that he comes back.'

7.4 Possessive constructions

Possessive constructions in Lahu are marked by the most important of all Lahu unrestricted particles, *ve*. The possessor comes before the thing possessed (possessor + *ve* + possessed): *ṇà ve ᳚-ch᳚* ‘my friend’; *Lāhū-ŷâ ve mû-mî* ‘the country of the Lahu’; *phí ve mē-tu* ‘a dog’s tail.’ Many possessive constructions may be shortened by omitting *ve*, becoming more like unitary noun compounds: *và? ve ᳚-śā* ‘the flesh of a pig’ vs *và?-śā* ‘pig flesh’; ‘pork.’ The *᳚-* prefix in the full genitive construction helps to set off the thing possessed as a separate noun. Similarly, *ve* is often omitted after a pronominal possessor; the following three ways to say ‘your father’ show a continuum of closeness of bonding between the two nouns: *nò ve ᳚-pa*, *nò ᳚-pā*, *nò-pa*.

7.5 Demonstratives and locatives

- (a) Spatial demonstratives: for people who live in the hills, it is important to know which way is up. Five demonstrative nouns function to indicate a general relative position, including two that specify space above or below the speaker: *chò* ‘here,’ *᳚* ‘there,’ *cō* ‘way over there’; ‘yonder,’ *nô* ‘up there,’ *mô* ‘down there.’
- (b) Spatial nouns: for more precise specification of relative location, a number of spatial nouns can be used after the noun that is the point of reference. Thus, *y᳚ ᳚-qh᳚* ‘in(side) the house,’ *y᳚ ᳚-h᳚* ‘under(neath) the house,’ *qhâ? ᳚-pâ-nê* ‘near the village,’ *he ᳚-q᳚-ji* ‘the middle of the swidden.’ Other spatial nouns are *᳚-qh᳚* ‘above’; ‘over’; ‘on top of’ (but not touching); *᳚-qh᳚-nô* ‘in back of; behind’; *᳚-thà?* ‘on top of; above and touching’; *᳚-na* ‘in front of and above’; *᳚-bâ* ‘outside’; *᳚-ḡû-ṣî* ‘in front of’; ‘before.’
- (c) Locative noun particles: Lahu has several noun particles of general locative meaning (*kâ?*, *᳚*, *lo* (< PTB **lam* ‘road’)) which are neutral with respect to directionality, and may all follow the equally vague spatial demonstratives. In the following examples,

lo does not specify direction of motion, or even motion vs rest; the interpretation depends on the built-in semantic features of the clause's verb:

<i>há-qō</i>	<i>lo</i>	<i>mi</i>	<i>chē</i>	<i>ve</i>
cave		sit	PROG	pu
'He's sitting in the cave.'				

<i>há.qō</i>	<i>lo</i>	<i>lō?</i>	<i>e</i>	<i>ò</i>
cave		enter	PV	PV
'He has already gone <i>into</i> the cave.'				

<i>há.qō</i>	<i>lo</i>	<i>tɔ?</i>	<i>e</i>	<i>ò</i>
cave		emerge	PV	PV
'He has already come <i>out of</i> the cave.'				

- (d) Determiners: the demonstrative *chi* 'this' can modify nouns in several ways. It may directly follow the noun it modifies (*N + chi*); or it may be connected to its head noun by the genitive particle *ve*, with this combination coming either before (*chi + ve + N*) or after (*N + chi + ve*) the head: *chɔ chi*, *chi ve chɔ*, *chɔ chi ve* all mean 'this person,' as does a fuller construction with numeral-plus-classifier added (*chɔ chi ve té gā*). Often *chi* has a weaker referential force than 'this,' merely referring back to something already mentioned or introduced into the discussion. *chi* also occurs with extentive nouns,¹⁴ to form 'nadvbrial' expressions that are intermediate between NPs and adverbials, e.g. *chi ma* 'this much,' *chi hi* 'this big,' *chi si* 'this long,' *chi fi* 'this far.' The Lahu expression for 'that' is *ò ve* (*lit.* 'of there') where the first element is one of the spatial demonstratives: 'that *N*' *ò + ve + N* or *N + ò + ve* (+ num + clf); 'that rock' *ò ve há-pi* or *há-pi ò ve (té si)*.

8 THE VERB PHRASE (VP)

A clause must have a VP, whereas NPs are optional. The core of a verb phrase is a verbal nucleus consisting of one or more verbs. This nucleus may be preceded by adverbials and/or followed by verb particles (PV). To be considered a verb, a morpheme must be negatable by means of the adverb *má* 'not': *thè?* *ve* 'to kick,' *má thè?* 'does not kick.'¹⁵ By this definition, adjectives in Lahu are really verbs, since they can be negated: *qhā ve* 'to be bitter'; *má qhā* 'is not bitter.' To express a higher than normal degree of a quality, the adverb *a-cí* 'more' or the extentive expression *šu a-ké* 'than others' may be used before the adjectival verb, but there is no clear contrast between the comparative and superlative degrees: *dà?* *ve* 'be good'; 'be pretty,' *a-cí dà?* *ve* 'more good; more pretty,' *šu a-ké dà?* *ve* 'better/pretty than others.'

Since Lahu sentences lack clearcut subjects and objects, the distinctions between transitive and intransitive verbs, or between active and passive voice, are basically alien to Lahu grammar. A given Lahu verb will receive an active or passive English translation according to which noun phrase is treated as the topic, something which must be inferred from the sentence as a whole:

- (a) *li?* *chi* *mí.chɔ* *qhɔ* *há?* *kə* *mē*
 book this shoulderbag inside quickly insert puf
 'Hurry and put these books into the shoulderbag.'
- (b) *li?* *chi* *mí.chɔ* *qhɔ* *kə* *tā* *ve* *yò*
 book this shoulderbag inside insert PV PV puf
 'These books have already been put into the shoulderbag.'

Sentence (a) is imperative, with the urging particle *mē* and the adverb *hā?* ‘quickly,’ which typically occurs in commands. Since *li?* *chi* ‘these books’ is inanimate, it cannot be the thing that initiates the action, so *kə* is naturally translated by an active verb. In (b), the VP contains the durative particle *tā*, which indicates previously completed action; the act of insertion is already accomplished, so *li?* *chi* is taken as the topic, and the English translation appropriately has a passive verb.

8.1 Verb particles (pv)

There are more than 20 important PVS, which may be divided into four subclasses. Conspicuously absent are any that refer to tense:

1. Directional: the PVS of this class include: *dà?* ‘reciprocal,’ ‘mutual’: *dʒ?* *dà?* ‘strike one another’; *və* ‘transportatory motion’: *fá* *və* ‘take something and hide it’; *e* ‘motion away’: *qɔ?* *e* ‘go back to’; *la* ‘motion towards’: *qɔ?* *la* ‘come back to’; and *lā* ‘non-third person benefaction.’ *la* and *e* are also used in a figurative sense, to indicate becoming, or a gradual approach to a present or future state of affairs: *pə* *la* ‘be almost finished’; *pə* *e* ‘be finished,’ ‘all used up.’
2. Experiential: this group of PVS express subjective attitudes towards the nature of one’s experience. They include: *qhe* ‘excessive repetition’; *gā* ‘desiderative’; *jɔ* ‘experiential’; and *à* ‘asseverative.’
3. Aspectual: these PVS include: *tā* ‘durative’; ‘perfective’; *tū* ‘future’; ‘hypothetical’; ‘purposive’; *šɔ* ‘durative’; *še?* ‘inchoative’; and *ò* ‘completed action’; ‘change of state.’
4. Imperative: e.g. *a* ‘mild imperative or suggestion’; *šā* ‘intended action of the first person’; *yà* ‘brusque imperative’; *vì* ‘hortatory’; *lō* ‘urge someone with insistence’; *-?* ‘unmarked imperative.’

8.2 Verb concatenation

Lahu is remarkable for the ease with which two or more verbs may be concatenated by simple juxtaposition to form complex verbal nuclei, even though the verbs refer to a series of separate, temporally consecutive actions, like the verbs ‘jump,’ ‘bite,’ and ‘eat’ in the following sentence:

<i>lā</i>	<i>pɔ?</i>	<i>chè?</i>	<i>cā</i>	<i>pə</i>	<i>še</i>	<i>ve</i>	<i>cē</i>
tiger	jump	bite	eat	finish	PV	PU	Puf
v ¹	v ²	v ³	v ⁴				

‘The tiger jumped (out) and bit (into them) and ate (them) all up!’

The fourth verb, *pə* ‘finish,’ is here used in an aspectual sense, indicating that the whole series of actions was carried through to completion.

The most interesting strings of verbs (“true concatenations”) are those which form a single verbal idea, so that they all belong to the same clause. One of the verbs in each concatenation is the main verb or verb-head (vh), which maintains its basic meaning; the other verb(s) undergo semantic “bleaching,” acquiring more abstract grammatical meanings so that they modify the verb-head. Several dozen Lahu verbs have the ability to appear before or after other verbs in concatenations; they are called “versatile verbs” (vv) as a tribute to their flexible nature. Pre-head versatile verbs are symbolized as vv, and

post-head versatiles as vv. In the following example of a five-verb concatenation, all four of the versatile verbs follow the verb-head:

<i>ci</i>	<i>ḡò</i>	<i>tɔ?</i>	<i>mā</i>	<i>p̄i</i>	<i>c᷑</i>
	vh	vv	vv	vv	vv
tooth	pull	emerge	teach	give	correct
'(They) ought to show them how to pull out teeth.'					

The verbs *tɔ?*, *mā*, *p̄i*, and *c᷑*, which as main verbs mean ‘emerge,’ ‘teach,’ ‘give,’ and ‘be correct,’ have much more abstract meanings as versatile verbs, translated here by ‘out’ (*tɔ?*), ‘show how’ (*mā*), ‘them’ (*p̄i*), and ‘ought’ (*c᷑*).

In the following concatenation, one vv precedes the vh, while three vvs follow it:

<i>ŋàhi</i>	<i>ga</i>	<i>yù</i>	<i>qay</i>	<i>ci</i>	<i>c᷑</i>	<i>ve</i>	<i>lā</i>
	vv	vh	vv	vv	vv	pu	pu
1pl	help	take	go	send	correct		
'Should (we) make (them) help to take (it) away?'							

8.3 Causation and benefaction

The oldest way of forming causative verbs in the Tibeto-Burman family was by an **s*-prefix. Although this prefix has long since disappeared from the Loloish languages, its effects survive in over a dozen Lahu verbs of causative meaning that differ only in tone and/or initial consonant from a corresponding non-causative verb:

<i>d᷑</i>	drink	<i>t᷑</i>	give to drink
<i>m᷑</i>	see	<i>m᷑</i>	show
<i>c᷑</i>	eat	<i>c᷑</i>	feed someone
<i>v᷑</i>	be far	<i>f᷑</i>	separate; demarcate
<i>v᷑?</i>	hide oneself	<i>f᷑</i>	hide something
<i>t᷑?</i>	be burning	<i>t᷑</i>	kindle; set on fire

For all other verbs, Lahu now must form causatives by means of verb concatenations, using such versatile verbs as *yù* ‘take,’ *te* ‘do,’ *p̄i* ‘give,’ and especially *ci* ‘send on an errand,’ e.g. *yù khá* (vv + vh) ‘make blocked up’; *te qè* (vv + vh) ‘make wide,’ ‘widen’; *ce p̄i* (vh + vv) ‘cause to fall, drop’; *qay ci* (vh + vv) ‘cause to go,’ *chu ci* (vh + vv) ‘cause to be fat.’

Lahu is careful to specify for whose benefit the verbal action is performed. This is done by two morphemes, the vv *p̄i* ‘give’ and the pv *lā* (< *lā* ‘come’). The outer-directed *p̄i* is used to indicate that the action affects a third person, while the inner-directed *lā* shows that the action affects a non-third person, e.g. *ch᷑ lā* (vh + pv) ‘chop for me/us/you’; *ch᷑ p̄i* (vh + vv) ‘chop for him/her/them.’

9 NOMINALIZATION AND RELATIVIZATION

Lahu has five nominalizing particles. Four of them have highly specific meanings, as we can illustrate by combining them with the clause *i-kâ?* *hé* ‘bathe’:

- (a) *pā* agentive nominalizer; one who v’s; a v’er
i-kâ? *hé pā* bather; one who bathes
- (b) *kî* locative nominalizer; the place where v
i-kâ? *hé kî* bathing place; bathroom

- (c) *thā* temporal nominalizer; the time that one v's
i-kā? *hé thā* the time for bathing

(d) *tù* purposive nominalizer; something for v'ing
i-kā? *hé tù* bathing suit; thing for bathing

The fifth nominalizing particle, *ve*, is the most important particle in the whole language. It forms the most general kind of nominalization: *i-kâ? hé ve* ‘bathing; to bathe.’ Frequently *ve* marks a clause that is embedded as the topic of a larger sentence:

nyà mâ qò?te ve tê chi qhò? gà ô
'I haven't done that for ten years now.'
(My not doing that again has now reached ten years.)

With great frequency this particle also occurs in final clauses to nominalize whole sentences:

a-šwè-thâ chɔ-qhâ nî ȣâ cɔ ve yò cê
‘Once upon a time there were two thieves.’
(. . . it was a case of there being two thieves)

We have seen how *ve* marks the relationship of possession, subordinating one noun (the possessor) to the thing possessed. Similar to this function is the role of *ve* as the marker of relative clauses. The relative clause (enclosed in square brackets) usually comes directly before the noun head (nrh):

In some cases, the relative clause may be shifted to the position directly after the nrh:

và ɻ-ó-qō [c᷑ tā ve] mè jā
nrh
'A boiled pig's head is very tasty.'

There is occasionally ambiguity between nominalizing and relativizing *ve*:

{*tē-ghā?-tē-lō ū* ve} *a-pi-qu chi ū e ve yō*
‘What the whole village knows is that the old woman has died.’

[*tē-qhāʔ-tē-lò šī ve*] *a-pi-qu chi ši e ve yò*
 nrh
 ‘The old woman whom the whole village knew has died.’

NOTES

- 1 This chapter is a condensed and revised version of Matisoff (1992).
 - 2 The etymology of the name *Lahu* remains obscure, though I believe the second syllable *-hu* derives from a PTB root **s-lu* ‘people,’ that underlies the element *-su* in many TB ethnonyms (e.g. Lisu, Bisu, Nosu, Tosu), as well as the second syllable *-lo* of *Lolo* itself.

- 3 See Vuong Hoang Tuyêñ (1973). The Vietnamese exonym Kha Quy seems to refer specifically to the Yellow Lahu, who are also called Kwi in Thai and Shan. A group of Lahu in Yunnan are also known as Kucong Yi (see Fei Xiaotong 1990: 11–24).
- 4 See Nishida (1969), Bradley (1979). YL lacks the post-velar stops /q qh/ (BL *qɔ̄*, YL *kɔ̄⁵⁵* ‘nine’; BL *qhā*, YL *kha⁵⁵* ‘bitter’), and does not have the central vowels /i/ or /ə/, e.g. BL *phi*, YL *phi⁵⁵* ‘dog’; BL *ka*, YL *ke³³* ‘put in.’
- 5 These checked syllables descend from older syllables with final /-p -t -k/. The high-rising tone, now unchecked, also descends from stop-finalled syllables of a certain type (Matisoff 1970).
- 6 Nasalized vowels occur in loanwords, allophonically in syllables with low vowels and laryngeal initials (rhinoglottophilia), and in various sound-symbolic functions (see Matisoff 1989). These are conventionally transcribed by a syllable-final *-n*, but are not to be confused with real nasal final consonants.
- 7 Medial *-w-*, in words like *bwē* (~*bj*) ‘room,’ *chi-pi-qwè* ‘barking deer,’ is to be considered part of the vocalic nucleus. See Matisoff (1973/82: 15–21).
- 8 The citation form of verbs takes the nominalizer *ve*, much like the English infinitive nominalizer ‘to’: *qay ve* ‘to go,’ *dà?* *ve* ‘to be good.’ See section 9.
- 9 This could also have the object + verb interpretation: *ú-gê te ve* ‘make a pillow.’
- 10 Other interrogatives which require *le* at the end of the sentence include: *à-thò?-ma te le* ‘why?’, *qhà ve* ‘which?’, *qhà-qhe* ‘how?’, *qhà-qhe ve* ‘what kind of?’, *qhà-thâ?* ‘when?’, *qhà-ní* ‘how many?’, *qhà-ma* ‘how much?’, *qhà-hi* ‘how big?’, *qhà-ši* ‘how long?’, *qhà-fí* ‘how far?’, *qhà* ‘where?’, *qhà-phâ* ‘what direction?’
- 11 The boundary between clauses is indicated by double crosshatches (##). The meaning of *le* is sometimes causal, not just conjoining. In the appropriate context the first clause could also be translated ‘Since he walked on the road all day’ The particle *pa-tɔ* (below) has a more specific causal sense.
- 12 The same morpheme occurs in the interrogative pronoun *a-šu* ‘who?’
- 13 These higher numerals all have final nasal consonants in Shan, and many Lahu speakers still pronounce them with nasalized vowels.
- 14 These extensives also combine with the interrogative *qhà*, e.g. *qhà-ma* ‘how much?’, *qhà-hi* ‘how big?’, etc.
- 15 Nouns are negated by means of the verbal expression *mâ hê?* ‘is not so’; ‘is not the case,’ e.g. *yô Lâhū-yâ mâ hê?* ‘He is not a Lahu.’

REFERENCES

- Bradley, David (1979) *Lahu Dialects*, Canberra: Australian National University Press.
- Fei, Xiaotong (1990) ‘Ethnic identification in China’, in Gehan Wijeyewardene (ed.) *Thai-Yunnan Project Newsletter*, No. 11. Canberra: Australian National University, Research School of Pacific Studies, Department of Anthropology, 11–24.
- Matisoff, James A. (1970) ‘Glottal dissimilation and the Lahu high-rising tone: a tonogenetic case-study’, *Journal of the American Oriental Society* 90.1: 13–44.
- Matisoff, James A. (1973/82) *The Grammar of Lahu*, University of California Publications in Linguistics No. 75, Berkeley and Los Angeles: University of California Press (reprinted 1982).
- Matisoff, James A. (1988) *The Dictionary of Lahu*, University of California Publications in Linguistics No. 111, Berkeley and Los Angeles: University of California Press.
- Matisoff, James A. (1989) ‘The bulging monosyllable, or the mora the merrier: echo-vowel adverbialization in Lahu’, in Jeremy H. C. S. Davidson (ed.) *South-East Asian Linguistics: Essays in Honour of Eugénie J. A. Henderson*, London: School of Oriental and African Studies, 163–197.

- Matisoff, James A. (1992) ‘The Lahu people and their language’, in Judy Lewis (ed.) *Minority Cultures of Laos*, Rancho Cordova, CA: Southeast Asia Community Resource Center, Folsom Cordova Unified School District, 125–247.
- Matisoff, James A. (2006) *English–Lahu Lexicon*, University of California Publications in Linguistics Vol. 139, Berkeley, Los Angeles, and London: University of California Press.
- Matisoff, James A. (2009) ‘Syntactic parallelism and morphological elaboration in Lahu religious poetry,’ in Anthony R. Walker (ed.) *Pika-pika: The Flashing Firefly; Essays for Pauline Walker by her Friends in the Arts and Sciences*, New Delhi: Hindustan Publishing Corporation, 117–136.
- Matisoff, James A. (2011) ‘Stung by a bee, you fear a fly: areal and universal aspects of Lahu proverbial wisdom’, *Bulletin of the School of Oriental and African Studies* 74.2: 273–302.
- Matisoff, James A. (2014) ‘Using native lexical resources to create technical neologisms for minority languages,’ in Richard Van Ness Simmons and Newell Ann Van Auken (eds) *Studies in Chinese and Sino-Tibetan Linguistics*, Taipei: Institute of Linguistics, Academia Sinica, 185–224.
- Nishida, Tatsuo (1969) [‘A study of Yellow Lahu: preliminary report on the language of the Yellow Lahu of Chiang Rai Province, Thailand’] (in Japanese) *Tonan Azia Kenkyu* (Kyoto) 7.1: 2–39.
- Vuong, Hoang Tuyêñ (1973) ‘De quelques groupes ethniques arrachés de justesse à l’extinction aux confins de la zone du Nord-Ouest’, in Lam Tam, Dang Nghiêm Van and Vuong Hoang Tuyêñ (eds) *Etudes Vietnamienヌes 36*, Données ethnographiques (II), Hanoi, 149–200 (section entitled ‘Les Co Sung ou La Hu,’ 176–185).

§3.8 Karenic

CHAPTER FORTY-NINE

EASTERN KAYAH LI

David Solnit

1 INTRODUCTION

Eastern Kayah Li is spoken in the area of Kayah State (Karen) of Burma east of the Pun river (west of and parallel to the Salween) and extending a short distance into Mae Hong Son province, northwestern Thailand. ‘Kayah Li’ corresponds to the self-designation /kajē li phú/ (for the people; the language is /kajē li ñò/). *kajē* is ‘person’, ‘Kayah’ and *li* is ‘red’. An older term is Karen, which is anglicized Burmese; Burmese <ni> is also ‘red’. ‘Karen’ is still preferred over ‘Kayah State’ for the geographic/political area by those affiliated with or sympathetic to the local insurgent organization.

1.1 Subgrouping and dialects

Eastern Kayah Li belongs to the Karenic branch of Tibeto-Burman. Karenic (or simply Karen) is most usefully divided into three geographic groupings: Southern, consisting of Pwo and Sgaw; Central, consisting of Kayah Li along with Brè, Yintale, Palaychi, Mopwa and many more; and the Pa-O language, which makes up a Northern group on its own. Padaung languages seem to be transitional, having affinities to both Northern and Central.

The Kayah Li are numerically the dominant group in the Central Karen area, with perhaps 100,000 speakers, all in Burma except for some 1,500 in Mae Hong Son province of Thailand.

There is variation, mostly staying within the bounds of mutual intelligibility, across the Kayah Li speech area. Kayah Li speakers in Thailand recognize two mutually intelligible dialects, *kè khu* ‘upper’ and *kè kē* ‘lower’. The language described in this chapter is the Lower variety. See Solnit (1997) for a summary of what little information is available about other local variations.

1.2 Language history

Karen languages are spoken along a north–south axis roughly coinciding with the Thailand–Burma border, reaching southwards nearly to the Isthmus of Kra, westwards into the Irrawaddy delta, and eastwards into Lampang and Chiang Rai provinces of Thailand. The centre of diversity is in Burma, in western Kayah State (Karen) and the adjoining area of Karen State.

Karen is thus located on the periphery of the Tibeto-Burman area. It is the southernmost and, along with Lolo-Burmese and Tujia, the easternmost of the Tibeto-Burman subgroups. This peripheral location has put Karen in extensive contact with non-Tibeto-Burman languages, of the Tai and Mon-Khmer families. The Tai influence is evident in

Karen historical phonology, most strikingly so in the Proto-Karen consonant system and its interaction with the evolution of tones. Loanword evidence points to contact with the Palaungic and Monic branches of Mon-Khmer. Karen's verb-medial syntactic typology, in which it is quite unusual in Tibeto-Burman (apart from Karen, only Bai is strongly verb-medial), is surely another result of contact with Tai, Mon-Khmer or both.

The areal factors make Karen's distinctness in syntactic type a weak reason for separating it out from (the rest of) Tibeto-Burman as is done in Benedict (1972; Benedict in fact does note the influence of Tai on Karen). Benedict (1972: 128 note 350) also cites lexical evidence for his grouping, namely a few words that Karen shares with Chinese and not with the rest of Tibeto-Burman. In a related vein, Jones (1975) has pointed to a set of Proto-Karen etyma having neither Tibeto-Burman etymologies nor identifiable loan sources, from which he draws the eccentric conclusion that Karen is an isolate. But in the absence of any demonstration that Karen has significantly more of such special lexical features than other Tibeto-Burman subgroups, Karen must be considered just another of those subgroups – as in Benedict's (1976) revised view.

The Central Karen area is the centre of Karen linguistic diversity and so the presumptive historical centre from which the languages have spread.

2 PHONOLOGY

2.1 Syllable structure

The maximal phonological shape of Kayah Li lexical items can be represented by this formula: [c v/t c1 (c2) (G) v/t v/t]. c or c = consonant, v or v = vowel, t or t = tone, G = glide (v and t are separated by a slash in the formula to reflect simultaneous occurrence). Parenthesized upper-case elements are optional. Lower-case elements (c v t) represent prefix and suffix: they are also optional, and are realized by a small subset of the full inventory of their type (for example, of 20 consonants in the full inventory, only four can appear in prefixes). The minimal full syllable is considered to require an initial consonant, the only exceptions being a very small number of morphemes that I write with no initial.

2.2 Vowels

Since Kayah Li has no final consonants, the vowel system equals the rhyme system. Simple rhymes are / i u u e ɔ o ε ʌ ə a/. The symbols have standard IPA values, with the following modifications:

- 1 /a/ is a low central [ʌ].
- 2 /e o/ are slightly higher than cardinal. When unstressed, /e/ may lower.
- 3 /u/ is a centralized high back unrounded vowel, but not a fully central [i].
- 4 /ɤ/ is a slightly centralized upper-mid back unrounded vowel, occurring largely in Thai or Shan loanwords.
- 5 /ʌ/ is a centralized version of cardinal [ʌ], similar to the English vowel often transcribed with the same symbol.

The phonemic status of the three-way contrast /u ɤ ʌ/ is solid only in the high tone, as in the minimal triplet *dú* 'cut', *'slice'*, *d̥'at'*, *d̥'give'*. In other tones, words with /ɤ/ are rare.

Vowels preceded by onglides make up a system of compound rhymes: /wi wa we ja jo (ju)/. The onglides /j- w-/ are usually closer to [ɛ] and [ɔ] respectively. /wa/ occurs with

all initials except /v/. /wi/ occurs with all simple initials except /v t n h 0/. /ja/ occurs only after labial obstruents and simple (non-cluster) velars. /ju jo we/ are relatively rare.

Affixes may contain a vowel whose quality copies that of the associated full syllable; I write this copy vowel as /ə/.

2.3 Consonants

Simple initial consonants are the following: voiceless unaspirated obstruents /p t c k ʔ/, voiceless aspirated obstruents /ph th ch kh/, voiced obstruents /b d/, voiced nasals /m n ŋ/, voiceless fricatives /s ʃ/, voiced non-nasal continuants /v l j r/. Note the following: /c ch/ are alveo-palatal affricates; /j/ may have fricative noise and occasional prenasalization; /ŋ/ is [ŋ̊] before front vowels and glide /j/; /r/ is a retroflex approximant; glottal stop may elide when preceded by a full syllable and in other prosodic environments not clearly understood.

Initial consonant clusters are /pl phr kl khr/ – note the complementary distribution. Medial /r/ is largely or completely devoiced by simultaneous aspiration, approaching a retroflex fricative in quality.

2.4 Tone

Eastern Kayah Li has four contrastive tones. Using the Chao five-level numerical pitch scale, their pre-pause realizations are mid (33), low level (11) plus final glottal stop, low falling (21) plus final creak and glottal stop, and high (55) plus glottal stop. In connected speech the low level and high tones omit the final glottal stops. I write the tones as follows, using /a/ as example: /ā/ mid, /a/ (no mark) low level, /à/ low falling, /á/ high.

There is also a high falling tone (52 plus glottal stop) which is the realization of a suffix that I write -ə. In Lower Eastern Kayah the suffix merges with the vowel of the preceding full syllable such that the resulting vowel has the quality of the full-syllable vowel and the tone of the suffix e.g. /chwa khriə/ [chwa²² khri⁵²] ‘crab’. In Upper, the suffix remains as an offglide -μ plus the high falling tone; compare Lower *chwa khri*⁵², Upper *chwa khrey*⁵² ‘crab’; Lower *təple*⁵², Upper *təpley*⁵² ‘bat’.

2.5 Phonological alternations

The vowel /ə/ exhibits vowel harmony. It occurs only in affixes, where it copies the vowel of the full syllable to which it is affixed. In prefixes, this vowel is reduced towards schwa as speech tempo increases; in suffixes it assimilates completely to the vowel of the full syllable.

Tonal dissimilation occurs in prefixes ʔi- and ʔi-, and in a few other morphemes and morpheme-like elements. With the two prefixes, the lexical tone is realized only before mid tone syllables. Before non-mid tone full syllables, dissimilatory conditioning applies: the prefix is high tone if the full syllable is low-level or low-falling, and low if the full syllable is high (or high falling). Examples: ʔiθuu ‘post’, ʔilò ‘to plant (seeds)’ (conditioned high before low-level and low-falling); ʔikhré ‘to winnow’ (conditioned low before high); ʔilù ‘the Kayah New Year festival’ (lexical low before mid); ʔivī ‘to whistle’ (lexical high before mid).

Something similar can be seen in *khe~khē* ‘leg’. Low-level *khe* occurs before mid tone in compounds like *khe mā* ‘knee’ (*mā* ‘joint’) and *khe rē* ‘paw’. Mid tone *khē* occurs before low-level, as in *khē do* ‘lower leg’ and *khē le* ‘foot’. Low-level *khe* is the

etymologically regular form, so in this case again the mid tone syllables allow the more basic form (historically so here) to surface.

3 WORD FORMATION

3.1 Derivational morphology

Although affixes exist in Kayah Li, they play a marginal role in word formation, which is predominantly by compounding. Even the verb complex, the core of the clause (see later) can be analysed as made up partly of verbal compounds.

Prefixes are of two types: *?i*, occurring in high and low-level tones; and *cə*, where *c* = /p t k/ and tone is low level. Both types include instances with and without identifiable function (semantic and morphological). An example with identifiable function is *?i* in names of instruments: *sé ?iché* ‘sewing machine’ (*sé* ‘large machine’, *ché* ‘sew’), *?ithá* a plough (*thá* ‘to’ plough). Examples in which the prefix has no identifiable function are *?ise* ‘salt’, *?ikhré* ‘to winnow’.

The suffix *-ə* is described above. A second suffix has the form *phú~hú~ú* (*phú* is also a noun ‘child’ and a postposed modifier ‘small’), meaning ‘member of a class’, ‘instance of a category’. Examples are *kajé li phú ~ kajé liú* ‘the (red) Kayah’ (*kajé* ‘person’, ‘Kayah’, *li* ‘red’), *klā́ú ~ klā́yphú* ‘soldier’ (*klā́* ‘army’) and *tḗú* ‘fish’. *tē* is also ‘fish’, but it cannot occur on its own; it must be followed by either this suffix or a second syllable specifying the particular type of fish.

Kayah Li uses reduplication of the last syllable in a clause to mean ‘also’, ‘too’, ‘either’. For example,

- (1) v̥é ma ɻe k̥á p̥hé thé ja ja
1sg be.so eat COMITATIVE simply pig flesh REDUPLICATION
'I ate only pork, too (as did he).'

The reduplicated morpheme may be of any grammatical category except for one type of clause-final particle.

3.2 Compounding

Kayah Li morphemes are usefully divided into free and bound types. Free morphemes are those capable of functioning as a major clause constituent such as Subject, Object or main verb; bound morphemes are those that cannot. Compound expressions may contain all possible combinations of free and bound morphemes, as exemplified in the following compound nouns:

- free + free: *síne thi* gun + penis: trigger
free + bound: *pù po* ox + enclosure: cattle-pen
bound + free: *tē bū* fish + white: a kind of large white fish

The general word for ‘fish’, *tēú*, with suffix *ú*, exemplifies bound + bound.

In addition to the foregoing type of compound noun, there are two types of compounds that combine one verbal member with one nominal, and function as predicates. One type is Verb-Object compounds such as *nō jechuə* ‘enter + Jesus: be a Christian’. A second type is Subject-Predicate, such as *mi du* ‘name + big: famous’ or *se ɻo* ‘fruit + exist: useful, meaningful’.

For verb-verb compounds, see section 4.2.2 ‘Serial verbs: the verb complex’.

4 SYNTAX

4.1 Structure of the noun phrase

There are three major types of nouns: (1) ordinary nouns, including common nouns, names and pronouns; (2) classifiers, with several subdivisions; and (3) localizers, which cover much of the semantic territory of English prepositions. Noun phrases headed by each major type have their own characteristics and seldom co-occur within any larger NP.

Ordinary-headed NPs consist of an ordinary noun plus optional modifiers. Nominal modifiers precede the head, and verbal modifiers (including verbs translated as English adjectives) follow, as in *vē dīpō du* ‘I + pot + big: my big pot’ or *phremō hā jē* ‘woman + lower.garment + tattered: tattered woman’s skirt’. The postposed verbal modifier is in fact a clause, and so may include multiple constituents. For example, *kajē [pā tha]* ‘people to cut sesame (plants), sesame-choppers’ consists of a head noun *kajē* ‘person’ modified by a verb-object construction made up of *pā* ‘cut’ and *tha* ‘sesame’.

There are three exceptions to the pre-head position of nominal modifiers. First, material follows artifact, as in *dīpō tɔthé* ‘pot + iron: iron pot’ and *bēlū tɔphé* ‘cloth + cotton: cotton cloth’. Second, many names of plant and animal species begin with a general term for the kind, as *thu tō* ‘drongo (*Dicrurus spp.*)’ (*thu* ‘bird’) or *sɔ leha* ‘teak tree’ (*sɔ* ‘tree’). Third, ethnic designations follow the head in the meaning ‘X-style’, as in *haca phrē* ‘clothes + Shan: Shan-style clothes’.

