

# The Sino-Tibetan Languages

Edited by  
Graham Thurgood  
and Randy J. LaPolla

ROUTLEDGE LANGUAGE FAMILY SERIES

**Also available as a printed book  
see title verso for ISBN details**

THE  
SINO-TIBETAN  
LANGUAGES

---

---

## ROUTLEDGE LANGUAGE FAMILY SERIES

---

---

Each volume provides a detailed, reliable account of every member language, or representative languages of a particular family. Each account is a reliable source of data, arranged according to the natural system of classification: phonology, morphology, syntax, lexis, semantics, dialectology and socio-linguistics. Each volume is designed to be the essential source of reference for a particular linguistic community, as well as for linguists working on typology and syntax.

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

The Bantu Languages  
*Edited by Derek Nurse & Gerard Philippson*

The Languages of the Caucasus  
*Edited by Alice Harris*

The Indo-Aryan Languages  
*Edited by George Cardona & Dhanesh Jain*

The Iranian Languages  
*Edited by Gernot Windfuhr*

The Khoesan Languages  
*Edited by Rainer Vossen*

The Manchu-Tungusic Languages  
*Edited by Alexander Vovin*

The Mongolic Languages  
*Edited by Juha Janhunan*

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

The Sino-Tibetan Languages  
*Edited by Graham Thurgood & Randy J. LaPolla*

THE

---

**SINO-TIBETAN**

---

LANGUAGES

Edited by

Graham Thurgood 杜冠明

and

Randy J. LaPolla 罗仁地

First published 2003  
by Routledge  
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada  
by Routledge  
29 West 35th Street, New York, NY 10001

*Routledge is an imprint of the Taylor & Francis Group*

© 2003 Graham Thurgood and Randy J. LaPolla, selection and editorial matter; the contributors, their own chapters

This edition published in the Taylor & Francis e-Library, 2006.

“To purchase your own copy of this or any of Taylor & Francis or Routledge’s collection of thousands of eBooks please go to [www.eBookstore.tandf.co.uk](http://www.eBookstore.tandf.co.uk).”

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

The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

*British Library Cataloguing in Publication Data*

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

*Library of Congress Cataloging in Publication Data*

A catalog record for this book has been requested

ISBN 0-203-22105-2 Master e-book ISBN

ISBN 0-203-27573-X (Adobe eReader Format)

ISBN 0-7007-1129-5 (Print Edition)

---

# CONTENTS

---

<i>List of illustrations</i>	xiv
<i>Preface</i>	xviii
<i>List of abbreviations</i>	xix
<b>PART 1 OVERVIEW CHAPTERS</b>	<b>1</b>
<b>1 A subgrouping of the Sino-Tibetan languages: the interaction between language contact, change, and inheritance</b> <i>Graham Thurgood</i>	<b>3</b>
1 Introduction	3
2 Sino-Tibetan	6
3 Chinese	6
4 Tibeto-Burman	7
References	20
<b>2 Overview of Sino-Tibetan morphosyntax</b> <i>Randy J. LaPolla</i>	<b>22</b>
1 Sino-Tibetan	22
2 Chinese	29
3 Tibeto-Burman	29
References	36
<b>3 Word order in Sino-Tibetan languages from a typological and geographical perspective</b> <i>Matthew S. Dryer</i>	<b>43</b>
1 Introduction	43
2 Word order in Tibeto-Burman	43
3 Chinese	48
4 Conclusion	54
References	54
<b>PART 2 OLD CHINESE AND CHINESE DIALECTS</b>	<b>57</b>
<b>4 A sketch of late Zhou Chinese grammar</b> <i>Derek Herforth</i>	<b>59</b>
1 Introduction	59
2 Late Zhou Chinese	59
3 Predication	60
4 Modification, clause demotion and nominalization	63
5 Basic constituent order	66

6	Conclusion	69
	Additional abbreviations	70
	References	70
<b>5</b>	<b>The Chinese dialects: phonology</b> <i>Jerry Norman</i>	72
	References	82
<b>6</b>	<b>Chinese dialects: grammar</b> <i>Anne O. Yue</i>	84
	1 Introduction	84
	2 Grammatical categories	84
	3 Sentence structure and word order	94
	4 Major sentence types	96
	Further reading	124
	Bibliography	124
<b>7</b>	<b>The characteristics of Mandarin dialects</b> <i>Dah-an Ho</i>	126
	1 Mandarin dialects in historical development	126
	2 Some characteristics of Mandarin dialects	128
	3 The regional characteristics of Mandarin dialects	129
	References	130
<b>8</b>	<b>Shanghai</b> <i>Eric Zee and Liejiong Xu</i>	131
	1 Shanghai phonology <i>Eric Zee</i>	131
	2 Shanghai morphology and syntax <i>Liejiong Xu</i>	138
	References	145
<b>9</b>	<b>Cantonese</b> <i>Robert S. Bauer and Stephen Matthews</i>	146
	1 Introduction	146
	2 Phonology	146
	3 Lexicon: colloquial Cantonese vocabulary	150
	4 Grammar	151
	5 Conclusion	154
	References	154
<b>10</b>	<b>Chinese writing</b> <i>Mark Hansell</i>	156
	1 Structure and function	156
	2 Historical development	160
	3 Writing as evidence in historical reconstruction	163
	4 Relation to other languages	164
	References	165
	Further reading	165

<b>PART 3</b>	<b>TIBETO-BURMAN LANGUAGES AND DIALECTS</b>	<b>167</b>
<b>11</b>	<b>The Tibeto-Burman languages of Northeastern India</b> <i>Robbins Burling</i>	169
1	Introduction	169
2	Typology	173
3	Classification	173
4	The central area	175
5	The northern area	178
6	The eastern border area	182
7	Conclusions	189
	References	189
<b>PART 4</b>	<b>LOLO-BURMESE LANGUAGES</b>	<b>193</b>
<b>12</b>	<b>Burmese</b> <i>Julian K. Wheatley</i>	195
1	Background	195
2	Phonology	197
3	Word classes	200
4	Syntax (constructions and particles)	202
	Additional abbreviations	206
	References	206
	Further reading	207
<b>13</b>	<b>Lahu</b> <i>James A. Matisoff</i>	208
1	Introduction	208
2	Phonology	209
3	Lahu vocabulary and word formation	210
4	Clause structure	211
5	Form classes	212
6	Argument–verb relations	212
7	The noun phrase	214
8	The verb phrase (VP)	217
9	Nominalization and relativization	219
	References	220
<b>14</b>	<b>Lisu</b> <i>David Bradley</i>	222
1	Introduction	222
2	Lisu phonology	223
3	Lisu morphosyntax	226
	References	235
<b>15</b>	<b>Akha</b> <i>Inga-Lill Hansson</i>	236
1	Introduction	236
2	Phonology	237
3	Word formation	239
4	Syntax	241



5	Sentence particles	247
6	Noun incorporation	249
7	Final particles	251
	References	251
<b>PART 5 BODISH LANGUAGES</b>		<b>253</b>
<b>16</b>	<b>Classical Tibetan</b> <i>Scott DeLancey</i>	255
1	The Tibetan language	255
2	Phonology	256
3	Word classes and inflections	258
4	Word formation	262
5	The noun phrase	263
6	Clause and sentence	265
	References	267
<b>17</b>	<b>Lhasa Tibetan</b> <i>Scott DeLancey</i>	270
1	Phonology	270
2	Nouns, adjectives, and nominal morphology	273
3	The verb	276
4	Word formation	281
5	Syntax	282
	References	286
<b>PART 6 TGTM LANGUAGES</b>		<b>289</b>
<b>18</b>	<b>Tamang</b> <i>Martine Mazaudon</i>	291
1	Introduction	291
2	Phonology	291
3	Typological summary	293
4	Basic sentences	293
5	The noun phrase	297
6	The verb phrase	300
7	Complex sentences	302
8	Comparison	309
9	Questions	310
10	Information structure	311
	Additional abbreviations	313
	References	313
<b>19</b>	<b>Chantyal</b> <i>Michael Noonan</i>	315
1	Introduction	315
2	Phonology	315
3	Morphology	318
4	Word formation	327
5	Syntax	329
	References	334

<b>20</b>	<b>Nar-Phu</b> <i>Michael Noonan</i>	336
	1 Introduction	336
	2 Phonology	337
	3 Morphology	340
	4 Word formation	347
	5 Syntax	348
	References	352
<b>PART 7 NEWAR DIALECTS</b>		<b>353</b>
<b>21</b>	<b>Dolakhā Newār</b> <i>Carol Genetti</i>	355
	1 Introduction	355
	2 Phonology	356
	3 Morphology	357
	4 Verbs	359
	5 Syntax	361
	6 Narrative text	364
	Additional abbreviations	369
	References	370
<b>22</b>	<b>Kathmandu Newar (Nepāl Bhāśā)</b> <i>David Hargreaves</i>	371
	1 Introduction	371
	2 Phonology	372
	3 Inflectional morphology	373
	4 Word formation	377
	5 Syntax	378
	References	383
<b>PART 8 NORTHEASTERN INDIA</b>		<b>385</b>
<b>23</b>	<b>Garō</b> <i>Robbins Burling</i>	387
	1 Introduction	387
	2 Phonology	388
	3 Verbs	390
	4 Noun phrases	391
	5 Adverbs and reduplication	397
	6 Complex sentences	397
	7 Language contact and language maintenance	399
	References	400
<b>24</b>	<b>Jinghpo</b> <i>Dai Qingxia and Lon Diehl</i>	401
	1 Phonology	401
	2 Grammar	403

<b>25 Hakha Lai</b>	<i>David A. Peterson</i>	409
1	Introduction	409
2	Phonology	409
3	Inflectional morphology	410
4	Syntax	419
	References	424
	Further reading	426
<b>26 Meithei</b>	<i>Shobhana L. Chelliah</i>	427
1	Location and classification	427
2	Phonetics and phonology	427
3	Morphology	429
4	Major lexical categories	431
5	Evidentiality	432
6	Syntax	433
	References	437
	Further reading	438
<b>27 Tshangla</b>	<i>Erik Andvik</i>	439
1	Introduction	439
2	Phonology	439
3	Morphophonemics	440
4	Noun phrase	441
5	Syntactic roles	443
6	Case marking	444
7	Tense/aspect	446
8	Negation	446
9	Mirativity	447
10	Copular clauses	448
11	Modality and sentence-types	448
12	Relative clauses	450
13	Complementation	451
14	Adverbial clauses	452
15	Clause chains	452
16	Concatenation	453
17	Grammaticalized non-final verbs	453
	References	455
	Further reading	455
<b>28 Tani languages</b>	<i>Jackson T.-S. Sun</i>	456
1	Background	456
2	Phonology	456
3	Morphology	458
4	Syntax	462
	Additional abbreviations	465
	References	466

<b>PART 9</b>	<b>GYALRONG LANGUAGES</b>	<b>467</b>
<b>29</b>	<b>Cogtse Gyarong</b> <i>Yasuhiko Nagano</i>	469
1	Introduction	469
2	Outline of phonology	470
3	Morphology and morphosyntax	471
	References	489
<b>30</b>	<b>Caodeng rGyalrong</b> <i>Jackson T.-S. Sun</i>	490
1	Background	490
2	Phonology	490
3	Morphology	491
4	Syntax	497
	Additional abbreviations	502
	References	502
<b>PART 10</b>	<b>KIRANTI LANGUAGES</b>	<b>503</b>
<b>31</b>	<b>Kiranti languages: an overview</b> <i>Karen H. Ebert</i>	505
1	Introduction	505
2	Typological characteristics	505
3	Genetic and areal groupings	516
	Additional abbreviations	517
	References	517
<b>32</b>	<b>Hayu</b> <i>Boyd Michailovsky</i>	518
1	Introduction	518
2	Phonology	518
3	Morphophonology	519
4	Verb morphology	520
5	Non-verbal morphology and word classes	523
6	Syntax	524
7	Information structure and discourse particles	530
8	The noun phrase	530
	Additional abbreviations	531
	References	532
<b>33</b>	<b>Camling</b> <i>Karen H. Ebert</i>	533
1	Introduction	533
2	Phonology	533
3	Nominals	534
4	The verb	537
5	Syntax	542
	Additional abbreviations	545
	References	545

<b>34</b>	<b>Belhare</b> <i>Balthasar Bickel</i>	546
1	Introduction	546
2	Phonology	546
3	Inflectional morphology	548
4	Derivational morphology and compounding	559
5	Syntax	561
	Additional abbreviations	569
	References	569
<b>PART 11 QIANGIC LANGUAGES</b>		<b>571</b>
<b>35</b>	<b>Qiang</b> <i>Randy J. LaPolla</i>	573
1	The phonological system	573
2	The noun phrase	575
3	Nominal relational morphology	577
4	The verb complex	579
5	Adverbials	583
6	The clause	584
	References	586
<b>36</b>	<b>Prinmi</b> <i>Picus Sizhi Ding</i>	588
1	Introduction	588
2	Phonology	588
3	Morphology	591
4	Syntax	594
	References	601
<b>37</b>	<b>Tangut</b> <i>Gong Hwang-Cherng</i>	602
1	Introduction	602
2	Phonology	602
3	Morphology	606
4	Word formation	610
5	Syntax	612
	References	619
<b>PART 12 KAREN LANGUAGES</b>		<b>621</b>
<b>38</b>	<b>Eastern Kayah Li</b> <i>David Solnit</i>	623
1	Introduction	623
2	Phonology	624
3	Word formation	625
4	Syntax	626
	References	631

<b>39 Pwo Karen</b> <i>Atsuhiko Kato</i>	632
1 Introduction	632
2 Sounds	633
3 Parts of speech	634
4 Morphology	637
5 Syntax	639
Additional abbreviations	647
References	648
 <b>PART 13 OTHER LANGUAGES</b>	 <b>649</b>
<b>40 Yunnan Bai</b> <i>Grace Wiersma</i>	651
1 Introduction	651
2 Phonology	654
3 Word classes	661
4 Word formation processes	666
5 Syntax	669
References	672
 <b>41 Dulong</b> <i>Randy J. LaPolla</i>	674
1 Introduction	674
2 Phonology	674
3 Morphosyntax	675
References	682
 <b>42 Kham</b> <i>David E. Watters</i>	683
1 Introduction	683
2 Phonology	685
3 Inflectional morphology	688
4 Adjectives and adjectivals	699
5 Noun phrase syntax	700
6 Nominalizations	701
7 Clause chains and switch reference	702
8 Evidentials	703
References	703
 <b>43 Lepcha</b> <i>Heleen Plaisier</i>	705
1 Introduction	705
2 Phonology	705
3 Nominal morphology	707
4 Verbal morphology	712
5 Text	715
Additional abbreviations	715
References	716
 <i>Index</i>	717

---

---

# ILLUSTRATIONS

---

---

## FIGURES

9.1	Structure of the Cantonese syllable	148
10.1	Compound characters	159
10.2	Characters which include the 'horse' radical	159
10.3	The historical development of Chinese characters	161
10.4	The development of 'form-sound' characters	162
10.5	The realization of a phonetic in different dialects	164
11.1	Relationships among the Bodo-Konyak-Jinghpaw languages	175
11.2	The Tani Languages (after J.T.-S. Sun 1993)	181
11.3	Classifications of the eastern border languages	183
11.4	Relationships among the languages of the eastern border	184
12.1	Vowels of major syllables	198
12.2	Tonal contrasts	199
12.3	Initial contrasts	199
12.4	Standard Burmese pronouns	201
12.5	Main 'case' marking postpositions (colloquial)	204
13.1	The seven tones of Black Lahu	209
13.2	Black Lahu consonants and vowels	210
15.1	Relations among the Southern Yipho languages	236
16.1	The Bodic branch	255
19.1	Relations among the Tamangic languages	315
19.2	The tense-aspect system	324
29.1	Independent personal pronouns	472
29.2	Morphosyntax of prefixes	475
29.3	Layers of prefixes	475
29.4	Semantic layers of prefixes	475
36.1	The structure of Prinmi syllables	589
36.2	A layered analysis of the noun phrase	595
36.3	The structure of the simplex sentence	595
40.1	Jianchuan Bai syllable structure	655
40.2	Jianchuan Bai tones	658
42.1	The dialects of Kham	684
42.2	A partial listing of village dialects belonging to the Takale regional dialect	684
42.3	The syllable in Kham	686
42.4	The contrastive 'four-tone' system in Kham	687
42.5	Person hierarchy relevant to Kham case marking	690

**MAPS**

3.1	Order of adjective and noun	46
3.2	Order of adjective and noun in Asia	46
3.3	Worldwide distribution of two orders of relative clause and noun among OV languages	52
11.1	Northeastern India	170
11.2	Bodo-Konyak-Jinghpaw languages	176
11.3	Languages of Arunachal Pradesh and neighbouring Assam	179
11.4	Languages of the eastern border	185
31.1	Map of Kiranti languages	506

**TABLES**

5.1	Diagnostic list for identifying Sinitic languages (I)	74
5.2	Diagnostic list for identifying Sinitic languages (II). Lesser known languages	75
5.3	The division of Chinese dialects: Northern, Central, and Southern	76
5.4	The register split in the <i>píng</i> , <i>shàng</i> , <i>qù</i> , and <i>rù</i> tones	77
5.5	Lexical differences between Mandarin and Wu dialects	79
5.6	Double correspondences of lower register tones in Min	81
5.7	A diagnostic ensemble of Min lexical items	82
7.1	Comparisons of the attributes in Chinese	126
7.2	Comparisons of the attributes of Modern Chinese dialects	127
7.3	Tone sandhi in the Changzhi dialect	129
7.4	Comparisons of Mandarin dialects by region	129
8.1	Consonants	132
8.2	Syllable structures	135
8.3	Tone categories	136
8.4	Lexical tone melodies	136
9.1	Cantonese initial consonants	147
9.2	Cantonese vowel phonemes (between //) and allophones (between [ ])	147
9.3	Fifty-six Cantonese rhymes	148
9.4	Lexical items contrasting seven tones on open syllable <i>ji</i>	149
9.5	Register stratification in Cantonese syntax	151
21.1	Affirmative paradigm of <i>nar</i> -‘eat’	360
23.1	Syllable initial consonants and consonant clusters	389
23.2	Codas: syllable final consonants and clusters	389
23.3	Pronouns	392
24.1	Syllable onsets	402
24.2	Lexical contrasts among onsets	402
24.3	Rhymes	402
24.4	Lexical contrasts among rhymes	404
24.5	Personal pronouns	405
24.6	A sample subset of SFWs in the declarative mood	408
25.1	Segmental inventory	410
25.2	Pronominals	411
25.3	Verbal agreement markers	414
25.4	Jussive agreement and negation	415
25.5	Primary relativization strategies	422



26.1	Chart of consonant phonemes	428
26.2	Chart of vowel phonemes	428
27.1	Tshangla consonants	440
27.2	Tshangla vowels	440
27.3	Tshangla verb classes	440
27.4	Exceptional nasal/liquid-final verb roots	441
27.5	Exceptional vowel-final verb roots	441
27.6	Demonstratives	442
27.7	Affirmative verb paradigm	446
27.8	Negative verb paradigm	447
30.1	Caodeng possessive prefixes	494
30.2	Caodeng person-number markers on intransitive verbs	495
30.3	Basic Caodeng orientation prefixes	496
31.1	Pronouns and possessive prefixes	507
31.2	Case markers	509
31.3	Thulung person-number markers	510
31.4	Bantawa person-number markers	510
32.1	Verb morphology	521
33.1	Consonant phonemes	534
33.2	Personal pronouns, possessives and demonstratives	535
33.3	Camling person-number markers NW dialects	538
34.1	Belhare phonemes	546
34.2	Pronouns and demonstratives	548
34.3	Case markers and their use	549
34.4	Possessive prefixes	550
34.5	Finite verb inflection	551
34.6	Affirmative (upper form in each cell) and negative (lower form) non-past paradigm ( <i>luma</i> ‘to tell’ and <i>khatma</i> ‘to go’)	552
35.1	The Qiang consonants	573
35.2	The Qiang personal pronouns	576
35.3	The Qiang person marking suffixes for intransitive verbs	581
35.4	The Qiang non-actor person marking suffixes	581
36.1	Monophthongs of Niuwozi Prinmi	589
36.2	Consonants of Niuwozi Prinmi	590
36.3	Suprasegmental categories in Niuwozi Prinmi	590
36.4	The directional prefixes	592
36.5	The conjugation of the copula and verbs	592
36.6	Attachment patterns of the interrogative and negative clitics	594
37.1	Reconstruction of Tangut finals	603
37.2	Inventory of consonants	605
40.1	Jianchuan Bai syllabary	656
40.2	Representation of Jianchuan tones after orthographic revision	659
40.3	Distribution of Chinese tone groups in literary and colloquial Chinese strata of Jianchuan Bai	660
40.4	Sentence particles in a Jianchuan Bai text	665
40.5	Case inflection of personal pronouns in Jianchuan Bai	666
40.6	Incipient semantic marking in Jianchuan Bai	667
41.1	The Dulong consonants	675
41.2	The Dulong personal pronouns	675

---

42.1	Vowel phonemes in Takale Kham	685
42.2	Consonant phonemes in Takale Kham	686
42.3	Number and case marking	689
42.4	Position classes for nominal affixes	691
42.5	Position classes for complex locatives	693
42.6	Spilt ergative agreement patterns based on a hierarchy of person	694
42.7	Agreement forms in five TB languages	695
42.8	Position classes in the Takale Kham verb	699
43.1	Consonants	706
43.2	Vowels	706
43.3	Initial consonant clusters	707
43.4	Personal pronouns	711
43.5	Demonstrative and interrogative pronouns	712

---

# PREFACE

---

In deciding what to put into this volume on the Sino-Tibetan languages, Randy LaPolla and I have had been guided by several considerations. First, the volume provides 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. For Sino-Tibetan as a whole, there are three overviews, one on both genetic classification and language contact (Graham Thurgood), one on Sino-Tibetan syntax and morphology (Randy LaPolla), and one on Sino-Tibetan word order typology (Matthew Dryer).

For the Sinitic side of the family, there are general articles on the phonology (Jerry Norman), the syntax (Anne Yue), and the writing system (Mark Hansell). Supplementing this overview are articles on Mandarin dialects (Dah-an Ho), Shanghainese (Eric Zee and Leijiong Xu), Cantonese (Bob Bauer and Stephen Matthews), and Late Zhou Chinese (Derek Herforth).

For the Tibeto-Burman side of the family, in addition to the Thurgood, LaPolla, and Dryer articles, there are several overviews of geographical or genetic subgroups, specifically, Burling discusses Northeast India, Karen Ebert discusses Kiranti, and Jackson Sun discusses Tani (Abor-Miri-Dafla). Complementing the overviews are detailed descriptions of more than thirty individual Tibeto-Burman languages, in all cases by one of the world's leading experts. 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.

All the languages mentioned in the individual chapters are found with tentative subgroupings in the Thurgood overview. In that article, the ninety-three languages which Dryer based his typological survey on are followed by •, thus, Lahu• indicates that Lahu was one of his source languages. Alternative names for languages are listed in various ways: where the alternative name is an older designation, it is put into square brackets; where it is simply an alternative, it is listed on the same line.

Tibeto-Burman subgrouping is an area rife with controversy, but despite this, Randy and I were, with quibbling here and there, able to agree. While there were, of course, differences of opinion, the degree of consensus was striking. Where the subgrouping was clear, we agreed on it; where it was unclear, we agreed that it was unclear.

Various people have helped make the volume possible: Jonathan Price, who has been supportive and helpful; the individual contributors, who have put up with Randy and me harassing them and have helped out whenever we asked; and others, such as Frank Li, who checked aspects of several papers for me. The frequent suggestions and helpful comments from contributors and non-contributors alike have added to the value of this unique collection.

Finally, Randy LaPolla has left his distinct imprint on the volume. His command of Chinese, of Tibeto-Burman, and of syntax have influenced every chapter in the volume. Working with him has been a pleasure and, in the most positive of senses, a learning experience.

---

# ABBREVIATIONS

---

BLSn	<i>Proceedings of the nth Annual Meeting of the Berkeley Linguistic Society</i>
BSOAS	<i>Bulletin of the School for Oriental and African Studies</i>
BIHP	<i>Bulletin of the Institute of History and Philology</i>
LTBA	<i>Linguistics of the Tibeto-Burman Area</i>
A	'actor' of transitive clause
ABL	ablative
ABS	absolutive
ACC	accusative
ADV	adverb
AFF	affix
AGT	agentive
AMG	location in or among
ANIM	animate
ANT	antipassive
ANTC	anticausative
APPLIC	applicative
APPR	approximative
APUD	location near
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
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
ICSTLL	International Conference on Sino-Tibetan Languages and Linguistics
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	level-pitch type
LINK	linker
LOC	locative
M	masculine/male

---

MAL	malefactive
MID	middle/middle voice
MIR	mirative (just discovered)
NF	non first person actor
N-PAST	non-past affirmative
NR	near
N-SG	nonsingular
NEG	negation
NGR	nasalizing grade
NOM	nominative
NOMZR	nominalizer
NRPAST	near past
OBJ	object case
OBLQ	oblique (non-subject) case
OCHBP	body part
OPT	optative
P	'undergoer' of transitive clause
PART	participle
PASS	passive
PAST	past
pauc	paucal
PERF	perfect
PFV	perfective
pl	plural
PN	pronoun
POSI	positional
POSS	possessive
POT	potential
PREF	prefix
PREP	preposition
PRES	present
PRSNTV	presentative
PROG	progressive
PROH	prohibitive
PURP	purposive
Q	interrogative/question
R	co-referential
R/M	reflexive/middle
REC	recent
RECIP	reciprocal
REDUP	reduplicated
REFL	reflexive
REL	relative
REM	remote
REQU	request marker
RES	resultative
REV	reverential second person
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
1	1st person
2	2nd person
3	3rd person

**PART 1**

---

# **OVERVIEW CHAPTERS**

---





# A SUBGROUPING OF THE SINO-TIBETAN LANGUAGES: THE INTERACTION BETWEEN LANGUAGE CONTACT, CHANGE, AND INHERITANCE

*Graham Thurgood*

## 1 INTRODUCTION

The Sino-Tibetan speaking people are associated in the literature with the Neolithic Yang-shao culture which originated in Yellow River valley in the central plains of northern China. Eventually this group of Sino-Tibetan speakers split into Sinitic (essentially Chinese) and Tibeto-Burman. No records of the original language exists, of course, and what was once a single language has been split into a family of languages. What was originally a single language has developed into a vast, diverse family of languages under the pressures of natural change, intermingled with frequent and often intimate contact with speakers of other languages. Much of this historical overview is concerned with describing, analysing, and cataloguing the results of natural change and language contact.

### 1.1 Recovering the linguistic history

The common origin – the parent language from which the Modern Sino-Tibetan languages have descended – is at least partially recoverable through historical reconstruction – as is much of the language contact – from the patterns of shared innovations found among the modern languages. Reconstruction gives us a picture of what the older language looked like. And, those linguistic elements that are present in the modern languages but not reconstructable often tell us about the nature of earlier language contact. From the interaction of the two, much about the linguistic and non-historic history can be recovered.

### 1.2 The genetic subgroups

The subgroups listed here are the well-established lower-level subgroups, that is, those for which there is significant supporting evidence. Most of these are substantiated by the existence of preliminary reconstructions of the group, establishing them beyond reasonable doubt.

Listed in this volume are the best established of these subgroupings. Burling has written an extensive chapter on languages found in Northeast India, Ebert has written on the Kiranti languages, Sun has written on the Tani languages, and LaPolla has discussed some higher level groupings. The proposed higher-level relationships are based on shared innovations. The lack of this type of evidence has sometimes forced writers to make subgrouping guesses on the basis of geography, a practice not followed here.

### 1.3 The languages, their names, and their locations

Bradley's 1997 *Tibeto-Burman Languages and Classification* is the most recent and the most extensive up-to-date survey of Tibeto-Burman languages. 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. This work is made even more valuable by its extensive and remarkably up-to-date bibliography.

The language names listed in this chapter are, with the occasional exception, limited to those mentioned somewhere in this volume. For a more exhaustive listing, see Matisoff's invaluable compendium *Languages and Dialects of Tibeto-Burman* (1996), which contains 118 double-columned pages of names of Tibeto-Burman languages and dialects, supplemented by the information in Bradley (1997). Even the language names present their own problems. Many are referred to in a myriad of ways, with different names used by different outsiders and still another name used by the people themselves. Sometimes certain groups even 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. No serious attempt to straighten out all the names has been undertaken here, but see the relevant discussions in Burling, in Sun, and in Ebert; for still further information, see Matisoff (1996).

### 1.4 Contact with other languages

The discussion of language contact – the contact with other languages and the influences of this contact – is presented against the background of what we know about the genetic subgroups. It is usually against the background of our knowledge of the genetic subgroups that we can see the linguistic effects of contact with other peoples. And, in the case of the Sino-Tibetan languages, a major role has been played by contact with peoples, often followed by incorporation, and sometimes to complete assimilation (see LaPolla 2001, for an extended discussion of migration and contact).

Dryer's chapter sketches the geographical dimensions of Sino-Tibetan word order variation, correlating the variation with contact with other language families of the region. This overview also notes broad areal contact, with the majority of the chapters in this volume on specific languages also including comments on the local effects of language contact.

### 1.5 The methodology of subgrouping

It is well known among practitioners that establishing genetic subgroups among related languages is often far more difficult than the simple discovery of a genetic relationship.

Subgrouping is complicated by the fact that many similarities between closely-related languages are the product not of common inheritance but of what Sapir called 'drift';

that is, the common starting point provided by a common origin often conspires with universal tendencies to provide parallel but historically quite independent development among genetically-related languages (Thurgood 1985: 378)

Further, if long-term contact can produce striking typological similarities even between unrelated languages – the typological similarities between Vietnamese and Chinese being a case in point, the influence of such long-term contact on genetically-related languages is even more marked. Thus, an understanding of genetic subgrouping requires a historical base that includes knowledge of the relevant contact relationships. Finally, nothing is made simpler by the fact that the database is enormous, uneven, and complicated.

The standard assumptions about subgrouping, not always followed in the Tibeto-Burman literature, are worth 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 re-examined carefully. Within the linguistic data, only shared innovations constitute evidence of an earlier period of common development; the corollary is that 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.

Thus, for example, a language is included in Lolo-Burmese if the membership is supported by reconstructions, particularly correspondences with the three Lolo-Burmese open tones. As with the contact information some of the information on subgrouping is from the chapters in this volume. For subgroupings discussed elsewhere in this volume, unless there something new is being added to the analysis, the languages and groups are simply listed; for more detail there is a reference to the relevant chapter. Burling treats the languages of North-east India in considerable detail. Sun's chapter on Tani, like Burling's chapter, thoroughly discusses the relevant dimensions of Tani; thus, little is given below beyond the subgroup name and a listing of language names. Similarly, Noonan's chapter discusses Tamangic and Norman discusses Chinese dialects, both beyond what is found here. Again, with Ebert's Kiranti chapter, the discussion of contact and subgrouping is thorough and complete so only the subgroup name and the language names are given here, except for a brief comment on shared innovations.

The subgroups presented here vary in the quality of the evidence used to support their existence. At one end are subgroups like Lolo-Burmese, which are backed by full reconstructions of the lexicon and with descriptions of the corresponding sound changes, or, LaPolla's Rung, which is backed by shared morphological innovations. Thus, subgroups based on full reconstructions or on cognate morphological innovations are more solidly established than those based only on lexical considerations. At the other end are clusters, that is, subgroups supported only by shared lexical correspondences; such clusters are less well-established because of the difficulties of ascertaining whether the shared material is the result of contact, something that is often difficult to establish without a thorough reconstruction.

The groupings labelled here as subgroups are backed by shared innovations – not just lexical similarities, usually attested in the literature but sometimes simply in my own files. These groupings are conservative – with speculation, labelled as such, but nonetheless thrown in here and there. At a higher level, some of these groups undoubtedly are subgrouped together; however, without the evidence of shared innovations, it is not possible to determine precisely. Lower level groupings only supported by lexical similarities are also present in the chapter but these are best thought of as clusters, rather than subgroups.

## 2 SINO-TIBETAN

The Sino-Tibetan family consists of two major subgroups, Chinese and Tibeto-Burman. By and large the distinction between the two is unambiguous and widely accepted, despite a dwindling number of older scholars who still see the connection as not yet proven. Questions remain, however, about the status of Bai and, to a lesser degree, Tangut, although this overview provides a tentative classification of both: Tangut is tentatively classified in this work as Tibeto-Burman, more specifically, as Qiangic, while Bai remains an unclassified Sino-Tibetan language.

## 3 CHINESE

Chinese is really a family of related dialect groups, with the diversity within the different dialect groups often marked enough to make even dialects within the same dialect group mutually unintelligible. The grouping below into six dialect families matches Norman's practice, aside from the placement of Hakka. Alternate groupings are common in the literature.

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

Each of these dialect groups is further subdivided into subgroups, but the larger groupings above 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 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, in this work it typically refers to the national language with more specific references used to refer to the dialects or to a specific language.

### 3.1 Contact and Chinese

Throughout its history, Chinese has been under the influence of language contact. In fact, even the basic SVO (Subject-Verb-Object) word order of Modern Chinese seems to be the product of contact. 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, describing clear, typological correlates of the contact.

The earliest recoverable Chinese vocabulary already contains interesting borrowings reflecting various forms of contact. Words for 'honey' and 'goose,' for instance, probably reflect early Indo-European contact of a limited nature not found elsewhere in Sino-Tibetan. Norman (1988) notes borrowings from Mon-Khmer including 'tiger,' 'ivory,' 'crossbow,' and the word for 'river,' found in the early name for the Yangtze (< \*kronɡ), undoubtedly indicating not just contact but also a significant early Austroasiatic (Mon-Khmer) presence in that area. Various calendric terms also look to have been borrowed from Austroasiatic, indicating a significant cultural influence on the early Chinese.

In addition to language groups that must have been absorbed without a trace after they came into contact with the Chinese, there are other absorbed groups that are reflected in the historical records. In the north, Chinese came 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 northern China, or at least parts of it, at various times. These kingdoms included speakers of Altaic (Tungusic, Mongolian, and so on) and Tibeto-Burman languages, and others and began the process of the Sinicization of these speakers at the same time 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 thirty 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. Taken together, these languages have also had a significant impact on word order (see Dryer).

In fact, 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 endangered and 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 TIBETO-BURMAN

The contact situation of early Tibeto-Burman speakers is far from fully understood; however, it is clear that Mon-Khmer speakers were once prevalent in many of the areas that Tibeto-Burman speakers now occupy and that in many cases the Mon-Khmer speakers were

absorbed into Tibeto-Burman speaking communities but not without some effects on the language structures.

#### 4.1 Lolo-Burmese branch [Burmese-Lolo]

##### 4.1.1 *Burmish*

Burmese• (see Wheatley, Chapter 12, this volume) (Yangon dialect; other dialects:  
Arakanese (dialect), Tavoyan)  
Zaiwa (Atsi)  
Lachi  
Bola  
Maru•  
Achang•

##### 4.1.2 *Loloish (= Yi)*

Northern Loloish: (Bradley 1997)

Nusu•  
Nasu  
Nosu  
Nisu

Central Loloish:

Sani, Ahi  
Lahu• (see Matisoff, Chapter 13, this volume)  
Lisu• (dialects: Black Lisu, Flowery Lisu, Southern Lisu) (see Bradley, Chapter 14, this volume)  
Lipho, Lolopho (closely related to Lisu)  
Micha, Lamo (closely related to the Lisu, Lipho, Lolopho complex)  
Jinuo, Jino•

Southern Loloish:

Akha•, Hani•, Khatu, Pijo, Haoni (see Hansson, Chapter 15, this volume)  
Bisoid: Phunoi, Bisu, Pyen  
Mpi

Other: (precise relationship unknown)

Gokhy  
Zauzou

Lolo-Burmese (= Burmese-Lolo) is a well-defined subgroup of Tibeto-Burman. Lexical reconstructions of Lolo-Burmese exist and these correspondence patterns are usually more than sufficient to establish beyond doubt whether a language is Lolo-Burmese or not. The most striking phonological innovation is proto-Lolo-Burmese tone 3. The core Lolo-Burmese languages all share reflexes of this particular tone and, of course, of the other two open tones reconstructed for this subgroup. It has been claimed that Lolo-Burmese itself is part of a slightly larger subgroup, not characterized by the presence of these tonal reflexes, but, if so, the bases for a wider subgroup remains to be established.

As the diagrammatic representation indicates, Lolo-Burmese is divided into Burmish languages, including Burmese, the language of Burma (= Myanmar), and Loloish languages,

including some of the best-described Tibeto-Burman languages, including Lahu (with Matisoff's exquisite dictionary and extensive grammar), Lisu, Akha. The basic division between Burmish and Loloish is again well-defined, with clear patterns of shared innovations.

The Lolo-Burmese languages have always been subject to contact pressures. Early Mon-Khmer influence on Proto-Lolo-Burmese, the proto-language, 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).

The Loloish languages are divided into Northern, Central, and Southern, on what initially might appear to be an exclusively geographical basis; however, here, the subgrouping finds substantiation in patterns of shared innovations. All the Loloish languages show Mon-Khmer influence and some Chinese influence, with those Loloish languages now located in Thailand containing numerous Thai loanwords.

## 4.2 The Bodic branch [= Tibetan]

The Tibetan division, a subset of Shafer's Bodic division (1955), contains at least three unquestionable members: the numerous Tibetan dialects; the Tamang-Gurung-Thakali-Manang languages, a subgroup called by various names even within this volume; and the small group of languages that includes Takpa. Speculatively, Tshangla has been added to this group.

LaPolla notes in Chapter 2, this volume, that the Bodic languages constitute a subgroup based on the shared innovation 'an \*-s ablative/ergative suffix on nouns'.

### 4.2.1 *Tibetan dialects*

Classical Tibetan (see DeLancey, Chapter 16, this volume)

Central (or U-Tsang)

Modern Lhasa• (see DeLancey, Chapter 17, this volume)

Shigatse

Sherpa•

Jirel

Lhomi•

Nyamkad•

Jad•

Kagate, etc.

Western Archaic

Balti•

Ladakhi•

Purik, Purki•

Western Innovative

Lahul

Spiti

Tod•



## Southern dialects of Sikkim and Bhutan

Khams  
Amdo  
Sikkimese•  
Dzongkha

4.2.2 *Tamang-Gurung-Thakali-Manang languages*

Tamang• (Mazaudon, Chapter 18, this volume)  
Gurung•  
Chantyal (see Noonan, Chapter 19, this volume)  
Nar Phu (see Noonan, Chapter 20, this volume)  
Manangba  
Thakali (including the dialects of Marpha, Thini, and Syang)  
Tangbe

4.2.3 *Takpa (=Northern Monpa) [Dwags]*

Bumtang, Cuona Menba

4.2.4 *Other*

Tshangla (= Cangluo Monpa, Motuo Menba, Central Monpa•)  
(see Andvik; for subgrouping, see Burling)  
Sharchockha

In discussing Tibetan dialects (Chapters 16 and 17), DeLancey describes their distribution as 'in a broad area reaching from northern Pakistan to Qinghai, Sichuan, and Yunnan provinces in China'. Noting that much work remains to be done on Tibetan dialectology, he cites Nishi (1986), who distinguishes the six major groups above (not counting Classical Tibetan, of course). Without exception, all members of the Tibetan group listed above share the innovation of a second person pronoun \**khyot* 'thou' and, although the reconstruction leaves much to be desired, the innovation of a third person singular, roughly \**kho*.

The Tamang-Gurung-Thakali-Manang group of languages are referred to in various ways by those who study them (see Mazaudon (Chapter 18); see Noonan (Chapters 19 and 20)). The subgroup, substantiated by reconstruction, has been worked on extensively and is reasonably well described. (Further information on the internal subgrouping is found in individual chapters in this volume.) Within the subgroup determining the internal relationships is complicated both by contact between various dialects and by the existence of a considerable Nepali influence on some of these languages.

As for Takpa, it shares an innovated second person pronoun that resembles one found in TGMT languages, apparently descending from \**khyot* 'thou'. Several writers have suggested that Takpa belongs with the Tibetan group (see DeLancey (Chapters 16 and 17); see Andvik (Chapter 27)). Thus, tentatively, Takpa has been put here.

Other: the membership of this group in Bodic is tenuous. Unlike the other languages in Bodic, Tshangla (Cangluo Monpa, Central Monpa, and Motuo Monpa (Sun, Hong Kai *et al.* 1980)) has uniquely innovated a first person pronoun, while failing to share the innovation of the second person pronoun. Nonetheless, it has been tentatively included, as DeLancey and Andvik see it as closely related to Tibetan. Additional support for these proposals is needed.

### 4.3 The ‘Sal’ languages: [Bodo-Konyak-Jinghpaw] (see Burling, Chapter 11)

Burling has established this subgrouping on the basis of shared lexical innovations, particularly the innovation of *sal* ‘sun’ (Burling 1983). Elsewhere in this volume, Burling discusses the membership and internal subgrouping of this branch in detail, so here it will suffice to simply list his three branches:

#### 4.3.1 *Bodo-Koch* [*Bodo, Bodo-Garo, Barish*]

- Bodo: Bodo•
- Boro
- Deori [Deuri•], Chutia
- Dimasa•
- Tiwa [Lalung]
- Mech
- Kachari•
- Hill Kachari
- Kokborok• [Tripuri]
- Garó: Garó•
- Koch: Koch: Tintinkiya Koch, Wa’ngang Koch and Pani Koch
- Rabha
- A’tong
- Ruga

#### 4.3.2 *Konyak group* (cf. also French (1983))

- Tangsa•
- Yogli [= Jugli•]
- Lungcang• (Lungchang, Longchang)
- Nocte• [Namsangia]
- Wancho [Banpara]
- Konyak [Tableng]
- Phom (Chingmengnu, Tamlu)
- Chang•
- Khiamngan
- Moshang
- Wakching

#### 4.3.3 *Jinghpaw*• [= *Jinghpo*], *Singhpo* possibly including *Luish*:

- Sak
- Kadu
- Andro<sup>†</sup>
- Sengmai<sup>†</sup>

Jinghpaw is the term for the speakers; Kachin is the designation for the ethnic group. Jinghpaw has been subgrouped in various places. Burling places it into his *sal* languages, that is, Bodo-Konyak-Jinghpaw, because, like the other languages in this group, it has a *sal* reflex for the word for ‘sun’.

The frequent connection of Jinghpaw with Lolo-Burmese in part reflects ethnic considerations, not linguistic evidence. Often, the Jinghpaw speakers are bilingual in Zaiwa (Atsi), a Burmish language, in many cases living interspersed with Zaiwa speakers. Other Jinghpaw speakers also know Maru, Lachi, or Bola, all Burmish languages. The intensity of contact has, of course, led both to borrowing and to convergence; however, neither of these constitutes evidence for genetic subgrouping.

#### 4.4 The Kuki-Chin-Naga branch

It is often assumed that the Naga constitute a genetic subgroup [Kuki-Chin-Naga], but as Burling notes, there is still no clear evidence that these subgroups comprise a larger supergroup, that is, it is not yet clear that the Ao, the Angami-Pochuri, the Zeme, the Tangkhul, and the Mizo-Kuki-Chin groups as well as Meitei and Karbi constitute a subgroup. (For further data on these groups, see Burling.)

##### 4.4.1 *The Ao group (see Burling, Chapter 11)*

Ao-Chungli  
Ao-Mongsen  
Yacham-Tengsa  
  
Sangtam [Tukumi]  
Yimchungrii [Yachumi]  
Lotha [Lhota]

Like Shafer and the LSI, Burling groups the last three languages with Ao.

##### 4.4.2 *The Angami-Pochuri group (see Burling, Chapter 11)*

Angami  
Chokri  
Kheza  
Mao [Sopvoma]  
  
Pochuri [southern Sangtam, eastern Rengma]  
Ntenyi  
Meluri [Anyo], Pochuri  
Simi/Sema  
Rengma proper (see Burling, Chapter 11, this volume)

##### 4.4.3 *The Zeme group (see Burling, Chapter 11)*

Zeme [Empeo, Kachcha●]  
Mzieme  
Liangmai [Kwoireng]  
  
Nruanghmei [Rongmei, Kabui]  
Puiron  
Khoirao  
Maram

#### 4.4.4 *The Tangkhul group (see Burling, Chapter 11)*

Tangkhul  
Maring

#### 4.4.5 *The Mizo-Kuki-Chin group (Kukish)*

##### 4.4.5.1 Chin

Northern Chin  
Tiddim Chin•  
Siyin•  
Thado•  
Ralte  
Paite  
Gangte  
Pawi  
Chiru  
Simte

##### Central Chin

Mizo• [Lushai]  
Lai•/Laizo, Hakha Lai (see Peterson, Chapter 25)  
Zahao  
Bawm• [Banjogi]  
Mara [Lakher•/Maram]

##### Southern Chin:

Shö  
Daai [Nitu]  
Khyang

##### 4.4.5.2 Kuki languages (or dialects) [Old Kuki]

Kom  
Aimol  
Bete  
Hallam  
Langrong  
Anal  
Chothe/Chote/Chawte  
Hmar

Peterson notes that the Chin languages are subgrouped into northern, central, and southern branches on the bases of phonological and grammatical innovations. The Mizo-Kuki-Chin subgroup as a whole reconstructs an innovative pronoun system that consists of *\*kai* 'I', *\*nang* 'thou', and *\*a-mi* 'third person' and a prefixal subject-verb agreement system that consists of *\*ka-* 'first', *\*na-* 'second', and *\*a-* 'third'. Thus, this subgroup shares innovations in the *\*kai* 'I', the *\*a-mi* 'third person', and the agreement markers. Note that these innovations are quite independent of those found in the Rung group.

#### 4.4.6 Karbi/Arleng [Mikir]

The inclusion of Karbi/Arleng [Mikir] within the so-called Naga group reflects three beliefs. The Naga group itself is not yet established as a genetic subgroup and so far contains various not yet interconnected subgroups, hence, the use of the term ‘so-called’ to describe it. Second, there is no positive subgrouping evidence to place Karbi with any particular subgroup; that is, we know what it is not; for example, it is neither Lolo-Burmese nor Tibetan, but we do not know what it is. And, third, given its geographical location and the fact that various earlier scholars have placed it in this group, it has been put here until someone figures out what to do with it. In short, following Burling, I leave it unclassified.

#### 4.4.7 Meithei/Meitei [formerly Manipuri, among various older names] (see Chelliah; for subgrouping, see Burling, Chapter 11)

The placing of Meithei/Meitei with the so-called Naga group has been done on essentially the same bases as the inclusion of Karbi (immediately above): that is, it has been put here until some clear evidence about its placement is found. The language has some lexical similarities to Kuki and to Tangkhul, but interpretation of these similarities is difficult, given the long-time contact between these groups. 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.

### 4.5 The Rung branch

By Rung is meant LaPolla’s large group consisting of the rGyalrong languages, the Dulong languages (T’rung, Rawang), the Kiranti languages, the West Himalayan languages (Kinauri, Almora), Kham, and, less probably, Magar, and Chepang. As LaPolla notes, this Rung is an expansion and, I might add, an improvement of two earlier subgrouping proposals in the literature, one by me, the other by Ebert. LaPolla (Chapter 2) bases the group on, among other considerations, the sharing of a complex, innovated person marking system, as LaPolla notes, as well as a \*-*si* reflexive/middle marking verb suffix except in proto-rGyalrong (see LaPolla).

#### 4.5.1 The rGyalrong• (WT rGyalrong) subgroup (see Sun, Chapter 30; Nagano, Chapter 29)

rGyalrong proper (to use J. Sun’s term)

eastern:	Situ
northeastern:	Chabao
northwestern:	Sidaba

Lavrung (= Guanyinqiao)

Ergong (Sun Hongkai), (= Daofu, = Horpa-Shangzhai), [Horpa]

This representation of rGyalrongic contains five languages, three subgrouped together into rGyalrong proper, plus two others: Lavrung, also known in the literature as Guanyinqiao, is now called Lavrung by Huang Bufan and by J. Sun. Sun Hongkai’s Ergong, known in the older literature as Horpa, has been termed Daofu by Huang Bufan and Horpa-Shangzhai by J. Sun, but all are the same language.

There is clear evidence of contact influences on rGyalrong, sometimes leading to a misunderstanding of its subgrouping. In particular, it is sometimes subgrouped with Tibetan, but as Nagano quite correctly points out in his contribution (and Sun Hongkai *et al.*), the Tibetan-like vocabulary is most likely borrowed. Nagano notes that the more basic vocabulary is typically general Tibeto-Burman, while the Tibetan-like forms are cultural items. Thus, the Tibetan resemblances are undoubtedly the product of contact.

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, Nagano, and LaPolla. More work remains to be done in this area.

In addition to Tibetan influences, there is a widespread Chinese influence, with rGyalrong being used at home and Chinese in public places, as Nagano notes.

#### 4.5.2 *Dulong and related languages (see LaPolla, Chapter 41) [Nungish]*

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

Anong  
Rawang•

The literature often suggests a genetic connection with Jinghpaw, but neither LaPolla nor I have been able to find substantiating evidence.

#### 4.5.3 *Kiranti (see Ebert, Chapter 31 for an overview)*

In her overview Ebert is tentative about the very existence of a Kiranti subgroup, referring to these languages as the Kiranti cluster, a group consisting of some thirty-two languages. Like Winter, she suggests it is premature to term these a subgroup. My own evidence for designating them as a subgroup is also limited: shared innovations in the pronouns and the related verbal agreement system and the beginnings of some historical reconstruction showing regular correspondences (Michailovsky 1994). Here, as elsewhere in the family, it is often difficult and at times impossible to differentiate similarities inherited from a common proto-language from those due to a long period of mutual contact and interaction.

Athpare•  
Bahing  
Bantawa  
Belhare (see Bickel, Chapter 34)  
Camling• (see Ebert, Chapter 33)  
Chilling  
Dumi•  
Hayu• (see Michailovsky, Chapter 32)  
Khaling•  
Kiranti  
Kulung  
Limbu•  
Nachereng  
Sunwar  
Thulung•  
Umbule

Yakkha  
Yamphu

Contact influence has played a significant role in shaping these languages. Most of these are endangered – by contact; all are heavily influenced by Nepali. The languages themselves seem to be the result of relatively recent migration into the area and share features with languages outside the geographical area. Within Tibeto-Burman Ebert notes similarities between the Southeastern Kiranti languages and the Naga and Chin languages, namely, in the prefixed person markers and the participle formation with *ka-*, similarities whose origins are not yet clear. She also has noted here and elsewhere similarities in rGyalrong to the inverse marking of Camling and Bantawa; these similarities might very well be inherited. Non-Tibeto-Burman influences also seem apparent but are restricted to subsets of Kiranti: Ebert suggests interesting 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 (see Ebert, Chapter 31).

#### 4.5.4 *West Himalayan group*

Connected to the Kiranti group is a second group of languages sometimes termed the Kinauri–Almora subgroup and sometimes referred to as West Himalayan. This subgroup is characterized by shared innovations in the pronouns and in the innovated agreement markers:

1sg	1sg AGR	2sg	2sg AGR	3sg	3sg AGR
*gai	*-ga	*gan	*-na	*du	*u

The reflexes of these two innovations nicely divide Kinauri–Almora into two subgroups.

#### 4.5.5 *Kinauri*

Chamba Lahuli  
Kanashi  
Rangloi, Gondla, Tinan•  
Pattani• [Manchati]  
Bunan, Ga(h)ri•  
Kinauri•, Kanauri  
Marchcha•  
Jahri•

#### 4.5.6 *Almora*

Byangsi•  
Chaudangsi•  
Darmiya•  
Rangkas

#### 4.5.7 *The Kham•, Magar•, and Chepang• languages*

Kham, Magar, and Chepang have been subgrouped in various places by various people for various reasons. The innovations among the agreement markers tentatively suggest placement within Rung.

Kham• (see Watters, Chapter 42)  
 Magar•  
 Chepeng

The inclusion of Magar and Chepeng is more problematic than the inclusion of Kham. Note that their placement together here is not the claim that these three represent any sort of subgroup.

#### 4.5.8 *The Qiangic languages*

The Qiangic group is a complex and still not fully-understood subgroup of Tibeto-Burman languages. The subgrouping provided here is tentative and is based in large part on the work of various authors, including four contributors to this volume, but also in a small part on cognate patterns in my own files.

The consensus that Qiang proper and Prinmi belong in the same subgroup is easily and fully substantiated by careful examination of cognate sets. The inclusion of Muya in this group (suggested in Sun Hongkai 1988: 67) is also strongly supported by the cognate sets, but Sun's rationale for the inclusion of the now extinct Tangut is not, as yet, clear to me.

Qiang• (see LaPolla, Chapter 35)  
 Northern Qiang  
 Southern Qiang

Prinmi (Primi, Pumi•)  
 Northern Prinmi  
 Southern Prinmi (Niuwozi Prinmi; see Ding, Chapter 36)

Muya  
 Tangut (see Gong, Chapter 37)

A second cluster of languages is also termed Qiangic. For these the definitive subgrouping evidence remains to be presented, but an inspection of the vocabulary suggests these are also part of this subgroup. Sun Hongkai (1998: 62) further breaks them up into the groups given here.