Pronouns are as follows:

	Sg	Pl	Undifferentiated
1	<i>vē</i>	<i>pe</i>	
2	<i>ne</i>	<i>sī</i>	
3	<i>jesī</i>	<i>?a, lū, ?ū</i>	

jesī is rare. The more common third-person pronouns are *?a*, *lū*, and *?ū*, which are unmarked for number. *?a* and *lū* alternate within a clause, with *lū* appearing only if *?a* or some other non-coreferential third-person NP appears before it in the clause. *?ū* is indefinite and backgrounded, often translatable as ‘other people’ or ‘they’ as in *they say...*

Classifier-headed NPs have the structure [(Dem) (s) CLF]. Dem is a Demonstrative, either *?x* ‘this’ or *na* ‘that’. s is a clause, in this function known as a preposed attributive clause (see later). CLF is a phrase made up of a classifier and an obligatory quantifier, with the ordering determined by the choice of classifier and quantifier (usually Q-CLF'). Quantifiers include the numerals and a few other morphemes such as *pwā* ‘every’ and *bá* ‘how many?’. Types of classifiers are:

- Time, including units of time such as *nā* ‘day’ and *na* ‘year’, and classifiers for instances of an action, such as *phó* ‘a time’.
- Measure, mostly names of containers (*pē* ‘bottle’, *khri* ‘packbasket’); also conventional units such as *plē* ‘cubit’ (also a Common noun ‘arrow’).
- General, including the familiar type that relates to the shape or other characteristic of the thing being counted. Examples: *pls* ‘Clf for small round things’, *bō* ‘Clf for lengths’, *be* ‘Clf for faceted or winglike things’.
- Localizer-headed NPs consist of a localizer obligitorily modified by a preceding ordinary-headed NP, the resulting NP in turn serving as object of a preposition. An example is *d̄s khrā kū* ‘inside the bottle-gourd’, consisting of the preposition *d̄s* followed by an NP made up of the localizer *kū* ‘inside-part’ modified by *khrā* ‘bottle-gourd’.

Localizers (the term is borrowed from Chao 1968) often also occur as ordinary nouns. *kū*, from the preceding example (which could more literally be glossed ‘at the bottle-gourd’s inside part’), also occurs in *thā kū* ‘spring, well’ (*thā* ‘water’) and *?akū* ‘hole’.

4.2 Structure of the clause

4.2.1 Word order and clause structure

Clause structure is as follows:

[NP₁ [[VC NP₂ NP₃ PP₁]V PP₂ CLF’]VP]_s SPTC

VC = verb complex, PP = prepositional phrase, CLF’ = quantifier + classifier, SPTC = sentence particle. Grammatical relations are NP₁ = subject, NP₂ = indirect object, and NP₃ = direct object. The remaining phrasal categories represent various oblique relations, with the one represented by the clause-final CLF’ termed Extent.

An example of indirect and direct objects:

- (2) *sárá ?iswá phúcè li*
teacher teach child writing
'The teacher teaches the children (their letters).'

An example of pp₁ (inside of v') and pp₂ (outside of v'):

- (3) *?a phā ná ?o [dř hi thū]₁ [hú pe ?o ?ā]₂ ná*
3 grandmother PTC exist at house edge like 1pl exist this PTC
'His grandmother lived at the edge of the village, as we do here.'

Prepositions are of three types:

- Locative, of which there are three, all glossable as ‘at, when’ but differing in evidential meaning: *mú* (known by inference or hearsay), *bí* (in sight), *dř* (unmarked for evidentiality).
- Extensive. Examples are *bá* ‘as much as’ and *tí* ‘as big as’.
- Miscellaneous. Examples are *phú~hú* ‘like, as’ and *né*, which indicates a backgrounded clause participant.

An example of extent expression, with alternate classifiers:

- (4) *phúcè cwá dř hohé sō nā /sō phó /sí sō*
child go at school three day three time three CLF:human
'The children went to school for three days/three times/three children [went to school].'

Verb serialization in Kayah Li is extensive, but it is confined to the verb complex (next section), in the form of immediate concatenation of verbs, with no intervening arguments.

There is one exception to the preceding statement: a construction involving a sequence of two v' within a single clause, having the general meaning ‘v₂ from having v₁-ed’. When v₁ is ?o ‘exist’, this construction is the most usual way to express the notion of source:

- (5) *?o dř sō kū tā the nō to*
exist at tree inside fall go.out at.all NEG

‘It didn’t fall out of the trees.’

This utterance is from a narrative in which the context makes it clear that the point of the sentence is that ‘it’ didn’t fall out from anywhere; i.e. both ‘exist’ and ‘fall’ are negated. Note that the negative too has scope over both verbs in the sequence, illustrating the monoclausal nature of the construction.

4.2.2 Serial verbs: the verb complex

The verb complex (vc) is the core of the v' and hence of the vp. Its general structure is:

(PTC) (PTC) . . . v (v) . . . (PTC) . . . (PTC)

In a verb string the first verb is the head and is the only required constituent. Any other sequence of verb-containing constructions is a clause series. What is not found is noun (or np) arguments intervening between the members of a series of verbs, the single exception being the fairly restricted source expression.

The following example illustrates the major properties of the verb complex (the bracketed string):

- (6) *?a [vī jo cwá] chā mō bē nA rA*
 3 throw fly go chicken female yellow ASSERT PAUSE
 ‘She threw the yellow hen so that it flew away.’ (376.3)

- 1 The vc has the properties of a single word. The complex expression *vījo cwá* takes a single Subject (*?a* ‘she/he’) and a single (Direct) Object (*chāmō bē* ‘yellow hen’).
- 2 The semantic arguments (participant roles, theta-roles) of the vc’s constituent verbs map onto the grammatical relations specified by the vc in a complex but partially predictable way. In this example the Subject *?a* has the agent role of *vī* ‘throw’ (it is the one that does the throwing). The Direct Object *chāmō bē* has three overlapping roles: the patient role of *vī* ‘throw’ (it is thrown), the agent role of *jo* ‘fly’ (it does the flying), and the agent role of *cwá* ‘go’ (it is the thing that goes [away]).
- 3 The predicates associated with the verbs in a vc relate to each other, pairwise, in a limited number of ways. In the example, ‘fly’ and ‘go’ are related by causation, and ‘throw’ relates to the unit ‘fly-go’ also by causation. Other possible interpredicate relations are those of sequence and modification.

An example of sequence:

- (7) *?a ka déhā rā ?aphā*
 3 return ask PTC grandmother
 ‘He went home and asked his grandmother.’

The actions of *ka* ‘go towards home’ and *déhā* ‘ask’ take place in temporal sequence.

An example of modification:

- (8) *ké rō chwa*
 AMB cold strong
 ‘It’s very cold.’

The predicate *rò* ‘cold’ is modified by *chwa* ‘be strong’.

4.3 Major sentence types

The Kayah Li sentence (defined as any construction that can stand on its own as an utterance bearing an illocutionary force) may be either one or more clauses (verbal sentence) or an NP (nominal sentence). Both types may end in a sentence particle. An example of a verbal sentence:

- (9) *nō nō hóhé to ē ŋ*
enter at.all school NEG QUES HUH
'Aren't you going to school, hey?'

An example of a nominal sentence:

- (10) *thwá ke*
cat PRH
(on hearing a noise:) 'Maybe it was the cat.'

4.3.1 Affirmative

Sentences may end with any of a battery of final particles, with meanings mostly in the areas of polarity, realis/irrealis and illocutionary force. However use of a final particle is not required for affirmative sentences.

The most neutral final particle is the nominalizer *na* (it is homophonous with a topic marker and the distal demonstrative ‘that’). Autonomous (non-embedded) clauses ending with *na* are considered to be nominal sentences. In these sentences *na* marks the propositional content as old information; for example, in narratives a *na* sentence often asserts the occurrence of an anticipated event.

Here are some examples of affirmative sentences marked with other final particles:

he lest, possible undesirable event

- (11) *sí de khrā, sí?ichē ?a ?3 he*
want put dry afraid 3 mildew
'I want to put it out to dry; I'm afraid it may mildew.'
- pó~pō* urging, ‘why don’t you’:

- (12) *ne th̄ t̄-mē ?a d̄ jū v̄e p̄o*
2sg drum one-CLF 3 give easy 1sg
'Your drum, won't you just give it to me?'

mē mild counter-assertion, also used in answers to questions:

- (13) *?a me s̄l̄ á l̄i m̄e*
3 do die New.Situation mutually
'They would have killed each other [if I hadn't stopped them].'

4.3.2 Negation

Negation is expressed by two final particles, the most common and neutral being *to*, exemplified in (5) earlier. The other is the negative imperative *me*.

4.3.3 Questions

Yes-no questions are signalled by the final particle *ɛ̄* (this is one of the few Kayah Li lexical items with zero initial). Question-word (Wh-) questions use *?itē* ‘what’ and a set of morphemes incorporating its abbreviated form *tē*: *me tē* ‘why’, *tō ú tē* ‘where (general)’, *b̄tē* ‘where (nearby)’, *bó kē tē* ‘when’, *bá CLF tē* ‘how many’, *hú tē* ‘how’. None of the foregoing question types differs in syntax from its affirmative counterpart. There are also two interrogative expressions made up of two discontinuous constituents, *d̄s tē tə* ‘CLF . . . ɛ̄’ ‘which one’ and *?ū . . . p̄e* ‘who’. An example of the former is

- (14) *d̄s tē tə-cr 2a v̄t l̄í ɛ̄*
 at what one-CLF 3 delicious more QUES
 ‘Which kind tastes better?’

4.3.4 Subordinate clauses

Clauses may function as the objects of prepositions, and of certain verbs, with no special marking. The verbs relate to perception and cognition (‘see’, ‘hear’, ‘suspect’ – but not ‘think’, for which see later), as in (clause in square brackets):

- (15) *2a cwá nihō kā d̄s [kəjḗ kḗ təhā́] n̄a*
 3 go hear PTC at person harvest thatch.tree PTC
 ‘They went and heard people harvesting thatch-tree [leaves].’

Like those with NP objects, prepositional phrases with clausal objects may express location and time, as well as other notions, such as:

- (16) *2a d̄i láteá li phá hú [phḗ ?iro] hé n̄a*
 3 give instead book skin like father sing say PTC
 ‘He gave a hide book instead, as Father sang, it’s said.’

Reported speech or thought is not expressed with verb + clause. The pattern used is a clause sequence, with the first clause reporting what was said or thought and the second denoting the act of speaking or thinking. Example:

- (17) *[ne chá mā phrē tē Sethmphē] 2a hé n̄a*
 2sg fight PCT Shan PTC (name) 3 say PTC
 ‘“Attack the Shans, Sethuephe!” she said.’

Clauses may modify nouns. The function is that of a relative clause in other languages, but since there are no relative pronouns I call the Kayah Li construction an attributive clause. Attributive clauses may be nominalized, preceding the modified noun, or verbal, following the noun.

Preposed attributive clauses occur in the structure P S *n̄a* CLF’, where P is a preposition, S is the modifying clause, and CLF’ is the modified nominal expression. These structures may be quite complex, as in this example from a conversation (attributive clause in brackets, modified nominal underlined):

- (18) *kwā n̄a ma d̄s [?a muū lípəna kukl̄s] n̄a təbe*
 axe PTC be.so at 3 hit nail head that one-CLF
 ‘The axe is the one in whose head he pounded a nail’ (‘the one that he pounded a nail into its head’).

Postposed attributive clauses are more restricted in size and structural complexity. They may contain either Subject or Object NP but not both, and often consist only of a

single verb, as in *dypɔ du* ‘pot + big: big pot’. Their clausehood is evident in their ability to contain nominal arguments and even sentence particles. An example of the latter is *kəjɛ mò ʔo to* ‘person + mother + exist + not: a person with no mother’, in which the attributive clause includes the sentence particle *to* ‘negative’.

BIBLIOGRAPHY

- Benedict, Paul K. (1972) *Sino-Tibetan: A Conspectus* (Princeton-Cambridge Studies in Chinese Linguistics 2), contributing editor James A. Matisoff. New York: Cambridge University Press. (Classic work on the Sino-Tibetan family, reiterating Benedict’s contention that Tai and Miao-Yao are not genetically related to Chinese and Tibeto-Burman. Treats Karen separately, as coordinate with Tibeto-Burman in a Tibeto-Karen group that is then coordinate with Chinese.)
- (1976) ‘Sino-Tibetan: Another look’, *Journal of the American Oriental Society* 96.2, 167–197.
- Chao, Yuen-ren (1968) *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
- Jones, Robert B. (1975) ‘The question of Karen linguistic affiliation’, presented at the 8th International Conference on Sino-Tibetan Languages and Linguistics, Berkeley.
- Karenny Literature Department (1994) *The Modern Western Kayah Li–English Lexicon*. Burma: Karenny Literature Department (sponsored by Payap Research and Development Institute, Payap University, Chiang Mai; and Summer Institute of Linguistics). (Short lexicon, each entry written in the indigenous script and in phonetic transcription, many with example sentences. Includes appendixes of vocabulary in special semantic fields, such as body parts, animals, illnesses, classifiers, birds.)
- Lehman, Frederic Kris (1967a) ‘Ethnic categories in Burma and the theory of social systems’, in Peter Kunstadter (ed.) *Southeast Asian Tribes, Minorities, and Nations* v. 1, Princeton: Princeton University Press.
- (1967b) ‘Kayah society as a function of the Shan-Burma-Karen context’, in Julian H. Steward (ed.) *Contemporary Change in Traditional Societies* v. 1, Urbana: University of Illinois Press, 1–102. (Extensive ethnographic information on the Western Kayah Li of Kyèbogyi.)
- (1979) ‘Who are the Karen, and if so, why? Karen ethnohistory and a formal theory of ethnicity’, in Charles F. Keyes (ed.) *Ethnic Adaptation and Identity: The Karen on the Thai Frontier with Burma*, Philadelphia: Institute for the Study of Human Issues, 215–253. (Discussion of points of anthropological theory, proposals concerning Kayah Li history, and ethnographic description of the Eastern Kayah Li of Mae Hong Son, Thailand.)
- Solnit, David B. (1997) *Eastern Kayah Li: Grammar, Texts, Glossary*. Honolulu: University of Hawaii Press. (Errata list available by emailing dbsolnit@sbcglobal.net. Errata are especially numerous in the transcription of vowels in Glossary items.)
- (2006) ‘Verb serialization in Eastern Kayah Li’, in Alexandra Y. Aikhenvald and R.M.W. Dixon (eds) *Serial Verb Constructions: A Cross-Linguistic Typology* (Explorations in Linguistic Typology), Oxford: Oxford University Press, 144–159. (Revises the analysis of verb serialization.)

CHAPTER FIFTY

PWO KAREN

Atsuhiko Kato

1 INTRODUCTION

Pwo Karen is a member of the Karenic branch of Tibeto-Burman. Its closest relative is probably Sgaw Karen with more distant relationships to the other languages of the Karenic branch, such as Pa-O, Kayan, and Karen (Kayah—see Solnit, this volume, on Eastern Kayah Li).

Given findings reported in the studies of Purser and Saya Tun Aung (1922), Jones (1961), Kato (1995, 2009b), Phillips (1996, 2000), and Dawkins and Phillips (2009a, 2009b), I tentatively divide the dialects of Pwo Karen into four groups: Eastern Pwo Karen, Western Pwo Karen, Htoklibang Pwo Karen, and Northern Pwo Karen. The Eastern Pwo dialects are spoken in Karen State, Mon State, and Tenasserim Division of Myanmar, and in western Thailand. The Western Pwo dialects are spoken around the Irrawaddy delta. Htoklibang Pwo, which most likely consists of a single dialect, is spoken in Mon State, and the Northern Pwo dialects are spoken in northwestern Thailand. The dialect treated in this chapter is one of the Eastern Pwo dialects spoken around Hpa-an, referred to here as the Hpa-an dialect. The description here is based on my Pwo Karen grammar (Kato 2004).

The above-mentioned dialect groups differ in many aspects, and are mutually unintelligible largely due to phonological and semantic differences. For example, in the eastern dialects /cáiñ/ [teain⁵⁵] means ‘to walk,’ but the cognate word in the western dialects is /sáiñ/ [sain¹¹], which means ‘to run.’ The dialects are not particularly different at the syntactic level, although there are some striking differences. For example, the eastern dialects have a causative construction that takes a complement clause, such as /jə ʔánmén lá ʔəwé yê/ [1sg order COMP 3sg come] ‘I ordered him to come,’ but the western dialects have no such construction. Instead, in the western dialects, the same thing is expressed by using the causative auxiliary /pánmè/ that is cognate with eastern /pánmén/, i.e. /jə ʔánmè yái ʔəwé/ [1sg CAUS come 3sg].

The exact number of Pwo Karen speakers is unknown. According to the estimated population statistics published by the Myanmar government in 1993, there were 2.86 million Karen in Myanmar, but there is no indication how many of these were Pwo Karens. An estimate would be that over one million Pwo Karens live in Myanmar, and as there are also a number of Pwo Karens living in Thailand, the whole population is probably between one and two million. Increasingly, however, Pwo Karens are shifting their language to Burmese and Thai under the influence of these neighbors.

Several writing systems have been created for the Pwo Karen dialects (Stern 1968; Kato 2006). The most widely used in Myanmar are the Monastic script and the Mission script. The Monastic script was created for one of the eastern dialects (perhaps for the Hpa-an dialect) and is largely based on the Mon script. Its history is poorly understood, but surviving records written in the Monastic script date back to the middle of the nineteenth century (U Phon Myint 1975). It is now coming into widespread use throughout

Karen State, as it is taught in many Buddhist monasteries. The Mission script, based on the Christian Sgaw Karen alphabet, was originally invented for one of the eastern dialects by American missionaries around the 1840s, but it is not commonly used in the eastern areas where Buddhists overwhelmingly outnumber Christians. The mission script is, however, presently popular in the Irrawaddy delta, where Christians have been increasing in number. Unfortunately, the script does not altogether suit the phonological systems of the Western dialects.

2 SOUNDS

The syllable structure of the Hpa-an dialect can be represented as $c_1(c_2)v_1(v_2)(c_3)/\tau$, where c_1 is an onset consonant, c_2 is the second member of a cluster, v is a vowel, c_3 is a coda, and τ is a tone. In this structure, $-v_1(v_2)(c_3)$ is referred to as the rhyme.

The consonant phonemes are as follows:

Stops

p	θ	t	c	k	?
ph		th	ch	kh	
b		d			

Fricatives

ε	x	h
y	v	

Nasals

m	n	(ŋ)	(ɳ)	n
---	---	-----	-----	---

Semivowels

w	j
---	---

Liquids

l (r)

The consonants /p/, /ŋ/, and /r/ are bracketed because they occur mainly in loanwords from Mon and Burmese. The /θ/ is realized either as an interdental stop or as a fricative. Younger generations tend to pronounce it as a stop, whereas older generations more often use a fricative. Both /c/ and /ch/ are affricates: [tε] and [tεʰ]. The /b/ [b] is consistently imploded and the /d/ [d] is sometimes realized as an egressive [d]. All of the consonants except /n/ can occur as onsets, but only the consonants /w/, /l/, /t/, and /j/ can occur as the second member of a cluster. In the coda, only /n/, which is a uvular nasal with very weak closure or sometimes simply nasalizes the preceding vowel, can occur. In some dialects, including the western Kyonbyaw and the eastern Tavoy dialects, /?/ also can occur as a coda (Kato 1995). In the Hpa-an dialect, however, the final /?/ has already disappeared.

The rhymes are as follows:

(a) Plain rhymes

Simple vowels	Diphthongs
i [ʃi]	i
u	ai
ɛ [ɪ]	au
e	o
a	ɔ

(b) Nasalized rhymes

Simple vowels	Diphthongs
əN an	ein ain
ON	əUN oUN

(Note: The rhyme /an/ is pronounced as [əN] or [aN]. The final /N/ of /ein/, /əUN/, and /oUN/ is frequently dropped, especially when pronounced at natural speed.)

There are four tones:

High-level	/má/ [ma ⁵⁵]	Pronounced with normal voice.
Mid-level	/mā/ [ma ³³ ~ma ³³⁴]	Pronounced with breathy voice, and sometimes accompanied by a rising contour, especially in utterance-final position.
Low-level	/mà/ [ma ¹¹]	Pronounced with normal voice.
Falling	/mâ/ [ma ⁵¹]	Pronounced with normal voice, and sometimes with slightly creaky voice.

When two falling tones are morphologically juxtaposed in a single word, the first falling tone often changes to the low-level tone:

- (1) /thî/ ‘water’ + /khlân/ ‘hot (of water)’ > /thikhhlân/ ‘green tea’

The Hpa-an dialect has atonic syllables, in which /ə/ is the only permitted vowel, and which never occur in utterance-final position. They are represented with no tone mark.

Intonation sometimes distorts the pitch contours of the tones. For instance, tones of verbs before the perfective particle /jào/ often have a rising contour [224], which resembles the contour of the mid-level tone:

- (2) /phlā bá [224] jào/
arrow hit PERF
‘The arrow has hit (the mark)!’

Since the verb /bá/ has the high-level tone, it may be pronounced as [55], but when it is pronounced as [224], the sentence clearly indicates that the speaker hoped the event would occur. All of the tones of verbs may be pronounced as [224] before the particle /jào/. The influence of this intonation pattern can also be observed in the verbs of yes–no interrogative sentences and the verbs of sentences answering them, as in (3).

- (3) /nə mə yê [224] ɪâ/
2sg IRR come Q
‘Are you really coming?’
/mə yê [224]/
IRR come
‘Yes, of course!’

This contour indicates doubt or suspicion in a question, and a strong belief or confidence in an answer.

3 PARTS OF SPEECH

Two Karen words can be divided into those that can constitute an utterance in isolation (i.e. verbs, nouns, adverbs, and interjections) and those that cannot (i.e. particles).

Beginning with a description of words, other than interjections, that can constitute an utterance in isolation, verbs can either be preceded or followed by verb particles. For example, /yê/ ‘to come’ is a verb, as it can be preceded by the preverbal particle /mə/ (irrealis):

- (4) /mə yê/
IRR come ‘(I) will come.’

Nouns cannot be preceded or followed by verb particles, as shown by the word /phlòUN/ ‘person; Karen’ in (5), and a noun can be the argument of a verb, as in (6).

- (5) */mə phlòUN/
IRR person
- (6) /phlòUN ?ó lə yà/
person exist one CLF:person
‘There is one person.’

Adverbs can neither co-occur with verb particles nor be the argument of a verb, as shown by /lèpòUN/ ‘much’ in (7) and (8).

- (7) */mə lèpòUN/
IRR much
- (8) */lèpòUN ?ó/
much exist

In contrast to the above word classes, particles cannot constitute an utterance in isolation. In this regard, they are similar to affixes (discussed later). In this chapter, these relatively independent morphemes are referred to as particles, but ultimately it may be impossible to draw a strict dividing line between particles and affixes. Some of the important particles are given in 3.1 to 3.7 (these lists are not exhaustive).

3.1 Adpositions

Pwo Karen has a number of prepositions, such as those illustrated below (see Kato 2010 for detailed descriptions).

/lá/ (~/lé/) ‘in; at’ (location), ‘to’ (goal), or ‘from’ (source):

- (9) /?əwê ?ó lá ?ə yéin/
3sg exist LOC 3sg house
‘He is in his house.’

/dē/ ‘with’ (instrument or accompanier) or ‘and’ (conjunction):

- (10) /jə ?án mì dē nótòUN/
1sg eat rice with spoon
‘I eat rice with a spoon.’

- (11) /khòθá dē θákwiθá/
mango and banana
‘a mango and a banana’

There is also the circumposition /bê . . . ðò/ ‘like, as’:

- (12) /?án mì bê pəjàn ðò/
eat rice like Burman like
‘(He) eats rice like a Burman.’

3.2 Demonstratives

Two Karen has three demonstratives, illustrated in (13) to (15) below. They follow nouns.

- (13) /χéin jò/
house this
'this house'
- (14) /χéin nó/
house that
'that house'
- (15) /χéin ?ò/
house that
'that house' (very far)

Demonstratives are often used as topic markers (especially /nó/):

- (16) /?ə χéin nó jə lì ?é/
3sg house TOP 1sg go NEG
'To his house, I didn't go.'

3.3 Adverbial particles

Adverbial particles function like adverbs but cannot constitute an utterance in isolation. For example:

- (17) /χì mā/
good very
'(It) is very good.'
- (18) /χì pəθài ?é/
good not.so NEG
'(It) is not very good.'

3.4 Preverbal particles

Kato (2004) lists 11 preverbal particles. These indicate various meanings, including irrealis, obligation, and causation:

/mə/ is the irrealis marker. There is no marker for realis.

- (19) /jə mə lì/
1sg IRR GO
'I will go.'

/bá/ means 'must, have to':

- (20) /jə bá χé kâ/
1sg must come Q
'Do I have to come?'

/dà/ is a causative marker that is used when the verb is a volitional verb, as in (21). When a causative marker is used, the causee occurs after the verb, and the original object follows it.

- (21) /jə dà ?án ?əwê khòθá/
1sg let eat 3sg mango
'I let him eat a mango./I made him eat a mango.'

/mà/ is a causative marker that is used when the verb is a non-volitional verb, as in (22), where /thé/ 'to be cut' is non-volitional.

- (22) /jə mà thé phli/
1sg make be.cut rope
'I cut the rope.'

As stated in Kato (2009a: 89), Pwo Karen lacks semantically and phonologically related pairs of intransitive and transitive verbs, as are found in many other Tibeto-Burman languages, e.g. Burmese /pyaʔ/ 'to be cut' vs /phyaʔ/ 'to cut.' Pwo Karen usually has only an intransitive verb, and the causative marker /mà/ is used in order to express the corresponding transitive situation, as in (22). Other examples include /θí/ 'to die' vs /mà θí/ 'to kill'; /lànθé/ 'to fall' vs /mà lànθé/ 'to drop'; /yàyòn/ 'to break (intr.)' vs /mà yàyòn/ 'to destroy.'

3.5 Postverbal particles

Kato (2004) lists 50 postverbal particles. These particles add a range of meanings, including 'to try,' 'to do in advance,' and various directions:

/jō/ 'try to (do)' (< /jō/ 'to look at'):

- (23) /jə kè jō lái/
1sg write try alphabet
'I tried to write the alphabet.'

/wè/ '(to do) in advance':

- (24) /jə lò wè ?əwê/
1sg tell in.advance 3sg
'I told him in advance.'

/thán/ 'upward movement,' 'increase,' or 'change to a better condition,' etc. (< /thán/ 'to ascend, climb'):

- (25) /khlàin thán/
speak upward
'He spoke upward.'

/làn/ 'downward movement,' 'decrease,' or 'change to a worse condition,' etc. (< /làn/ 'to descend'):

- (26) /kè làn/
write downward
'to write down'

3.6 Subordinate clause markers (see also section 5.7.3)

Subordinate clause markers form adverbial clauses, as shown by /yòn/ 'after' in (27).

- (27) /jə ?án mì yòn, ?ánlū thî/
1sg eat rice after bathe water
'After eating, I took a shower.'

3.7 Sentence-final particles

The sentence-final particles denote various speaker attitudes. For example, /nê/ is used when the speaker seeks the agreement of the hearer:

- (28) /?əjò chêñ mā nê/
this sweet very isn't.it
'This is very sweet, isn't it?'

4 MORPHOLOGY

4.1 Inflectional morphology

Two Karen nouns and verbs do not inflect. Nonetheless, the pronoun paradigm might be termed inflectional. Each pronoun has two forms, labeled Form I and Form II:

	Form I	Form II
1sg	/jə/	/jè/
1pl	/pə/ ~ /hə/	/pà/ ~ /hà/
2sg	/nə/	/nà/
2pl	/nəθí/	/nəθí/
3sg	/?əwê/	/?ə/ ~ /?əwê/
	/?ə/	
3pl	/?əθí/	/?əθí/

A Form I pronoun is used before a verb for subjects, or before a noun to denote a possessor. There are two forms for the Form I 3sg pronoun, namely /?əwê/, used before both verbs and nouns, and /?ə/, used only before nouns. However, /?ə/ is sometimes used for the subject of a subordinate clause. The 3pl form /?əθí/ sometimes becomes /?əθí?ə/ before nouns. Form II pronouns are used after verbs and prepositions, or when pronouns are topicalized. Each pronoun also has an emphatic form: /jəwê/ (1sg), /pəwê ~ həwê/ (1pl), /nəwê/ (2sg), /nəθíwê/ (2pl), /?əwê/ (3sg), /?əθíwê/ (3pl).

4.2 Derivational morphology

As mentioned earlier, it is difficult to draw a strict dividing line between particles and affixes. In this chapter, those morphemes that cannot constitute an utterance in isolation and that are morphologically less independent are termed affixes. The most commonly used affixes are listed below.

4.2.1 Suffixes

/-chā/ is used to form a noun that denotes an owner or an expert:

- (29) /yéinchā/ ‘owner of a house’ < /yéin/ ‘house’
(30) /tōonchā/ ‘dancer of Karen dance’ < /tōon/ ‘Karen “Don” dance’

/-phó/ originated from the noun /phó/ ‘child,’ and is used to form nouns that denote members of a group:

- (31) /təwâñphó/ ‘villagers’ < /təwâñ/ ‘village’
(32) /yéinphó/ ‘family’ < /yéin/ ‘house’

4.2.2 *Prefixes*

/pə-/ (also pronounced as /hə-/) is used to form nouns denoting certain kinds of people:

- (33) /pəθàbán/ ‘young people’ < /θàbán/ ‘young’
(34) /pəphlòon/ ‘Pwo Karen people’ < /phlòon/ ‘Pwo Karen; person’

/?ə-/ is used to form nouns from verbs:

- (35) /?əθá/ ‘fruit’ < /θá/ ‘to bear fruit’
(36) /?ədí/ ‘egg’ < /dí/ ‘to lay (an egg)’

/?è-/ (also pronounced as /?ə-/) is used to form adverbs from verbs:

- (37) /?èphlé/ ‘fast, swiftly’ < /phlé/ ‘to be fast’
(38) /?èblè/ ‘as to be satisfied’ < /blè/ ‘satisfied’

4.2.3 *Reduplication*

Reduplication is used to form adverbs from verbs:

- (39) /phléphlé/ ‘fast, swiftly’ < /phlé/ ‘to be fast’ (/?èphlé/ shown in (37) is more formal than /phléphlé/.)
(40) /yìyì/ ‘well’ < /yì/ ‘to be good’

4.3 Compounding

As in many of the monosyllabic languages of the region, compounds are common in Pwo Karen. Some examples of these are given below (N and v denote nouns and verbs respectively).

4.3.1 *N < N + V*

- (41) /thî?ɔ/ ‘drinking water’ < /thî/ ‘water’ + /?ɔ/ ‘to drink’
(42) /mídwài/ ‘matchstick’ < /mí/ ‘fire’ + /dwài/ ‘to light’

4.3.2 *v < N + V*

- (43) /θàthán/ ‘to be angry’ < /θà/ ‘heart’ + /thán/ ‘to ascend’
(44) /nāγhôN/ ‘to hear’ < /nā/ ‘ear’ + /γhôN/ ‘to hear’

4.3.3 *N < N + N* (In N + N compounds, the second element is generally the head.)

- (45) /méthî/ ‘tear’ < /mé/ ‘eye’ + /thî/ ‘water’
(46) /θítào/ ‘hospital’ < /θí/ ‘medicine’ + /tào/ ‘building’

4.3.4 *v < V + N*

- (47) /chênnâ/ ‘comfortable to listen’ < /chôn/ ‘sweet’ + /nâ/ ‘ear’
(48) /kèpərəN/ ‘to write (a letter)’ < /kè/ ‘write’ + /pərəN/ ‘news’

4.3.5 $v < v + v$

- (49) /chèinxâ/ ‘clean’ < /chêin/ ‘clean’ + /xî/ ‘beautiful’
(50) /χəpnkhən/ ‘stable’ < /χəpn/ ‘stable’ + /khən/ ‘hard’

Concatenated serial verbs (see section 5.5.1) might also be considered verb compounds, as no clear criteria have been found to distinguish serial and compound verbs.

5 SYNTAX

5.1 Word order

The basic Pwo Karen word order is Subject–Verb–Object, as illustrated in (51) and (52).

- (51) /θà?wà thè thwí/
Thawa kick dog
'Thawa kicked the dog.'
- (52) /jə mwē phlòuN/
1sg COP Karen
'I am a Karen.' (/mwē/ is a copulative verb. It cannot be omitted.)

In a ditransitive clause, the recipient occurs directly after the verb, and the theme follows it:

- (53) /jə phílân ?əwê lái?ào/
1sg give 3sg book
'I gave him a book.'

The order of Subject and Verb is retained even in existential and possessive sentences, as can be seen in (54) and (55).

- (54) /lái?ào lə béin ʔɔ́ lí cəpwé ?əphânkhoú/
book one CLF:flat thing exist LOC table top
'There is a book on the table.'
- (55) /jə phó ʔɔ́/
1sg child exist
'I have a child.'

Adverbs and adpositional phrases occur after the verb and the object (if there is one), as in (56).

- (56) /?án mì ʔéphlé/
eat rice fast
'Eat rice fast!'

5.2 Negation

There are two particles that denote negation, namely /?é/ and /lə/. The form /?é/ is an adverbial particle (see section 3.3) and is used in main clauses, as in (57).

- (57) /?əwê ?án mì xèxè ʔé/
3sg eat rice slowly NEG
'He does not eat rice slowly.'

The form /lə/ is a preverbal particle, used in subordinate clauses. When /lə/ occurs before the verb, the syllable /bá/ is usually placed after the verb phrase or immediately after the verb, probably in order to reinforce the phonetic weight of /lə/, as in (58).

- (58) /?əwê ?è lə ?án mì bá nó, jə mə ?án/
 3sg if NEG eat rice NEG that 1sg IRR eat
 (=/?əwê ?è lə ?án bá mì nó, jə mə ?án/)
 ‘If he doesn’t eat the rice, I will eat it.’

5.3 Interrogative sentences

Yes–no interrogative sentences are formed with the sentence-final particle /bâ/:

- (59) /nə mə thàin bâ/
 2sg IRR return Q
 ‘Are you going back?’

In sentences that include an interrogative word, the sentence-final particle /l̩/ occurs. Interrogative words including /chənō/ ‘what’ and /phlòon/ ‘who’ appear in the same position as in assertive sentences, that is, no movement occurs:

- (60) /nə ?án chənō l̩/
 2sg eat what Q
 ‘What did you eat?’

5.4 Noun phrases

Stative verbs and demonstratives follow nouns, as in (61).

- (61) /phlòon yì jò/
 person good this
 ‘this good man’

Pwo Karen has many numeral classifiers that follow numerals. The numeral classifier construction occurs after the noun:

- (62) /thò θən dù/
 pig three CLF:animal
 ‘three pigs’

In possessive constructions, the possessor precedes the possessed. The 3sg pronoun (Form I) /?ə/ is often placed before the possessed noun.

- (63) /θà?wà (?ə) yéin/
 Thawa 3sg house
 ‘Thawa’s house’

Relative clauses are introduced by the relative marker /l̩ə/ (~ /l̩é/). The relative clause follows the head noun when the head noun is identical to the subject of the relative clause, as in (64). In such cases, a pronoun that refers to the head noun occurs in the relative clause. However, both the relative marker and the following pronoun tend to be omitted in informal speech. Relative clauses are often followed by demonstratives.

- (64) /phlòUN (ló ?ə) lì lá ?ə yéin nò/
 person REL 3sg go LOC 3sg house that
 ‘the person who went to his house’

When the head noun is not identical to the subject of the relative clause, the relative clause may either precede or follow the head noun. As is seen in (65), the relative marker /ló/ must occur when the relative clause follows the head noun, and the co-referential pronoun may occur in the relative clause only when the head noun is animate. When the relative clause precedes the head noun, as in (66), the relative marker does not occur. The pattern in (66) is preferred in informal speech.

- (65) /?əkhwâ ló jø dá (?ò) ló dàò phèN nò/
 male REL 1sg see 3sg LOC room inside that
 ‘the man that I saw in the room’
- (66) /jø dá ló dàò phèN ?əkhwâ nò/
 1sg see LOC room inside male that
 ‘the man that I saw in the room’

5.5 Verb serialization

Like other Southeast Asian languages, verb serialization occurs in Pwo Karen. Serialization involving two verbs is examined here, as it is basic in Pwo Karen verb serialization. The first verb is referred to as v1 and the second as v2. Verb serialization in Pwo Karen is more limited than in other Southeast Asian verb-medial languages such as Thai and Vietnamese. In Thai, a neighboring language, it is possible for a noun phrase or adpositional phrase to intervene between serialized verbs, as in /pay talàat súuu plaa/ [go market buy fish] ‘(I) went to the market and bought a fish.’ In Pwo Karen, however, the corresponding serialization is not acceptable, as is illustrated in (67).

- (67) */jø lì phjâ xwè já/
 1sg go market buy fish

In Pwo Karen dialects that have close contact with Thai, this type of verb serialization appears to be acceptable to some speakers, but it is unacceptable in the Hpa-an dialect. To express the equivalent of ‘I went to the market and bought a fish’ using verb serialization in the Hpa-an dialect, it is necessary to put the word /phjâ/ ‘market’ somewhere other than between the two verbs, making it a non-argument, as in (68).

- (68) /jø lì xwè já ló phjâ/
 1sg go buy fish LOC market
 ‘I went and bought a fish at the market.’

Of course, the same scenario can be expressed without using verb serialization at all, for example by means of combined clauses:

- (69) /jø lì phjâ yòn, xwè já/
 1sg go market after buy fish

In Pwo Karen, a noun phrase may only separate two verbs when the second verb denotes the result of the first verb or ability (see the separated type described in section 5.5.2). In such cases, a noun phrase or adpositional phrase may occur between v1 and v2. Example (67) is unacceptable because the serialized verbs express consecutive actions, and v2 denotes neither the result of the first verb nor ability.

Serialized verbs in Pwo Karen can be divided into two groups, namely concatenated and separated serialization, according to the position of the negative marker /lə/ (see section 5.2) when they occur in a subordinate clause, as set out below:

- 1 Concatenated type: the negative marker /lə/ occurs before v1.

- (70) /jə lə lì xwè bá ?əkhúcòn, . . . /
1sg NEG go buy NEG because
'Because I did not go to buy (something), . . .'

- 2 Separated type: the negative marker /lə/ occurs before v2.

- (71) /jə ?ɔ̄ thî lə nī bá ?əkhócòn, . . . /
1sg drink water NEG get NEG because
'Because I couldn't drink water, . . .'

The characteristics of each of these types are discussed below.

5.5.1 *Concatenated type*

In concatenated serialization, v1 and v2 are tightly combined and no other elements intervene between them. Therefore, it might be tempting to view serialized verbs of this type as compound verbs, but the combination of v1 and v2 is relatively free and productive. v1 and v2 are generally arranged in accordance with the temporal order of events. The semantic relationships denoted by the sequence of v1 and v2 include means-purpose and manner-action. There are four possible combinations of v1 and v2, determined according to whether the verbs are intransitive or transitive:

- (a) Intransitive + intransitive: the whole concatenated verb functions as intransitive.

- (72) /?əwê thàin mî/
3sg return sleep
'He went back and slept.'
- (73) /jə chînàñ kòeà/
1sg sit shout
'I sat and shouted.'

- (b) Intransitive + transitive: the whole concatenated verb functions as transitive.

- (74) /jə chînàñ ?ókhò ?əwê/
1sg sit wait 3sg
'I waited for him while seated.'
- (75) /jə lì xwè já lə bén/
1sg go buy fish one CLF:flat thing
'I went and buy a fish.'

- (c) Transitive + transitive: the whole concatenated verb functions as transitive.

- (76) /jə xwè ?án kó/
1sg buy eat cake
'I bought a cake and ate it.'

- (77) /jə phón phəyôon ?əwê/
1sg catch hug 3sg
'I caught and hugged her.'

(d) Transitive + intransitive: the whole concatenated verb functions as transitive.

- (78) /jə dó θí thwí/
1sg strike die dog
'I struck the dog to kill it.'

- (79) /?əwê thào chéin ?ə mé/
3sg brush clean 3sg tooth
'He brushed his teeth to clean them.'

In the patterns (a), (b), and (c), both v1 and v2 are volitional verbs (agentive verbs), and the subject arguments of v1 and v2 are identical. The object arguments, i.e. the arguments which occur as the objects when v1 or v2 are used individually, in pattern (c) are usually, but not always, identical for v1 and v2. Consider the example in (80), where the object arguments are non-identical:

- (80) /jə chûlân ?án mì dē thîlá/
1sg put.in eat rice with salt
'I ate rice after putting salt in it.'

In (80), the object argument of v1 is /thîlá/ 'salt,' and that of v2 is /mì/ 'rice.' Thus, they are non-identical. In such cases, the object of the concatenated verbs must be the object argument of v2. In the example above, /mì/ 'rice' is an argument of v2. The word /thîlá/ 'salt,' which is the argument of v1, cannot occur in the object position for v1 + v2:

- (81) */jə chûlân ?án thîlá/
1sg put.in eat salt
Intended meaning: 'I ate rice after putting salt in it.'

Pattern (d) differs from the remaining patterns in two respects. First, the subject arguments of v1 and v2 are not identical. In (78), the subject argument of v1 is 'I,' which is the subject of the whole v1 + v2 construction, whereas the subject argument of v2 is the 'dog,' which is the object of whole v1 + v2. The subject argument of v2 is always identical to the object argument of v1. Second, the volitionality of v1 and v2 differ in pattern (d), in that v1 is volitional but v2 is non-volitional.

Concatenated verbs usually occur in temporal order. However, there are two exceptions. One such exception is illustrated in (82).

- (82) /jə mî kòn thadòñ/
1sg sleep put.on sarong
'I put on my sarong and went to sleep.'

In this example, the order /kòn mî/ [put.on sleep] is not allowed. It appears that Pwo Karen prohibits the combination transitive + intransitive when the agents of the two verbs are identical, and in such cases the verbs need to be reversed. The second exception is where one of the consecutive events involves movement:

- (83) /jə yê ?án mi/
1sg come eat rice
'I came after having lunch.'

Here again, the order /pán yê/ [eat come] is not allowed. Verbs denoting movement must occur as v1. As this sentence can also mean ‘I came to eat lunch,’ it is semantically ambiguous.

5.5.2 Separated type

In verb serialization with separated verbs, a noun phrase or adpositional phrase may occur between v1 and v2. In such cases, the co-referentiality of the arguments of v1 and v2 vary from case to case. Only non-volitional verbs can occur as v2, as in the examples below. In separated verb serialization, v2 denotes either the result of v1, as in (84) to (86), or ability, as in (87) and (88).

- (84) /pəwê lânthé ðî/
3sg fall die
'He fell and died.'
- (85) /jə pán mì blè/
1sg eat rice full
'I ate rice and got full.'
- (86) /jə dó thwí ðî pōon/
1sg strike dog die suddenly
'When I struck the dog, it happened to die.'
- (87) /nə ʔɔ̄ thî nī/
2sg drink water get
'You can drink water.'
- (88) /jə khlain phlōon bá ?é/
1sg speak Karen right NEG
'I cannot speak Karen.'

The verbs that denote ability constitute a closed class, including /nī/ ‘get,’ /bá/ ‘be right,’ /ké/ ‘become,’ /ðí/ ‘be capable,’ /khlau/ ‘be free,’ and /bawin/ ‘dare.’ These verbs each denote a semantically different kind of ability, and the appropriate verb is determined by the context.

v2 is usually intransitive because most non-volitional verbs are intransitive. However, transitive verbs may occur as v2, as long as the verb is non-volitional. The verb /dá/ ‘find’ is such an example:

- (89) /jə pánxû khánphài dá ?é/
1sg look.for sandal find NEG
'I looked for sandals but I couldn't find (ones).'

What is notable about Pwo Karen verb serialization is the semantic difference between concatenated serialization and separated serialization. Consider the examples below:

- (90) /jə dó ðî thwí/ (Concatenated type)
1sg strike die dog
'I struck the dog to kill it.'
- (91) /jə dó thwí ðî pōon/ (Separated type)
1sg strike dog die suddenly
'When I struck the dog, it happened to die.'

In (90), the death of the dog was expected by the actor from the beginning; thus, the death was the purpose of the striking. In (91), in contrast, however, the death occurred unexpectedly or accidentally. Thus, sentence (91) automatically implies the death of the dog, whereas (90) does not necessarily do so. This fact can be demonstrated by placing a clause meaning ‘but it did not die’ after the clauses (90) and (91), as in (92) and (93).

- (92) /jə́ dó̄ ðí̄ thwí, lānān θí̄ ðí̄ ?é/
1sg strike die dog but also die NEG
'I struck the dog to kill it, but it did not die.'
- (93) */jə́ dó̄ thwí̄ ðí̄, lānān θí̄ ðí̄ ?é/
1sg strike dog die but also die NEG

As is clear from the ungrammaticality of (93), the addition of this clause to a clause containing separated serialization yields a contradiction. This is because the serial verbs in (91) imply a result, whereas the concatenated verbs as in (90) do not. This is similar to verb serialization in Kayah, where concatenated verbs like (90) also do not necessarily imply a result (Solnit 1997: 68).

5.6 Agent-defocusing

Pwo Karen does not possess a passive construction. As Myhill (1997) and Sansò (2006) discussed, one of the most important functions of the passive voice in many languages is to make the agent less prominent, i.e. the agent-defocusing function. In Pwo Karen, agent-defocusing is achieved by using the noun /chə/ ‘thing,’ as the subject, as in (94).

- (94) /chə́ dó̄ ?əwē/
thing hit 3sg
'He was hit (by someone).' (Literally: 'A thing hit him.')

The use of /chə/ in the subject slot makes the agent less prominent. This construction is used when the speaker does not need to refer to the agent, does not know the agent, or wants to emphasize the patient.

5.7 Sentences with multiple clauses

5.7.1 Complement sentences

Complement sentences can be embedded in matrix sentences without any marking. The examples in (95) and (96) illustrate complement sentences embedded as the subject.

- (95) /[hə́ ?án chədòchəlā] yì̄ mā/
1pl eat vegetable good very
'It is very good for us to eat vegetables.'
- (96) /[?əwḗ klí] phlé/
3sg run fast
'He runs fast.' (Literally: 'His running is fast.')

The complementizer /lə/ (~ /lé/) may occur before a complement sentence that is embedded as the object of a matrix sentence, as in (97).

- (97) /jə dá [(ló) ?əwê klí]/
1sg see COMP 3sg run
'I saw him running.'

5.7.2 Coordinated clauses

Conjunctions are used to coordinate clauses, as illustrated by the use of /lānān/ 'but' in (98) to mark the coordination of two clauses. This form can also be used as a subordinate clause marker (see section 5.7.3). When used as a conjunction, it may be preceded by a pause.

- (98) /jə lì pəjān khān, lānān θí ?əwê lì θāiN khān/
1sg go Burma country but also 3sg go Thai country
'I went to Burma, but he went to Thailand.'

5.7.3 Adverbial clauses

Adverbial clauses usually precede main clauses. They are formed by using various subordinate clause markers. Each marker occurs in a different position in the subordinate clause. In the examples in (99) to (101), subordinate markers occur clause-finally.

- (99) /jə ?ó lə máu bá ?əkhócòn, lənjò jə lì ké ?é/
1sg exist NEG healthy NEG because today 1sg go become NEG
'Since I'm not fine, I can't go today.'
- (100) /pàothán pàitâlân yòn, chəphóxā náo làn ləpòUN mā l̥é/
open window after insect enter downward much very EMPHASIS
'After (I) opened the window, many insects came in.'
- (101) /lái?àu ?ó ləpòUN lānān θí, jə pō máu ?é/
book exist much but also 1sg read comfortable NEG
'Although there are many books, they aren't fun.'

The particle /lānān/ in (101) may occur immediately after the verb, as in (102).

- (102) /lái?àu ?ó lānān ləpòUN θí, jə pō máu ?é/
book exist although much also 1sg read comfortable NEG

The subordinate clause marker /kəlā/ occurs clause-initially, as in (103).

- (103) /kəlā lə ?án mì dài bá, əjá khwái wī nə cú/
before NEG eat rice still NEG wash thoroughly first your hand
'Wash your hands before you eat rice.'

The subordinate clause marker /?è/ occurs before the verb, as in (104).

- (104) /màncò əbjò ?è yê, dà ?ó ləxi/
uncle Thabyaw if come let drink don't
'If uncle Thabyaw comes, don't let him drink (liquor).'

Finally, the subordinate clause marker /bē . . . θò/ 'so as to' surrounds a clause, as in (105). Note that /bē . . . θò/ that denotes 'like, as' (see (12) in section 3.1) is an adposition whereas /bē . . . θò/ that denotes 'so as to' is a subordinate clause marker. I consider them to be homonyms.

- (105) /bê jə mə ðí ðò, jə klicì châ mā/
 so.as.to 1sg IRR capable so.as.to 1sg endeavor much very
 'I endeavored so much so that I could (do it).'

REFERENCES

- Dawkins, Erin and Audra Phillips (2009a) *A Sociolinguistic Survey of Pwo Karen in Northern Thailand*. Chiang Mai: Linguistic Department, Payap University.
- Dawkins, Erin and Audra Phillips (2009b) *An Investigation of Intelligibility between West-Central Thailand Pwo Karen and Northern Pwo Karen*. Chiang Mai: Linguistic Department, Payap University.
- Jones, Robert B. Jr. (1961) *Karen Linguistic Studies*. Berkeley and Los Angeles: University of California Press.
- Kato, Atsuhiro (1995) 'The phonological systems of Pwo Karen dialects', *LTBA* 18.1: 63–103.
- Kato, Atsuhiro (2004) A Pwo Karen Grammar. PhD dissertation, University of Tokyo. (In Japanese)
- Kato, Atsuhiro (2006) 'Dōitsugengonai ni okeru mojihukyūjōkyō no sai ni tsuite: Pōkarengō no jirei [Difference in prevalence of writing systems in the same language: a case of Pwo Karen]', in: Shiohara Asako and Kodama Shigeaki (eds) *Hyōki no Shūkan no Nai Gengo no Hyōki [Writing Unwritten Languages]*, 89–110. Tokyo: Research Institute for Languages and Cultures of Asia and Africa. (In Japanese)
- Kato, Atsuhiro (2009a) 'Valence-changing particles in Pwo Karen', *LTBA* 32.2: 71–102.
- Kato, Atsuhiro (2009b) 'A basic vocabulary of Htoklibang Pwo Karen with Hpa-an, Kyon-byaw, and Proto-Pwo Karen forms', *Asian and African Languages and Linguistics* 4: 169–218.
- Kato, Atsuhiro (2010) "Case" in Pwo Karen', in: Sawada Hideo (ed.) *Grammatical Phenomena of Tibeto-Burman Languages I: Case-marking and Related Matters*, 311–330. Tokyo: Research Institute for Languages and Cultures of Asia and Africa. (In Japanese)
- Myhill, John (1997) 'Toward a functional typology of agent defocusing', *Linguistics* 35: 799–844.
- Phillips, Audra (1996) 'Dialect comparison among the Pwo Karen of central Thailand', *Proceedings of the Fourth International Symposium on Languages and Linguistics* Vol. III: 1122–1162.
- Phillips, Audra (2000) 'West-central Thailand Pwo Karen phonology'. Paper presented at the 33rd ICSTLL, Ramkhamhaeng University, Bangkok.
- Purser, W.C.B. and Saya Tun Aung (1922) *A Comparative Dictionary of the Pwo-Karen Dialect*. Rangoon: American Baptist Mission Press.
- Sansö, Andrea (2006) "Agent defocusing" revisited', in: Werner Abraham and Larisa Leisiö (eds) *Passivization and Typology, Form and Function*, 232–273. Amsterdam/Philadelphia: John Benjamins.
- Solnit, David B. (1997) *Eastern Kayah Li: Grammar, Texts, Glossary*. Honolulu: University of Hawaii Press.
- Stern, Theodore (1968) 'Three Pwo Karen scripts: a study of alphabet formation', *Anthropological Linguistics* 10.1: 1–39.
- U Phon Myint (1975) *Buddha Bhaasaa Pui: Karang Pecaa Samuing (A History of Palm-leaf Inscriptions of Buddhist Pwo Karen)*. Rangoon: Sapreuu:caapetuik. (In Burmese)

§3.9 Other languages

CHAPTER FIFTY-ONE

LEPCHA

Heleen Plaisier

1 INTRODUCTION

The Lepcha people call themselves *róngkup* ‘children of the Róng’, in full *mítuncí róngkup rumkup* ‘Children of the Róng and of God’, and their language is called *róngring*. The term ‘Lepcha’ derives from Nepali *Lāpce* or *Lāpcā*, which originally had the derogatory connotation of ‘inarticulate speech’. Nowadays, the term ‘Lepcha’ is widely used without this connotation.

The Lepcha language is spoken in Sikkim and Darjeeling district in West-Bengal of India, the Ilām district of Nepal and in south-western Bhutan, altogether by upwards of 30,000 speakers. The Lepcha people divide themselves into four main groups according to the region they inhabit, i.e. *támsángmú* (from Kalimpong, Kurseong, Mirik and Darjeeling), *renjongmú* (from Sikkim), *Pilammú* (from Ilām) and *promú* (from Bhutan). The four groups do not represent different dialects; the regional differences between the Lepcha spoken in different areas are largely lexical. The genetic position of Lepcha within Tibeto-Burman is still unclear, despite the efforts of Hodgson (1857), Waddell (1899), Konow (1909), Shafer (1950), Forrest (1962), Benedict (1972) and Bodman (1988, 1989), with some questioning whether it is Tibeto-Burman or not.

Archibald Campbell published a first note on the Lepcha language in 1840 (Campbell 1840). Since 1845, translations of parts of the Bible into Lepcha have been published. In 1876, Colonel (later General) George Byres Mainwaring published a grammar of Lepcha. His romantic view of Lepcha as the *Ursprache* greatly influenced later studies of Lepcha, and his extensive work stimulated an interest in the language throughout the following century. In 1898, the Lepcha–English dictionary compiled by George Mainwaring was edited and published by Albert Grünwedel (Mainwaring 1898). These works were followed by different short accounts of the Lepcha language and by several anthropological studies on the Lepchas, the most important one being the work of Halfdan Siiger and Jørgen Rischel (1967). Prabhakar Sinha (1966) wrote a grammar of Lepcha as an unpublished PhD dissertation at Deccan College in Pune. Richard Keith Sprigg has written numerous valuable articles on the Lepcha language, the indigenous Lepcha script and Lepcha history (Sprigg 1966a, 1966b, 1982, 1983, 1986 and many others). Native Lepcha scholars such as Khárpú Támsáng and Dóngtshen Luksóm (1986) have published important studies on the Lepcha language and culture written in the Lepcha language. This chapter is based on fieldwork conducted by myself in the Lepcha-speaking area between 1994 and 1998. A longer grammatical description, papers on the native Lepcha orthography, descriptions of Lepcha manuscripts and translations of Lepcha texts have also been published (Plaisier 2003, 2005, 2006, 2011, 2012).

2 PHONOLOGY

The inventory of Lepcha consonant phonemes is given in Table 51.1, with a phonetic transcription between square brackets and the transcription used throughout this chapter in italics. The distinction between certain pairs of consonants is fading, i.e. /ph/ versus /f/, /j/ versus /z/, and /v/ versus /w/. The distinction between /s/ and /sh/ is neutralized before the vowel /i/.

Lepcha has eight phonemic vowels, which are listed in Table 51.2. There is no phonological vowel length, although a vowel in an open syllable tends to have a longer realization than the same vowel in a closed syllable. The distinction between the phonemes /o/ and /ö/ is clearly phonetically differentiated in reading pronunciations and the cultivated enunciations of many speakers. However, the distinction is lost in the speech of other speakers, particularly those highly fluent in Nepali, where no comparable phonological distinction exists.

The transcription used in this chapter is a faithful transliteration of the native Lepcha orthography, which is consistent with the way Lepcha is written in traditional texts. This makes it possible to derive the spelling in original Lepcha orthography from the transliteration. However, from Table 51.2 it can be seen that the native orthography is not wholly phonological. The native orthography distinguishes between symbols that do not or no longer represent a phonological distinction, i.e. *i* vs *î* and *a* vs *â*. The circumflex accent used in the transcription of the vowel phoneme /a/ is a convention in transliterating Lepcha script which dates back to Mainwaring and stems from the transcription of the diacritic flourish known as the *rān* or ‘circumflex’ sign in the native Lepcha orthography. In Lepcha writings, the circumflex sign is mainly present in closed syllables, where it may have originally indicated stress or pitch (Plaisier 2003: 28–9, 2006: 37; Sprigg 1983: 316).

Each syllable in Lepcha contains a vowel phoneme at its core. The syllable contains an initial consonant or consonant cluster, and may or may not contain a final consonant. While all consonants may occur as syllable-initial consonants, not all consonants may be combined with the post-consonantal glides -*y*- and -*r*-, nor with post-consonantal -*l*-. All possible combinations for initial consonant clusters are listed in Table 51.3. The following consonants may appear in syllable-final position: /k/, /t/, /p/, /ng/, /n/, /m/, /r/, /l/.

Lepcha morphemes are monosyllabic. Many Lepcha words are composed of different syllables, with the stress usually on the second syllable. Richard Keith Sprigg analyses Lepcha as having contrastive stress (1966b: 199–200). The prosody of Lepcha is characterized by the clustering of syllables into groups; often these are phrases or groups of words marked by suffixes, postpositions or particles. The prosody of Lepcha from Northern Sikkim is markedly different to that of Lepcha from other regions, in that it is much slower and much less monotonous. This type of intonation is believed to represent an older and more elegant style of speaking, less influenced by neighbouring languages such as Nepali or Dränjoke.

3 NOMINAL MORPHOLOGY

Nominals comprise the following parts of speech: nouns, pronouns, adjectives and numerals. Nouns lack a grammatical gender distinction and show no agreement with articles, adjectives or verbs. However, Lepcha has rich derivational nominal morphology, which includes gender-specific derivational suffixes. Affixing, compounding and reduplicating represent the major derivational processes. Nominals may take case endings, i.e. suffixes

TABLE 51.1 CONSONANTS

	Voiceless stop	Aspirated voiceless stop	Voiced stop	Voiceless affricate	Aspirated voiceless affricate	Voiceless fricative	Voiced fricative	Voiced nasal	Voiced trill	Voiced approximant	Voiceless approximant
Labial	[p] <i>p</i>	[p ^b] <i>p̪</i>	[b] <i>b</i>	[d] <i>d</i>	[s] <i>s</i>	[f] <i>f</i>	[v] <i>v</i>	[m] <i>m</i>	[w] <i>w</i>		
Dental	[t] <i>t</i>	[t ^b] <i>t̪</i>			[θ] <i>θ</i>			[n] <i>n</i>	[l] <i>l</i>		
Alveolar											
Palato-alveolar						[s] <i>s</i>	[z] <i>z</i>				
Retroflex	[ʈ] <i>tr</i>	[ʈ ^b] <i>tr̪</i>	[ɖ] <i>dr</i>		[ʃ] <i>ʃ</i>	[ʃ] <i>ʃ</i>	[ʒ] <i>ʒ</i>				
Palatal	[c] <i>c</i>	[c ^b] <i>c̪</i>						[ɲ] <i>ny</i>			
Velar	[k] <i>k</i>	[k ^b] <i>k̪</i>		[g] <i>g</i>				[ŋ] <i>ng</i>			
Glottal	[?] <i>?</i>								[h] <i>h</i>		

TABLE 51.2 VOWELS

Vowel phoneme	Phonetic realizations	Transliteration
/a/	[ə ~ ʌ ~ u̯]	a, â
/á/	[a ~ a̯]	á
/i/	[i]	i, í
/o/	[o ~ u]	o
/ó/	[ɔ̯]	ó
/u/	[u̯ ~ i]	u
/ú/	[u]	ú
/e/	[e ~ ɪ ~ ε̯]	e

TABLE 51.3 INITIAL CONSONANT CLUSTERS

	-y-	-r-	-ry-	-l-	-ly-
k	ky	kr	kry	kl	kly
kh	khy	tr	try		
g	gy	gr	gry	gl	gly
ng		ngr			
t	ty				
th	thy				
d	dy				
p	py	pr	qry	pl	ply
ph	phy				
f	fy	fr	fry	fl	fly
b	by	br		bl	bly
m	my	mr	mry	ml	mly
r	ry				
l	ly				
h	hy	hr	hry		hy
		thr	thry		
v	vy				

or postpositions attached to the nominal. Two or more case endings may co-occur attached to a single noun when this yields a desired and semantically plausible combination.

Plural number in nouns is expressed by the suffixes <-pang> and <-sang>. The non-human plural suffix <-pang> (PL.NH) is used to express plurality of animals, things, indeed, all entities except humans, including evil spirits and demons. The human plural suffix <-sang> (PL.H) indexes the plurality of human beings and personified, anthropomorphic beings such as benevolent heavenly creatures, gods, goddesses and good spirits. The plural morphemes <-sang> (PL.H) and <-pang> (PL.NH) are directly suffixed to the nominal constituent which they modify. This entails that the suffixes precede all case endings, postpositions, articles or demonstrative pronouns which modify the nominal, but that when a nominal consists of a noun immediately followed by a modifying adjective, the morphemes <-sang> (PL.H) and <-pang> (PL.NH) follow the adjective, i.e. they follow the nominal constituent as a whole.

- (1) *bik-pang*
cow-PL.NH
'cows'
- (2) *rum-sang*
god-PL.H
'gods'
- (3) *?álóng muró-?átím-sang ma-nyí-ne*
now man-big-PL.H NEG-have-NEG
'There are no tall men around at the moment.'

I have not been able to corroborate the existance of a dual suffix <-nyum>, as mentioned by Mainwaring (1876: 27), Sinha (1966: 67) and Tamsang (1978: 10). My findings indicate that the Lepcha plural category denotes manifoldness, and that the Lepcha plural suffixes are not used when a nominal is modified by a numeral denoting a quantity greater than one, e.g. *cho nyet* 'two books', *?akup fali* 'four kids'. The meaning of the morpheme <-pang> is actually 'thing, things', whereas the meaning of the morpheme <-sang> is 'whole, entire'. The Lepcha plural ending <-sang> can be used with proper names or kinship terms, where it denotes a group of more than one, e.g. *Nyíma-sang* 'Nyima and his friends'.

The numeral *kát* 'one' is not an indefinite article as such, but can be used in some of the same functions as the indefinite article in English, to express an explicit singular number. Lepcha *kát* 'one' conveys the sense of 'a certain' or 'a', e.g. *muró-kát* 'a man, one man, a certain man'. The suffix <-re> (DEF) is the Lepcha definite article. The suffix <-re> (DEF) is attached to nouns, e.g. *muró-re* 'the man'. Like the definite article in English, the Lepcha definite article fulfils a thematic or even contrastive function. The fact that the numeral *kát* 'one' is a numeral is underscored by instances such as *kalók-kát-re* 'the one mouse' or 'that mouse', in which *kát* is used in combination with the definite article <-re>. The Lepcha definite article <-re> (DEF) is etymologically cognate with the root found in the Lepcha demonstratives *?áre* 'this' and *?ore* 'that', and it seems obvious that the meaning of the definite article is derived from the deictic function of the demonstrative. Indeed, many Lepcha speakers use the definite article <-re> as a stopgap when they pause mid-sentence to formulate their thoughts. The morpheme <-mu>, which is analysed by Mainwaring (1876: 23) and Tamsang (1978, 1994) as another definite article, is in fact an individuative suffix for people, evidently cognate with the first element in both Lepcha *muzu* 'body' and *muró* or *maró* 'person'.

A number of Lepcha postpositions can be combined with constituents of different syntactic status, i.e. with nouns, with verbs and with entire clauses. The apparent difference in sense expressed by such morphemes, e.g. the locative morpheme <-ká> (LOC) and the ablative suffix <-nun ~ -nu> (ABL), is a straightforward function of the difference in syntactic status of the constituent which they modify, and it can be shown that the morpheme in question still expresses the same basic meaning.

The meaning of the ablative suffix <-nun ~ -nu>, phonetically [nun ~ nui], is one of source or cause. Some native speakers and Lepcha authors consistently use the form <-nun> after nouns and the form <-nu> after verbs. This is generally regarded as proper or correct usage. However, this would appear to be a cultivated norm, because in practice most speakers and writers do not in fact observe this distinction. One sense which the suffix <-nun ~ -nu> expresses when affixed to nouns and adverbs denoting a place or time is the ablative, namely to mark the source, origin or point of departure of the activity denoted by the main verb, in contexts comparable in function to English 'from'.

-
- (4) *saróng-nun go cholí-ká nón-sho*
today-ABL 1sg school-LOC go-NPST
'As of today I shall go to school.'
- (5) *darjyúlyáng-nun kalenpúng-tet . . .*
Darjeeling-ABL Kalimpong-till
'From Darjeeling to Kalimpong . . .'
- (6) *hó sabá-nu*
2sg where-ABL
'Where are you from?'

In this ablative sense, the ablative suffix <-nun ~ -nu> contrasts with the root <lóm> (< PTB *lam ‘road, direction’), which occurs as a lative suffix ‘via, by means of which, from’, as a noun with the meaning ‘road, way’ and as a verb with the meaning ‘walk’.

There are a number of Lepcha suffixes which are similar in meaning but distinct from both the ablative and the lative suffixes. One of these is the suffix <-ren>, which has the straightforward temporal meaning ‘since’, e.g. *tasó-ren* ‘since yesterday’. Another such suffix is the ending <-kón> ‘side, towards, in the direction of’. The noun *lyáng* ‘place, land, country’ is used as a general locational noun.

- (7) *kásu-lyáng kóm ma-nyi-ne*
1sg-place money NEG-have-NEG
'I don't have money on me.'

The ablative suffix <-nun ~ -nu> can also be affixed to nouns to fulfil an agentive function. The suffix <-nun ~ -nu> is used to highlight the agentive character of an inanimate entity which acts as the agent of an activity (9), whether this be transitive or intransitive. In this function, <-nun ~ -nu> is more likely to occur when the verb is transitive and when the subject is animate, because, in these cases the agentive character of the subject is either more obvious or more pronounced (8). The agentive suffix is more likely to occur in the past tense because the agentive meaning of the Lepcha ergative is more likely to be appropriate in contexts where the activity expressed has actually taken place already (10).

Another Lepcha morpheme which is appropriate to discuss in this context is the intensifier <-do> ‘self’, which accentuates the identity of the referent denoted by the constituent to which it is suffixed, in the sense of ‘this very one, by himself’, which can follow the semantic role marker.

- (8) *?ályu-nun-do kalók-kát sót ma*
cat-ABL-self mouse-one kill AST
'The cat killed a mouse.'
- (9) *saróng-sá so-nu kásu-sá nyót lók-hát*
today-GEN rain-ABL 1sg-GEN field damage-lose
'Today's rain ruined our fields.'
- (10) *tungvyeng thok-nón, tú-nu thok*
door close-go who-ABL close
'The door is closed, who closed it?'