Zhaba  
 Queyu  
 Guiqiong  
 Ersu  
 Shixing  
 Namuzi

Questions about the wider affiliations of Qiangic also exist. These languages are often assumed to subgroup with the rGyalrong languages, but as LaPolla quite accurately points out, the rGyalrong languages subgroup more strongly with the rest of the Rung group, than with Qiang. On the other, an examination of cognate sets suggests a special relationship, but one that is not yet clear.

The influence of Chinese on Qiangic is obvious.

#### 4.6 **The Karenic branch (see Solnit, Chapter 38; see Kato, Chapter 39)**

Northern  
 Pa-O



## Central

Kayah Li• (Karenni), Brè (= Bwe•), Yintale, Palayachi, Mopwa, and so on

## Southern

Pwo Karen and Sgaw•

## Other

Padaung

Karenic is a well-defined subgroup, supported by historical reconstructions. Its internal subgroups similarly find support in the patterns of shared innovations. The tentative homeland for the subgroup, Solnit argues, is the central Karen area since it is the area of greatest linguistic diversity. Following general linguistic practice, Solnit concludes that, because the area of greatest diversity is in western Kayah State (Karenni) and the adjoining area of Karen State, this is most probably the homeland from which the languages have spread.

Its distinctive SVO (Subject-Verb-Object) word order sets Karenic off from the rest of Tibeto-Burman and seems to have led earlier scholars such as Benedict to subgroup Karenic outside of Tibeto-Burman proper. More modern scholarship, such as Solnit (this volume), views the word order as due to the contact with non-Tibeto-Burman SVO languages. Solnit points out that the Karenic languages are found at the southeastern edge of Tibeto-Burman territory (along with Lolo-Burmese and Tujia speakers). As Solnit observes, this peripheral location has brought Karen into close contact with members of two characteristically SVO language families, Tai-Kadai and Mon-Khmer. Solnit reports loanwords in Karen from Mon-Khmer (specifically, the Palaungic and Monic branches) and from Tai-Kadai languages. The Mon-Khmer influence is pervasive culturally and linguistically, including considerable bilingualism. Solnit characterizes the Tai influence as strikingly evident in the proto-Karen consonant system and its interaction with the evolution of tones. Thus, despite the opinions sometimes expressed in older scholarship, one might side with Solnit in assuming that Karen's SVO word order is the result of contact and that Karenic is simply a subgroup of Tibeto-Burman.

## 4.7 Other small subgroups

### 4.7.1 Tani (see Sun, Chapter 30; see Burling, Chapter 23) [Abor-Miri-Dafla; Mirish]

Various scholars have argued that the Tani languages constitute a distinct Tibeto-Burman subgroup, with Sun's recent reconstruction of proto-Tani (1993) placing its status as a subgroup on a firm footing. For a fuller discussion of the internal relationships within Tani, see Sun.

Mising• (Plains Miri): Padam-Mising, Mirish, [formerly Abor-Miri-Dafla]

Adi [formerly Abor]: Padam Adi, Padam-Minyong Adi

Tagin

Hill Miri

Bokar

Damu

Apatani•/Apa Tani

Milang•

Na Bengni

Bengni

Gallong•

Nishi• [formerly Dafla], Nishing

Nyisu

The nature of Tani's subgrouping relationships with neighbouring Tibeto-Burman languages is not evident yet. As Sun notes, while the Tani languages bear similarities to Taraan and Idu languages on the east and Hruso, Dhammai, and Levai or Bangru on the west, these similarities may be the result, not so much of a close subgrouping relationship, as due to long-term contact. Certainly contact has had a significant effect on Tani languages. Tani languages are also under contact pressure from Indo-Aryan and even from English, citing the importation of loanwords with the otherwise absent codas *-s*, *-l*, and *-m*.

#### 4.7.2 *Sherdukpen-Bugun-Sulung-Lishpa (see Burling, Chapter 11)*

Sherdukpen  
Bugun•/Khoa  
Lishpa  
Sulung

Sulung is aberrant enough to have led some scholars to question whether or not it was even Tibeto-Burman. Its position here reflects nothing more than my imitation of Burling.

#### 4.7.3 *Hrusish (see Burling, Chapter 11)*

Hruso [Aka]  
Dhammai/Miji  
Bangru/Levai

#### 4.7.4 *The Idu-Digarú group (see Burling)*

Idu [Chulikatta]  
Idu Mishmi•  
Digarú [Digaro Mishmi•/Taraan]

#### 4.7.5 *The Miju/Kaman group (see Burling)*

Miju/Kaman

### 4.8 Unsubgrouped languages

The number of languages that could be characterized as unsubgrouped is enormous. Rather than put them all in this section, they have been placed tentatively in various subgroups. However, this tentative placement should not be mistaken for understanding. The work on Sino-Tibetan contact and subgrouping that remains to be done far surpasses the work that has been done.

The three languages immediately below are unsubgrouped Tibeto-Burman languages, that is, while they are Tibeto-Burman, it is unclear where they belong in Tibeto-Burman. The status of Bai is still less determinant; it is unclear whether it is Sinitic or Tibeto-Burman.

Pyu  
Naxi•  
Tujia  
Bai

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.

The position of Naxi (various alternative spellings) is still unclear despite much speculation.

Tujia is on the southeastern edge of Tibeto-Burman and, as such, has come under considerable contact pressure. Its subgrouping remains a mystery.

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 three thousand 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. Dryer (Chapter 3) for instance, notes that some of the morphological oddities shared by Chinese and Bai are likely to be contact induced, just as might some of the shared lexical (and morphological items) detailed by Norman in Chapter 5.

## REFERENCES

- 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.
- Burling, Robbins (1983) 'The Sal languages', *Linguistics of the Tibeto-Burman Area* 7.2: 1–32.
- French, Walter T. (1983) 'Northern Naga: a Tibeto-Burman mesolanguage', unpublished PhD dissertation, City University of New York.
- 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.
- 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, 225–54.
- Michailovsky, Boyd (1994) 'Manner vs place of articulation in the Kiranti initial stops', Hajime Kitamura, Tatsuo Nishida, and Yasuhiko Nagano (eds) *Current Issues in Sino-Tibetan Linguistics*, The Organizing Committee, the 26th International Conference on Sino-Tibetan Languages and Linguistics, Osaka 1994, 766–72.
- Matisoff, James A., Baron, Stephen P. and Lowe, John (1996) *Languages and dialects of Tibeto-Burman*, Center for Southeast Asia Studies: University of California, Berkeley.
- Nishi, Yoshio (1986) 'A classification of Tibetan dialects [in Japanese]', *National Minority Studies Museum Research Report* 11.4: 837–901.
- Norman, Jerry (1988) *Chinese*, Cambridge Language Surveys: Cambridge University Press.
- Shafer, Robert (1955) 'Classification of the Sino-Tibetan languages', *Word* 2: 94–111.
- Sun, Hongkai, Lu, Shaozun, Zhang, Jichuan and Ouyang, Jueya (1980) *The Languages of Monba, Luoba and Deng* (in Chinese), Beijing: Nationalities Publishing House.
- Sun, Hongkai (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*, Kyoto: Shokado, 61–73.
- 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.* Pacific Linguistics Series C. No. 87. Canberra: Australian National University, 376–400.

Wiersma, Grace (1990) 'A study of the Bai (Minjia) language along historical lines', unpublished PhD dissertation, University of California, Berkeley.

# 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 eight thousand years ago (Thurgood 1994), the proto-language was monosyllabic. Matisoff (1991a: 490) reconstructs the syllable canon as \*(P) (P) C<sub>i</sub> (G) v (: (C<sub>f</sub>) (s)).<sup>1</sup> 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 *tǎrüng* ‘T’rung/Dulong’, the vocalic element of the *tǎ-* prefix is very slight).

There was no relational morphology (LaPolla 1990, 1992a,b, 1993b, 1994b, 1995a,b), but there was derivational morphology in the form of prefixes, suffixes, and voicing alternations of the initial consonants (Wolfenden 1929; Benedict 1972; Pulleyblank 1962–3, 1972, 1973a,b, 1977–8, 1991, 2000; Bodman 1980; Mei 1980, 1988, 1989; LaPolla 1994c; Gong 2000). Following are examples of several types of derivational morphology.<sup>2</sup>

### 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). Mei (1989) argues all of these functions are manifestations of a more general directive function. For example, Old Chinese (OC) *\*mjang* (亡) ‘be gone’ : *\*smangs* (喪) ‘to lose’; OC *\*mək* (墨) ‘ink’ : *\*smək* (黑) ‘black’; Written Tibetan (WT) *smag* ‘dark’; OC *\*C-rjəʂ* (吏) ‘clerk’, ‘minor official’ : *\*srjəʂ?* (使) ‘to cause (someone to be an emissary)’, ‘to send’; *\*tju?* (帚) ‘broom’ : *\*stu?* (掃) ‘to sweep’; *\*ljek* (易) ‘to exchange’ *\*sljeks* (賜) ‘to give’, ‘gift’; WT: *grib* ‘shade’, ‘shadow’ : *sgrib-pa* ‘to shade, to darken’; *gril* ‘a roll’ : *sgril-ba* ‘to roll together’, ‘to form into a roll’.

---

1 P=prefix, C<sub>i</sub>=initial consonant, G=glide, :=vowel length, C<sub>f</sub>=final consonant, s=suffixal \*-s; parentheses mark that the item does not appear in all syllables.

2 This list is probably not exhaustive, and the necessarily brief discussion glosses over many controversies and details. As is always the case in attempting to find Sino-Tibetan correspondences, the lack of a single standard for the reconstruction of Old Chinese (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. 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 Sargart 1998 for discussion of some of the recent changes).

## 1.2 \*Voicing alternation

In both OC and Tibeto-Burman (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. Benedict (1972: 124) discussed this for TB, but argued that in Chinese no consistent pattern of morphological alternation could be recognized. Most scholars now would see the Chinese 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 \**ǎ-* (a non-syllabic pharyngeal glide) which voiced the initial of the original transitive roots. Mei (1989) includes this prefix in a paradigm with the \**s-* directive prefix and the \**-s* direction of action changing suffix (below).<sup>3</sup> Both Pulleyblank and Mei base the idea for the prefix mainly on the Written Tibetan *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 Chinese forms, and uses \**f-* for the form of the prefix,<sup>4</sup> e.g. \**kens* (見) 'see': \**fikens* (> \**gens*) (現) 'appear/be visible'. While this analysis is attractive from a systemic point of view, Benedict (*ibid*) 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'; '*khriil-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 Chinese. We also find the voicing alternation in TB languages independent of prefixation, e.g. \**kh(r)jok* (曲) 'bend', 'bent': \**fikh(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 Chinese 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 *fe-phe* 'tear (clothes)': *de-pe* 'be torn'; TB \**pe* ~ \**be*; Benedict 1972: 59) (:: OC \**phajs* (破) 'to break': \**pajʔ* (跛) 'lame'). Other examples: OC \**prats* (敗) 'to defeat': \**fprats* (\**brats*) (敗) 'to be defeated'; \**krujs* (壞) 'to destroy', 'ruin': \**fikrujs* (\**grujs*) (壞) 'to be ruined'; \**irjangʔ* (長) 'grow tall', 'increase'; 'elder': \**ftrjang* (\**drjang*) (長) 'long'; Bodo *ben* 'to be straight': *phen* 'to make

3 In a slightly earlier paper, Mei (1988) argues for reconstructing a voiced initial rather than a prefix.

4 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 now reconstructs with \**ml-* clusters (e.g. Gong: \**N-ljək* (食), Baxter \**mljək* (see Matisoff 1995, footnote 1; originally \**Ljək* in Baxter 1992).

straight' (TB *\*bleŋ* ~ *\*pleŋ*; OC *\*bren* (平) 'level'?). It seems there were intransitivizing (and nominalizing) prefixes in PTB and possibly STC, 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), and possibly *\*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)' (there is also a separate intransitivizing/stativizing *ə-* prefix in T'rung as well: *tāl* 'roll (vt)' > *ətāl* 'roll (vi)' (LaPolla 1995c; see also LaPolla 2000a)). These are independent of the voicing alternation.

Quite a few scholars have assumed that the *\*s-* causative prefix was responsible for all of the voicing distinctions now found in the family (e.g. Dai 2001), but, while this is true for some languages, particularly within Lolo-Burmese, the examples given in the discussion of this and the previous section show that the two are separate phenomena.

### 1.3 *\*-t* suffix<sup>5</sup>

The *\*-t* suffix most often has the function of transitivizing an intransitive verb, as in WT *fibye-ba* 'open', 'separate' (vi) : *fibyed-pa* 'open', 'separate' (vt), Rawang *ŋū* 'weep' : *ŋut* 'mourn', 'cry for someone (vt)', but in some cases seems to nominalize intransitive verbs, as in WT *ŋu-mo* 'weep' : *ŋud-mo* 'a sob' (see also Benedict 1991), and in still other cases seems not to have had any affect 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). In Chinese 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').

### 1.4 *\*-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), e.g. Proto-Tibeto-Burman (PTB) *\*r-mi* 'person' : OC *\*mjn* (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 *\*ŋja* (語) 'speak' ~ *\*ŋjan* (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) and an *\*-n*-initial pronoun, possibly *\*njəjʔ* (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 *\*ʔa*

5 Although not often explicitly mentioned, except by Jin (1998a, b), 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.

(烏) : \*ʔan (安) both ‘interrogative pronoun’ (‘where’), and possibly \*ʔak (惡) ‘interrogative pronoun’.<sup>6</sup>

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, 1957). 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’. 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 (言者) from \*tjə ʔja (之於) ‘3rd person object pronoun’ + adposition ‘in’, ‘at’, ‘to’; \*njəʔ (耳) from \*njə ʔləʔ (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’.

### 1.5 \*-s suffix

The \*-s suffix generally had a nominalizing function (Pulleyblank 1973b; Mei 1980, 1989), 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 (1980) suggests these two functions derive from two different homophonous suffixes, which he equates with the Tibetan nominalizing and ablative suffixes respectively. In Modern Chinese this suffix is now reflected in the ‘departing’ tone. 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’: \*djus (授) ‘to give’.

### 1.6 \*-j suffix

Matisoff (1989, 1995) discusses etyma that show palatal-final and non-palatal-final variants, and posits three different sources for variants with morphological differences: PST \*s-way ɤ \*s-yay ‘go’; ‘motion away’, for transitive motion/motion away from the deictic centre or emergent quality in stative verbs; PST \*ya (ɤ \*za ɤ \*tsa ɤ \*dza) ‘child’, ‘son’ for a diminutive or affective sense; and PST \*way ɤ \*ray 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).

6 The usual reading of this last character when used as an interrogative pronoun is \*ʔa, but it is written using a character that is in other contexts pronounced \*ʔak. If it is the same pronunciation as the one otherwise written (烏), it seems odd to use a character that normally is read with a stop final.



### 1.7 \*-ʔ suffix

Chinese seems to have had a glottal stop suffix which developed into the rising tone category, e.g. \*trjang (張) ‘to make long’, ‘stretch’: \*trjangʔ (長) ‘grow tall’, ‘increase’; ‘elder’; \*wək (或) ‘someone’: \*wəkʔ (有) ‘there is’; \*kak (各) ‘each’: \*k(rf)akʔ (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. rGyalrong, T’rung), but it is not clear that there is any relation between the forms in these languages and those in Chinese.

### 1.8 \*-k suffix

There may have been a \*-k suffix as well, as we find a large number of lexical items in both TB and Chinese 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’; \*djuj (誰) ‘who’ : \*djuk (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: \*wəkʔ (有) ‘there is’ : \*wək (或) ‘someone’; \*k(rf)akʔ (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 \*nja (如) ‘like’ : \*njak (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).<sup>7</sup> 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 \*-∅ ~ \*-k variants.

It has long been known that within Sino-Tibetan we must deal with word families rather than isolated words (Karlgren 1933, 1956; Wolfenden 1936, 1937, 1939). 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): \*kat (割) ‘to injure’, ‘to harm’ (vt) : \*fikat (\*gat) (害) ‘to suffer harm or injury’ (vi) : \*fikats (\*gats) (害) ‘harm’, ‘injury’ (n); \*trjang (張) ‘to make long’, ‘stretch’ (vt) : \*trjangʔ (長) ‘grow tall’, ‘increase’; ‘elder’ (intransitive active verb) : \*firjang (\*drjang) (長) ‘long’ (stative verb).

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.

<sup>7</sup> Our answer to this question will affect our understanding of certain word families. For example, Pulleyblank (1991: 30) suggests that \*k(rf)akʔ (舉) ‘all’; ‘lift’ as an allofam \*kjat (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.

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*<sup>33</sup> ‘duck’ > *am*<sup>35</sup> ‘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* (𠵼) (夫/父/甫) was used as an extra-syllabic suffix for creating agentive nouns, just as in TB (e.g. *\*din p(r)ja* (田夫) ‘farmer’), and this may be the cognate of PTB *\*-pa*.

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 *\*lja* (𠵼), Newar *lā*, Burmese *lā*, Meithei *la* (Matisoff 1995: 73–4). As mentioned above, though, the Chinese form has been said to be a fusion form, from *\*le hwa* (也呼) (Norman 1988: 95).

Unmarked clausal negation in PST took the form of a preverbal particle *\*ma-*. For PTB we can also reconstruct a prohibitive (negative imperative) particle *\*ta-* (see ex. (1) below for a Lahu example), but this is not found in Chinese. Chinese instead had two negative imperative particles *\*mja* (毋), which was homophonous with the unmarked negator but written with a different character, and *\*mjət* (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)).<sup>8</sup>

Most languages in the family have not grammaticalized grammatical relations, but many have grammaticalized semantic marking.<sup>9</sup> For detailed arguments against the existence of syntactic functions in particular Sino-Tibetan 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.

Classifiers were not part of PST, but evolved individually in quite a few of the languages in the family (see for example Xu 1987, 1990; Dai 1994, 1997a,b for prosodic reasons for some languages developing classifiers, and some not). There was no definite marking in PST, and only a few languages in TB, such as Qiang (see LaPolla forthcoming) 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

8 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.

9 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, Chapter 6, for the concept of pivot and its relation to grammatical relations. See Dixon 1995, Chapter 2, and also Hale and Watters 1973 on semantic marking vs grammatical marking.

classifiers, both in TB and among Chinese dialects, 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 *yar<sup>55</sup> ga<sup>33</sup> che<sup>55</sup>* [one CL vehicle] ‘one car’, *ga<sup>33</sup> che<sup>55</sup>* (roughly) ‘the car’. This feature is an areal feature of part of Southeast Asia (Baron 1973).

In terms of word order, 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). Originally there were no nominalizers or relative markers in relative clauses, but various languages have developed one or the other since that time. In cases where the relative clause is nominalized, this construction then is also a subcase of the N-N modifier-modified construction. 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. At present, the Sinitic languages (Chinese dialects), 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. 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 Tibeto-Burman languages other than Bai and Karen, we assume that it was Bai, Karen, and Chinese that changed rather than all the other Tibeto-Burman languages. As argued in LaPolla 1993a, these three languages show a remarkable similarity in the particular patterns they developed. In Old Chinese, 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 be due to an earlier preverbal focus position. In Karen and Bai, we have the same situation as in Old Chinese 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 Old Chinese: 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 Sincized 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 follow 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 Old Chinese, 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 Tibeto-Burman languages mainly in terms of the position of the NP representing the under-goer 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; also Dryer, this volume).

## 2 CHINESE

In Chinese, there was a gradual loss of productivity of the derivational morphology sometime around the formation of the characters (roughly 3500 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). There has been no grammaticalization of grammatical relations; the basic structure of the clause is topic-comment rather than subject-predicate (Chao 1955, 1959, 1968; Lyu 1979; LaPolla 1990, 1993b). Information structure is the chief determinant of word order in Chinese (LaPolla 1995d). 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 Anne Yue-Hashimoto (1993, this volume), Huang Borong (1996), and Chappell (2001), 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. 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 Tibeto-Burman languages.

## 3 TIBETO-BURMAN

After the split-up of Sino-Tibetan into Sinitic and Tibeto-Burman, 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 2000b, 2001), I divide TB into the following groups:<sup>10</sup>

The Bodic group: Tibetan 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 a conjunct-disjunct system in some languages (see for example Hale and Watters 1973: 207ff. on Jirel, Newar, and Sherpa; Hale 1980 on Newar), and a lack of a bound pronominal person marking system or reflexive/middle marking.

---

10 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 2000b 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.

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). Languages of this group characteristically have a set of 5–10 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 1981; Huang Butan 1991). 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). 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 (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 Bufan 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 may simply be areal influence.

The Rung group: rGyalrong, T’rung (Dulong), Rawang, Kiranti, Kham, and Western Himalayan (Byangsi, Darma, Chaudangsi, Kinauri), languages that (except for rGyalrong) migrated down along the eastern edge of the Himalayas and then across Burma and into Northern India and Nepal.<sup>11</sup> These languages have clearly cognate complex person marking systems, and all but rGyalrong have a \*-si reflexive/middle marking suffix on the verb:

	<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 rGyalrong splits off from the others, as it does not share the innovation of the reflexive/middle marking. This accords well with the migrations assumed. A second branching of Western Himalayan off from Rawang and

11 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 2000b). See also Grierson (1909, vol. III), for particular characteristics shared between the Western and Eastern Himalayan pronominalized languages not shared by the Tibetan 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. The name Rung was coined by Thurgood (1984), but used for a somewhat larger grouping of languages. That original grouping is no longer recognized, and so I have used the name for this grouping.

Kiranti is assumed, as the latter two share the innovation of a non-first-person-actor marking prefix.<sup>12</sup>

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', *\*naŋ* '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. Within Lolo-Burmese, some Loloish languages have been greatly influenced by contact with Tai, while Burmese has been more influenced by Mon (see Bradley 1980). 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).

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.

The Tani group: Sun (1993a, Chapter 5; 1993b) argues convincingly that the Tani group (formerly called Mirish or Abor-Miri-Dafla, including the languages of the Adi, Nishi, Bengni, Apatani, and Mishing 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.<sup>13</sup>

12 Thurgood (1984) discusses the fact that rGyalrung, T' rung, and Kham all have a preverbal 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.

13 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.

The Bodo-Konyak-Jinghpaw group: This group, which includes the Luish, Bodo-Garo, Koch, Konyak, and Jinghpaw/Singpho languages (Burling, this volume), 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. Within the larger grouping, only Nocte and Jinghpaw have person marking systems, and they do not appear to be cognate.

Aside from these genetic groupings, and a split in prosodic type between a Southeastern iambic stress area and a Northern trochaic stress area, 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 Sino-Tibetan 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).

### 3.2 Person marking

Several branches of TB have independently innovated person marking, possibly due to areal influence (LaPolla 1992a, 1994b, in press). 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 and Naga, have recently developed such systems in some dialects (see for example the Delugong dialect of Sgaw Karen discussed in Dai *et al.* 1991).<sup>14</sup> The pattern discussed most often is that of the Rung group, because of its wide geographic distribution. This pattern has been associated with the Tangut pattern, but it is not clear whether the Rung pattern developed out of the simpler Tangut pattern (1sg  $*\text{-}\eta\alpha^2$ , 2sg  $*\text{-}n\alpha^2$  (the same forms as for the free pronouns), first and second person plural  $n\eta^2$ ).<sup>15</sup> Attempts to associate the Rung pattern with other patterns in the family and reconstruct it to PTB have been unsuccessful (see LaPolla 1992a 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 1983a: 72)

---

14 Independent innovation of bound pronominal paradigms in various languages in a family is not unique to TB, but occurred also in Amerind (Mithun 1991) and Australian (Dixon 1980) languages.

15 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.

a difference between an animate (Idu  $t^{55}$ ) and an inanimate (Idu  $kh a^{55}$ ) 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,  $d\sigma^{31}$  for liquids,  $de^{31}$  for general animates,  $k\chi^{31}$  for existence within a group, and one existential verb,  $s\sigma^{55}$ , which is used only in the poetic language (Li and Wang 1986: 54). In Queyu there are seven existential verbs (Wang 1991: 61):  $t\dot{f}i^{55}$ , for animals;  $t\dot{c}y^{13}$ , for location in a vessel or certain area;  $bo^{31}$ , for non-movable objects;  $\dot{c}i^{13}$ , for movable objects;  $lo^{13}$ , for an object mixed up in another object;  $ru^{13}$ , for abstract objects; and  $t\dot{f}e^{13}$ , 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:  $nji^{51}$ , which seems to mark the existence or long-term location of animate beings and has the causative form  $nji^{51}$ ;  $lu\eta^{55}$ , for short-term location of animate beings and has the causative form  $lu\eta^{55}$ ;  $vo^{55}$ , for possession by a person;  $t\dot{f}o\dot{p}^{31}$ , for inanimates;  $po^{51}$ , for containment within a vessel; and  $to\eta^{51}$ , for roads and footprints. Other languages that have this feature are Jinghpaw, Apatani, Tamang, Naxi, Nusu, Pumi, rGyalrong, Qiang 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  $t\dot{c}y^{13}$ , Zaiwa  $po^{51}$ ), ‘possession by a person’ (Hani  $t\dot{f}e^{13}$ , Zaiwa  $vo^{55}$ ), the forms used in these languages are clearly not cognate.

### 3.4 Causative marking

The PST  $*s$ -causative prefix and voicing alternations are no longer productive in most TB languages, and so languages throughout the family (more than eighty 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 thà? 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  $xu^{55}$ , Xiandao Achang  $san^{31}$ , Bola  $n_2^{55}$ , and Leqi/Langsu  $l_2\dot{p}^{55}$ . In each case we have the independent grammaticalization of a free verb into a post-verbal causative marker.

### 3.5 Benefactive marking

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 ( $-t\dot{f}a^{33}$ ), Tamang ( $pín$ ), Tsangla ( $bi$ ), Camling ( $bi$ ), Belhare ( $-per$ ), and Lahu ( $p\dot{r}$ ; for third person benefactives; Matisoff 1991b). As can be seen from these examples, the verb used in this construction is often the PST verb  $*biy$  ‘give’, but the constructions themselves were independently innovated. A fully morphological etymologically opaque benefactive such as is found in Rawang, where the suffix  $-\bar{a}$  has



an applicative benefactive function (LaPolla 2000a; e.g. *rí-ā-ò-ē* (carry-BEN-TR.NPAST-NPAST) ‘(He) is carrying (something) for him’) is rare.

### 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; LaPolla 1994b). 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; 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) 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’ marking (LaPolla 1992b), 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). 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 (which is unidirectional) corresponding to degree of grammaticalization.

From the survey of ‘object’ marking in Tibeto-Burman, it was found that out of 126 languages surveyed, twenty-two languages had no nominal object marking, twenty 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 eighty-four 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), *thàʔ* marks a patient argument in (1a), but a recipient argument in (1b).

- |        |                 |             |           |             |    |                      |            |           |             |             |
|--------|-----------------|-------------|-----------|-------------|----|----------------------|------------|-----------|-------------|-------------|
| (1) a. | <i>ɲà</i>       | <i>thàʔ</i> | <i>tâ</i> | <i>dʂʔ.</i> | b. | <i>liʔ</i>           | <i>chi</i> | <i>ɲà</i> | <i>thàʔ</i> | <i>pîʔ.</i> |
|        | 1sg             | OBJ         | NEG.IMP   | hit         |    | book                 | that       | 1sg       | OBJ         | give        |
|        | ‘Don’t hit me.’ |             |           |             |    | ‘Give me that book.’ |            |           |             |             |

I refer to this type of marking as ‘anti-ergative’ marking, as the crucial function of this type of marking is to mark an animate argument that might otherwise be interpreted as an actor as 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.<sup>16</sup> 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 *thàʔ*, 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-ergative markers (e.g. Lahu (*thàʔ*), Akha (*áŋ*)), or differ in terms of having anti-ergative marking or not (e.g. Akha, which has anti-ergative 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-ergative 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 Tibeto-Burman. Both are semantically based, but one (ergative and patient marking) is based on what semantic role a referent has, and the other (anti-ergative 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) 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 Tibetan (DeLancey 1986; Woodbury 1986; Sun 1993c; Hongladarom 1993). Other languages which have evidential marking are Qiang (LaPolla to appear), Newar (Hargreaves 1983), 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

---

16 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, *dechticaetiative*. 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 Tibeto-Burman 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.

reflexives gets 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). 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, *àŋ sat-çǐu* 'He is hitting himself' and *àŋ et-çǐu* '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 1995c, 2000a; 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), 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 intransitivity (Nagano 1984: 55; Jin *et al.* 1958: 81). Mizo (Changte 1993; Lorrain and Savidge 1898) has a verb prefix *-in* 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.

## REFERENCES

- 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', ICSTLL 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, '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.
- 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: 149–53.
- Bhat, D.N.S. (1991) *Grammatical Relations: The Evidence Against Their Necessity and Universality*, London: Routledge.
- Blansitt, E.L., Jr. (1984) 'Deicticative 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: 259–67.
- Bradley, David (1997) 'Tibeto-Burman languages and classification', in D. Bradley (ed.) *Tibeto-Burman languages of the Himalayas*, Canberra: Australian National University.
- Burling, Robbins (1983) 'The Sal languages', *LTBA* 7.2: 1–32.
- Caughey, Ross (1978) 'Participant rank and verbal cross reference in Chepang', in Grimes (ed.) 1980, 163–78.

- Chafe, Wallace and Nichols, Johanna (eds) (1986) *Evidentiality: The Linguistic Coding of epistemology*, Norwood, NJ: Ablex.
- Chang, Kun and Chang, Betty Shefts (1975) 'Gyarong historical phonology', *BIHP* 46: 391–524.
- 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: 229–43.
- Chappell, Hillary (2001) 'Language contact and areal diffusion in Sinitic languages', in R.M.W. Dixon and A.Y. Aikhenvald (eds) *Areal Diffusion and Genetic Inheritance: Case Studies in Language Change*, Oxford: Oxford University Press.
- 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, 237–49.
- 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, 250–9.
- Chao, Yuen Ren (1968) *A Grammar of Spoken Chinese*, Berkeley and Los Angeles: University of California Press.
- Chhange, Lalnunthangi (1993) 'A grammar of Mizo', unpublished PhD dissertation, University of Oregon.
- 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, 112–21.
- Comrie, Bernard (1978) 'Ergativity', in W.P. Lehmann (ed.) *Syntactic Typology: Studies in the Phenomenology of Language*, Austin: University of Texas Press, 329–94.
- Dai, Qingxia (1994) 'Mian-Zang yu geti liangci yanjiu (a study on numeral classifiers in Tibeto-Burman)', in *Mian-Zang yu xin lun (Recent Contributions to Tibeto-Burman Studies)*, Beijing: Zhongyang Minzu Xueyuan Chubanshe, 166–81.
- Dai, Qingxia (1997a) 'A study on count-noun classifiers in Tibeto-Burman languages', *Studies on Yi-Burmese languages*, ed. by the Editorial Committee of the International Yi-Burmese Conference, Chengdu: Sichuan Nationalities Publishing House, 355–73.
- 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, Xu, Xijian (1992) *Jingpoyu yufa (Jinghpaw grammar)*, Beijing: Zhongyang Minzu Xueyuan Chubanshe.
- Das, Sarat Chandra (1902) *A Tibetan-English Dictionary*, Calcutta: The 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 (1985) 'The analysis-synthesis-lexis cycle in Tibeto-Burman: a case study in motivated change', in J. Haiman (ed.) *Iconicity in Syntax* (Typological studies in language 6), Amsterdam and Philadelphia: John Benjamins Pub. Co, 367–89.
- DeLancey, Scott (1986) 'Evidentiality and volitionality in Tibetan', in Chafe and Nichols 1986, 203–13.
- DeLancey, Scott (1987) 'The Sino-Tibetan languages', in B. Comrie (ed.) *The World's Major Languages*, New York: Oxford University Press, 799–810.
- 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.
- Downer, G.B. (1959) 'Derivation by tone change in Classical Chinese', *BSOAS* 22.2: 258–90.
- Driem, George van (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, 157–67.
- Dryer, Matthew S. (1986) 'Primary objects, secondary objects, and antidative', *Language* 62: 808–45.
- Ebert, Karen H. (1987) 'Grammatical marking of speech act participants in Tibeto-Burman', *Journal of Pragmatics* 11.4: 473–82.
- 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 Thurgood, Matisoff, and Bradley (eds) (1985) 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., 227–55.
- 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.
- Graham, Angus (1983) "'Yun" and "yue" as verbs and particles', *Acta Orientalia Hafniensia* 44: 33–71.
- Grierson, George (ed.) (1909) *Linguistic survey of India*, Calcutta: Superintendent of Government Printing.
- Grymes, Joseph E. (ed.) (1980) *Papers on Discourse*, Dallas: Summer Institute of Linguistics.
- 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.
- Hargreaves, David (1983) 'Evidentiality in Newari', unpublished MA thesis, University of Oregon.
- 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 Chubanshe.
- Huang, Bufan (1991) 'Qiangyuzhi (the Qiangic branch)', in M. Xueliang (ed.) *Hanzangyu gailun (A general introduction to Sino-Tibetan languages)*, Beijing: Peking University Press, 208–370.
- Jin, Lixin (1998a) 'Hanzangyu de mingci houzhui \*-n (the noun suffix \*-n in Sino-Tibetan)', *Minzu Yuwen* 1998, 1: 43–8.
- 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.
- 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), *Yuyin Yanjiu* 3: 71–108.

- 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.
- 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.
- 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, January 19–21, 1994.
- 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: 131–73.
- LaPolla, Randy J. (1995a) 'Ergative marking in Tibeto-Burman', in Nishi, Matisoff and Nagano (eds) (1995) 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: 1149–85.
- LaPolla, Randy J. (1995c) 'Reflexive and middle marking in Dulong/Rawang', *ICSTLL* 28, Charlottesville, VA, October 6–9. To appear in *Himalayan Linguistics*, ed. by George van Driem. Berlin: Mouton de Gruyter, (also published in Chinese in *Zhongguo Minzu Yuyan Luncong* (1) (*Collected essays on Chinese minority languages*, 1) ed. by Dai Qingxia *et al.*, 13–34. Central University of Nationalities Press, 1996).
- LaPolla, Randy J. (1995d) 'Pragmatic relations and word order in Chinese', in P. Downing and M. Noonan (eds) *Word order in discourse*, Amsterdam and Philadelphia: Benjamins Pub. Co., 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. (2000a) '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. (2000b) 'Subgrouping in Tibeto-Burman: can an individual-identifying standard be developed? How do we factor in the history of migrations and language contact?', *ICSTLL* 33, Bangkok and Trang, October 2–6, 2000.
- 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, 225–54.
- LaPolla, Randy J. with Huang Chenglong (to appear) *A Grammar of Qiang, with Annotated Texts and Glossary*, Berlin: Mouton de Gruyter.
- LaPolla, Randy J. and Lowe, John B. (eds) (1994) *Bibliography of the Papers from the International Conferences on Sino-Tibetan Languages and Linguistics, I–XXV*. STEDT Monograph Series, No. 1, second edition. Berkeley: Center for East and Southeast Asian Studies.
- Li, Yongsui and Wang, Ersong (1986) *Haniyu jianzhi* (*Brief description of the Hani language*), Beijing: Minzu 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.

- 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 Pub. Co.
- 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, 413–42.
- Matisoff, James A. (1989) 'Palatal suffixes in Sino-Tibetan', *ICSTLL 22*, University of Hawaii, Oct. 6–8, 1989.
- Matisoff, James A. (1990) 'On megalocomparison', *Language* 66.1: 106–20.
- 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 Nishi, Matisoff and Nagano (eds) (1995) 35–91.
- Mei, Tsu-lin (1980) 'Sisheng bieyi zhong de cengci' (Strata in the use of tones to distinguish meaning), *Zhongguo Yuwen* 1980.6: 427–43.
- Mei, Tsu-lin (1988) 'Neibu nigou Hanyu sanli' (Three examples of internal reconstruction in Chinese), *Zhongguo Yuwen* 1988.3: 169–81.
- Mei, Tsu-lin (1989) 'The causative and denominative functions of the \*s- prefix in Old Chinese', *Proceedings of the Second International Conference on Sinology*, Taipei: Academia Sinica.
- Michailovsky, Boyd (1975) 'Notes on the Kiranti verb (East Nepal)', *LTBA* 2.2: 183–218.
- Michailovsky, Boyd (1985) 'Tibeto-Burman dental suffixes: evidence from Limbu (Nepal)', in Thurgood, Matisoff and Bradley (eds) (1985) 334–43.
- 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 Pub. Co.
- Nagano, Yasuhiko (1984) *A historical study of the rGyarong 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.
- Nishi, Yoshio, Matisoff, James A. and Nagano, Yasuhiko (1995) *New Horizons in Tibeto-Burman Morpho-Syntax* (Senri Ethnological Studies 41), Osaka: National Museum of Ethnology.
- Norman, Jerry (1988) *Chinese*, Cambridge: Cambridge University Press.
- Pulleyblank, Edwin G. (1962–3) 'The consonantal system of Old Chinese (2 parts)', *Asia Major* (new series) 9: 58–144, 206–65.
- 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: 113–14.
- Pulleyblank, Edwin G. (1973b) 'Some further evidence regarding Old Chinese \*-s and its time of disappearance', *BSOAS* 36.2: 368–73.
- 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, ILL: Open Court Press.
- Pulleyblank, Edwin G. (2000) 'Morphology in Old Chinese', *JCL* 28.1: 26–51.
- Schuessler, A. (1985) 'The function of *qusheng* in Early Zhou Chinese', in Thurgood, Matisoff and Bradley (eds) 1985: 344–62.

- Shafer, Robert (1955) 'Classification of the Sino-Tibetan languages', *Word* 11.1: 94–111.
- Simon, Walter (1941) 'Certain Tibetan suffixes and their combinations', *Harvard Journal of Asiatic Studies* 5: 388–9.
- Simon, Walter (1942) 'Tibetan *daŋ*, *ciŋ*, *kyin*, *yin*, and *fiam*', *BSOAS* 10: 954–75.
- 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: 51–7.
- Solnit, David (1997) *Eastern Kayah Li: Grammar, Texts, Glossary*, Honolulu: University of Hawai'i Press.
- Sun, Hongkai (1981) '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' in *Minzu Yuwen* Editorial Board (ed.) (on the question of affiliation of the Qiang language), *Minzu Yuwen yanjiu wenji* (collected papers from *Minzu Yuwen*), Xining: Qinghai Minzu Chubanshe, 189–224.
- Sun, Hongkai (1983a) 'Yidu Luobayu gaiyao' (A brief description of the Idu Lhoba Language), *Minzu Yuwen* 1983.6: 63–79.
- Sun, Hongkai (1983b) 'Woguo Zang-Mianyu dongci de rencheng fanchou' (The person category of verbs in the Tibeto-Burman languages of China), *Minzu Yuwen* 1983.2: 17–29.
- 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), Eguchi, *et al.* (ed.) *Languages and history in East Asia: A Festschrift for Tatsuo Nishida on the occasion of his 60th birthday*, Vol. I, 61–73. Kyoto: Shokado.
- 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: 157–81.
- Sun, Tianshin Jackson (1993a) 'A historical-comparative study of the Tani (Mirish) branch', unpublished PhD dissertation, University of California, Berkeley.
- Sun, Tianshin Jackson (1993b) 'The linguistic position of Tani (Mirish) in Tibeto-Burman: a lexical assessment', *LTBA* 16.2: 143–88.
- Sun, Tianshin Jackson. (1993c) 'Evidentials in Amdo Tibetan', *BIHP* 63.4.
- Tayang, Aduk (1983) *A Phrase book in Padam*, Shillong: The Director of Information and Public Relations, Arunachal Pradesh.
- Thurgood, Graham (1984) 'The "Rung" languages: a major new Tibeto-Burman subgroup', *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society*, 338–49. University of California, Berkeley.
- Thurgood, Graham (1985) 'Pronouns, verb agreement systems, and the subgrouping of Tibeto-Burman', Thurgood, Matisoff and Bradley (eds) 1985: 376–400.
- Thurgood, Graham (1986) 'The nature and origins of the Akha evidential system', in Chafe and Nichols (eds) 1986: 214–22.
- Thurgood, Graham (1994) 'Tai-Kadai and Austronesian, the nature of the historical relationship', *Oceanic Linguistics* 33.2, 345–68.
- Thurgood, Graham, Matisoff, James A. and Bradley, David (eds) (1985) *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.
- Tournadre, Nicolas (1991) 'The rhetorical use of the Tibetan ergative', *LTBA* 14.1: 93–108.
- 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, Tianxi (1991) 'Queyu', in Dai *et al.* (1991) 46–63.
- Wang, Li (1980) *Shijing yundu* (The rhymes of the Book of Poetry), Shanghai: Guji Chubanshe.



- 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: 350–59.
- Wheatley, Julian K. (1985) 'The decline of verb-final syntax in the Yi (Lolo) languages of southeastern China', in Thurgood, Matisoff and Bradley (eds) 1985: 401–20.
- 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: 401–16.
- 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: 625–55.
- Wolfenden, Stuart (1939) 'On the restitution of final consonants in certain word types of Burmese' *Acta Orientalia* 17: 153–68.
- Woodbury, Anthony C. (1986) 'Interactions of tense and evidentiality: a study of Sherpa and English', in Chafe and Nichols (eds) 1986: 188–202.
- Xu, Lin and Yansun, Zhao (1984) *Baiyu jianzhi (A brief description of the Bai language)*, Beijing: Nationalities Press.
- Xu, Xijian (1987) 'Classifiers in Jingpo', *Minzu Yuwen* 1987.5.
- Xu, Xijian (1990) 'Jingpoyu liangci de chansheng he fazhan', *Zhongyang Minzu Xueyuan Xuebao* 1990.2.
- 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).

# WORD ORDER IN SINO-TIBETAN LANGUAGES FROM A TYPOLOGICAL AND GEOGRAPHICAL PERSPECTIVE<sup>1</sup>

*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 1, 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 2, I discuss word order in Chinese, identifying some typological 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 (forthcoming). It is based on examination of descriptions of ninety-three 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

---

<sup>1</sup> 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.



And although Bai has some prepositions, it also has some postpositions, like *no*<sup>33</sup> marking the indirect object in (6).

- (6) *ŋa*<sup>55</sup> *si*<sup>31</sup> *nu*<sup>55</sup> *no*<sup>33</sup> *pe*<sup>21</sup>*xo*<sup>55</sup> *ku*<sup>55</sup>  
 1PL give 2SG OBJ flower CLSFR  
 ‘We gave you a flower.’ (Xu and Zhao 1984: 51)

## 2.2 Noncorrelating word order characteristics

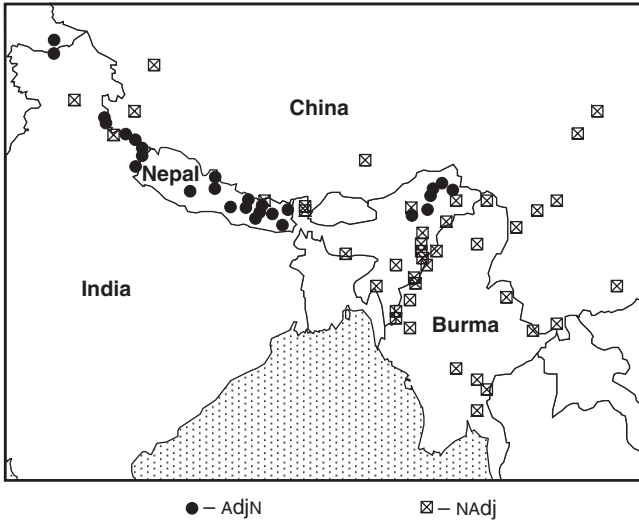
As discussed in detail in Dryer (1992), there are a number of word order characteristics which, contrary to widespread belief, do not correlate with the order of object and verb. These include the order of adjective, numeral, and demonstrative with respect to a modified noun and the order of degree words 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 (forthcoming).

### 2.2.1 Order of adjective and noun

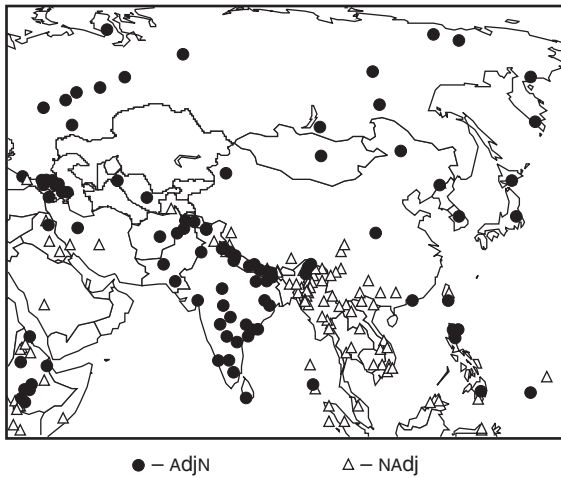
Both orders of adjective and noun are well-attested as preferred orders among TB languages. Among the eighty-five TB languages for which I was able to obtain information on this, the preferred order is AdjN in twenty-eight languages and NAdj in forty languages, and in seventeen languages, both orders 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



MAP 3.1 ORDER OF ADJECTIVE AND NOUN



MAP 3.2 ORDER OF ADJECTIVE AND NOUN IN ASIA

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 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 seventy-nine TB languages for which I was able to obtain information on the order of demonstrative and noun, fifty-one are DemN, twenty 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 noun, as in the example in (7) from Nishi.

- (7) *sa nyem sî*  
 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 twenty-five languages in which DegAdj is preferred, eleven languages in which AdjDeg is preferred, and four in which both orders occur with no evidence that one is preferred. The subgroup with the most AdjDeg 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 AdjDeg while Nyamkad (Sharma 1992) is DegAdj.

### 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 xǐhuan Susan*      b. *Susan xǐhuan wǒ*  
 1SG like Susan      Susan like 1SG  
 ‘I like Susan.’      ‘Susan likes me.’

SOV word order is also possible, with the object marked by the object marker *ba*, as in (9), though this word order is less common (see Sun and Givón 1985 for evidence from narrative and oral text).

- (9) *wǒ bǎ 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. *wǒ-men 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 grammaticization from their original status as verbs to their current status as more preposition-like elements. I will assume here that this grammaticization has proceeded far enough for some of them to justify calling themselves prepositions.

Mandarin also has postpositions, as illustrated by *qǐ* ‘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 jiǎohúá*  
 3SG like fox like sly  
 ‘He is sly like a fox.’

Note how both of these postpositions optionally co-occur with a preposition; however, the postposition in these constructions is obligatory. The number and frequency of both prepositions and postpositions in Mandarin suggests 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 AD. 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: 142)

The construction in (13) is Marker-Standard-Adjective (*bǐ+tā+pàng*). Again, this order is quite rare crosslinguistically; 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 above in Section 2.1.

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 ěrduō  
 rabbit LINK ear  
 ‘the rabbit’s ear’ (Li and Thompson: 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: 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 Southeast 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) above. 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

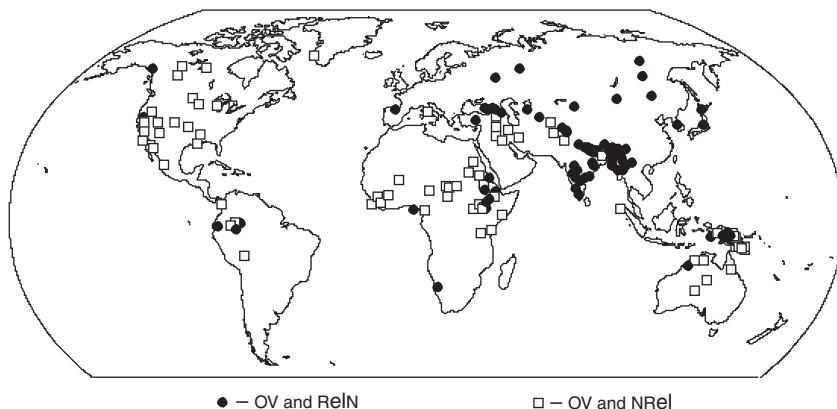
modifiers on the other. Mandarin typically places the head before a complement, a dependent which is required grammatically and semantically to complete the meaning of the phrase: verb before object, adposition before object, copula before predicate, verbs with meanings like 'want' or 'able' before their 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; 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 simply 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



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

and noun among OV languages. Map 3.3 shows clearly how ReIN 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 ReIN order of Chinese resembles the common ReIN order to the north and to the west, in TB. But there are other respects, however, in which Chinese word order 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 above 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 above, 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 premodifying 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.

#### REFERENCES

- Bradley, David (1994) ‘Tibeto-Burman’, in Christopher Moseley and R.E. Asher (eds) *Atlas of the World’s Languages*, London: Routledge 168–82.
- 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 Vocabulary*, Rangoon: American Baptist Mission Press.
- Dai, Qingxia (1985) *Achanggu Jianzhi* (A brief description of the Achang language), Beijing: Minzu Chubanshe.
- Dryer, Matthew S. (1988) ‘Object-verb order and adjective-noun order: dispelling a myth’, *Lingua* 74: 77–109.
- 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. (2001) ‘Mon-Khmer word order from a crosslinguistic perspective’, in a volume of papers from the 1996 meeting of the South East Asian Linguistics Society, Eugene, Oregon.
- Dryer, Matthew S. (forthcoming) ‘Word Order in Tibeto-Burman’.
- Giridhar, Puttushetra Puttushwamy (1980) *Angami Grammar*, Manasagangotri, Mysore: Central Institute of Indian Languages.

- Greenberg, Joseph H. (1963) 'Some universals of grammar with particular reference to the order of meaningful elements', in Joseph Greenberg (ed.) *Universals of Language*, Cambridge, MA: MIT Press 73–113.
- Hamilton, R.C. (1900) *An Outline Grammar of the Dafla Language*, Shillong: The Assam Secretariat Printing Office.
- Harriehausen, Bettina (1990) *Hmong Njua*, Tübingen: Max Niemeyer Verlag.
- 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.
- Hawkins, John A. (1983) *Word Order Universals*, New York: Academic Press.
- 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 (1976) *Kokborok Grammar*, Central Institute of Indian Languages, Grammar Series 3.
- LaPolla, Randy J. (1994) 'On the change to verb-medial word order in Proto-Chinese: evidence from Tibeto-Burman', *Current Issues in Sino-Tibetan Linguistics* 98–104.
- LaPolla, Randy J. (1998) 'The role of migration and language contact in the development of the Sino-Tibetan language family', paper presented at the workshop on the connection between areal diffusion and the genetic model of language relationship.
- 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.
- Masica, Colin P. (1976) *On Defining a Linguistic Area*, Chicago: University of Chicago Press.
- Matthews, Stephen and Yip, Virginia (1994) *Cantonese: A Comprehensive Grammar*, New York: Routledge.
- Rangan, K. (1979) *Purki Grammar*, Manasagangotri, Mysore: Central Institute of Indian Languages.
- Saul, Janice E. and Wilson, Nancy Freiburger (1980) *Nung Grammar*, 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**

---

**OLD CHINESE AND**

---

**CHINESE DIALECTS**

---





# A SKETCH OF LATE ZHOU CHINESE GRAMMAR

*Derek Herforth*

## 1 INTRODUCTION

The Chinese, or Sinitic, languages are widely, if not unanimously, regarded as genetically-related to the Tibeto-Burman (TB) family whose major members are described in this volume. Evidence for the Sino-Tibetan (ST) hypothesis consists of phonological and semantic correspondences between the basic word-stocks and reconstructed morphological systems of Old Chinese (thirteenth century BCE–third century CE) and TB languages, especially the earliest attested of these, Old Tibetan (eighth century) and Old Burmese (twelfth century). Recent discussions of these aspects of the ST hypothesis can be found in Norman (1988: 12–16) and Baxter (1995) for ST lexicon, and Norman (1988: 84–7) and Baxter and Sagart (1998) for morphology.

The morphology so far reconstructed for Old Chinese (OC) is exclusively derivational. While the analysis of OC data reveals some evidence of regular phonological alternations that derive nouns from verbs (*\*bənʰ* ‘a share’ from *\*pən* ‘to divide’), or transitives from intransitives (*\*kinʰ* ‘to see’ from *\*ginʰ* ‘to appear’), etc. no traces have been found of subsyllabic markers of tense, number (singular vs plural) or core participant relations (subject vs direct object, etc.). Languages such as OC, Mandarin, and Modern English that fail to mark the latter relations by either morphology or case particles make up for that lack by requiring strict word order within the clause. (Cf. Latin or Japanese, where the phrase-internal marking of subject and object by case endings/particles permits greater freedom of ordering.) After brief remarks about the Late Zhou Chinese (LZC) corpus (2), this chapter will sketch some of the basic syntactic constructions of the language attested in that corpus.

## 2 LATE ZHOU CHINESE

LZC is not the earliest form of written Chinese, but the language preserved in texts dating primarily from the Warring States period (475–221 BCE), the last historiographic division within the long Zhou dynasty (eleventh century to 221 BCE). Chinese is attested at least as early as the thirteenth century BCE, late in the Shang dynasty, (sixteenth to eleventh century BCE), but the extant corpus of pre-Warring States times consists of a limited number of relatively formulaic discourse genres, chiefly divinatory, ceremonial, and poetic texts. For two reasons this corpus fails to provide a comprehensive picture of the syntactic resources of early Chinese: (a) there are considerable problems in the decipherment and precise interpretation of much of the earliest data, and (b) due to their largely formulaic nature, surviving early texts appear to display a rather limited repertory of constructions. It is only from Warring States times that a corpus of narrative and expository prose texts has survived sufficiently intact to allow a relatively detailed account of early Chinese syntax.

This corpus is written in a script which for the most part does not reflect the sort of derivational morphology mentioned above. Thus, a single graph was typically used to write the distinct, but morphologically related syllables meaning ‘see’ and ‘appear’, or ‘a share’ and ‘divide’. (Cf. the single spelling <suspect> for both the noun ‘SUSpect’ and the verb ‘susPECT’, or <read> for the phonologically distinct present and past tense forms of the verb.) This homography renders LZC text considerably more ambiguous than either the spoken language assumed to underlie it or text in an alphabetic script. The impression of ambiguity is further strengthened by the lack of punctuation in most premodern editions of LZC texts, i.e. the failure to represent graphically any of the disambiguating prosodic features, such as juncture or pause length, which can be assumed to have been part of spoken LZC.

By comparing the morpheme glosses to the translations in the examples that follow, readers will find that written LZC seems in many ways a minimal linguistic system. Many obligatory distinctions made lexically and/or morphologically and/or syntactically in modern European languages are simply not overt in LZC text; rather, they must be construed compositionally from the immediate linguistic co-text or inferred from the larger context. Lest it be assumed that such minimality is entirely or even primarily an artifact of written LZC, interested readers may consult David Gil’s description of Riau Indonesian, a spoken language in many ways as ‘underspecified’ (Riddle and Stahlke 1992) as the written language analysed here.

The following brief survey will focus on three major areas of LZC grammar: (1) predication, (2) modification and its relation to clause demotion and nominalization, and (3) basic word order and its variations. Brief treatment of several aspects of LZC grammar not covered here may be found in Norman (1988: 83–108). Pulleyblank (1995), Gassmann (1997) and the other works listed in the bibliography provide more detailed description and analysis.

### 3 PREDICATION

In common with other languages, the majority of LZC clauses are organized around a lexical verb; however, several other lexical classes are also used as predictors.

#### 3.1 Nouns in predicate function

Time-stable, classificatory predication is expressed [*x*, *y* 也 *yě* (assertive particle)], where, in the simplest case both *x* and *y* are noun phrases.

- (1) 周 公 弟 也 Me 4.9/23.11  
*Zhōu gōng dì yě*  
 PN sire younger.brother PRT  
 ‘The sire of Zhou was the younger brother.’

The relation of *x* to *y* need not be that of strict identification or class inclusion. In each clause in (2), *y* has an explanatory construal, the second clause showing in addition the negative form of [*x*, *y* 也], namely [*x* 非 *fēi* *y* 也].

- (2)                     也;           非           也 Me 10.1/51.13  
*qí zhì ěr lì yě qí zhòng fēi ěr lì yě*  
 3PO arrive 2p strength PRT 3PO hit.on is.not 2p strength PRT  
 (It’s like target shooting from a distance of over one hundred paces.)  
 ‘Its (=the arrow’s) arrival (at the target) is (a matter of/due to) your strength;  
 its hitting the bull’s eye is not (due to) your strength.’

Under the explanatory construal of *y*, both *x* and *y* can accommodate full clauses, *x* often nominalized by the determiner-like 者 *zhě* ‘the one who.../situation that...’.

- |     |             |            |           |             |           |             |              |            |           |
|-----|-------------|------------|-----------|-------------|-----------|-------------|--------------|------------|-----------|
|     | x           |            |           |             |           |             |              | 者          |           |
| (3) | 然           | 至          | 日         | 晚           | 必         | 歸           | 饑            | 者          |           |
|     | <i>rán</i>  | <i>zhì</i> | <i>rì</i> | <i>wǎn</i>  | <i>bì</i> | <i>guī</i>  | <i>xiǎng</i> | <i>zhě</i> |           |
|     | be.thus     | arrive     | sun       | late        | must      | return.home | ration       | det        |           |
|     | y           |            |           |             |           |             |              | 也          |           |
|     | 塵           | 飯          | 塗         | 羹           | 可         | 以           | 戲            | 而          | 不         |
|     | <i>chén</i> | <i>fàn</i> | <i>tú</i> | <i>gēng</i> | <i>kě</i> | <i>y•</i>   | <i>xì</i>    | <i>ér</i>  | <i>bù</i> |
|     | dust        | food       | mud       | stew        | can.be    | INS         | play         | CN         | NA        |
|     |             |            |           |             |           |             |              | can.be     | eat       |
|     |             |            |           |             |           |             |              |            | PRT       |
- ‘However, the reason that, when the day is done, (children playing house) always return home for chow is that, while dust victuals and mud stew can be played with, they cannot be eaten.’ Hf 32.21.5

The meaning here translated ‘the reason... is that...’ is not expressed lexically, but is a part of the inferred semantics of this construction when *x* refers to an activity or situation.

### 3.2 Numerical predicates

Number phrases, Number (+ optional NP), can function as predicates, in which case they will be negated by the same adverbial particle 不 *bù* (NA) used with prototypical verbs, already seen immediately above in 不可食 *bù kě shí* ‘cannot be eaten’, and again in both clauses of (4).

- |     |           |           |            |            |   |           |           |           |            |              |
|-----|-----------|-----------|------------|------------|---|-----------|-----------|-----------|------------|--------------|
| (4) | 醫         | 不         | 三          | 世          | , | 不         | 服         | 其         | 藥          | Lj 2.18/9.19 |
|     | <i>yī</i> | <i>bù</i> | <i>sān</i> | <i>shì</i> |   | <i>bù</i> | <i>fú</i> | <i>qí</i> | <i>yaò</i> |              |
|     | medic     | NA        | 3          | generation |   | NA        | apply     | 3PO       | medicine   |              |
- ‘If a doctor (is) not (of the) third generation, (one) does not take his medicine.’

Like verb-centred predicates, numerical predicates accommodate aspect markers, as shown by the presence in (5) of 已 *yǐ* ‘already’ and 矣 *yǐ*, roughly, ‘as of reference time’.

- |     |             |           |           |            |           |              |
|-----|-------------|-----------|-----------|------------|-----------|--------------|
| (5) | 年           | 已         | 七         | 十          | 矣         | Me 9.9/50.14 |
|     | <i>nián</i> | <i>yǐ</i> | <i>qī</i> | <i>shí</i> | <i>yǐ</i> |              |
|     | year        | already   | 7         | 10         | ASP       |              |
- ‘(His) years (were) already seventy (in number).’

### 3.3 Nouns in verbal function

In (6), the negative adverbial 不 *bù* precedes what is intuitively a noun, 大夫 *dà fū* ‘magnate’.

- |     |            |             |           |     |           |              |           |        |
|-----|------------|-------------|-----------|-----|-----------|--------------|-----------|--------|
| (6) | 楚          | 公           | 子         | ... | 不         | 大夫           | 矣         | Gy 195 |
|     | <i>Chǔ</i> | <i>gōng</i> | <i>zǐ</i> |     | <i>bù</i> | <i>dà fū</i> | <i>yǐ</i> |        |
|     | PN         | sire        | son       |     | NA        | magnate      | ASP       |        |
- ‘The scion of Chu... no longer (acts like) a (mere) magnate.’

As in Modern English, in LZC many nouns also function as (in)transitive verbs. Zhang (1999) lists the following semantic classes in which this sort of dual function is commonly found: items of material culture, especially tools and structures (衣 *yī* ‘upper garment’, 衣 *yì* ‘wear’, ‘clothe’; 館 *guǎn* ‘(to) lodge’), body parts (指 *zhǐ* ‘finger’, 指 *zhǐ* ‘point’; 目 *mù* ‘(to) eye’, ‘designate’), social identities (王 *wáng* ‘king’, 王 *wàng* ‘rule as king’; 子 *zǐ* ‘son’, 子 *zǐ* ‘act like/treat

as son'), ambient natural phenomena (雨 *yǔ* 'rain', yù 'to precipitate'; 蝗 *huáng* 'locust', 'suffer a plague of locusts'), and positional words, which pattern as nouns in LZC, (上 *shàng* 'upper part', 'above', *shàng* 'go up'; 後 *hòu* 'space behind', 'time after', 'fall behind', 'arrive late'). The distinct Mandarin readings for some of these 'functional conversions' are vestiges of derivational processes along the lines of English 'cloth' > 'clothe', 'food' > 'feed', etc. In other cases, again with parallels in English, the two syntactic functions share the same form. However, just as English denominal verbs resemble other verbs in the language by typically inflecting for past tense and for third person singular in the present, so their LZC counterparts behave like prototypical verbs in that language, co-occurring with the 'dispersed verb morphology' of negative 不 *bù* and aspectual markers, as in (6) above.

### 3.4 Adjectives as predicates

In common with their counterparts in many other East and Southeast Asian languages, LZC adjectives occur as predicates without a linking verb. Again, the negative is the adverbial 不 *bù*, not the negative copula 非 *fēi*. Such 'adjectives', in fact, constitute a class of stative verbs.

- (7) 紇 不 佞                      Zz 278.3  
*Hé bù nǐng*  
 PN NA well-spoken  
 'I, He, am not well-spoken.'

### 3.5 Noun predication markers in verb-based clauses

The data presented above reflect the treatment of numerical, denominal, and adjectival predicates in the grammar as verbal rather than nominal. Such predicates do not ordinarily co-occur with clause-final 也 *yě*, nor are they negated by 非 *fēi*. Under certain conditions, however, the two markers of noun predication, 也 and 非, do occur in verb-based clauses: [Subject (非) Verbal Predicate 也]. In such 'mixed' clauses, the explanatory force of [(非) y 也] seen in (2) and (3) above is typically preserved, as suggested by the contrast between the next two examples. (8) consists of two clauses, both with the trappings of verb-based clausehood: no final 也 and negation in the first clause by 不 *bù*. Note the straightforward, descriptive/narrative construal.

- (8) 紇 不 佞                      失 守 宗                      祧                      Zz 278.3  
*Hé bù nǐng                      shī shǒu zōng                      tiāo*  
 PN NA well-spoken, lose protect ancestral temple forebear  
 'I, He, am not well-spoken and have failed to preserve the cult of our distant forebears.'

By contrast, while the two clauses of (9) are verb-based, they also show the markers of noun predication: negation by 非 and final 也 in the first, and in the second, 也 following a verbal predicate negated by 不. These clauses come from a passage in which the writer of (8) is at pains to explain what motivated his conduct on a previous occasion.

- (9) 紇 非 能 害 也 知                      不 足 也                      Zz 278.5  
*Hé fēi néng hài yě zhī                      bù zú yě*  
 PN is.not able harm PRT knowledge NA sufficient PRT  
 '(The explanation) is not (as you might think) that I, He, am capable of causing harm; (rather,) it's that I lack savoir-faire.'

As this example suggests, the [Subject (非) Verbal predicate 也] construction is used to assert the speaker's explanation of a situation, as in the second clause, or to deny an assumed inference about motives or causes, as in the first. Functionally, the construction is directly comparable to English 'It's (not) that + CLAUSE', or 'What happened was + CLAUSE', and Japanese 'CLAUSE nominalized by *no* + copula *da/dewanai*'.

#### 4 MODIFICATION, CLAUSE DEMOTION AND NOMINALIZATION

In this section, we sketch three closely related areas of LZC phrase structure. Short (or 'possessive') modification of NPs provides a template for the demotion of full clauses to constituents: [NP1 之 NP2<sub>head</sub>] → [Subject 之 Predicate]. Two strategies for nominalizing predicates participate in the short modification template to yield long modification of the English relative clause type: [(Predicate nominalized by 者 *zhě* or 所 *suǒ*)<sub>=NP1</sub> 之 NP2]. Finally, while the general rule in LZC is that 'modifiers precede modifieds', there exists an important class of exceptions when the modified is a verb or predicate phrase.

##### 4.1 Modification of NP by NP and clause demotion

Modification of an NP by a single morpheme is generally achieved by simple juxtaposition, as in 塗羹 *tú gēng* 'mud stew' (NP-NP) of (3), 凶歲 *xiōng suì* 'bad year' (Adj-NP), and in the nested NP modification 楚公子 *chǔ gōng zǐ*, probably to be analysed (Chu (sire (son))) 'son of the sire of Chu' of (6). Longer modifiers tend to be linked to the head NP by the 'possessive' marker 之 *zhī*, here glossed 'PO', possession being one of the most commonly attested semantic relations between modifier and modified. [NP1 之 NP2<sub>head</sub>] structures are shown in the first and third phrases of (10), where the relation of NP1 to NP2 is clearly not that of possessor.

- (10) 萬 乘 之 國， 弑 其 君 者， 必 千  
*wàn shèng zhī guó shì qí jūn zhě bì qiān*  
 10k vehicle PO state assassinate 3PO lord DET must 1k  
 乘 之 家 Me 1.1/1.6  
*shèng zhī jiā*  
 vehicle PO family

'A state of ten thousand chariots, the person who assassinates its ruler will surely be (from) a family of a thousand chariots.'

The third person possessive pronoun, '3PO', has the special form, 其 *qí* 'her/his/its/their', already seen in (4) 其藥 *qí yào* 'his medicine', and again in the second phrase of (10).

'Possessive' 之 *zhī* and 其 *qí* also mark demoted clauses, i.e. clauses that cannot stand alone, but function as constituents, embedded within superordinate clauses. The possessive particles mark the subject as 'modifier' of the predicate, the same sort of strategy found in English 'Mary/she lost > Mary's/her loss'. The pattern occurs twice in (2): 其至 *qí zhì* 'its arrival' and 其中 *qí zhōng* 'its hitting the mark'. The bare circumstantial clause introducing (11) shows a clause with full NP subject demoted by '之 *zhī*-insertion'. Note the embedding function of clause-final 也 *yě*, a further example of which occurs in (20) below.

- (11) 丈 夫 之 冠 也， 父 命 之 Me 6.2/31.18  
*zhàng fū zhī guān yě fù mìng zhī*  
 2.25m man PO headgear PRT father order 3O

'When a young man is capped, his father ordains him.'

## 4.2 Nominalization and modification of NP by clause

LZC has two nominalizing particles, 者 *zhě* and 所 *suǒ*, which convert predicates into free relative NPs roughly analogous to the underlined phrases in ‘Who steals my purse steals trash’ for 者 *zhě*, and ‘She questioned the accused’ for 所 *suǒ*. As NPs, such derived free relatives can participate as either modifier or head in the [NP1 之 NP2] modification structure illustrated in (10).

The so-called ‘subject relativizer’ 者 *zhě* is a quasi-determiner (DET) suffixed to a predicate to form a free relative, often on the absent subject, yielding the meaning ‘(the) one(s) who/thing(s) that PREDICATE’. An example has already appeared in (10): 弑其君者 *shì qí jūn zhě* ‘(the) one who assassinates its lord’. (12) shows Predicate + 者 functioning as the head NP2 in an [NP1 之 NP2] structure.

- (12) 居 是 邦 也, ... 友 其 士 之 仁 者 Ly 15.10/42.23  
*jū shì bāng yě yǒu qí shì zhī rén zhě*  
 reside DET state PRT friend 3PO officer PO humane DET  
 ‘When residing in a state, befriend the humane among its officers.’  
 (lit. its officers’ are.humane the.ones)

DET 者 *zhě* does not, in fact, always stand for the subject of the predicate it nominalizes, even if no subject is expressed. Immediate context will often require construal of the derived NP as a free abstract event or situation nominal, as already exemplified in the *x* constituent of (4), ‘When day grows late, the inevitable return home for rations (is explained by) ...’. Construal of 必歸饜者 ‘must return.home ration DET’ as a subject nominalization, ‘those who must return home for rations’, does not cohere with the explanatory predicate which follows: ‘(the fact that,) while dust rice and mud gruel can be used to play with, they cannot be eaten’. This example illustrates how construing decontextualized LZC sentences often requires looking ahead to ascertain what sort of semantic structure awaits the integration of an underspecified initial NP or clause. (See again Gil on Riau Indonesian.)

The second relativizing particle, ‘OR’ 所 *suǒ*, nominalizes on the object (or locative phrase) of the predicate to which it is prefixed. Thus, 所得 *suǒ dé* ‘OR get’, means ‘what (x) gets/got’. (Two further examples of 所 *suǒ* object nominalization occur in (23).)

- (13) 季武子以所得於齊之兵作林鍾 Zz 266.16  
*Jì Wǔ zǐ yǐ suǒ dé yú Qí zhī bīng zuò lín zhōng*  
 PN INS OR get LP PN PO weapon make forest bell  
 ‘Ji Wuzi, with the weapons he had obtained from Qi, made a set of “forest bells”.’  
 (i.e. he cast the bells using the metal obtained from recycled weapons.)

In (13), the free relative, 所得於齊 ‘what he got from Qi’, functions as the modifying NP1 in an [NP1 之 NP2] construction, literally ‘[what (he) got from Qi]’s weapons’, the sort of complex modification which English is constrained to express by a relative clause, ‘the weapons (which) he obtained from Qi’. This sort of example suggests that in LZC the modification of a noun by a clause requires prior nominalization of that clause. The resulting NP is then attached as NP1 to the head noun, NP2, such that ‘modifier always precedes modified’. The language shows no direct, post-head attachment of clauses to NPs by means of relative pronouns.

## 4.3 Predicate modification

A basic distinction can be drawn between ‘adpredicate’ particles, a relatively closed class in synchronic description (but subject to enrichment through grammaticization), and the open

class of full words whose occasional predicate-modifying function is effected by positioning them immediately left of the predicate. The latter ‘syntactically derived adverbials’ (but not the particles) all have more basic functions, chiefly adjectival, verbal or nominal, each of which correlates with non-adverbial positions in the clause.

Two examples of the particle type occur in (14): the aspectual negative (AN) 未 *wèi* ‘not.yet/not.quite’ and the existential quantifier (XQ) 或 *huò* ‘there.exists.a.specific.case.of’, here rendered ‘anyone’, in the context of 未 *wèi* ‘not.yet’ and 北方之學者 *běi fāng zhī xué zhě* ‘northern scholar(s)’.

- (14) 北 方 之 學 者 未 能 或 之 先 也 Me 5.4/29.21  
*běi fāng zhī xué zhě wèi néng huò zhī xiān yě*  
 north quarter PO study DET AN able XQ 3O get.ahead PRT  
 ‘Among scholars of the north, there has not yet been anyone able to surpass him.’  
 (Cf. a quantified NP version, ‘So far, no northern scholar has been able to surpass him’.)

Some expressions translatable as English adverbials are LZC intransitive verbs rather than adverbs. In (15), the subject of 久 *jiǔ* ‘last.long’ is the clause 天棄商 *tiān qì Shāng* ‘Heaven abandoned Shang’, demoted by 之-insertion to ‘Heaven’s abandonment of Shang’.

- (15) 天 之 棄 商 久 矣 Zz 98.30  
*tiān zhī qì Shāng jiǔ yǐ*  
 heaven 3PO abandon PN last.long ASP  
 ‘Heaven’s abandonment of Shang has now lasted a long time’ =  
 ‘Heaven long ago abandoned Shang’.

When the verb of (15), 久 *jiǔ*, occurs squarely between subject and predicate, as in (16), however, there can be little question of its adverbial function.

- (16) 慶 克 久 不 出 Zz 221.17  
*Qīng Kè jiǔ bù chū*  
 PN long NA emerge  
 ‘Qing Ke for a long time did not go out.’

A second example of this type is found in the first clause of (17), the full version of (6).

- (17) 楚 公 子 甚 美 不 大 夫 矣 Gy 195  
*Chǔ gōng zǐ shèn měi bù dà fū yǐ*  
 PN sire son be.extreme be.fine NA magnate ASP  
 ‘The scion of Chu (is) very grand; (he) no longer (acts like) a (mere) magnate.’

Here the adjective 甚 *shèn* ‘be.extreme’ modifies a second adjective serving as predicate, 美 *měi* ‘be.fine’.

Quantifying adverbials, derived by position from nouns and numbers, are attested clustering in the prepredicate site, as in (18).

- (18) 吾 日 三 省 吾 身 Ly 1.4/1.11  
*wú rì sān xǐng wú shēn*  
 1p day 3 examine 1p body  
 ‘I daily examine myself three times/ways.’

Deverbal adverbials do not cluster as readily, however, those expressing manner or degree often moving to a postverbal position.



- (19) 黍 種 嘗 貴 甚 Hf 31.34.2  
 shǔ zhǒng cháng guì shèn  
 millet seed once be.priized extreme  
 ‘Millet seed was once extremely expensive.’

嘗 *cháng* ‘once’ is derived from the transitive verb ‘taste’ by its position preceding the predicate 貴 *guì* ‘be.priized’. 甚 *shèn* ‘extreme’, which we observed in prepredicate position in (17), has moved to the right periphery of the clause in (19).

As a site for adverbials, the right periphery of the clause is more capacious than prepredicate position, accommodating phrases with their own arguments or modifiers.

- (20) 夫人鄭 袖知王 悅 愛之也亦悅 愛之甚於王  
 fū rén Zhèng Xiū zhī wáng yuè ài zhī yě yì yuè ài zhī shèn yú wáng  
 Lady PN know king pleased dote 3O PRT also pleased dote 3O extreme LP king  
 ‘Lady Zheng Xiu knew the king doted on her (=the new girl in the harem) and she (=Zheng Xiu) too doted on the girl *even more than the king (did)*’. Hf 31.28.25

- (21) [益] 施 澤 於 民 未 久 Me 9.6/49.10  
 Yì shī zé yú mǐn wèi jiǔ  
 PN dispense bounty LP folk AN last.long  
 ‘(Yi) dispensed bounty to the people *for not very long*.’

In (21), 未久 *wèi jiǔ* ‘not.quite last.long’ might be analysed as the main predicate, but as there are no signs of clause-demotion such as seen in (15) – the first clause of (21) is subjectless, so 其施澤於民 *qí shī zé yú mǐn* ‘his dispensing of bounty to the people’ would be expected – we assign the negated predicate 未久 to the clause-final ‘secondary predicate’ position required to account for the syntax of examples (19) and (20).

While the details of predicate modification are not well understood, the LZC clause clearly has two positions for such modifiers. The internal site between subject and predicate is of limited capacity, allowing some ‘stacking’ of monosyllabic modifiers, as in (14) and (18), but closed to predicators with their own complements, as in (20), or modifiers, as in (21). The right-peripheral site for such ‘heavy’ modification of the predicate appears to provide the single, principled exception to the generalization about modifiers preceding their heads in LZC.

## 5 BASIC CONSTITUENT ORDER

Mention of syntactic positions leads to issues of basic constituent order and its variations. All data cited above bear out the description of LZC as a verb medial language with basic S(ubject) V(erb) O(bject) order. We have further observed adpositional phrases both in ‘A(djunct)1’, between subject and verb, as in (13), and in ‘A2’, following the object, as in (21). This data is repeated below in abbreviated form.

- |      | P        | R | E                        | V | E        | R | B |  | P        | O | S         | T | V   | E           | R                    | B |
|------|----------|---|--------------------------|---|----------|---|---|--|----------|---|-----------|---|-----|-------------|----------------------|---|
|      | <u>S</u> |   | <u>A1</u>                |   | <u>V</u> |   |   |  | <u>O</u> |   | <u>A2</u> |   |     |             | <u>2nd predicate</u> |   |
| (13) | 季武子      |   | 以所得於齊之兵                  |   | 作        |   |   |  | 林        |   | 鍾         |   |     |             |                      |   |
|      | Ji Wuzi  |   | with weapons got from Qi |   | made     |   |   |  | forest   |   | bells.    |   |     |             |                      |   |
| (21) | [益]      |   |                          |   | 施        |   |   |  | 澤        |   | 於良        |   | 未   | 久           |                      |   |
|      | ‘Yi      |   |                          |   | dispense |   |   |  | bounty   |   | to folk   |   | not | very long.’ |                      |   |

The LZC postverb appears to be limited to the three positions indicated, maximally occupied by two argument/adjunct/complement phrases (O, A2) and a single secondary predicate. (21) thus exemplifies one sub-type of fully saturated postverb. Discussion of the preverb below will introduce further, semantic constraints on the postverbal part of the clause.

### 5.1 Topicalization

A definite NP may occupy a site to the left of the subject as the topic (T) of a lower predication. LZC topics are of two types, gapped and gapless. A gapped topic phrase retains a grammatical relation (S, O, or A) to the predicator, the position associated with that relation often occupied by a pronoun co-referential with the topic phrase, such as the third person oblique 之 *zhī* ‘it’ in the O position of (22).

- |      |               |   |              |             |            |             |
|------|---------------|---|--------------|-------------|------------|-------------|
|      | <u>T</u>      | S | <u>A1</u>    | v           | O          |             |
| (22) | 吾道            |   | 一以           | 貫           | 之          | Ly 4.15/8.7 |
|      | <i>wú dào</i> |   | <i>yī yǐ</i> | <i>guàn</i> | <i>zhī</i> |             |
|      | my way        |   | 1 INS        | link        | 3O         |             |
- ‘My way, (I) link it together with a SINGLE (principle).’

(For the inversion of prepositional 以 *yǐ yī* ‘with one (thing)’ to postpositional 一以 ‘with ONE SINGLE (thing)’, see below.)

A topic is ‘gapless’ if it does not bear a relation S, O, or A to the main predicator.

- |      |                    |            |              |                      |                 |              |          |
|------|--------------------|------------|--------------|----------------------|-----------------|--------------|----------|
|      | <u>T</u>           |            |              | <u>S</u>             | v               | <u>O</u>     |          |
| (23) | 子重                 | 於          | 是            | 役也                   | 所獲              | 不如           | 所亡       |
|      | <i>Zhichóng yú</i> | <i>shì</i> | <i>yì yě</i> | <i>suǒ huò bù rú</i> | <i>suǒ wáng</i> |              | Zz 230.2 |
|      | PN                 | LP         | DET job PRT  | OR bag NA            | be.like         | OR disappear |          |
- (lit. Zichong on this campaign the caught did not equal the lost =)  
‘On this campaign, Zichong’s gains did not equal his losses.’

The topic in (23) is gapless according to the above definition as the predicator, ‘be.like’, affords no S, O, or A slot for a third NP such as ‘Zichong’. The English strategy is to treat ‘Zichong’ as possessive modifier (= demoted agent) of the verb’s two arguments, ‘the caught’ and ‘the lost’, as shown in the idiomatic translation. If we were to recognize gaps in the structure of LZC NPs, then a sentence like (23) would be described as having a ‘gapped possessor topic’.

### 5.2 Scope-based constraints on SVO

In (23), a second phrase occurs to the left of the main predication in both LZC and English – the ‘locative’ adjunct 於是役 *yú shì yì* ‘on this campaign’. In both languages, the unmarked position for such phrases is in the postverb, following the object, as shown in the predicate of (21): 施澤於民 *shī zé yú mín* ‘dispensed bounty to the folk’. The fronting of the adjunct in (23) ensures that the entire main predication is construed within the scope of the temporal expression; that is, both the gains and the losses referred to in the SVO kernel are understood as having occurred ‘on this campaign’. (Cf. ‘Zichong’s gains did not equal his losses on this campaign’, where the scope of the adjunct can be construed as narrow, restricted to the NP immediately preceding it.) Constituent order in both LZC and English here respects a principle of ‘scope iconicity’ whereby a phrase tends to precede other phrases whose meanings fall within its scope. This principle, exemplified in both languages in (23), is adhered to much

more strictly in LZC than in English, accounting for a large family of LZC variations on SVO that English lacks. In the remainder of the discussion, we will sketch the extent to which scope-iconicity determines LZC constituent order.

Analysis reveals that constituent questions (as in ‘Who did he see?’) and statements with identificational focus (as in ‘It was PAT he saw’) have inter-phrasal scope relations which differ from those of their plain declarative cousins (‘Lou saw Pat’). A constituent question word asks for the identification of an unknown entity, and so constitutes a kind of focus. LZC reflects this difference between plain and focused O and A2 phrases by inverting the focused type to a preverbal position: default S V O A2 + Focus/Question on O/A2 → S O/A2 V. This is shown in the question and answer sequence (24a, b).

- |         | O   | V  | O complement     |              |
|---------|---|--|------------------|--------------|
| (24) a. | 何<br><i>hé</i><br>what  | 謂<br><i>wèi</i><br>call                  | 辰<br><i>chén</i> |              |
| (24) b. | 日 月 之 會 <sub>i</sub><br><i>rì yuè zhī huì</i><br>sun moon PO meet | 是 <sub>i</sub> 謂<br><i>shì wèi</i><br>PF | 辰<br><i>chén</i> | Zz 340.25–26 |
- ‘What is it (we) refer to as *chen*?’  
‘It is the conjunction of the sun with the moon that (we) refer to as *chen*.’

In (24a), 何 *hé* ‘what’ is the object of the verb 謂 *wèi* ‘call’, but occurs to its left. (辰 *chén* is an object complement, unaffected by the syntactic operation.) In the answer, 日月之會 *rì yuè zhī huì* ‘the conjunction of the sun and moon’, object of the repeated verb, again precedes that verb and is marked as focused by the resumptive proform 是 *shì* occurring between it and the verb. Focus is reflected in the English translation by the cleft construction, ‘It BE NP that CLAUSE’.

When the object of an adposition, such as instrumental 以 *yǐ* ‘with’, is questioned or in focus, the same sort of inversion occurs. Since O, A2 and ‘object of adposition’ are all types of complement, the rule governing inversion in both the postverb and adpositional phrases can be stated: default [Head + Complement] + Focus/Question on Complement → [Complement + Head]. This rule is further exemplified in (22), our earlier example of a gapped topic, repeated here, where the complement of 以 *yǐ* is in focus.

- (22) 吾 道 一 以 貫 之 Ly 4.15/8.7  
my way I INS link 3O  
‘My way, it is with a SINGLE (principle) that (I) link it together.’

In (25), the complement of the same instrumental adposition is questioned, the question word, like each of the focused complements in (22) and (24a, b), preceding its head.

- (25) 何 以 報 德 Ly 14.34/40.15  
*hé yǐ bào dé*  
what INS repay kind.deed  
‘With what does one repay a kindness?’

There are numerous other sub-conditions, all related to the semantic relation ‘scope over’, that motivate the appearance of O and A2 phrases in the preverb such that S O/A2 V occurs as a regular, semantically-determined alternant to basic S V O A2 order.

Phrases with the English meanings ‘anywhere (they went)’ or ‘whatever (he says)’ are of this type, setting a generic condition on the situation described in (the rest of) the predicate.

'She met with success wherever she went' means 'If she went somewhere/anywhere at all, then she invariably met with success'. LZC lacks words like 'anywhere' or 'whatever' that mark wide scope phrase-internally. (Recall the lack of the existentially quantified NP 'no northern scholar' in the LZC version of (14).) The LZC strategy is to signal wide scope by position, placing underspecified NPs like 所往者 *suǒ wǎng zhě* 'where (x) went' clause-initially to mean 'wherever (x) went'.

	<u>Wide-scope</u>			NP <sub>i</sub>	<u>adverbial</u>				
(26)	叔孫	所	館	者	雖	一	日		
	<i>Shūsūn</i>	<i>suǒ</i>	<i>guǎn</i>	<i>zhě</i>	<i>suī</i>	<i>yī</i>	<i>rì</i>		
	PN	OR	lodge	DET	though	1	day		
	S	A1	adv	V	<u>O</u>			<u>A2</u>	<u>2nd pred</u>
			必	葺	其 <sub>i</sub>	牆	屋		
			<i>bì</i>	<i>qì</i>	<i>qí</i>	<i>qiáng</i>	<i>wū</i>		
			must	thatch	3PO	wall	roof		
			去	之				如	始
			<i>qù</i>	<i>zhī</i>				<i>rú</i>	<i>shǐ</i>
			leave	3O				be.like	1st arrive

Wherever Shusun lodged, even for a single day, he always rethatched the walls and roof, leaving it as it had been when he first arrived. Zz 382.23

Cf. the further syntactic option afforded by phrase-internal marking of wide scope in English: 'Shusun invariably rethatched the walls and roof of wherever/any place where he lodged, even if only for a day, leaving it as it had been when he first arrived'.

LZC lacks 'free-choice' NPs like 'wherever' and wide-scope determiners like 'any', both exploited in the alternative translations of (26). A phrase like 叔孫所館者 *Shūsūn suǒ guǎn zhě* in a postverb site can have only the narrow-scope, referential construal '(a/the place) where Shusun lodged'.

## 6 CONCLUSION

In spite of a shared basic word order, SVO, the clause structures of LZC and English differ significantly in fundamental ways. First, the number of phrases that can follow the LZC verb is strictly limited such that 'bottom-heavy' sentences like (27) are never attested.

- (27) Fred buttered a scone for his Labrador methodically with a putty knife in the pantry at 3:27a.m. last Monday.

We suspect that, partly for reasons related to scope, such chains of postverb phrases would be ungrammatical in LZC, as they are in modern Chinese languages.

Second, as we have seen, under certain conditions, either semantic (wide scope) or discourse-pragmatic (identificational/interrogative focus), O and A2 phrases must occur in the LZC preverb. This means that, as a 'basic word order', SVO is less entrenched in LZC than in English. Stated another way, the LZC clause has room for a greater number and variety of phrase-types in its 'top-heavy' preverb than does English. The left periphery accommodates gapless topics (23) and wide-scope 'any' phrases (26), while there are also slots for a variety of constituents between the subject and verb: adjuncts (14), object question phrases (24a), and focused elements (24b). Leaving aside the complex issue of adverbials and secondary

predication, a rough representation of these major differences between LZC and English clause structure can, for mnemonic purposes, be sketched as in (28).

- (28) a. LZC: ... S A1 ... V O A2  
 b. English: ... S V O A1 A2 ...

### ADDITIONAL ABBREVIATIONS

AN	aspectual negative 未
CN	connective 而
DET	determiner 者, 是, etc.
INS	instrumental adposition 以
LP	locative-path adposition 於
LZC	Late Zhou Chinese
NA	negative adverb 不
OR	object relativizer 所
P	person (grammatical category)
PF	deictic proform 是
PN	proper noun
PO	possessive, clause-demoting 之
PRT	particle
XQ	existential quantifier 或
3O	third person oblique pronoun 之
3PO	third person possessive 其

### REFERENCES

- Baxter, William H. (1995) ‘“A stronger affinity ... than could have been produced by accident”’: a probabilistic comparison of Old Chinese and Tibeto-Burman’, in William S.-Y. Wang (ed.) *The Ancestry of The Chinese Language* (Journal of Chinese Linguistics monograph series, 8) Berkeley, CA: Journal of Chinese Linguistics.
- Baxter, William H. and Sagart, Laurent (1998) ‘Word formation in Old Chinese’, in J.L. Packard (ed.) *New Approaches to Chinese Word Formation*, Berlin: Mouton.
- Boltz, William G. (1999) ‘Language and writing’, in Michael Loewe and Edward L. Shaughnessy (eds) *The Cambridge History of Ancient China: From the Origins of Civilization to 221 BC*, Cambridge: Cambridge University Press.
- Cikoski, John S. (1978) ‘Three essays on classical Chinese grammar: (one) an outline sketch of word-classes and sentence structure in classical Chinese’, *Computational Analyses of Asian and African Languages* 8: 17–152.
- Gabelentz, Georg von der (1[881]960) *Chinesische Grammatik*, 4th edn, Halle: Niemeyer.
- Gassmann, Robert H. (1997) *Grundstrukturen der antikchinesischen Syntax: eine erklärende Grammatik*, Bern: Peter Lang.
- Gil, David (1994) ‘The structure of Riau Indonesian’, *Nordic Journal of Linguistics* 17: 179–200.
- Harbsmeier, Christoph (1981) *Aspects of Classical Chinese Syntax*, London: Curzon Press.
- Norman, Jerry (1988) *Chinese*, Cambridge: Cambridge University Press.
- Peyraube, Alain (1988) *Syntaxe diachronique du Chinois: évolution des constructions datives du XIV<sup>e</sup> siècle av. J.-C. au XVIII<sup>e</sup> siècle* [Mémoires de l’Institut des Hautes Études Chinoises, v. XXIX], Paris: Collège de France.
- Pulleyblank, Edwin G. (1995) *Outline of Classical Chinese Grammar*, Vancouver, BC: University of British Columbia Press.
- Riddle, Elizabeth and Stahlke, Herbert (1992) ‘Linguistic typology and Sinospheric languages’, in Martha Ratliff and Eric Schiller (eds) *Papers from the First Annual Meeting of the Southeast Asian Linguistics Society*, Tempe AZ: Arizona State University 351–66.
- Zhang Jiawen (1999) 張家文. ‘古漢語名詞活用說再認識’ *古漢語研究* 44.3: 7–11.

### Primary sources

References to several LZC texts cited above are to the editions of *The Chinese University of Hong Kong Institute of Chinese Studies Ancient Chinese Text Concordance Series*, 香港中文大學中國文化研究所先秦兩漢古籍逐字索引叢刊, hereafter 'ACT', edited by D.C. Lau 劉殿爵 and Chen Fong Ching 陳方正, and published by The Commercial Press, Hong Kong 商務印書館 (香港), 1992 -. Passages cited from ACT are identified according to the following convention: (chapter[.verse]/)page.line.

Gy = *Guóyǔ*. 國語. Discourses of the states. 上海師範大學古籍整理組校點. 上海: 上海古籍, 1978.

Hf = *Hánfēizǐ*. 韓非子. The Han Fei corpus. 周鍾靈, 施孝過, 許惟賢主編韓非子索引. 北京: 中華書局, 1982.

Lj = *Lǐjì*. 禮記. The Record of rites. ACT, 1992

Ly = *Lúnyǔ*. 論語. The Confucian analects. ACT, 1995.

Me = *Mèngzǐ*. 孟子. The Mencius corpus. ACT, 1995.

Zz = *Zuǒzhuàn*. 春秋左傳. The Spring and autumn annals, with the Zuo tradition. ACT, 1995.

# THE CHINESE DIALECTS: PHONOLOGY

*Jerry Norman*

Here the term dialect is used simply in the sense of a distinct local form of speech. None of the purely local forms of Chinese has the status of a standard language. The national standard, Pǔtōnghuà, is based on the dialect of Běijīng but is by no means identical with it (Hú 1987). The August 1999 issue of the *National Geographic* contains a map on which it is stated that the Sinitic branch of Sino-Tibetan family 'includes eight mutually unintelligible languages, often mistakenly called dialects'. This statement, as it stands, is highly misleading. If one takes mutual intelligibility as the criterion for defining the difference between dialect and language, then one would have to recognize not eight but hundreds of 'languages' in China; moreover, the eight 'languages' referred to in the quote are actually groups of dialects. Wú is not a language but a grouping of numerous non-mutually intelligible local forms of speech. The differences among the Wú dialects are in many cases considerable and it is hard to see how such disparate forms of speech could be considered a single language. The same is true of the other dialect groups: Mandarin, Mǐn, Hakka, Yuè, Gàn, and Xiāng. For the comparativist, Chinese is a vast dialectal complex containing hundreds of mutually unintelligible local varieties, each of which can be viewed as a distinct object for comparison. Transcending the local dialects is the national standard language, which, although officially based on the Běijīng dialect, must be recognized as a distinct entity. It is fundamentally difficult to apply the terms 'language' and 'dialect' derived as they are from a different linguistic context in Europe, in a perfectly consistent way in the case of China. In the present chapter, I will refer to the national standard language as Chinese; local forms of speech whatever their socio-linguistic status, will be referred to as dialects of a particular place. Dialects in turn are considered to belong to a small number of more or less well-defined dialect groups like Wú and Mǐn. For want of a better term, I will also refer to the totality of all forms of Chinese, ancient and modern, local and standard, as Chinese. This is admittedly awkward on occasion, but it is established usage and any attempt to remedy the situation with newly created terms would, I fear, only increase the confusion.

In the context of Sino-Tibetan, a basic question is, how are we to identify a Sinitic language? I would like to propose a simple test. It consists of a short list of words which, as far as I know, occur in all forms of Chinese; these words are further divided into eight sets based on common tonal categories. The list contains forty words; a Sinitic form of speech should contain all of these words (with perhaps a very small number of exceptions) and exhibit basically the same tonal patterns. The list is given below. Each word is given what I call a Common Chinese (CC) form; these forms have been established on the basis of modern dialect forms; unlike the traditional *Qièyùn* syllabic inventory, Common Chinese contains only distinctions attested in modern dialects. To keep this Common Chinese system simple, I have left out of consideration the Mǐn dialects which are widely recognized as being divergent in many ways. Following the CC forms are dialect forms from Běijīng (Bj), Sūzhōu

(Sz), Nánchāng (Nc), Méixiàn (Mx), and Guǎngzhōu (Gz). Forms are cited from Běijīng Dàxué (1995). (See Table 5.4 below and accompanying discussion for the tone categories.)

We can employ the forms in Table 5.1 as a diagnostic test for determining whether a particular language or dialect is Sinitic or not. In fact, it is probably the case that a shorter list could be employed for this purpose; a larger number of forms simply reduces the operation of chance. The words given here have all been remarkably stable throughout Chinese history. To apply the test, the words on the list should be compared to the pertinent dialect forms. A Sinitic dialect should have etymologically related forms for all or most of the items on the list. Obviously the large, metropolitan dialects presented on the list meet this criterion perfectly. As a further illustration, in Table 5.2 I give forms from five dialects less well known. Hépíng (Hp), Zhènqián (Zc) and Jiāngshān (Js) are from my own field notes. Jiāngyǒng (Jy) forms are from Huáng (1993) and Jiànchuān (Jc) forms are from Xú and Zhào (1984).

It is clear that Jiāngyǒng, Hépíng, Zhènqián, and Jiāngshān fully meet the test of being Sinitic dialects, but what about Jiànchuān? The most arresting feature of this dialect is that its tonal system is apparently aberrant from the point of view of Sinitic dialects. However, closer inspection reveals that if we disregard the numerical designations of the tone, we see that there is a fairly high degree of correspondence among tonal classes. Common Chinese tone one, for example, corresponds regularly to Jiànchuān tonal category two, Common Chinese tone three to Jiànchuān tone one, and so forth. In fact there is more than 70 per cent agreement between the Jiànchuān forms and those of the other unquestionably Sinitic dialects. Jiànchuān, nevertheless, remains the least Sinitic of the dialects displayed here.

Let us examine Jiànchuān a bit more closely; it is generally considered to be a dialect of the Bái language and is placed by many on the Tibeto-Burman side of the Sino-Tibetan family. Such indeed is the official view among linguists in China. Some Western scholars, however, have viewed Bái as a variety of Chinese. R.A.D. Forrest in his 1948 book, *The Chinese Language*, for example, attempts to link Bái to the Mǐn dialects, influenced no doubt by the variant designation for the Bái, Mǐnjiā. Forrest was severely hampered by his Bái material, which apparently did not register tones. In *The Tibeto-Burman Lexicon*, edited by Huáng Bùfán (1992) Bái is included as one of the Tibeto-Burman languages, yet even a cursory inspection of this work shows that Bái shows only a small number of clear links with Tibeto-Burman. While it would probably be going too far to consider Bái a Sinitic dialect, its close links to Sinitic cannot be easily dismissed.

As indicated above, Chinese proper is divided into a multitude of dialects, often differing very dramatically from one another. It has long been observed that these dialects fall into a small number of groups whose members bear a number of common traits. The conventional scheme recognizes seven or eight such groups. To some extent this classification is based on intuition and much more needs to be done to validate it. Below I give a brief survey of these dialect groupings.

In my 1988 book, *Chinese*, I divided Chinese dialects into three zones – Northern, Central and Southern. This division was based on a set of ten characteristics plotted on a table. Here I would like to repeat this exercise, adding a few more traits, as well as removing the Mǐn dialects from the dialects illustrated. The traits used in Table 5.3 are the following:

- 1 The third person pronoun is *tā*, or cognate to it.
- 2 The subordinative particle is *de*, or cognate to it.
- 3 The copula is *shì*, or cognate to it.
- 4 Velars palatalize before high front vowels.
- 5 Words like *rǎn* ‘dye’ and *rè* ‘hot’ have a non-nasal initial.
- 6 Words like *wěi* ‘tail’ and *wén* ‘mosquito’ have a non-nasal initial.



TABLE 5.1 DIAGNOSTIC LIST FOR IDENTIFYING SINITIC LANGUAGES (I)

	CC	Bj	Sz	Nc	Mx	Gz
<i>Tone 1</i>						
1 sky	*thian <sup>1</sup>	t <sup>h</sup> iɛn <sup>1</sup>	t <sup>h</sup> iɪ <sup>1</sup>	t <sup>h</sup> iɛn <sup>1</sup>	t <sup>h</sup> iɛn <sup>1</sup>	t <sup>h</sup> in <sup>1</sup>
2 three	*sam <sup>1</sup>	san <sup>1</sup>	se <sup>1</sup>	san <sup>1</sup>	sam <sup>1</sup>	sam <sup>1</sup>
3 chicken	*kiaɪ <sup>1</sup>	tɕi <sup>1</sup>	tɕi <sup>1</sup>	tɕi <sup>1</sup>	ke <sup>1</sup>	kɛi <sup>1</sup>
4 liver	*kon <sup>1</sup>	kan <sup>1</sup>	kø <sup>1</sup>	kon <sup>1</sup>	kon <sup>1</sup>	kon <sup>1</sup>
5 deep	*shim <sup>1</sup>	ʂən <sup>1</sup>	sən <sup>1</sup>	sən <sup>1</sup>	ts <sup>h</sup> əm <sup>1</sup>	səm <sup>1</sup>
<i>Tone 2</i>						
6 skin	*bi <sup>2</sup>	p <sup>h</sup> i <sup>2</sup>	bi <sup>2</sup>	p <sup>h</sup> i <sup>2</sup>	p <sup>h</sup> i <sup>2</sup>	p <sup>h</sup> ei <sup>2</sup>
7 come	*loi <sup>2</sup>	lai <sup>2</sup>	le <sup>2</sup>	lai <sup>2</sup>	loi <sup>2</sup>	lɛi <sup>2</sup>
8 flow	*liou <sup>2</sup>	liu <sup>2</sup>	ly <sup>2</sup>	liu <sup>2</sup>	liu <sup>2</sup>	lɛu <sup>2</sup>
9 cow	*ŋiou <sup>2</sup>	niu <sup>2</sup>	nɻ <sup>2</sup>	niu <sup>2</sup>	niu <sup>2</sup>	ŋɛu <sup>2</sup>
10 long	*jioŋ <sup>2</sup>	tʂ <sup>h</sup> aŋ <sup>2</sup>	zaŋ <sup>2</sup>	ts <sup>h</sup> oŋ <sup>2</sup>	ts <sup>h</sup> oŋ <sup>2</sup>	ts <sup>h</sup> æŋ <sup>2</sup>
<i>Tone 3</i>						
11 water	*shyi <sup>3</sup>	ʂui <sup>3</sup>	sɥ <sup>3</sup>	suɪ <sup>3</sup>	suɪ <sup>3</sup>	sæy <sup>3</sup>
12 fire	*huo <sup>3</sup>	xuo <sup>3</sup>	həu <sup>3</sup>	fo <sup>3</sup>	fo <sup>3</sup>	fo <sup>3</sup>
13 hand	*shiou <sup>3</sup>	ʂou <sup>3</sup>	sɻ <sup>3</sup>	səu <sup>3</sup>	su <sup>3</sup>	sɛu <sup>3</sup>
14 early	*tsou <sup>3</sup>	tsau <sup>3</sup>	tsæ <sup>3</sup>	tsau <sup>3</sup>	tsau <sup>3</sup>	tsou <sup>3</sup>
15 wait	*teŋ <sup>3</sup>	təŋ <sup>3</sup>	tən <sup>3</sup>	tɛn <sup>3</sup>	tɛn <sup>3</sup>	tɛŋ <sup>3</sup>
<i>Tone 4</i>						
16 rise	*zhioŋ <sup>4</sup>	ʂaŋ <sup>5</sup>	zɛŋ <sup>6</sup>	soŋ <sup>6</sup>	soŋ <sup>1</sup>	sæŋ <sup>4</sup>
17 go down	*hha <sup>4</sup>	ɕia <sup>5</sup>	ɦo <sup>6</sup>	ha <sup>6</sup>	ha <sup>1</sup>	ha <sup>4</sup>
18 heavy	*jiuŋ <sup>4</sup>	tsuŋ <sup>3</sup>	zoŋ <sup>6</sup>	ts <sup>h</sup> iuŋ <sup>6</sup>	ts <sup>h</sup> iuŋ <sup>1</sup>	ts <sup>h</sup> iuŋ <sup>4</sup>
19 sit	*dzuo <sup>4</sup>	tsuo <sup>5</sup>	zəu <sup>6</sup>	ts <sup>h</sup> o <sup>6</sup>	ts <sup>h</sup> o <sup>1</sup>	ts <sup>h</sup> o <sup>4</sup>
20 near	*gin <sup>4</sup>	tɕin <sup>5</sup>	dzin <sup>6</sup>	tɕ <sup>h</sup> in <sup>6</sup>	k <sup>h</sup> iu <sup>1</sup>	k <sup>h</sup> ɛn <sup>4</sup>
<i>Tone 5</i>						
21 go	*khie <sup>5</sup>	tɕ <sup>h</sup> y <sup>5</sup>	tɕ <sup>h</sup> i <sup>5</sup>	tɕ <sup>h</sup> iɛ <sup>5</sup>	hi <sup>5</sup>	hæy <sup>5</sup>
22 broken	*pho <sup>5</sup>	p <sup>h</sup> uo <sup>5</sup>	p <sup>h</sup> u <sup>5</sup>	p <sup>h</sup> o <sup>5</sup>	p <sup>h</sup> o <sup>5</sup>	p <sup>h</sup> o <sup>5</sup>
23 four	*si <sup>5</sup>	sɪ <sup>5</sup>	sɪ <sup>5</sup>	sɪ <sup>5</sup>	si <sup>5</sup>	sei <sup>5</sup>
24 half	*pon <sup>5</sup>	pan <sup>5</sup>	pø <sup>5</sup>	pon <sup>5</sup>	pan <sup>5</sup>	pun <sup>5</sup>
25 laugh	*siau <sup>5</sup>	ɕiau <sup>5</sup>	siæ <sup>5</sup>	ɕiɛu <sup>5</sup>	siau <sup>5</sup>	siu <sup>5</sup>
<i>Tone 6</i>						
26 sick	*bian <sup>6</sup>	piŋ <sup>5</sup>	beŋ <sup>6</sup>	p <sup>h</sup> iaŋ <sup>6</sup>	p <sup>h</sup> iaŋ <sup>5</sup>	pɛŋ <sup>6</sup>
27 graph	*dzi <sup>6</sup>	tsɪ <sup>5</sup>	zɪ <sup>6</sup>	ts <sup>h</sup> ɪ <sup>6</sup>	tsɪ <sup>5</sup>	tsi <sup>6</sup>
28 old	*giou <sup>6</sup>	tɕiu <sup>5</sup>	dzi <sup>6</sup>	tɕ <sup>h</sup> iu <sup>6</sup>	k <sup>h</sup> iu <sup>5</sup>	keu <sup>6</sup>
29 tree	*zhy <sup>6</sup>	ʂu <sup>5</sup>	zɥ <sup>6</sup>	ɕy <sup>6</sup>	su <sup>5</sup>	sy <sup>6</sup>
30 two	*nhi <sup>6</sup>	ə <sup>5</sup>	ni <sup>6</sup>	ə <sup>6</sup>	ni <sup>5</sup>	ji <sup>6</sup>
<i>Tone 7</i>						
31 blood	*hioɪ <sup>7</sup>	ɕiɛ <sup>3</sup>	ɕyɻ <sup>7</sup>	ɕyot <sup>7</sup>	hiat <sup>7</sup>	hyɪ <sup>7</sup>
32 seven	*tshit <sup>7</sup>	tɕ <sup>h</sup> i <sup>1</sup>	tɕ <sup>h</sup> iɪ <sup>7</sup>	tɕ <sup>h</sup> iu <sup>7</sup>	ts <sup>h</sup> iu <sup>7</sup>	ts <sup>h</sup> ɛ <sup>7</sup>
33 bamboo	*ciuk <sup>7</sup>	tsu <sup>2</sup>	tso <sup>7</sup>	tsuk <sup>7</sup>	tsuk <sup>7</sup>	tsuk <sup>7</sup>
34 go out	*chiut <sup>7</sup>	ts <sup>h</sup> iu <sup>1</sup>	ts <sup>h</sup> yɻ <sup>7</sup>	ts <sup>h</sup> ɛ <sup>7</sup>	ts <sup>h</sup> iu <sup>7</sup>	ts <sup>h</sup> ɛ <sup>7</sup>
35 nail	*kap <sup>7</sup>	tɕia <sup>3</sup>	ka <sup>7</sup>	kat <sup>7</sup>	kap <sup>7</sup>	kap <sup>7</sup>
<i>Tone 8</i>						
36 month	*ŋioɪ <sup>8</sup>	yɛ <sup>5</sup>	ŋyɻ <sup>8</sup>	ɲyot <sup>8</sup>	ɲyat <sup>8</sup>	ɲyt <sup>8</sup>
37 white	*bak <sup>8</sup>	pai <sup>2</sup>	bɛɻ <sup>8</sup>	p <sup>h</sup> ak <sup>8</sup>	p <sup>h</sup> ak <sup>8</sup>	pak <sup>8</sup>
38 ten	*zhip <sup>8</sup>	ʂɪ <sup>2</sup>	zyɻ <sup>8</sup>	sə <sup>8</sup>	səp <sup>8</sup>	sɛp <sup>8</sup>
39 stone	*zhiak <sup>8</sup>	ʂɪ <sup>2</sup>	zɛɻ <sup>8</sup>	sak <sup>8</sup>	sak <sup>8</sup>	sɛk <sup>8</sup>
40 mat	*ziak <sup>8</sup>	ɕi <sup>2</sup>	zil <sup>8</sup>	tɕ <sup>h</sup> iak <sup>8</sup>	sit <sup>8</sup>	tsɛk <sup>8</sup>

Notes: The Common Chinese reconstructions are provisional; nonetheless, I think they encapsulate what information we have on non-Min dialects. The word 'cow' appears to have an alternate CC form, \*niou. No. 35 'nail' refers to finger- and toenails.