The genitive suffix <-sá> (GEN) has both a genitive and an instrumental function. As a marker of genitive relationships the ending <-sá> expresses possession, part-whole

relationships and related semantic functions. Genitive usage of the suffix <-sá> is straightforward and ubiquitous in the language.

- (11) *?áre kásu-sá lí go ma*
this 1sg-GEN house be AST
'This is my house.'
- (12) *kásu-sá ?ákâ nyet nyí ma*
1sg-GEN hand two have AST
'I have two hands.'

The ending <-sá> can be used to indicate the instrument or means by which an action is enacted or takes place, e.g. (13)–(15).

- (13) *go ?úng-sá ?ákâ cóng-sho*
1sg water-GEN hand wash-NPST
'I wash my hands with water.'
- (14) *go nyúgú-sá pi-sho*
1sg pen-GEN write-NPST
'I write with a pen.'
- (15) *go ?ámík-sá ngák-bám*
1sg eye-GEN look-PROG
'I look with my eyes.'

The instrumental sense of the genitive morpheme <-sá> may be contrasted with the ablative suffix <-nun ~ -nu>. The ablative suffix marks the agent of the action denoted by the main verb, rather than an instrument by which the action is performed. The genitive ending <-sá> is especially used when the referent of the constituent it modifies is the obvious choice by means of which to enact the situation denoted by the main verb, i.e. when the activity denoted by the verb, in a manner of speaking, belongs to the implement. In the following examples, both Lepcha <-sá> and <-nun> are used. The difference in meaning is that in example (16) the pen is depicted as a means or point of origin and therefore highlighted as the means by which the writing takes place, whereas in example (17) the pen is merely mentioned as the obvious means with no special highlighting of its function as the implement.

- (16) *nyúgú-nu shú zúk-sho*
pen-ABL what make-NPST
'What do you do with/by means of a pen?'
- (17) *nyúgú-sá shú zúk-sho*
pen-GEN what make-NPST
'What do you do with a pen?'

The postposition *dep* 'together with' indicates accompaniment, and is often used in conjunction with or as an alternative to the Lepcha genitive suffix <-sá>.

- (18) *hó hu-dep nú*
2sg 3sg-with go
'You, go with him!'

The meaning of the locative suffix <-ká> (LOC) can be attached to nominals, verbs or entire clauses. When the locative morpheme <-ká> is suffixed to a noun, it denotes the site of an activity or the destination toward which an activity is directed.

- (19) *sáthang* *sáryók-ká* *bám-sho*
 tiger jungle-LOC reside-NPST
 'The tiger lives in the jungle.'

- (20) *go* *Kalenpíung-ká* *nóng-sho*
 1sg Kalimpong-LOC go-NPST
 'I am going to Kalimpong.'

The dative suffix <-m> (DAT) indicates the goal or site of an activity. When an activity expressed by a verb is directed 'to' or 'for' someone or something, the goal is marked by the dative suffix. When affixed to a nominal denoting an animate referent, the locative suffix <-ká> can be contrasted with the dative suffix <-m>, which also marks an entity towards which the action or situation expressed by the verb is directed, but in a much more personal way than the locative suffix <-ká>.

- (21) *go* *kaju-ká* *kám* *zo* *Bi*
 1sg dog-LOC little.bit food Give
 'I gave some food to the dog.'
- (22) *go* *kaju-re-m* *kám* *zo* *Bi*
 1sg dog-DEF-DAT little.bit food Give
 'I gave the dog some food.'

The personal pronouns differentiate three persons and three numbers. The personal pronouns are listed in Table 51.4. The third person denotes human referents only. When the referent is non-human, a demonstrative is used, except in the case of fables or other tales in which animals are personified, where the third person pronoun is used to refer to these animals. The form *ka*, which is found in certain expressions such as *ka mítuncí róngkup rumkup* 'we children of the Róng and of God', is a basic or reduced form of the first person plural form *káyú*.

The Lepcha genitive suffix <-sá> is used to derive possessive pronouns. Singular possessive pronouns are formed by suffixing the marker <-sá> to the oblique forms of the singular pronouns, e.g. *kasu-sá lí* 'my house', *ʔádo-sá cho* 'your book', *hudo-sá yuk* 'her letter'. To form dual or plural possessive pronouns, the genitive suffix is added to the dual and plural personal pronouns listed in Table 51.4, e.g. *káyú-sá cho* 'our book', *huyú-sá cho* 'their book', etc.

The major demonstrative and corresponding interrogative pronouns are listed in Table 51.5. The proximate deictic morpheme <ʔá-> refers to something or someone near to the speaker and the distant deictic morpheme <ʔo-> indicates something or someone far from the speaker. Three other deictic elements commonly encountered are <pe-> 'over here', <me-> 'down there' and <tá-> 'up there'. These elements may also be prefixed to the article *re*, e.g. *pere* 'that over there', *mere* 'that down there', *táre* 'that up there', or to other deictic elements such as the morpheme <-bá>, which indicates a

TABLE 51.4 PERSONAL PRONOUNS

	Singular	Singular oblique	Dual	Plural
1st person	<i>go</i>	<i>kásu</i>	<i>kányí</i>	<i>káyú</i>
2nd person	<i>hó</i>	<i>ʔádo</i>	<i>ʔányí</i>	<i>ʔáyú</i>
3rd person	<i>hu</i>	<i>hudo</i>	<i>hunyí</i>	<i>huyú</i>

TABLE 51.5 DEMONSTRATIVE AND INTERROGATIVE PRONOUNS

<i>?áre</i> ‘this’	<i>?ore</i> ‘that’	<i>sare</i> ‘which’
<i>?ábá</i> ‘here’	<i>?obá</i> ‘there’	<i>sabá</i> ‘where’
<i>?álom</i> ‘like this’	<i>?olom</i> ‘like that’	<i>salom</i> ‘how’
<i>?átet</i> ‘this much’	<i>?otet</i> ‘that much’	<i>satet</i> ‘how much’

location relative to the speaker, e.g. *mebá* ‘there below’. The interrogative morpheme <sá-> is used interrogatively and in relative constructions. The interrogative pronouns may also be used as relative pronouns. In addition to the interrogative pronouns listed here, Lepcha speakers make frequent use of the question words *tú* ‘who’ and *shú* ‘what’. Notions such as ‘everywhere’ and ‘everyone’ are expressed in Lepcha by using interrogative pronouns reinforced with the morpheme *lá* ‘also’, e.g. *sabálá* ‘everywhere’, *sathálá* ‘always’. When these are combined with a negative verb they convey the senses ‘nowhere’, ‘never’, etc.

The Lepcha equivalent to an English adjective may be an adjective or a verb expressing a state or a condition. Many adjectives are formed from verbs by adding the prefix <-?á-> and sometimes suffixing <-m> to the verb root, e.g. *ryú* ‘be good’, *?áryúm* ‘good’. Adjectives may be used adnominally, predicatively, or independently as nominal heads. Adjectives may be modified by intensifiers, such as *?ágýáp* ‘much, very’, or *do* ‘self’.

Comparative constructions are formed by means of the postposition *len* ‘than, compared to’, which follows the element it modifies. A superlative meaning is expressed by comparison with a totality expressed by *gun* ‘all, every’, e.g. *gun-len nahan* [all-than before] ‘before everyone else’.

4 VERBAL MORPHOLOGY

All Lepcha verbs have in principle two stem forms, namely a regular and an inflected stem. Most verb stems end in a consonant. All of the verbs with a stem-final consonant and a minority of the verbs with an open stem are invariable. In other words, for a majority of verbs, the regular and the inflected stem are one and the same. However, the majority of the verbs which regularly show an open stem exhibit an inflected stem with a final consonant before auxiliary verbs. The consonants which occur as final in such inflected stems are /t/, /n/ and /m/, e.g. inflected *lín* vs regular *lí* ‘speak’, *zom* vs *zo* ‘eat’ and *dit* vs *di* ‘reach’. The inflected stem cannot be predicted on the basis of the form of the verb.

The verb *bi* ~ *bo* ‘give’ also has two different stems, but in this case the choice for one or the other stem depends on the recipient of the verb. The stem *bo* is used when something is given to a first or second person singular or plural recipient, e.g. *kasum bo* ‘give to me’. The stem *bi* is used when something is given to a third person singular or plural recipient, e.g. *hudom bi* ‘give to him/her’.

The Lepcha verb does not conjugate to show agreement for person and number. Tense, mood and other meanings of the verb are expressed by the use of endings and auxiliary verbs. Mood can be signalled through the use of modal auxiliaries, verbs like ‘can’, ‘may’, ‘shall’ and ‘must’, which indicate a wide range of moods, such as permission, possibility, intention or necessity. The modal of exigency is the verb *gát* ‘must’, which expresses a need or desire to do something. To be able to do so in the sense of being in a position to do something or being physically capable of doing something is expressed

by the verb *khu* ~ *khut* ‘can’. To be able to do something in the sense of knowing how to perform a certain task or activity is expressed by the verb *yâ* ‘know, know how to’. To be allowed or permitted to do something is expressed by the verb *kón* ‘let, allow’. The verb *ngün* can be used as a main verb expressing the meaning ‘become, happen, occur’, and when it is used as a modal verb it expresses whether or not an action is seen as necessary, allowed or all right.

The unmarked form of the verb in Lepcha indicates a preterite tense, which denotes actions anterior to the speech moment without reference to result or duration. The zero form of the verb may express just the transpiration of an event in past time, or it may express that the situation has only just started to take place, essentially a change of state. When the subject is marked by the ablative suffix <-nun ~ -nu>, sentences can often be interpreted as having a past meaning.

- (23) *hu nón ma*
3sg go AST
'He went away.'
- (24) *go-nun yuk pi ma*
1sg-ABL letter write AST
'I wrote a letter.'
- (25) *hlo-ká sozóng-sho*
peak-LOC be.cold-NPST
'It is cold on the hill.'
- (26) *go vám-kát theng-sho*
1sg song-one sing-NPST
'I will sing a song.'

A non-preterite or non-past tense is expressed by adding the verbal ending <-sho> (NPST) to the verb. The non-past tense is used to describe a situation or activity which is taking place at the present time, an activity which the speaker is planning to perform, or an event or situation which the speaker is certain or convinced is going to take place soon.

The verbs in Lepcha which cover the senses of English ‘be’ are *go* ‘be’ and *nyí* ‘exist’. The verb *go* ‘be’ is used as an identity marker, to say that X is Y, and so to express the identity or inherent quality of a person, entity or thing. Incidentally, the verb *go* is homophonous with the first person singular pronoun *go*. The form *gum*, glossed as ‘be + AST’, is a contracted form of <*go-ma*>. The assertive particle *ma* (AST) adds force to a statement, and can be translated into English as ‘it is so’, ‘it is the case that’. The assertive particle is commonly used with both copulas, even with the fully lexicalized form *gum*. The verb *nyí* ‘have’ covers the attributive, locative and existential senses of English ‘to

be', and so it may be used to ascribe a quality to someone or something, to indicate the whereabouts of the subject of a sentence, or in an existential sense it may indicate the presence or availability of a person, commodity or thing.

- (27) *?áre kásu-sá lí go ma*
this 1sg-GEN house be AST
'This is my house.'
- (28) *lyáng ?áre rong-kup bám-lyáng gum*
land this Lepcha-child live-land be + AST
'This land is the homeland of the Lepchas.'
- (29) *vom ?ákrím nyí ma*
salt bitter have AST
'Salt is bitter.'
- (30) *kásu-sá lí ngase-ká nyí ma*
1sg-GEN house Ngase-LOC have AST
'My house is in Ngase.'
- (31) *kásu-sá ?ákup nyet nyí ma*
1sg-GEN child two have AST
'I have two children.'

There are a number of verbal auxiliaries expressing *Aktionsarten* in Lepcha which in other contexts may occur as main verbs, such as *hát* 'lose, leave behind', *tho* 'put', and *lel* 'complete'. The exhaustive auxiliary *tho* indicates that an activity or action has come to an end. The completive auxiliary *lel* indicates that an activity was carried out to completion. The resultative auxiliary *nón* indicates that a state or an event exists as a result of a transition in the past.

When the locative suffix <-ká> is affixed to a verb, it exhibits two different functions. The meaning of direction yields the notion of a supine when the locative suffix <-ká> is affixed to a verb, and produces an adhortative meaning when suffixed to an entire clause. The supine expresses the sense 'in order to' and is attached to a main verb which thereby becomes the verbal complement, denoting a situation towards which the activity denoted by the main verb is directed. Supine forms marked by the locative suffix <-ká> may appear as complements of verbs of motion as well as of other verbs.

- (32) *go lyem-ká non-det ma*
1sg play-LOC go-move AST
'I am going to play.'

The second function of the locative morpheme <-ká> in combination with a verb is an adhortative function. The supine verbs serve as complements of the main verb of a syntagm or clause, whereas adhortative forms in <-ká> serve as main verbs themselves.

- (33) *kanyí-dep theng-ká*
1dl-together sing-LOC
'Come on, sing along with the two of us!'
- (34) *róng cho rok-ká*
Lepcha book read-LOC
'Let's read Lepcha!'

With verbs, too, the suffix <-nun ~ -nu> simultaneously expresses the notions of cause and source. Verbs in Lepcha inherently involve a temporal dimension and verbs marked with <-nun ~ -nu> consequently have temporal implications such as anteriority of the situation denoted by the verb.

- (35) *?ázóm zóm-lel-nu rok-ká*
rice eat-complete-ABL read-LOC
‘After you have eaten your food, you should study.’
- (36) *?ámik ?ok-nu ngák gang gun shi*
eye open-ABL look if all see
‘If you open your eyes and look, you will see everything.’

Negation of verbs is expressed by means of the negative prefix <ma-> in combination with the negative suffix <-ne>, e.g. *hu ma-nóng-ne* [3sg NEG-go-NEG] ‘he did not go’. The negative affixes can be seen as a single discontinuous morpheme and are attached to the regular stem of the verb. When the stem of the verb is open, the negative suffix <-ne> may be shortened to <-n>, e.g. *ma-nyí-n* [NEG-be-NEG] ‘there is not’.

5 CLAUSE-FINAL PARTICLES, COORDINATION AND SUBORDINATION

Several clause-final particles are used in Lepcha to express the mood or emotional attitude of the speaker towards what he or she is saying. The clause-final particle *?o*, glossed as declarative (DECL), marks the end of a statement and is used predominantly in written text. The clause-final assertive particle *ma* (AST) has been mentioned above. The request particle *le* expresses a polite request and is used when a person wishes to express reassurance towards the addressee or request the addressee to do something. The authoritative particle *ce* expresses authority on the part of the speaker and is used to urge the addressee to do something. The dubitative particle *te* indicates doubt, uncertainty or even confusion about the precise nature or result of an event. The possibility particle *pú* expresses a hypothetical possibility. Inference or assumption is expressed by the morpheme *lyók*, which can also be used as a main verb meaning ‘resemble, look like’. The certainty particle *pá* combines an element of direct perception or direct observation with an element of certainty. The discovery particle *yámbá* marks information that the speaker has come to know, whether acquired directly through observation or through inference. Something that was said by someone else may be marked by the reported speech particle *mere*, whereas the quotative particle *yang* marks a direct quote.

The coordinative suffix *?án* ‘and’ coordinates arguments. This suffix connects two or more elements of a sentence, such as noun phrases, verb phrases or independent clauses. The alternative conjunction *yángne* ‘or, either’ indicates the existence of an alternative. The adversative conjunction *shenlá* ‘but’ expresses a contrast or opposition between two propositions. The most common subordinating conjunctions in Lepcha are *gang* ‘if’ and *gorünglá* ‘although, even’.

The most important nominalizer in Lepcha is the factitive marker <-bú> (NOM), which nominalizes verbs as well as clauses, and these nominalized constituents serve as adnominal attributes or as subordinated clauses. When used with verbs, the nominalizing factitive marker <-bú> (NOM) conveys an imperfective meaning and indicates a state or a situation as such. When a nominalized verb or clause functions as a nominal head it can be pluralized and take case endings if the result is semantically plausible.

6 TEXT

[*renjong lyáng ?áre-re*] [*ká mútunci-róng-kup rum-kup-sang-sá*
Renjong land this-DEF 1pl arch-Lepcha-child god-child-PL.H-GEN
?yá ?yá lúngtyen-nu-do thop-yut-bú-sá
long.ago long.ago legend-ABL-self obtain-send.down-NMLZ-GEN
?áyít ?ágerek ?álát ?ábám ?án ?ámák lyáng] *gum ma ?o*
creation birth arrival residence and death land be+AST AST DECL

This land of Renjong has been the land of creation, birth, residence and death of us Lepchas, sons of God, ever since the oldest times and legends onwards.

[*lyáng-?áre-sá nahán-mú ?áyít róng ?ábryáng-re*]
land-this-GEN before-most creation Lepcha name-DEF
[ne-mayel máluk renjong lyáng] gum ma ?o
Ne-Mayel Máluk Renjong land be + AST AST DECL
?án róng-sang-nu-re huyú-do-sang-re-m mútunci-róng-kup-rum-kup
and Lepcha-PL.H-ABL-DEF they-self-PL.H-DEF-DAT arch-Lepcha-child-god-child
yang lúngprýá-nu lí-bám ma ?o
thus repeat-ABL say-PROG AST DECL

The oldest Lepcha name of this land is Ne Mayel Máluk Renjong, and the Lepchas always call themselves Mútunci Róngkup Rumkup.

róng lungtyen-sá ?áyít-sung ká-róng-ság-sá pímnyothing-re-m-re
Lepcha legend-GEN creation-story 1pl-Lepcha-PL.H-GEN ancestor-DEF-DAT-DEF
kóngchen kónghlo-cú-bi dóngyóng-sá gun-lá len ?áthó-bú
Himalayas-mountain-peak mountain.range-GEN all-also then high-NMLZ
cú kíngsúm zóngbú-cú-sá ?ányák-sá ?átsóng-bú sanóng-lóm
mountain Kanchenjunga-mountain-GEN summit-GEN pure-NMLZ snow-via
?áyítbú-debú-rum-nu zük-nu ?áyít-thóm-bú ma ?o
Áyítbú-Debú-god-ABL make-ABL creation-end-NMLZ AST DECL

According to the traditional history of origin, the forefathers of us, Lepchas were created and made by the God Ayítbú Dyebú out of pure snow of the top of the highest mountain of the Himalayas, mount Kanchenjunga.

(The full Lepcha text of this traditional myth can be found in the book *Róng lúngtyen sung. Lepcha Myths*, written by Kharpu Támsáng and privately published by Lyángsóng Támsáng (Támsáng 1996).)

ADDITIONAL ABBREVIATIONS

- AST assertive
PL.H plural human
PL.NH plural non human

REFERENCES

- Benedict, Paul K. (1972) *Sino-Tibetan. A Conspectus*, Cambridge: Cambridge University Press.
- Bodman, Nicholas (1988) ‘On the place of Lepcha in Sino-Tibetan. A lexical comparison’, *LTBA* 11 (1): 1–26.
- Bodman, Nicholas (1989) ‘Some remarks on Lepcha vowels’, in David Bradley, Eugénie J. A. Henderson and Martine Mazaudon (eds) *Prosodic Analysis and Asian Linguistics. To Honour R. K. Sprigg*, Canberra: Pacific Linguistics, 137–141.
- Campbell, Archibald (1840) ‘Note on the Lepchas of Sikkim, with a vocabulary of their language’, *Journal of the Asiatic Society of Bengal* 9: 379–393.
- Forrest, R.A.D. (1962) ‘The linguistic position of Róng (Lepcha)’, *Journal of the American Oriental Society* 82: 331–335.
- Hodgson, Brian Houghton (1857) ‘Comparative vocabulary of the languages of the broken tribes of Népal’, *Journal of the Asiatic Society of Bengal* 17: 73–83.
- Konow, Sten (1909) Vol. III ‘Tibeto-Burman family’ in G. A. Grierson (ed.) *Linguistic Survey of India*, Calcutta: Superintendent of Government Printing, India.
- Luksóm, Dóngtshen (1986) *Mútuncí Ríngthryum un Ríngsuktóm. A Lepcha Grammar and Composition for Classes IX & X*, Gangtok: Department of Education, Government of Sikkim.
- Mainwaring, George Byres (1876) *A Grammar of the Rong (Lepcha) Language as it Exists in the Dorjeling and Sikim Hills*, Calcutta: Baptist Mission Press.
- Mainwaring, George Byres (1898) *Dictionary of the Lepcha Language*, revised and completed by Albert Grünwedel, Berlin: Unger Brothers.
- Plaisier, Heleen (2003) *Catalogue of Lepcha Manuscripts in the van Manen Collection*, Leiden: Kern Institute Miscellanea.
- Plaisier, Heleen (2005) ‘A brief introduction to Lepcha orthography and literature’, *Bulletin of Tibetology* 41 (1): 7–24.
- Plaisier, Heleen (2006) *A Grammar of Lepcha*, Leiden: Brill.
- Plaisier, Heleen (2011) ‘A key to four transcription systems of Lepcha’, in Mark Turin and Bettina Zeisler (eds) *Himalayan Languages and Linguistics: Studies in Phonology, Semantics, Morphology and Syntax*, Leiden: Brill, 41–53.
- Plaisier, Heleen (2012) ‘Two Lepcha delük texts’, in Nathan W. Hill (ed.) *Medieval Tibeto-Burman Languages IV*, Leiden: Brill, 23–109.
- Shafer, Robert (1950) ‘Classification of some languages of the Himalayas’, *Journal of the Bihar Research Society* 34: 192–214.
- Siiger, Halfdan and Jørgen, Rischel (1967) *The Lepchas: Culture and Religion of a Himalayan People*, Copenhagen: Gyldenal, Nationalmuseets Skriften, Etnografisk række 11.
- Sinha, Prabhakar (1966) *A Descriptive Grammar of Lepcha*, unpublished PhD dissertation, Pune: Deccan College.
- Spragg, Richard Keith (1966a) ‘The glottal stop and glottal constriction in Lepcha and borrowing from Tibetan’, *Bulletin of Tibetology* 3: 5–14.
- Spragg, Richard Keith (1966b) ‘Lepcha and Balti-Tibetan; tonal or non-tonal languages?’, *Asia Major* new series 12 (2): 185–201.
- Spragg, Richard Keith (1982) ‘The Lepcha language and three hundred years of Tibetan influence in Sikkim’, *Journal of the Asiatic Society* 24: 16–31.
- Spragg, Richard Keith (1983) ‘Hooker’s expenses in Sikkim, an early Lepcha text’, *Bulletin of the School of Oriental and African Studies* 46 (2): 305–325.

- Sprigg, Richard Keith (1986) ‘The syllable finals of Tibetan loanwords in Lepcha orthography’, *Linguistics of the Tibeto-Burman Area* 9 (1): 27–46.
- Tamsang, Khárpú (1978) *A Grammar of the Lepcha Language*, Kalimpong: Lyangsong Tamsang.
- Tamsang, Khárpú (1994) *The Lepcha–English Encyclopedic Dictionary*, Kalimpong: Mayel Clymit Tamsang.
- Támsáng, Kharpu (1996) *Róng língtyen sung. Lepcha Myths*, Kalimpong: Lyángsóng Támsáng.
- Waddell, Lawrence Austine (1899) ‘The Lepchas or Róng and their songs’, *Internationales Archiv für Ethnographie* 12: 41–57.

CHAPTER FIFTY-TWO

TUJIA

Shixuan Xu, Meiyuan Lu and Hongyan Hu

1 INTRODUCTION

Tujia, a language of the Tibeto-Burman language family, derives its name from the official designation of its speakers, who were in 1957 recognized as one of the official minority nationalities in China, the Tujia Nationality. Tujia is distributed in the northwest of Hunan province in China. It comprises two dialects: the northern dialect, the speakers of which call themselves *bizika* [pi²⁴tsi⁵⁵kʰa²¹], and the southern dialect, whose speakers call themselves *mojihe* [mo²¹tei³³xio³⁵]. The two dialects differ greatly in terms of their phonology, lexicon and grammar. Hence, they are mutually unintelligible.

The Tujia ethnic minority is an aboriginal ethnic group with a very large population, which, according to the 6th National Census of 2010, is 8,353,900. They are distributed in an area across the shared borders of provinces of Hunan, Hubei, and Guizhou, as well as Chongqing Municipality. Owing to the long-term influence of Chinese and Chinese culture, most of the Tujia people have abandoned the Tujia language and shifted to Chinese, and the remaining native speakers are all bilingual, though to different degrees. The Tujia speakers are becoming fewer and fewer in number. Today, the northern dialect is only spoken in a few villages and townships, straddling the common borders of Longshan, Yongshun, Guzhang, and Baojing counties, with a total population of about 50,000. The southern dialect is distributed in only six or seven villages of Tanxi Township, Luxi County, with only a few hundred speakers. Tujia has become highly endangered.

It still remains unclear as to the specific genetic affiliation of Tujia within the Tibeto-Burman language family. It was subsumed under the Yi branch of the Lolo-Burmese subgroup in the earliest publications by Chinese scholars (Wang 1955). Tian *et al.* (1986) considered the specific genetic affiliation of Tujia to be inconclusive within the Tibeto-Burman family of languages. More recently, He Tianzhen (2003), after a comparison of cognate words and patterns of phonological change, postulated that Tujia could be a member of the Qiangic branch. Xu Shixuan (2011), synthesizing the history of the Tujia ethnic group and the Tujia language, and the historical evidence, inferred that Tujia was a unique language that split off from the parent language of all the Tibeto-Burman languages very early. She also pointed out that Tujia had been subjected to heavy influence from the Yi language and Chinese successively, so that the exact historical origin of Tujia will require in-depth research, weeding out influence on the lexicon layer by layer.

In this chapter, we will offer a description of the northern dialect spoken in Tasha Rural Township, Longshan County.

2 PHONOLOGY

Tujia is a syllable tone language. A syllable usually consists of an onset, a rhyme, and a tone. Onsets are always single consonants. Rhymes consist of monophthongs, diphthongs,

triphtonghs, and monophthongs or diphthongs with a nasal coda. There are four contrastive tones, with rising and falling contours.

2.1 Onsets

Tujia has 19 onsets, which are all single consonants. Plosives and affricates are all voiceless and can be divided into aspirated and unaspirated ones. There are five fricatives, two of which are voiced. More details are shown in Table 52.1.

Notes on the onset system:

1. In a zero initial syllable which begins with a vowel sound (not a glide), there is a mild glottal stop at the beginning of the syllable.
2. In the pronunciation of young speakers, the phonological contrast between the two onsets /n/ and /l/ are often neutralized before rhymes of high vowels, due to the influence of Southwestern Mandarin.

2.2 Rhymes

Tujia has 26 rhymes, comprising seven monophthongs, eight diphthongs, and three triphthongs, as well as eight other compound rhymes of monophthongs or diphthongs plus a nasal coda, as shown in Table 52.2. The nasal coda can be realized phonetically as two different nasals, one dental and one velar, but there is no phonological contrast between them in coda position.

TABLE 52.1 THE TUJIA ONSETS

Manner	Voice	Phonation	Place			
			Labial	Dental	Palatal	Velar
Plosives	voiceless	unaspirated	p	t		k
		aspirated	p ^h	t ^h		k ^h
Affricates	voiceless	unaspirated		ts	tɕ	
		aspirated		ts ^h	tɕ ^h	
Fricatives	voiceless			s	ç	x
	voiced			z		ɣ
Nasals			m	n		ŋ
Approximants			l			

TABLE 52.2 THE TUJIA RHYMES

Monophthongs	ɿ	i	a	i	o	u	ɤ
Diphthongs		ie	ei				
		ia		ai		ua	
		iu		au	ou		
Triphthongs		iau				uai	uei
Nasal codas		in	en	an		uan	uen
		ian					
			ion		on		

Notes on the rhyme system:

1. In zero onset syllables, the initial vowel /i/ or /u/ in compound vowel rhymes is pronounced with a certain degree of friction, similar to the semivowels [j] or [w] in this case.
2. The nasal coda has two conditioned variants: after the vowel /o/, its actual realization is [ŋ]; after the other vowels, it is pronounced as [n].

2.3 Tone

There are four tones in Tuja: high level (55), high rising (24), low falling (21), and high falling (53). Words with these four tones are illustrated in Table 52.3.

2.4 Syllabic structure

Every syllable of Tuja bears one of the four tones mentioned. The syllabic structure can be divided into two kinds: a consonant onset plus a rhyme, and a single syllabic rhyme. Illustrating examples are as follows:

Consonant onset plus a rhyme:

li²¹ ‘speak’ kai⁵³ ‘how much / many’ liau²⁴ ‘throw’,
xen²⁴ ‘hate’ tian⁵⁵ ‘crazy’

Single syllabic rhymes:

a²⁴ ‘write’ ai⁵⁵ ‘that’ uei⁵⁵ ‘thin’
an²⁴ ‘we’ ian⁵⁵ ‘tobacco’.

2.5 Sandhi

This section deals with external sandhi in Tuja, that is, phonological processes that occur at word boundaries. Two morphophonological processes are prominent, tone sandhi and fusion of syllables across word boundaries.

2.5.1 Tone sandhi

In Tuja, tone sandhi constitutes one major type of morphophonemic alternation. Regular tone sandhi occurs in classifiers and particles. Generally, following words of tone 55 and tone 24, classifiers and particles are uttered with their base tones, whereas after words of tone 53 and tone 21, their tones are subjected to change. Two groups of examples are provided as illustration. For instance, *nie⁵⁵ o⁵⁵* ‘two + CL,’ *xi²⁴ o⁵⁵* ‘ten + CL,’ *o²¹ o²¹* ‘six +

TABLE 52.3 THE TONES OF TUJA

Tone description	Tone numbers	Examples	
High level	55	s ₁ ⁵⁵ ‘month’	pa ⁵⁵ ‘to paste’
High rising	24	s ₁ ²⁴ ‘to give’	pa ²⁴ ‘to strip’
Low falling	21	s ₁ ²¹ ‘to grow’	pa ²¹ ‘slope’
High falling	53	s ₁ ⁵³ ‘to pull’	pa ⁵³ ‘to look’

CL,’ and *ka⁵³ o²¹* ‘how many + CL.’ In the preceding four examples, the classifier *o⁵⁵* keeps the underlying tone 55 after *nie⁵⁵* and *xi²⁴* in the first two examples, while after *o²¹* and *ka⁵³* in the last two examples it bears the tone of 21, which is the sandhi tone. In the other groups of examples, *z⁷⁵⁵ lu²¹* ‘to have gone to do,’ *ka²⁴ lu²¹* ‘to have gone to eat,’ *a²¹ lu⁵⁵* ‘to have gone to fetch,’ and *py⁵³ lu⁵⁵* ‘to have gone to carry,’ the particle *lu²¹* keeps the base tone 21 after *z⁷⁵⁵* and *ka²⁴* in the first two examples, whereas it bears the sandhi tone 55 after *a²¹* and *py⁵³* in the last two examples.

2.5.2 Fusion of syllables

In Tujia, the post-verbal elements are sometimes subjected to fusion of syllables due to the grammatical function of adjacent words. Specifically, the grammatical words, dropping their onsets and tones, are reduced to a bare rhyme or a partial rhyme, which is then fused to the preceding syllable to form a new word. For example the perfect aspect particle *liau⁵⁵* may be reduced in this way, and then two new words, *tau²⁴* and *t^hau⁵³*, are formed by merging the reduced phonetic form of *liau⁵⁵* with the preceding negation particles *ta²⁴* and *t^ha⁵³*. Another example is the fusion of *liau⁵⁵* with the preceding modal verbs, e.g. *t^ha⁵⁵ to²¹ liau²¹* → *t^ha⁵⁵tau²¹* ‘need not’; *t^ha⁵⁵ts^hy²¹ liau²¹* → *t^ha⁵⁵ts^hau²¹* ‘be not willing to.’ The morphophonemic alternation, syllable fusion, is common in Tujia, and is indicated in corresponding sections in the following.

3 MORPHOSYNTAX

3.1 Morphology

According to the number of morphemes in a word, the Tujia words can be divided into monomorphemic words and polymorphemic words. Monomorphemic words contain only one morpheme, and their meanings and structures are not analyzable. In terms of syllables contained, monomorphemic words may be classified into monosyllabic and multi-syllabic ones, e.g. *teie²⁴* ‘hand,’ *pa⁵³* ‘to see,’ *su²⁴su⁵⁵* ‘snow,’ *ts^ha²⁴ts^ha⁵⁵* ‘magpie,’ *mi²⁴mi⁵⁵* ‘bamboo shoot,’ and *teian²⁴teian²⁴on²¹* ‘cicada.’ Polymorphemic words, or compound words, will be dealt with later in this section.

Words can be formed by affixation or compounding. Many nouns are formed using affixes. In Tujia, suffixes, which far outnumber prefixes, are very productive. The prefix *a⁵⁵-* is frequently used to form nouns, e.g. *a²⁴ko⁵⁵* ‘friend,’ *a⁵⁵ma⁵⁵* ‘grandmother,’ and *a²⁴ta⁵⁵* ‘sister.’ The following suffixes are derived from grammaticalization of words designating concrete things, maintaining only broad generalized meanings. Besides performing a certain classification function, they can supplement the roots in meaning. To enumerate just a few Tujia suffixes, the suffix *-la⁵⁵* signifies long and thin objects: e.g. *yr²⁴la⁵⁵* ‘string,’ and *pi²⁴la⁵⁵* ‘intestines’; *-pu⁵⁵* signifies granular things: e.g. *li²⁴pu⁵⁵* ‘rice,’ and *te^hi²⁴pu⁵⁵* ‘soybean’; *-ma⁵⁵* is an agentive nominalizer, indicating a person who performs the action described by the root: e.g. *ts^h7²⁴ma⁵⁵* ‘the person who is crying,’ *li²¹ma²¹* ‘speakers’; *-ts^ho⁵⁵* is a locative suffix following the verbal root and indicates the place where the action is taking place: e.g. *ka²⁴ ts^ho⁵⁵* ‘a place for eating’; and *-ts^hy⁵⁵* is an objective suffix following the verbal root, indicating the thing which is involved in the action: e.g. *ka²⁴ ts^hy⁵⁵* ‘food,’ and *ta²⁴ts^hy⁵⁵* ‘clothing.’