**TABLE 5.2** DIAGNOSTIC LIST FOR IDENTIFYING SINITIC LANGUAGES (II). LESSER KNOWN LANGUAGES

	<i>CC</i>	<i>Jy</i>	<i>Hp</i>	<i>Zq</i>	<i>Jc</i>	<i>Js</i>
1 sky	* <i>thian</i>	<i>tʰəŋ¹</i>	<i>hən¹</i>	<i>tʰiŋ¹</i>	<i>xɛ²</i>	<i>tʰiɛn¹</i>
2 three	* <i>sam</i>	<i>soŋ¹</i>	<i>sam¹</i>	<i>saŋ¹</i>	<i>sɑ²</i>	<i>saŋ¹</i>
3 chicken	* <i>kiai</i>	<i>tɕi¹</i>	<i>kei¹</i>	<i>kai¹</i>	<i>ke¹</i>	<i>iɛ¹</i>
4 liver	* <i>kon</i>	<i>kaŋ¹</i>	<i>hon¹</i>	<i>huaiŋ¹</i>	<i>kɑ²</i>	<i>kəŋ¹</i>
5 deep	* <i>shim</i>	<i>ɕie¹</i>	<i>tɕʰim¹</i>	<i>tsʰeŋ¹</i>	<i>sɛ¹</i>	<i>tsʰøn¹</i>
6 skin	* <i>bi</i>	<i>pø²</i>	<i>pʰui²</i>	<i>pʰue²</i>	<i>pe⁸</i>	<i>pʰei²</i>
7 come	* <i>loi</i>	<i>lø²</i>	<i>li²</i>	<i>lœ²</i>	<i>yu⁴</i>	<i>li²</i>
8 flow	* <i>liou</i>	<i>liou²</i>	<i>liu²</i>	<i>lo²</i>	<i>ku⁸</i>	<i>lu²</i>
9 cow	* <i>ŋiou</i>	<i>ŋou²</i>	<i>ny²</i>	<i>niu²</i>	<i>ŋu⁸</i>	<i>ŋuɔ²</i>
10 long	* <i>jioŋ</i>	<i>tɕiaŋ²</i>	<i>hoŋ²</i>	<i>tauŋ⁹</i>	<i>tsø⁸</i>	<i>tʰæ²</i>
11 water	* <i>shyi</i>	<i>ɕya³</i>	<i>sui³</i>	<i>sui³</i>	<i>ɕui¹</i>	<i>yi³</i>
12 fire	* <i>huo</i>	<i>fu³</i>	<i>kʰuei³</i>	<i>hue³</i>	<i>xui¹</i>	<i>xuei³</i>
13 hand	* <i>shiou</i>	<i>ɕiou³</i>	<i>ɕiu³</i>	<i>siu³</i>	<i>su¹</i>	<i>tɕʰye³</i>
14 early	* <i>tsou</i>	<i>tsau³</i>	<i>tʰoɽ⁴</i>	<i>tsø³</i>	<i>tsu¹</i>	<i>tsiau³</i>
15 wait	* <i>teŋ</i>	<i>lai³</i>	<i>ten³</i>	<i>teuŋ³</i>	<i>tu¹</i>	<i>tæ³</i>
16 rise	* <i>zhioŋ</i>	<i>ɕiaŋ⁴</i>	<i>ɕioŋɽ⁴</i>	<i>ioŋ⁵</i>	<i>tsø¹</i>	<i>dziaŋ⁴</i>
17 go down	* <i>hha</i>	<i>fuɔ⁴</i>	<i>haɽ⁴</i>	<i>ha⁵</i>	<i>tʰu²</i>	<i>ɔ⁴</i>
18 heavy	* <i>jiuŋ</i>	<i>tɕiaŋ⁴</i>	<i>hŋɽ⁴</i>	<i>toŋ⁶</i>	<i>tsv¹</i>	<i>dziuŋ⁴</i>
19 sit	* <i>dzuo</i>	<i>tsəu⁴</i>	<i>tʰoiɽ⁴</i>	<i>tsuai⁵</i>	<i>kv⁷</i>	<i>si⁴</i>
20 near	* <i>gin</i>	<i>tɕie⁴</i>	<i>kʰuanɽ⁴</i>	<i>kyeŋ⁵</i>	<i>tɕi¹</i>	<i>gɛ⁴</i>
21 go	* <i>khie</i>	<i>hu⁵</i>	<i>kʰo⁵</i>	<i>kʰo⁵</i>	<i>ŋe⁸</i>	<i>kʰuɔ⁵</i>
22 broken	* <i>pho</i>	<i>pʰu⁵</i>	<i>pʰai⁵</i>	<i>pʰua⁵</i>	<i>pʰo³</i>	<i>pʰai⁵</i>
23 four	* <i>si</i>	<i>sa⁵</i>	<i>si⁵</i>	<i>si⁵</i>	<i>ɕi⁵</i>	<i>si⁵</i>
24 half	* <i>pon</i>	<i>paŋ⁵</i>	<i>pon⁵</i>	<i>puaiŋ⁵</i>	<i>po⁸</i>	<i>piɛn⁵</i>
25 laugh	* <i>siau</i>	<i>siu⁵</i>	<i>sieu⁵</i>	<i>sio⁵</i>	<i>so³</i>	<i>tɕʰiau⁵</i>
26 sick	* <i>biaŋ</i>	<i>pioŋ⁶</i>	<i>kʰue⁵</i>	<i>paŋ⁶</i>	<i>pɛ³</i>	<i>ban⁶</i>
27 graph	* <i>dzi</i>	<i>tsuɔ⁶</i>	<i>tʰe⁶</i>	<i>tsi⁹</i>	<i>tsu³</i>	<i>dzuɔ⁶</i>
28 old	* <i>giou</i>	<i>tɕiou⁶</i>	<i>kʰy⁶</i>	<i>kiu⁶</i>	<i>ku³</i>	<i>gu⁶</i>
29 tree	* <i>zhy</i>	<i>ɕyu⁶</i>	<i>tɕʰy⁵</i>	<i>tsʰiu⁶</i>	<i>tsu³</i>	<i>dzuɔ⁶</i>
30 two	* <i>nhi</i>	<i>na⁶</i>	<i>ni⁶</i>	<i>ni⁶</i>	<i>ne⁵</i>	<i>ni⁶</i>
31 blood	* <i>hiot</i>	<i>ɕy⁷</i>	<i>fæ⁷</i>	<i>huai³</i>	<i>suɑ⁵</i>	<i>xyɛɽ⁷</i>
32 seven	* <i>tshit</i>	<i>tsʰa⁷</i>	<i>tʰi⁷</i>	<i>tsʰi³</i>	<i>tɕʰi⁵</i>	<i>tɕʰəɽ⁷</i>
33 bamboo	* <i>ciuk</i>	<i>liou⁷</i>	<i>ty⁷</i>	<i>ty³</i>	<i>ke⁵</i>	<i>ta⁷</i>
34 go out	* <i>chiut</i>	<i>ɕya⁷</i>	<i>tsʰui⁷</i>	<i>tsʰue³</i>	<i>tɕʰi⁵</i>	<i>tɕʰyɛɽ⁷</i>
35 nail	* <i>kap</i>	<i>kuɔ⁷</i>	<i>kam⁷</i>	<i>ka³</i>	<i>ke⁵</i>	<i>kaɽ⁷</i>
36 month	* <i>ŋiot</i>	<i>ŋy⁶</i>	<i>væɽ⁴</i>	<i>ŋye⁵</i>	<i>ŋua⁵</i>	<i>ŋuoɽ⁸</i>
37 white	* <i>bak</i>	<i>puɔ⁶</i>	<i>pʰaɽ⁴</i>	<i>pa⁶</i>	<i>pɛ⁷</i>	<i>baɽ⁸</i>
38 ten	* <i>zhip</i>	<i>suɔ⁶</i>	<i>ɕimɽ⁴</i>	<i>tsi⁶</i>	<i>tsɛ⁷</i>	<i>søɽ⁸</i>
39 stone	* <i>zhiak</i>	<i>ɕyɔ⁶</i>	<i>ɕioɽ⁴</i>	<i>tsio⁶</i>	<i>tsø⁷</i>	<i>siɛɽ⁸</i>
40 mat	* <i>ziak</i>	<i>tsie⁶</i>	<i>tʰiaɽ⁴</i>	<i>sio⁶</i>	<i>pʰu³</i>	<i>siɛɽ⁸</i>

*Note:* The ‘Lesser known languages’ part could be left out, if wanted for space reasons.

- 7 The *qù* tone lacks a register distinction.
- 8 The verb ‘to wear (clothing)’ is *chuān*, or cognate to it.
- 9 The word for ‘(cooking) pot’ is *guō*, or cognate to it.
- 10 The word for ‘house’ is *fāng(zi)*, or cognate to it.
- 11 The word for ‘son’ is *ér(zi)*, or cognate to it.
- 12 The word for ‘stand’ is *zhàn*, or cognate to it.

- 13 The verb in the expression ‘to rain’ is *xià*, or cognate to it.  
 14 The verb for ‘to walk’ is *zǒu*, or cognate to it.  
 15 The gender marker for animals is prefixed.

The dialects given on Table 5.3 are Běijīng (Bj), Tàiyuán (Ty), Yángzhōu (Yz); these three dialects belong to the Mandarin group. Sūzhōu (Sz) and Wēnzhōu are Wú dialects. Chángshā (Cs) and Shuāngfēng are considered to be Xiāng dialects. Nánchāng represents the Gàn group. Méixiàn is a Hakka dialect. Guǎngzhōu is the best known of the Yuè dialects. Jiāngyǒng is a so-called Tǔ dialect, a group of Xiāng dialects spoken an area where Southwest Mandarin is also spoken, mainly in southern Húnán. Sources of data are Běijīng Dàxué 1995 and for Jiāngyǒng, Huáng (1993).

There are some cases where both a plus and minus value are given; this means that there are competing forms; one of the forms usually reflects a Mandarin origin and very likely represents a late borrowing. In such cases, I consider the minus value more basic. The dialects of Běijīng, Tàiyuán, and Yángzhōu all have fifteen plus values. This state of affairs defines the Northern or Mandarin zone. Some Mandarin dialects may have one or two negative values, but not more. The dialects of Méixiàn, Guǎngzhōu and Jiāngyǒng have either all minus values or only one plus. These dialects belong to the Southern zone. The remaining dialects all have a mixture of plus and minus values; these dialects comprise the Central zone. Dialects in the Central zone originally most likely were of a more Southern type, but centuries of Northern influence have given rise to the situation that we see today. This division of Chinese dialects into zones is not intended to be a classification *per se*, but a framework within which to discuss classification.

Before proceeding to a discussion of the various groups of dialects, it should be made clear that Chinese is not entirely amenable to a Stammbaum formulation. The fact that these dialects have been spoken in close proximity to one another for two millennia and the pervasive influence of various quasi-standards and koinés on all Chinese dialects over a very long period easily obscures underlying relationships. Nonetheless, to go the other extreme and claim that Chinese dialects cannot be classified in any meaningful way would be equally

**TABLE 5.3** THE DIVISION OF CHINESE DIALECTS: NORTHERN, CENTRAL, AND SOUTHERN

	Bj	Ty	Yz	Sz	Wz	Cs	Sf	Nc	Mx	Gz	Jy
1	+	+	+	-	-	+	+	-	-	-	-
2	+	+	+	-	-	-	-	-	-	-	-
3	+	+	+	+	+	+	+	-	-	-	-
4	+	+	+	+	+	+	+	+	-	-	+
5	+	+	+	-	-	+	-	-	-	-	-
6	+	+	+	+	-	+	+	+	-	-	-
7	+	+	+	-	-	-	-	-	+	-	-
8	+	+	+	-	-	+	+	-	-	-	-
9	+	+	+	-	-	+	+	+	-	-	-
10	+	+	+	+	-	-	-	-	-	-	-
11	+	+	+	+	+	-	-	-	-	-	-
12	+	+	+	-	-	±	-	-	-	-	-
13	+	+	+	-	-	-	-	-	-	-	-
14	+	+	+	±	±	+	-	+	-	-	-
15	+	+	+	+	±	-	-	-	-	-	-

Notes: (1) Our Jiāngyǒng source has *təu*<sup>6</sup> for the third person pronoun; both the initial and the tone argue against its being cognate to *tā* [thɛ]. (2) Wú dialects have a labiodental initial (*Sz fɔʔ*<sup>7</sup>); I consider such forms cognate to *bù*.

erroneous. Below I will discuss each of the major groups in turn, paying especial attention to how each group is delimited from the others.

Mandarin dialects are spoken by more than 70 per cent of all Chinese speakers. Dialects of this group are spoken everywhere north of the Yangtze River as well as in the southwestern provinces of Sichuān, Yúnnán and Guìzhōu, and in parts of Guǎngxī and Húnán. Looked at historically, the Mandarin dialects represent a radical transformation. Phonologically the Common Chinese voiced obstruents have all become voiceless, a feature also found in several other dialect groups. The original Chinese tonal system has in most cases been drastically altered and the vocalism has become simplified (Norman 1999). In addition, Mandarin dialects possess an innovative lexical pattern, much of which is of rather recent provenance.

The variety of Chinese reflected in early lexical sources like the *Qièyùn* and the *Jingdiǎn shìwén* had a three-way distinction of obstruent initials: aspirated and unaspirated voiceless as well as a series which is widely recognized as having been voiced in some sense. One of the most salient features of the Chinese phonological series is the devoicing of this third series in a majority of dialects, the exception being the Wú dialects and some varieties of the Xiāng group. In Mandarin dialects the most common pattern is the one in which the earlier voiced obstruents become voiceless and aspirated in the *píng* tonal category and voiceless and unaspirated in the other tonal categories. This pattern is found in the vast majority of Mandarin dialects and is found almost everywhere in major political and cultural centres. However, it is not universal by any means. There are a number of dialects in Shānxī which do not exhibit this pattern. Línfén, for example, has aspirated initials in non-*píng* tones and Tàigǔ shows non-aspirates in the *píng* tone (Wēn and Hóu 1993).

Medieval sources like the early seventh century *Qièyùn* rhyme dictionary reflect a stage of Chinese in which there were four tonal categories: *píng*, *shàng*, *qù* and *rù*. No modern dialect has preserved such a system intact; in all known cases, there is evidence that each of the medieval categories has split into two contrasting registers; one is from voiceless initials (the upper or *yīn* register) and one from voiced initials (the lower or *yáng* register); this process yielded a system of eight tonal categories, conventionally designated as shown below:

TABLE 5.4 THE REGISTER SPLIT IN THE PÍNG, SHÀNG, QÙ, AND RÙ TONES

upper:	1 <i>yīnpíng</i>	3 <i>yīnshàng</i>	5 <i>yīnqù</i>	7 <i>yīnrù</i>
lower:	2 <i>yángpíng</i>	4 <i>yángshàng</i>	6 <i>yángqù</i>	8 <i>yángrù</i>

The merger of tone four with the *qù* tone category is universal in Mandarin dialects. Words with obstruent initials initially merged with the *yángqù* and words with sonorant initials merged with the *yīnshàng*. Subsequently in the vast majority of dialects, there was a merger of the upper and lower *qù* tones, but here and there conservative dialects can be found that keep the two *qù* tones separate. It is in regard to the *rù* tones that Mandarin dialects differ most dramatically. There is a basic division among dialects that retain the *rù* category and those in which it has disappeared. Many dialects in Shānxī and surrounding areas have maintained this category. Another large group of dialects that retains the *rù* tone is the area between the Jiāng (Yangtze) and Huái rivers. Another large group of dialects lack a *rù* tone category; this is the case with a majority of the Southwestern dialects. In these dialects, *rù* tone words merge with the *yángpíng* category. In the north, dialects without a separate *rù* tone category fall into several types. Central Plains dialects exhibit a split in which *rù* tone words divide according to initial category: words with original voiceless and sonorant initials merge with tone one (*yīnpíng*) while words with voiced obstruent initials go to tone two (*yángpíng*).

In the closely related Lányín dialects, words with voiceless and sonorant initials are found in tone five (*yīnqù*), but words with voiced obstruent initials in tone two (*yángpíng*). Three Northern groups – Běijīng, Jīnlǚ and Jiāoliáo – have in common a shift of words with sonorant initials into tone five and words with voiced obstruent initials into tone two. Words with voiceless initials show three different patterns: in the Jīnlǚ group, they merge into tone one, in the Jiāoliáo dialects with tone three, while in the Běijīng group such words are distributed into the remaining tonal categories without any clear condition for the split.

On the basis of *rù* tone development we can divide Mandarin dialects into three large groups and several subgroups:

Group 1 consists of dialects which retain the *rù* tone. Such dialects are found in several non-contiguous regions, the chief of them being Shānxī and the Lower Yangtze region.

Group 2 consists of dialects in which *rù* tone words merge unconditionally with the *yángpíng* category. Such dialects are found in the southwest in the provinces of Sìchuān, Yúnnán, Guìzhōu and parts of Húběi, Húnán, and Guǎngxī.

Group 3 embraces dialects in which the *rù* category has become split according to initial types. On the basis of merger patterns these dialects can be divided into several subtypes as described above.

The prototypical Mandarin dialect has lost tonal category four (*yángshàng*) and has merged categories five and six (the two *qù* tones). The development of *rù* tone words exhibits different patterns, allowing us to create a reasonably neat subclassification of this largest of Chinese dialect groups. It is probably not possible to define Mandarin with a single phonological feature. In Table 5.3 we see, however, how a set of features can be used successfully to set off Mandarin from other dialect groups. Notice that features used are weighted in favour of lexicon; this approach reflects the fact that the Mandarin dialects on the whole have a very distinctive set of vocabulary traits, many of which, judging from the historical record, are of rather recent origin. One can almost define Mandarin as a group of dialects which use *tā* as the third person pronoun along with *de* as the subordinative particle. Certainly these two features along with a set of other lexical items ('wear', 'pot', 'house', 'son', 'stand', 'rain', and 'walk') are a sufficient basis for classifying a dialect as Mandarin.<sup>1</sup>

Wú dialects are very different phonologically from Mandarin. The most notable typological feature is the retention of a distinctive set of voiced or murmured initials. But one must be careful not to treat this feature as the defining characteristic of Wú since voiced (or murmured) initials are also found in other dialect groups. The feature that most clearly sets Wú off from Mandarin is its tonal pattern; typically, all tonal categories are divided into two registers. The more conservative Wú dialects maintain a distinction between the upper and lower *shàng* tones – for example, Wúxī, Chángshú and Wēnzhōu. More innovative dialects (Shànghǎi, Chángzhōu, and Sūzhōu) have merged the *yángshàng* with *yángqù*. But even a very innovative dialect like Shànghǎi still maintains the register system very faithfully in that only in very rare cases do upper and lower register tones merge. The tonal conservatism of Wú serves to divide it from the those Xiāng dialects in which voiced obstruent initials are found but in which there is no *yángshàng* category nor is there a register distinction for the *rù* category which is usual in Wú dialects.

Lexically, Wú dialects generally lack the Mandarin innovations mentioned above. Table 5.5 compares some basic features of the Wú dialects with corresponding Mandarin forms.

---

1 For the classification of Mandarin dialects see Lǐ (1985), Simmons (1999a), and Ho (this volume).

TABLE 5.5 LEXICAL DIFFERENCES BETWEEN MANDARIN AND WU DIALECTS

	<i>Bj</i>	<i>Ty</i>	<i>Yz</i>	<i>Sz</i>	<i>Cm</i>	<i>Wl</i>	<i>Wz</i>
1	<i>t<sup>h</sup>a<sup>1</sup></i>	<i>t<sup>h</sup>a<sup>1</sup></i>	<i>t<sup>h</sup>a<sup>1</sup></i>	<i>li<sup>1</sup></i>	<i>fi<sup>2</sup></i>	<i>gie<sup>2</sup></i>	<i>gei<sup>2</sup></i>
2	<i>tə<sup>0</sup></i>	<i>tə<sup>0</sup></i>	<i>ti<sup>0</sup></i>	<i>ky<sup>0</sup></i>	<i>go<sup>0</sup></i>	<i>kie<sup>0</sup></i>	<i>ge<sup>0</sup></i>
3	<i>pu<sup>5</sup></i>	<i>pə<sup>7</sup></i>	<i>pə<sup>7</sup></i>	<i>fY<sup>7</sup></i>	<i>fə<sup>7</sup></i>	<i>və<sup>8</sup></i>	<i>fu<sup>3</sup></i>
4	<i>tʂ<sup>h</sup>uan<sup>1</sup></i>	<i>ts<sup>h</sup>uə<sup>1</sup></i>	<i>ts<sup>h</sup>uə<sup>1</sup></i>	<i>tsD<sup>7</sup></i>	<i>tsa<sup>7</sup></i>	<i>tciə<sup>7</sup></i>	<i>tciə<sup>7</sup></i>
5	<i>kuo<sup>1</sup></i>	<i>ku<sup>1</sup></i>	<i>ko<sup>1</sup></i>	<i>fiə<sup>8</sup></i>	<i>fiuo<sup>8</sup></i>	<i>fiuo<sup>8</sup></i>	<i>fiə<sup>8</sup></i>
6	<i>fan<sup>2</sup></i>	<i>fə<sup>2</sup></i>	<i>fan<sup>2</sup></i>	<i>vDŋ<sup>2</sup></i>	<i>vā<sup>2</sup></i>	<i>u<sup>7</sup></i>	<i>u<sup>7</sup></i>
7	<i>æ<sup>2</sup></i>	<i>ə<sup>2</sup></i>	<i>a<sup>2</sup></i>	<i>pi<sup>2</sup></i>	<i>n<sup>2</sup></i>	<i>fi<sup>2</sup></i>	<i>ŋ<sup>2</sup></i>
8	<i>tʂan<sup>5</sup></i>	<i>tsə<sup>5</sup></i>	<i>tsə<sup>5</sup></i>	<i>li<sup>8</sup></i>	<i>gei<sup>4</sup></i>	<i>dzi<sup>2</sup></i>	<i>ge<sup>4</sup></i>
9	<i>tsou<sup>3</sup></i>	<i>tsəu<sup>3</sup></i>	<i>tsy<sup>3</sup></i>	<i>bə<sup>2</sup></i>	<i>bə<sup>2</sup></i>	<i>tsy<sup>3</sup></i>	<i>fiə<sup>2</sup></i>

Notes: The lexical forms are (1) third person pronoun, (2) subordinative particle, (3) unmarked negative, (4) 'wear', (5) 'pot', (6) 'house', (7) 'son', (8) 'stand', (9) 'walk'.

In addition to Sūzhōu and Wēnzhōu, forms are given for Chóngmíng (Cm) and Wēnlíng (Wl); forms for the latter two dialects were kindly provided by Professor Zhāng Huìyīng and Professor Lǐ Róng respectively.

Table 5.5 highlights some of the lexical differences between Mandarin and Wú. Where the Mandarin dialects have an innovative lexicon, Wú reflects an older lexical stratum. A pattern similar to that of Wú is found throughout the Central dialects as well as in Hakka and Yuè. Only Mǐn has a different lexical profile.<sup>2</sup>

The framework proposed above places Wú, Gàn and Xiāng in the Central zone; the dialects of this zone, while clearly not Mandarin, have obviously absorbed a certain number of words of Mandarin origin in their historical development.

Gàn speakers comprise around 3 per cent of all Han peoples. The Gàn dialects are mostly spoken in north and central Jiāngxī province and in a few adjacent areas in Húnán, Húběi and Fújiàn. Some scholars believe that no meaningful distinction can be made between Gàn and Hakka; this question will be addressed below. In Gàn dialects (with only a few exceptions) the old voiced obstruents have become voiceless aspirates in all tonal categories, but this is clearly not a sufficient reason for classifying a dialect as Gàn, as we have seen above, because there are dialects clearly belonging to other groups that show this feature. Gàn dialects have between four and eight tones. The dialect of Dūchāng is reported to have ten tones but this is due to splits in several tones based on the presence of aspiration. A few dialects keep the upper and lower *shàng* tones separate; the more general pattern merges the lower *shàng* with original voiced obstruent initials with the lower *qū*. Some Gàn dialects have a few sporadic cases of the lower *shàng* going to the upper *píng* tone, as in Hakka.

The Gàn lexicon shows the typical Central zone profile. One prominent difference from Wú is the almost universal use of reflexes of Common Chinese \**tsoi*<sup>3</sup> for 'son' (where Wú uses reflexes of Common Chinese \**nhi*<sup>2</sup>). The Gàn unmarked negative is from \**pu*<sup>7</sup> whereas Wú dialects have forms with labiodental initials (see Table 5.5 for examples). It is difficult to find a small number of features by which the Gàn dialects can be uniquely characterized. While they clearly differ from Wú and Xiāng in the way they treat the old voiced obstruents, it is more difficult to keep them separate from Hakka where the two groups show the same development with regard to this feature. However, as we have seen, there are problems in using this criterion to define the major dialect groups.

2 On Wú classification see Yú (1999).

The Hakkas comprise about 3.5 per cent of the Hàn population. They are found to the south of Gàn in Jiāngxī province, in northern Guǎngdōng, southwestern Fújiàn, as well as in many scattered communities in south China. In addition to the development of the old voiced obstruents into modern voiceless aspirates in all tonal categories mentioned already, another striking Hakka feature is the shift of many lower *shàng* words into the upper *píng*. This feature is also found in some Gàn dialects in words with obstruent initials, but it is extremely rare in words with sonorant initials, while in Hakka dialects there are numerous cases of such a shift. In terms of lexicon, there are a few important differences between Gàn and Hakka. One is the use of *\*m<sup>2</sup>* rather than *\*pu<sup>7</sup>* for the unmarked negative and *\*zhik<sup>8</sup>* for ɣat rather than *\*chiak<sup>7</sup>*; the latter form is common in Northern and Central dialects. Another Hakka feature that distinguishes it from Gàn in most cases is the use of *\*hhei* for the copula (as in the Yuè dialects). In the final analysis, however, it is often difficult to draw a sharp line between Gàn and Hakka dialects. The dialect of Jiànníng in western Fújiàn, for example, has a reflex of *\*m<sup>2</sup>* for 'not' and *\*zhik<sup>8</sup>* for 'eat', as in Hakka, but uses *\*zhi<sup>4</sup>* for the copula, as in Gàn. Jiànníng also has very few examples of the lower *shàng* going to the upper *píng* in words with obstruent initials and no examples of the shift in question with sonorant initials. There are clearly some dialects that are transitional between Gàn and Hakka but I would maintain that on the whole the distinction is worth retaining, especially since the Hakka dialects clearly fall into my Southern zone while Gàn just as clearly belongs to the Central zone.<sup>3</sup>

The third dialect group found in the Central zone is the Xiāng group. Approximately 3 per cent of the Hàn population speak Xiāng dialects. Here two subtypes are generally recognized: one group, called New Xiāng, has lost the old voiced obstruents altogether, while the second group, Old Xiāng, retains voicing to some degree. In the New Xiāng dialects, the voiced obstruents have become voiceless non-aspirates in all tonal categories. Thus, where Shuāngfēng (an Old Xiāng dialect) has *dzu<sup>6</sup>* for 'sit', Chángshā (a New Xiāng dialect) has *tso<sup>6</sup>*. Once more, it is doubtful that the development of the old voiced obstruents is a sufficient reason to classify a dialect as Xiāng. However, this feature does distinguish them neatly from Wú and Gàn. But what about Xiāng and Mandarin? The behaviour of the old voiced obstruents is also, on the whole, a fairly reliable way to distinguish these two groups, but all Xiāng dialects (as far as we know) show the shift of the *yángshàng* to the (lower) *qù*. Unlike Wú, Xiāng almost never has both an upper and lower *rù* tone. A distinction between the upper and lower *qù* is common but not universal. The New Xiāng dialects, especially that of the provincial capital, Chángshā, already seem to be highly mandarinized. Nonetheless, in terms of lexicon, Xiāng dialects show numerous non-Mandarin features, suggesting that the mandarinization found in the region may be of rather recent origin. This group of dialects needs to be studied much more before anything definitive can be established.

Yuè dialects are spoken by about 4 per cent of all Chinese speakers. The most salient feature of the Yuè dialects is their tonal system. The upper and lower *shàng* tones are mostly retained as separate categories, but, more importantly, the upper *rù* category splits in two with vowel length as the conditioning feature. It is worth noting that a similar split can be observed in many Tai dialects. Yuè dialects do not show a consistent pattern of devoicing in obstruent initials. In the paradigmatic dialect of Guǎngzhōu, one finds aspiration in the *píng* and *shàng* tones. A notable lexical peculiarity is the use of *na<sup>3</sup>* as the female gender marker for animals.

<sup>3</sup> Two important recent works on Gàn and Hakka are Lǐ and Zhāng (1992) and Sagart (1993). On the Gàn–Hakka problem see Norman (1989) and Yán (1986).

**TABLE 5.6** DOUBLE CORRESPONDENCES OF LOWER REGISTER TONES IN MÍN

	<i>Bj</i>	<i>Sz</i>	<i>Sf</i>	<i>Nc</i>	<i>Fz</i>	<i>Xm</i>	<i>Jo</i>
<i>tone two (yángpíng)</i>							
level	$p^h i \eta^2$	$bi n^2$	$bi \eta^2$	$p^h i a \eta^2$	$pa \eta^2$	$p \bar{i}^2$	$pi a \eta^5$
skin	$p^h i^2$	$bi^2$	$bi^2$	$p^h i^2$	$p^h ui^2$	$p^h e^2$	$p^h y \epsilon^5$
<i>tone four (yángshàng)</i>							
-fold	$pei^5$	$be^6$	$be^6$	$p^h i^6$	$pui^6$	$pe^6$	$po^4$
quilt	$pei^5$	$bi^6$	$bi^6$	$p^h i^6$	$p^h ui^6$	$p^h e^6$	$p^h y \epsilon^6$
<i>tone six (yángqù)</i>							
pace	$pu^5$	$bu^6$	$bu^6$	$p^h u^6$	$puo^6$	$p \mathcal{J}^6$	$pi \mathcal{J}^6$
nose	$pi^2$	$b y \mathcal{J}^8$	$bi^6$	$p^h i^8$	$p^h ei^5$	$p^h \tau^6$	$p^h i^6$
<i>tone eight (yángrù)</i>							
white	$pai^2$	$b \mathcal{D} \mathcal{J}^8$	$pia^2$	$p^h ak^8$	$pa \mathcal{J}^8$	$pe \mathcal{J}^8$	$pa^6$
hail	$pau^2$	$b \mathcal{D}^6$	$p^h y^6$	$p^h \mathcal{J}k^8$	$p^h \epsilon y \mathcal{J}^8$	$p^h au \mathcal{J}^8$	$p^h au^6$

Notes: ‘-fold’ as in three-fold, ten-fold, etc. Note that ‘nose’ comes from a *yáng rù* tone in Běijīng, Sūzhōu and Nánchāng while the other dialects reflect a *yáng qù* origin; likewise, ‘hail’ reflects a *yáng rù* tone in Běijīng, Nánchāng and Mǐn forms while Sūzhōu and Shuāngfēng reflect a *yáng qù* origin.

Yuè dialects appear to have the most in common with Hakka but are clearly distinguished from them by tonal development.<sup>4</sup>

It is generally recognized that the Mǐn dialects, spoken by over 5 per cent of all Chinese, lie outside the mainstream of Chinese linguistic development. The defining characteristic of Mǐn is the presence of lower register tones of both aspirated and unaspirated correspondences to the old voiced stops and affricates of Common Chinese. Where in non-Mǐn dialects there is a generally regular development of the voiced obstruents, Mǐn dialects show a double correspondence within a single tone category, one aspirated the other not. Moreover, the lexical incidence of aspiration or the lack thereof is highly consistent from one dialect to another. Table 5.6 illustrates this situation with voiced labial stops along with their reflexes in several non-Mǐn dialects.

The dialects used for illustration are Běijīng (*Bj*), Sūzhōu (*Sz*), Shuāngfēng (*Sf*), Nánchāng (*Nc*), Fúzhōu (*Fz*), Xiàmén (*Xm*) and Jiàn’ōu (*Jo*). All forms are taken from Běijīng Dàxué (1995).

Tonal systems in Mǐn dialects provide few distinctively Mǐn characteristics. The lower *shàng* survives in some Southern Mǐn dialects (Lóngyán and some Quánzhōu dialects, for example) and in the Cháozhōu region of Guǎngdōng.<sup>5</sup> In Xiàmén and all of the coastal region north of Fúzhōu it has merged with the lower *qù* as in Mandarin; whether this is due to Mandarin influence or is a parallel development is difficult to say. Tonal development in the Northwestern Mǐn dialects (Jiàn’ōu, Jiànyáng, Shàowǔ, etc.) is very complex due the necessity for reconstructing a three-way distinction of stops and affricates for these dialects (Norman 1974, 1996).

Another highly distinctive characteristic of the Mǐn dialects is their lexicon. While some of this peculiarly Mǐn vocabulary can be found sporadically elsewhere, it is doubtful that the entire ensemble can be found in any non-Mǐn dialect. Table 5.7 presents such an ensemble compared to forms found in dialects of other groups. (*Gz* stands for Guǎngzhōu.)

4 For Yuè classification see Yú (1991).

5 Branner (1999) is a good source for Lóngyán.



TABLE 5.7 A DIAGNOSTIC ENSEMBLE OF MIN LEXICAL ITEMS

	Sz	Sf	Nc	Gz	Fz	Xm	Jo
wear	tsɔʔ <sup>7</sup>	t <sup>h</sup> u <sup>1</sup>	ts <sup>h</sup> ɔn <sup>1</sup>	tsæk <sup>7</sup>	sæyŋ <sup>6</sup>	ts <sup>h</sup> iŋ <sup>6</sup>	tsæyŋ <sup>6</sup>
pot	fɔ ʔ <sup>8</sup>	ku <sup>1</sup>	uɔ <sup>1</sup>	wɔk <sup>8</sup>	tiaŋ <sup>3</sup>	tiā <sup>3</sup>	tiaŋ <sup>3</sup>
house	vɔŋ <sup>2</sup>	əu <sup>7</sup>	uk <sup>7</sup>	uk <sup>7</sup>	ts <sup>h</sup> uo <sup>5</sup>	ts <sup>h</sup> u <sup>5</sup>	ts <sup>h</sup> iɔ <sup>5</sup>
son	ŋi <sup>2</sup>	tse <sup>3</sup>	tsai <sup>3</sup>	tsai <sup>3</sup>	kian <sup>3</sup>	kiā <sup>3</sup>	kyiŋ <sup>3</sup>
walk	bæ <sup>2</sup>	ɣaŋ <sup>2</sup>	tseu <sup>3</sup>	haŋ <sup>2</sup>	kian <sup>2</sup>	kiā <sup>2</sup>	kian <sup>3</sup>
foot	tɕiɔʔ <sup>7</sup>	tʉ <sup>2</sup>	tɕiɔk <sup>7</sup>	kæk <sup>7</sup>	k <sup>h</sup> a <sup>1</sup>	k <sup>h</sup> a <sup>1</sup>	k <sup>h</sup> au <sup>1</sup>
mouth	tsɿ <sup>3</sup>	tɕy <sup>3</sup>	tsui <sup>3</sup>	tsæy <sup>3</sup>	ts <sup>h</sup> ui <sup>5</sup>	ts <sup>h</sup> ui <sup>5</sup>	ts <sup>h</sup> ui <sup>5</sup>

Note: Although the Min forms for ‘walk’ are in some sense cognate to the Sf and Gz forms, they possess a very distinctive phonological form.

Modern spoken Chinese is a collection of hundreds of dialects, the majority of which are to some degree mutually unintelligible. Along with Bái, they form a major subgroup within the Sino-Tibetan family. Historically speaking, the present-day Chinese dialects have their roots in the great Qín and Hàn imperial expansion (BC 221–AD 220) and thus have a time depth similar to that of the Romance languages, and, indeed, the degree of diversity among the Chinese dialects is in general similar to that found in the Romance family. Karlgren and others have placed the dialectalization of Chinese much later (eighth to tenth centuries AD) but this seems unrealistic given the degree of diversity found at present. The influence of a standard written language and various quasi-official koinés on all varieties of spoken Chinese has also been akin to the influence of Latin and later standards on the Romance languages. Although our knowledge of dialectal Chinese has grown enormously in the last fifty years, a vast amount of work still remains to be done before a reasonably accurate picture of their historical development can be drawn.

## REFERENCES

- Běijīng Dàxué (1995) *Hànyǔ fāngyán cǐhuì*, Beijing: Yuwen Chubanshe.
- Branner, David (1999) ‘The classification of Longyan’, in Simmons (ed.) (1999a).
- Forrest, R.A.D. (1948) *The Chinese Language*, London: Faber and Faber.
- Hú, Míngyáng (1987) *Běijīnghuà chū'àn*, Beijing: Shangwu.
- Huáng, Bùfán (ed.) (1992) *Zàng-Miǎn yǔzú yǔyán cǐhuì*, Beijing: Zhongyāng Mǐnzú Xuéyuan.
- Huáng, Xuězhēn (1993) *Jiāngyǒng fāngyán yánjiū*, Beijing: Shehui Kēxuéyuan.
- Lǐ, Róng (1985) ‘Guānhuà fāngyán de fēnqū’, *Fāngyán* 1: 2–5.
- Lǐ, Rúlóng and Zhāng, Shuāngqīng (1992) *Kè-Gàn fāngyán diàochá bàogào*, Xiāmen: Xiāmen Dàxué Chubanshe.
- Norman, Jerry (1974) ‘Tonal development in Mǐn’, *Journal of Chinese Linguistics* 2: 27–36.
- Norman, Jerry (1988) *Chinese*, Cambridge: Cambridge University Press.
- Norman, Jerry (1989) ‘What is a Kējī dialect?’ *Proceedings of the Second International Conference on Sinology, Academia Sinica*, 323–44.
- Norman, Jerry (1996) ‘Tonal development in the Jennchyan dialect’, *Yuen Ren Society Treasury of Chinese Dialect Data* 2: 7–41.
- Norman, Jerry (1999) ‘Vocalism in Chinese dialect classification’, in Simmons (ed.) (1999a).
- Sagart, Laurent (1993) *Les dialectes Gan: études sur la phonologie et le lexique d'un groupe de dialectes chinois*, Paris: Langues Croisées.
- Simmons, Richard Vanness (ed.) (1999a) *Issues in Chinese Dialect Description and Classification, Journal of Chinese Linguistics*, monograph series number 15.
- Simmons Richard Vanness (1999b) ‘On Chinese dialect classification – a case study examining the relationships of the Harngjou and Jennjiang dialects’, in Simmons (ed.) (1999a).

- Xú, Lín and Zhào, Yǎnpíng (1984) *Báiyǔ jiǎnzhi*, Beijing: Renmin Chubanshe.
- Wēn, Duānzhèng and Hóu, Jīngyī (1993) *Shānxī fāngyán diàochá bàogào*, Taiyuan: Shanxi Gaoxiao Lianhe.
- Yán, Sēn (1986) 'Jiāngxī fāngyán de fēnqū', *Fāngyán* 1: 19–38.
- Yán, Sēn (1993) *Líchuān fāngyán yánjiū*, Beijing: Shehui Kexueyuan.
- Yú, Aǐqín [Anne Yue-Hashimoto] (1991) 'Yuè fāngyán wèntí chūtàn', *Fāngyán* 3: 164–81.
- Yú, Zhìqiáng (1999) 'Issues in selecting features for genetic classification of Wú dialects', in Simmons (ed.) (1999a).

# CHINESE DIALECTS: 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). 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 subgroups of the Northern dialect such as Southeastern Mandarin or Jianghuai, Southwestern Mandarin (including Hubei), Huizhou, Shanxi (including the Jin group), 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 neighbouring 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, not to say detailed grammars of each major dialect group or of individual dialects.

## 2 GRAMMATICAL CATEGORIES

The lack of inflectional morphology in Modern Chinese renders it very difficult to define grammatical categories in terms of morphological features. The traditional classification of grammatical categories into 實字 ‘full words’ and 虛字 ‘empty words’ is based on the degree of semantic content 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 pattern with other categories.

### 2.2 Substantives

#### 2.2.1 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) that contains a measure word or a classifier (CL)<sup>1</sup>

---

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.

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.<sup>2</sup> For example: 一個橘子 [*ji*<sup>55</sup> *kə* *tɕy*<sup>35</sup> *tsɿ*] ‘an orange’ or 兩斤橘子 [*ljan*<sup>214>21</sup> *tɕin*<sup>55</sup> *tɕy*<sup>35</sup> *tsɿ*] ‘two catties of oranges’, but 三杯水 [*san*<sup>55</sup> *pej*<sup>55</sup> *swɛj*<sup>213</sup>] ‘three glasses of water’.<sup>3</sup>

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 Lánzhou 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 and its cognates, 隻 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.

### 2.2.2 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 [*tʰa*<sup>55</sup>] and its cognates are widely used only in the Northern, most Xiang, and a small number of Wu (for example, Yíxing, Lìyáng, Jinhúa, Danyáng, Jìngjiang, Chángzhou, Wúxi) dialects;<sup>4</sup> 佢 [*k'*]<sup>5</sup> or 其 [*ki*] and its cognates are used in the Yue, the Hakka, most Gan, some Wu – especially southern Wu (Chángshú, Húzhou Shuanglín, Zhuji, Yúyáo, Níngbo, Huángyán, Wenzhou, Qúzhou, Jinhúa, Yǒngkang), the Huizhou (Jixi, Shèxiàn Túnxi, Xiuníng, Yixiàn, Qímén, Wùyuan), some Western Min (Yǒngán, Shaxiàn, Sanmíng Sanyuán, Jiàn'ou, Nánpíng Xiáyáng, Jiànyáng, Songxi, Zhènghé, Shùnchang Yángdun, Pūchéng, Shùnchang, Jiānglè, Míngxi) and a small number of Xiang dialects (Suíníng, Chéngbù, Wūgang, Xùpǔ, Xīnhuà, Qíyáng, Máyáng); while 伊 [*i*] and its cognates are used in most Min and Wu dialects.

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). While the suffix 們 [*mən*] and its variants are widely used

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.

in most Northern dialects, a great variety of other suffixes exist across dialects, including 都 [tɔu] (Hándan, Wǔ'an of Hebei; Huòjia of Henan), 兜 (Anyáng of Henan, Hakka dialects such as Dongguǎn Qíngxi, Cóngguà, Xiàngǎng Shénzhèn, Méngshan Xihé), 岸 [ɲan] (Táoyuán of Hunan), [ $y^{31} \sim i^{31} \sim A^{31} + ts^{31}$ ] (Fúfeng of Shaanxi), [te]/[ne] (Gan dialect of Cháolíng, Wu dialect of Wenzhou, Gan dialect of Yángxin), [t'i]/[ti]/[li] (Píngjiāng Chángshòu of Hunan, Hakka dialects of Lùchuan, Gan dialects of Lǐlíng, Xínyú, Yífeng, Píngjiāng, Xiushūi, Hakka dialect of Liánnán), [na] (Línwǔ of Hunan), [no<sup>21</sup>]/[tɕi<sup>44</sup>]/[tɕi<sup>21</sup>] (Rùchéng of Hunan), [tse] (Pǔqí of Hubei), [kɔli] (Nánchāng of Jiangxi), [uo<sup>34</sup>] (Tàihé of Jiangxi), [tej<sup>22</sup>] (majority of the Yue dialects), and [ŋ] (Shàntóu of Guangdong). It is not unusual to form the plural by suffixing certain nouns such as 人 / 儂 / 伙 'person' (Northwestern Min, Hakka dialects such as Tónggǔ Sandu, Huizhou dialects such as Shèxiàn, Jìxi, Túnxì, Xiúníng, Yìxiàn, Dàitián of central Min) or 等 (人) ~ 登 [ten] (Hakka dialects of Méixiàn, Wēngyuán, and Língxiàn in Hunan) or 儕 (Dányáng of Wu, Hakka dialects such as Anyì of Jiangxi, Chángtíng of Fujian) or 多 (Yíyáng of Gan, Northwestern Min such as Jiànníng, Shàowǔ) or 大家 'everyone' (Shùnchāng of Min, Qímén of Huizhou). After all, plural suffixes are probably derived from earlier nouns.<sup>6</sup> The plural may also be indicated through suffixing a phrase such as: 多人 'several people' (Hakka dialects such as Níngguà, Gan dialects such as Yúgān, Nánchéng), 些人 / 幾個 / 一下 'several' (Guǎngjì of Hubei for second and third persons) or 幾個人 'several people' (Sùqián of Jiāngsu) or 各人 (Eastern Min dialects) or 個人 (Xianju of Wu) or 勒人 (Yǒngkāng of Wu).

Sometimes the root and the suffix undergo contraction and produce a composite form (for example, 俺 < 我 + 們<sup>7</sup> for 1pl is common among Shandong dialects such as Máo píng, Píngdù, Wéifāng, Zībó; also in Guǎngjì of Hubei), found in some Hebei (Wèixiàn has 俺 for 1pl, 您 for 2pl and [nie<sup>43</sup>]/[nɛ<sup>43</sup>] for 3pl), Shanxi (Hóngtòng and Línfén have [ɲua] < 我家 for 1pl, [nia] < 你家 for 2pl; Língchuan has [uə<sup>213</sup>] < 我們 for 1pl, [nə<sup>213</sup>] < 你們 for 2pl and [t'ə<sup>33</sup>] < 他們 for 3pl; Shanyin may have [ua<sup>52</sup>] for 1pl and [niə<sup>52</sup>] for 2pl with 們 optionally suffixed), and Jiangsu dialects ([t'am] for 他們 in Gànyú). Southern Min dialects such as Shàntóu and Cháozhōu of Guangdong have contracted forms ending in [ŋ] in the plural ([naŋ<sup>53</sup>] for 1incl, [uŋ<sup>53</sup>] for 1ex pl, [niŋ<sup>53</sup>] for 2pl and [iŋ<sup>33</sup>] for 3pl).

Note that the plural marker may differ for first vs second vs third persons, especially in the Wu dialects: the Suzhou dialect of Jiangsu uses the suffix [toʔ] for the second and the third persons, but a contraction form [ni] for the first person, Shanghai uses the suffix 呢 for the first person, the suffix [A<sup>53>44</sup>] for the third person, but a contraction form [nA<sup>23</sup>] for the second person, the Hǎiyán 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 (Zengchéng, Yángjiāng, Héshān and the Siyi dialects of Xinhui, Táishān, Kaipíng, Enpíng, use a glottalized low falling tone), Wu (Shanghai Fèngxián), as well as some Northern dialects of Shanxi (Línýi may use a falling-rising tone besides lengthening the rhyme, Yùchéng uses a long low rising tone optionally followed by the suffix [ti]) and Shaanxi (Xí'an and Bǎoji use a low falling tone; Shāngxìàn).

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 for [+visible] and 他 for [-visible].

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

6 See for example Lyu (1985): 59–61 for a possible etymological derivation of 們 from 輩.

7 See Lyu (1985): 54.

Chángshòu), some Wu (for example, Wenzhou, Chángzhou, Jiangyin), some Gan (such as Yifeng, Duchang, Nánchéng), as well as most Min dialects (for example, Fúzhou, Jiàn'ou, Jiànníng, Shàowǔ, Shùnychang, Pútíán, Xiàmén, Shàntóu, Cháozhou, Dàtián).

Case distinction is almost unheard of, but it exists in the Xúnhuà dialect of Qinghai (Northwestern Mandarin) probably as a result of contact with minority languages in the region:<sup>8</sup>

	1	2	3		1	2	3
sg nom	<i>ŋə</i>	<i>ni</i>	<i>t'A</i>	pl nom	<i>ŋəmv</i>	<i>nimv</i>	<i>t'Amv</i>
sg acc	<i>ŋa</i>	<i>nia</i>	<i>tciɛ</i>	pl acc	<i>ŋamA</i>	<i>niamA</i>	<i>tciɛmA</i>

In the Hǎiyán dialect, personal pronouns distinguish between a preverbal and a postverbal form, reminiscent of a case distinction too perhaps.

### 2.2.3 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:

- (1) a. SM (我 *wɔ*<sup>213>35</sup> 買 *maj*<sup>213>11</sup> 了 *lə* 個 *kə* 收音機 *sow*<sup>55</sup> *jin*<sup>55</sup> *tci*<sup>55</sup>)  
 1sg buy pfv-asp CL radio  
 ‘I bought a radio.’
- (收音機 *sow*<sup>55</sup> *jin*<sup>55</sup> *tci*<sup>55</sup>) 多少 *twɔ*<sup>55</sup> *saw* 錢 *tɕ'jen*<sup>35</sup> ?  
 radio how-much money  
 ‘How much is (the radio)?’

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

- (1) b. SM 把 *pa*<sup>213>11</sup> 它 *t'a*<sup>55</sup> 關上 *kwan*<sup>55</sup> *ʂaŋ!*  
 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 postverbal position, for example:

- (2) a. Shanghai [*die?*] 碗 湯 倒 脫 伊  
 this CL soup pour out 3sg  
 ‘Throw out this bowl of soup.’

8 Examples taken from Huang (1996): 439.

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

- (2) b. 倒 *tow*<sup>35</sup> [*tʃ*<sup>35</sup>] [*ni*<sup>55</sup>] 碗 *wun*<sup>35</sup> 湯 *t'ɔŋ*<sup>55</sup> 佢 *k'øy*<sup>24</sup>  
 pour pfv-asp this CL soup 3sg  
 'Throw out this bowl of soup.'

#### 2.2.4 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 vs 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 the vast number of Northern dialects. The majority of the Min dialects also use a form with an affricate initial for the proximal (except Jiàn'ou, Jiànyáng, Songxi) but one with a velar fricative or [h] initial for the distal (except Jiàn'ou, Jiànyáng, Songxi, Shaxiàn, Yǒngán).

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 for the proximal and a form with a velar stop [k] initial for the distal deictic, the Siyi dialects, Yángjiāng, and Zhongshan 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 have 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 (Lúxi, Yuánlíng, Gùzhāng, Bǎojīng, Xiāngxiāng, Ānxiāng, 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 are by far the most varied 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 vs distal contrast. Dialects like Suzhou, Lìyáng, Dányáng, Chángzhōu, and Huángyán use forms with front vowels for the proximal and those with non-front vowels for the distal deictic. Dialects like Zhujì, Qúzhōu, and Jinhua use forms with velar initials for the proximal and those with labial initials for the distal deictic. This second category has a variant in dialects like Shèngxìàn Chónggrén, Shèngxìàn Tàipíng, and Níngbō, 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 [ʔ] to indicate the distal deictic, found in dialects such as Bǎoshān Shuāngcǎodūn, Bǎoshān Luódìàn, Nánhuì Zhōupǔ, Shanghai, and Songjiāng.

The feature visibility is reported to exist in the Shǎnxìàn dialect of Henan, where the distal demonstratives distinguish between [*vei*<sup>23</sup>] with the feature [+visible] and [*nai*<sup>23</sup>] with the feature [–visible].

A small number of dialects, some Northern dialects in Shandong (Zibó, Shòuguāng, Wéifāng), Shanxi (Yuánpíng, Shòuyáng, Yángchéng, Shílóu, Zhongyáng, Línxiàn, Liǔlín, Yúshè, Yúxiàn, Xiyáng, Jìnglè, Línfén, Xìnzhou, Héshùn, Yángqu, Tàiyuán, Wànróng) and Hubei (Zhongxiāng, Yingshan, Luótián, Xishǔi, Jiayú, Huáng'an, Qíchun, Gonggan, Hèfeng,

Zìgui), some Wu (Suzhou, Hǎiyán Tongyuán, Wúxi, Chángshù, Shèngxiàn Chánglè) and some Hakka (Lùfeng,<sup>9</sup> Xinfeng, Cháo Zhou) dialects have a tripartite distinction of proximal vs medial vs distal. For example, Yingshan has [tɛ<sup>35</sup>] vs [n<sup>35</sup>] vs [a<sup>33</sup>], Zibó has [tɕə<sup>214</sup>] vs [ŋiə<sup>214</sup>] vs [na<sup>214</sup>], Shanxi has [tɕ<sup>55</sup>]/[tɕl<sup>33</sup>]/[tɕə<sup>31</sup>]/[tɕai<sup>51</sup>]/[tɕei<sup>51</sup>]/[tɕəi<sup>35</sup>]/[tɕar<sup>51</sup>]/[tɕəp<sup>55</sup>]/[tɕəp<sup>4</sup>]/[tɕə<sup>21</sup>]/[tɕə<sup>31</sup>]/[tɕə<sup>53</sup>]/[tɕəi<sup>55</sup>] vs [na<sup>35</sup>]/[nA<sup>53</sup>]/[nA<sup>35</sup>]/[nə<sup>53</sup>]/[nə<sup>21</sup>]/[nai<sup>33</sup>]/[nai<sup>51</sup>]/[nəi<sup>35</sup>]/[nei<sup>53</sup>]/[nəi<sup>55</sup>]/[nə<sup>53</sup>]/[niou<sup>35</sup>]/[nəp<sup>4</sup>]/[vei<sup>53</sup>] vs [u<sup>33</sup>]/[uo<sup>53</sup>]/[uɿ<sup>55</sup>]/[uə<sup>21</sup>]/[ua<sup>53</sup>]/[uai<sup>51</sup>]/[uei<sup>51</sup>]/[uəi<sup>35</sup>]/[uəi<sup>55</sup>]/[uar<sup>51</sup>]/[uə<sup>53</sup>]/[vɿ<sup>53</sup>]/[və<sup>55</sup>]/[vei<sup>35</sup>]/[vəp<sup>4</sup>], Suzhou has [kE<sup>44</sup>kɿ<sup>44>21</sup>] vs [gɿ<sup>23</sup>kɿ<sup>44</sup>] vs [kuE<sup>44</sup>kɿ<sup>44>21</sup>]. In addition, the Shèngxiàn Chánglè dialect maintains this tripartite distinction together with a distinction between a stative vs an active state. A location is stative while a goal which some movement is directed toward is active. The stative demonstratives are [ku<sup>44</sup>] for proximal (visible and touchable), [lɔy<sup>35</sup>] for medial (nearby but invisible and untouchable), and [moŋ<sup>35</sup>] for distal; while the active demonstratives are [kua<sup>53</sup>] for proximal (visible and touchable), [lia<sup>53</sup>] for medial (nearby but invisible and untouchable), and [maŋ<sup>53</sup>]/[moŋ<sup>35</sup>fla<sup>53</sup>] 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 vs proximal distinction into one, that is, using the same form for both designations.

## 2.3 Verbs and aspects

### 2.3.1 Verbs

The verb (V) can be uniquely defined as negatable. The so-called adjectives are but stative verbs (Vstat), erroneously identified as equivalent 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’, 恨 ‘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.

### 2.3.2 Aspects

Aspect markers (ASP) in the modern dialects are mostly derived from verbs. In other words, certain verbs become grammaticalized into function words. This grammaticalization process took place at different times and 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 usage, the majority of dialects employ a mixture of suffixes (where grammaticalization

9 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 a distinction in Chinese.



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.

### 2.3.2.1 Perfective aspect

The perfective aspect (pfv-asp) signals the realization of a certain state or action in the past or in the future, most commonly expressed by a suffix on the verb. In the Northern dialects it is overwhelmingly 了 [lǎo], derived from a verb with the meaning of ‘to finish’, and its variants [li]/[lou]/[liou]/[lio]/[lɔ]/[liau]/[liɛ], although in the Huizhou dialects 著 [tɕio] or 到 [tɕ] or 得 [tɛ] is also used; in the Wu dialects it is mostly [tsəʔ] and its variants [dɔʔ]/[deʔ]/[teʔ]/[tsɿ]/[zɿ]/[la]/[da]/[ta]; in the Xiang dialects it is mostly [ta]/[tɔ] and its variants [tie]/[tu]/[tau] or 得 [te] or [ka]/[ga]/[kɔ]/[ku]/[ko]/[kua] or a combination of the form [kata]/[kota]/[kɔtɔ]; in the Gan dialects it is [i]/[e]/[ɛ]/[ə] or [te]/[tə]/[tɛ]/[tɛʔ] or [li]/[le]/[lo]/[lɔ]; in the Hakka dialect it is [e] or [le]/[li]/[lei]/[liu]/[liɔ]/[liɔ]/[luɔ]; and in the Yue dialects it is [tɕ] or [ə] (Siyi). In the Min dialects, there is a perfective aspect marker [ɛ] or [liu], grammaticalized to different degrees among different dialects, and such a concept may also be expressed with a complement verb such as 好 [ho<sup>53</sup>] in Shàntóu (with favourable connotation) or 去 [k’u] in Quánzhōu and Shàntóu (with unfavourable 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,

- (3) Tàixing 我 說 [ka] 服 [ka] 他 了  
 1sg talk pfv-asp convince pfv-asp 3sg FP  
 ‘I (have) convinced him.’

(FP=final phase particle) 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).

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 (QP=question particle):

- (4) a. Cantonese: 你 nei<sup>24</sup> 有 jɛw<sup>24</sup> 無 mow<sup>24</sup> 有 jɛw<sup>24</sup> (食 sik<sup>2</sup>)/無 mow<sup>24</sup>  
 (Yue) 2sg have not-have have (eat) /not-have  
           食 sik<sup>2</sup> 飯 fan<sup>22</sup> 啊 a<sup>44</sup>? (食 sik<sup>2</sup>)  
           eat rice QP (eat)  
           ‘Did you eat?’ Yes (I) did/No (I) did not.’

- b. 你 *nei*<sup>24</sup> 食 *sik*<sup>2</sup> [*tʃɔ*<sup>35</sup>] 食 *sik*<sup>2</sup> [*tʃɔ*<sup>35</sup>] 喇 *lɑ*  
 2sg eat pfv-asp eat pfv-asp FP/  
 飯 *fan*<sup>22</sup> 未 *mej*<sup>22</sup> 啊 *a*<sup>44</sup> ? (*[tʃoŋ*<sup>22</sup>]) 未 *mej*<sup>22</sup> 食 *sik*<sup>2</sup> 啊 *a*  
 rice not-yet QP (still) not-yet eat FP  
 ‘Have you eaten yet?’ ‘(I) have eaten/(I still) have not eaten.’
- c. SM: 你 *ni*<sup>213>11</sup> 吃 *tʃɿ*<sup>55</sup> 了 *lə* 吃 *tʃɿ*<sup>55</sup> 了 *lə* / 沒 *mej*<sup>35</sup>  
 2sg eat pfv-asp eat pfv-asp/ not-have  
 飯 *fan*<sup>51</sup> 沒 *mej*<sup>35</sup> 有 *ʔ*<sup>213</sup> ? 吃 *tʃɿ*<sup>55</sup>  
 rice 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 *ŋ*<sup>24</sup> 屋企 *ʊk*<sup>5</sup> *k*<sup>’</sup> *ei*<sup>35</sup> 有 *jɐw*<sup>24</sup> 養 *qøŋ*<sup>24</sup> 雞 *kɐi*<sup>55</sup>  
 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 Huòxiàn (Shanxi):

- (6) a. 你 吃 [*lou*] 飯 [*la*] 麼 ? 吃 [*la*] / (還) 無 吃 [*lie*]  
 2sg eat pfv-asp rice pfv-asp QP eat pfv-asp/(still) not-have eat FP  
 ‘Have you eaten?’ ‘(I) have eaten/(I still) have not eaten.’
- b. 你 吃 飯 [*li*] 麼 ? 吃 [*li*] / 無 吃  
 2sg eat rice PAST QP eat PAST/not-have eat  
 ‘Did you eat?’ ‘Yes (I) did/No (I) did not.’

The final 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 來 ‘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 Mishi* 元朝秘史, *Piaotongshi* 朴通事.<sup>10</sup>

In some dialects, the perfective aspect takes the form of rhyme change. For example, in the Huòjia dialect (Jin, 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 Shangxiàn dialect (Shaanxi, Northern), it is the lengthening and lowering of the vowel plus the change of each tonal segment into a falling contour. For example:

- (7) Huòjia 我去 *tɕ*<sup>’</sup> *yɛ*<sup>13</sup> < *tɕ*<sup>’</sup> *y*<sup>13</sup> 兩趟  
 1sg go two times  
 ‘I went/have gone twice.’

10 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 [*pin*<sup>55</sup>] 道來啊 ‘where have you been?’ (*lit.* ‘where did you go and come back?’) or in Yángjiang (Yue).

- (8) Shangxiàn 病 *piã̃*:<sup>55l</sup> < *pĩ*:<sup>55</sup> 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 Yǒngkāng (Southern Wu); or retroflexion of the verb, as in Hǎiyáng and Máo píng (Shandong).

### 2.3.2.2 Change-of-state *cum* new situation

Across the dialects, there is a phase particle 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. Very often, it occurs together with the perfective aspect in the same sentence. For this final particle, in the Northern dialects, a particle with the same phonetic form as the perfective aspect marker but derived from a different historical source is used,<sup>11</sup> for example, the following sentence:

- (9) SM: 吃 *tʂ*<sup>55</sup> 飯 *fan*<sup>5l</sup> 了 *lɔ*  
 eat rice FP

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 affairs 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 [*lɔʔ*]/[*liɿʔ*] 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*]/[*tɛ*]/[*tɿʔ*] or of 了 [*li*]/[*lɛ*]/[*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 particle, [*la*]/[*lɔ*], is distinct from the perfective aspect marker; and in the Min dialects, some weakened form of 了, [*lɔ*] or [*ou*], is used.

### 2.3.2.3 Experiential aspect

Except for Southern Min, almost all dialects use the suffix 過, [*kwɔ*] 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 Wèixiàn (Hebei), 著 [*tʂo*<sup>3l</sup>] is reported for Kunmíng (Yunnan) and [*ta*] for Jíngmén (Hubei). In Southern Min, the verb [*bat*] is used before another verb for the same function, although the suffix 著 [*tioʔ*] and 過 [*k(u)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(u)e*] is probably under the influence of either SM or Cantonese (which is the *lingua franca* in Guangdong province where many Southern Min dialects are spoken).

### 2.3.2.4 Progressive aspect

Like in many other languages of the world,<sup>12</sup> the majority of the dialects use a form or a phrase derived from an indefinite locative expression with a locative verb (Vloc) and

11 The final particle 了 is derived from a weak form of 來 which in turn may be derived from the combination of 了 + 也.

12 See Comrie (1978) Section 5.2.1.1.

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 Chéngdu (Sichuan). In the Northern Wu dialects, the crystallized locative phrase  $[l\partial\eta\varepsilon]/[l\partial\eta\bar{l}\partial]/[l\partial\eta\eta\partial]$  and their variants are suffixed to the verb; and in the Southern Wu dialects such as Wenzhou, it is realized as  $[z\eta\eta]$ .<sup>13</sup> In the Huizhou dialects, a similar crystallized phrase, 是  $[no\beta^3]/[k\gamma^{11}]/[lo^{11}]$ ,<sup>14</sup> sometimes with proximal (在  $[n^{22}]/[t\partial\eta]$ /在  $[n^{22}]/[na]$ /是  $[mo^{53}]/[le^{33}]/$ 是  $[ka^{53}]/[le]$ ) vs distal (是那  $[t\partial\eta]$ /是  $[m\bar{e}i^{22}]/[t\partial\eta]$ /在  $[na^{31}]/[na]$ /是  $[ka^{53}]/[le]$ /是  $[mo^{53}]/[le^{33}]$ ) 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  $[h\bar{v}i^{35}]$  道  $dow^{22}/[h\bar{v}i^{35}]$  處  $sy^{44}$  or  $[\partial^{33}]/[nei\eta^{35}]$  (Siyi dialects) may be used,<sup>15</sup> there is also a distinct progressive suffix  $[k\bar{v}n^{35}]/[kin^{55}]$  (Siyi). While  $[kin]/[t\bar{c}in]$  is also used in many Hakka dialects, which is either cognate with the Yue form or a loan from the Yue dialects,  $[ten^{31}]$  or  $[ts'anj^{52}]/[t'i^{52}]$  is used in the prestige Hakka dialect Méixiàn. In other dialects the progressive shares 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  $[t\bar{c}'i]$  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  $[t\bar{s}\partial]$  with a weakened form  $[t\partial]$ , although for the progressive aspect the final particle (FP)  $[n\partial]$  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 Southwestern Mandarin (Guìyáng of Guizhou, Northeastern Yunnan) and the same suffix  $[t\partial\eta]$  in many Shanxi dialects. 在  $[ts'ai]$  is also used as a prefix in the Gan dialects but a suffix in some Anhui dialects (Héféi) for the progressive; it can also be suffixed to other progressive suffixes to indicate the same (Huòqiu of Anhui, Pǔqi of Hubei, Sichuan, Chénxi of Hunan). Parallel to the situation in the Wu dialects, in the Mǐn dialects, 在  $[te]/[le]$  is prefixed ( $[le^{21}]$  in Fúzhou) to a verb to indicate the progressive aspect and suffixed ( $[le]$  in Fúzhou) 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 Southern Mǐn dialect of Shàntóu renders support to this assumption. In this dialect, among the entire array of some seven variants,  $[lo]/[to]/[toko]$  and  $[na]/[nako]/[pan]/[panko]$ , 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  $[hu^{53} ko]$  'there' in Cháozhou).

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,  $v[t\bar{s}\partial]v[t\bar{s}\partial]$ , indicates such a progression. In the Xiang dialects,  $v[ta]v[ta]$ ,  $v[tau]/[tou]v[tau]/[tou]$ ,  $v[ts'an]v[ts'an]$  (Yiyáng, Xiangtán),  $v[t\bar{c}'y]/[dy]v[t\bar{c}'y]/[dy]$  (Lóudǐ, Xiangxiang), are all reduplicated progressive forms. In the Gan dialects,  $v[a]v$  is the predominant form, while  $v\text{住}[k'y]v[k'y]$   $v\text{到}[tau]v[tau]/v[s\bar{a}]v[s\bar{a}]$  are also used. In the Yue dialects, a suffix 下  $[ha^{35}]$  is used with a reduplicated verb to indicate the same, namely,  $VV [ha^{35}]$  or  $v[ha^{24}]v[ha^{24}]$ . Some Northern

13  $[l\partial\eta]$  and  $[z\eta]$  are  $vloc$  while  $[h\varepsilon]/[l\bar{l}]/[to\eta]$  is probably an indefinite locative deictic and  $[ta]$  is the distal locative 'there'.

14 In several Huizhou dialects such as Jixi, Túnxi, Xiuníng, Yixiàn,  $vcop$  assumes the function of  $vloc$  too.

15 Both  $[h\bar{v}i^{35}]$  and  $[\partial^{33}]$  are  $vloc$  and  $[nei\eta^{35}]$  is the distal locative 'there'.

(Huòqiū, Shèxiàn of Anhui) and Wu dialects (Shanghai) simply use a reduplication of the verb, *VV*, for such purpose.

### 2.3.2.5 Durative aspect

In many Northern and Central dialects, the same marker for the perfective aspect is also used for the durative aspect (*dur-asp*). However, in many of these dialects, more than one marker is used for the durative. In the Wu dialects, the same form for the progressive aspect is suffixed to the verb to express another form of the durative aspect, namely,  $V + [l\partial\partial h\epsilon]/[l\partial\partial?l\partial]/[l\partial\partial to?]$ . In the Xiang dialects, while the perfective aspect marker  $[ta]$  or  $[t\epsilon]$  and 得  $[t\chi]$  also marks the durative, the progressive suffixes, 起  $[t\epsilon'i]$  and 倒  $[tau]/[tou]$ , are so used too. Similarly, in some Southwestern Mandarin dialects, the progressive suffix 起 or 倒, as mentioned in the previous section, also functions as the durative. The suffix 倒 is also used throughout the Gan dialects as  $[tau]/[tou]/[t\partial]$ , as well as many of the Hakka dialects as  $[tau]/[t\tau]/[t\tau]$ . In Méixiàn however, the durative aspect marker is not distinct from the progressive aspect. Most of the Yue dialects use 住  $[t/y^{22}]$  as the durative suffix, but the Sipi 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 → (NP) (TIME (PRT)) VP ({FP, QP})  
 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 last structure; the PP has a head first structure, while the VP has a head middle structure.

The most frequently occurring unmarked word order in a sentence is Subject (S) Verb (V) 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 an 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.<sup>16</sup> On the other hand, in this type of

16 This is the case in the majority of the Wu dialects. In some, such as Lìyáng, Jintán, Danyáng, Jiangyin, Chángzhou, Suzhou, Jiaxìng, 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 explanation below in connection with postverbal adverbs.

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,

- (10) a. Suzhou, Kunshan: 衣裳 脫 脫 伊  
 b. Wúxi, Jiangyin: 衣裳 脫 落 它  
 c. Wángjiājǐng: 衣裳 脫 其 掉  
 garment take off 3sg off  
 'Take off the garment.'

Furthermore, in some Northwestern dialects (Northern), such as those in Qinghai or Gansu, SOV word order does obtain under the influence of neighbouring 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 postverbal 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:

- (11) a. 叫 架 車 [la]  
 all CL car FP  
 'Call a car.'  
 b. 架 車 來 [tʃɔ̃<sup>35</sup>] [la]  
 CL car come PFV-ASP 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 the Wu (for example, Suzhou, Shanghai, Jinhua, Yǒngkāng, Shuānglín, Chángshú, Shèngxìàn, Wenzhou, Yǒngkāng) and 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, 支筆 designates 'the pen' in 支筆不好寫 'the pen does not write well'.

The double-asterisked word order for certain adverbs (postverbal) 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, Shèxiàn Túnxi, Xiuning, Yixiàn, Qímén, Wùyuan), Jianghuai, Southwestern Mandarin, and even a few Northern dialects in Hebei (Mǎnchéng), Qinghai, and Inner Mongolia (Huhehote). Postverbal 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)<sup>17</sup> 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,

17 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 '21 January', 'today', etc.

among the Wu dialects, when the Adv is 快, *v+Adv* is the norm in Yíxing, Suzhou, Chángshú, Bǎoshan Shuangcǎodun, Nánhuì Zhoupǔ, Songjiang, Húzhou Shuanglín, Shàoxing, Yúyáo, and Níngbo; but both *v+Adv* and *Adv+v* are used in Jingjiang, Wúxi, Kunshan, Bǎoshan Luódiàn, Shanghai, Wújiāng Lílǐ, Wújiāng Shèngzé, Jiāxing, Hángzhou, and Zhujì; while *Adv+v* appears in Lìyáng, Jīnhuá, Dányáng, Jiāngyin, Chángzhou, Shèngxiàn Chóngren, Shèngxiàn Tàipíng, Huángyán, Wenzhou, Qúzhou, and Yǒngān. 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 is 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 Expressions dependent on whether the sentence is affirmative or negative. This will be discussed in Section 4.3 below.

##### 4.1 The copular sentence NP1 + *vcop* + NP2

The great majority of the dialects use a form of the verb 是 [*ʃr*<sup>51</sup>] as the copula (*vcop*), which is derived from the demonstrative in the same form in Old Chinese. Only the Yue and the Hakka dialects employ a different form, 係 [*hɛj*<sup>22</sup>] in Cantonese.<sup>18</sup>

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 *ta*<sup>55</sup> 北京人 *pei*<sup>214>11</sup> *tɕiŋ*<sup>55</sup> *zən*<sup>35</sup>  
 3 SG 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 where neg = negative marker, 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, sentences such as:

- (13) Shùyáng 那 也 圖書館書  
 that also library-book  
 ‘That (is) also [a] library book.’
- (14) Jīnhuá 這 本 書 弗 我 的  
 this CL book neg 1sg NOMZR  
 ‘This book (is) not mine.’  
 (NOMZR = nominalizer)

are perfectly grammatical but ungrammatical in other dialects, Northern or Southern.

18 There are some exceptions. The Yue dialect of Liánzhou uses 是 while according to Yang (1974) the Jiahé dialect of Hunan uses both 是 and 係.

#### 4.2 Existential sentences (TIME) (PLACE) Vex + NP1 / V + dur-asp + NP2

Sentences describing existence fall into three major types: those that carry an existential verb such as 有 (Vex, [jɔw] in SM) or the copula, those relating to natural phenomena, and those whose verbs are either marked with the durative aspect or limited to expressing appearance or disappearance. They are all characterized by an NP of indefinite reference<sup>19</sup> 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<sup>213>11</sup> mjɛn 有 jow<sup>213>11</sup> 什麼 sɔm<sup>35</sup> mɔ?  
 inside exist what  
 ‘What is inside?’ (not knowing whether there is anything inside)
- (16) 裏面 li<sup>213>11</sup> mjɛn 是 sɿ<sup>51</sup> 什麼 sɔm<sup>35</sup> mɔ?  
 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,

- (17) a. SM: 下 ɟja<sup>51</sup> 雨 qy<sup>213>11</sup> 了 lɔ  
 b. Cantonese: 落 lɔk<sup>2</sup> 雨 qy<sup>24</sup> [lɔ]  
 fall rain FP  
 ‘It is raining now.’

The third type 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 laj<sup>35</sup> 了 lɔ 客人 k'ɿ<sup>51</sup> zɛn<sup>35</sup>  
 come pfv-asp guest  
 ‘Some guest(s) came.’
- (19) Cantonese: 有 jɔw<sup>24</sup> 人客 jɛn<sup>11</sup> hak<sup>4</sup> 來 lɛj<sup>11</sup> [tɿ<sup>35</sup>]  
 Vex guest come pfv-asp  
 ‘Some guest(s) came.’
- (20) SM: 牆 tɟ'jaŋ<sup>35</sup> 上 saŋ<sup>51</sup> 掛 kwa<sup>51</sup> 著 tɟɔ/了 lɔ 一 ji<sup>55</sup> 幅 fu<sup>35</sup> 畫兒 xwar<sup>51</sup>  
 wall-on hang dur-asp/pfv-asp one CL painting  
 ‘A picture hung on the wall.’

19 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 (emph=emphatic marker)  
 ‘There is this kind of person everywhere.’



- (21) Cantonese: [pʊŋ<sup>22</sup>] 牆 tʃ'øŋ<sup>11</sup> 上高 søŋ<sup>22</sup> kow<sup>55</sup> 掛 kwa<sup>44</sup> [tʃɔ̃<sup>35</sup>]  
 CL wall above hang pfv-asp  
 有 jɛw<sup>24</sup> 一 jɛt<sup>5</sup> 幅 fʊk<sup>5</sup> 畫 wa<sup>35</sup>  
 vex one CL painting 'ibid.'  
 [pʊŋ<sup>22</sup>] 牆 tʃ'øŋ<sup>11</sup> 上高 søŋ<sup>22</sup> kow<sup>55</sup> 有 jɛw<sup>24</sup> 幅 fʊk<sup>5</sup>  
 CL wall above vex CL  
 畫 wa<sup>35</sup> 掛 kwa<sup>44</sup> [hɛt<sup>35</sup>] 道 dow<sup>22</sup>  
 painting hang vloc place/pro-asp<sup>20</sup> 'ibid.'

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 [(loʔ) hoʔ]= stative  
 wall-there hang dur-asp (one CL) painting (vloc) place  
 'a picture hung on the wall'  
 (23) 牆頭浪 掛 了 (一張) 畫= active  
 wall-there hang pfv-asp (one CL) painting 'ibid.'

### 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) stop 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 [pje<sup>35</sup>] 'don't!' or 不 [pu<sup>51>35</sup>] + 要 [jäw<sup>51</sup>]/用 [jðŋ<sup>53</sup>]/准 [tʃwən<sup>214</sup>]/許 [ɕy<sup>214</sup>] 'do not!' / 'need not!' / 'not allowed!' etc., 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/aspects, 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-asp) 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-asp)<sup>21</sup> or 未 + 曾 is used for

20 The indefinite locative expression functions as the progressive aspect. See Section 2.3 above.

21 In some Northern Wu dialects, the neg is 勿 with a [p]/[f] initial rather than 未 with an [m] initial.

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:<sup>22</sup>

- |                               |  |   |  |
|-------------------------------|--|---|--|
| (24) a. Xiàmén (Southern Min) | 伊 無 去 :  | : | 伊 未 去  |
| b. Méixiàn (Hakka)            | 其 無 去 :  | : | 其 未 去  |
| c. Cantonese (Yue)            | 佢 $k'\phi y^{24}[mou^{24}]$<br>去 $h\phi y^{44}$            | : | 佢 $k'\phi y^{24}$ 未 $mey^{22}$<br>去 $h\phi y^{44}$                   |
| d. Suzhou (Wu)                | 伊 $[m^{24>22} p\gamma^{24}]$ 去                             | : | 伊 勿 曾 去  |
| e. SM (Northern)              | 他 $ta^{55}$ 沒 $mej^{35}$<br>去 $t\zeta y^{51}$              | = | 他 $ta^{55}$ 沒 $mej^{35}$<br>去 $t\zeta y^{51}$                        |
| f. Shuangfeng (Xiang)         | 他 $[m\gamma^{33}]$ 去                                       | = | 他 $[m\gamma^{33}]$ 去   |
| g. Nánchang (Gan)             | 佢 $[mau^{21}]$ 去<br>3SG NEG + VEX go<br>'S/he did not go.' | = | 佢 $[mau^{21}]$ 去<br>3SG NEG + PFV-ASP go<br>'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, 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  $[be^{42}]$  and negated with  $[m^{33}]$ , possibility/probability is expressed by  $[ue^{21}]$  and negated with  $[bue^{21}]$ ; necessity is expressed by 著  $[tio^{24}]$  and negated with the prohibitive 免  $[bian^{42}]$ , the perfective is expressed by 了  $[diau^{42}]$  and negated with 未  $[be^{22}]$ , and existence/affirmation is expressed by 有  $[u^{22}]$  and negated with 無  $[bo^{23}]$ . 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  $[m^{33}]$  but simple negation of Vstat is expressed with  $[bo^{23}]$ , since  $[m^{33}]$  is indicative of subjective volition and  $[bo^{23}]$  is used for objective, factual negation.

In general, the scope of negation is rightward, namely, it includes all elements occurring to the right of the neg. Thus, depending on the location of neg, different scopes of negation may result. Compare the following in SM:

- |                                 |                           |                             |                             |                             |                         |
|---------------------------------|---------------------------|-----------------------------|-----------------------------|-----------------------------|-------------------------|
| (25)                            | 你<br>$ni^{213>11}$<br>2sg | 不<br>$pu^{51}$<br>neg       | 可以<br>$k'e^{213>35}$<br>may | 看<br>$ji^{213>11}$<br>watch | $k'an^{51}$             |
| 'You are not allowed to watch.' |                           |                             |                             |                             |                         |
| (26)                            | 你<br>$ni^{213>11}$<br>2sg | 可以<br>$k'e^{213>35}$<br>may | 不<br>$ji^{213>11}$<br>neg   | 看<br>$pu^{51>35}$<br>watch  | $k'an^{51}$             |
| 'You may choose not to watch.'  |                           |                             |                             |                             |                         |
| (27)                            | 你<br>$ni^{213>11}$<br>2sg | 不<br>$pu^{51}$<br>neg       | 可以<br>$k'e^{213>35}$<br>may | 不<br>$ji^{213>11}$<br>neg   | 看<br>$pu^{51}$<br>watch |
| 'You cannot but watch.'         |                           |                             |                             |                             |                         |

22 This distinction also prevails in the Huizhou dialects. See Hirata (1998): 274.

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 postverbally in an affirmative sentence but preverbally in a negative sentence. For example in SM:

- (28) 他 *t'a*<sup>55</sup> 病 *piŋ*<sup>51</sup> 了 *lə* 三 *san*<sup>55</sup> 天 *tjan*<sup>55</sup> 了 *lə*  
 3sg be-sick pfv-asp three day FP  
 'He has been sick for three days.'
- (29) 他 *t'a*<sup>55</sup> 三 *san*<sup>55</sup> 天 *tjan*<sup>55</sup> 没 *mej*<sup>35</sup> 病 *piŋ*<sup>51</sup> 了 *lə*  
 3sg three day neg + pfv-asp be-sick FP  
 'He has not been sick for three days.'
- (30) 我 *wɔ*<sup>214>11</sup> 來 *laŋ*<sup>35</sup> 了 *lə* 三 *san*<sup>55</sup> 次 *ts'i*<sup>51</sup> 了 *lə*  
 1sg come pfv-asp three time FP  
 'I have come three times.'
- (31) 我 *wɔ*<sup>214>11</sup> 三 *san*<sup>55</sup> 次 *ts'i*<sup>51</sup> 没 *mej*<sup>35</sup> 來 *laŋ*<sup>35</sup> 了 *lə*  
 1sg three time neg + pfv-asp 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 (QP) 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 Línxià (Gansu), this kind of question always ends with [zɑ] and in Zhongníng (Ningxia) 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 [*seɪ*<sup>35</sup>] and its cognates in many Northern and even some Yue (such as Siyi) dialects but 哪個, literally 'which one', and its variants in Southwestern Mandarin, Jianghuai, Xiang, Gan, 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 [a<sup>21</sup>] with an extra low falling tone or [mɛ<sup>55</sup>] or [mɔk<sup>21</sup>] 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.<sup>23</sup> For example, in SM:

- (32) 你 ni<sup>214>11</sup> 去 tɕ'y<sup>51</sup> 嗎 ma?  
 2sg go QP  
 'Are you going?'
- (33) 你 ni<sup>214>11</sup> 不 pu 去 tɕ'y<sup>51</sup> 嗎 ma?  
 2sg neg go QP  
 '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 nei<sup>24</sup> 去 høy<sup>44</sup> [mɛ<sup>55</sup>]/[a<sup>21</sup>]?  
 2sg go QP  
 'Are you going (I thought you're not going)?'
- (35) 你 nei<sup>24</sup> [m] 去 høy<sup>44</sup> [mɛ<sup>55</sup>]/[a<sup>21</sup>]?  
 2sg neg go QP  
 '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<sup>214>11</sup> 去 tɕ'y<sup>51</sup> / 是 sɿ<sup>51</sup>, 我 wɔ<sup>214>11</sup> 不 pu 去 tɕ'y<sup>51</sup>  
 no Isg go / yes Isg 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 [fiɛzɿ]), in the Yue dialects it is 定 (係) [tɿŋ<sup>22</sup>(hɐj<sup>22</sup>)] (Cantonese) or [ma<sup>35</sup>] 又 [jiw<sup>32</sup>] (Siyi) – although a pause particle may follow or punctuate the first disjunct. For example, in Suzhou:

- (37) 俚 是 蘇州人 啲 還是 溫州人  
 ʔij zɿ sɜutsɔŋim nɛ fiɛzɿ ʔuəntsəʔɿim  
 3sg be Suzhou-person PRT cnj Wenzhou-person  
 'Is s/he from Suzhou or from Wenzhou?'

23 There is a historical explanation for this asymmetry. According to Lyu (1942), 嗎 is derived from the combination of a negative plus a final particle.

In some dialects, for example, in Zǎozhuang (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 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, Southwestern, Jianghuai) as well as Central (Northern Wu) and Southern dialects (Southern 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  $tɕ'y^{51}$  不  $pu$  去  $tɕ'y^{51}$   
 2sg go neg go  
 ‘Are you going?’

When the 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 (Yùchéng), 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 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 historical comparison of Pekinese spoken at the beginning of the twentieth century, according to Zhang (1990: 72), and 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 Southwestern Mandarin, especially Hubei and Sichuan, but also in Shandong (Jiaodong Peninsula) and Manchuria. For example, in Èzhou:

- (42) 吃 冒 吃 飯 啊  
 eat neg-pfv-asp eat rice QP  
 ‘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 Fúzhou (Min), Liánchéng (Hakka) and the latter in some Wu dialects such as Shàoxīng, Zhūjì, Wǔyì, Jīnhuá, Shèngxiàn.

The VP-neg type has a long recorded history since at least Qin times<sup>24</sup> and is native to many Southern and Central dialects as well as the peripheral areas (Shandong, Northwestern dialects of Shanxi, Shaanxi, Gansu, Qinghai) of the Northern dialects. The neg takes different forms in different dialects and depending on whether the VP contains ASP. In the Northern dialects, it is often 不 or 没 (if the VP is marked with ASP). In the Gan dialects, it is mostly [mo]/[mɔ] or [po]/[pɔ] and in the Hakka dialects it is 無 [mɔ]/[mɔu]. In the Yue dialects, it is mostly 嗎 [ma] (which is derived from a contraction of the general negative marker [m] plus the final particle [a]) or 未 [mei<sup>22</sup>] (neg for the perfective)/[mian<sup>11</sup>] (a contracted form of 未曾 in for example, the Siyi dialects), for example:

- (43) Kaiping (Yue) 佢 講 得 現 嗎  
 3sg speak COMPL-mkr clear neg + QP  
 'Can s/he speak clearly?'

In the Wu dialects, neg is mostly 勿 [vɛʔ] or 伐 [vɑ] irrespective whether VP contains ASP (in some dialects, neg is [vəʔ] (<勿+曾) if VP contains ASP). For example:

- (44) Jiaxìng (Wu) 伊 ʔi 來 IEɛ 哩 li 伐 va  
 3sg come pfv-asp NEG + QP  
 'Has s/he come?'

In the Southern 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 types of verbs and on aspectual/modal types used in the VP: [m] (non-aspectual), [bo] (past), [be] (perfective), or [bue] (probability/possibility) in for example Quánzhōu. For example:

- (45) Quánzhōu 你 明 [na<sup>24</sup>] [byʔ<sup>4</sup>] 來 [m<sup>21</sup>]  
 2sg tomorrow will come neg  
 'Are you coming tomorrow?'

- (46) Xiàmén 有 u 芳 p'ɑŋ 無 bo  
 af-asp fragrant neg  
 'Is it fragrant?'

The ADV-VP type, prevalent along the southeastern coast of China,<sup>25</sup> 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 [aʔ], [a], [aʔ], [a], [əʔ], [ɿʔ], [həʔ] or a velar initial in such forms as [kəʔ], [k'əʔ], [kə], [k'aʔ],

24 不 serving as neg in VP-neg appears in the bamboo slips unearthed from Qin tombs in Shuǐhǔdì of Yúnmèng in Hubei.

25 Only Yunnan geographically lies outside this southeastern region. However, according to Zhang (1990), Han immigrants into Yunnan are mostly from Anhui province.

[ka], [kiʔ], [kʊ], [kɔu], [ku], [kɻu], [xaʔ], [xɑ], [xaʔ], [xa], [xɐe], [xɛ], [xəʔ]. The question may also take an optional question particle at the end. For example:

- (47) Suzhou (Wu) 俚 *li* 阿 *ʔaʔ* 曉得 *ɕiətəʔ* [tɕiɔ]  
 3sg ADV know QP  
 ‘Does s/he know?’

In some dialects, the ADV is followed by a negative marker. For example, in Lóngnán (Jiangxi Hakka) the ADV is [æ<sup>55</sup>] followed by the negative marker 不 :

- (48) a. 你 的 瓜 [æ<sup>55</sup>] 不 甜  
 2sg attr-mkr melon ADV neg sweet  
 ‘Is your melon sweet?’  
 (attr-mkr = attributive marker)

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

- b. Mópíng 可 不 來  
 ADV neg come  
 ‘(Are you) coming?’  
 c. Mópíng 飯 是 没 熟  
 rice ADV neg + pfv-asp ripe  
 ‘Is the rice done?’

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

- (48) d. 飯兒 (是) 没 涼  
 rice ADV neg + pfv-asp 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 Yílan (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 Shàntou dialect (Min) illustrate this complex usage:

- (49) (i) 伊 是 [nia<sup>33</sup>] 弟 (啊) [mi<sup>35</sup>]  
 (ii) 伊 [k'aʔ<sup>2</sup>] 是 [nia<sup>33</sup>] 弟  
 (iii) 伊 [k'aʔ<sup>2</sup>] 是 [nia<sup>33</sup>] 弟 (啊) [mi<sup>35</sup>]  
 3sg ADV be 2gen younger-brother (cnj) neg  
 ‘Is he your younger brother?’  
 (iv) 伊 是 [m<sup>35</sup>] 是 [nia<sup>33</sup>] 弟  
 3sg be neg be your younger-brother  
 (v) 伊 [k'aʔ<sup>2</sup>] 好 來 (啊) [m<sup>35</sup>] 好  
 3sg ADV willing come PRT neg willing  
 ‘Is s/he willing to come?’





- (53) Xiangxiang (Xiang) 狹 我 錢 / 狹 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 係 / 我 撥 係 本 書  
*ŋeu pəʔ pən sɿ nɛ / ŋeu pəʔ nɛ pən sɿ*  
 1sg give CL book 2sg/1sg give 2sg CL book  
 'I gave him 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 [*pa*<sup>213</sup>] 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:

- (55) 他 *ta*<sup>55</sup> 把 *pa*<sup>213 > 11</sup> 那 *na*<sup>51</sup> 個 *kə* 蘋果 *p'iy*<sup>35</sup> *kwɔ* 吃 *tʂ'i*<sup>55</sup> 了 *lə*  
 s/he BA that CL apple eat FP  
 'S/he ate that apple.'
- (56) 他 *ta*<sup>55</sup> 把 *pa*<sup>213 > 11</sup> 正經事 *tʂəŋ*<sup>51</sup> *tʂiŋ*<sup>55</sup> *ʂi*<sup>51</sup> 當作 *taŋ*<sup>51</sup> *tswɔ*<sup>51</sup>  
 3sg BA serious matter regard  
 兒戲 *ər*<sup>35</sup> *ci*<sup>51</sup>  
 child-play  
 'S/he regards serious matters as jokes.'
- (57) \* 他 *ta*<sup>55</sup> 把 *pa*<sup>213 > 11</sup> 三 *san*<sup>55</sup> 個 *kə* 蘋果 *p'iy*<sup>35</sup> *kwɔ* 吃 *ʂ'i*<sup>55</sup> 了 *lə*  
 3sg BA three CL apple eat FP  
 'S/he ate three apples.'
- (58) \* 他 *ta*<sup>55</sup> 把 *pa*<sup>213 > 11</sup> 那 *na*<sup>51</sup> 個 *kə* 蘋果 *p'iy*<sup>35</sup> *kwɔ* 吃 *tʂ'i*<sup>55</sup>  
 3sg BA that CL apple eat  
 'S/he eats that apple.'

the last two are ungrammatical since the NP2 ('three apples') in (57) is indefinite and the VP of (58) is not specific.

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

- (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 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) 我 把 鎖 打 得 開  
 1sg BA lock hit can open  
 'I can open the lock.'

and the Lánzhou (Gansu) and the Wèinán dialect (Shaanxi) tolerate the negative in the VP:

- (61) Wèinán 把 門 不 要 老 閉 著  
 BA door neg need always closed dur-asp  
 '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 Lánzhou, for example, the following sentences are ungrammatical in standard Mandarin but perfectly grammatical in Lánzhou:

- (62) 我 把 他們 的 名字 知道  
 1sg BA 3pl attr-mkr 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: 給 (Luòyáng of Henan; Jiaochéng, Hándan of Shanxi), [*pə*] and its variants (Línfén of Shanxi, some Xiang and many Wu dialects), [*ma*] (Qingjiàn of Shaanxi, Huángméi of Hubei), 幫 (many Southwestern Mandarin dialects in Yunnan, some Huizhou and Wu and Min dialects), 挨 (many Southwestern Mandarin dialects in Yunnan), [*na*] and its variants (some Gan, Xiang, Wu, Min dialects), 將 (Yue, Hakka, Min), etc.

#### 4.7 The 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 described as the passive construction in Chinese. One type has the general structure of NP<sub>patient</sub>+PASS (+NP<sub>agent</sub>)+VP, characterized by the option of having the agent expressed or unexpressed. This type occurs only in Northern Chinese where the passive marker (PASS) is [*pej*<sup>51</sup>] or [*kej*<sup>213</sup>] in SM. This passive marker is often mis-analysed as an agent marker. The fact that it can occur without an expressed agent argues against this interpretation. For example in SM:

- (63) 我 wə<sup>213>11</sup> 被 *pej*<sup>51</sup> 騙 p'iɛn<sup>51</sup> 了 lə  
 1sg PASS deceive FP  
 'I was deceived.'

Therefore [*pej*<sup>51</sup>] or [*kej*<sup>213</sup>] is best understood as a passive marker. This also accords well with the historical development of the passive marked with 被 [*pej*<sup>51</sup>] in which the agentless form predates the same with the agent appearing.

The other type of passive construction has the general structure of NP<sub>patient</sub>+PASS+NP<sub>agent</sub>+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:

- (64) a. 我  $\eta\alpha^{24}$  畀  $pej^{35}$  人  $j\eta n^{11}$  [ $\eta ak^5$ ] [ $t\alpha^{35}$ ] (我  $\eta\alpha^{24}$ )<sup>28</sup>  
 1sg PASS people deceive pfv-asp 1sg  
 ‘I was deceived (by someone).’

The passive marker may be regarded as the agent marker. This second type of passive 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 [ $z\alpha\eta^{51}$ ] or 叫 [ $t\alpha jaw^{51}$ ] or 教 [ $t\alpha jaw^{55}$ ] and is popular in colloquial speech. Its equivalent to (63) is:

- b. 我  $w\alpha^{213 > 11}$  讓  $z\alpha\eta^{51}$  / 叫  $t\alpha jaw^{51}$  人  $z\alpha n^{35}$  騙  $p'i\epsilon n^{51}$  了  $l\alpha$   
 1sg PASS people deceive pfv-asp  
 ‘I was deceived (by someone).’

A prominent characteristic of these two types of passive construction is their association with undesirable events, especially in colloquial speech. In some dialects, only unfavourable events can be expressed with this type of passive while favourable events must be expressed in the active. For example in the Yantái dialect (Shandong), while (65a) is grammatical:

- (65) a. 他 叫 他 婆 打 了  
 3sg PASS 3sg wife hit FP  
 ‘He was hit by his wife.’

(65b) is not and must be rendered as (65c):

- b. \*他 叫 他 婆 誇 了  
 3sg PASS 3sg wife praise FP  
 ‘He was praised by his wife.’  
 c. 他 婆 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 passive sentence, it takes on an unfavourable meaning. For example in SM, the verb ‘to hear’ has no undesirable meaning in (66a):

- (66) a. 我  $w\alpha^{213 > 11}$  說  $s w\alpha^{55}$  的  $t\alpha$  話  $x w\alpha^{51}$  他  $t'a^{55}$   
 1sg say attr-mkr word 3sg  
 聽見  $t'i\eta^{55} t\alpha j\epsilon n$  了  $l\alpha$   
 hear FP  
 ‘What I said was overheard by him/her.’

but carries unfavourable connotation in (66b):

- b. 我  $w\alpha^{213 > 11}$  說  $s w\alpha^{55}$  的  $t\alpha$  話  $x w\alpha^{51}$  讓  $z\alpha\eta^{51}$  他  $t'a^{55}$   
 1sg say attr-mkr word PASS 3sg  
 聽見  $t'i\eta^{55} t\alpha j\epsilon n$  了  $l\alpha$   
 hear FP  
 ‘What I said was (unfortunately) overheard by him/her.’

28 Notice that the patient may be repeated after the main verb.

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

- (67) 這  $t\zeta\partial^{51}$  孩子  $xaj^{35}tsɿ$  常常  $t\zeta'an^{35}t\zeta'an$  被  $bej^{51}$  老師  $law^{213>11} \zeta r^{55}$   
 this child often PASS teacher  
 稱讚  $t\zeta'\partial\eta^{55}tsan^{51}$   
 praise  
 'This child is often praised by the teacher.'
- (68) 他  $t'a^{55}$  被  $pej^{51}$  選  $\zeta\eta\epsilon n^{213>11}$  爲  $wej^{35}$  總統  $ts\partial\eta^{213>35} t'\partial\eta^{213}$   
 3sg PASS elect be president  
 'S/he was elected president.'

It is often observed that the broadened usage of the passive 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 passive has certainly widened, it should not be forgotten that even in ancient times certain verbs such as 寵 'to favour' have always been freely used in the passive construction.<sup>29</sup>

An interesting aspect that relates the Chinese dialects to their neighbouring languages is the identity of the passive marker with the verb 'to give' and/or with the causative marker. In the Northern dialects, the passive 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 (Bāngbù, Língbǐ, Sùxiàn, Suixi, Dangshan, Fùyáng, Fùnán, Línquán, Jièshǒu, Tàihé, Háoxiàn, Woyáng, Fèngtái, Shòuxiàn, Yíngshàng, Huòqiú, Jìnzài) and Jianghuai dialects (Wúwéi, Lújiāng, Shuchéng, Huáinán, Huáiyuǎn, Dìngyuǎn, Jiāshan, Lái'an, Tongchéng). 讓 is used in both functions in Jianghuai (Mǎ'ānshan), Xiang (Nánlíng), and Gan (Wàngjiāng, Tàihú, Qiánshan, Yuèxi, Sùsong) dialects. For these double functions there are other markers such as 得 or [dei] (Línfén in Shanxi), [tso] (Jìchéng in Shanxi) or [tsau] (Dìngxi, Tongwèi, Lóngdié of Gansu) and in Southwestern Mandarin [tsau] (Chéngdū) or [tsuo] (Kunmíng) or [tsok] (Nánfíng of Guangxi).

In the Southern, the Wu, some Xiang (Línwǔ) and Gan dialects the passive 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 [tɛ] (Xíníng in Qinghai), or in Jianghuai [te] (Rúgāo in Jiangsu) or [ha] (Nántóng in Jiangsu) and 給 in Jianghuai (Héféi, Féidong, Féixi, Hánshan, Héxiàn, Lújiāng, Huòshan, Huáinán, Dìngyuǎn, Chúxiàn, Quánjiāo, Tongchéng), Huizhou (Shèxiàn, Jìxi, Túnxi, Xiúníng, Qímén, Níngguó) and Gan (Huáiníng, Sùsong). In some dialects, the different usages may correlate with different pronunciations: for example [kɛ<sup>11</sup>] as passive marker but [kɛ<sup>24</sup>] as the verb 'to give' in the Huáiyin dialect (Jianghuai). In a number of Northern dialects, one and the same marker 給 (Jiāochéng of Shanxi) or 把 (Xìshǔi of Hubei, Yángzhou of Jiangsu, Lángxi of Anhui, also the Gan dialect of Wàngjiāng in Anhui), serves as the verb 'to give', the passive marker, and the disposal marker.

29 Wang (1989: 285) considers these exceptions as falling into a fixed category of 'favour from above' which, just as calamities, cannot be opposed.

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 passive marker, and the causative marker. For example in Cantonese:

- (69) 佢 *k'øy*<sup>24</sup> [*m*<sup>11</sup>] 畀 *pej*<sup>35</sup> 我 *ŋɔ*<sup>24</sup> 畀 *pej*<sup>35</sup> 錢 *tʃ'in*<sup>35</sup> 你 *nej*<sup>24</sup>  
 3sg neg let 1sg give money 2sg  
 ‘S/he does not let me give you money.’
- (70) 我 *ŋɔ*<sup>24</sup> [*mou*<sup>24</sup>] 畀 *pej*<sup>35</sup> 佢 *k'øy*<sup>24</sup> 打 *ta*<sup>35</sup> 我 *ŋɔ*<sup>24</sup>  
 1sg neg + pfv-asp PASS/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’ (passive) is ‘to cause to give’ (causative).

#### 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 are topicalized. The semantic as well as syntactic relationship between these two NPs is diverse. It may be genitive or partitive, falling into roughly four types. For example:

- (71) SM 象 *ɕjan*<sup>51</sup> 鼻子 *pi*<sup>35</sup> *zɪ* 長 *tʂ'an*<sup>35</sup>  
 elephant nose long  
 ‘The elephant has a long nose.’
- (72) SM 那 *na*<sup>51</sup> 個 *kə* 孩子 *xaj*<sup>35</sup> *zɪ* 爸爸 *pa*<sup>51</sup> *pa* 是 *ʂ'*<sup>51</sup> 大夫 *taj*<sup>51</sup> *fu*  
 that CL child father be doctor  
 ‘That child has a doctor father.’
- (73) Cantonese 中國 *tʃuŋ*<sup>55</sup> *kwɔk*<sup>44</sup> 人口 *jən*<sup>11</sup> *hɛw*<sup>35</sup> 多 *tɔ*<sup>55</sup>  
 China population numerous  
 ‘China has a big population.’
- (74) SM 三 *san*<sup>55</sup> 個 *kə* 孩子 *xaj*<sup>35</sup> *zɪ* 兩 *jan*<sup>213</sup> *>11* 個 *kə* 病 *piŋ*<sup>51</sup> 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 *ɕjan*<sup>51</sup> 鼻子 *pi*<sup>35</sup> *zɪ* 長 *tʂ'an*<sup>35</sup> is not the same in structure as 象 *ɕjan*<sup>51</sup> 的 *tə* 鼻子 *pi*<sup>35</sup> *zɪ* 長 *tʂ'an*<sup>35</sup> ‘the trunk of the elephant is long’. There are at least three differences. Often an adverb can occur between the NPs, for example:

- (73) a. Cantonese 中國 *tʃuŋ*<sup>55</sup> *kwɔk*<sup>44</sup> 的確 *tuk*<sup>5</sup> *k'ɔk*<sup>44</sup> 人口 *jən*<sup>11</sup> *hɛw*<sup>35</sup> 多 *tɔ*<sup>55</sup>  
 China certainly population numerous  
 ‘China certainly has a big population.’

A pause particle may occur between them:

- (72) a. SM 那 *na*<sup>51</sup> 個 *kə* 孩子 *xaj*<sup>35</sup> *zɪ* 呀 *ja* 爸爸 *pa*<sup>51</sup> *pa* 是 *ʂ'*<sup>51</sup> 大夫 *taj*<sup>51</sup> *fu*  
 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):

- (71) a. SM 鼻子  $pi^{35}zi$  長  $t\dot{s}'a\eta^{35}$  的  $t\dot{\theta}$  象  $\epsilon ja\eta^{51}$   
 nose long attr mkr elephant  
 'Elephants which have long trunks.'

The partitive type, on the other hand, may be understood as derived from a phrase with the meaning of 'among ...', namely, (74) may be interpreted as follows:

- (74) a. 三  $san^{55}$  個  $k\dot{\theta}$  孩子  $xaj^{35}zi$  裡頭  $li^{213>11}$   $tow$  兩  $ljaj^{213>11}$   
 three CL child inside two  
 個  $k\dot{\theta}$  病  $pi\eta^{51}$  了  $l\dot{\theta}$   
 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. The major difference between the Northern and Central (but including northern Min and most Hakka) dialects 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+CPR+Vcomp+NP2+(MP) or NP1+(ADV)+Vcomp+CPR+NP2+(MP), where MP stands for Measure Phrase. The ADV is limited to those with the meaning of 'even'. The CPR is most commonly 比 [ $pi^{213}$ ] in SM) in Northern Chinese but varies in Southern Chinese: [ $k'a?$ ] in Southern Min, 過 [ $kw\dot{\sigma}^{44}$ ] in Yue, 起 in Shandong (Móupíng, Zhuchéng, Píngdù, Wéifāng, Zibó). The following examples provide contrast in word order between the North and the South:

- (75) a. SM 他  $t'a^{55}$  比  $pi^{213>35}$  我  $w\dot{\sigma}^{213>11}$  高  $kaw^{55}$   
 b. Xiàmén 伊  $i$   $k'a?$   $lo$   $gwa$   
 (Min)  
 c. Cantonese  $k'\emptyset y^{24}$   $kow^{55}$  過  $kw\dot{\sigma}^{44}$  我  $\eta\dot{\sigma}^{24}$   
 d. Móupíng 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):

- (76) a. SM 他  $t'a^{55}$  高  $kaw^{55}$  我  $w\dot{\sigma}^{213>11}$  一  $ji^{55}$  個  $k\dot{\theta}$  頭  $t'ow^{35}$   
 3sg be-tall 1sg one CL head  
 'S/he is a head taller than I am.'  
 b. Xiangxiang 我 大 你 三 歲  
 1sg old 2sg three year  
 'I am three years older than you.'

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 (a):

- (77) SM 我  $wɔ^{213>11}$  書  $su^{55}$  比  $pi^{213>35}$  你  $ni^{213>11}$  多  $twɔ^{55}$   
 1sg book CPR 2sg numerous

while Socrates’ famous saying must be given with alternative (b):

- (78) SM 我  $wɔ^{213>11}$  愛  $aj^{51}$  我  $wɔ^{213>11}$  的 老師  $law^{213>11}$   $ʂi^{55}$ ,  
 1sg love 1sg ATTR teacher,  
 可是  $k'ɔ^{11}$   $ʂi$  我  $wɔ^{213>11}$  更  $kɲŋ^{51}$  愛  $aj^{51}$  真理  $tʂɻn^{55}$   $li^{213}$   
 but 1sg even more love truth  
 ‘I love truth more than my teacher.’