Compound words comprise two or more roots which are combined according to certain structural relations, mainly the following four types: two morphemes of similar meaning juxtaposed, e.g. *t^ha²⁴k^hu⁵⁵* [bake + roast] ‘wok’; modifying morphemes

preceding roots, e.g. *za²¹lie²¹* [hen + egg] ‘egg,’ and *ts^hɔ²¹la²¹* [water + canal] ‘ditch’; modifying morphemes following roots, e.g. *u²⁴pa⁵⁵* [cattle + male] ‘ox,’ and *s⁷²¹k^hu⁵⁵t^hu²¹* [meat + fat] ‘fat’; nominal morphemes combined with verbal morphemes, e.g. *tcie²⁴p^ho⁵⁵* [hand + release] ‘stop working for the day.’

3.2 Lexical categories

3.2.1 The grammatical categories of words

Tujia has 11 lexical categories, among which, nouns, verbs, pronouns, classifiers, adjectives, and particles constitute the six major lexical categories.

Nouns are defined by their ability to appear with a quantifier or a numeral-classifier phrase. Verbs are defined by their ability to occur with negation and to be a predicate on their own. More details about nouns and verbs will be stated in the following sections (3.3.1 for nouns and 3.3.2 for verbs).

Pronouns include personal pronouns, demonstrative pronouns, and interrogative pronouns. Personal pronouns differentiate three persons, each including a singular and a plural form. All three plural pronouns have alternative simple forms. Details are shown in Table 52.4, where the bracketed forms are abbreviated forms of plural pronouns. Demonstrative pronouns comprise proximal demonstrative *kai²⁴* (this) and distal demonstrative *ai⁵⁵* (that). They can be used as free pronouns and demonstrative adjectives, and can also be combined with nouns, quantifiers, and combinations of classifiers and numerals. Locative demonstratives can be divided into five scales, distinguished by morphological inflection or affixes. More details can be found in Table 52.5.

The interrogative pronouns include *tc^hie⁵³cie²¹* ‘what,’ *a²¹sy⁵³* ‘who,’ *k^hei²¹tu⁵⁵ / k^hou²⁴ / k^hei²¹* ‘where,’ *tc^hie⁵³sy²¹cie²¹* ‘when,’ *tc^hi⁵⁵tu⁵⁵* ‘which date,’ *kai⁵³* ‘how many / much,’ *tc^hin⁵³mo²¹nie²¹* ‘why.’

Tujia is rich in classifiers. A classifier is a word that reflects the conceptual classification of nouns according to the features of the referents. The classifiers in Tujia are almost all monosyllabic morphemes. Syntactically, a classifier can occur with a noun only by way of a numeral-classifier adnominal phrase. The choice of a classifier is determined by the referent of the noun that precedes it. Nouns indicating items with

TABLE 52.4 TUJIA PERSONAL PRONOUNS

Person	Singular	Plural
1st	ŋa ²⁴	a ²⁴ ni ⁵⁵ (an ²⁴)
2nd	ni ²⁴	sy ²⁴ ni ⁵⁵ (sy ²⁴)
3rd	ko ²⁴	ky ⁵³ tsy ²¹ (ky ⁵³)

TABLE 52.5 TUJIA DEMONSTRATIVE PRONOUNS

Proximal demonstratives	Distal demonstratives		
this / these	kai ²⁴	that / those	ai ⁵⁵
here / near	ky ²¹	there / far	au ⁵⁵
nearer	kau ²⁴	farther	en ⁵⁵ ky ⁵⁵
		much farther	a ²⁴ ky ⁵⁵

similar characteristics often take an identical classifier. Classifiers are characterized by the feature of tone sandhi when combining with a preceding numeral, as shown in section 2.5.

Adjectives are similar to verbs in many ways. For example, both of them have the aspectual category and the directional category. They can both take complements and can be modified by adverbs. However the adjectives cannot take any objects. One prominent characteristic of adjectives is that each adjective has two forms: a base form and a variant version formed from the adjective base by adding suffixes, such as *pa⁵⁵*, *ka⁵⁵la²¹*, *ka⁵⁵ts^ha²¹*, *k^hu⁵⁵t^hu²¹*. For example:

The base forms: *ts^hl⁵³* ‘big’ *uei⁵⁵* ‘thin’

The variant forms: *ts^hl⁵³pa⁵⁵* ‘big’ *uei⁵⁵ka⁵⁵ts^ha²¹* ‘thin’

Both forms of adjective can be used as predicates, and their position in the sentence is the same. So they can be used interchangeably in most cases (1a). However, the variant forms of adjectives have neither aspectual nor directional categories. When used as arguments in a sentence, the basic forms must be nominalized by the nominalizing structural particle *ei⁵⁵* ((1b), (1c)), while the variant forms can be used as arguments directly ((1d), (1e)). These two forms also differ from each other when used as modifiers. The base adjective must be followed by the nominalizing structural particle *ei⁵⁵* to constitute a modifier, which always precedes the head noun (1f). The variant forms can be used as modifiers directly, but they must follow the head noun (1g).

- (1) a. *kai²⁴ ts^ho⁵³ la⁵⁵ ts^ho⁵⁵ ts^hl⁵³ / ts^hl⁵³pa⁵⁵.*
this house one CL big
'This house is very big.'
- b. *uei⁵⁵ ei⁵⁵ pa⁵³ po²¹ ts^ha²⁴.* c. *ŋa²⁴ s¹²⁴ ei⁵⁵ k^hy⁵³.*
thin NMLZ look STA well 1sg fat NMLZ afraid
'Being thin looks very beautiful.'
- d. *uei⁵⁵ka⁵⁵ts^ha²¹ pa⁵³ po²¹ ts^ha²⁴.* e. *ŋa²⁴ s¹²⁴k^hu⁵⁵t^hu²¹ ti⁵³ ei²¹.*
thin look STA well 1sg fat need only
'Being thin looks very beautiful.'
- f. *ts^hl⁵³ nie²¹ ei²⁴ ts^ho⁵³* g. *ts^ho⁵³ ts^hl⁵³pa⁵⁵.*
big ASSOC NMLZ house house big
'big houses'

Clitic particles are considered to be a separate word class because they are bound morphemes with diverse grammatical functions and are very important clausal constituents. All particles are subject to tone sandhi when following a word with tone 53 or 21. *nie⁵⁵* is an associative particle indicating nominal modifiers. It is located between the modifier and the head noun. Modifiers can be pronouns, nouns, or nominal phrases. *nie⁵⁵* can often be omitted when the modifier is a pronoun. In contexts where the head noun following *nie⁵⁵* is well understood by all parties involved, it can be left out as well. Another particle *nie⁵⁵* exists in Tuja. It is a comparative construction marker homophonous with the associative particle *nie⁵⁵*, and will be dealt with in section 3.4.3. *lie⁵⁵* is used to connect two or more verbs serving as the predicate in a sentence. These verbs are equal in syntactic functions but are in order of sequence. The agentive particle *ko⁵⁵* is in some cases optional. However, wherever confusion may arise in differentiating the agent and the patient, *ko⁵⁵*

is obligatory, as in (2a). Aside from this, when there is a need to emphasize the agent, *ko⁵⁵* is sometimes used as well, as in (2b). *ko⁵⁵* has evolved from the grammaticalization of the third person singular pronoun *ko²⁴*.

- (2) a. k_y⁵³ ko⁵⁵ a²⁴ta⁵⁵ xa²¹ liau²¹.
 3pl AGT elder sister hit PRF
 ‘They hit my elder sister.’
- b. k_y⁵³ ko⁵⁵ ie²¹ ka²⁴ tei⁵³ liau²¹.
 3pl AGT food eat finish PRF
 ‘It is they that ate all the food.’

There is an emphatic particle, *to²¹*, which is only used after pronouns acting as subjects, objects, or modifiers, intensifying what the preceding pronouns refer to, as in (3a), (3b), (3c). For example:

- (3) a. ko²⁴ to²¹ ko²⁴ xa²¹. b. n_y²⁴ n_y²⁴ to²¹ lo²¹.
 3sg EMPH 3sg hit 1sg 1sg EMPH scold
 ‘He himself hit him.’ ‘I scolded myself.’
- c. kai²⁴ an²⁴ to²¹ nie²¹ sa²¹.
 this 1pl EMPH ASSOC thing
 ‘This is our own thing.’

There is also a reciprocal particle, *ta⁵³*, which will be dealt with in a later section.

Apart from the aforementioned main lexical categories, Tujia also has five minor categories, which are numerals, postpositions, adverbs, conjunctions, and interjections.

As for numerals, the cardinal numerals in the Tujia lexicon are mostly native Tujia words, while all ordinal numerals, and numerals indicating ‘zero,’ ‘100,’ and ‘100,000,000’ are all Chinese loans.

Postpositions are enclitic to the NPs, denoting diverse meanings such as *lie²¹* (instrumental), *po²¹* (allative or benefactive), *lie⁵⁵* (ablative), *ta⁵⁵* (comitative), as well as *ka²¹*, *tei²¹*, and *t^hu²⁴* (locative) given in Table 52.6.

- (4) a. ni²⁴ te^hie²¹pi²⁴ lie²¹ ka²⁴! b. n_y²⁴ san²⁴xai⁵³ po² yi-²⁴.
 2sg bowl INST eat 1sg Shanghai ALL go
 ‘Please eat with this bowl.’ ‘I am going to Shanghai.’
- c. ko²⁴ pr²¹tein⁵⁵ lie⁵⁵ en²¹ts_y²¹. d. n_y²⁴ ta⁵⁵ k_y²¹ts^h_y²⁴!
 3sg Beijing ABL come 1sg COM play
 ‘He came from Beijing.’ ‘Please play with me!’
- e. ts^h_y⁵⁵phei⁵⁵ s_y²¹t^hie²⁴ ka²¹. f. an²⁴ ts^ho⁵³ k^hu⁵³ts^ha⁵³ tei²¹.
 book desk LOC 1pl house hill LOC
 ‘The book is on the desk.’ ‘Our house is at the bottom of the hill.’
- g. po⁵³li²¹ xu²¹p^ha²¹ t^hu²⁴.
 child river LOC
 ‘The child is in the river.’

In specific contexts, postpositions are frequently subject to tone sandhi.

Adverbs are words modifying verbs or the base forms of adjective. Their positions in the clause are variable depending on their functions. Tujia adverbs form a diverse category, including adverbs of manner, time, and frequency.

TABLE 52.6 LOCATIVE POSTPOSITIONS

Postposition	
ka ²¹ ‘on’	kai ²⁴ s ²¹ t ^h ie ²⁴ ka ²¹ [this desk + on] ‘on this desk’
tei ²¹ ‘under’	ai ⁵⁵ k ^h uai ²¹ tei ²¹ [that chair + under] ‘under that chair’
th ^h u ²⁴ ‘in’	ŋa ²⁴ pei ⁵⁵ pei ⁵⁵ th ^h u ²⁴ [my cup + in] ‘in my cup’

Conjunctions are words connecting words, phrases, or clauses. In Tujia, conjunctions are not indispensable: two phrases and clauses may simply be juxtaposed without any linking conjunctions, and the context is crucial to determine the relationship between them. However, specific conjunctions have to be used to clarify the relationship in some cases. Tujia has a specific conjunction *nie*⁵⁵ (meaning ‘and’) for linking two or more nouns.

Interjections are exclamative words expressing a variety of emotions, such as surprise, shock, and compliments. They have no fixed tones and usually occur at the beginning of a clause.

3.2.2 Grammaticalization of content words

Some content or notional words have undergone grammaticalization as they became more abstract in meaning and came to serve grammatical functions in certain linguistic contexts. For example, some content nouns have evolved into classifiers when they have lost some or all of their precise lexical meaning and started to fulfill a general grammatical function of denoting a kind of object, e.g. a⁵⁵mi²¹ ‘petal’ → ts^hŋ²⁴s⁵⁵nie⁵⁵ mi²¹ ‘two pieces of orange.’ Some locative nouns, due to high frequency of usage following nouns, have undergone semantic bleaching and phonetic reduction and become locative postpositions, with their grammatical functions enhanced, such as ka²¹xa²⁴ ‘top’ → ka²¹ ‘on,’ tei²¹t^ha²¹ ‘bottom’ → tei²¹ ‘under,’ and o²⁴t^hu⁵⁵ ‘inside’ → t^hu²⁴ ‘in.’

In addition to content nouns, notional verbs may also undergo grammaticalization. The following is an example illustrating grammaticalization of the directional verb yi²⁴ ‘go.’ The directional verb yi²⁴ could directly follow action verbs indicating the direction of verbs, which is different in word order from other verbs. This usage of the verb yi²⁴ occurred very frequently. Over time, the high frequency of usage gradually led to the desemanticization and generalization in meaning of this word. Meanwhile, it was fused more and more closely to preceding verbs, which led to phonetic reduction of this word; specifically, it dropped its onset and tone, with only the rhyme -i left. The truncated form -i is never uttered on its own. Instead, it always adheres to the rhyme of the preceding syllable, e.g. pa⁵³ ‘to look’ > pa-i⁵³ ‘go to look.’ What is worth mentioning, furthermore, is that when the tone of the preceding syllable is 21, -i makes it change into 24, e.g. ta²¹ ‘descend’ > ta-i²⁴ ‘go down.’

On the other hand, though reduced phonetically, the reduced form -i does not entirely lose the lexical meaning of the directional verb yi²⁴ ‘go,’ and in some cases still retains a direction implication, implying the actor has to be displaced in order to accomplish the action which the verb of the clause indicates (5a); in other cases, -i concurrently indicates shift of action position and prospective aspect (5b); and still in other cases it only marks prospective aspect (5c) (see Xu and Lu 2014 for a detailed discussion).

- (5) a. ηa^{24} ie^{21} $ka-i^{24}$ xu^{55} . b. ηa^{24} $lau^{24}tsj^{55}$ $ts^h an^{21}kan-i^{53}$.
 1sg food eat INCH 1sg tomorrow go to market-PROS
 'I will go to eat.' 'I will go to the market tomorrow.'
- c. kai^{24} $ts^h \gamma^{24}stj^{55}$ $a^{55}teie-i^{55}$.
 these orange ripe-PROS
 'These oranges will be ripe.'

The above description offers the full story whereby the directional verb yi^{24} 'go' transformed from a full lexical word to an absolute aspectual marker through semantic bleaching and phonetic erosion. The three examples vividly reconstruct the procedure of grammaticalization of the directional verb yi^{24} 'go': first, it suffered a partial loss of its lexical meaning and a phonological attrition, serving little grammatical function, then it was reinforced in grammatical function besides assuming certain lexical meaning, and finally it consummated its grammaticalization by evolving into a pure grammatical marker, indicating an event is about to or going to happen. Compared with the inchoative aspect particle xu^{21} , $-i$ conveys an implication of subjectivity or realis.

3.3 Phrases

3.3.1 Noun phrases

In Tujia, a large number of noun phrases consist of nouns juxtaposed to form a complex NP. The juxtaposed nouns may be connected using the conjunction nie^{55} , but this is optional:

- (6) a. an^{24} $a^{24}ta^{55}$ (nie^{55}) zon^{53} ts^hu^{55} .
 1pl elder sister (and) younger sister home
 'My elder sister and younger sister are at home.'
- b. ηa^{24} $ts^h \gamma^{55}p^hei^{55}$ $ts^h \gamma^{55}k^hi^{55}tha^{55}$ (nie^{55}) pi^{21} ti^{53} .
 1sg book paper (and) pen need
 'I need a book, a piece of paper and a pen.'

Other noun phrases consist of a head noun and modifying elements. Modifying elements include adnominal associative structures, demonstratives, adnominal classifier phrases, nominalized verbs or adjectives, modifying nouns, and/or the plural particle tie^{55} .

Tujia has an associative particle nie^{55} . It is used to form an associative structure by linking words together. This structure precedes the head noun that is being modified. When the modifying word is a pronoun, the particle is often not used. The semantics of the associative structure include possession and modification, e.g.:

- (7) a. ηa^{24} (nie^{55}) o^{21} b. $a^{21}pa^{21}$ nie^{21} ts^ho^{53}
 1sg (ASSOC) back basket stone ASSOC house
 'my back basket' 'houses made of stones'

Within a noun phrase, the two demonstratives kai^{24} 'this / these' and ai^{55} 'that / those' can function either as determiners or pronouns. As the former, they occur as adnominals preceding the head nouns; as the latter, they always function as the head of a noun phrase, e.g.:

- (8) a. kai^{24} $ts^h \gamma^{55}p^hei^{55}$ b. ai^{55} $ts^h o^{53}$
 this / these book that / those house
 'this book / these books' 'that house / those houses'

- c. ni²⁴ nie⁵⁵ kai²⁴ xo²¹ lie⁵⁵! d. ko²⁴ nie⁵⁵ ai⁵⁵ ña²⁴ ti⁵³.
2sg ASSOC this take DIR 3sg ASSOC that 1sg need
'Take this of yours away!' 'I need that of his.'

In a Tuja noun phrase, the head noun precedes the numeral-classifier adnominal phrase, which consists of a numeral preceding a classifier, e.g.:

- (9) k^ha²¹mon²¹ la⁵⁵ mon⁵⁵
tree one CL
'a tree'

There are also other modifying elements in noun phrases, including nouns, verbs, and adjectives. When verbs and the basic forms of adjectives act as modifiers, they must take the associative particle *nie⁵⁵* and the nominalizing structural particle *ei⁵⁵*. The variant forms of adjectives do not need any particles when serving as adnominal modifiers. As far as the position of modifying elements is concerned, nouns and nominalized verbs or the basic forms of adjectives precede the head noun, as in (10a), (10b), (10c), whereas the variant forms of adjectives follow the head noun, as in (10d).

- (10) a. k^ha²¹ khⁱ⁵⁵t^hi⁵⁵ b. tein⁵³ s^j⁵³ nie²¹ ci²⁴ t^hu⁵⁵k^hu⁵⁵
wood basin often.use ASSOC NMLZ cutting knife
'wooden basins' 'the frequently used cutting knife'
c. zr⁵³ nie²¹ ei²⁴ s^j⁵⁵pa⁵⁵ d. s^j⁵⁵pa⁵⁵ zr⁵³k^ho²¹k^huei²⁴
beautiful ASSOC NMLZ clothes clothes beautiful
'beautiful clothes' 'beautiful clothes'

The plural or collective particle *tie⁵⁵* only appears at the end of a noun phrase with either an animate or inanimate referent. *tie⁵⁵* is not obligatory. It is often used to emphasize that there is more than one person or thing, e.g.:

- (11) a. ai⁵⁵ ts^hl⁵⁵p^hei⁵⁵ ka⁵⁵ p^hei⁵⁵ tie⁵⁵ b. kai²⁴ po⁵³li²¹ tie⁵⁵
those book some CL PL these child PL
'those books' 'these children'

When all the elements described above co-occur in a noun phrase, their order can be summarized as follows:

ASSOC modifier → demonstrative → HEAD → nominalized adjective → CL phrase → PL

For example:

- (12) an⁵⁵ja⁵⁵ nie⁵⁵ ai⁵⁵ ts^hei⁵⁵p^hei⁵⁵ a²⁴p^hai⁵⁵ ka⁵⁵ p^hei⁵⁵ tie⁵⁵
younger brother ASSOC this book old some CL PL
'Some of those old books of my younger brother's'

The noun phrase can be followed by a semantic-relation marking postposition to mark the semantic roles of the referent. See section 3.2.1 for examples.

3.3.2 Verbal phrases

In Tuja, verbal phrases consist of verbs and following verbal particles denoting such concepts as aspect, directionality, relevance, modality, and negation.

Tujia has four aspectual verb particles: *xu²¹*, *la⁵⁵*, *lie⁵⁵*, and *liau⁵⁵*, respectively indicate a certain state in event progress or action taking.

xu²¹ marks inchoative aspect, indicating starting or performing of a future event or action described by the verb. When used as the only particle after a verb, it signals an action which is imminent, as in (13a).

la⁵⁵ is used to indicate progressive aspect, i.e. presenting an action as being ongoing rather than referring to its beginning or its end, irrespective of it being a past, present, or future action. In addition, it is also used for habitual actions which are in effect timeless, with the emphasis being placed on the continual performance of the action ((13b), (13c)). When *la⁵⁵* follows the state particle *po⁵⁵*, it indicates that the given situation in this sentence is still going on/continuing (13d).

lie⁵⁵ indicates perfective aspect. It is often used for definite or specific events that have happened in the past. This particle indicates that the event is being viewed in its entirety, which is bounded temporally, spatially, or conceptually. It is also frequently used as a particle in narrative texts indicating the first event in a sequence of actions, the occurrence of which is bounded by subsequent actions. Furthermore, this particle is also used in procedural texts signaling the first step of the speaker's instruction to the hearer, which ought to be completed before the undertaking of the second stage (13e).

liau⁵⁵, a perfect aspect particle, indicates that an action or an event has already occurred. Compared with *lie⁵⁵*, it emphasizes the influence on the present that the result of action brings about (13f). Besides, *liau⁵⁵* can be used after a non-verb predicate, indicating that the given event or situation has appeared. When the clauses with non-verb predicates are negated, the verb must be supplied, which suggests that there should be an implicit/ opaque verb in an affirmative nominal predicate. Hence, the aspect marker *liau⁵⁵* at the end of the sentence still indicates the completion of an action or situation change (more in section 3.4.1); for example (13g).

Two of the four aspectual verb particles, *la⁵⁵* and *xu²¹* can appear simultaneously after a verb, signifying more minutely a state of composite meanings. For example, the particle combination *la²¹xu²¹*, comprising the progressive aspect particle *la²¹* and the inchoative aspect particle *xu²¹*, indicates that the action has begun or the situation has appeared. It also emphasizes that the happening or appearing of the given action or event precedes that indicated by *xu²¹* alone, e.g. (13h), (13i).

- | | |
|---|---|
| (13) a. <i>ŋa²⁴ tɕie²¹pi²¹ o⁵⁵ xu⁵⁵</i> | b. <i>o²¹thɑ²¹ mx²⁴tsy²¹ la²¹</i> |
| 1sg bowl wash INCH | outside rain PROG |
| 'I will wash the bowls soon.' | 'It is raining outside.' |
| c. <i>ŋa²⁴ tsau⁵⁵ku⁵⁵tie⁵⁵ tɕin⁵³ xu⁵⁵tsʰɑ²¹ la²¹</i> | d. <i>ko²⁴ nie²⁴ po⁵⁵ la⁵⁵!</i> |
| 1sg morning often run PROG | 3sg lie STA PROG |
| 'I often run in the morning.' | 'He is lying in bed!' |
| e. <i>ko²⁴ ŋa²⁴ la²⁴ ton⁵⁵ pa⁵³ lie⁵⁵.</i> | f. <i>ŋa²⁴ ie²¹ ka²⁴ liau⁵⁵.</i> |
| 3sg 1sg one CL look PFV | 1sg food eat PRF |
| 'He took a look at me.' | 'I have eaten.' |
| g. <i>po⁵³li²¹ on⁵⁵ sɿ⁵⁵ tʰon⁵⁵ liau⁵⁵.</i> | h. <i>an²⁴ni⁵⁵ zɿ⁵⁵ la²¹ xu²¹.</i> |
| child five year old PRF | 1pl do= PROG INCH |
| 'The child is five years old now.' | 'We are going to do it soon.' |
| i. <i>ko²⁴ nie²⁴ po⁵⁵ la⁵⁵ xu²¹.</i> | |
| 3sg lie STA PROG INCH | |
| 'He has been lying in bed.' | |

Tujia has five verb particles indicating directionality, which are postposed to the aspectual particles described earlier. The concept of direction may be geographical or relational. The particles *cie⁵⁵*, *ti⁵⁵*, and *tiu⁵⁵* indicate actions towards the speaker, while *lie⁵⁵* and *lu²¹* mark actions directed away from the speaker.

cie⁵⁵ is most frequently used in imperative clauses (14a), and it can also be used in declarative clauses describing past events to indicate directionality. In the latter, it indicates that the speaker has returned to his starting point after having performed the action stated (14b).

ti⁵⁵ is used in non-imperative clauses, indicating direction towards the speaker (14c).

tiu⁵⁵, a fusion of *ti⁵⁵* and the perfect marker *liau⁵⁵*, is used when perfect aspect has to be indicated as well as direction. For actions which have already happened, *ti⁵⁵* and *tiu⁵⁵* can both be used, but the latter conveys an additional sense of emphasis (14d).

lu²¹ is used in place of *lie⁵⁵* when perfect aspect has to be signaled. It is a fusion of *lie⁵⁵* and *liau⁵⁵*. They can both be used in all kinds of sentences to indicate direction away from the speaker, as in (14e), (14f).

- (14) a. tɕʰie²¹pi²¹ la²⁴ ta⁵⁵ xo²¹ εie²⁴.
 bowl one CL bring DIR
 ‘Bring a bowl here.’
- b. ɳa²⁴ lai⁵³ kai⁵⁵san⁵⁵ sɿ⁵⁵pa⁵⁵ pʰi⁵⁵ εie⁵⁵.
 1sg today market clothes buy DIR
 ‘Today I went to the market to buy clothes and came back (home).’
- c. ni²⁴ ie²¹ ka²⁴ ka²⁴ ti⁵⁵ tʰa⁵⁵?
 2sg food eat eat DIR NEG
 ‘Will you come to eat or not?’
- d. kr⁵³ o²¹ so⁵⁵ ta⁵⁵ xo²¹ tiu²⁴.
 3pl back basket three CL bring DIR:PRF
 ‘They have brought three back baskets.’
- e. ɳa²⁴ po⁵³li²¹ tsʰo⁵³ lie⁵⁵ xu²¹.
 1sg child take DIR INCH
 ‘I will take the child away.’
- f. ko²⁴ ko⁵⁵ po⁵³li²¹ tsʰo⁵³ lu²¹.
 3sg AGT child take DIR:PRF
 ‘He has taken the child away.’

The directional particles have the tendency to further grammaticalize, undergoing semantic changes from expressing concrete spatial meaning pertaining to actions to abstract temporal meaning. For example, *tsʰan²⁴ tiu⁵⁵* [sing + DIR:PRF] ‘to have begun to sing,’ and *l²¹ lu²¹* [say+ DIR:PRF] ‘to have continued to say.’

There are relevance particles in Tujia, which convey communicative functions. These particles are used by speakers pragmatically to convey an implicit request or event command, or the implicit consequences of an action or state; while on the other hand, they indicate to the hearers the somewhat relevance of the given state or action to the situation where the speech act is taking place. These particles always occur at the end of a string of particles following the verb. For example, *nie⁵⁵* indicates the continuing relevance of an action. In English, it can probably be best rendered by the adverb ‘still,’ as in (15).

- (15) ko²⁴ nie²⁴ po⁵⁵ la⁵⁵ nie⁵⁵.
 3sg sleep STA PROG still
 ‘He is still sleeping.’

ku⁵⁵ is used to indicate present relevance. It may be used on its own in imperative clauses conveying hortative as in (16a). This present relevance particle can best be represented in English by the adverb ‘now’ as in (16b).

- (16) a. an²⁴ ka²⁴ ku⁵⁵. b. an²⁴ nie²⁴ ku⁵⁵
 1pl eat CSM 1pl sleep CSM
 ‘Let’s eat.’ ‘Let’s sleep now!’

Apart from the aforementioned particles indicating aspect, directionality, and relevance, some other elements also follow verbs to supply complementary characterizations of the result, status, or frequency of the action. For example: *tci⁵³* ‘finish’ is used to focus on the completion of the action described by the verb, as in (17).

- (17) ko²⁴ ie²¹ ka²⁴ tci⁵³ xu⁵⁵.
 3sg food eat finish INCH
 ‘He will finish eating soon.’

s7²¹ is used to convey the intenseness of an action or state which has caused an extreme result, which means ‘extremely’ or ‘... to death.’ It is still not clear etymologically whether it is related to the Tuja verb *s7²⁴* ‘to die’ or the Chinese word *s7²¹*, which is identical to the Chinese verb ‘to die.’ In terms of use, the Tuja *s7²¹* is virtually equivalent to the Chinese *s7²¹*, it is thus more likely a Chinese loan. More has to be done in this respect, e.g.:

- (18) na²⁴ o⁵³ la²⁴ ts¹₁⁵⁵ i²¹ lie⁵⁵ ky⁵³ s¹₁⁵⁵ la²¹ xu²¹.
 1sg snake one CL see PRF fear die PROG INCH
 ‘I saw a snake and was frightened to death.’

to²¹, as a particle, indicates accomplishment, such as the reaching of a goal or the obtaining of a desired result. In such contexts, its use is comparable to that of the adverbial use of the word *tau²⁴* in Southwest Mandarin, which may be a derived use of the verb meaning ‘to arrive,’ e.g.:

- (19) an²⁴ ni⁵⁵ ko²⁴ tsa⁵³ to²¹ th^a₂₁ thⁱ₂₄.
 1pl him catch fulfill cannot
 ‘We cannot manage to catch him.’

In some cases, a resultative complement follows the main verb, marked by the adverbial particle *mo⁵⁵*, as in (20), which is identical in form with the adverbial particle as illustrated in (22b) in the following:

- (20) ai⁵⁵ xa⁵⁵ lie²¹ li²⁴ a⁵³ mo²¹ s¹₁²⁴ liau⁵⁵.
 that dog hungry ADV die PRF
 ‘That dog died of hunger.’

Some particles follow the verb to indicate frequency of actions, as in (21).

- (21) lon²¹ pai²¹ my²⁴ ts¹₁²¹ xen²¹.
 this year rain often
 ‘It always rains this year.’

Adverbial elements usually appear in preverbal position (22a). Preverbal adjectival adverbials that are not reduplicated usually take the adverbial particle *mo⁵⁵* (22b).

- (22) a. an²⁴ py²¹ tcein⁵⁵ lie⁵⁵ en²¹ ts¹₁²¹.
 1pl Beijing ABL come
 ‘We came from Beijing.’

- b. $s\bar{v}^{24}$ lai^{24} $xuan^{53}$ mo^{21} yi^{24} .
 2pl slow ADV go
 ‘You please go slowly.’

Tujia has two negation particles, ta^{55} and $t^h a^{55}$, which appear in post-verbal position. The former indicates objective negation and the latter expresses a subjective decision by the speaker not to perform a certain action. Both of them have corresponding forms indicating perfect aspect: tau^{55} and $t^h au^{55}$, which indicate that the actions will no longer be performed or the attributes will not exist. As mentioned in section 2.5.2, tau^{55} and $t^h au^{55}$ are fusions of negation particles and the reduced form of the perfect aspect particle $liau^{55}$. Following are the examples:

- (23) a. lai^{53} ki^{21} ta^{24} .
 today hot NEG
 ‘It is not hot today.’
- b. kai^{24} $s\bar{l}^{21}$ $teie^{21}$ ki^{21} tau^{24} .
 now hot NEG:PRF
 ‘It is not hot now.’ (Implying it was hot previously.)
- c. ηa^{24} zy^{24} xu^{21} $t^h a^{53}$.
 1sg wine drink NEG
 ‘I do not drink wine.’ (Implying I am not willing to drink wine.)
- d. ηa^{24} ai^{55} lie^{55} zy^{24} xu^{21} $t^h au^{53}$.
 1sg hereafter wine drink NEG:PRF
 ‘I will not drink wine from now on.’

Words can be formed with $t^h a^{53}$ and a following verbal or adjectival morpheme, for example, pie^{53} ‘favorably comparable to’— $t^h a^{55}$ pie^{53} ‘inferior to, unfavorably comparable to’; tie^{55} ‘up to (sb)’— $t^h a^{55}$ tie^{55} ‘beyond (sb)’; ei^{53} ‘sharp’— $t^h a^{24}$ ei^{53} ‘blunt’; xon^{24} ‘like’— $t^h a^{21}$ xon^{21} ‘do not want.’ It is likely that such words are the residues of the constructions of “ $t^h a^{53}$ + verbal or adjectival morphemes” in the history of Tujia, which have been lexicalized due to their frequent co-occurring and such vestigial constructions have been fossilized into words.

In terms of the marking of illocutionary force, the declarative and imperative are unmarked. The prohibitive uses the subjective negation particle $t^h a^{55}$, though in this construction it appears before the main verb,¹ e.g.:

- (24) $t^h a^{55}$ ka^{24} !
 NEG eat
 ‘Don’t eat!’

In Tujia sentences, verbs can take a single argument, two arguments, or three arguments. The single argument can be an agent or an experiencer as in (25a). Some verbs can take two arguments, an agent and a patient, as in (25b), and some verbs can take three arguments, an agent, a patient, and a recipient, in which the recipient is located before the patient and the verb, as in (25c).

- (25) a. nie^{24} pi^{55} za^{55} lu^{21} .
 bird fly DIR
 ‘The birds have flown.’
- b. ηa^{24} $s\bar{l}^{55}$ pa^{55} tsa^{24} la^{55} .
 1sg clothes wash PROG
 ‘I am washing clothes.’