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  $t'a^{55}$  高  $kaw^{55}$  (還是  $xaj^{35}$   $ʂi$ ) 你  $ni^{213>11}$  高  $kaw^{55}$ ? – 他  $t'a^{55}$   
 3sg be-tall (or) 2sg be-tall – 3sg  
 高  $kaw^{55}$   
 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 implicates 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) a. SM 我  $wɔ^{213>11}$  跟  $kɻm^{55}$  你  $ni^{213>11}$  一樣  $ji^{55}$   $jaŋ^{51}$  高  $kaw^{55}$   
 b. Shanghai 我 搭 儂 一樣 高  
 (Wu)

c. Méixiàn (Hakka)	我	同	你	一般	高
d. Cantonese (Yue)	我 $\eta\tau^{35}$	同 $t'u\eta^{11}$	你 $nei^{24}$	一樣 $j\epsilon t^5 j\theta\eta^{22}$	高 $kow^{55}$
e. Xiàmén (Min)	我	合	汝	平(平)	懸
	1sg	cnj	2sg	same	be-tall
	'I am of your same height.'				

and with 'resemble':

(81) a. SM	你 $ni^{213 > 11}$	像 $\zeta i a \eta^{51}$	他 $t'a^{55}$	那樣 $na^{51} ja \eta$	健壯 $t\zeta i e n^{51} t\zeta w a \eta^{51}$
b. Cantonese (Yue)	你 $nei^{24}$	似 $ts'i^{24} k' \phi y^{24}$		[ $k\epsilon m^{44}$ ]	壯健 $t\zeta \eta^{44} kin^{22}$
c. Quánzhōu (Min)	汝	親像	若伊		奘勇
	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, since the affirmative forms sometimes convey the same meaning between these two types. For example:

(82) a. SM	我 $w\tau^{213 > 11}$	没有 $mej^{35} jow$	你 $ni^{213 > 11}$	(那麼 $na^{51} m\theta$ )	高 $kaw^{55}$
b. Cantonese	我 $\eta\tau^{35}$	mow <sup>24</sup>	你 $nei^{24}$	[ $k\epsilon m^{44}$ ]	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.

### 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-mkr + NPh where NPh signifies the Head Noun. The attr-mkr 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 (Huáiyin, Shùyáng) and Wu (Chángzhou, Suzhou) dialects as well as Southern Min and Yue, where the NPh must contain a DEM in at least the Min and the Yue dialects. For example:

(83) a. Huáiyin	人家	是	漂漂亮亮	大	姑娘
	IMPERS	be	beautiful	big	girl
	'She is a beautiful young girl.'				



- b. Kaipíng (Yue) 暖暖  $n\text{ɔn}^{44}n\text{ɔn}^{35}$  [neɪŋ<sup>21</sup>] [kaw<sup>32</sup>] 番薯  $fan^{44}si^{35}$   
 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:

- (84) a. Héngyáng 你 手 上 隻 菌子 吃不得, 有 毒  
 (Xiang) 2sg hand on CL mushroom eat-not-can, have poison  
 ‘The mushrooms in your hand are inedible; they are poisonous.’  
 b. 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 clauses can be distinguished when the head noun is marked with a determiner that contains a demonstrative. When the modifying clause is preceded by 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:

- (85) a. 他  $t'a^{55}$  把  $pa^{213>11}$  那  $na^{51}$  枝  $tʂ^{55}$  新  $ʃin^{55}$  買  $maj^{213>11}$   
 3sg BA that CL new buy  
 的  $tə$  筆  $pi^{213>11}$  丟  $tiv^{55}$  了  $lə$   
 attr-mkr pen lost FP  
 ‘He lost that newly bought pen.’  
 b. 他  $t'a^{55}$  把  $pa^{213>11}$  新  $ʃin^{55}$  買  $maj^{213>11}$  的  $tə$  那  $na^{51}$   
 3sg BA new buy attr-mkr that  
 枝  $tʂ^{55}$  筆  $pi^{213>11}$  丟  $tiv^{55}$  了  $lə$   
 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:

- (86) a. \*新  $ʃin^{55}$  買  $maj^{213>11}$  的  $tə$  一  $ji^{55}$  枝  $tʂ^{55}$  筆  $pi^{213>11}$   
 new buy attr-mkr one CL pen  
 丟  $tiv^{55}$  了  $lə$   
 lost FP  
 b. \*有  $jow^{213>11}$  新  $ʃin^{55}$  買  $maj^{213>11}$  的  $tə$  一  $ji^{55}$  枝  $tʂ^{55}$   
 exist new buy attr-mkr one CL  
 筆  $pi^{213>11}$  在  $tsaj^{51}$  這兒  $tʂər^{51}$   
 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:

- (87) a. 他  $t'a^{55}$  買  $maj^{213>11}$  的  $tə$  原因  $qen^{35} jin^{55}$   
 3sg buy attr-mkr reason  
 'The reason he bought (something)'.
- b. \* 他  $t'a^{55}$  買  $maj^{213>11}$  的  $tə$   
 'for which (reason) he bought'
- (88) a. 他  $t'a^{55}$  買  $maj^{213>11}$  的  $tə$  東西  $ɔŋ^{55} ci$   
 3sg buy attr-mkr thing  
 'The thing he bought'.
- b. 他  $t'a^{55}$  買  $maj^{213>11}$  的  $tə$   
 3sg buy attr-mkr > NOMZR  
 'that which he bought'.

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 d-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 attr-mkr in (87a) or (88a), which occurs with the head noun, and attr-mkr in (88b), which does not occur with the head noun, the former should be called a clause marker (attr-mkr) and the latter a nominalizer. In most dialects, the attr-mkr and the NOMZR share the same form, either a classifier or some form of 的. 的 assumes the phonetic shape of [tə] or [ni]/[nə]/[ne] (Southwestern 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 ([kɔ<sup>44</sup>] in Cantonese) and its cognates with a velar initial are widely used in the Southern and Central dialects as well as a number of Southwestern Mandarin dialects distributed mainly in the southern periphery of Hunan (Xintián, Gùiyáng, Yízhāng, Línwǔ, Lánshān, Níngyuǎn, Dàoxiàn, Jiānghuá, Jiāngyǒng, Fènghuáng). 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 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] subjects, 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  $tɕ'y^{51}$  可以  $k'y^{213>35}$   $ji$  嗎  $ma$   
 1sg neg go to-be-all-right QP  
 '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 sentence as object. For example in Cantonese:

- (90) 大家 *taj*<sup>22</sup>*ka*<sup>55</sup> 都 *tow*<sup>55</sup> 知道 *tʃi*<sup>55</sup> *tow*<sup>44</sup> [*mow*<sup>24</sup>] 電話 *tin*<sup>22</sup> *wa*<sup>35</sup>  
 every one all know have-not telephone  
 [*m*<sup>11</sup>] 方便 *fɔŋ*<sup>55</sup> *pin*<sup>22</sup>  
 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 complementisers (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 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*]/[*təʔ*]/[*tiəʔ*]/[*lei*]/[*ləʔ*]/[*ně*] (Shanxi dialects) or [*te*] (Guìyáng of SW Mandarin) or [*tsəʔ*]/[*tseʔ*]/[*tsɛʔ*]/[*təʔ*]/[*tɛʔ*]/[*dəʔ*]/[*tɐʔ*]/[*ləʔ*]/[*lɐʔ*]/[*læ*]/[*lɛ*]/[*ɔ*] (Wu) or [*tɐk*<sup>5</sup>] (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:

- (91) a. SM 他 *t'a*<sup>55</sup> 跑 *p'aw*<sup>213>11</sup> 得 *tə* 很 *xən*<sup>213>11</sup> 快 *k'waj*<sup>51</sup>  
 b. Chángsha 他 跑 起 好 快  
 c. Cantonese 佢 *k'øj*<sup>24</sup> 走 *tʃw*<sup>35</sup> 得 *tɐk*<sup>5</sup> 好 *how*<sup>35</sup> 快 *faj*<sup>44</sup>  
 d. Kaipíng 佢 *k'uy*<sup>44</sup> 走 *taw*<sup>55</sup> 快 *faj*<sup>35</sup> 快 *faj*<sup>44</sup>  
 e. Miáoli 其 走 [*t'ɔŋ*<sup>24</sup>] 快 *faj*<sup>44</sup>  
 f. 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.<sup>30</sup> For example:

- (92) a. SM: (吃 *tʃɪ*<sup>55</sup>) 藥 *jaw*<sup>51</sup> 吃 *tʃɪ*<sup>55</sup> 得 *tə* 多 *twɔ*<sup>55</sup>  
 b. Chángsha 吃 得藥 有  
 c. Cantonese 吃 *sik*<sup>2</sup> 得 *tɐk*<sup>5</sup> 藥 *jøk*<sup>22</sup> 多 *tɔ*<sup>55</sup>  
 eat medicine eat COMP medicine much  
 ‘Take a lot of medicine.’

30 In the Yue dialects, the Northern option is also used.

- (93) Chángshú (Wu) 菜 *ts'æ* 燒 *ɣɔ* 來 *læ* 交關 *tɕiɔkuə* 好 *xɔ*  
 food cook COMP very good  
 'The food is very well cooked.'

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 [*tow*<sup>44</sup>] in Yue, [*kaʔ*<sup>32</sup>] or [*kaw*<sup>31</sup>] in S. Min. Within Northern Chinese, some Shanxi dialects assign different markers for these two types of complements. Manner COMPLs are found in examples (94) while Extent COMPLs are found in (95).

- (94) a. Jíxiàn 你 說 得 *tei* 好  
 b. Xiangyuán 你 說 得 *təʔ*<sup>5</sup> 真 好  
 c. Qinxian 你 說 勒 *ləʔ*<sup>5</sup> 楞 好  
 d. Cantonese 你 *nej*<sup>24</sup> 講 *kɔŋ*<sup>35</sup> 得 *təʔ*<sup>5</sup> 好 *how*<sup>35</sup>  
 2sg speak COMP very good  
 'You speak very well.'
- (95) a. Jíxiàn 重 [*li*] 我 都 荷 不 起 啦  
 b. Xiangyuán 重 得 *tiəʔ*<sup>3</sup> 連 我 也 [*xəŋ*<sup>213</sup>] 不 動 了  
 c. Qinxian 扎實 得 *tiəʔ*<sup>3</sup> 連 我 也 荷 不 動 了  
 d. Cantonese 重 *tʃ'ʊŋ*<sup>24</sup> 到 *tow*<sup>44</sup> 我 *ŋɔ*<sup>24</sup> 都 *tow*<sup>55</sup> [*ŋ*<sup>11</sup>] 抬 *t'ɔj*<sup>24</sup>  
 heavy COMP 1sg also neg lift  
 得 *tək*<sup>5</sup> [*jok*<sup>5</sup>]  
 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. For example, the subject NP of VP1 can appear before VP2 if VP1 and VP2 share the same subject NP:

- (96) a. SM 他 *t'a*<sup>55</sup> 跑 *p'aw*<sup>213 > 11</sup> 得 *tə* 累 *lej*<sup>51</sup> 極 *tɕi*<sup>35</sup> 了 *lə*  
 3sg run COMP tired extreme FP  
 'He is extremely tired from running.'  
 b. SM 跑 *p'aw*<sup>213 > 11</sup> 得 *tə* 他 *t'a*<sup>55</sup> 累 *lej*<sup>51</sup> 極 *tɕi*<sup>35</sup> 了 *lə*  
 'The running makes him extremely tired.'

but this switch is impossible for constructions with the Manner COMPL; therefore (91a) cannot be turned into (91f):

- (91) f. SM \*跑 *p'aw*<sup>213 > 11</sup> 得 *tə* 他 *t'a*<sup>55</sup> 很 *xən*<sup>213 > 11</sup> 快 *k'waŋ*<sup>51</sup>  
 \*'The running makes him extremely fast.'

#### 4.12.2 Resultative and directional

The Resultative COMPL is also called the Causative COMPL. Constructions with this complement consist 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



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, Chángzhou, Suzhou, Kunshan, Bǎoshan, Shanghai, Wújiāng, Jiāxìng, Húzhōu, Hángzhōu, Shàoxìng, Zhūjì, Shèngxiàn, Yúyáo, Níngbō, Huángyán, Wenzhou, Qúzhōu, Jīnhuá) and Yue dialects,<sup>32</sup> for example:

- (100) a. Dòngkǒu 果 隻 塘 裡 捉 得 蠻 多 魚 倒  
 this CL pond-in catch POT very many fish obtain  
 ‘Many fish can be caught in this pond.’
- b. Jīnhuá 佢 寫 得 字 來  
 3sg write POT word obtain  
 ‘S/he can write.’
- c. Cantonese 搞 *kaw*<sup>35</sup> 得 *tek*<sup>5</sup> 佢 *køy*<sup>24</sup> [*tim*<sup>22</sup>]  
 make POT 3sg straighten  
 ‘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. While the first alternative is most popular, the second alternative occurs in Chángsha and the Yue dialects and the last one in Chángsha and some Wu (Shèngxiàn Chónggrén, Níngbō) dialects only. For example:

- (101) a. Héngyáng 狗 追 貓 不 到  
 (Xiang) dog pursue cat neg reach  
 ‘The dog cannot catch the cat.’
- b. Shàoxìng 我 *ŋo* 打 *taŋ* 伊 *fi* 勿 *vi?* 過 *ku*  
 c. Cantonese 我 *ŋo*<sup>24</sup> 打 *ta*<sup>35</sup> 佢 *køy*<sup>24</sup> [*m*<sup>11</sup>] 過 *kwɔ*<sup>44</sup>  
 1sg hit 3sg neg surpass  
 ‘I am no match in fighting him/her.’
- (102) a. Chángsha 冒 買 得 肉 到  
 neg- pfv-asp buy POT meat obtain  
 ‘Did not succeed in buying meat.’
- b. Cantonese 你 *nej*<sup>24</sup> [*m*<sup>11</sup>] 打 *ta*<sup>35</sup> 得 *tek*<sup>5</sup> 佢 *køy*<sup>24</sup> 死 *sej*<sup>35</sup>  
 2sg neg hit POT 3sg die  
 ‘You cannot beat him/her to death.’
- (103) a. Chángsha 他 吃 不 飯 進  
 3sg eat neg rice enter  
 ‘S/he cannot eat (send food down).’
- b. Shèngxiàn Chónggrén 我 *ŋɿ* 打 *tā* 勿 *ve?* 伊 *fi* 過 *kr*  
 1sg hit neg 3sg surpass  
 ‘I am no match in fighting him/her.’

32 In the modern Yue dialects, this word order obtains only if O is a pronoun; while in the Wu dialects, except for Shàoxìng, Yúyáo, and Wenzhou, this word order co-exists with the Northern word order.

The Directional COMPL is actually another type of resultative complement where v2, followed by a directional suffix (*dir*), indicates the direction of the movement of v1 and is a directional verb (*vdir*). v1 is usually a motion verb (*vm*) or action verb. The *dir* is either 來 ‘towards’ 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:

- (104) a. SM 跑  $p'aw^{213>11}$  回  $xwej^{35}$  來  $laj$   
 b. Cantonese 走  $tsɛw^{35}$  翻  $fan^{55}$  來  $lɛj^{11}$   
 run return towards  
 ‘Run back (towards speaker).’
- (105) a. SM 跑  $p'aw^{213>11}$  回  $xwej^{35}$  家  $jia^{55}$  來  $laj$   
 b. Cantonese 走  $tsɛw^{35}$  翻  $fan^{55}$  屋企  $ɔk^5k'ej^{35}$   
 run return home towards  
 ‘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:

- (106) a. SM 拿  $na^{35}$  兩  $na^{213>35}$  本  $pən^{213>11}$  書  $su^{35}$  出  $ts'u^{55}$  來  $laj$   
 b. SM 拿  $na^{35}$  出  $ts'u^{55}$  兩  $na^{213>35}$  本  $pən^{213>11}$  書  $su^{35}$  來  $laj$   
 c. SM 拿  $na^{35}$  出  $ts'u^{55}$  來  $laj$  兩  $na^{213>35}$  本  $pən^{213>11}$  書  $su^{35}$   
 d. Cantonese [nɪŋ<sup>55</sup>] 兩  $lɔŋ^{24}$  本  $pun^{35}$  書  $sy^{55}$  出  $tʃøt^5$  來  $lɛj$   
 take exit towards two CL book exit towards  
 ‘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 mislabelled ‘prepivotals’ 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:

- (107) a. SM 你  $ni^{213>11}$  勸  $tɕ'qɛn^{51}$  他  $ta^{55}$  千萬  $tɕ'jɛn^{55}$  wan 要  $jaw^{51}$   
 來  $laj^{35}$   
 b. Cantonese 你  $nej^{24}$  勸  $hyn^{44}$  佢  $k'øj^{24}$  千祈  $tʃin^{55}k'ej^{11}$   
 2sg persuade 3sg by-all-means  
 要  $jiw^{44}$  來  $lɛj^{11}$   
 must come  
 ‘You persuade him to come by all means.’

where ‘by all means’ can only be used in an imperative sentence.

- (108) a. SM 你  $ni^{213>11}$  叫  $t\zeta jaw^{51}$  我  $\eta\zeta^{213>11}$  別  $pj\epsilon^{35}$  給  $kej^{213>11}$  他  $ta^{55}$   
 2sg ask 1sg neg-imp give 3sg  
 牛奶  $niw^{35}naj$  吃  $t\zeta^{55}$   
 milk eat
- b. Shanghai 儂 叫 我 勿要 撥 牛奶 伊 吃  
 2sg ask 1sg neg-imp give milk 3sg eat  
 ‘You asked me not to give him milk to drink.’

where the negative is the neg-imp and not one that can occur in non-imperative sentences.

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  $ta^{55}$  千萬  $t\zeta'j\epsilon n^{55}wan$  要  $jaw^{51}$  來  $laj^{35}$  or \*佢  $k'oy^{24}$  千祈  $t'f'in^{55}k'ej^{11}$  要  $jiw^{44}$  來  $l\epsilon j^{11}$  or \*我  $ni^{213>11}$  別  $pj\epsilon^{35}$  給  $kej^{213>11}$  他  $ta^{55}$  牛奶  $niw^{35}naj$  吃  $t\zeta^{55}$  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:

- (109) a. SM 自從  $tsr^{51}ts'o\eta$  認識  $z\epsilon n^{51}sl$  你  $ni^{35}$  以來  $ji^{11}laj$  李四  $li^{11}sr^{51}$   
 from know 2sg since Li Si  
 沒  $mej^{35}$  喝  $x\gamma^{55}$  過  $kw\zeta$  酒  $t\zeta jow^{213}$   
 neg-per-asp drink exp-asp liquor
- b. SM 李四  $li^{11}sr^{51}$  自從  $tsr^{51}ts'o\eta$  認識  $z\epsilon n^{51}sl$  你  $ni^{35}$  以來  $ji^{11}laj$   
 Li Si from know 2sg since  
 沒  $mej^{35}$  喝  $x\gamma^{55}$  過  $kw\zeta$  酒  $t\zeta jow^{213}$   
 neg-per-asp drink exp-asp liquor  
 ‘Since knowing you, Li Si has not drunk any liquor.’
- (110) a. Cantonese 因為  $j\epsilon n^{55}w\epsilon j^{22}$  個  $k\zeta^{44}$  仔  $z\epsilon^{35}$  病  $p\epsilon\eta^{22}$  所以  $s\zeta^{35}ji^{24}$   
 because CL son sick therefore  
 李四  $lej^{24}sej^{44}$  [mow<sup>24</sup>] 來  $l\epsilon j^{11}$   
 Li Si neg-PAST come
- b. Cantonese 李四  $lej^{24}sej^{44}$  因為  $j\epsilon n^{55}w\epsilon j^{22}$  個  $k\zeta^{44}$  仔  $z\epsilon^{35}$  病  $p\epsilon\eta^{22}$   
 Li Si because CL son sick  
 所以  $s\zeta^{35}ji^{24}$  [mow<sup>24</sup>] 來  $l\epsilon j^{11}$   
 therefore neg-PAST come  
 ‘Li Si did not come because his/her son was sick.’
- (111) a. SM 要不是  $jaw^{51}pu\zeta l$  他  $t'a^{55}$  來  $laj^{35}$  我  $w\zeta^{213>11}$  就  $t\zeta jow^{51}$   
 if-not 3sg come 1sg then  
 同意  $t'o\eta^{35}ji$  了  $l\epsilon$   
 agree FP
- b. SM 我  $w\zeta^{213>11}$  要不是  $jaw^{51}pu\zeta l$  他  $t'a^{55}$  來  $laj^{35}$  就  $t\zeta jow^{51}$   
 1sg if-not 3sg come then  
 同意  $t'o\eta^{35}ji$  了  $l\epsilon$   
 agree FP  
 ‘If he had not come, I would have agreed.’



- (112) a. Cantonese 雖然 *søy<sup>55</sup>jin<sup>11</sup>* 我 *ŋɔ<sup>24</sup>* 畀 *pej<sup>35</sup>* [*tʃɔ<sup>35</sup>*] 錢 *tʃin<sup>22</sup>*  
 although 1sg give pfv-asp money  
 佢 *k'øɣ<sup>24</sup>* 都 *tou<sup>55</sup>* [*tʃuŋ<sup>22</sup>*] [*nɐw<sup>55</sup>*]  
 3sg also still angry
- b. Cantonese 佢 *k'øɣ<sup>24</sup>* 雖然 *søy<sup>55</sup>jin<sup>11</sup>* 我 *ŋɔ<sup>24</sup>* 畀 *pej<sup>35</sup>* [*tʃɔ<sup>35</sup>*]  
 3sg although 1sg give pfv-asp  
 錢 *tʃin<sup>35</sup>* 都 *tou<sup>55</sup>* [*tʃuŋ<sup>22</sup>*] [*nɐw<sup>55</sup>*]  
 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) and (112a), the conjunction of the Dependent Clause must precede its subject NP. When both clauses share the same subject NP, as in (109a) and (110a), the subject NP appears only in the Main Clause but elided in the Dependent Clause. If the subject NP should appear in the Dependent Clause, as in (109b) and (110b), it must precede the conjunction, just as in (111b) and (112b) 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 with disjoint reference, for example, if a 3sg PN is added in the Embedded Clause of (109a) or (110a), ‘他 *t'a<sup>55</sup>*,’ and ‘佢 *k'øɣ<sup>24</sup>*,’ will be understood as a person different from Li Si:

- (109) c. SM 自從 *tsɿ<sup>51</sup>ts'uŋ* 他 *t'a<sup>55</sup>* 認識 *zən<sup>51</sup>sɿ* 你 *ni<sup>35</sup>* 以來 *ji<sup>11</sup>laj*  
 from 3sg know 2sg since  
 李四 *li<sup>11</sup>sɿ<sup>51</sup>* 沒 *mej<sup>35</sup>* ...  
 Li Si neg-PAST ...
- (110) c. Cantonese 因為 *jən<sup>55</sup>wɛj<sup>22</sup>* 佢 *k'øɣ<sup>24</sup>* 個 *kɔ<sup>44</sup>* 仔 *ɿ<sup>35</sup>* 病 *pɛŋ<sup>22</sup>*  
 because 3sg CL son sick  
 所以 *sɔ<sup>35</sup>ji<sup>24</sup>* 李四 *lej<sup>24</sup>sej<sup>44</sup>* [*mow<sup>24</sup>*] ...  
 therefore Li Si neg-PAST

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 individual conjunctions. For example, Time (Embedded) clauses usually carry the conjunction, unless they designate ‘after’ in which case if the pfv-asp appears, no conjunction is needed but the Matrix Clause will carry some adverb such as ‘then’. For example:

- (113) a. SM 你 *ni<sup>213>11</sup>* 吃 *ts'ɿ<sup>55</sup>* 了 *lɔ* 飯 *fan<sup>51</sup>* 再 *tsaj<sup>51</sup>* 去 *tɕ'y<sup>51</sup>*  
 b. Cantonese 你 *nej<sup>24</sup>* 食 *sɿk<sup>2</sup>* [*tʃɔ<sup>35</sup>*] 飯 *fan<sup>22</sup>* [*tʃi<sup>44</sup>*] 去 *høɣ<sup>44</sup>*  
 c. Suzhou (Wu) [nɛ] 吃 *tɕ'iɔ?* [*tsɿ*] 飯 *vɛ* 再 *tɕɛ* 去 *tɕ'ij*  
 2SG eat PFV-ASP rice then go  
 ‘You go after eating.’

With Conditional clauses, sometimes no conjunction appears in either clause at all:

- (114) SM 你 *ni<sup>213>11</sup>* 不 *pu<sup>51</sup>* 來 *laj<sup>35</sup>* 我 *wɔ<sup>213>11</sup>* 不 *pu<sup>51>35</sup>* 去 *tɕ'y<sup>51</sup>*  
 2sg neg come 1sg neg go  
 ‘If you don’t come, I won’t go.’

#### 4.14 Serial verbal constructions

Since Chinese need not use markers to distinguish Dependent from Main or Embedded from Matrix clause, VPs can occur in a series. For example:

- (115) a. SM 他  $t'a^{55}$  穿上  $t\zeta'wan^{55}\zeta a\eta$  衣服  $ji^{55}fu$  跑  $p'aw^{213>11}$  出  $t\zeta'u^{55}$   
 3SG put-on clothes run exit  
 大門  $ta^{51}m\partial n$  來  $laj$   
 front-door towards  
 'He put on his clothes and ran out of the front door (towards 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:

- b. SM 他  $t'a^{55}$  跑  $p'aw^{213>11}$  出  $t\zeta'u^{55}$  大門  $ta^{51}m\partial n$  來  $laj$   
 3sg run exit front-door towards  
 穿上  $t\zeta'wan^{55}\zeta a\eta$  衣服  $ji^{55}fu$   
 put-on clothes  
 'He ran out of the front door (towards the speaker) and put on his clothes.'

The so-called Pivotal construction is a kind of serial verbal construction in which a postverbal NP serves as a pivot – object of the preceding verb and subject of the following verb. For example:

- (116) Cantonese 你  $nej^{24}$  派  $p'aj^{44}$  佢  $k'oy^{24}$  去  $h\phi y^{44}$  [pin<sup>55</sup>] 道  $tow^{22}$  啊  $a^{44}$   
 2SG send 3SG go which way QP  
 'Where do you send him?'

where '佢  $k'oy^{24}$ ' is at once the object of 'to send' and subject of 'to go'.

#### 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:

- (117) a. SM 一面  $ji^{35}mi\epsilon n$  看  $k'an^{51}$  報  $paw^{51}$  一面  $ji^{35}mi\epsilon n$  吃  $t\zeta l^{65}$  飯  $fan^{51}$   
 b. Shanghai 一頭 看 報 一頭 吃 飯  
 simultaneously read newspaper simultaneously eat rice  
 'Read newspaper and eat at the same time.'
- (118) a. SM 一面  $ji^{55}mi\epsilon n^5$  吃  $t\zeta l^{55}$  飯  $fan^{51}$  一面  $ji^{55}mi\epsilon n^{51}$  看  $k'an^{51}$  報  $paw^{51}$   
 b. 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:

- (119) a. SM 要嘛  $jaw^{51}ma$  吃  $t\zeta l^{55}$  飯  $fan^{51}$  要嘛  $jaw^{51}ma$  吃  $t\zeta l^{55}$   
 麵  $mj\epsilon n^{51}$   
 b. Cantonese 一係  $j\epsilon t^5 h\epsilon j^{22}$  食  $sik^2$  飯  $fan^{22}$  一係  $j\epsilon t^5 h\epsilon j^{22}$  食  $sik^2$   
 either eat rice or eat  
 麵  $min^{22}$   
 noodle  
 'Either eat rice or noodles.'



- Yuan, Jiahua (1983) *Hanyu Fangyan Gaiyao (Outline of Chinese Dialects)*, Beijing: Wenzhi Gaige Press.
- Yue-(Hashimoto), A. (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)*, Beijing University, unpublished PhD dissertation.
- Zhang, Shuangqing (ed.) (1996) *Dongci de Ti (Verbal Aspects)*, Hong Kong: Chinese University.

# 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. 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,<sup>1</sup> 學 *ɕye* A2, 色 *sɿ* C1, 更 *kəŋ* A1 in Putonghua are *pau* A2, *ɕiau* A2, *ʂai* C1, *tɕiŋ* 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 modifier prior to the head, and 7 SVO word order. The results are given in Table 7.1.

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 7.2), we can see that Mandarin is the dialect with the least number of syllables, which signifies that the sound system of Mandarin is the simplest and so the pace of development is the most rapid.

**TABLE 7.1** COMPARISONS OF THE ATTRIBUTES IN CHINESE

	1	2	3	4	5	6	7
Proto-Chinese	+	-	-	±	-	-	-
Modern Chinese	+	+	+	+	+	+	+

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.

TABLE 7.2 COMPARISONS OF THE ATTRIBUTES OF MODERN CHINESE DIALECTS

<i>Dialect</i>	<i>Representative area</i>	<i>Number of initials</i>	<i>Number of finals</i>	<i>Number of tones</i>	<i>Possible number of syllables</i>
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

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. 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 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 variation 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 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*, *kɛr*, *kio* and *kɛ*, depending on the derivation or inflection. The form *kai* is the unmarked reading (or reference reading); *kɛr* 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; *kɛ* 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) 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 of the 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 thirty-five tone contour and the second changes to a fifty-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 7.3.

**TABLE 7.3** 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)	ʔ54		ʔ45 + 53	ʔ45 + 535	ʔ54 + 53 ʔ4 + 44

3 Infixable 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 *kaŋ* (桿) ‘pole’, ‘stick’ produces the infixed word *kaʔl laŋ* (格懶); adding it to *paŋ* (拌) ‘to stir’ produces *paʔl laŋ* (潑浪). 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 7.4:

**TABLE 7.4** 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	-ʔ 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



TABLE 7.4 (CONTINUED)

	<i>Common dialect (Northern Mandarin)</i>	<i>Chengdu (Southwestern Mandarin)</i>	<i>Nanjing (Jiang-Huai Mandarin)</i>
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 back fricative with <i>u</i> vowel <i>xu</i> ?	Yes	No, <i>x</i> before <i>u</i> pronounced as <i>f</i>	Yes
Has opposition of əŋ· <i>ij</i> and ən· <i>in</i> ?	Yes	No, merged as ən· <i>in</i>	No, merged as əŋ· <i>ij</i>

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 subdivisions 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.

## REFERENCES

- Hou, Jingyi (1987) *Hanyu fangyan wengao ji* (Collection of papers on Chinese dialects), Tokyo: Tokyo University of Foreign Languages.
- Yuan, Jiahua (1980) *Hanyu fangyan gaiyao* (Overview of the Chinese dialects), Beijing: Wenzhi Gaige Chubanshe.
- Mei, Tsu-lin (1997) 'Hanyu qi ge leixing tezheng de laiyan (The origin of seven of the typological characteristics of Chinese)'. *Language and Linguistics of China*, Vol. 4, edited by Tseng Chiu-yu, Taipei: Academia Sinica, 81–104.
- He, Wei (1989) *Jinyu Yanjiu* (Studies on the Jin dialect), Tokyo: Tokyo University of Foreign Languages.
- Jiang, Shaoyu (1994) *Jindai Hanyu yanjiu gaikuang* (Overview of studies on Modern Chinese), Beijing: Beijing University Press.

# SHANGHAI

## 1 SHANGHAI PHONOLOGY *Eric Zee*

The Shanghai dialect, a member of the Wu dialect family, is spoken by approximately twelve 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 forties.

### 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 forties 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<sup>h</sup> t<sup>h</sup> k<sup>h</sup>]. The breathiness was described as being realized not before but 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 ʒ dʒ g<sup>w</sup>] are more often pronounced as voiceless than voiced, and, as for [f], it is always voiceless. The voiced impression of the historical syllable-initial consonants [b d g g<sup>w</sup> dʒ v z ʒ 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<sup>w</sup> tʃ f s ʃ 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 [f<sup>m</sup> f<sup>n</sup> f<sup>ɲ</sup> f<sup>ŋ</sup> f<sup>l</sup> f<sup>w</sup>] and those followed by a non-breathy vowel as [ʔ<sup>m</sup> ʔ<sup>n</sup> ʔ<sup>ɲ</sup> ʔ<sup>ŋ</sup> ʔ<sup>l</sup> ʔ<sup>w</sup>]. The transcriptions are misleading, as the sonorants are neither preceded by the voiced glottal frication, nor are they glottalized. The consonant system in Shanghai consists of twelve plosives [p p<sup>h</sup> b t t<sup>h</sup> d k k<sup>h</sup> g k<sup>w</sup> k<sup>wh</sup> g<sup>w</sup>], one glottal stop [ʔ], four nasals [m n ɲ ŋ], nine fricatives [f v s z ʃ ʒ h f<sup>h</sup> h<sup>w</sup>], five affricates [ts ts<sup>h</sup> tʃ tʃ<sup>h</sup> dʒ], two approximants [w j], and one lateral liquid [l] as shown in the following consonant chart.

Phonologically, [b d g g<sup>w</sup>] are treated as the allophones of /p t k k<sup>w</sup>/ and [v z ʒ dʒ f] as the allophones of /f s ʃ tʃ h/. The historical [dʒ] in Shanghai is now [z]. The consonant phonemes in Shanghai include /p p<sup>h</sup> b t t<sup>h</sup> d k k<sup>h</sup> g k<sup>w</sup> k<sup>wh</sup> ʔ m n ɲ ŋ f s ʃ h h<sup>w</sup> ts ts<sup>h</sup> tʃ tʃ<sup>h</sup> w j l/.

The voiced initial obstruents in Shanghai are associated with the tones having a low tonal onset, such as [1<sup>3</sup>] and [12<sup>?</sup>] (‘?’ indicating that the tone is associated with the syllable with a glottal ending), whereas the voiceless counterparts are associated with the tones having a non-low tonal onset, such as [5<sup>1</sup>], [3<sup>4</sup>], and [5<sup>?</sup>]. Such an association of consonant type with

TABLE 8.1 CONSONANTS

	<i>Bilabial</i>	<i>Labio-dental</i>	<i>Alveolar</i>	<i>Alveolo-palatal</i>	<i>Velar</i>	<i>Glottal</i>
Plosive	[p p <sup>h</sup> b]		[t t <sup>h</sup> d]		[k k <sup>h</sup> g] [k <sup>w</sup> k <sup>wh</sup> g <sup>w</sup> ]	[ʔ]
Nasal	[m]		[n]	[ɲ]	[ŋ]	
Fricative		[f v]	[s z]	[ç ʒ]		[h f i h <sup>w</sup> ]
Affricate			[ts ts <sup>h</sup> ]	[tʃ tʃ <sup>h</sup> dz]		
Approximant	[w]			[j]		
Lateral			[l]			

tone does not hold for the syllable-initial sonorants [m n ɲ ŋ w j l], as the sonorants co-occur with any tone. The historical four-way place contrast of the nasals in the syllable-initial position is maintained in Shanghai, for example, [mɛ<sup>13</sup>] ‘slow’, [nɛ<sup>13</sup>] ‘difficult’, [ɲjɛ<sup>13</sup>] ‘twenty’, and [ŋɛ<sup>13</sup>] ‘slow-witted’. Examples of the monosyllables containing the consonants in the syllable-initial position in Shanghai are listed as follows (with [ʔ] after the vowel [a] indicating that the velar nasal is weakened, but not dropped, and [v] representing a vowel or diphthong).

/p/	[pɪ <sup>51</sup> ] edge, [pɪ <sup>34</sup> ] flat, [pɪʔ <sup>57</sup> ] pen [bɪ <sup>13</sup> ] skin, [bɪʔ <sup>127</sup> ] other
/p <sup>h</sup> /	[p <sup>h</sup> ɪ <sup>51</sup> ] to criticize, [p <sup>h</sup> ɪ <sup>34</sup> ] to lie, [p <sup>h</sup> ɪʔ <sup>57</sup> ] to cleave
/t/	[tɪ <sup>51</sup> ] low, [tɪ <sup>34</sup> ] shop, [tɪʔ <sup>57</sup> ] to fall [dɪ <sup>13</sup> ] sweet, [dɪʔ <sup>127</sup> ] enemy
/t <sup>h</sup> /	[t <sup>h</sup> ɪ <sup>51</sup> ] sky, [t <sup>h</sup> ɪ <sup>34</sup> ] to substitute, [t <sup>h</sup> ɪʔ <sup>57</sup> ] to kick
/k/	[kɛ <sup>51</sup> ] crafty, [kɛ <sup>34</sup> ] to select, [kɔʔ <sup>57</sup> ] pigeon [gɛ <sup>13</sup> ] to lean against, [gɔʔ <sup>127</sup> ] crowded
/k <sup>h</sup> /	[k <sup>h</sup> ɔ <sup>51</sup> ] to knock, [k <sup>h</sup> ɔ <sup>34</sup> ] to rely on, [k <sup>h</sup> ɔʔ <sup>57</sup> ] to cough
/k <sup>w</sup> /	[k <sup>w</sup> ɛ <sup>51</sup> ] to close, [k <sup>w</sup> ɛ <sup>34</sup> ] accustomed to, [k <sup>w</sup> ɔʔ <sup>57</sup> ] bone [g <sup>w</sup> ɛ <sup>13</sup> ] to throw, [g <sup>w</sup> vʔ <sup>127</sup> ] (non-occurring)
/k <sup>wh</sup> /	[k <sup>wh</sup> a <sup>51</sup> ] to boast, [k <sup>wh</sup> a <sup>34</sup> ] fast, [k <sup>wh</sup> ɔʔ <sup>57</sup> ] wide
/ts/	[tsɿ <sup>51</sup> ] pig, [tsɿ <sup>34</sup> ] paper, [tsaʔ <sup>57</sup> ] to tie
/ts <sup>h</sup> /	[ts <sup>h</sup> ɿ <sup>51</sup> ] to take out from between, [ts <sup>h</sup> ɿ <sup>34</sup> ] stench, [ts <sup>h</sup> aʔ <sup>57</sup> ] ruler
/tʃ/	[tʃɿ <sup>51</sup> ] pointed, [tʃɿ <sup>34</sup> ] to mail, [tʃɿʔ <sup>57</sup> ] to catch/receive [dʒɿ <sup>13</sup> ] to ride (a horse), [dʒɿʔ <sup>127</sup> ] extreme
/tʃ <sup>h</sup> /	[tʃ <sup>h</sup> ɿ <sup>51</sup> ] to bully, [tʃ <sup>h</sup> ɿ <sup>34</sup> ] to go, [tʃ <sup>h</sup> ɿʔ <sup>57</sup> ] seven
/f/	[fɿ <sup>51</sup> ] to fly, [fɿ <sup>34</sup> ] lung, [fɔʔ <sup>57</sup> ] fortune [vɿ <sup>13</sup> ] fat, [vɔʔ <sup>127</sup> ] to obey
/s/	[sɔ <sup>51</sup> ] to cook, [sɔ <sup>34</sup> ] scant, [sɔʔ <sup>57</sup> ] uncle [zɔ <sup>13</sup> ] to build, [zɔʔ <sup>127</sup> ] cooked
/ç/	[çja <sup>51</sup> ] fragrant, [çja <sup>34</sup> ] to think, [çɪʔ <sup>57</sup> ] snow [ʒja <sup>13</sup> ] to resemble, [ʒɪʔ <sup>127</sup> ] mat
/h/	[ha <sup>51</sup> ] laugh (onomatopoeic), [ha <sup>34</sup> ] crab, [haʔ <sup>57</sup> ] blind [fi <sup>13</sup> ] shoe, [fiʔ <sup>127</sup> ] box
/h <sup>w</sup> /	[h <sup>w</sup> ɛj <sup>51</sup> ] dust, [h <sup>w</sup> ɛj <sup>34</sup> ] to repent, [h <sup>w</sup> aʔ <sup>57</sup> ] to crack
/m/	[mej <sup>51</sup> ] every, [mej <sup>34</sup> ] good-looking, [mVʔ <sup>57</sup> ] (non-occurring) [mej <sup>13</sup> ] plum, [mɔʔ <sup>127</sup> ] ink
/n/	[nɛ <sup>51</sup> ] to hand over, [nV(ŋ) <sup>34</sup> ]/[nVʔ <sup>57</sup> ] (non-occurring) [nɛ <sup>13</sup> ] difficult, [nɔʔ <sup>127</sup> ] promise

/ɲ/	[ɲjɿ <sup>51</sup> ] to pinch, [ɲjV(ŋ) <sup>34</sup> ]/[ɲjVɿ <sup>52</sup> ] (non-occurring) [ɲjɿ <sup>13</sup> ] cow, [ɲjɔɿ <sup>122</sup> ] meat
/ŋ/	[ŋu <sup>51</sup> ] I (literary), [ŋV(ŋ) <sup>34</sup> ]/[ŋVɿ <sup>52</sup> ] (non-occurring) [ŋu <sup>13</sup> ] I, [ŋɔɿ <sup>122</sup> ] crane
/l/	[ljɿ <sup>51</sup> ] to sneak out, [lV(ŋ) <sup>34</sup> ]/[lVɿ <sup>52</sup> ] (non-occurring) [ljɿ <sup>13</sup> ] flowing, [lɿ <sup>122</sup> ] to stand
/w/	[wəŋ <sup>51</sup> ] lukewarm, [wəŋ <sup>34</sup> ] steady, [waɿ <sup>52</sup> ] to scoop [wəŋ <sup>13</sup> ] dizzy, [waɿ <sup>122</sup> ] slippery
/j/	[ja <sup>51</sup> ] to hide, [ja <sup>34</sup> ] elegant, [jɿ <sup>52</sup> ] one [ja <sup>13</sup> ] father (term of reference), [jɿ <sup>122</sup> ] leaf
/ʔ/	[ʔəŋ <sup>51</sup> ] favour, [ʔV(ŋ) <sup>34</sup> ] (non-occurring), [ʔaɿ <sup>52</sup> ] duck, [ʔV(ŋ) <sup>13</sup> ]/[ʔVɿ <sup>122</sup> ] (non-occurring)

## 1.2 Vowels

The vowel system in Shanghai consists of [i ɿ y ɛ ə ø a ɔ ɤ o u ʊ] which are listed in the vowel chart below. Phonologically, /i y ɛ ə ø a ɔ ɤ o u/ are the vowel phonemes and [ɿ ɿ ə ʊ] are allophones of /i y ɛ u/, respectively. The derivation of [ɿ ɿ ə ʊ] requires the following phonological rules, /i/ → [ɿ]/\_ [ŋ ʔ]; /y/ → [ɿ]/\_ [ŋ ʔ]; /ɛ/ → [ə]/\_ [ŋ ʔ]; 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 closed-mid [o] and closed [u]. In the speech of speakers in their early forties and younger, the rhymes [-a<sup>0</sup>] and [-ɔ<sup>0</sup>] have merged

Close	i y		u
	ɿ ɿ	ʊ	
		o ɤ	
Mid	ɛ ø ə ɔ		
Open	a		

and become [-ã<sup>0</sup>], with the vowel being heavily nasalized. Examples of the monosyllabic words containing the vowels are as follows.

/i/	[i]	[ɕi <sup>51</sup> ] ahead, [ti <sup>34</sup> ] shop, [mi <sup>13</sup> ] rice [ɿ]	[pɿ <sup>52</sup> ] pen, [dɿ <sup>122</sup> ] enemy [ɕiŋ <sup>51</sup> ] new, [tiŋ <sup>34</sup> ] peak, [biŋ <sup>13</sup> ] bottle
/y/	[y]	[ɕy <sup>51</sup> ] weak, [tɕy <sup>34</sup> ] expensive, [ɲy <sup>13</sup> ] female [ɿ]	[tɕyɿ <sup>52</sup> ] military, [dʒyɿ <sup>122</sup> ] to dig [tɕyŋ <sup>51</sup> ] military, [ɕyŋ <sup>34</sup> ] training, [jyŋ <sup>13</sup> ] cloud
/ɛ/	[ɛ]	[kɛ <sup>51</sup> ] sly, [ts <sup>h</sup> ɛ <sup>34</sup> ] vegetable, [mɛ <sup>13</sup> ] slow [ə]	[k <sup>w</sup> əɿ <sup>52</sup> ] bone, [həɿ <sup>122</sup> ] box [kəŋ <sup>51</sup> ] root, [təŋ <sup>34</sup> ] to wait, [bəŋ <sup>13</sup> ] stupid
/ø/	[ø]	[sø <sup>51</sup> ] sour, [k <sup>h</sup> ø <sup>34</sup> ] to see, [mø <sup>13</sup> ] full	
/a/	[a]	[k <sup>h</sup> a <sup>51</sup> ] to wipe, [ts <sup>h</sup> a <sup>34</sup> ] to tear, [da <sup>13</sup> ] to wash [ʔaɿ <sup>52</sup> ] duck, [baɿ <sup>122</sup> ] white [sa <sup>51</sup> ] uncooked, [sa <sup>34</sup> ] thrifty, [za <sup>013</sup> ] long	
/ɔ/	[ɔ]	[sɔ <sup>51</sup> ] to cook, [hɔ <sup>34</sup> ] good, [lɔ <sup>13</sup> ] old [k <sup>h</sup> ɔɿ <sup>52</sup> ] to cry, [dɔɿ <sup>122</sup> ] poisonous [fɔ <sup>51</sup> ] square, [p <sup>h</sup> ɔ <sup>34</sup> ] weighty, [gɔ <sup>013</sup> ] foolish	
/o/	[o]	[po <sup>51</sup> ] scar, [p <sup>h</sup> o <sup>34</sup> ] fear, [bo <sup>13</sup> ] to climb	

/ɣ/	[ɣ]	[sɣ <sup>51</sup> ] to collect, [k <sup>h</sup> ɣ <sup>34</sup> ] mouth, [lɣ <sup>13</sup> ] leaking
/u/	[u]	[tu <sup>51</sup> ] abundant, [pu <sup>34</sup> ] cloth, [gu <sup>13</sup> ] to squat
	[ʊ]	[sʊŋ <sup>51</sup> ] loose, [t <sup>h</sup> ʊŋ <sup>34</sup> ] sore, [nʊŋ <sup>13</sup> ] you

### 1.3 Syllabic consonants

The syllabic consonants in Shanghai are [ŋ ŋ'] and [ɹ]. While there are only a few monosyllables which contain a syllabic nasal as the vowel nucleus, for example, [ŋ] in the bisyllabic compound [ŋ ma<sup>5-3</sup>] 'mother' and [ŋ<sup>13</sup>] 'fish', a sizable number of monosyllables contain a syllabic alveolar approximant [ɹ], for example, [ts<sup>h</sup>ɹ<sup>51</sup>] 'to blow', [tsɹ<sup>34</sup>] 'paper', and [zɹ<sup>13</sup>] 'self'. [ɹ] which occurs only after the apico-alveolar sibilants [ts ts<sup>h</sup> s z] is treated as an allophone of the vowel phoneme /i/. The approximant [ɹ] is called an 'apical vowel' in Xu and Tang (1988).

### 1.4 Diphthongs

Four diphthongs /ja jɔ jɣ ej/ are identified in Shanghai. The first element of the phonetic representations of /ja jɔ jɣ/ 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/	[tja <sup>51</sup> ] father (term of address), [tɕja <sup>34</sup> ] to borrow, [dʒja <sup>13</sup> ] talented [tɕ <sup>h</sup> ja <sup>51</sup> ] gun, [ɕja <sup>34</sup> ] to think, [lja <sup>13</sup> ] bright [ɕja <sup>52</sup> ] to whittle, [ʒja <sup>122</sup> ] to chew
/jɔ/	[p <sup>h</sup> jɔ <sup>51</sup> ] to drift, [pjɔ <sup>34</sup> ] wrist watch, [djɔ <sup>13</sup> ] to exchange [tɕjɔ <sup>52</sup> ] orange, [ɲjɔ <sup>122</sup> ] meat
/jɣ/	[ljɣ <sup>51</sup> ] to sneak out, [tɕjɣ <sup>34</sup> ] wine, [ɲjɣ <sup>13</sup> ] cow
/ej/	[h <sup>w</sup> ej <sup>51</sup> ] dust, [tej <sup>34</sup> ] correct, [wej <sup>13</sup> ] stomach

Other diphthongs, such as [jɛ] and [jʊ] 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 ʒ tç tç<sup>h</sup> dz] or [ɲ]. A general on-glide-insertion rule,  $\emptyset \rightarrow [j]/[+ \text{coronal} - \text{anterior}] \_\_\_ \text{v}$ , where  $\text{v} \neq [i \text{ } \text{ɪ} \text{ } y \text{ } \text{ɻ}]$ , is formulated for inserting [j] before any vowel, except [i], [ɪ], [y], or [ɻ], when the vowel is preceded by an alveolo-palatal sound. The 'true' diphthongs, such as [ja jɔ jɣ 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 the consonant chart shown earlier, 'V' a vowel, a diphthong, a syllabic alveolar approximant [ɹ], a syllabic bilabial nasal [ŋ], or a syllabic velar nasal [ŋ'], 'S' a glottal stop [ʔ], and 'N' a velar nasal [ŋ]. The syllabic alveolar approximant [ɹ] occurs only after an apico-alveolar fricative [s z] or affricate [ts ts<sup>h</sup>], and it is not followed by a nasal or glottal ending. The syllabic bilabial nasal [ŋ] or velar nasal [ŋ] occurs singly, not preceded nor followed by a consonant. The possible combinations of the syllable-initial consonant with the rhyme in Shanghai are listed in the following chart. The consonants which occur before a particular rhyme are listed under the rhyme in the same cell.

TABLE 8.2 SYLLABLE STRUCTURES

v	v + [ŋ]	v + [ʔ]	D	D + [ŋ]	D + [ʔ]
[-i] [p p <sup>h</sup> b m f v t t <sup>h</sup> d l ç z tç tç <sup>h</sup> dz j]	[-iŋ] [p p <sup>h</sup> b m t t <sup>h</sup> d l ç z tç tç <sup>h</sup> dz j]	[-iʔ] [p p <sup>h</sup> b m t t <sup>h</sup> d l ç z tç tç <sup>h</sup> dz j]			
[-ɿ] [ts ts <sup>h</sup> s z]					
[-y] [l ç z tç tç <sup>h</sup> dz ɲ j]	[-yŋ] [ç z tç tç <sup>h</sup> dz j]	[-yʔ] [ç tç tç <sup>h</sup> dz j]			
			[-ej] [p p <sup>h</sup> b w t t <sup>h</sup> d l]		
[-ɛ] [p p <sup>h</sup> b m f v t t <sup>h</sup> d l s z ts ts <sup>h</sup> k k <sup>h</sup> g k <sup>w</sup> k <sup>wh</sup> ç <sup>w</sup> ŋ h h <sup>w</sup> fi j w]	[-əŋ] [p p <sup>h</sup> b m f v t t <sup>h</sup> d l s z ts ts <sup>h</sup> k k <sup>h</sup> g k <sup>w</sup> k <sup>wh</sup> h h <sup>w</sup> fi w]	[-əʔ] [p p <sup>h</sup> b m f v t t <sup>h</sup> d l s z ts ts <sup>h</sup> k k <sup>h</sup> g k <sup>w</sup> k <sup>wh</sup> ŋ h h <sup>w</sup> fi j w ʔ]	[-jɛ] [ɲ]		
[-∅] [p p <sup>h</sup> b m t t <sup>h</sup> d l s z ts ts <sup>h</sup> k k <sup>h</sup> k <sup>w</sup> k <sup>wh</sup> ŋ h h <sup>w</sup> fi w]					
[-a] [p p <sup>h</sup> b m t t <sup>h</sup> d n l s z ts ts <sup>h</sup> k k <sup>h</sup> g k <sup>w</sup> k <sup>wh</sup> ŋ h h <sup>w</sup> fi j w]	[-a <sup>ŋ</sup> ] [p p <sup>h</sup> b m t l s z ts ts <sup>h</sup> k k <sup>h</sup> ŋ fi w]	[-aʔ] [p p <sup>h</sup> b m f v t t <sup>h</sup> d l s z ts ts <sup>h</sup> k k <sup>h</sup> g k <sup>w</sup> k <sup>wh</sup> ŋ h h <sup>w</sup> fi j w ʔ]	[-ja] [t ç z tç tç <sup>h</sup> dz]	[-ja <sup>ŋ</sup> ] [l ç z tç tç <sup>h</sup> dz ɲ]	[-jaʔ] [l ç z tç tç <sup>h</sup> ɲ]
[-ɔ] [p p <sup>h</sup> b m t t <sup>h</sup> d n l s z ts ts <sup>h</sup> k k <sup>h</sup> g ŋ h fi j]	[-ɔ <sup>ŋ</sup> ] [p p <sup>h</sup> b m t t <sup>h</sup> d l s z ts ts <sup>h</sup> k k <sup>h</sup> g ŋ h h <sup>w</sup> fi j w]	[-ɔʔ] [p p <sup>h</sup> b m t t <sup>h</sup> d n l s z ts ts <sup>h</sup> k k <sup>h</sup> ŋ h fi ʔ]	[-jɔ] [p p <sup>h</sup> b m v t t <sup>h</sup> d l ç z tç tç <sup>h</sup> dz ɲ]		[-jɔʔ] [ç z tç tç <sup>h</sup> dz ɲ]
[-o] [p p <sup>h</sup> b m n s z ts ts <sup>h</sup> k k <sup>h</sup> ŋ h w]					
[-ɻ] [p <sup>h</sup> m f v t t <sup>h</sup> d l s z ts ts <sup>h</sup> k k <sup>h</sup> g ŋ h fi j]			[-jɻ] [l ç z tç tç <sup>h</sup> dz ɲ]		
[-u] [p p <sup>h</sup> b m f v t t <sup>h</sup> d n l s z ts ts <sup>h</sup> k k <sup>h</sup> g ŋ h fi]	[-uŋ] [p p <sup>h</sup> b m f v t t <sup>h</sup> d n l s z ts ts <sup>h</sup> k k <sup>h</sup> g h fi j]			[-juŋ] [ç tç tç <sup>h</sup> dz ɲ]	

Notes: v = vowel or syllabic approximant; D = diphthong.

## 1.6 Tones

Five of the eight historical tones (see chart) are preserved in Shanghai. 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. The numerical values for the five citations tones in Shanghai are [51], [34], [13], [5ʔ], and [12ʔ] as shown in the following chart. The superscript [ʔ] indicates that the tone is associated with the short syllables which end with a glottal stop.

**TABLE 8.3** TONE CATEGORIES

<i>Historical tone category</i>		<i>Citation form on monosyllables</i>
<i>yin</i>	<i>ping</i>	[ <sup>5</sup> 1]
<i>yin</i>	<i>shang</i>	[ <sup>3</sup> 4]
<i>yin</i>	<i>qu</i>	
	<i>ping</i>	[ <sup>1</sup> 3]
<i>yang</i>	<i>shang</i>	
	<i>qu</i>	
<i>yin</i>	<i>ru</i>	[ <sup>5</sup> 7]
<i>yang</i>	<i>ru</i>	[ <sup>12</sup> 7]

**TABLE 8.4** LEXICAL TONE MELODIES

<i>Types of lexical tone melody</i>	<i>Citation form on mono-syllables</i>	<i>TM on bisyllabic compounds</i>	<i>TM on trisyllabic compounds</i>	<i>TM on quadrisyllabic compounds</i>	<i>TM on quintesyllabic compounds</i>
I /51/	[ <sup>5</sup> 1]	[ <sup>5</sup> -1]	[ <sup>5</sup> -3-1]	[ <sup>5</sup> -3↑-3↓-1]	[ <sup>5</sup> -4-3-2-1]
II /351/	[ <sup>3</sup> 4]	[ <sup>3</sup> -4]	[ <sup>3</sup> -5-1]	[ <sup>3</sup> -5-3-1]	[ <sup>3</sup> -5-3↑-3↓-1]
III /151/	[ <sup>1</sup> 3]	[ <sup>1</sup> -3]	[ <sup>1</sup> -5-1]	[ <sup>1</sup> -5-3-1]	[ <sup>1</sup> -5-3↑-3↓-1]
IV /451/	[ <sup>5</sup> 7]	[ <sup>4</sup> -5]	[ <sup>4</sup> -5-1]	[ <sup>4</sup> -5-3-1]	[ <sup>4</sup> -5-3↑-3↓-1]
V /12/	[ <sup>12</sup> 7]	[ <sup>1</sup> -12]	[ <sup>1</sup> -1-12]	[ <sup>1</sup> -1-1-12]	[ <sup>1</sup> -1-1-1-12]
		([ <sup>1</sup> -3])	([ <sup>1</sup> -5-1])	([ <sup>1</sup> -5-3-1])	([ <sup>1</sup> -5-3↑-3↓-1])

Note: TM = tone melody.

## 1.7 Tone melodies

There is a fixed number of lexical tone melodies on the monosyllables and compounds in Shanghai, as shown in the following chart. The type of the lexical tone melody realized on the bisyllabic and longer compounds is triggered by the lexical tone melody on the initial monosyllable in the compounds. The lexical tone melodies are elastic, in that their length is adjustable to the size of the compounds, as can be seen in the tone melody chart below. This is especially obvious with the Types I and V Lexical Tone Melodies. The Type I Lexical Tone Melody /<sup>5</sup>1/ is realized as [<sup>5</sup>-1], [<sup>5</sup>-3-1], [<sup>5</sup>-3↑-3↓-1], and [5-4-3-2-1] on the bisyllabic, trisyllabic, quadrisyllabic, and quintesyllabic compounds, respectively (the vertical arrows after a numeral indicating that the tone value is either raised or lowered). Similarly, the Type V Lexical Tone Melody /<sup>1</sup>2/ is realized as [<sup>1</sup>-12], [<sup>1</sup>-1-12], [<sup>1</sup>-1-1-12], and [<sup>1</sup>-1-1-1-12] on the bisyllabic and longer compounds. The Type V Lexical Tone Melody /<sup>1</sup>2/ may be optionally realized as [<sup>1</sup>-3] on some of the bisyllabic compounds and [<sup>1</sup>-5-1], [<sup>1</sup>-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 was analysed as an auto-segmental process of rightward spreading of the lexical tone on the initial syllable in the compounds, and the lexical tones on the following syllables were deleted. The lexical tones were posited as /HM/, /MH/, /LH/, /H/, and /LH/ which correspond to the Types I, II, III, IV, and V Lexical Tone Melodies, respectively, in this study. The tone, [<sup>1</sup>], on the last syllable in the compounds was considered as a default low tone. For comparison, see Zee and Maddieson (1980).

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 types of 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 Melody /<sup>51</sup>/:  
 [ɕja<sup>ŋ</sup> ji<sup>5-1</sup>] cigarette  
 [t<sup>h</sup>i vəŋ dɛ<sup>5-3-1</sup>] observatory  
 [sɛ ɕm lja<sup>ŋ</sup> ji<sup>5-3↑-3↓-1</sup>] indecisive  
 [kɛ li və(ʔ) kɛ ka<sup>5-4-3-2-1</sup>] awkward
- Type II Lexical Tone Melody /<sup>351</sup>/:  
 [hu ts<sup>h</sup>o<sup>3-4</sup>] train  
 [ta<sup>ŋ</sup> zɿ tɕi<sup>3-5-1</sup>] typewriter  
 [tjɔ y lɔ<sup>ŋ</sup> tɔ<sup>ŋ</sup> tɔ<sup>ŋ</sup> tɔ<sup>ŋ</sup> tɔ<sup>ŋ</sup>] hanging around and doing nothing  
 [ɕjɔ t<sup>h</sup>i d y sɿ vu<sup>3-5-3↑-3↓-1</sup>] little barber
- Type III Lexical Tone Melody /<sup>151</sup>/:  
 [djɔ ka<sup>ŋ</sup>1-3] spoon  
 [zəŋ tɕm bŋ<sup>1-5-1</sup>] nutty  
 [zja t<sup>h</sup>i zja di<sup>1-5-3-1</sup>] thank heaven and earth  
 [zja<sup>ŋ</sup> sa(ʔ) j y ka zɿ<sup>1-5-3↑-3↓-1</sup>] pretentious and pompous
- Type IV Lexical Tone Melody /<sup>451</sup>/:  
 [ʔo(ʔ) ts<sup>h</sup>o(ʔ) 4-5] dirty  
 [tɕja(ʔ) da(ʔ) ts<sup>h</sup>o<sup>4-5-1</sup>] bicycle  
 [ha(ʔ) sɛ wo sɿ<sup>4-5-3-1</sup>] talking nonsense  
 [tɔ(ʔ) dŋ mo sɿ k<sup>h</sup>ɛ<sup>4-5-3↑-3↓-1</sup>] confident
- Type V Lexical Tone Melody /<sup>12</sup>/ (optionally /<sup>151</sup>/ on some compounds):  
 [za(ʔ) ljy<sup>1-12</sup>] pomegranate  
 [lɔ(ʔ) jŋ tɕi<sup>1-1-12</sup>] tape recorder  
 [vɔ(ʔ) tɛ da fɿɔ<sup>1-1-1-12 (1-5-3-1)</sup>] Fudan University  
 [la(ʔ) li d y ŋi tsɿ<sup>1-1-1-1-12 (1-5-3↑-3↓-1)</sup>] 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, for example,

[ts<sup>h</sup>ɔ p<sup>h</sup>jɔ<sup>3-4</sup>] [ŋu<sup>13</sup>] [wɛ pə(ʔ) ji lə<sup>1-5-3-1</sup>]  
 money I return to him ASP  
 ‘I have returned the money to him.’

[ŋu<sup>13</sup>] [və(ʔ) tə(ʔ) ji kɔ<sup>ŋ</sup>1-5-3-1]  
 I (NEG) to him/her tell  
 ‘I will not tell him/her.’



[ŋu<sup>13</sup>] [tɕ<sup>34</sup>] [zɔ<sup>ŋ</sup> hɛ tɕ<sup>h</sup>i<sup>1-5-1</sup>]  
 I arrive Shanghai go  
 'I am going to Shanghai.'

[ji<sup>13</sup>] [tɕ<sup>h</sup> i(?) hɔ lə<sup>4-5-1</sup>]  
 He/she eat finished ASP  
 'He/she has finished eating.'

For more about the phrasal phonology in Shanghai, see Jin (1986) and Selkirk and Shen (1990).

### Concluding remarks on phonology

The foregoing presents a précis of the main aspects of the Shanghai phonology based on the speech of the language consultants in the late 1940s. 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, Beijing Mandarin. The educated young adults in their early twenties find difficulty in pronouncing a sizable number monosyllabic words in Shanghai. They would utter the monosyllables in Beijing Mandarin when pressed. A phonology of Shanghai based on the speech of the younger speakers is expected to differ slightly 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:

gɔʔ pəŋ [1-3] sɿ [51]  
 this CL book  
 'this book'

liə<sup>ŋ</sup> pej [1-3] sɿ [34]  
 two CL water  
 'two cups of water'

gɔʔ gɔʔ [1-3] liə<sup>ŋ</sup> pej [1-3] sɿ [34]  
 these CL two CL water  
 'these two cups of water'

As in Mandarin, the numeral *jiʔ* '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 in 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 the front position or immediately precede the head noun:

*ji* [<sup>13</sup>] *cia gəʔ* [<sup>3-4</sup>] *sɛ pəŋ* [<sup>5-1</sup>] *sɿ* [<sup>51</sup>]  
 he write MOD-MARKER three copy book  
 'the three books he wrote'

*sɛ pəŋ* [<sup>5-1</sup>] *ji* [<sup>13</sup>] *cia gəʔ* [<sup>3-4</sup>] *sɿ* [<sup>51</sup>]  
 three copy he write MOD-MARKER book  
 'the three books he wrote'

## 2.2 Verb phrases

The object follows the verb in the canonical form.

*dʒʔ tsɿ* [<sup>1-3</sup>] *pəŋ* [<sup>3</sup>] *sɿ* [<sup>51</sup>]  
 read ASP copy book  
 'have read a book'

The typical Shanghainese order for the double object construction is: V-Direct Object-Indirect Object, where the indirect object is a pronoun (Xu and Tang 1988: 479):

*soŋ pəŋ* [<sup>3-4</sup>] *sɿ noŋ* [<sup>5-1</sup>]  
 give copy book you  
 'give you a book'

It can be described as derived from the oblique construction by preposition deletion:

*soŋ pəŋ* [<sup>3-4</sup>] *sɿ* [<sup>5-1</sup>] *pəʔ noŋ* [<sup>3-4</sup>]  
 give copy book to you  
 'give you 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* [<sup>3-4</sup>] *pəŋ* [<sup>3</sup>] *sɿ* [<sup>51</sup>]  
 give you 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 a numerically quantified NP.

*dʒʔ tsɿ* [<sup>1-3</sup>] *liɑ<sup>ŋ</sup> gəʔ* [<sup>1-3</sup>] *tsɔŋdɿ* [<sup>5-1</sup>]  
 read ASP two CL hours  
 'have read for two hours'

*dʒʔ tsɿ* [<sup>1-3</sup>] *sɛ t<sup>h</sup> ʒ<sup>ŋ</sup>* [<sup>5-1</sup>]  
 read ASP three times  
 'have read three times'

Such postverbal adjuncts are analysed as complements in traditional grammar. In contrast, preverbal adjuncts denoting time, place, manner, etc., are assigned to the category of adverbial.

As in Mandarin, the OV order is also found. The object is often preceded by the preposition *nɛ*, equivalent to *ba* in the so-called *ba*-construction in Mandarin.

*nɛ* [51] *gəʔŋɛ* [1-3] *sɿ* [51] *gʷɛ tʰəʔ* [1-3]  
 PREP these 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.

*nɛ* [51] *gəʔŋɛ* [1-3] *sɿ* [51] *gʷɛ tʰəʔji* [1-5-1]  
 PREP these 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 *nɛ* is optional, while omission of its Mandarin counterpart *ba* often results in sentences less acceptable to Mandarin speakers.

The resultative construction consists of a verb plus a resultative complement plus a noun:

*ɕia tsʰo tsɿ* [3-5-1] *ditsɿ* [1-3]  
 write wrong ASP address  
 'Write the address wrongly.'

*tɕʰjaʔ wəʔ 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 predicative 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 till it is all gone. Thematically, the final noun phrase performs a dual role, 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.

*tɕʰjaʔji vəʔ wəʔ* [4-5-3-1]  
 eat it not finish  
 'can't eat it up'

*tɕʰjaʔ vəʔ wəʔ* [4-5-3-1]  
 eat not 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, a clause, etc.

*tsɔvɛ* [3-4] *ji* [1-3] *zɑ<sup>ŋ</sup>zɑ<sup>ŋ</sup>* [1-3] *vəʔtɕ<sup>h</sup>jaʔ* [1-3]  
 breakfast he often not eat  
 'Breakfast, he often doesn't eat.'

*nətɕiŋ lu* [1-5-3-1] *ŋu* [13] *ɕiɔtəʔji* [3-5-1] *zɿ tsɿ* [1-3] *zɑ<sup>ŋ</sup>jy* [1-3]  
 Nanjing street on I know he live ASP long  
 'On that street I know he has lived for a long time.'

*sɔ* [5] *vɛ* [13] *ji* [13] *tsəʔ* *wej* [3-4] *ts<sup>h</sup>ɔ* [3] *dɛ* [13]  
 cook he 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.

*tsɔvəʔ* [3-4] *ji* [13] *zɑ<sup>ŋ</sup>zɑ<sup>ŋ</sup>* [1-3] *tɕ<sup>h</sup>jaʔ* [5] *mipɔ* [1-3]  
 breakfast he often eat bread  
 'For breakfast, he often eats bread.'

*nətɕiŋ lu* [1-5-1] *ŋu* [13] *ɕiɔtəʔji* [3-5-1] *la* [12] *imi* [5-1]  
 Nanjing street I know he in there  
*zɿ tsɿ* [1-3] *zɑ<sup>ŋ</sup>jy* [1-3]  
 live ASP long  
 'On that street I know he has lived for a long time.'

There is an anaphoric relation between *tsɔvəʔ* and *mipɔ*, and between *nətɕiŋ lu* and *imi*. 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.

*biŋku* [1-3] *zəʔk<sup>wh</sup>ej* [1-12] *jiʔkoŋtɕiŋ* [4-5-1]  
 apples ten dollars one kilo  
 'As for apples, it costs 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 the sentence final position. Shanghainese has two particles, *mə* and *zɿ*, used exclusively to mark topics.

*dinɔməʔ* [1-5-1] *ŋu zɿ* [1-3] *ŋa fiɔ<sup>ŋ</sup>* [1-3]  
 computer TOP I am layman  
 'As for computers, I am a layman.'

*gəʔgəʔ* [1-3] *ŋiŋ* [13] *zɿ* [1] *məməʔ* [1-3] *ŋiŋ* [13] *kø* [3] *za* [13]  
 this CL person TOP no one 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* [<sup>13</sup>] *doŋdi* [<sup>1-3</sup>] *m̄məʔ* [<sup>1-3</sup>] *fu ku* [<sup>3-4</sup>]  
 he money not pay ASP-  
 'Money, he has not paid.'

*noŋ* [<sup>13</sup>] *zo* [<sup>13</sup>] *jɔ* [<sup>3</sup>] *tɕ<sup>hi</sup>ja?* *va* [<sup>4-5</sup>]  
 you tea would-like 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*:

*noŋ* [<sup>13</sup>] *ma tsɿ* [<sup>1-3</sup>] *saməʔzɿ* [<sup>3-5-1</sup>]  
 you bought ASP what?  
 'What did you buy?'

*ji* [<sup>13</sup>] *sazəŋk<sup>w</sup>ɔ<sup>n</sup>* [<sup>3-5-1</sup>] *le ku* [<sup>1-3</sup>]  
 he 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:

*noŋ* [<sup>13</sup>] *tɕ<sup>hi</sup>i vəʔ* [<sup>3-4</sup>]  
 you go Q?  
 'Did you go?'

*noŋ* [<sup>13</sup>] *tɕ<sup>hi</sup>i vəʔ tɕ<sup>hi</sup>i* [<sup>3-5-1</sup>]  
 you go not go?  
 'Did you go or not?'

*noŋ* [<sup>13</sup>] *aʔ* [<sup>5</sup>] *tɕ<sup>hi</sup>i* [<sup>34</sup>]  
 you 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):

*noŋ* [<sup>13</sup>] *aʔ* [<sup>5</sup>] *tɕ<sup>hi</sup>i vəʔ* [<sup>3-4</sup>]  
 you Q go Q?  
 'Did you go?'

*noŋ* [<sup>13</sup>] *aʔ* [<sup>5</sup>] *tɕ<sup>hi</sup>i vəʔ tɕ<sup>hi</sup>i* [<sup>3-5-1</sup>]  
 you Q go not go?  
 'Did or didn't you go?'

When the V-not-V form is used and the verb is transitive, Shanghaiese is different from Mandarin in word order. While the object can follow either V in Mandarin, in Shanghaiese it must follow the second V:

*noŋ* [13] *ts<sup>h</sup>ɣ* *vəʔts<sup>h</sup>ɣ* [5-3-1] *ʒua<sup>ŋ</sup>i* [5-1]  
 you smoke not smoke cigarette?  
 'Did you smoke or not?'

\**noŋ* [13] *ts<sup>h</sup>ɣ* [51] *ʒua<sup>ŋ</sup>i* [5-1] *vəʔts<sup>h</sup>ɣ* [1-3]  
 you smoke cigarette not smoke?  
 'Did you smoke or not smoke?'

Occasionally, the object as well as the verb can be reduplicated.

The interrogative particle *aʔ* cannot immediately precede the negative particle *vəʔ*. To ask a negative question, Shanghaiese uses *aʔz<sub>ɿ</sub>* instead of *aʔ*. The following sentence means: 'Is it the case that you won't go?'

*noŋ* [13] *aʔz<sub>ɿ</sub>* [4-5] *vəʔtʂ<sup>h</sup>i* [1-12]  
 you Q not go?  
 'Did you go or not?'

## 2.5 Negative structures

Negative sentences are formed by using one of the two negators, *vəʔ* and *ŋməʔ*.

The negative adverb *vəʔ* precedes and modifies an adjective or a verb:

*ji* [13] *vəʔtʂ<sup>ŋ</sup>ʒiŋ* [1-5-1]  
 he not careful  
 'He is not careful.'

*ji* [13] *vəʔlɛ* [1-3]  
 he not 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 *ŋməʔ* 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]<sub>s<sub>ɿ</sub> [34] *ji* [1] *kʂ<sup>ŋ</sup>* [34] *lipa* [1-3] *sɛ* [51] *ŋməʔ* [1-3] *tʂ<sup>h</sup>i* [34]  
 Thursday he said Wednesday not went  
 'On Thursday he said he had not gone on Wednesday.'</sub>

\**lipa* [1-3]<sub>s<sub>ɿ</sub> [34] *ji* [1] *kʂ<sup>ŋ</sup>* [34] *lipa* [1-3] *ŋ* [13] *ŋməʔ* [1-3] *tʂ<sup>h</sup>i* [34]  
 Thursday he said Friday not went  
 'On Thursday he said he would not go on Friday.'</sub>

The adverb *ŋməʔ* can be used to negate an event (i.e. action, achievement, or accomplishment), but not a state. To negate a state, *vəʔ* is used.

*ŋu* [<sup>13</sup>] *vəʔ* *ɕiɔtəʔ* [<sup>1-5-1</sup>] *gəʔtsɔʷ* [<sup>1-3</sup>] *zɰt<sup>h</sup>i* [<sup>1-3</sup>]  
 I not know this CL matter  
 'I don't know this matter.'

\**ŋu* [<sup>13</sup>] *ŋməʔ* [<sup>1-3</sup>] *ɕiɔtəʔ* [<sup>1-3</sup>] *gəʔtsɔʷ* [<sup>1-3</sup>] *zɰt<sup>h</sup>i* [<sup>1-3</sup>]  
 I not 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* *gəʔ* [<sup>1-3</sup>] *mik<sup>h</sup>oŋ* [<sup>3-4</sup>] *vəʔ* *fiŋ* [<sup>1-3</sup>]  
 his face not red  
 'His face is not red.'

*ji* *gəʔ* [<sup>1-3</sup>] *mik<sup>h</sup>oŋ* [<sup>3-4</sup>] *ŋməʔ* [<sup>1-3</sup>] *fiŋ* [<sup>13</sup>]  
 his face not-become red  
 'His face has not turned red.'

The latter is acceptable only if it is interpreted as involving a change of state.

The adverb *vəʔ* tends to occur immediately before the expression it modifies. Compare the following pair:

*ji* [<sup>13</sup>] *zɑ<sup>ŋ</sup>zɑ<sup>ŋ</sup>* [<sup>1-3</sup>] *vəʔ* *kɔʷ* [<sup>1-3</sup>] *fiɛwo* [<sup>1-3</sup>]  
 he often not speak speech  
 'He often does not speak.'

*ji* [<sup>13</sup>] *vəʔ* [<sup>12</sup>] *zɑ<sup>ŋ</sup>zɑ<sup>ŋ</sup>* [<sup>1-3</sup>] *kɔʷ* [<sup>34</sup>] *fiɛwo* [<sup>1-3</sup>]  
 he not 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* [<sup>13</sup>] *vəʔ* *kɔʷ* [<sup>1-3</sup>] *tɕiŋsɔʷ* [<sup>3-4</sup>]  
 he not speak clearly  
 'He does not speak clearly.'

*ji* [<sup>13</sup>] *kɔʷ* *vəʔ* *tɕiŋsɔʷ* [<sup>3-5-3-1</sup>]  
 he speak not clearly  
 'He can't speak clearly.'

It has been suggested by grammarians (Xu and Shao 1998: 148) that cliticized to the verb *kɔʷ*, *vəʔ* in the first sentence cannot extend its scope of negation to cover *tɕiŋsɔʷ*. In the second sentence, *vəʔ* directly negates *tɕiŋsɔʷ*, which is the focus of the sentence.

If *ŋməʔ* is used, the scope of negation may cover the entire verb phrase, not just the expression immediately following it.

*ji* [<sup>13</sup>] *ŋməʔ* [<sup>1-3</sup>] *kɔʷ* *tɕiŋsɔʷ* [<sup>3-5-1</sup>]  
 he not speak clearly  
 'He didn't speak clearly.'

This sentence means that he fails to make himself clear.

The older generation of Shanghainese speakers use *vəʔzəŋ* as an alternative of *ŋməʔ*. This expression comes from the dialects spoken in the region between Shanghai and the

Yangtze River. Although interchangeable in most contexts,  $\eta m \partial \eta$  and  $v \partial \eta z \partial \eta$  have some difference. The former can negate an existential state but the latter cannot.

$i g \partial \eta$  [<sup>5-1</sup>]     $z \partial \eta k^w \partial \eta$  [<sup>1-3</sup>]     $no \eta$  [<sup>13</sup>]     $\eta m \partial \eta$  [<sup>1-3</sup>]     $lah \epsilon$  [<sup>1-3</sup>]  
 that CL    time                    you            not            in  
 ‘You were not in at that time.’

\* $i g \partial \eta$  [<sup>5-1</sup>]     $z \partial \eta k^w \partial \eta$  [<sup>1-3</sup>]     $no \eta$  [<sup>13</sup>]     $v \partial \eta z \partial \eta$  [<sup>1-3</sup>]     $lah \epsilon$  [<sup>1-3</sup>]  
 That CL    time                    you            not            in  
 ‘That time you were not in.’

## REFERENCES

- Chao, Yuan Ren (1928) *Studies in the Modern Wu-Dialects*, Beijing: Tsing Hua College Research Institute, Monograph No. 4.
- Chao, Yuan 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, 106–10.
- Cao, Jianfen and Maddieson, Ian (1992) ‘An exploration of phonation types in Wu dialects of Chinese’, *Journal of Phonetics* 20: 77–92.
- Jin, Shunde (1986) *Shanghai Morphotonemics*, Bloomington: Indiana University Linguistics Club.
- Nespor, Marina and Vogel, Irene (1986) *Prosodic Phonology*, Dordrecht: Foris.
- Ren, Nianqi (1995) ‘A fiberoptic and transillumination study of Shanghai stops’, in Eric Zee (ed.) *Studies of Wu Dialects*, Hong Kong New Asia College, the Chinese University of Hong Kong, 261–7.
- 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).
- 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, 313–37.
- Sherard, Michael (1972) ‘Shanghai Phonology’, unpublished PhD dissertation, Cornell University.
- Xu, Liejiong and Shao, Jingmin (1998) *Studies on Shanghainese Grammar*, Shanghai: East China Normal University Press (in Chinese).
- Xu, Baohua and Tang, Zhenzhu (1988) *Shanghai Shiqu Fangyan Zhi (A Study of the Metropolitan Shanghai Dialect)*, Shanghai: Shanghai Education 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.



# CANTONESE

*Robert S. Bauer and Stephen Matthews*

## 1 INTRODUCTION

The term ‘Cantonese’ designates the speech of 廣州 *kwɔŋ<sup>3</sup>-tsɐw<sup>1</sup>* Guangzhou, or Canton, the name by which the provincial capital has been known in the West; it also refers to the Chinese variety spoken by the majority of the ethnic Chinese population of the Hong Kong Special Administrative Region which borders southern Guangdong, as well as that of many overseas Chinese communities in Europe, North America, and Australasia. The Cantonese spoken in 西關 *sɛŋ<sup>1</sup>-kwən<sup>1</sup>*, a district of Guangzhou, has been traditionally regarded as the prestige form; although both the Guangzhou and Hong Kong varieties are quite similar, differences of pronunciation, vocabulary, and grammar are to be found. Cantonese speakers themselves call their speech by a variety of names, e.g. 廣州話 *kwɔŋ<sup>3</sup>-tsɐw<sup>1</sup>-wa<sup>6\*</sup>* ‘speech of Guangzhou’, 廣東話 *kwɔŋ<sup>3</sup>-toŋ<sup>1</sup>-wa<sup>6\*</sup>* ‘speech of Guangdong’, 香港話 *hœŋ<sup>1</sup>-kɔŋ<sup>3</sup>-wa<sup>6\*</sup>* ‘speech of Hong Kong’, or 白話 *pak<sup>8</sup>-wa<sup>6\*</sup>* ‘plain speech’ (in the romanization 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; the asterisk marks the high rising changed tone). In the popular imagination Cantonese has regional dialect status even among speakers of other varieties of Chinese, but 普通話 *phow<sup>3</sup>-thoŋ<sup>1</sup>-wa<sup>6\*</sup>* ‘common speech’, which is based on 北方官話 *pɛk<sup>7a</sup>-fɔŋ<sup>1</sup> kwun<sup>1</sup>-wa<sup>6\*</sup>* ‘northern Mandarin’ dialects, is esteemed as China’s national language. In China, Cantonese is classified as a 方言 *fɔŋ<sup>1</sup>-jin<sup>2</sup>* ‘dialect’ of 粵 *jy<sup>8</sup>* or Yue, one of the seven major Chinese dialect families recognized by Chinese dialectologists; however, given the mutual unintelligibility among these so-called dialects, Western sinologists have usually regarded Cantonese, along with 客家話 Hakka, 吳語 Wu, 贛語 Gan, 湘語 Xiang, 閩語 Min, and 官話 Mandarin as related languages. The Yue dialects are distributed across most of Guangdong and the eastern part of Guangxi. The worldwide number of speakers of Yue dialects exceeds 40 million.

As noted by Norman (1988: 217–18), the defining characteristic of a Yue dialect (with only a few exceptions) appears to be the division of the 陰人 Yin Ru tone category into two subcategories which has been conditioned by vowel length in the reading pronunciation of standard Chinese characters (colloquial pronunciations of some lexical items may not follow this generalization); e.g. 上陰人 upper Yin Ru with high stopped tone and short vowel: 黑 *hɛk<sup>7a</sup>* ‘black’, 濕 *sɛp<sup>7a</sup>* ‘wet’; and 下陰人 lower Yin Ru with mid stopped tone and long vowel: 客 *hak<sup>7b</sup>* ‘guest’, 颯 *sap<sup>7b</sup>* ‘sound of wind’. Given the early historical contact between Proto-Yue and the Tai languages (which contrast long and short vowels) in Southern China, this distinctive development in Yue may represent an early Tai substratum (Bauer 1996).

## 2 PHONOLOGY

### 2.1 Initial and final consonants

As indicated in Table 9.1 below, the phonological inventory of standard Hong Kong Cantonese comprises nineteen initial consonants (or twenty if we include the so-called

**TABLE 9.1** CANTONESE INITIAL CONSONANTS

<i>Manners of articulation</i>	<i>Places of articulation</i>				
	<i>labial</i>	<i>dental/ alveolar</i>	<i>post-alveolar/ palatal</i>	<i>velar</i>	<i>glottal</i>
stop					
unaspirated	/p/	/t/		/k/ /kw/	(ʔ)
aspirated	/ph/	/th/		/kh/ /khw/	
nasal	/m/	/n/		/ŋ/	
fricative	/f/	/s/	[ç]		/h/
affricate					
unaspirated		/ts/	[tç]		
aspirated		/tsh/	[tçh]		
approximant	/w/	/l/	/j/		

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 these palatalized consonants are the regular realizations of the sibilant fricative and affricate initials in Guangzhou Cantonese. 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 very skewed distribution. The following nine consonants occur in syllable-final position: *-m, -n, -ŋ, -p, -t, -k, -w, -j, -ŋ* (final semi-vowel *-ŋ* occurs only after the mid central round vowel *ɵ* 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 9.2 below, Cantonese includes eleven vowel phonemes and fourteen vowel allophones.

Some vowels are long and some short in duration. Most analyses treat only *ɐ* and *a* as contrasting in length; however, one acoustic study by Li (1985) indicates the length difference is the primary cue for distinguishing *e<sup>i</sup>* / *ɛ* and *o<sup>w</sup>* / *ɔ* when they occur before velar consonants *-ŋ, -k*. The vowels join with the nine final consonants to form fifty-six rhymes which are listed below in Table 9.3.

**TABLE 9.2** CANTONESE VOWEL PHONEMES (BETWEEN / /) AND ALLOPHONES (BETWEEN [ ])

	<i>Front</i>	<i>Central</i>	<i>Back</i>
High	/i/ [i:] /y/ [y:]		/u/ [u:]
High-mid	/e/ [e], [e <sup>i</sup> ]	/ɵ/ [ɵ]	/o/ [o], [o <sup>w</sup> ]
Mid	/ɛ/ [ɛ:]	/ɐ/ [ɐ]	/ɔ/ [ɔ:]
Low	/œ/ [œ:]	/a/ [a:]	

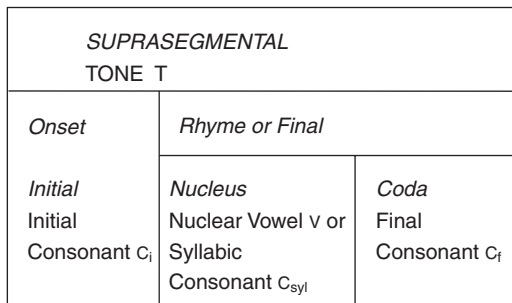
**TABLE 9.3** FIFTY-SIX CANTONESE RHYMES

	<i>-j (-ŋ)</i>	<i>-w</i>	<i>-m</i>	<i>-n</i>	<i>-ŋ</i>	<i>-p</i>	<i>-t</i>	<i>-k</i>
[i:]		[i:w]	[i:m]	[i:n]		[i:p]	[i:t]	
[y:]				[y:n]			[y:t]	
*[e <sup>j</sup> ]					[e <sup>j</sup> ŋ]			[e <sup>j</sup> k]
*[e]	[e:j]							
[ɛ:]		[ɛ:w]	[ɛ:m]	[ɛ:n]	[ɛ:ŋ]	[ɛ:p]	[ɛ:t]	[ɛ:k]
[œ:]					[œ:ŋ]			[œ:k]
*[ø]	[øɥ]			[ø:n]			[ø:t]	
*[ø]	[ø:j]	[ø:w]	[ø:m]	[ø:n]	[ø:ŋ]	[ø:p]	[ø:t]	[ø:k]
[a:]	[a:j]	[a:w]	[a:m]	[a:n]	[a:ŋ]	[a:p]	[a:t]	[a:k]
[u:]	[u:j]			[u:n]			[u:t]	
*[o <sup>w</sup> ]					[o <sup>w</sup> ŋ]			[o <sup>w</sup> k]
*[o]		[ow]						
[ɔ:]			[ɔ:j]	[ɔ:n]		[ɔ:p]	[ɔ:t]	[ɔ:k]

Vowels preceded by the asterisk in the above table do not occur as independent rhymes in open syllables. The above rhymes in turn combine with the initial consonants to form 763 syllables (out of a possible 1122) which occur in the pronunciations of standard Chinese characters, colloquial Cantonese words, and English loanwords.

### 2.3 Cantonese syllable

The Cantonese syllable takes a relatively simple structure: every syllable must comprise at least a rhyme or final as its nucleus which may be preceded and/or followed by a consonant (there are no consonant clusters except under special circumstances, and then only for some speakers); 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 the diagram below:

**FIGURE 9.1** STRUCTURE OF THE CANTONESE SYLLABLE

### 2.4.1 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 tones in Cantonese is either 6 or 7, or 9 or 10; it depends on whether one is talking about Hong Kong or Guangzhou Cantonese, and whether one regards the tone contours on the 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*. The historical devoicing of originally voiced initial consonants in Ancient Chinese doubled the number of tones in some dialects; in Modern Cantonese upper and lower 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, Mid-Low Rising, Mid-Low Level, and Mid-Low Stopped.

Some older, educated speakers of Guangzhou Cantonese contrast the High Falling and High Level tones, with the latter tone functioning as a morphological device to mark certain types of concrete nouns (this is one of so-called changed tones or *pin<sup>5</sup>-jəm<sup>1</sup>* 變音). Most speakers of Hong Kong Cantonese, on the other hand, only have the High Level tone. Table 9.4 below illustrates the ten tone contours with their corresponding Chao tone letters.

### 2.4.2 Changed tones

In addition to the tones listed above, there are two so-called changed tones or *pin<sup>3</sup>-jəm<sup>1</sup>* 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 indicate the two changed tones; e.g. the Mid-Low Falling tone on the second syllable of *a-ji<sup>20</sup>* 阿姨 ‘auntie’ typically changes to the High Level *pin<sup>5</sup>-jəm<sup>1</sup>* which is marked with the small raised circle; the Mid-Low Level tone of the second syllable of *a-ji<sup>6\*</sup>* 阿二 ‘Number Two servant’ changes to the High Rising *pin<sup>5</sup>-jəm<sup>1</sup>* and is marked with the asterisk. As noted by Y.R. 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’.

## 2.5 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

**TABLE 9.4** LEXICAL ITEMS CONTRASTING SEVEN TONES ON OPEN SYLLABLE *ji*

Lexical item	Tone contour	Tone letter	Tone category
1a <i>ji<sup>1a</sup></i> 衣 clothes	High Level	˥55	上陰平 High Upper Even
1b <i>ji<sup>1b</sup></i> 醫 to cure	High Falling	˥52	下陰平 Low Upper Even
2 <i>ji<sup>2</sup></i> 疑 suspicious	Mid-Low Falling	˨21	陽平 Lower Even
3 <i>ji<sup>3</sup></i> 椅 chair	High Rising	˥25	陰上 Upper Rising
4 <i>ji<sup>4</sup></i> 耳 ear	Mid-Low Rising	˨23	陽上 Lower Rising
5 <i>ji<sup>5</sup></i> 意 idea	Mid Level	˥33	陰去 Upper Going
6 <i>ji<sup>6</sup></i> 二 two	Mid-Low Level	˨22	陽去 Lower Going
7a <i>jek<sup>7a</sup></i> 益 benefit	High Stopped	˥5	上陰人 High Upper Entering
7b <i>jak<sup>7b</sup></i> 喫 to eat	Mid Stopped	˨33	下陰人 Low Upper Entering
8 <i>jek<sup>8</sup></i> 亦 also	Mid-Low Stopped	˨2	陽人 Lower Entering

literary form has a short vowel and the colloquial form a long vowel; e.g. 生 *sɛŋ<sup>l</sup>*, *saŋ<sup>l</sup>* ‘life’, 耕 *kɛŋ<sup>l</sup>*, *kaŋ<sup>l</sup>* ‘plough’, 行 *hɛŋ<sup>2</sup>*, *haŋ<sup>2</sup>* ‘walk’, 爭 *tsɛŋ<sup>l</sup>*, *tsaŋ<sup>l</sup>* ‘to quarrel’, 平 *phɛŋ<sup>2</sup>*, *phɛŋ<sup>2</sup>* ‘flat, cheap’, 瓶 *phɛŋ<sup>2</sup>*, *phɛŋ<sup>2</sup>* ‘vase’, 頂 *tɛŋ<sup>2</sup>*, *tɛŋ<sup>2</sup>* ‘top’, ‘extremity’, 聽 *thɛŋ<sup>l</sup>*, *thɛŋ<sup>l</sup>* ‘hear, listen’.

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 twenty years or so, several sociolinguistic studies (reviewed in Bauer and Benedict 1997: 327–42) have investigated the following sound changes:  $\eta \rightarrow m$ ;  $kw- \rightarrow k-$ ,  $khw- \rightarrow kh-$ ;  $kh\epsilon y^5 \rightarrow h\epsilon y^5$  (only in third person pronoun),  $\eta- \rightarrow \emptyset-$ ;  $n- \rightarrow l-$ ;  $-V\eta \rightarrow -Vn$ ,  $-Vk \rightarrow -Vt$ . 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 *lan<sup>4</sup>* *jɛm<sup>l</sup>* ‘lazy pronunciation’ which reflects their ignorance of the norms of Cantonese speech.

### 3 LEXICON: 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 may have dialectal characters as their written forms, while others have no written form (as indicated by the empty box), e.g. *pɛw<sup>6</sup>* □ ‘jostle with hips’, *phɛ<sup>d</sup>* □ ‘stagger’, *tɛm<sup>5</sup>* 掙聚 ‘droop, hang down’, *tɛp<sup>8</sup>* 才泵 ‘beat, pound’, *thɛm<sup>4</sup>* 𪗇 ‘puddle’, *tɛw<sup>6</sup>* □ ‘throw away’, *tɛŋ<sup>5</sup>* 掙定 ‘throw (at target)’, *kat<sup>8</sup>-tsat<sup>8</sup>\** 𪗇 𪗇 ‘cockroach’, *khɛ<sup>l</sup>* □ ‘shit’, *khwak<sup>7a</sup>* 𪗇 𪗇 ‘loop’, ‘circle’, *ni<sup>l</sup>* 呢 ‘this’, *nɛm<sup>2</sup>* 脛 ‘soft’, *ŋam<sup>l</sup>* 𪗇 岩 ‘all right, good’, *jaŋ<sup>5</sup>* □ ‘to kick off’. Such items may represent a non-Han, viz. Austro-Tai and/or Austro-Asiatic, substratum that reflects the historical contact Cantonese has had with languages of these families.

#### 3.1 English loanwords

No Chinese dialect has been more influenced by a European language than Cantonese; as a result of its historical contact with English, which dates back 300 years, a large number of English words have been borrowed into Hong Kong and Guangzhou Cantonese. Many of these loanwords have become so assimilated that they have written forms with Chinese characters and many Cantonese speakers assume they are Chinese words; e.g. *pa1-si<sup>6</sup>\** 巴士 ‘bus’, *tsi<sup>l</sup>-si<sup>6</sup>\** 芝士 ‘cheese’, *tɔ<sup>l</sup>-si<sup>6</sup>\** 多士 ‘toast’, *khat<sup>7a</sup>* 咭 ‘card’, *fej<sup>l</sup>-lɛm<sup>2</sup>\** 菲林 ‘film’, *thip<sup>7a</sup>-si<sup>6</sup>\** 𪗇 士 ‘tips’. The phonetic adaptation of these loanwords has influenced the structure of the Cantonese syllabary in that some syllables only occur in English loanwords.

#### 3.2 Written Cantonese

Hong Kong Cantonese distinguishes itself from some other Chinese dialects through the informal development of a written form in which there is a one-to-one correspondence between speech and writing. Writing in Cantonese has become a pervasive phenomenon in Hong Kong newspapers, magazines, comic books, subway advertisements, etc., but its conventions remain unstandardized and inconsistent (cf. Bauer 1982, 1988; Cheung and Bauer forthcoming). Helping to promote the extensive use of written Cantonese in Hong Kong has been the tradition of teaching students to read the standard Chinese characters with Cantonese pronunciation; by contrast, schoolchildren in Cantonese-speaking areas of the mainland learn to read in Putonghua. Many colloquial Cantonese words are not etymologically related to standard Chinese, so these words may not have Chinese characters

associated with them; as a consequence, Cantonese writers have had to create characters to transcribe these words.

#### 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 1972; Matthews and Yip 1994; Matthews 1998a). 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 (1991; 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 9.5 below.

TABLE 9.5 REGISTER STRATIFICATION IN CANTONESE SYNTAX

	<i>High Cantonese</i>	<i>Low Cantonese</i>
Passive voice	<i>pej</i> <sup>6</sup> 被 (NP) V	<i>pej</i> <sup>3</sup> 俾 NP V
Comparative degree	<i>pej</i> <sup>3</sup> 比 NP <sub>1</sub> Adj NP <sub>2</sub>	NP <sub>1</sub> Adj <i>kwɔ̃</i> <sup>5</sup> 過 NP <sub>2</sub>
Excessive degree (too)	<i>thaj</i> <sup>5</sup> 太 Adj	Adj <i>kwɔ̃</i> <sup>5</sup> <i>thɛw</i> <sup>2</sup> 過 頭, Adj <i>tɛk</i> <sup>7a</sup> <i>tsɛj</i> <sup>6</sup> 得 滯
Possessive marker	NP <i>ke</i> <sup>5</sup> 嘅 N	NP Classifier N
Relative clause	[s...] <i>ke</i> <sup>5</sup> 嘅 N	[s...] <i>kɔ̃</i> <sup>3</sup> 個 Classifier N

#### 4.1 Passive construction

The passive construction with colloquial *pej*<sup>3</sup> 俾 superficially resembles Mandarin *bèi* 被, but its origin as a grammaticalized form of the verb *pej*<sup>3</sup> 俾 ‘give’ is reflected in syntactic and semantic differences. In particular, the Cantonese passive with *pej*<sup>3</sup> 俾 requires an agent phrase. In the following sentence *jɛn*<sup>2</sup> 人 ‘person’ is supplied as the generic agent (PERF indicates the perfective aspect marker).

- (1) *khəu*<sup>4</sup> *pej*<sup>3</sup> *jɛn*<sup>2</sup> *laj*<sup>1</sup> *tsɔ̃*<sup>3</sup>  
 佢 俾 人 拉 左  
 s/he by person arrest-PERF  
 ‘S/he was arrested by someone.’

#### 4.2 Comparative constructions

Among excessive constructions, for example, *thaj*<sup>5</sup> *jit*<sup>8</sup> 太熱 ‘too hot’ corresponds to Mandarin *tài rè* and is used in neutral or formal register, while *jit*<sup>8</sup> *kwɔ̃*<sup>5</sup> *thɛw*<sup>2</sup> 熱過頭 and *jit*<sup>8</sup> *tɛk*<sup>7a</sup> *tsɛj*<sup>6</sup> 熱得滯 are more colloquial. There is evidence for change from the indigenous toward the pan-Chinese model: Adj. + *kwɔ̃*<sup>5</sup> *thɛw*<sup>2</sup> 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 *thaj*<sup>5</sup> is head-final. This applies equally to comparative structures with *kwɔ*<sup>5</sup> 過 and *pej*<sup>3</sup> 比, respectively head-initial and head-final:

- (2) *khəŋ*<sup>4</sup> *lek*<sup>7a</sup> *kwɔ*<sup>5</sup> *ŋɔ*<sup>4</sup>  
 佢 叻 過 我  
 s/he smart -er I  
 ‘She’s smarter than I.’
- (3) *khəŋ*<sup>4</sup> *pej*<sup>3</sup> *ŋɔ*<sup>4</sup> *tshoŋ*<sup>1</sup> *meŋ*<sup>2</sup>  
 佢 比 我 聰 明  
 s/he compared I 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 *kwɔ*<sup>5</sup> 過 appears to be cognate with Thai *kwàx*, 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’ (Stassen 1985). This type of comparative is typically found in SVO languages.

#### 4.3 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 1994): one with the particle *kɛ*<sup>5</sup> 嘅 corresponding to Mandarin *dè* 的, and the other using a demonstrative and classifier:

- (4) *khəŋ*<sup>4</sup> *tshæŋ*<sup>5</sup> *kɛ*<sup>5</sup> *kɔ*<sup>1</sup> *how*<sup>3</sup> *səm*<sup>1</sup> *kɛ*<sup>5</sup>  
 佢 唱 嘅 歌 好 深 嘅  
 s/he sing PRT song very deep PRT  
 ‘The song(s) she sings is/are rather deep.’
- (5) *khəŋ*<sup>4</sup> *tshæŋ*<sup>5</sup> *kɔ*<sup>3</sup> *sɛw*<sup>3</sup> *kɔ*<sup>1</sup> *how*<sup>3</sup> *lan*<sup>2</sup> *tshæŋ*<sup>5</sup> *kɛ*<sup>5</sup>  
 佢 唱 個 首 歌 好 難 唱 嘅  
 s/he sing that CL song very hard sing  
 ‘The song she sings is hard to sing.’

The classifier construction has a specific reference and belongs to the colloquial register (Matthews and Yip 2001). Similarly, possessive constructions use either *kɛ*<sup>5</sup> or a classifier to link the possessor and possessed noun:

- (6) *khəŋ*<sup>4</sup> *kɛ*<sup>5</sup> *thɛj*<sup>3</sup> *fat*<sup>7b</sup> *thaj*<sup>5</sup> *pej*<sup>1</sup> *kun*<sup>1</sup> *kɛ*<sup>5</sup>  
 佢 嘅 睇 法 太 悲 觀 嘅  
 s/he POSS attitude too pessimistic PRT  
 ‘Her attitude is too pessimistic.’
- (7) *khəŋ*<sup>4</sup> *kan*<sup>1</sup> *ok*<sup>7a</sup> *how*<sup>3</sup> *taj*<sup>6</sup> *kan*<sup>1</sup>  
 佢 間 屋 好 大 間  
 she CL house very big CL  
 ‘Her house is very big.’

In addition to the possessive use of the classifier in *khəŋ<sup>4</sup> kan<sup>1</sup> ok<sup>7a</sup>*, we should note here the use of the classifier following the adjective *taj<sup>6</sup>*. 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) *kan<sup>1</sup> ok<sup>7a</sup> how<sup>3</sup> taj<sup>6</sup>*  
 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.

#### 4.4 Verb phrase

Some word order patterns in the verb phrase are distinctively Cantonese. A small number of adverbs, notably *sin<sup>1</sup>* 先 ‘first’ and *thim<sup>1</sup>* 忝 ‘additionally’ (cf. *kwɔ<sup>5</sup> thəw<sup>2</sup>* and *tək<sup>7a</sup> tsəj<sup>6</sup>* above), follow the verb:

- (9) *ŋɔ<sup>4</sup> -tej<sup>6</sup> tsəw<sup>3</sup> sin<sup>1</sup>*  
 我 哋 走 先  
 we go first  
 ‘We’re leaving now.’

In the double object construction the indirect object of *pej<sup>3</sup>* ‘give’ follows the direct:

- (10) *ŋɔ<sup>4</sup> pej<sup>3</sup> tshin<sup>2\*</sup> lej<sup>4</sup>*  
 我 俾 錢 你  
 I give money you  
 ‘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) *ŋɔ<sup>4</sup> pej<sup>3</sup> lej<sup>4</sup> kəm<sup>5</sup> tɔ<sup>1</sup> tshin<sup>2\*</sup>, lej<sup>4</sup> tow<sup>1</sup> m<sup>2</sup> jiw<sup>5</sup>?*  
 我 俾 你 咁 多 錢 你 都 唔 要  
 I give you so much money you all not want  
 ‘I give you so much money (and) you don’t want it?’

The pretransitive or ‘disposal’ construction uses *tsæŋ<sup>1</sup>* 將 (OM, object marker; cognate with Mandarin *jiāng*). Like Mandarin *bǎ* 把, this serves to place a definite or specific object before the verb, but the construction with *tsæŋ<sup>1</sup>* is more restricted in function, typically retaining a sense of displacement:

- (12) *ŋɔ<sup>4</sup> tsæŋ<sup>1</sup> til sam<sup>1</sup> paj<sup>3</sup> həj<sup>3</sup> lej<sup>4</sup> tow<sup>6</sup>*  
 我 將 D 衫 擺 係 你 度  
 I OM PL clothes put at you there (PL = plural marker for following noun)  
 ‘I’m putting the clothes in your place/room.’



#### 4.5 Utterance-final particles

The utterance-final particles (PRT) of Cantonese are particularly rich both in sheer number and in their combinatorial possibilities. Some thirty basic particles have been identified (Kwok 1984), and most of them occur in several variant forms and combinations (Bauer and Benedict, 1997: 291–295; Leung 1992; Yao 1980). One example involves a set of particles having evidential functions, including  $w\text{ɔ}^5$  denoting surprising or notable information (Luke 1990; Matthews 1998b) and  $w\text{ɔ}^4$  ‘hearsay’ evidentiality:

- (13)  $kh\text{ə}y^4-tej^6$   $jaw^6$   $pun^1$   $ok^{7a}$   $w\text{ɔ}^5$   
 佢 咁地 又 搬 屋 喎  
 they again move house PRT  
 ‘You know, they’re moving house again (believe it or not).’

- (14)  $kh\text{ə}y^4-tej^6$   $jaw^6$   $pun^1$   $ok^{7a}$   $w\text{ɔ}^4$   
 佢 咁地 又 搬 屋 喎  
 they 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

- Bauer, Robert S. (1982) ‘D for two in Cantonese’, *Journal of Chinese Linguistics* 10.2: 277–81.  
 Bauer, Robert S. (1988) ‘Written Cantonese of Hong Kong’, *Cahiers de Linguistique Asie Orientale* XVII.2: 245–93.  
 Bauer, Robert S. (1996) ‘Identifying the Tai substratum in Cantonese’, *Proceedings of the Fourth International Symposium on Languages and Linguistics, Pan-Asiatic Linguistics V*: 1806–44, Bangkok: Institute of Language and Culture for Rural Development, Mahidol University.  
 Bauer, Robert S. and Benedict, Paul K. (1997) *Modern Cantonese Phonology*, 摩登廣州話語音學, Berlin: Mouton de Gruyter.  
 Cheung Kwan-hin and Bauer, Robert S. (forthcoming). *The representation of Cantonese with Chinese Characters*, *Journal of Chinese Linguistics*, Monograph, Berkeley, CA.  
 Chao Yuen Ren (1968) *A Grammar of Spoken Chinese*, Berkeley: University of California Press.  
 Hashimoto Mantaro (1983) *Yuyan dili leixingxue*, Beijing: Beijing daxue chubanshe.  
 Hashimoto Mantaro (1987) ‘Question and its presupposition in Chinese’, in *Wang Li Memorial Volumes* (English volume), Chinese Language Society of Hong Kong, 147–77.  
 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.  
 Leung Chung Sam (1992) *A Study of the Utterance Particles in Cantonese as Spoken in Hong Kong*, unpublished MPhil thesis, Hong Kong Polytechnic University.  
 Li Xingde (1985) ‘Guangzhouhua yuanyinde yinzhì jì chángduàn duìlì (phonetic quality and long-short opposition of Cantonese vowels)’, *Fangyan* 1: 28–38.  
 Lucas, Alain, and Xie Honghua (1994) ‘There is practically one universal Chinese grammar: A 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 (1999) 'Y.R. Chao and universal Chinese grammar', in D. Cram *et al.* (eds) *History of Linguistics 1996*, vol. 1. Amsterdam: John Benjamins, 27–224.
- 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 and Patrizia Pacioni (1997) 'Specificity and genericity of NPs in Cantonese and Mandarin', in Xu Liejiong (ed.) *The Referential Properties of Chinese Noun Phrases*, Paris: EHESS, Centre de Recherches Linguistiques sur l'Asie Orientale.
- Matthews, Stephen and Yip, Virginia (1994) *Cantonese: A Comprehensive Grammar*, London: Routledge.
- 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, 266–81.
- Norman, Jerry (1988) *Chinese*, Cambridge: Cambridge University Press.
- O'Melia, Thomas (1941) *First Year Cantonese*, Second Edition, Hong Kong: Maryknoll House.
- Stassen, Leon (1985) *Comparison and Universal Grammar*, Oxford: Blackwell.
- Yao Shun-chiu (1980) 'Sentential connotations in Cantonese', *Fangyan* 1: 35–52.
- Yue-Hashimoto, A. (1991) 'Stratification in comparative dialectal grammar: a case in Southern Min', *Journal of Chinese Linguistics* 20: 172–200.
- Yue-Hashimoto, A. (1993) *Comparative Chinese Dialectal Grammar: A Handbook for Investigators*, Paris: EHESS, Centre de Recherches Linguistiques sur l'Asie Orientale.
- Zee, Eric (1999a) 'An acoustical analysis of the diphthongs in Cantonese', *Proceedings of the XIV Congress of Phonetic Sciences*, 1–4.
- Zee, Eric (1999b) 'Resonance frequency and vowel transcription in Cantonese', *Proceedings of the 10th North American Conference of Chinese Linguistics and the 7th Annual Meeting of the International Association of Chinese Linguistics 1999*, 90–7.