- c. p^ha²¹p^hu⁵⁵ nja²⁴ po²¹ t^{ch}i²⁴pu⁵⁵ ka⁵⁵ s^l₁²¹ lie²⁴ lie⁵⁵.
 grandpa 1sg ALL soybean some CL give PRF
 'Grandpa gave me some liters of soybeans.'

The verbs that often take an agent and a patient can undergo valency decreasing by using the reciprocal particle *ta*⁵³, as in (26a). It is the same morpheme as the adverb ‘together.’ The homophonous reciprocal particle and adverb *ta*⁵³ can co-occur in the same clause, with the adverb preceding the reciprocal particle, as in (26b).

Reciprocals can also be formed on an intransitive verb as well, and in this case the co-occurrence of the reciprocal particle *ta*⁵³ and the adverb *ta*⁵³ are often compulsory, as in (26c).

- (26) a. an²⁴ p^hon²¹ni²¹ ta⁵³ i²¹ liau²¹.
 1pl yesterday RECP see PRF
 'We saw each other yesterday.'
- b. kx⁵³ ta⁵³ ta⁵³ xa²¹ la²¹.
 3pl together RECP hit PROG
 'They are hitting each other.'
- c. an²⁴ ta⁵³ ta⁵³ ko⁵³ liau⁵⁵.
 1pl together RECP quarrel PRF
 'We quarreled with each other.'

Tujia has a number of modal verbs, which constitute a special subtype of verb. The modal verbs always occur with notional verbs to express necessity, ability, intentions, desires, and truthfulness of the assertions. Most of them have corresponding negative forms, and corresponding perfect forms, as shown later. As mentioned in section 2.5.2, the perfect forms are fusions of modal verbs and the reduced form of the perfect aspect particle *liau*⁵⁵. The modal verbs can be further divided into the following three subgroups according to their semantic associations:

1. deontic modality, indicating necessity and obligation, e.g. *to*⁵⁵ 'must/ should,' and its corresponding negative form *t^ha⁵⁵to*²¹ 'need not';
2. epistemic modality, indicating the speaker's degree of certainty and the truthfulness of the assertion, e.g. *tr⁵⁵ei*²¹ 'can,' and its corresponding negative form *t^ha⁵⁵t^ht⁵⁵* 'not be able to'; *yr⁵⁵* 'be good at,' and the negative form *t^hγ⁵⁵yr⁵⁵* 'not be good at'; *x⁵⁵ t^hy⁵⁵* (negative) 'be not strong enough to'; as well as *pie*⁵⁵ 'have time to';
3. desiderative modality, indicating desire or willingness, e.g. *t^ha⁵⁵* 'want to'; *t^hy⁵⁵* 'be willing to,' and the corresponding negative form *t^ha⁵⁵t^hy²¹* 'be not willing to.'

3.4 Special constructions

3.4.1 Copular constructions

In Tujia, affirmative copular constructions don't require a copula. They are expressed by means of clauses with non-verb predicates, that is, nouns, pronouns, or noun phrases are juxtaposed to indicate a judgment. While in negative copular constructions, a negative particle must be added at the end of sentences.

Copular constructions consist of two subsets. One subset indicates a judgment on facts ((27a), (27b)); constructions of this type are negated by adding the Chinese loan *pu*²⁴*s^l*⁵⁵ at the end of sentences, as in (27c), (27d).

- (27) a. $\eta a^{24} pi^{24} tsj^{55} k^h a^{21}$.
 1sg Tujia
 'I am Tujia.'
- b. $an^{24} k^h a^{21} pu^{21} p^h o^{24} ka^{55}$.
 1pl grandmother teacher
 'My grandmother is a teacher.'
- c. $lau^{24} tsj^{55} ts^h l^{53} k^h a^{55} pu^{24} sij^{55}$.
 tomorrow New Year be not
 'Tomorrow isn't New Year.'
- d. $po^{53} li^{21} on^{55} sy^{55} t^h on^{55} ie^5 ta^{21} nie^{55}$.
 child five year old reach not still
 'The child is less than five years old.'

The other subset indicates a judgment on positions or locations. There are no verbs expressing 'be located in' or 'lie in' in Tujia. The position or location of someone or something is conveyed by a place noun phrase or a locative noun phrase acting as a predicate to indicate where the involved entity is ((28a), (28b)). The negative locative copular constructions are formed by adding the verbal construction *kau⁵³ta²¹* at the end of sentences, as in (28c), (28d) *kau⁵³ta²¹* consists of the Chinese loan *kau⁵³* and the negation particle *ta²¹* of Tujia.

- (28) a. $an^{24} ts^h u^{55} t^h a^{55} sau^{55}$.
 1pl house (a place name)
 'Our home is in Tasha.'
- b. $mo^{21} sij^{21} t^h ie^{24} tsj^{55} ky^{55}$.
 cat desk front
 'The cat is in front of the desk.'
- c. $t^h u^{21} su^{55} kuan^{53} xu^{21} p^h a^{21} t^h a^{21} kau^{53} ta^{21}$.
 library lake LOC be not
 'The library is not beside the lake.'
- d. $sy^{24} a^{21} nie^{53} an^{24} ts^h u^{55} kau^{53} ta^{21}$.
 2pl mother 1pl house be not
 'Your mother is not in our house.'

3.4.2 Existential construction

There is only one verb of possession / existence, *cie²⁴* 'have,' the negative of which is *t^hai²⁴* 'do not have.' The two verbs are illustrated as in (29a), (29b). Both of them have perfect forms, which are *ciau²⁴* 'already have' and *t^hau²⁴* 'do not have any more' respectively. As indicated in section 2.5.2, these two perfect forms are fusions of the two verbs with the reduced form of the perfect particle *liau⁵⁵*. *cie²⁴* indicates the possession by a certain possessor, or existence of a referent at a certain place. It may also be used to signal a noun phrase which introduces an entity into a discourse, as in (29c).

- (29) a. $\eta a^{24} [a^{24} ko^{55} on^{53} yi^{21}] cie^{24}$.
 1sg friend five CL have
 'I have five friends.'

- b. kau²⁴ ts^hy⁵⁵ts¹⁵⁵ t^hau²⁴.
here bus DON'T.HAVE
'There is no bus any more here.'
- c. uan⁵³thian²¹, zen⁵⁵ka⁵⁵tie⁵⁵ nie⁵⁵ xu⁵⁵ eie²⁴.
long ago brother and sister two CL have
'Once upon a time, there was a brother and a sister.'

3.4.3 *The comparative construction*

Generally, the comparative construction in Tujia has the word order of Standard–Comparee–Parameter–Marker, wherein the standard is always the topic of the comparative construction. The marker *nie*⁵⁵, identical in form to the particle *nie*⁵⁵ ‘still,’ is located at the end of the comparative sentence (30a). The verb or adjective functioning as the parameter can be in the plain form or take a classifier phrase (30b).

- (30) a. ηa²⁴ an²⁴ a²⁴ta⁵⁵ s¹²⁴ nie⁵⁵.
1sg 1pl sister fat COMPAR
'My elder sister is fatter than me.'
- b. ko²⁴ ni²⁴ la²⁴pie⁵⁵ ts^h1⁵³ nie²¹.
3sg 2sg much old COMPAR
'You are much older than him.'

3.5 Clause

3.5.1 *Clausal word order*

The word order in the Tujia clauses is verb final, while the NPs are ordered with the elements of higher degree of topicality located nearer to the beginning of the clause (31a). Topics are often followed by a pause in speech, which sets the topical elements apart from the rest of the utterance. The pause may just be silence, or may also be signaled by specific particles, which are *m*⁵⁵ and *lie*⁵⁵ in Tujia ((31b), (31c)). The immediate preverbal position is the unmarked focus position (the unmarked position for introducing ‘new’ referents / information).

- (31) a. ie²¹, ηa²⁴ t^ha²¹ xon²¹ liau⁵⁵.
food 1sg not need PRF
'As for food, I do not need [it] anymore.'
- b. ηa²⁴ m⁵⁵, ts^h1⁵⁵t^hu⁵⁵ po⁵⁵ ta⁵⁵.
1sg TOP study STA NEG
'I never went to school.'
- c. kai²⁴ po⁵³li²¹ lie²¹, nie²¹pa²¹ nie²¹ sa²¹ zu²¹ t^ha⁵⁵.
this child TOP parents ASSOC word listen NEG
'The child doesn't follow his parents' advice.'

3.5.2 *Polar and Wh- interrogatives*

Polar interrogatives are generally formed using the post-verbal question particle *man*⁵⁵ or *ma*⁵⁵, the latter being a Chinese loan (32a). Another type of polar interrogative is

formed by juxtaposing positive and negative alternatives (i.e. an A-not-A question), as in (32b). Wh questions have interrogative pronouns *in situ*, and do not require a final particle (32c).

- (32) a. kai²⁴ ni²⁴ nie⁵⁵ s⁵⁵pa⁵⁵ ma⁵⁵/ man⁵⁵?
this 2sg ASSOC coat Q
'Is this your coat?'
b. sy²⁴ ie²¹ ka²⁴ ka²⁴ t^ha⁵⁵?
2pl food eat eat NEG
'Will you eat or not?'
c. ai⁵⁵ti⁵⁵ te^hie⁵³cie²¹?
that what
'What is that?'

3.5.3 Subordinate clauses

The semantic relations between two or more clauses can be divided into coordinate and subordinate relationships. Coordinate constructions do not involve a conjunction; the relationship between the clauses is expressed through juxtaposition. Subordinating constructions often need a conjunction or subordinator. The subordinators include clause-final *mr*⁵⁵ 'if,' and clause-initial *ai*⁵⁵*kan*²¹*mo*²¹ 'so.'

- (33) a. lau²⁴ts⁵⁵ mr²⁴ts²¹ mr²¹, nja²⁴ ts^hu⁵⁵ on²¹ po-i²⁴.
tomorrow rain if 1sg home sit STA-PROS
'If it rains tomorrow, I will stay at home.'
b. su²⁴su⁵⁵ts²⁴ la⁵⁵, ai⁵⁵kan²¹mo²¹ sa⁵³.
snow PROG so cold
'It is snowing, so it is cold.'

ABBREVIATIONS

ASSOC	associative particle
CSM	present relevance particle (change of state)
INCH	inchoative aspect particle
PROS	prospective aspect particle/marker
STA	state particle

NOTE

- 1 Editor's note: This may be the older pattern for Tuja, as it is the common pattern in Tibeto-Burman, and seems to involve the Proto-Tibeto-Burman prohibitive marker *ta.

REFERENCES

- He, Tianzhen. 2003. Tujiayu de zhishu wenti (The question of the affiliation of Tuja). *Zhongnan minzu daxue xuebao (Journal of South-Central University for Nationalities)* 23.1: 88–97.
Tian, Desheng, He, Tianzhen, Chen, Kang, Li, Jingzhong, Xie, Zhimin, and Peng, Xiumo. 1986. *Tujiayu Jianzhi (A Brief Sketch of Tuja)*. Beijing: Minzu Publishing House.

- Wang, Jingru. 1955. Guanyu Xiangxi Tujiayu de chubu yijian (A preliminary opinion about the Tujiya of Xiangxi). Research Department of the Central College for Nationalities. *Zhongguo Minzu Wenti Yanjiu Jikan (Collected Papers on Ethnic Issues in China)* 4: 135–174.
- Xu, Shixuan. 2011. Tuji zuqun he yuyan de liubian (The Tuji ethnic group and the development of the Tuji language). Paper presented at the Second International Meeting of Linguistic Evolution and Genetic Evolution, Fudan University, Shanghai, September 2011.
- Xu, Shixuan, and Lu, Meiyuan. 2014. Tujiayu dongci tibiaoji -i de laiyuan he yufahua guocheng (On the origin and grammaticalization of the aspectual verbal marker -i in Tuji]. *Minzu Yuwen (Minority Languages of China)* 6: 18–26.

CHAPTER FIFTY-THREE

JINGHPAW¹

Keita Kurabe

1 INTRODUCTION

Jinghpaw is a Tibeto-Burman (TB) language spoken in a broad region stretching from the upper Brahmaputra valley of Northeast India across northern Burma, and beyond the Burma–China border into far western Yunnan. The population of the Jinghpaw people is estimated to be approximately 630,000 in Burma (Bradley 1996), 37,000 in China (Dai 2012), and 5,000 to 6,000 in India (Morey 2010). Jinghpaw is closely related to the Luish [Asakian] languages, such as Sak, Cak, Kadu, Ganan, Andro, Sengmai, and Chairel (Matisoff 2013). Together these constitute the Jinghpaw-Luish branch of TB, which covers a widespread but discontinuous area in the northwestern part of Greater Mainland Southeast Asia. A special relationship with Bodo-Koch [Bodo-Garo] and Konyak [Northern Naga] languages has also been suggested for Jinghpaw, especially on the basis of distinctive roots, such as ‘sun’ and ‘fire’ (Burling 1983). Benedict (1972: 5–6), recognizing the geographic centrality and linguistic diversity of Jinghpaw within TB, provides a diagram of the Sino-Tibetan family where all TB major subgroups (except Karenic) are associated with Jinghpaw in the center, remarking that Jinghpaw stands at the linguistic “crossroads” of TB.

The Jinghpaw-speaking area is a site of intensive contact in which language contact among intra- and extra-TB languages has been a long-standing phenomenon. Of particular importance is the fact that Jinghpaw serves as a lingua franca among the Kachin people, who form a socio-cultural complex of shared cultural traits including a marriage-alliance system. This Kachin grouping consists of speakers of languages belonging to several TB branches, such as Jinghpaw, Zaiwa [Atsi] (Burmish), Lhaovo [Maru] (Burmish), Lacid [Lachi] (Burmish), Ngochang (Burmish) and Rawang (Nungish), and includes some Lisu (Loloish) speakers in Burma as well. The Kachin, as such, provide a case of one-to-many correspondence between ethnicity and languages. Aside from common inheritance, members of the Kachin grouping share linguistic as well as cultural traits that have arisen as a result of intensive contact. The direction of contact-induced language change is multilateral, directed to Jinghpaw itself as well as to non-Jinghpaw Kachin languages, leading to more or less convergence among some Kachin languages. Jinghpaw influence on other Kachin languages varies from language to language. Zaiwa is most influenced by Jinghpaw, as reflected in its phonological similarity to Jinghpaw as well as in its abundant Jinghpaw loanwords, including kinship terms. The Kachin people, including the Jinghpaw, have also had a long-term symbiotic relationship with Tai-speaking Shan peoples, as reflected, in part, in a number of Shan loanwords in their languages. Leach (1954) proposes that Kachin communities in the first half of the twentieth century were “oscillating” between an egalitarian system and a Shan feudal system.

Jinghpaw consists of a number of “dialects,” not all of which are mutually intelligible. The Jinghpaw dialects spoken in Northeast India are collectively referred to by the name

“Singpho” (Morey 2010). On the grounds of shared innovations in phonology and lexicon, the Jinghpaw dialects may be divided into two major groups: Southern and Northern (Kurabe 2014). The former group, which includes Standard Jinghpaw, Nhkum, Gauri, Htingnai, Hkahku, and Shadan, is distributed in the southern part of the Jinghpaw-speaking region of southern and central Kachin State, northern Shan State, and western Yunnan. The latter group, including Duleng, Dingga, Shang, Numhpuk, and Turung, is distributed in the northern part of the Jinghpaw-speaking region, covering northern Kachin State and Northeast India. The Jinghpaw data presented in this chapter is from Standard Jinghpaw spoken in and around Myitkyina and Bhamo of Burma, unless otherwise noted.²

2 PHONOLOGY

2.1 Phoneme inventory

Jinghpaw consonants and vowels are given in Table 53.1.

Vowel length is not phonemic in Jinghpaw. The phoneme /h/ is marginal, occurring mostly in loanwords or onomatopoeia. The fact that many of the /h/ of Shan or Burmese loanwords are adapted as /kh/ in Jinghpaw indicates its marginality, e.g. *khoj* ‘shellfish’ (< Shan *hɔj'*). The phonemes /s/ and /c/ may be interpreted phonologically as aspirated counterparts of /ts/ and /tʃ/ based on the phonological gaps in the inventory as well as on the fact that they, together with aspirates, trigger the alternation of the causative prefix, i.e. *ɛɔ-* → *jɛ-* / _ [+aspirate], as in *ɛɔ-cay* [CAUS-black] ‘blacken,’ *ɛɔ-màt* [CAUS-lost] ‘lose,’ *ɛɔ-ráy* [CAUS-COP] ‘adjust,’ *jɛ-phrò* [CAUS-white] ‘whiten,’ *jɛ-sán* [CAUS-clean] ‘clear,’ and *jɛ-ɛày* [CAUS-enter] ‘join,’ which can be seen as a kind of dissimilation process. The phonemic status of the preglottalized sonorant series is established by such minimal or near minimal pairs as: *maj* ‘corpse’ vs *?maj* ‘dark’; *nay* ‘dull’ vs *?nay* ‘sticky’; *yay* ‘1SG’ vs *?yay* ‘bear’; *ràt* ‘scratch’ vs *?rát* ‘elder sister of wife’; *lày* ‘swinging’ vs *?lày* ‘once’; *wàn* ‘coiled’ vs *?wàn* ‘fire’; *yàt* ‘hang down’ vs *?yàt* ‘slow.’ Vowels following preglottalized sonorants tend to be creaky. These vowels, together with vowels preceded by voiceless consonants, are treated as creaky vowels in Jinghpaw spoken in China (Dai and Xu 1992, among others).

Jinghpaw is a tone language with four contrastive tones in unchecked syllables and two in checked syllables, as evidenced by the minimal or near minimal pairs in Table 53.2.

The falling tone is infrequent, occurring in such specific words as kinship terms, sentence-final particles and interjections. Some of these are derived from an underlying

TABLE 53.1 JINGHPAW PHONEME INVENTORY

Consonants						Vowels			
p	t	ts	c	k	?	i			u
ph	th			kh		e	ə		o
b	d	dz	j	g			a		
		s	ɛ		(h)				
m	n				ŋ				
?m	?n				?ŋ				
w	l	r	y						
?w	?l	?r	?y						

TABLE 53.2 TONES IN JINGHPAW

Tone	Pitch	Examples		
High	55	<i>lá</i>	‘take’	<i>khám</i>
Mid	33	<i>la</i>	‘male’	<i>kham</i>
Low	31	<i>là</i>	‘wait’	<i>khàm</i>
Falling	51	<i>?wā</i>	‘father’	<i>tím</i>
				‘but’

low by means of tone spread, where a high spreads to a low from the left to right, e.g. *ń-lù* [NEG-get] → *ń-lū*. Note additionally that the negative prefix changes the following low tone in a checked syllable to a high, e.g. *ń-sát* [NEG-kill] — *ń-sát*, and that the underlying high on the negative prefix assimilates with a following mid into a mid, e.g. *ń-tsun* [NEG-say] → *n-tsun*. Jinghpaw also has some irregular tonal changes triggered by prefixation, the nature of which is not yet thoroughly understood.

2.2 Word and syllable structure

Monomorphemic words are almost always monosyllabic or disyllabic, with verbs tending to be monosyllabic and nouns tending to be disyllabic in basic vocabulary. Monosyllables are always heavy, being headed either by phonetically long vowels or phonetically short vowels followed by coda consonants. A large number of disyllables take the form of the iambic or ‘sesquisyllabic’ structure (Matisoff 1973), consisting of a major (heavy) syllable preceded by a minor (light) syllable with reduced phonemic possibilities. The examples given in Table 53.3 illustrate the Jinghpaw word structures.

The major syllable allows up to two prenuclear consonants and one postnuclear consonant, so that the maximal structure is $c_1c_2vc_3$, plus a tone. c_1 may be any consonant in the inventory when c_2 is not filled. Sonority must increase in the onset. Two sonorants /r/ and /y/ may occur as c_2 when the preceding consonant is a stop (bilabial, velar) or nasal (bilabial, alveolar), except for the combination of a nasal plus /r/. The optional coda consonant (c_3) can be /p, t, k, ?, m, n, ń, w, y/, where /k/ is mostly restricted to loanwords, as the Proto-Tibeto-Burman (PTB) coda *k has developed into /?/ in modern Jinghpaw (Benedict 1972: 14). The four diphthongs, *ui*, *oi*, *ai*, *au*, can be interpreted phonologically as a sequence of a vowel plus a glide, i.e. /uy/, /oy/, /ay/, /aw/, based on the fact that they never occur in closed syllables. The vowel-coda combination (vc_3) is basically free, except when v is a schwa or when c_3 is a glide.

The minor syllable is headed either by a phonetically short schwa /ə/ or a syllabic nasal, represented by /n/ in this chapter, which assimilates to place of articulation of the following consonants. The fact that the minor syllable never occurs word-finally, together with the fact that the first syllable of a disyllabic word is sometimes reduced to a minor syllable, e.g. *ginsúp* ‘play’ → *gəsíup*, as well as the fact that a monosyllabic prefix, when prefixed to a monosyllabic base, is sometimes reduced to a minor syllable retaining its original tone, e.g. *?á-lóy* [ADV-easy] → *?ó-lóy* ‘easily,’ indicates a consistent iambic pattern for Jinghpaw prosody. The schwa-headed minor syllable, which is always light and open and does not allow complex onsets, has a reduced set of onset consonants, the most frequent of which include /g, ?, ɛ, m, l/. The consonants /dz, n, ń, r/ never occur in this position. The syllabicity of the syllabic nasal can be evidenced by observing: (a) that it bears a tone; (b) that it cannot be prefixed by monosyllable-targeting prefixes; (c) that it is not copied in partial reduplication; (d) that it is assigned one musical beat, just like a

TABLE 53.3 JINGHPAW WORD STRUCTURES

Monosyllabic words		Sesquisyllabic words		Fully disyllabic words	
?ù	‘bird’	gəwá	‘bite’	gìnrù	‘ancestry’
jan	‘sun’	lətá?	‘hand’	gìnsúp	‘play’
khri	‘sour’	cəpre	‘bean’	gùmphró	‘silver’
khruj	‘live’	məkhrèt	‘scratch’	gùmpyét	‘flatten’

full syllable; and (e) that some have fully syllabic variants, e.g. *n̥may* ~ *n̥ŋmay* ‘tail,’ *n̥pòt* ~ *n̥ŋpòt* ~ *n̥umpòt* ‘beginning.’

Aside from sesquisyllabic words, Jinghpaw has fully disyllabic words consisting of two major syllables, the phonological shape of which is also restricted; they only allow simple onsets and their rhymes are largely restricted to high vowels followed by nasal codas. Some of these rhymes are found in complementary distribution, namely /i/ basically occurs before /n/, and /u/ before /m/. There are some words which display allomorphy between /in/ and /um/, e.g. *ginthuŋ* ~ *gùmthuŋ* ‘bridle.’ Note also that a rhyme /in/ in this position tends to be preceded by coronal consonants, e.g. *ciŋná* ‘cane.’

3 LEXICON

The Jinghpaw lexicon exhibits various kinds of Southeast Asian areal traits. As in other languages of East and Southeast Asia, Jinghpaw has a rich lexicon of elaborate expressions and psycho-collocations (Matisoff 1986), e.g. *gà-tsò-gà-nèm* [word-high-word-low] ‘tones,’ *myít dij* [mind-straight] ‘honest,’ *məsin gədùn* [liver-short] ‘be irritable’ (for many more examples, see Hanson 1906; Yue *et al.* 1981; Xu *et al.* 1983; Maran 1979). Jinghpaw also has multiple words for particular actions, such as ‘washing,’ ‘carrying,’ and ‘cutting,’ as is the case with other Southeast Asian languages, e.g. *phay* ‘carry on the shoulder,’ *lay* ‘carry in the hand,’ *bà* ‘carry a child tied on the back,’ and *gun* ‘carry something suspended by a strap from the head.’

In addition to a large number of words inherited from PTB, as reflected in its basic vocabulary, Jinghpaw has absorbed many words from Shan, with which Jinghpaw has been in intensive contact (Kurabe 2016). A large number of Jinghpaw words relating to wet-rice cultivation, trading, and feudalism are of Shan origin, reflecting the economic and political situation in northern Burma. These include *kháw* ‘paddy’ (< Shan *khaw*³), *gát* ‘market’ (< *kaat*²), *múj* ‘country’ (< *mryŋ*⁴), *khokhám* ‘king’ (< *hɔ¹kʰam*⁴ ‘royal palace’), and two person names *dzàw* and *nàj* (< *tsaw*³ ‘lord,’ *naay*⁴ ‘princess’). Jinghpaw also adopted words from Burmese and Chinese, but in a more limited fashion, e.g. *lákmat* ‘certificate’ (< Written Burmese *lakmhat*), *làwbàn* ‘boss’ (< Chinese *lǎobǎn*). As a lingua franca in the Kachin socio-cultural complex, Jinghpaw has provided many words to non-Jinghpaw Kachin languages, which form the areal lexicon in the Kachin cultural area.

4 MORPHOLOGY

Productive affixes are not abundant in modern Jinghpaw, although there are various kinds of fossilized affixes, many of which are retentions from PTB (Wolfenden 1929: 70–86; Benedict 1972: 96–121), including the nominalizing suffix *-t*, e.g. *cà-t* [eat-NMLZ] ‘rice, food.’ The vast majority of affixes are prefixes, many of which take the phonological shape

of a minor syllable. Some derivational affixes impose constraints on the phonological make-up of the base they are added to. The kinship prefix *ŋá-* and causative prefix *ɛə-*, for example, can only be added to monosyllabic bases. There are some special kinds of affixes, including a meaningless prefix *ŋə-* which is employed only to add phonological bulk to monosyllabic bases, e.g. *sày ~ ŋə-sày* ‘blood,’ and a nominalizing prefix *má-* ‘everything that,’ originating from a full verb *má?* ‘be exhausted’ (Hanson 1896: 32), which co-occurs with reduplication, e.g. *ŋá-má-ŋá* [live-NMLZ-live] ‘everything that exists.’ Observe in the following examples that Jinghpaw has both class-maintaining (a, b, c) and class-changing affixes (d, e, f), and that only nouns and verbs are involved in affixation.

- (1) a. *cə-nà* [CAUS-hear] ‘inform’ b. *ń-lá* [NEG-take] ‘not take’
 c. *ɬə-nū* [KIN-mother] ‘my mother’ d. *ɬó-tsóm* [ADV-beautiful] ‘well’
 e. *ɬə-éá* [HAB-eat] ‘always eat’ f. *cə-si* [NMLZ-die] ‘dead person’

Compounding is a productive morphological process of word formation. Jinghpaw exhibits both endocentric (a, b, c, d, below) and exocentric compounds (e), the former of which may be both right-headed (a, b) and left-headed (c, d). Only nouns and verbs are productively involved in compounding. All logically possible combinations of a noun and a verb are attested in Jinghpaw: noun-noun (a, g, h); verb-noun (b); noun-verb (c, d, e); verb-verb (f). Jinghpaw has a rich lexicon of appositional compounds (f, g, h). They have a fixed order, and the order of the members is largely predictable. Two major rules determine the ordering: higher first rule (Dai and Xu 1992: 400–5)—if the syllable length of the two members is equal, then the member including the higher vowel comes first (f, g); shorter first rule—if the syllable length of the two members is not equal, then the shorter member comes first (h).

Recall that the first syllable of a disyllabic word is sometimes reduced to a minor syllable due to the predominance of the iambic pattern of Jinghpaw prosody. When applied to a disyllabic compound, this process of ‘sesquisyllabization,’ as discussed by Dai and Wu (1995), obscures the etymology of the compound. To illustrate this, consider the compound *süt-dék* [wealth-depository] ‘box’ which is often reduced to *sədék*. As pointed out by Matisoff (2003: 130), Jinghpaw has a number of arm- or leg-related nouns and verbs with a minor syllable *lə-*, which has its diachronic source in the PTB etymon **lak* ‘hand, arm,’ e.g. *ləkhrá* ‘right-hand’ (cf. *khrá* ‘right’), *ləkhát* ‘kick’ (cf. *khát* ‘be kicking’). Note additionally that the minor syllable of a base is sometimes deleted in compounding, e.g. *nàw-čərà* [dance-place] → *nàw-rà* ‘dancing floor,’ *čəday-dò?* [navel-cut] → *dày-dò?* ‘birth place.’ This process, together with sesquisyllabization, further obscures the etymology of given compounds. Consider the cases of *bəren* ‘dragon’ which has its diachronic source in a compound *ləpu-ren* [snake-long] and *sərín* ‘ginned cotton’ which originates from *pəsi-rín* [cotton-roll].

Jinghpaw reduplication is manifested mainly as partial reduplication, copying the last syllable of the base. Jinghpaw employs reduplication in order to mark habituality

(a, below), distributivity (b), indefiniteness (c), and plurality of demonstratives (d). Reduplication can be employed to form adverbs or adverbial clauses as well (e, f). Related to this is the reduplication-like noun-verb construction which may involve loanwords, such as *dəmŷà? myà?* ‘rob’ (< Burmese *dāmŷā* ‘robber’) and *làwbàn bàn* ‘become a boss’ (< Chinese *lǎobān* ‘boss’). Although it superficially resembles morphological reduplication, this construction is beyond the scope of morphology, as can be seen in the fact that fully syntactic elements, such as adverbs, can be interposed.

- | | | | | | |
|-----|------------------------------|---------------|--------------|------------------------------|--------------------|
| (3) | a. <i>məci?</i> ~ <i>ci?</i> | [sick-REDUP] | ‘often sick’ | b. <i>məsum~sum</i> | [three-REDUP] |
| | | | | | ‘three each’ |
| | c. <i>ləŋây~ŋây</i> | [one-REDUP] | ‘some’ | d. <i>gəday~day</i> | [who-REDUP] |
| | | | | | ‘who (pl.)’ |
| | e. <i>jəkhriŋ~khrin̩</i> | [delay-REDUP] | ‘often’ | f. <i>thù?</i> ~ <i>thù?</i> | [rain-REDUP] |
| | | | | | ‘whether it rains’ |

5 SYNTAX

5.1 Word classes

This chapter divides Jinghpaw words into the following five word classes: nouns, verbs, adverbs, particles, and interjections. Nouns are defined by their ability to be the heads of NPs. Verbs are words which share a bundle of properties, such as the ability to be negated by the negative prefix and the ability to inflect for person, number, aspect, and mood. As such, verbs denoting property concepts, including adjectives, form a subclass of verbs. Particles are words which cannot form utterances on their own, and interjections are words which do not hold any syntactic relationship with other words. Adverbs do not have any of the properties given above, typically occurring in immediate pre-verbal position. Many Jinghpaw adverbs are of verbal origin, both synchronically and diachronically, being or having been derived from lexical verbs by means of affixation, reduplication, and adverbialization of serial verbs.

5.2 The noun phrase

Within an NP, genitive and relative modifiers occur in the pre-head position, while modifiers such as an adjective, plural marker *ni*, numeral and classifier occur in the post-head position, as in (4). A demonstrative may occur in either pre- or post-head position, e.g. *nday mənaŋ* [this-friend] ~ *mənaŋ nday*. Note that, although verbs denoting property concepts may occur after nouns, e.g. *pù-khá* [intestine-bitter] ‘gallbladder,’ many of them cannot occur in this position freely, e.g. **nàmsì-khá* [fruit-bitter]. This fact suggests that these examples are compounds, not syntactic phrases. Most verbs, including verbs expressing property concepts, syntactically modify nouns only by means of relativization (nominalization). There is, nevertheless, a small set of verbs which freely modify nouns in post-nominal position. These verbs express the four core semantic types of adjectives (Dixon 1977), i.e. DIMENSION, AGE, VALUE, and COLOR, such as *gəbà* ‘big,’ *diysà* ‘old,’ *gaja* ‘good,’ and *phrò* ‘white.’ This chapter treats only these words as ‘adjectives,’ which are a subclass of verbs.

- | | | | |
|-----|-----------------------------|------------------|---|
| (4) | a. <i>nyé?</i> <i>mənaŋ</i> | <i>my friend</i> | b. <i>ɛ́at</i> <i>ɛ́á</i> <i>ɻay</i> <i>mənaŋ</i> |
| | | | <i>food</i> <i>eat</i> <i>NMLZ</i> <i>friend</i> |
| | | | ‘friend who ate food’ |

c. <i>mənay gəja</i> friend good 'good friend'	d. <i>gəja ʔay mənay</i> good NMLZ friend 'good friend'
e. <i>mənay gəja ni məray məsum</i> friend good pl CLF:HUMAN three 'three good friends'	

In contrast to common nouns, which take a full range of adnominals, closed classes of nouns cannot take the same range of adnominals. Demonstratives and personal pronouns, for example, only allow post-adnominals. Table 53.4 summarizes the properties of noun subclasses in terms of whether a given noun: (a) can be the head of an NP; (b) can occur on its own; (c) can be pluralized by the plural marker; (d) can be quantified by a numeral; (e) can be determinated by a pre-head demonstrative; (f) can be preceded by a relative clause; (g) can be preceded by a genitive phrase; (h) can be determinated by a post-head demonstrative; (i) can be modified by an adjective; or (j) can be involved in partial reduplication.

5.2.1 Demonstratives

Jinghpaw demonstratives are distinguished in terms of speaker/addressee orientation, relative distance, and height from the speaker: *nday* [proximal; speaker-centered]; *day* [proximal; addressee-centered]; *wó-rà* [distal; level]; *thó-rà* [distal; up]; *lé-rà* [distal; down], the latter three of which involve *ɛərà* 'place.' Demonstratives may occur in pre- or post-head positions, except when they modify locator nouns or postpositions, in which case only pre-head position is grammatical. Only proximal demonstratives may be used for anaphora or cataphora.

5.2.2 Personal pronouns

The personal pronoun system exhibits three-way splits in person (1st, 2nd, 3rd) and number (singular, dual, plural), yielding the paradigm given in Table 53.5. There is no gender or inclusive/exclusive distinction. Separate genitive forms exist for singular, which are likely to have their diachronic sources in the contraction of singular personal pronouns and the genitive case marker *ʔà?*. Duals appear to involve the obsolete numeral *ni* 'two' (Wolfenden 1929: 73). Plural pronouns are segmentable into duals plus an element *the* 'all, whole.'

Related to this is the three-way distinction of possessive bound pronouns which are prefixed only to inalienable kinship terms, illustrated below. Although limited in scope, these pronominal possessive prefixes imposed on the possesum can be seen as instances of head-marking. Note that the 2nd person prefix *ni-* apparently came from *nay* 'you' or

TABLE 53.4 PROPERTIES OF NOUN SUBCLASSES

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
common nouns	yes	yes	yes	yes	yes	yes	yes	yes	yes	some
numerals	yes	yes	yes	no	yes	yes	yes	yes	no	yes
demonstratives	yes	yes	yes	yes	no	no	no	no	no	yes
personal pronouns	yes	yes	some	yes	no	no	no	no	no	no
interrogative pronouns	yes	yes	yes	yes	no	no	no	no	no	yes
locator nouns	yes	yes	no	no	yes	yes	yes	no	no	no
postpositions	yes	no	no	no	yes	yes	yes	no	no	some

TABLE 53.5 PERSONAL PRONOUNS

	Singular (NOM)	Singular (GEN)	Dual	Plural
1st	<i>yay</i>	<i>nyé?</i>	<i>?án</i>	<i>?ánthe</i>
2nd	<i>nay</i>	<i>ná?</i>	<i>nán</i>	<i>nánthe</i>
3rd	<i>ei</i>	<i>ei?</i>	<i>eán</i>	<i>eánthe</i>

ná? ‘your.’ Less apparent, the 3rd person prefix *gə-* seems to have a historical connection to the 3rd person pronoun which is pronounced *khyi* in some Jinghpaw varieties (Dai and Xu 1992: 23), suggesting a sound change **khyi* > *ei* in Standard Jinghpaw.

- (5) a. *nyé? ?ɔ-phû* ‘my brother’ b. *?ánthe ?à? ?ɔ-phû* ‘our brother’
 c. *ná? n̥-phû* ‘your (sg) brother’ d. *nánthe ?à? n̥-phû* ‘your (pl) brother’
 e. *ei? gə-phû* ‘his/her brother’ f. *eánthe ?à? gə-phû* ‘their brother’

5.2.3 Numerals and classifiers

Jinghpaw has a decimal numeral system, as is the case with many other TB languages. Jinghpaw numerals from 3 to 100, including a distinct root *khun* ‘20’ (cf. Written Burmese *akun* ‘all’), are inherited from PTB. The usual numerals *ləŋây* ‘one’ and *ləkhoŋ* ‘two’ are Jinghpaw innovations, although more general TB roots for ‘one’ and ‘two’ also survive in certain environments, e.g. *məray mi* [CLF:HUMAN-one] ‘one person’ and *ni-ná?* [two-night] ‘two nights’ (Matisoff 1994). An interrogative *gədè* ‘how many/much’ can be shown to be a numeral on the grounds that it shows the same distribution as other numerals. It is of interest to note that round numbers such as *khyij* ‘1,000,’ *mùn* ‘10,000,’ *sèn* ‘100,000,’ *wàn* ‘1,000,000,’ and *ri* ‘10,000,000,’ all of which are of Shan or Chinese origin, are distinguished from inherited round numbers *ei* ‘ten’ and *tsa* ‘100’ in terms of the relative position in which they occur. Compare: *məli-ei* ‘40,’ *məli-tsa* ‘400,’ *khyij məli* ‘4,000,’ *mùn məli* ‘40,000,’ *sèn məli* ‘400,000,’ etc. These borrowed round numbers can be interpreted as classifiers based on their position. This is supported, in part, by the fact that Jinghpaw has borrowed many classifiers from neighboring languages (Xu 1987).

Jinghpaw is not a classifier-rich language. Numerals can occur in the absence of classifiers, and nouns can be directly quantified by numerals. There is, however, a small set of classifiers that can precede numerals. The classifier numeral phrase can occur without a head noun. Consider:

- (6) a. *mà məsum* b. *mà məray məsum*
 child three child CLF:HUMAN three
 ‘three children’ ‘three children’
- c. *məray məsum* d. *nday məray məsum*
 CLF:HUMAN three this CLF:HUMAN three
 ‘three persons’ ‘these three persons’

5.2.4 Interrogative pronouns

Most Jinghpaw interrogatives are analyzable into an element *gə-* and following morphemes. Interrogative pronouns include: *pha* ‘what,’ *gə-day* [*gə*-that] ‘who,’ and *gə-rà* [*gə*-place] ‘where, which.’ The interrogative *gə-day* ‘who’ has a distinctive genitive form *gə-dé?* ‘whose,’ which results from a contraction of *gə-day* plus *?à?* ‘GEN.’ Interrogative

pronouns, together with interrogative adverbs such as *gə-lóy* [gə-then] ‘when,’ *gə-náy* [gə-here] ‘where,’ *gə-dè?* [gə-ALL] ‘to where,’ *gə-níy* [gə-thus] ‘how,’ form interrogative pro-forms. These interrogative pro-forms all trigger the occurrence of the mood marker for wh-questions, thereby constituting a natural class. Moreover, they can form question sentences on their own, unlike other nouns or adverbs. Jinghpaw also frequently uses lexicalized expressions involving interrogatives as their elements, e.g. *pha bò?* [what-kind] ‘what,’ *pha majo?* [what-because] ‘why,’ *gə-rà khu* [which-like] ‘how,’ *gə-rà tèn* [which-time] ‘what time.’

5.2.5 Locator nouns

Jinghpaw has a set of locator nouns which specify spatio-temporal relations of arguments, including *cøy* ‘before,’ *phay* ‘behind,’ *ntsā* ‘above,’ and *ípú?* ‘below.’ Some locator nouns express abstract meanings, as exemplified by *ntsā* ‘above,’ which also carries the sense of ‘regarding to,’ and *phay* ‘behind,’ which occurs obligatorily when animate nouns stand as the goal of movement marked by the allative *dè?*

5.2.6 Postpositions

Jinghpaw has a small set of postpositions which function like case markers, such as *dzòn* ‘like,’ *mətu* ‘for,’ *məlāy* ‘instead,’ *majo* ‘because.’ The nounhood of these postpositions can be seen from the fact that they can be modified by a genitive modifier unlike case markers, e.g. *ná?mətu* [your-for] ‘for you,’ *ná?məlāy* [your-instead] ‘instead of you,’ *ná?majo* [your-because] ‘because of you,’ etc.

5.3 Case markers

Jinghpaw is a double-marking language at the clausal level and in general dependent-marking at the NP level. The case marking pattern is the nominative-accusative type (S/A vs. P), in which S and A occur without any overt marker, in contrast to P which, in a transitive clause, is marked by an accusative marker when there is a possibility that it may be misconstrued with A, displaying the ‘anti-ergative’ marking (LaPolla 1992), as illustrated below. In possessive phrases, the possessor is basically marked by a genitive.

- (7) *nyé?* *?wâ* *ŋay* *phé?* *tsó?rà?* *ŋy-ay.*
1sg.GEN father 1sg ACC love-like 1sg-DECL
‘My father loves me.’
- (8) *nay* *phé?* *ŋay* *jəphu* *jò?* *na* *ŋy-ay.*
2sg ACC 1sg price give IRR 1sg-DECL
‘I will give you wages.’

The inventory of Jinghpaw case markers is provided in Table 53.6. Case markers, aside from the two genitive cases, occur at the clausal level. Genitive, ablative, and comitative cases can occur at the NP level. The locative *thà?* and perative *khu* are of nominal origin, coming from *ləthà?* ‘upper’ and *khu* ‘hole.’ The ablative *nná*, which is also used as a sequential subordinator and often pronounced *ná* in colloquial speech, usually co-occurs with a preceding locative case. The genitive *ná*, which seems to have been derived from the ablative *nná*, marks a possessive relationship when the possessor noun is a location or time.

TABLE 53.6 CASE MARKERS

<i>phé?</i>	accusative	p (patient, recipient, causee, applied object); path
<i>kó?</i>	locative	location of existence and event; time of event; goal of movement; possessor
<i>thà?</i>	locative	location of existence and event; time of event; goal of movement; standard of comparison
<i>ʔè</i>	locative	location of existence and event; time of event
<i>dè?</i>	allative	goal of movement; location of existence and event; time of event
<i>ñná</i>	ablative	source of movement; location of emergence
<i>khu</i>	perlicative	path; means
<i>ʔá?</i>	genitive	possessor
<i>ná</i>	genitive	possessor
<i>thè?</i>	comitative	co-participant of s/a (companion and reciprocant); instrument; means; material

5.4 The verb complex

The verb complex of Jinghpaw is fairly complex, the most expanded form of which may consist of a sequence of verbs followed by a string of various kinds of auxiliaries, which in turn are followed by verbal endings marking verbal categories such as direction, as well as such inflectional categories as number, person, aspect, and mood.

5.4.1 Verbal endings

Verbal endings are manifested as affixes or phonologically independent words, all of which occur after verbs in a fixed linear order. The template for verbal endings is given below. Person, number, aspect, and mood are obligatory categories in verbal endings. Optional morphemes marking direction, intensity, or possessor agreement, represented by direction in the template below, may intervene between the inflectional categories.

(9) Verb–Number–Aspect–(Direction)–Person–Mood

As an illustration, consider the following sentence, which illustrates all categories in the template. Observe in the example that the verb agrees in number and person with a possessor argument, the agreement of which is signaled by the possessor agreement marker *l-* ‘PA’ as well as that the cross-referencing is not always with core arguments.

(10)	<i>nánthe</i>	<i>ʔá?</i>	<i>myit-məsín</i>	<i>thà?</i>	<i>yon-khyen</i>	<i>ʔay</i>
	2pl	GEN	mind-liver	LOC	mourn-distress	NMLZ
	<i>phriŋ</i>	<i>màt</i>	<i>mə-sə-l-it</i>		<i>d-ay.</i>	
	full	COMPL	pl-SCM-PA-2nd		2nd-DECL	

‘Your heart is filled with grief.’

5.4.1.1 Number

Number marking is manifested by means of prefixation. Plural number (including dual) is overtly marked by a prefix *mə-* (*m-* before a vowel). There is no overt marker for singular. Plural number may also be marked by *gá?* for 1st person in certain environments (see section 5.4.1.4).

(11)	<i>cánthe</i>	<i>ŋay</i>	<i>phé?</i>	<i>ní-mədàt</i>	<i>na</i>	<i>m-á?</i>	<i>ʔay.</i>
	3pl	1sg	ACC	NEG-listen	IRR	pl-3rd	DECL

‘They will not listen to me.’

5.4.1.2 Aspect

Jinghpaw is an aspect-prominent language with a binary aspect value of change of state and non-change of state. The change of state aspect, manifested by a prefix *sə-* (*s-* before vowels), marks a change of state, whether it is the onset or endpoint. The non-change of state aspect, which is not marked overtly, marks any other situation. The change of state prefix changes the following tone to high when it is a checked syllable.

- (12) *khokhám gò si màt s-Ø-ay.*
 king TOP die COMPL SCM-3rd-DECL
 ‘The king has passed away.’

5.4.1.3 Direction

Optional morphemes marking such categories as direction, intensity, and possessor agreement may intervene between the verb and morphemes marking inflectional categories. All the morphemes which occupy this slot consist of a single consonant, including *r-* ‘VEN,’ *s-* ‘AND,’ *l-* ‘PA,’ and *n-* ‘INTNS.’ Jinghpaw has grammaticalized a binary category of direction. The venitive (cislative) *r-* marks motion toward the deictic center and the andative (translocative) *s-* marks motion away from the deictic center, as exemplified by a minimal pair below (see DeLancey 1980, for further discussion). Observe in the examples that the Jinghpaw verb *sa* is deictically neutral, meaning both ‘go’ and ‘come,’ and the direction of movement is morphologically marked by the directional markers.

- (13) a. *sa r-it-Ø.* b. *sa s-it-Ø.*
 go/come VEN-2nd-IMP go/come AND-2nd-IMP
 ‘Come here!’ ‘Go away!’

Other morphemes that occupy the same slot include *l-* ‘PA’ and *n-* ‘INTNS.’ The morpheme *l-* ‘PA’ is used to signal agreement with a possessor of an argument, as in (10). Argument indexation itself is achieved by number and personal indices affixed to *l-*, and the function of *l-* is simply to signal that the agreement is with a possessor. The same slot may be occupied by *n-* ‘INTNS,’ the nature of which is not fully understood although it expresses intensity in imperative and hortative moods.

As pointed out by DeLancey (1980, 2011) and Dai (2012), the morphemes occupying the direction slot have their sources in auxiliary verbs, which ultimately originate from full verbs, having been reduced both phonologically and semantically. Dai (2012: 212–13) proposes that the marker *l-* ‘PA’ originates from a Jinghpaw verb *lù* ‘have, possess.’ DeLancey (1980: 165, 2011: 72) proposes that directional morphemes *r-* ‘VEN’ and *s-* ‘AND’ are from verbs denoting ‘come’ and ‘go.’ DeLancey (2011: 64) suggests that the morpheme *n-* ‘INTNS’ probably reflects an old copula.

5.4.1.4 Person

Person marking is the most complicated category in the verbal ending due to the fact that the personal indices display allomorphy, and that person indexation is based on a mixture of systems drawing on grammatical roles as well as on a person hierarchy where speech act participants (SAPs) outrank a 3rd person. Person marking on verbs is achieved by means of affixation of personal indices which consist of four sets, as summarized in Table 53.7 (tones are not indicated).

TABLE 53.7 PERSONAL INDICES

	Set Ia	Set Ib	Set II	Set III
1sg	- <i>iyy-</i>	- <i>a?</i>	- <i>i?</i>	- <i>e?</i>
1pl	<i>ga?</i>	<i>ga?</i>	- <i>i?</i>	- <i>e?</i>
2nd	- <i>ind-</i>	- <i>ind-</i>	- <i>itd-</i>	<i>d-</i>
3rd	\emptyset / - <i>a?</i>	- <i>a?</i>	- <i>u?</i>	<i>w</i> / \emptyset

Observe in the table that the number category for 1st person is formally distinguished only in Set I, and the category is neutralized in Sets II and III. Recall that in the latter case, plural number is marked by *ma-* regardless of person (section 5.4.1.1). Notice also that the forms of Sets Ia and Ib are nearly identical except the syncretism of 1st person singular and 3rd person in Set Ib, where the former patterns after the latter, and the 3rd person of Set Ia which displays two forms, where \emptyset - is used to mark a 3rd person singular in declarative mood with the direction slot unfilled and -*a?* elsewhere.

The choice between Set I and Set II is determined according to mood types or existence/non-existence of morphemes in the direction slot, regardless of aspect or verb types. This can be roughly summarized as given in Table 53.8. Set III indices are used to explicitly mark two participants (see below).

As an illustration, consider the paradigms of *sa* ‘go/come’ and *rāy* ‘COP’ in declarative (marked by *ay*) or interrogative (marked by *ni*) mood in Table 53.9. (Note that phonemes -*y*- and -*ə*- are inserted by epenthesis. As for allomorphy conditions, see below.)

As can be seen, the most complex personal indices which take the phonological shape VC_1C_2 are realized by various allomorphs. The allomorphy conditions can be schematized as follows: $VC_1C_2 \rightarrow VC_1 / C _ C$; $VC_1C_2 \rightarrow C_1C_2 / \#_V$; $VC_1C_2 \rightarrow C_1 / \#_C$. As an illustration, consider the index -*ind-* ‘2nd’ in Table 53.10 which can be realized as the following four allomorphs in non-change of state aspect depending on the allomorphy conditions given above. As can be seen, from the underlying VC_1C_2 complex, C_2 is deleted when followed by a consonant, and V is deleted when there are no preceding syllables, in which case the underlying low tone on the vowel is associated with the following coda nasal which is realized as a syllabic nasal in surface representation.³

The Jinghpaw person indexation system is based on both grammatical roles and a person hierarchy in which SAPs outrank a 3rd person (SAP > 3rd). Verb agreement is with s in intransitive clauses, as in (14), and with A in polyvalent clauses when the participant-configurations are SAP → SAP, SAP → non-SAP, and non-SAP ↔ non-SAP, as in examples (15) to (17). When the participant-configuration is non-SAP → SAP, however, verb agreement is either with SAP or non-SAP, as in examples (18) and (19). The former agreement pattern here is that of hierarchical person marking (DeLancey 1980: 25–7), in which the person marking is determined by a person hierarchy SAP > 3rd, presumably because of the greater salience of SAPs. The situation thus leads to a conclusion that the coding property of verbal agreement cannot be used to define ‘subject’ in Jinghpaw since s may align with both A and P in terms of person in the non-SAP → SAP configuration. For concreteness, consider the following examples:

- (14) *ŋay* *gò* *ʔwâ* *phay* *dè?* *wà* *na* *ŋy-ay*.
 1sg TOP father behind ALL return IRR 1sg-DECL
 ‘I am going to the Father.’

TABLE 53.8 CONDITIONS FOR THE CHOICE BETWEEN SET I AND SET II

Conditions		Exceptions
Set Ia	declarative	2nd person plural (marked by Set II)
Set Ib	interrogative, speculative, exclamative	2nd person plural (marked by Set II)
Set II	imperative, hortative, or the direction slot is filled with PA/INTNS	1st person in hortative (marked by Set Ia)

TABLE 53.9 PARADIGMS OF SA ‘GO/COME’ AND RÁY ‘COP’

	Declarative	Declarative with r- ‘VEN’	Interrogative	Declarative with l- ‘PA’
1sg	<i>sa jŋ-ay</i>	<i>sa r-ŋ̩ n-ay</i>	<i>sa ?à? n̄i</i>	<i>ráy l-i? ?ay</i>
2sg	<i>sa ñd-ay</i>	<i>sa r-ñin d-ay</i>	<i>sa ñ-n̄i</i>	<i>ráy l-it d-ay</i>
3sg	<i>sa Ø-?ay</i>	<i>sa r-à? ?ay</i>	<i>sa ?à? n̄i</i>	<i>ráy l-ù? ?ay</i>
1pl	<i>sa gà? ?ay</i>	<i>sa r-ə-gà? ?ay</i>	<i>sa gà? n̄i</i>	<i>ráy mə-l-i? ?ay</i>
2pl	<i>sa m-y-it d-ay</i>	<i>sa mə-r-in d-ay</i>	<i>sa m-y-it n̄i</i>	<i>ráy mə-l-it d-ay</i>
3pl	<i>sa m-à? ?ay</i>	<i>sa mə-r-à? ?ay</i>	<i>sa m-à? n̄i</i>	<i>ráy mə-l-ù? ?ay</i>

TABLE 53.10 ALLOMORPHS OF THE INDEX -IND- ‘2ND’

Allomorphs	Examples	Glosses
-ind-	<i>sa r-in d-ay</i>	[go-VEN-2nd-DECL]
-in	<i>sa r-in n̄i</i>	[go-VEN-2nd-Q]
ñd-	<i>sa ñd-ay</i>	[go-2nd-DECL]
ñ-	<i>sa ñ-n̄i</i>	[go-2nd-Q]

- (15) *ŋay naŋ phé?* *ɛəkòn* *jŋ-ay.*
 1sg 2sg ACC praise 1sg-DECL
 ‘I thank you.’
- (16) *ŋay ei phé?* *gùmróŋ* *ŋà* *jŋ-ay.*
 1sg 3sg ACC boast CONT 1sg-DECL
 ‘I had boasted to him (about you).’
- (17) *ei eánthe phé?* *pha mùŋ* *n-tsun* *dán yu Ø-?ay.*
 3sg 3pl ACC what also NEG-say show try 3rd-DECL
 ‘He did not say anything to them.’
- (18) *ei ŋay phé?* *ní-jə-thi?* *káw* *jŋ-ay.*
 3sg 1sg ACC NEG-CAUS-left thoroughly 1sg-DECL
 ‘He has not left me alone.’
- (19) *ei ŋay phé?* *ɛìŋ* *ŋa* *thèt* *dàt* *Ø-?ay.*
 3sg 1sg ACC thus say order release 3rd-DECL
 ‘He had said these things to me.’

Aside from the personal indices which mark only one participant discussed above, Jinghpaw also has a set of agreement indices consisting of bimorphemic indices taken from Set III which explicitly mark two participants. The difference between these two index systems seems to be based on focus: the former system is exploited when the

speaker puts one participant in focus, while the latter is used when the speaker focuses on two participants (Scott DeLancey, personal communication). Notice that, in the latter system, number agreement is with *p*, as illustrated by (23).

- (20) *ŋay nay phé?* *gərum d-è?* *?ay.*
1sg 2sg ACC help 2ndP-1stA DECL
'I helped you.'
- (21) *ŋay ei phé?* *ce w-è?* *?ay.*
1sg 3sg ACC know 3rdP-1stA DECL
'I know him.'
- (22) *mədù?* *gò ei phé?* *ɛkòn w-ù?* *?ay.*
master TOP 3sg ACC command 3rdP-3rdA DECL
'The master commanded him.'
- (23) *gà nday ŋay nánthe phé?* *tsun dá mə-d-è?* *?ay.*
word this 1sg 2pl ACC say RES pl-2rdP-1stA DECL
'I have told you this.'

5.4.1.5 Mood

Jinghpaw has grammatical mood markers which occupy the last slot in the verbal ending and manifest six different paradigmatic values: declarative marked by *?ay*, interrogative marked by *nî* or *tâ* according to whether it is a polar or wh-question, exclamative marked by *kha*, speculative marked by *doy*, imperative marked by *Ø*, and hortative marked by *gà?*. According to which mood they manifest, Jinghpaw verbal sentences can be divided into six types. Negated commands are formed by imperative mood with the adverb *khùm*, which originates from the verb *khùm* 'prohibit.' It is of interest to note that commands can be directed at any person. Canonical imperatives directed at the addressee are formed by the imperative marker *Ø*, and non-canonical imperatives directed at the speaker or 3rd person are formed by the hortative marker *gà?*.

- (24) *?nâŋ dè?* *sa r-it-Ø.*
here ALL go/come VEN-2nd-IMP
'Come here.'
- (25) *yá?* *gò ŋay báy tsun ï-gà?.*
now TOP 1sg again say 1sg-HORT
'Now let me tell you again.' (Dai and Xu 1992: 296)
- (26) *ei nánthe phaŋ dè?* *sa ?ù?-gà?.*
3sg 2pl behind ALL go/come 3rd-HORT
'Let him go to you.'

5.4.1.6 Simplification of verbal endings

The complex verbal endings described earlier, full exhibition of which can be seen in early Jinghpaw writings, are simplified in modern Jinghpaw, especially in the spoken style. For example, modern spoken Jinghpaw usually does not encode the person category on the verb. In this style, the 3rd person declarative mood marker *?ay*, which also marks verb citation and clausal nominalization, plays a versatile role, being used irrespective of

person of an argument (see Kurabe 2012, for an exemplification of this style). Using the personal indices in the spoken language sounds unnatural, judged to be highly elevated style. The simplification of verbal endings seems to be caused by language contact. As van Driem (2001: 394) puts it:

The use of a somewhat pidginised and grammatically simplified Kachin Jinghpaw throughout northern Burma as a lingua franca between various Kachin communities is a long-standing phenomenon, and the existence of this pidgin clouds the original picture of the Jinghpaw languages, as many of the dialects have been influenced by the morphologically simplified lingua franca.

The simplification of the verbal ending is also reported in Jinghpaw spoken in China, especially in the variety spoken by younger speakers (Dai 2012), as well as in Singpho in Northeast India (Morey 2010).

5.4.2 Verb serialization

Jinghpaw verbs can be serialized productively, without any overt marker of coordination or subordination, constituting a single predicate. Serial verbs are contiguous, and the syntactic and semantic relationship holding between serialized verbs may be sequential, manner, purpose, or complementation. One constraint imposed on Jinghpaw serial verbs is that A and/or S must be shared between serialized verbs with the exception of the complementation type. This behavioral property provides a piece of evidence to posit the category of ‘subject’ in Jinghpaw. Sharing of P is not obligatory.

- (27) *yay khàʔ-thuyj dè? n̄tsin sa jàʔ lù? n̄thóm...*
1sg water-tub ALL water go draw drink after
'After I went to the well and drew water and drank it...'
- (28) *ei cóm gó nàm-məlì dè? ləgyím sa m̄at n̄ná...*
3sg TOP forest-forest ALL conceal go COMPL SEQ
'He withdrew to the forest and...'
- (29) *là ləkhóy gó tsan ʔay gá dè? jùm məri sa m-à? ʔay.*
man two TOP far NMLZ land ALL salt buy go pl-3rd DECL
'Two men went far to buy salt.'
- (30) *ʔù-gənù gó gəcà ni phé? pyen cərín ya na mətu...*
bird-mother TOP child pl ACC fly teach BEN NMLZ for
'The mother bird (brought out her children from the nest) to teach them how to fly...'

Jinghpaw has a set of auxiliary-like verbs which occur in verb serialization, semantically modifying the other verbs. The fact that they, in serialization, can be negated by the negative prefix suggests that they are genuine verbs. These auxiliary-like verbs include *ce* ‘know, be able to, be in the habit of,’ *lù* ‘get, can,’ *may* ‘be good, may,’ *má?* ‘exhausted, entirely,’ and *kam* ‘be willing.’ Many of these verbs are special in that they can occur both in pre- and post-head positions. As an illustration, consider the verb *ce* ‘know’:

- (31) *jan-cà ni cà cəŋ ce mənaw m-à? ʔay.*
sun-people pl only first know dance pl-3rd DECL
'The first time, only people of the sun are able to dance the Mānau dance.'

TABLE 53.11 AUXILIARIES

<i>ŋà</i>	'CONTINUOUS'	< 'live'	<i>bù?</i>	'INTENSIFIER'	< 'infected'
<i>to</i>	'CONTINUOUS'	< 'lie down'	<i>si</i>	'INTENSIFIER'	< 'die'
<i>màt</i>	'COMPLETIVE'	< 'lost'	<i>ɛɔŋjún</i>	'CAUSATIVE'	< 'dispatch'
<i>yu</i>	'CONATIVE'	< 'see'	<i>lóm</i>	'COMITATIVE APPLICATIVE'	< 'accompany'
<i>khát</i>	'RECIPROCAL'	< 'dispute'	<i>ya</i>	'BENEFACTIVE APPLICATIVE'	< 'give'

- (32) *day* *?ù-?əmyú* *gò* *grày* *gəthèt* *?ay* *dzàybrù-praŋ* *ni* *thà?* *çà*
 that bird-race TOP very hot NMLZ sand-plain pl LOC only
ŋà *ce* *m-à?* *?ay.*
 live know pl-3rd DECL

'That kind of bird only lives in a very hot desert.'

5.4.3 Auxiliaries

Auxiliaries, which are optional for the verb complex, occur after (serialized) verbs, expressing such meanings or functions as aspectuality, modality, evidentiality, intensity, and voice. Auxiliaries can be chained in a single verb complex, the order of which is not strictly constrained. Many Jinghpaw auxiliaries are of verbal origin, derived through grammaticalization processes. A few random examples are provided in Table 53.11.

5.5 Nominalization and subordination

The pervasive use and multifunctionality of clausal nominalization is a prominent feature of Jinghpaw grammar. Nominalized clauses are formed chiefly by *?ay*, which also marks verb citations and declarative mood. A nominalized clause plays a versatile role, functioning as a complement clause (verbal or nominal), as in (33), relative clause (headed or headless), as in (34), (35), adverbial clause, as in (36), or non-embedded, independent clause followed by a copula, as in (37). This kind of special relationship between nominalization and subordination is a widespread phenomenon within TB languages (Matisoff 1972). Consider the following Jinghpaw examples (square brackets are used to enclose nominalized clauses):

- (33) *khà?* *cədón* *yu* *yày* *[fləlam khun sùŋ ?ay]* *ce* *m-à?* *?ay.*
 water measure try when fathom 20 deep NMLZ know pl-3rd DECL
 'They found that the water was 120 feet deep when they measured it.'
- (34) *[yay ?ó-tsóm* *çà* *nó?* *ń-myít* *yu* *?ay]* *[cəkhyon* *ṛà?* *khəlém*
 1sg ADV-beautiful only still NEG-think try NMLZ wolf GEN deceive
?ay] *gà* *phé?* *mədàt* *cút* *káw* *s-Ø-è?* *?ay.*
 NMLZ word ACC listen mistake thoroughly CSM-3rdp-1sta DECL
 'I mistakenly had listened to the wolf's lying words which I did not think about carefully.'
- (35) *[day ni thà? jòn ?ay]* *ni* *phé?* *ŋay* *mù* *ŋy-ay.*
 that pl LOC ride NMLZ pl ACC 1sg see 1sg-DECL
 'I saw those who rode them (horses).'

- (36) *day mà gò [ní-thá? lá ?ay] cà ?wâ phay dè?*...
that child TOP NEG-pick take NMLZ only father behind behind ALL
'The child (went) to his father without picking it up...'
- (37) *[ŋay nay phay dè? cəŋún dàt ?ay] ráy nná.*...
1sg 2sg behind ALL send release NMLZ COP SEQ
'I sent (him) for you, and...'

ADDITIONAL ABBREVIATIONS

CONT	continuous
CSM	change of state marker
HAB	habitual
HORT	hortative
INTNS	intensity
REDUP	reduplicant
SEQ	sequential
VEN	venitive

NOTES

- 1 I would like to express my gratitude to Professor Atsuhiko Kato and the editors for their valuable comments on an earlier version of this paper. My fieldwork was supported in part by a Grant-in-Aid for JSPS Fellows (Nos. 24-2938 and 26-2254) from the Japan Society for the Promotion of Science (JSPS).
- 2 Example sentences are mostly taken from *Jinghpaw Readers* and the Jinghpaw Bible, which provide a rich exemplification of verbal endings lost in modern spoken Jinghpaw (see section 5.4.1.6).
- 3 Because of the complex conditions and allomorphy, personal indices are treated as portmanteau with morphemes marking direction, intensity, possessor agreement, and mood in most previous studies, including Hanson (1896), Dai and Xu (1992), and Dai (2012), which employ quite a large number of paradigms or lists to represent them, with the exception of DeLancey (1980, 2011).

REFERENCES

- Benedict, Paul K. 1972. *Sino-Tibetan: A Conspectus*, contributing editor James A. Matisoff. New York: Cambridge University Press.
- Bradley, David. 1996. Kachin. In Stephen A. Wurm, Peter Mühlhäusler and Darrell T. Tryon (eds), *Atlas of Languages of Intercultural Communication in the Pacific, Asia, and the Americas*, Vol. 2.1. Berlin: Mouton de Gruyter, 749–751.
- Burling, Robbins. 1983. The Sal languages. *Linguistics of the Tibeto-Burman Area* 7.2: 1–32.
- Dai, Qingxia. 2012. *Jingpoyu cankao yufa (A Jingpo Reference Grammar)*. Beijing: China Social Sciences Press.
- Dai, Qingxia, and Xu Xijian. 1992. *Jingpoyu yufa (The Grammar of Kachin)*. Beijing: Central Nationalities University Press.
- Dai, Qingxia, and Wu Hede. 1995. Jinghpaw prefixes: Their classification, origins, and implications for general morphology. *Senri Ethnological Studies* 41: 93–131.
- DeLancey, Scott. 1980. Deictic categories in the Tibeto-Burman verb. Unpublished PhD dissertation, Indiana University.

- DeLancey, Scott. 2011. Nocte and Jinghpaw: Morphological correspondences. In Gwendolyn Hyslop, Stephen Morey and Mark Post (eds), *North East Indian Linguistics 3*. New Delhi: Foundation/Cambridge University Press India, 61–75.
- Dixon, R. M. W. 1977. Where have all the adjectives gone? *Studies in Language* 1: 19–80.
- Hanson, Ola. 1896. *A Grammar of the Kachin Language*. Rangoon: American Baptist Mission Press.
- Hanson, Ola. 1906. *A Dictionary of the Kachin Language*. Rangoon: American Baptist Mission Press.
- Kurabe, Keita. 2012. Jingpho dialogue texts with grammatical notes. *Asian and African Languages and Linguistics* 7: 121–153.
- Kurabe, Keita. 2014. The reflexes of the Proto-Jingpho glides in modern dialects. *Linguistics of the Tibeto-Burman Area* 37.2: 181–197.
- Kurabe, Keita. 2016. A grammar of Jinghpaw, from northern Burma. Unpublished PhD dissertation, Kyoto University.
- LaPolla, Randy J. 1992. Anti-ergative marking in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 15.1: 1–9.
- Leach, Edmund R. 1954. *Political Systems of Highland Burma: A Study of Kachin Social Structure*. London: G. Bell and Sons.
- Maran, La Raw. 1979. A dictionary of modern spoken Jingpho. Unpublished manuscript.
- Matisoff, James A. 1972. Lahu nominalization, relativization, and genetivization. In John P. Kimball (ed.), *Syntax and Semantics* 1. New York: Academic Press, 237–257.
- Matisoff, James A. 1973. Tonogenesis in Southeast Asia. In Larry M. Hyman (ed.), *Consonant Types and Tone*. Los Angeles: University of Southern California, 71–95.
- Matisoff, James A. 1986. Hearts and minds in South-East Asian languages and English: An essay in the comparative lexical semantics of psycho-collocations. *Cahiers de Linguistique Asie Orientale* 15: 5–57.
- Matisoff, James A. 1994. Watch out for number ONE: Jingpho *ŋai* ‘I’ and *ləŋai* ‘one’ (with some speculations about Jingpho number TWO). *Linguistics of the Tibeto-Burman Area* 17.1: 155–165.
- Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley and Los Angeles: University of California Press.
- Matisoff, James A. 2013. Re-examining the genetic position of Jingpho: Putting flesh on the bones of the Jingpho/Luish relationship. *Linguistics of the Tibeto-Burman Area* 36.2: 15–95.
- Morey, Stephen. 2010. *Turung: A Variety of Singpho Language Spoken in Assam*. Canberra: Pacific Linguistics.
- van Driem, George. 2001. *Languages of the Himalayas: An Ethnolinguistic Handbook of the Greater Himalayan Region*. Leiden: Brill.
- Wolfenden, Stuart N. 1929. *Outlines of Tibeto-Burman Linguistic Morphology*. London: Royal Asiatic Society.
- Xu, Xijian. 1987. Jingpoyu de liangci (Classifiers of Jingpo). *Minzu Yuwen* 1987.5: 27–35.
- Xu, Xijian, Xiao Jiacheng, Yue Xiangkun and Dai Qingxia (eds) 1983. *Jing-Han cidian (Jingpo-Chinese Dictionary)*. Kunming: Yunnan People’s Press.
- Yue, Xiangkun, Dai Qingxia, Xiao Jiacheng and Xu Xijian (eds) 1981. *Han-Jing cidian (Chinese-Jingpo Dictionary)*. Kunming: Yunnan People’s Press.

INDEX

Page numbers in italics refer to figures and tables.

- ablaut 262–3, 376
Achang 74
AdjN order 72–3
adpositions (postpositions, prepositions) 52–3;
 Bai 71; Garo 252–3; Mandarin 75–6
adverbial affixes 246
adverbial clauses: Chantyal 512–13; Lhasa
 Tibetan 399; Manange 530; Nar-Phu 554–5;
 Tshangla 431–2; Wambule 753
adverbs: Chinese 130–1; Garo 253–4
affirmative aspect 124
affricate series 105–6, 107
agreement marking 263, 264
Akha 54, 885–900
aktionsart *see* aspect marking
Altaic languages 8, 101
‘Amri Karbi’ 302, 318
Angami-Pochuri subgroup 21–2, 231
Anong 25, 635, 644
anti-agentive/anti-ergative marking 53–4, 57
Ao subgroup 218, 230, 277; *see also*
 Mongsen Ao
Apatani 52, 323
Arakanese dialect 856
Arunachal Pradesh 213, 215, 322, 418
Asakian subgroup (proposed) 20, 225
aspect marking: Belhare 705; Chinese 124,
 155; Hakha Lai 265
Assam 214
Assamese 227, 324
Athpare 648, 649, 652, 664
attributive constructions 150–2
Austroasiatic languages: contact with ST
 languages: Chinese 7, 8, 101; Karbi 23, 302,
 319; presence in NE India 213, 214
Austronesian languages: contact with Chinese
 8; proposed supergroup relations with ST 6
Bahing 648, 649, 650, 651, 654, 658, 663,
 665, 670
Bai 5, 7, 30–1, 47, 50, 71, 81; relationship to
 Chinese 87
Bantawa 26, 648, 649, 652, 654, 667, 669, 675
Beijing dialect 164
Belhare 52, 648, 649, 652, 653, 654, 659,
 696–719
benefactive constructions 52
Bengali 227, 256
Bhutan 404, 406, 418
Bodish subgroup 9–10, 48; word order 74, 79;
 see also Classical Tibetan, Dolakha Newar,
 Kathmandu Newar, Kurtöp, Lhasa Tibetan,
 Tshangla
Bodo-Koch 20
Borgohain, Bipin 350
Boro subgroup 226
Boro-Garo subgroup 215, 218, 224, 225–6,
 227, 235
borrowings *see* loanwords
Buddhism 406
Bugun 220
Bumthap 13
Burmese 9, 30, 856–75
Burmese writing 9, 858, 862–6
Burmese-Yipho subgroup *see* Lolo-Burmese
 subgroup
Camling 26, 52, 722–35
Campbell, Archibald 959
Cantonese 169–70; contact influence 176;
 grammar 177–80; historical development
 170; phonology 171–5; vocabulary 175–6;
 writing 176–7
case marking 52–3; Belhare 699–700; Camling
 725–6; Cogtse Rgyalrong 591–2; Garo
 252–3; Hakha Lai 261; Jinghpaw 1001–2;
 Kathmandu Newar 455–7; Khroskyabs 611;
 Kiranti 653; Kurtöp 410; Lhasa Tibetan
 389–90, 398–9; Manange 524–5; Mongsen
 Ao 289–92; Sangla Kinnauri 759–60; Stau
 604–5, 611; Tamang 482; Tangut 611;
 Tshangla 423–5; Wambule 740–1
causative constructions 52
Central Naga subgroup 21, 230

- Chakpa village 299
Chamchang orthography 355–7
Chamling 648, 653, 669, 675
Changki 277, 278, 279
Changzhi dialect 166–7
Chantyal 468, 494–513, 514
Chao tone numbers 92
Chepang 27
Chhathare 648
Chhintang 648, 649, 669
Chinese (Sinitic subgroup) 6–7; contact with other ST languages 7–8; dialect classification 85, 87, 88, 101, 114; grammar 48, 74–80, 114–31; historical development 96–101; morphosyntax 48; origin 85–7; phonology 87–96, 101–10; sentence types 131–60; *see also* Cantonese, Mandarin dialects, Shanghainese
Chinese writing 85, 200–9
Chug 220
Chungli 277, 278, 279
Classical Chinese 76, 208
Classical Newar 56
Classical Tibetan 10, 369–83, 840
classifiers 46; Burmese 869; Camling 725; Chantyal 500; Chinese 115, 160; Garo 249–50; Jinghpaw 1000; Lahu 924; Lizu 830
causal morphology 45–6
clause-chaining: Classical Tibetan 382–3; Dolakha Newar 444–5; Karbi 317; Kathmandu Newar 463–4; Kurtöp 412–13; Tshangla 432–3
Cogtse Rgyalrong 572–92
comparative degree 147–9, 150
comparison constructions: Chinese 76, 147–50; Cantonese 178–9; Japhug 628–9; Kathmandu Newar 465; Lai 266; Tamang 487
complement clauses: Chinese 151–2; Japhug 631–3; Lhasa Tibetan 399–400; Stau 606–8; Tshangla 430–1
compounding 266
computers 205
conjoining constructions 160
consonants: Akha 886–7; Belhare 697–8; Camling 722–3; Chantyal 495–6; Chinese 89–91, 110; Cantonese 171; Old/Middle Chinese 110; Shanghainese 185–7; Cogtse Rgyalrong 573; Dolakha Newar 437; Dulong 635; Garo 244–5; Hayu 681; Japhug 614; Karbi 303; Kathmandu Newar 454; Kayah Li 934; Kiranti 648–9; Kurtöp 407, 408; Lahu 919; Lepcha 960; Lisu 903–4; Lizu 823–5; Nar-Phu 537–8; Prinmi 791–2; Pwo Karen 943; Qiang 773–4; Sangla 756; Tamang 469; Tangsa 357–8; Tangut 806–8; Tani subgroup 324; Wambule 736, 738–9; Yongning Na 841–2; Zaiwa 877
constituent order 46–7
contact influence 4; *see also* loanwords
converbs 295–8, 662–3
copular sentences/verbs: Chantyal 505, 510; Chinese 131–2; Dolakha Newar 441; Hakha Lai 273; Hayu 688; Karbi 306; Kiranti 659–60; Lhasa Tibetan 393; Nar-Phu 549–50, 552–3; Tshangla 427–8
Count of Monte Cristo 857
covers 75

Dakpa 13, 14
Datsang dialect 572
DemN order 74
demonstratives: Chantyal 501–2; Chinese 119–21; Classical Tibetan 374; Garo 248–9; Hakha Lai 260; Lhasa Tibetan 389; Tuja 988–9; Wambule 742
dentilabilization 104–5
Deuri 226
deverbal modifiers 440
devoicing 104
Dgebshes 597
dialect chains 3, 4
Digaru 28
diphthongs 189
directional marking 16, 19; Chinese 156; Hakha Lai 264
discourse status marking 261
disjunctive questions 136–7
disposal forms 141–2
Dolakha Newar 25, 30, 56, 436–51
double-object construction 140–1
double-subject construction 145–7
'drift' 3
duality of patterning 200
Dulong 25, 51, 54, 635–43, 644
Dulong-Rawang subgroup 24
Dumi 649, 658, 660, 664, 672
Dunganese dialect 115
Dungkar 406
Duoxu 17
durative aspect 128–9
Dzala 405
Dzatang dialect 572
Dzongkha 48, 404, 405, 406

East Bodish subgroup 11, 12–15
Eastern Kayah Li *see* Kayah Li
Eastern Zhou period 204
English influence 171, 176–7, 256

- epistemic modality marking 265
equal/equalling degree 149, 150
Ergong *see Stau*
Ersu 17
Ersuic subgroup 15, 16, 17
evidential marking 54; Hakha Lai 265; Japhug 617–20; Kurtöp 414–15; Lhasa Tibetan 392–5; Meitei 344; Stau 608–9
excessive degree 178
existential verbs 51–2; Chinese 132–3; Stau 603–4
experiential aspect 127
- final consonants *see* consonants
final position classes 310
finite suffixes 392
Flowery Lisu 903
Fraser script (Lisu) 902–3, 915
- Gan dialects: grammar 122, 127, 128, 144; phonology 101
Garo 213, 214, 226, 235, 243–4; contact influence 256; grammar 246–7; phonology 244–56
genealogical subgroups 5
genitives 249
Ghale 11
glottal stops 245
glyphs 201
Gongduk 15
grammar *see* morphosyntax, world order
grammaticalized grammatical relations 46
Grünwedel, Albert 959
Guabie Na 840
Gurung 48, 468
- Hǎiyán dialect 117
Hakha Lai 233, 258; grammar 259–73; phonology 258–9
Hakka dialects: grammar 122, 127, 128, 131, 138, 145; phonology 101
Han dynasty 205
Hanyu Pinyin romanization scheme 111
Hayu 648, 649, 650, 652, 654, 658, 661, 662, 663, 667, 669, 670, 673, 674, 680–94
Hindi 324
Hmar 233
Hmong-Mien languages: grammar 71, 77; influence on Chinese 8, 101; supergroup relations with ST 6
Hong Kong 169, 175, 177
honorifics: Lhasa Tibetan 397; Manange 521; Nar-Phu 544
horizontal word transfer *see* loanwords
- Horpa *see* Stau
Hruso 218, 222
Huojia dialect 126, 166
- ideographs 201
ideophonic elements 265–6
Idu 28, 51, 213, 217, 218, 223, 234
Idu-Digaru/Idu-Tawrā subgroup 28, 223–4, 323
imperative forms 41; Belhare 705; Camling 730; Chinese 157; Classical Tibetan 376, 377, 378, 380; Dulong 639; Hayu 684–5; Japhug 616; Karbi 316–17; Lizu 834; Sangla Kinnauri 768–9; Tshangla 428; Tshobdun Rgyalrong 567; Wambule 749; Zaiwai 881–2
- Indic languages: contact with Sal languages 51; word order 72
'Indo-Burmic' languages 280
Indo-European languages: contact with Chinese 7
Indo-sphere 51
infixed words 167
initial consonants *see* consonants
International Phonetic Alphabet (IPA) 674, 875
interrogative forms: Chantyal 511; Chinese 135–40; Shanghainese 196–7; Classical Tibetan 380; Cogtse Rgyalrong 592; Hakha Lai 273; Jinghpaw 1000–1; Karbi 316; Kayah Li 940; Hayu 686; Kathmandu Newar 465; Lhasa Tibetan 401; Lizu 836; Meitei 343, 345–6; Nar-Phu 553–4; Pwo Karen 951–2; Tamang 487–8; Tangut 819–20; Wambule 741–2
- Intha dialects 856–7
intransitivizing (verbs) 42
- Japanese 209
Japhug 11, 557, 602, 609, 610, 614–33
Jero 648, 650, 736
Jiaxiān dialect 119
Jin dynasty 8
Jinghpaw 20, 21, 25, 32, 51, 52, 280, 993–1009
Jingpho subgroup (proposed) 225, 227
Jingpho-Asakan 225, 227
Jinhuá dialect 131–2
jussive mood 273
- Kachin people 993
Kai script 205
Kaman 28, 223, 224
Kaman-Meyor subgroup (proposed) 28–9, 228
Kamengic (Kho-Bwa) subgroup 28, 220
Kamnyu dialect 614
Karbi 5, 21, 23, 217, 218, 231, 232, 302–18

- Karenic subgroup 27, 50, 81; *see also* Eastern Kayah Li, Pwo Karen
Kathmandu Newar 25, 30, 54, 453–65
Kathmandu Valley 436, 453
Kayah Li 932–41
Khaling 658, 669, 672, 675
Kham 27, 49, 54, 57
Kham-Magar subgroup 10
Khara Khoto 805
Khroskyabs (Lavrung) 557, 597, 601, 602, 609–11
Kinnaur 770
Kiranti cluster/subgroup 10, 18, 24, 25–6, 27, 49, 50, 54, 74, 370, 453, 646–74; *see also* Belhare, Camling, Hayu, Wambule
Kiranti-Rgyalrong-Rawang subgroup (proposed) 56
Konyakian subgroup 20, 280
Korean 209
Koro-Milang subgroup 29, 222, 223, 231, 235, 323
Kozlov, P.K. 805
Kuki-Chin subgroup 21, 22–3, 74, 230, 232
Kulung 648, 649, 650, 654, 662
Kurtöp 13, 404–15
- Ladefoged, Peter 739
Lahu 42, 52, 53, 886, 898, 916, 918–29
Lai 233, 258 *see also* Hakha Lai
language endangerment: Kurtöp 406–7; Tani subgroup 322–3
Lavrung *see* Khroskyabs
Lepcha 48, 230, 647, 959–71
lexical items 203–4
Lhasa Tibetan 385–401
Lhokpu 15
Liao dynasty 8
light verbs 397
Limbu 648, 649, 650, 652, 653, 654, 658, 660, 662, 663, 664, 667, 669, 670, 672, 673
lingua franca status 4–5
linguistic research (NE India) 215–16
Lipo 902, 903, 909, 910, 911, 915
Lish 220
Lisu 25, 50, 227–8, 902–15
Lizu 17, 823–37
loanwords/borrowings (horizontal word transfers): Akha 886; Bai 30, 50; Boro-Garo 214; Burmese 859, 861, 864, 865; Chinese 7; Cantonese 171, 176; Dolakha Newar 437; Garo 256; Japhug 11, 614; Jinghpaw 993, 994; Karen 27; Lizu 825, 827; Loloish 9; Manange 521; Naic 840; Nar-Phu 534; Puroik 222; Pwo Karen 943; Qiang 776; Stau 599; Yongning Na 840–1
- locative verbs 51
logographic writing 202
Lolo 81
Lolo-Burmese subgroup 4, 5, 8–9, 19, 21, 32, 46, 50, 227–8, 586, 611; *see also* Akha, Burmese, Lahu, Lisu, Zaiwa
Lolopo 902, 903
Luksóm, Dóngtshen 959
Lushai 586
- Magar 27
Mainwaring, George Byres 959
Manange 468, 514, 516–31, 534
Manangs 534
Manchu 8
Mandarin dialects: geographic distribution 87, 164; grammar 75–8, 123; historical development 164–6; lingua franca status 5; phonology 90, 93, 94–5, 101, 104, 105, 111; regional characteristics 167–8
Manipur 215, 232–3
Marrison, Geoffrey 230
Maru 74
McMahon Line 418
Meghalaya 214
Mei, Tsu-lin 164
Meitei/Meithei 5, 21, 23–4, 218, 231, 232–3, 291, 338–48
mergers (phonology) 105–6
Mey 220
Meyor 28–9, 224
Middle Burmese 859–60
Middle Chinese (MC) 85, 97–9, 105, 108, 109–10, 208
middle voice marking 54
Miji 218, 222
Min dialects: grammar 122, 125, 128, 134, 138; phonology 95, 101, 105, 108
mirativity: Kurtöp 414; Tshangla 426–7
Mising 214, 406
Mizo 214
Mizoram 214
modal elements 264
Modern Chinese 97
modifiers 250–1
modifying clauses 150–1
Mon 9, 856, 857, 859
Mon-Khmer languages: contact with ST languages 4; Chinese 7; Karenic 27, 932, 933; Lolo-Burmese 9; word order 71, 72, 77
Monastic script (Pwo Karen) 942
Mongol languages 8

- Mongsen Ao 218, 277
monosyllabicity 88
mood marking: Belhare 703; Hakha Lai 266;
Jinghpaw 1006; Mongsen Ao 292–4;
Tshangla 428–9
morphosyllabic writing 202–3
morphosyntax 40–55
Mossang, Lakhum 357
Mru 5
Mru-Hkongso 23
Mueshaungx 355
Muklom 352
Müping dialect 116
Muya 16
Myanmar 350, 856, 875
Myanmar Language Commission 858
- Na Dene languages 6
'Naga' languages/peoples 228–9
Nagaland 215, 277, 279
Naic/Naish (Naxi) subgroup 18–19, 32, 52,
611, 840; *see also* Yongning Na
Nar-Phu 468, 494, 507, 514, 525, 534–55
nasal codas 102–4
Naxi *see* Naic/Naish subgroup
NDem order 74
negative marking 45–6; Belhare 700–3;
Camling 730–1; Chantyal 510–11; Chinese
133–5; Shanghainese 197–9; Classical
Tibetan 380; Cogtse Rgyalrong 592; Karbi
310; Kathmandu Newar 465; Kayah Li 939;
Kiranti 658–9; Lhasa Tibetan 400–1; Lizu
835; Nar-Phu 553; Pwo Karen 950–1;
Sangla Kinnauri 768; Stau 601; Tangut
816–19; Tshangla 426
Nepal 418, 436, 468, 494, 516, 534, 722, 736
Nepali 437, 494, 534, 673, 736
neutral questions 137–40
Newar 25, 30, 48, 56, 436, 453; *see also*
Dolakha Newar, Kathmandu Newar
Ninglang Na 840
Nishi 11, 74
nomenclature (NE India) 216–17
nominalization 42–3, 44; Belhare 707; Chinese
152; Garo 252; Hakha Lai 266–7; Japhug
624–5; Jinghpaw 1008–9; Karbi 315–16;
Kiranti 671; Kurtöp 411–12; Lahu 928–9;
Lhasa Tibetan 390–1; Stau 605–6; Wambule
743, 750–2
non-final verbs 433–4
Norman, Jerry 109
Northeast India 213–34
Northern Naga 20, 225, 226, 227
noun complement clauses 151–2
- noun phrases: Akha 891–4; Belhare 713–14;
Burmese 871–2; Chantyal 508; Cantonese
179; Shanghainese 192–3; Classical Tibetan
374–5; Garo 247–53; Hakha Lai 268–9;
Hayu 693–4; Japhug 627–8; Jinghpaw
998–9; Karbi 311–15; Kathmandu Newar
455–7, 461; Kayah Li 936–7; Lahu 924–6;
Lepcha 960–7; Lisu 906–11; Lizu 827–30;
Mongsen Ao 288–9; Nar-Phu 551–2; Prinmi
797–8; Qiang 775–7; Sangla Kinnauri
757–62; Stau 604–5; Tamang 475–7; Tangut
815–16; Tshangla 420–3; Tuja 982–3;
Yongning Na 848–51; Zaiwa 880
- nouns: Burmese 867; Camling 724–5; Chantyal
497–9; Chinese 114; Classical Tibetan
372–3, 375; Cogtse Rgyalrong 574–5;
Dolakha Newar 438–9; Garo 251–2; Hakha
Lai 260–2; Hayu 685–6; Japhug 626–7;
Kiranti 651; Lhasa Tibetan 388; Nar-Phu
543–4; Tangut 809–10; Tani subgroup 328;
Yongning Na 845–7
- number marking: Belhare 699; Hakha Lai 261;
Kathmandu Newar 455; Sangla
Kinnauri 758
- numerals 74; Chantyal 500; Classical Tibetan
373–4; Cogtse Rgyalrong 590–1; Dolakha
Newar 439–40; Garo 249; Hakha Lai 261–2;
Hayu 686; Jinghpaw 1000; Karbi 232;
Kathmandu Newar 457; Kiranti 652–3;
Lahu 924; Lizu 830; Wambule 740
- Nusu 52
Nyishi 323
- Old Burmese 859
Old Chinese (OC) 40, 48, 85, 97, 109–10, 207
'Old Kuki' languages 233
Old Tibetan (Written Tibetan) 10, 40, 369–70
Ole 354
Oracle Bone inscriptions 204
ordinal numbers 267
OV languages 70, 72, 79
- palatalization 104, 105
Pali 9, 857, 858–9
Panchthar 652
Pangwa 352–4
paper 204
Pasi-Padam 219
passive (patient-as-subject) constructions
143–5; Cantonese 177–8
- perfective aspect 124–6
person marking 51, 56–7
phase particles 126–7
Phedappe 648

- Phong 352
phonology 85–110; *see also* consonants, syllable structure, vowels
phylogenetic subgroups 3–4
pictographs 201–2
Pingyao dialect 167
pivotal constructions 159
Pollard script (Lisu) 902, 915
polyandry/polygamy 406
possessive constructions 260
post-tense suffixes 246
postpositions *see* adpositions
Pre-Modern Chinese 97
prefixes 40–3; prepositions *see* adpositions
Prinmi 17, 790–803
Prins, Marielle 573
printing 204
progressive aspect 127–8, 129
pronouns/pronominalization 46; Burmese 867–9; Camling 723–4; Chantyal 501; Chinese 115–18, 162; Classical Tibetan 373; Cogtse Rgyalrong 575–7; Dolakha Newar 438; Garo 248; Hakha Lai 260; Hayu 686; Jinghpaw 999–1000; Karbi 306; Kiranti 651–2; Lhasa Tibetan 389; Lizu 829–30; Meitei 342; Nar-Phu 545–7; Sangla Kinnauri 761; Tangut 810; Wambule 741; Western Himalayan 26
Proto-Ao 21, 292
Proto-Bodish 53
Proto-Boro-Garo 227
Proto-Central-Naga 21
Proto-Ersuic 16, 17
Proto-Karenic 933
Proto-Kuki-Chin 50
Proto-Lolo-Burmese 8–9
Proto-Sino-Tibetan 24; morphosyntax 40, 43; word order 78, 79–80
Proto-Tamangic 11
Proto-Tangkhul 22
Proto-Tani 27, 323, 324
Proto-Tibetan 10
Puma 648, 649
Pumi 7, 15, 16, 17, 81, 52, 611
Puroik 28, 220, 222
Putonghua standard 6, 164, 170
Pwo Karen 942–58
Pyu 30, 857

Q-word questions 135–6
question particle 45, 135
questions *see* interrogative forms
Qiang 15, 16, 17, 46, 52, 54, 81, 586, 773–86
Qiāngic subgroup 15–17, 19, 24, 48–9; *see also* Lizu, Prinmi, Qiang, Tangut
Qing dynasty 8

‘Rai’ 646
Raute (Raji) 30
Rawang 21, 25, 42, 49, 51, 54, 586, 635, 637, 644
Rawang people 57
Rawang-Dulong-Anong subgroup *see* Dulong
rebus characters 206
reduplication: Chinese 128; Garo 253–4; Karbi 309–10; Kathmandu Newar 460; Lhasa Tibetan 396
reflexive voice marking 54–5; Karbi 311
relative clauses 47; Chinese 78–9; Cantonese 179; Garo 251; Hakha Lai 271–2; Hayu 693–4; Japhug 629–31; Kathmandu Newar 461–2; Kiranti 671–3; Lhasa Tibetan 390–1; Manange 529; Mongsen Ao 289; Nar-Phu 554; Tamang 478
relator nouns: Classical Tibetan 375; Karbi 305, 307; Lhasa Tibetan 390
retroflexes 111
Rgyalrong subgroup 7, 15, 16, 18, 24, 26, 49, 52, 54, 56, 57, 586; *see also* Cogtse Rgyalrong, Japhug, Stau, Tshobdun Rgyalrong
Rischel, Jørgen 959
Rung subgroup 24–5, 26, 49–50, 51, 56; *see also* Belhare, Camling, Cogtse Rgyalrong, Dulong, Hayu, Japhug, Kiranti subgroup, Sangla Kinnauri, Stau, Tshobdun Rgyalrong, Wambule
rhymes: Burmese 862, 864, 865; Chinese 97–9; Cantonese 171–2; Kayah Li 933; Pwo Karen 943–4; Stau 599; Tangut 806; Tuja 975–6

Sal (Bodo-Konyak-Jinghpaw) group 20–1, 50–1, 224–7, 280
Sanchu 220
Sangla Kinnauri 756–69, 770
Sanskrit 291, 373, 437, 454, 858
Sartang 220
sentence structure 129–31
Sgaw Karen alphabet 943
Shandong dialects 116
Shang dynasty 204
Shanghaiene: grammar 192–9; phonology 185–92
Shèngxiān Chánglè dialect 119
Showu Rgyalrong 557, 558

- Shuowen jiezi* dictionary 205
Siangic 222
Siiger, Halfdan 959
Sinha, Prabhakar 959
Sinitic languages *see* Chinese languages
Sino-sphere 51
Sino-Tibetan (ST) languages: morphosyntax 40–7; subgrouping 5–6; supergroup relations 6
Sino-Tibetan-Austronesian supergroup (proposed) 6
Situ 557, 566, 611
Sixteen Kingdoms period 8
Sprigg, Richard Keith 959
stative verbs 121
Stau (Ergong, Horpa) 557, 597–611
stop codas 102
stress patterns 305
subgroups *see* genealogical subgroups, phylogenetic subgroups
subjective evaluation marking 265
subjoining constructions 157–9
subordinate clauses/sentences 5; Belhare 718; Camling 733; Classical Tibetan 382–3; Garo 254–6; Haha Lai 270–1; Hayu 692; Jinghpaw 1008–9; Karbi 317; Kathmandu Newar 464; Kayah Li 940; Kiranti 662–4, 673–4; Lepcha 960; Lhasa Tibetan 399; Meitei 346–8; Prinmi 800–1; Tamang 480–2, 489, 490; Tangut 820–1; Tshobdun Rgyalrong 567–9; Yongning Na 853
suffixes 43–5
Sunwar 648, 649, 658, 669
supergroup relations 6
superlative degree 149, 150
SVO languages 77
syllabic consonants 188–9
syllable structure: Belhare 696–7; Burmese 860–2; Camling 723; Chinese 88; Cantonese 172–5; Old/Middle Chinese 109; Shanghainese 189–90; Classical Tibetan 371–2; Cogtse Rgyalrong 574; Dulong 646; Garo 244; Hakha Lai 259; Hayu 680; Jinghpaw 995–6; Karbi 304; Kathmandu Newar 454; Kayah Li 933; Kiranti 647; Kurtöp 409; Lahu 919; Lhasa Tibetan 387; Lisu 825–6; Manange 516–18; Meitei 339; Mongsen Ao 280–1; Prinmi 790–1; Pwo Karen 943; Qiang 774–5; Tamang 470–1; Tangut 805–6; Tshangla 418–19; Tshobdun Rgyalrong 557–8; Tujia 976; Yongning Na 841
syllable types 110
T’rung 49, 56, 57
Tai-Kadai languages: contact with ST languages: Chinese 8, 101; Karenic 27, 932–3; supergroup relations with ST 6; word order 71, 72, 77
Taixing dialect 125
Tamang 48, 52, 468–90
Tamangic subgroup 9, 11–12, 31, 370, 495; *see also* Chantyal, Manange, Nar-Phu, Tamang
Támsáng, Khárpú 959
Tangbe 468
Tangkhul subgroup 22, 231–2
Tangam 322
Tangsa 226–7, 235, 350–65
Tangut 7, 16, 49, 51, 611, 805–21
Tani languages 27, 50, 54, 57, 218, 222, 223, 322–35
Tanpa dialect 572
Tavoyan dialects 856
Tawrā 223
Tengsa 230
tenses 265
Thai 886
Thakali 468
Thulung 648, 649, 650, 653, 658, 660, 663, 664, 670, 673
Tibet 385, 418
Tibetan 10–11, 48, 54, 56, 80, 235; *see also* Classical Tibetan, Lhasa Tibetan
Tibeto-Burman languages: influence on Chinese 8; morphosyntax 48–55; subgrouping 5; NE India 217–19; word order 70–4
Tikhak 352
tone letters 111
tone melodies 190–2
tone sandhi 92–3, 111, 166–7, 976–7
tones: Chantyal 496; Chinese 91–4; Cantonese 172–5; Old/Middle Chinese 106–9; Shanghainese 190; Cogtse Rgyalrong 573–4; East Bodish 13; Hakha Lai 258–9; IPA representation 875; Jinghpaw 994–5; Karbi 304–5; Kayah Li 934; Kiranti 650–1; Lahu 919; Lhasa Tibetan 387–8; Lisu 905–6; Lisu 826–7; Lolo-Burmese 9, 10, 32; Manange 518–19; Meitei 339–40; Mongsen Ao 282–4; Nar-Phu 538–41; Prinmi 792–3; Pwo Karen 944; Tamang 470; Tangsa 359–61; Tangut 806, 809; Tani subgroup 326; Tshangla 419; Tujia 976; Yongning Na 842–3; Zaiwa 877–8
topic structures: Shanghainese 195–6; Tamang 489
Tournadre 11

- transitivizing (verbs) 43
Tripura 214
Tshangla 11, 15, 52, 213, 418–34
Tshobdun Rgyalrong 557–69
Tsu-lin, Mei 45
Tujia 5, 30, 974–91
Tungusic languages 8
- Upper Minyong Adi 219
Ursprache 959
utterance-final particles 180
- verb complements 153–7
verb phrase series 159, 952–5
verb phrases: Akha 894–6; Burmese 869–71; Chantyal 502–5; Cantonese 179–80; Shanghainese 193–4; Cogtse Rgyalrong 578–89; Hakha Lai 262–6; Jinghpaw 1002–7; Kathmandu Newar 457–9; Lahu 926–8; Lisu 911–12; Lizu 830–5; Qiang 780–4; Sangla Kinnauri 763–8; Tamang 478–80; Tujia 983–8; Yongning Na 847–8; Zaiwa 883
- Verbs: Belhare 700, 709–11; Burmese 866–7; Camling 726–32; Chinese 121; Classical Tibetan 376–8; Dolakha Newar 441–3; Garo 246–7; Hayu 682–5; Kiranti 654; Lepcha 967–70; Lhasa Tibetan 391–2; Manange 526–9; Meitei 342; Mongsen Ao 292–4; Nar-Phu 547–9; Tangut 811–13; Wambule 743–6; Yongning Na 843–5
- Vietnamese 209
- VO languages 71, 78
- vowel changes 166
- vowels: Camling 723; Chantyal 494–5; Cantonese 171–2; Mandarin 166; Shanghainese 187–8; Cogtse Rgyalrong 573; Dolakha Newar 437–8; Dulong 635–6; Garo 245–6; Hayu 680; Japhug 615; Jinghpaw 994; Karbi 303–4; Kathmandu Newar 454–5; Kayah Li 933–4; Kiranti 649–50; Kurtöp 407; Lahu 919; Lepcha 960; Lhasa Tibetan 387; Lisu 904–5; Lizu 825; Manange 518; Meitei 338; Mongsen Ao 281–2; Nar-Phu 535–7; Prinmi 791; Qiang 774; Sangla Kinnauri 756–7; Stau 599; Tamang 469–70; Tangsa 358–9; Tangut 806; Tani subgroup 325; Tibeto-Burman 218; Tshobdun Rgyalrong 538; Wambule 737–8; Yongning Na 842; Zaiwa 877
- Wambule 648–9, 650, 651, 654, 660, 664, 665, 669, 672, 736–53
- Wēnzhōu dialect 107, 108
- Western Himalayan subgroup 10, 24, 26, 49, 50, 54, 56, 370; word order 74; *see also* Sangla Kinnauri
- Western Naga languages 231
- Western Zhou bronzes 204
- Win, Rev. Gam 355, 357
- word order 46–7, 70–81
- writing *see* Burmese writing, Chinese writing
- Written Tibetan *see* Old Tibetan
- Wu dialects: grammar 117, 121, 122, 125, 127, 131, 133, 136, 138, 142, 144, 153, 156; phonology 95, 101, 108
- Xi'an dialect 107, 108
- Xiang dialects: grammar 119, 121, 123, 127, 128, 144, 153; phonology 101
- Xixia 209
- Xu Shen 205
- Yacham 230
- Yakkha 648, 649, 652, 653, 659, 669, 670, 672
- Yamphu 648, 649, 650, 661, 662, 663, 669, 670
- Yeniseian languages 6
- yes–no questions: Chinese 136; Shanghainese 196; Classical Tibetan 380; Meitei 345
- Yongning Na (Mosuo) 840–54
- Yuan dynasty 8
- Yue dialects 174–5; grammar 119, 122, 128, 131, 133, 136, 138, 145, 153, 156; phonology 96, 101, 108
- Zaiwa 21, 52, 877–84
- Zelingrong 22
- Zeme languages 21