# 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 discreet, 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 with an increasing number of strokes, from one to thirty

一 十 千 木 日 耳 尾 朋 亮 豈 梅 超 亂 認 鬧 壘 點 壘 懷 黨 蠟 竊 驗 癱 欖 鑽 鑿 鑿 鬱 鸞

Strokes are not only the basic unit in the physical act of writing, but also are 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, telephone books, etc.) are based on number and configuration of strokes in Chinese writing.

The majority of characters are graphic compounds, composed of two or more components (either other, simpler characters, or frequently used multi-stroke elements). This compounding adds a third layer of structure to the script, but since it is intimately related to the expression of sound and meaning, it will be dealt with below in the section on 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>=*eleven*); and it can represent different words that are synonyms (<1000>=*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, one of the Chinese vernaculars), 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; all are invited to match the graph with the referent by pictographic principles:

鯨 鮑 魷 鰻 鮪

### 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 with its own independent existence in the written language (or a reduced form of an independent character). 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 10.1, first row). The right-hand elements of these characters, if used alone, would be the characters given in the second row of Figure 10.1. They are used as phonetics in the characters in the first row because of phonetic similarity. 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

鯨 <i>jīng</i> whale	鮑 <i>bāo</i> abalone	魷 <i>yóu</i> cuttlefish	鰻 <i>mán</i> eel	鮪 <i>wěi</i> tuna
京 <i>jīng</i> capital city	包 <i>bāo</i> to wrap	尤 <i>yóu</i> especially	曼 <i>màn</i> graceful	有 <i>yǒu</i> to have

FIGURE 10.1 COMPOUND CHARACTERS

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 usefulness 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 twenty-two characters with the phonetic <堯> give fifteen different pronunciations (*yāo yáo yǎo jiāo jiǎo qiāo qiáo xiāo xiáo 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 10.2, <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 their distribution is not even. 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 unlearned, 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 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 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.

罵 <i>mà</i> to scold	碼 <i>má</i> symbol	瑪 <i>mǎ</i> agate	媽 <i>mā</i> mother	鎊 <i>mǎ</i> masurium
駝 <i>tuó</i> camel	駁 <i>bó</i> to refute	駐 <i>zhù</i> to station	騎 <i>qí</i> to ride	駕 <i>jià</i> to drive

FIGURE 10.2 CHARACTERS WHICH INCLUDE THE ‘HORSE’ RADICAL

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 2000 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 6–8000 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 *zhōnggǔ* 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, is 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 printing, Chinese fell behind the West technologically in typewriters and early computers, which could not handle the thousands of characters. More advanced computers obviate the screen display and printing problems, but use of the alphabetic keyboard for input remains a bottleneck,

perhaps to be resolved by the development of reliable speech recognition and handwriting recognition software.

## 2.2 Development of 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 PRC 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 by merely giving official blessing to folk shorthand characters). An example of the historical progression can be seen in Figure 10.3.

## 2.3 Development of orthography

Xu Shen's *c.*100 CE dictionary *Shuowen jiezi* establishes a sixfold classification of characters (or *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 *xiangxing* ('resembling form'): pictographs, as 'horse' in Figure 10.3.
- 3 *xingsheng* ('form-sound'): characters consisting of a radical and a phonetic.
- 4 *huiyi* ('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').

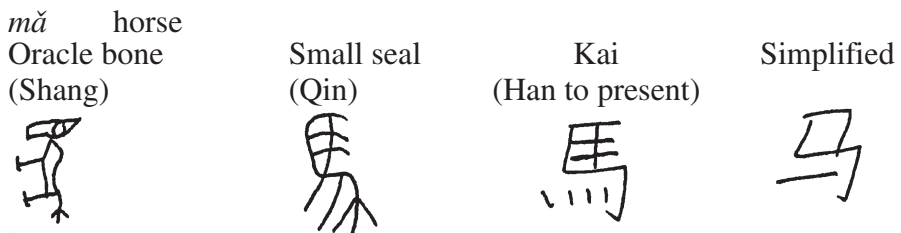


FIGURE 10.3 THE HISTORICAL DEVELOPMENT OF CHINESE CHARACTERS



- 5 *zhuanzhu* ('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 *jiajie* ('borrowed'): rebus characters. A character with a particular meaning is used to write a homophonous character of unrelated meaning. For example, 被 *bèi* '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 (*xiangxing*). 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 (*jiajie*) 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 the 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 (*xingsheng*) character. For example, in Figure 10.4, there is no written form for 'emperor' at Time 1. In Time 2, the use of 'stem' as a rebus allows 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 *gū* '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 (*zhuanzhu*), the above process accounts for all of the categories except type 4 (*huiyi*). The concept of character formation by combination 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 Boodberg (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 mysterious thought processes of the ancients. In a typically egregious example, Aria (1991) states that 星 *xīng* 'star' is

Time 1	帝	<i>dì</i>	stem, base of fruit	( )	<i>dì</i>	emperor
Time 2	帝	<i>dì</i>	stem, base of fruit	帝	<i>dì</i>	emperor
Time 3	蒂	<i>dì</i>	stem, base of fruit	帝	<i>dì</i>	emperor

FIGURE 10.4 THE DEVELOPMENT OF 'FORM-SOUND' CHARACTERS

... 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. (p. 87)

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 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 fifteen 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 Fayuan'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<sup>h</sup>] 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.

#### 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 the common 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 10.5 the phonetic <吳> is not intended to associate [u] with [gɔ]: because orthographies are language-specific, this phonetic's value in Mandarin is different from its value in Southern Min. The association between [u] and [gɔ] 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 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

		<u>Mandarin</u>	<u>Southern Min</u>
吳	<i>Wu (a surname)</i>	<i>u</i> <sup>35</sup>	<i>gɔ</i> <sup>24</sup>
誤	<i>mistake</i>	<i>u</i> <sup>51</sup>	<i>gɔ</i> <sup>22</sup>
娛	<i>amuse</i>	<i>y</i> <sup>35</sup>	<i>gɔ</i> <sup>22</sup>

FIGURE 10.5 THE REALIZATION OF A PHONETIC IN DIFFERENT DIALECTS

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.

## REFERENCES

- Aria, Barbara (1991) *The Nature of the Chinese Character*, New York: Simon and 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: 329–72.
- DeFrancis, John (1989) *Visible Speech: the Diverse Oneness of Writing Systems*, Honolulu: University of Hawaii Press.
- Gong, Hwang-chen (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.
- Zhou, Youguang (1989) '汉字文化圈的文字演变' (The evolution of writing in the Chinese character cultural sphere), *民族语文* 1989. 1: 37–55.

## 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.
- Norman, Jerry (1988) *Chinese*, New York: Cambridge 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, Yuh (1993) *A History of Chinese Calligraphy*, Hong Kong: The Chinese University Press.



**PART 3**

---

**TIBETO-BURMAN**

---

**LANGUAGES AND**

---

**DIALECTS**

---



# THE TIBETO-BURMAN LANGUAGES OF NORTHEASTERN INDIA<sup>1</sup>

*Robbins Burling*

## 1 INTRODUCTION

India's densest concentration of Tibeto-Burman languages is found in its far northeast where an almost disconnected fragment of the country is nearly encircled by Bhutan, Tibet, Myanmar, and Bangladesh (see Map 11.1). By comparison with most Indian states, the seven states of the northeast are small in area and low in population, but even by Indian standards they are very high in linguistic and ethnic heterogeneity. At their centre is Assam, the only state of the seven that is predominantly lowland. A hilly extension to the south is still a part of Assam, but the heart of the state consists of the low valley of the Brahmaputra River. Surrounding Assam are the six 'hill states,' an oddly modest way to describe not only the mountainous region that divides India from Myanmar, but even the eastern slope of the Himalayas where the state of Arunachal Pradesh rises all the way up to its border with Tibet.

The great majority of the indigenous people of Arunachal on the north and of Nagaland, Manipur, and Mizoram in the east speak Tibeto-Burman languages. Two or three generations ago Tripura State was also predominantly Tibeto-Burman, but so many Bengalis have now settled there that the indigenous Tibeto-Burman speaking Kokborok, once the dominant group, have been reduced to a minority in their own homeland. Meghalaya, south of western Assam, is the only hill state with a substantial and long established non-Tibeto-Burman population. The Khasis and the closely related Jaintia, War, and Lyngngam, who form a bit over half the state's population and who, together, occupy its eastern three-fifths, speak an Austroasiatic language that links the Mon-Khmer languages of southeast Asia with the Munda languages of

---

1 Thanks to a generous fellowship from the Fulbright Foundation, I was fortunate to be able to spend seven months in Northeast India during 1996–7 where I was affiliated with North-Eastern Hill University in Shillong in Meghalaya, and I made a shorter return visit in the spring of 1999. While there, I renewed an old interest in trying to unscramble the relationships among the Tibeto-Burman languages of the area. I have consulted a large proportion of the linguists, both Indian and foreign, who have had an interest in Northeastern Indian Tibeto-Burman languages, and I could not have written this chapter without the help of N.K. Achumi, David Bradley, Thangi Chhange, 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, and Jackson T-S. Sun and Graham Thurgood. All have been generous with their knowledge, their help, and their suggestions. Where the chapter remains inadequate it is because I have not been clever enough to follow their advice.





**MAP 11.1** NORTHEASTERN INDIA

central India. Except for the Khasis, the largest group of people in Meghalaya are the Tibeto-Burman Garos, who occupy the western two-fifths of the state.

Geographically, ethnically, and linguistically, Assam is different from the six hill states. Its major language is Assamese, an Indic language so close to Bengali that, much to the annoyance of the Assamese, the Bengalis often dismiss it as nothing but a corrupt form of Bengali. In the last century, a large number of Bengali settlers found their way to Assam, and these, together with the longer established Assamese, form the dominant population. Even in Assam, however, there are substantial Tibeto-Burman minorities. They live not only in the hilly southern wing of the state but in the Brahmaputra valley as well, and they must number well over a million in all. The minorities who live in the plains are presumed to be the remnants of the once dominant population. Most of the people who belong to these 'plains tribes' are able to speak Assamese, at least as a second language, and many of those who now speak only Assamese still trace their origins to a Tibeto-Burman tribal background. Some people who count themselves as Rabha, for example, now speak only Assamese, while others continue to learn the Tibeto-Burman Rabha language as their mother tongue.

The Tibeto-Burman languages of these northeastern states cannot really be counted. We still know too little about some regions to list all the languages, but even for better known areas it is often impossible to know whether two related speech forms should be considered separate languages, or dialects of a single language. Each of about twenty Kuki groups, for example,

most of them living in Manipur state, is usually said to have its own language, but these languages are very similar to one another, and some of them must be mutually intelligible. This makes it impossible to count the number of Kuki languages. Neighbours within some dialect areas can understand each other even though people living further apart cannot and, in these cases, it is futile to look for boundaries that would allow languages to be counted. Estimates of the number of languages spoken in Northeastern India run as high as two hundred (Downs 1994), but I suspect that this number can be reached only by counting a good many mutually intelligible dialects as separate languages. Nevertheless, even the most modest estimate would have to recognize at least fifty or sixty languages. The linguistic diversity is particularly great in the eastern and western extremes of Arunachal Pradesh, in the mountainous parts of Manipur and throughout Nagaland. There are fewer languages, each with more speakers, in Manipur's high central valley (where Meitei is dominant), in Mizoram (with Mizo [Lushai]), in Tripura (with Kokborok [Tripuri]), in western Meghalaya (with Garo), and in Assam (with, Rabha-Koch, Bodo, Mising [plains Miri], and Karbi [Mikir]), but smaller minority languages are found in all these states as well.

Sorting out these languages is made complicated by their confused and constantly changing nomenclature. The British began to contact speakers of the languages toward the end of the eighteenth century as their frontier gradually expanded northward and eastward from Bengal. The Assam valley was annexed to British India in 1824, but control reached into the hills more slowly. Even at the time of Indian independence in 1947 some parts of Nagaland and NEFA (the North-East Frontier Agency, now Arunachal Pradesh) were only feebly administered. The British would first learn about the tribes of the unadministered hinterland from people in areas already under their control, so the names that came to be used in English language publications were often the names used by the neighbours rather than by the people themselves. Since independence, more and more people have insisted upon using their own names or have even invented new names. The people formerly known as 'Lushai' are now called 'Mizos'. The Mikir have become Karbi, the Plains Miri, Mising. Some people have also realigned their ethnic affiliations. Recently, for example, some Rabhas and Koch, formerly regarded as separate 'tribes', have developed the slogan 'Rabha are Koch, Koch are Rabha', hoping to strengthen their political agenda by unity (Karlsson 1997). Such realignments mean that it is not simply the names of stable groups that change, but also the groups that need to be named.

In this chapter I will do my best to use the terms now in actual use in Northeastern India, but I will often add older names in square brackets in the hope of making my discussion easy to relate to the older literature. Where two names are both used, I give both, separated by a slash, starting with my guess for the one most likely to survive. It must be acknowledged that keeping up-to-date with the changing names of the northeast is a challenge that will not soon end, but courtesy demands that we do our best. Additional changes will surely be made. I have made no attempt to list all the alternative spellings that have been, and continue to be, used for tribal language names but have done my best to select the spellings most often used in Northeastern India today.

Terminology is made even more difficult by a confusion between ethnic and linguistic affiliation. To be sure, the people themselves often draw ethnic boundaries in the same places where we, and they, can recognize linguistic boundaries, but there are many exceptions. Some people in Northeastern India who cannot understand each other's languages describe themselves as belonging to the same 'tribe' (the term 'tribe' is not yet rejected in Northeastern India), and some people who claim to belong to different 'tribes' can speak to each other easily. Unfortunately, ethnic labels have too often been imagined to be equivalent to language labels, and the result can be total confusion.

Groups of languages have also been given inappropriate and misleading labels, and I find none more misleading than 'Kamarupan'. This term was first introduced by Matisoff (1991) simply as a convenient geographical catch all in which to toss these northeastern languages when their genetic subgrouping was unclear. This was a laudable admission of ignorance about an area that has lacked any convincing genetic classification but it invited misunderstanding. Others, who may have been influenced by the term 'Kamarupan', seem to have presumed that the Tibeto-Burman languages of Northeastern India really do form a genetic subgroup. Even the generally authoritative *International Encyclopedia of Linguistics* lists all the Tibeto-Burman languages of the region as 'Baric', a term never before used so widely (DeLancey 1991). There is not a shred of evidence for any special genetic relatedness of all the Northeastern Indian Tibeto-Burman languages, and ample evidence against such a grouping. It would be best to abandon 'Kamarupan' even if the only alternative is to say 'the Tibeto-Burman languages spoken in Northeastern India'.

Another term that has too often been used as if it designated a linguistic subgroup is 'Naga.' Today, the people known as 'Nagas' certainly recognize some common 'Naga' ethnicity, but this recognition may have come only after the British gave them the name 'Naga'. Most of the indigenous people of Nagaland, together with some ethnic groups in the bordering areas of Manipur, Assam, Arunachal Pradesh, and Myanmar are, by general consensus, now accepted as 'Nagas', but this term should not fool us into believing that they must have some linguistic unity. The languages spoken by 'Nagas' fall into at least two, and possibly several, completely distinct branches of Tibeto-Burman. I will continue to use scare quotes around 'Naga' to remind readers that it is not a linguistic label.

By the turn of the twentieth century, British officials and Western missionaries were beginning to produce grammars and word lists of the northeastern languages, and a few publications were sufficiently ambitious to be called 'dictionaries' (Lorrain and Savage 1898; Lorrain 1907; Mason 1905; see Huffman 1986 for additional references). These were immensely valuable to the local people, who were being introduced to literacy for the first time, but most fall short of a linguist's ideal. While most of these languages are tonal, for example, it would be difficult to discover that fact from most pre-independence publications.

The linguistic work of British officials and missionaries was interrupted by Indian independence in 1947, and for most of the years since then, political unrest has made it nearly impossible for foreigners to do scholarly work in Northeastern India. After a gap, however, publication resumed under new auspices. Linguists affiliated with the Central Institute of Indian Languages have done the most linguistically sophisticated work, but even the CIIL publications are uneven, and a few suffer from the isolation of its workers from international Tibeto-Burman scholarship. Two other energetic agencies have turned out a considerable body of publications. The Research Department of the Government of Arunachal Pradesh, and the Nagaland Bhasha Parishad (Nagaland Language Agency) have both produced wordlists, dictionaries and grammatical sketches of many of the languages of their respective states and sometimes of neighbouring states as well. Finally, there is a good deal of miscellaneous publication by speakers of the languages. There are schoolbooks and various other vernacular publications, and an irregular production of dictionaries. These appear in very small editions, often from local presses, and they often go out of print without the outside world ever hearing about them. Still, the material is steadily growing more abundant, 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. We have every reason to hope that the quality of the considerable volume of work being done on Northeastern Indian languages will gradually rise, and we must welcome it all as an important contribution to Tibeto-Burman studies.

## 2 TYPOLOGY

Few, if any, typological characteristics set the Tibeto-Burman languages of Northeastern India clearly apart from those spoken elsewhere. As in many Asian tonal languages, the syllable is a crucial phonological unit, and the inventories of syllable initial consonants are quite different and substantially larger than those of syllable finals. Bilabial, apical, and velar stops can usually occur at both the beginning and end of a syllable, but only initials allow contrasts in voicing or aspiration. Initial nasals seem always to include *m* and *n*, usually, and occasionally *ɲ*. At the end of a syllable, most languages have *m*, *n*, and *ŋ*, although few, notably Angami, lack final consonants entirely. As if in compensation, Angami has a particularly rich array of initials. Affricates and sibilants may occur in one or two positions initially and a *w* or *v*, an *r* or *l*, or both, occasionally a *y* or an *h* and one or two other consonants plus a few clusters are likely to round out the initials. A few languages along the eastern border allow such relatively exotic features as pre-consonantal or voiceless nasals. In addition to the stops and nasals, most languages allow a final (but not an initial) glottal stop, often transcribed by *h* in the now widely used Roman orthographies.

Vowel systems vary from a quite simple five-vowel pattern (e.g. Garo) to those with nine or ten simple vowels plus several diphthongs (e.g. Phom). High back unrounded vowels, often represented by *ü* in the Roman orthographies, are a characteristic feature of many of these languages.

Every eastern border language that I have ever heard has at least three tones, and most Bodo-Koch languages seem to have two. Garo lacks tonal contrast entirely, but in Garo the contrast between a glottal stop and its absence is clearly cognate with the tone contrast of other Bodo-Koch languages. The situation in Arunachal is not well known, although it has sometimes been said that tones become fewer as one moves westward through the state. However, Dhammai (Simon n.d.), Sherdukpen (Dondrup 1988), and Bugun (Dondrup 1990) in the far west are all reported to have at least some tone contrasts. The apparent rarity of tones in some areas may be due to the genuine rarity of people looking for them. Only very rarely are tones indicated in the available publications.

More often than not, syllable boundaries correspond to morpheme boundaries but enough morphemes (and syllables) can be combined in a single word to justify calling most of these languages 'agglutinative'. Noun compounding is ubiquitous and many languages have a number of noun and adjective 'prefixes', sometimes with clear semantic reference, sometimes much more generalized. Verbs can be particularly complex, with negative and interrogative particles and various sorts of adverbial and tense affixes, more often suffixes than prefixes. Sporadic examples of verb agreement can be found, but many languages lack verb agreement entirely. I have never encountered a Tibeto-Burman language in Northeastern India where the verb came anywhere except at the end of a sentence, and all the languages seem to have suffixed case markers supplemented by postpositions. Unambiguous case markers allow the order of noun phrases to be quite free. Many, though not all, of the languages have numeral classifiers, but the complexity of the classifier systems appear to be quite variable. Beyond such rather generalized statements, it is difficult to be specific about the characteristic features of these languages. For most languages, we remain ignorant about all but the most superficial aspects of the grammar.

## 3 CLASSIFICATION

The remainder of this chapter will be devoted to proposing a new classification. In the absence of reliable phonological analyses or good descriptions of morphology and syntax, the

comparativist rarely has anything to work with except for wordlists. We now have substantial lists from several dozen languages although reliability is variable and difficult to judge. Tones are rarely indicated, transcriptions are idiosyncratic, and misprints plentiful. Nevertheless, people with the skill and motivation to work with this material can learn a great deal. In particular, three important dissertations (Marrison 1967; French 1983; Sun 1993a) show how much can be learnt from the available material. These three dissertations have brought considerable order to masses of confusing data.

The first thing that a comparativist would like to know about these languages is their genetic subgrouping. Any classification offered today, has to be regarded as extremely tentative, but it is possible to place the languages more confidently than could be done with the materials available to Shafer (1955), whose classification, in the absence of anything better, continues to be cited. There have been attempts to do comparative phonology of these languages, but the ambitious ones (Shafer 1950; Weidert 1987) are very difficult to interpret, and those that are more modestly limited to a narrower range of languages (Burling 1959; French 1983; Sun 1993a) tell us little about the wider aspects of classification. At present, therefore, we have no choice but to rely primarily upon the lexicon for tentative judgements about which languages are genetically close and which are more distant. Languages which share the largest number of apparent cognates can be judged to be most closely related but when reliable comparative phonology is lacking, as it generally is in Northeastern India, even judging cognate status often requires an uncomfortable amount of guess work. In spite of these problems, some aspects of the classification now seem clear, and in the rest of this chapter I will describe the aspects in which we can have reasonable confidence, offer my best guesses where confidence is impossible, and try to suggest how much is reasonably reliable and how much is guess work.

I will begin with the largest reasonably well established subgroup of the northeast, a group that I have, in the past, called the 'Sal' languages (Burling 1983), but that I will here call more transparently though more clumsily, 'Bodo-Konyak-Jinghpaw'. It is convenient that this is the subgroup with which I am most familiar, and it is also convenient that, by occupying a central swath of the area, this group divides the rest of the northeast rather neatly into two geographical regions: the northern area, primarily Arunachal Pradesh, and the eastern hills that lie along the Myanmar border. This breaks the job down into three parts, and I will consider each in turn. I will use the term 'subgroup' and even number them, but I do not want to imply that all fourteen numbered subgroups have equivalent genetic status. The subgroups are not meant to be coordinate branches of a language family, but simply the largest groups that can now be reasonably proposed. Bodo-Konyak-Jinghpaw, for example, is a vastly broader and more diverse subgroup than Mizo-Kuki-Chin. The classification offered is deliberately conservative. I recognize subgroups only when the members appear to share innovations that carry them away from common Tibeto-Burman. Inevitably, this means that languages that are little known cannot be grouped as successfully as those that have been better studied. It is surely unlikely that Northeastern India harbours fourteen totally separate branches of Tibeto-Burman, and as we learn more about these languages, we should be able to merge some of the groups that I list into larger ones. Nevertheless, in our present state of ignorance, it seems more prudent to recognize too many groups than to toss languages together under arbitrary geographic labels that do more to hide than to clarify the relationships among the languages.

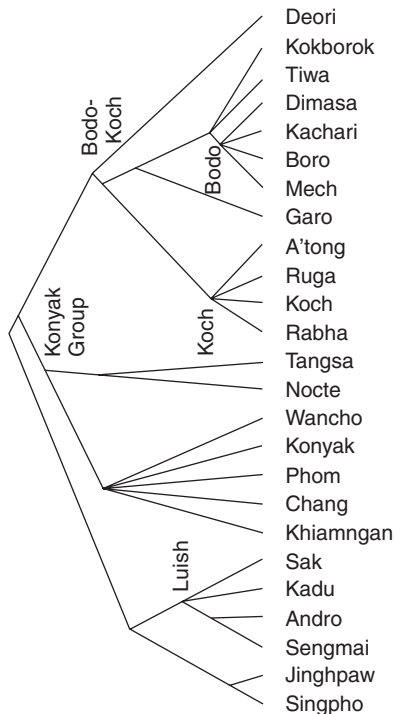
Numbers in parentheses following the names of languages give an estimate, in thousands, of the number of speakers. Unless specifically stated otherwise, the figures for languages that straddle an international border refer only to speakers in India. Most of the figures come from a publication of the India Missions Association (Anonymous 1997), and are estimates for the

year 1997. For a few languages I have used figures from Ethnologue (Grimes 1996). Figures lower than '11' are taken from the introductions to grammars of the languages. Where no figure is given it is because I have found no estimate, but these languages can be assumed to be spoken by quite small groups. Population estimates in Northeastern India are notoriously unreliable. Among other problems, it is often unclear exactly which subgroups are included within a 'language'. The figures should be taken as informed guesses, but they seem, as far as I can judge, quite plausible.

## 4 THE CENTRAL AREA

### i Bodo-Konyak-Jinghpaw

This is, by far, the most extensive recognizable subgroup of any that is based primarily in Northeastern India, and it constitutes a major subgroup of Tibeto-Burman. Benedict went so far as to suggest that the most basic cleavage within Tibeto-Burman divides Bodo-Konyak-Jinghpaw (represented in his analysis only by Garo and Jinghpaw) from everything else (Benedict 1976:177). A plausible stambaum for these languages, indicating my best judgement, is shown in Figure 11.1. Map 11.2 shows their approximate location.



**FIGURE 11.1** RELATIONSHIPS AMONG THE BODO-KONYAK-JINGHPAW LANGUAGES



MAP 11.2 BODO-KONYAK-JINGHPAW LANGUAGES

Bodo-Konyak-Jinghpaw has three main branches.<sup>2</sup> One of these, Bodo-Koch [Bodo, Bodo-Garo, Barish], is well established, having been recognized at least since the time of the *Linguistic Survey of India* (Grierson Vol. III, Part II, 1903). The Bodo-Koch branch, in turn, divides into four parts: Bodo, Garo, Koch, and Deori. The data given in Burling (1959) suggests that Garo is closer to Bodo, than to Koch, which is why I prefer to call the larger group 'Bodo-Koch', but most published classifications imply that Garo is closer to Koch. This is probably because of confusion over A'tong and Ruga. Speakers of these two languages live in the southern part of Garo Hills and are accepted by everyone as being ethnically Garo. Their languages, however, are by no means mutually intelligible with Garo. (Ruga is now moribund. I doubt if anyone younger than fifty would claim to speak the language. A'tong, apparently, is a bit healthier.) Since the speakers of all these languages are ethnically Garos, there has been an unfortunate tendency to consider A'tong and Ruga to be dialects of Garo. Benedict formalized the confusion by calling A'tong and Ruga 'Garo A' while he called Garo proper 'Garo B' (1972: 6 fn. 20). It has long been clear that the languages of the A'tong and Ruga are close to Rabha, which is spoken to the north and northwest of the Garo area, and Benedict duly included Rabha among the 'Garo A' languages, even though no one considers the Rabhas to be Garos. To call A'tong and Ruga 'Garo A' is to recognize common *ethnicity*; to group Rabha with A'tong and Ruga is to recognize *linguistic* similarity. To conclude that Rabha is a kind of Garo is like calling Welsh a form of French because it is so much like Breton, whose speakers are French. The Rabha do not count as Garos either by ethnicity or by language, any more than the Welsh count as French.

<sup>2</sup> I do not use 'branch', as Shafer did, to imply a specific level in a taxonomic hierarchy, but only as an informal way to label one part of a larger group.

Rabha (31), A'tong, Ruga, and Koch (23) (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 mislabeled 'Garo of Jalpaiguri' (Grierson 1903), who call themselves (more appropriately) 'Koch' or 'Rabha' (Karlssoon 1997). Locally, this group of languages is known as 'Koch', and that is the best term for linguists to use as well.

The Bodo branch includes several languages (or dialects) spoken both in the valley and in the hilly areas to the south. These have been known by such names as Boro (or Bodo) (39), Mech, Dimasa (106), Kachari (59), and Hill Kachari. At least some of these are probably different enough to count as separate languages, although some of their speakers now prefer to identify themselves as 'Bodo' rather than 'Mech', 'Dimasa', or 'Kachari'. Tiwa [Lalung] (23), also in Assam, and Kokborok [Tripuri] (691) in Tripura state to the south appear to be only marginally more distant from Boro, Dimasa, and Kachari, than the latter are from each other.

The third branch of Bodo-Koch consists of a single language, Garo (575). Though spoken over a wide area in Western Meghalaya and neighbouring areas of Assam and Bangladesh, all dialects of Garo proper (i.e. not including Rabha and A'tong) are reasonably mutually intelligible. Fourth, and last, is a little known language known as 'Deori' or 'Chutia' (19) that is spoken further east in Assam. Deori certainly belongs with the Bodo-Koch languages, although the exact point at which it joins the others is less certain. In some respects, it appears to be rather deviant. When speakers of these Bodo-Koch languages encounter one another, they easily recognize the similarity of their speech, but most are well beyond mutual intelligibility. My very loose impression is that they differ from each other about as much as the Romance languages.

The languages of the Konyak group, named for its largest member, are less well known than the Bodo-Koch languages, and they appear to be considerably more heterogeneous, although the unity of the group seems clear. Their similarities with the Bodo-Koch languages have often been noted (Shafer 1955; Benedict 1972; Burling 1983), most persuasively by French (1983). Six languages have generally been recognized as belonging to the Konyak group: Tangsa (17), Nocte [Namsangia] (33), and Wancho [Banpara] (45) in southeasternmost Arunachal Pradesh; Konyak [Tableng] (105), Phom [Chingmengnu, Tamlu] (34), and Chang (31) in the adjacent region of Northeastern Nagaland. French offers convincing evidence that Tangsa and Nocte form a subgroup within the larger Konyak group. More cautiously and a bit less convincingly he suggests that Wancho and Chang are particularly close to each other, as are Konyak and Phom. To these six, Khiamngan (25) can now be added as a somewhat divergent seventh member (Kumar, Meya, and Jon, n.d.). Tangsa, and perhaps some other members of the Konyak group, is sufficiently heterogeneous that some so-called 'dialects' lack mutual intelligibility.

The special similarity of Jinghpaw (652, mostly in Myanmar and Yunnan) to Bodo-Koch and the Konyak group has been noted by Benedict (1972, 1976) and Burling (1971, 1983) and, again, considered most carefully by French (1983). Jinghpaw has figured much more prominently in Tibeto-Burman studies than have the poorly known languages of the Konyak group, and it has been classified in a number of ways. Often it has been placed with Lolo-Burmese because it shares considerable vocabulary with the so-called 'northern Burmish' languages (Atsi, Maru, Hpon, etc.) and even with Burmese. It should be remembered, however, that Jinghpaw is used as a *lingua franca* in the linguistically heterogeneous Kachin State of northern Myanmar. This is an area of massive multilingualism, and Jinghpaw has had ample opportunity to borrow from the northern Burmish languages. (To endless confusion, people still insist on calling both Jinghpaw and the northern Burmish languages 'Kachin', but 'Kachin' is an ethnic term and its use as a linguistic label should be avoided.) The northern Burmish languages are the



closest linguistic relatives of Burmese but they preserve more syllable final consonants than does Burmese. A word could have been borrowed last week from a northern Burmish language into Jinghpaw, and still have a rather archaic look to someone who knows Burmese better than the other languages. This has given a spurious appearance of an ancient connection between Burmese and Jinghpaw. Borrowing can easily account for the similarities of Jinghpaw to Burmese and to the northern Burmish languages, while its similarities to Bodo-Koch and the Konyak group have to be older (Benedict 1976:178). If Jinghpaw seems to be geographically remote from Bodo-Koch and the Konyak group, it should be remembered that a form of Jinghpaw known as 'Singpho' is spoken in Arunachal Pradesh, just north of Tangsa, the northernmost language of the Konyak group. There is no geographic break at all. French's analysis does suggest that Jinghpaw is a bit less similar to Bodo-Koch and the Konyak group than these two are to each other, and that is how I have drawn them in Figure 11.1.

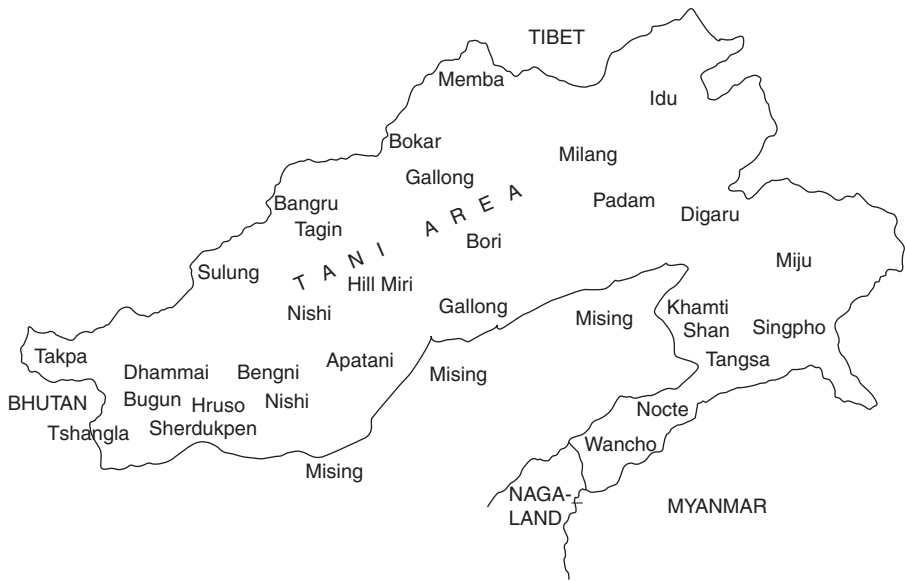
Finally, we should take note of a scattered group of minor languages known as 'Luish'. Grierson, long ago (1921), recognized the relationship among these languages, and Benedict (1972: 5) pointed out their similarity to Jinghpaw. The least poorly known of these poorly known languages, are Sak of the Chittagong Hill Tracts in Bangladesh (Bernot 1966), and Kadu of upper Burma (Brown 1920). Within Northeastern India, Luish was once represented by Andro and Sengmai of Manipur, which are known only from an ancient wordlist (McCulloch 1859). Andro and Sengmai villages still exist, and I was able to visit them both in 1999. Villagers prize the tradition that they once had their own languages but the only remnant that exists, apart from cherished photocopies of McCulloch's wordlist, are some short chants in Andro, whose meaning no one any longer knows, and a dozen or so words dimly recalled by one man in Sengmai, possibly learnt from McCulloch's list.

Bernot (1966) assembled the available evidence on the Luish languages and, modest though the data on Andro and Sengmai are, it seems clear that they are rather closely related to Sak and Kadu. Bernot's own data on Sak are the best that is available on any of these languages, and its special similarities to Jinghpaw are obvious. Its similarities to the Bodo-Koch languages are less strong, but still clear. The evidence remains thinner than we would like, but it is hard to doubt that the Luish languages should be placed on the Jinghpaw branch of Bodo-Konyak-Jinghpaw, and that is how I have drawn them on the tree.

The locations of the Bodo-Konyak-Jinghpaw languages are shown on Map 11.2. Except for Karbi and some settlements of Rengma, both of which are Tibeto-Burman, and Khasi, which is Austro-Asiatic, the gaps between the Bodo-Koch languages are filled by Assamese and Bengali. The usual and reasonable presumption is that the Bodo-Koch languages once had a more continuous distribution than they do now, but that they have been separated by the gradual spread of the Indic languages. It seems probable that earlier forms of Bodo-Koch were, at one time, the predominant languages of the Assam valley and perhaps of some parts of northern Bengal as well.

## 5 THE NORTHERN AREA

Almost all of the indigenous people of Arunachal Pradesh, (formerly NEFA, the North East Frontier Agency) are Tibeto-Burman speakers, and a single group of closely related languages, Tani, covers the central half or two thirds of the state. The languages of the western and eastern fringes of Arunachal are far more heterogeneous, less well known, and more problematic (see Map 11.3). I can do no more here than list them by whatever tentative groups other linguists have suggested, and lay out what still needs to be learnt. For all of Arunachal, I rely especially on the work of Sun (1993a,b). I begin with the westernmost groups and proceed eastward.



**MAP 11.3** LANGUAGES OF ARUNACHAL PRADESH AND NEIGHBOURING ASSAM

## ii Tshangla-Takpa

Tshangla is spoken in western Arunachal, near the border of Bhutan. The same language, sometimes known as 'Shar chopkha', is the predominant language of the neighbouring region of Southeastern Bhutan. In China the language is known as 'Cangluo Menba' or 'Motuo Menba', and in Arunachal it is often referred to as 'Central Monpa' 'Monpa', however, is derived from the Tibetan word for 'non-Tibetan' and it has been used in so many ways as to risk losing clear reference. 'Tshangla' is at less risk of ambiguity (Michailovsky 1994). 'Memba', sometimes used for a language spoken in Northernmost Arunachal, is probably an alternative spelling of the same word.

Takpa [Dwags] is spoken in the very western tip of Arunachal, where it is often referred to as 'Northern Monpa'. It resembles the Bumtang languages of central Bhutan. The language is also spoken in Tibet, where it is known as 'Cuona Menba'. (Michailovsky and Mazaudon 1994; van Driem 1992: 11).

Linguists who have worked in Bhutan suggest that Bumtang, and thus Takpa, are closer to Tibetan than is Tshangla (van Driem 1992; Michailovsky and Mazaudon 1994, Michailovsky personal communication). However, a comparison of vocabulary indicates that Tshangla and Takpa resemble each other more closely than either resembles Tibetan (Anonymous 1991). Tentatively, therefore, I group Tshangla and Takpa together. I can detect no special resemblance between either of these languages and the other languages of Arunachal, except for a few scattered similarities with Sherdukpen. The Sherdukpen like the Tshangla and Takpa are Buddhists and they must all have come under parallel cultural influence from Tibet. This could easily explain the few similarities to Sherdukpen. Otherwise Tshangla and Takpa seem quite different from the languages to their east.

### iii Sherdukpen, Bugun/Khoa, Sulung, Lishpa

Sun (1993a: 12 fn. 18) suggests that Sherdukpen (2) and Bugun (8), both of which are found just to the east of Tshangla, together with a little known language called 'Lishpa', might form a group of their own. To these he more tentatively adds Sulung (4), which is spoken well to the north, along the Tibetan border. In Arunachal, the Sulung people, who call themselves 'Puroit', are regarded as the original inhabitants. They have been severely oppressed by others who are supposed to have migrated in from elsewhere.

Relatively recent vocabularies of Sherdukpen (Dondrup 1988), Bugun (Dondrup 1990), and Sulung (Tayeng 1990; Anonymous 1991) allow a more systematic comparison among these three languages than would have been possible earlier, and there can now be little doubt that they belong together. Their vocabulary is remarkably divergent from other Tibeto-Burman languages, however, and a few of us have even wondered whether Sulung belongs to Tibeto-Burman at all. However, each of the three languages has a number of apparent Tibeto-Burman cognates so, if the three do form a group, their Tibeto-Burman affiliation seems secure. Simon (1976) offers fragmentary data from Lishpa and from a language that he calls 'But-pa', both of which, he says, are found to the west of the Bugun. His data suggest that these also belong with this group, and his handful of But-pa examples look so much like Sherdukpen that it might even count as a dialect.

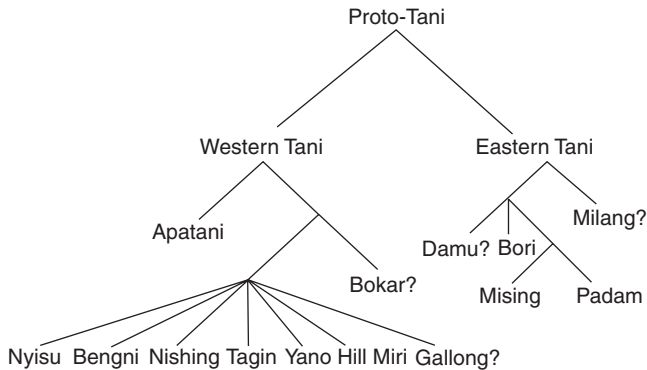
### iv Hrusish

Hruso [Aka] (3) and Dhammai/Miji (4) are spoken just east and north of Bugun and Sherdukpen. Shafer (1947) gave the name 'Hruso A' to the language of the people who call themselves 'Dhammai', while he called Hruso proper 'Hruso B'. In fact, these are quite different languages, as is evident from Simon's vocabularies of Dhammai (1979) and Hruso (1970), not dialects as seems to be implied by calling them both 'Hruso'. Nevertheless, they show enough similarities to let them be grouped together. In addition to these two, Sun has now added a third language to the Hrusish group, the little known Bangru/Levai, which is spoken further north along the Tibetan Border (1993b: 159). Sun describes it as very similar to Dhammai.

### v Tani [Mirish, Misingish, Abor-Miri-Dafla]

The languages of central Arunachal Pradesh belong to a long recognized but variously named group. 'Abor', 'Miri', and 'Dafla' are no longer acceptable to the speakers and no longer used in Northeast India, so these names, as well as the derivative 'Mirish', should be avoided by linguists. Sun (1993a) calls this group 'Tani', a term that is supported by local usage, and I will adopt that term here. In addition to the languages of central Arunachal, Tani includes Mising [Plains Miri], which is spoken by a Hinduized population that lives in the plains of upper Assam. Sun (1993a: 8) guesses the total population of Tani speakers as 600,000, about half them Misings.

The membership of the Tani group is clear, but the internal relationships among its languages have been muddled by an ethnic classification that is only vaguely related to the language classification. The western Tani, and their languages, are often called 'Nishi' [formerly Dafla], although Apa Tani apparently stands apart, while the eastern Tani are known as 'Adi' [formerly Abor]. Unfortunately, 'Nishi' has also occasionally been used for languages that are not Tani at all and 'Adi' has been used for quite diverse Tani Languages (Sun 1993a: 9–10). Although the Mising live in the plains of Assam rather than in the hills of Arunachal,



**FIGURE 11.2** THE TANI LANGUAGES (AFTER J.T.-S. SUN 1993)

their language does not seem to stand apart from the eastern Tani languages that are spoken in the hills. Sun, who has sorted through the confusion says, ‘The Tani area, barring a few aberrant linguistic islands, seems to consist of a continuum of mutually intelligible local varieties shading gradually into one another’ (1993a: 8). Even if neighbours can generally understand each other, those living further apart cannot. I reproduce Sun’s Tani stambaum as Figure 11.2. The languages on the left side of this tree include those that are sometimes called ‘Nishi’, while those on the right are more often called ‘Adi’.

#### vi Idu-Digaru

The languages of Northeastern Arunachal give us another example of confusion between ethnic and linguistic classification. Idu [Chulikatta], Digaru/Taraon, and Miju/Kaman have always been known as the ‘Mishmi’ languages and speakers of all three seem to acknowledge their ethnic affinity. Anonymous (1997) gives 33 (thousand) as the combined population of all Mishmis, which strikes me as low. These three languages are very different from each other, but it is generally accepted that Idu and Digaru, the more northern languages, share more with each other than either shares with Miju (Sun 1993b: 160). In a lexicostatistic comparison, Sun finds that the Tani languages share more cognates with Digaru (which he calls ‘Taraon’) than with the languages of any other non-Tani group. Sun makes it clear that Digaru and the related Idu are not, themselves, Tani languages and he presents the similarity between Tani and Idu-Digaru as no more than suggestive. In view of his caution, I hesitate to place Idu-Taraon into a larger group that includes Tani, but the resemblance should be noted.

#### vii Miju

Miju/Kaman is the southeasternmost of the three eastern Arunachal tribes known as ‘Mishmi’. Their language is very different from Idu and Taraon, however, and they must be placed in a group of their own. Unlike the other ‘Mishmi’ languages, Miju shows no more similarity with Tani than does any randomly selected Tibeto-Burman language.

Finally, before leaving Arunachal, brief mention should be made of Shan. A few Shan speakers, mostly of the variety known as ‘Khamti’, still live along the upper Brahmaputra in Eastern Arunachal. They are the westernmost surviving representatives of the Shans, and they speak the only indigenous language of Arunachal that is not Tibeto-Burman. Beyond the

Miju and the Khamti Shan are the Singpho and the members of the Konyak group that have already been discussed.

## 6 THE EASTERN BORDER AREA

The states along the Myanmar border comprise the linguistically most heterogeneous region of Northeastern India. In spite of valiant efforts by Marrison (1967) and French (1983), the languages of the people called 'Naga' are particularly confusing, and there may be no Tibeto-Burman region anywhere where the languages are so diverse. Some linguists have suggested that the original centre from which a language family spread can be recognized by the present diversity of its languages. If this were a valid generalization, then Nagaland and the regions that border it would make a good candidate for the Sino-Tibetan homeland. I have never heard such a claim, however, and many stories are reported from the 'Nagas' themselves about their migrations from other places.

Only two things are clear about the eastern area. First, as we have already seen, the Konyak group in the northeast is affiliated with Bodo-Koch and with Jinghpaw. Second, Mizo, Kuki, and Chin in the south, form a closely related group. Between these geographical extremes we find massive heterogeneity and uncertainty.

Except for the languages of the Konyak group, Shafer (1955) classified all the 'Naga' languages with those that I call 'Mizo-Kuki-Chin', placing them all in his 'Kukish' Section. Ever since, Tibeto-Burmanists have tended to presume that 'Naga' and Mizo-Kuki-Chin are joined into some sort of coherent branch of Tibeto-Burman. This may, in fact, be correct, but Shafer's discussion cannot be taken as conclusive. His most important paper on the subject has the odd title 'The Naga branches of Kukish' (1950). Since the languages called 'Naga' are vastly more diverse than the Mizo-Kuki-Chin languages, this is a bit like speaking of German, Danish, and Dutch as the 'continental branches of English'. Shafer gives large numbers of apparent cognates that relate Kuki to various groups of 'Naga' languages. Unfortunately, he offers no evidence that the similarities he finds between the 'Naga' groups and Kuki are more extensive than those between these groups and any other Tibeto-Burman language. This, and an often disorderly presentation, leave his claims very difficult to evaluate.

Nevertheless, in spite of their enormous diversity, several attempts to classify the 'Naga' languages have reached similar, though by no means identical, conclusions. Leaving aside classifications that simply parrot someone else's opinion, three deserve to be taken seriously: The Linguistic Survey of India (Vol. III, Part III, Grierson 1904), Shafer (1955), and Marrison (1967).

Geoffrey Marrison 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 genetic, and it is 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 my own judgement.

The classification that I offer depends more exclusively on lexical comparison than does Marrison's. This is no doubt the reason that my classification comes closer to Shafer's, than to Marrison's. Lexical comparison, however, is peculiarly difficult for the two dozen or more 'Naga' languages, for they seem almost perversely diverse, as if the people had spent several centuries deliberately manipulating their languages so as to set themselves apart from their neighbours. A word that is common in the north may be missing in the central part of Nagaland, only to turn up again in the south. Another will be splattered in irregular patches across the map, and still another confined to a few neighbouring languages that otherwise appear to have little in common. In the face of this extraordinary diversity, any attempt at classification

must be extremely tentative. Another complication is that most of the languages spoken near the Myanmar border are also spoken on the other side. Beyond these are still other people who identify themselves as ‘Naga’ but whose languages are not found in India at all. A full understanding of the ‘Naga’ languages will not be possible until these languages of Myanmar are taken into account.

I will be cautious, grouping together only those languages that seem clearly similar. In Figure 11.3, my judgement is compared to the classification of the Linguistic Survey of India, Shafer, and Marrison. Both the LSI and Shafer name their subgroups by compass terms, but the names they use differ so much from each other that they invite confusion. I will, therefore, abandon the compass and name each group by its most prominent language – The ‘Ao group’, the ‘Zeme group’ and so on. I will list the groups as if each is independent of the others, but I will briefly consider their possible interrelationships at the end. My judgement is summarized by the tree of Figure 11.4, where the solid lines at the right of the figure show relationships that seem most reliable. Broken lines represent probable relationships, and the dotted lines are speculative. The relative heights of the branching nodes offer my very rough guess about the degree of difference between the languages.

Burling (2000)		LSI (1903, 1904)		Shafer (1950)		Marrison (1967)	
Ao group	Yacham-Tengsa	Central	Tengsa	Northern Naga branch	Tengsa, Yatsam	B-1	Yacham-Tengsa
	Ao-Chungli		Aō-Chungli		Tsungli	B-1	Ao-Chungli
	Ao-Mongsen		Aō-Mongsen		Mongsen	B-1	Ao-Mongsen
	Sangtam		Thukumi		Thukumi	B-1	Sangtam
	Yimchungrū		Yachumi		Yatsumi	B-2	Yimchungrū
	Lotha		Lhota		Hlota	B-2	Lotha
Angami-Pochuri group	Rengma N.	Western		Eastern branch		B-2	Ntenyi
	Pochuri					B-2	Meluri
	Rengma		Rengmā/unzā		Rengma	C-2	Rengma
	Simi		Simi/Semā		Simi/Sema	C-1	Sema
	Angami		Angāmi/Tengimā		Angami	C-1	Angami
	Chokri				Tsākrima	C-1	Chokri
	Kheza		Kezhama		Kezama	C-1	Khezama
	Mao		NK Sopvomā/Māo		Sopvoma	C-1	Mao
Zeme group	Nruanghmei	Naga-Bodo	Kabui	Western branch	Kabui	C-2	Nruanghmei
	Puiron					C-2	Puiron
	Khoirao		Khoirao		Khoirao	C-2	Khoirao
	Zeme	Emp̄o/Kachchā	Empeo		C-2	Zeme	
	Mzieme				C-2	Mzieme	
	Liangmai	Kwoireng	Kwoireng		C-2	Liangmai	
	Maram	Marām	Maram		C-2	Maram	
T'khul group	Tangkhum	Naga-Kuki	Luhūpā	Luhupa branch	Tangkhum	B-3	Tangkhum
	Maring				Maring	B-3	Maring
	Karbi	NB	Mikir		Mikir Branch		
	Meithei		Meithei/Manipuri		Meithei Branch		
	Mizo-Kuki-Chin		Mizo-Kuki-Chin		Kuki Branches		

FIGURE 11.3 CLASSIFICATIONS OF THE EASTERN BORDER LANGUAGES

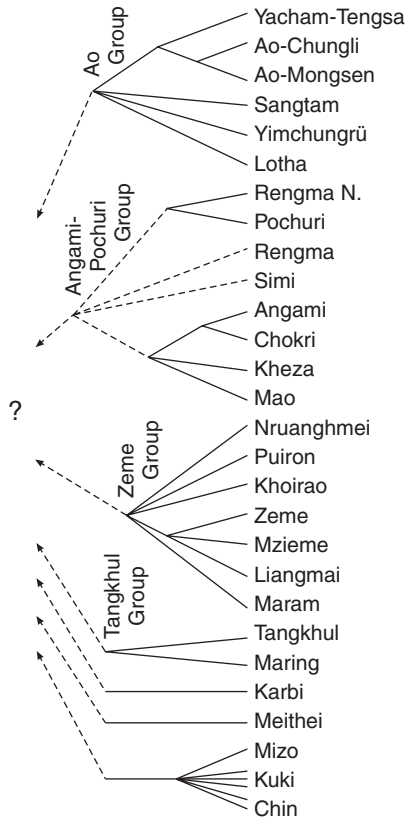


FIGURE 11.4 RELATIONSHIPS AMONG THE LANGUAGES OF THE EASTERN BORDER

Mizo-Kuki-Chin is listed as one group among many. Its internal diversity is probably less than that of the Ao group and certainly less than that of the Angami-Pochuri Group, and it shows roughly the same degree of similarity with the other groups as they show with each other. Continuing to move clockwise around the northeast, I will now proceed from northeast to southwest. The locations of the languages are shown on Map 11.4. Their locations and the boundaries between the groups are relatively accurate for Nagaland; Manipur has required much more guess work.

### viii Ao Group

These languages, 'Central' for the LSI and 'Northern' for Shafer, are spoken to the south and west of the Konyak group. Ao (141), the group's most important language, has at least two major dialects that border on mutual unintelligibility: Chungli and Mongsen. Yacham and Tengsa in the north are similar to Chungli and Mongsen but even more similar to each other. Like Shafer and the LSI, I group Lotha (80), Sangtam [Tukumi] (39), and Yimchungrü [Yachumi] (37) with Ao. As can be seen in Figure 11.3, Marrison separates Lotha, and Yimchungrü from the others, but he also points out, correctly I believe, that Sangtam, and



MAP 11.4 LANGUAGES OF THE EASTERN BORDER



Yimchungrü are closely allied, implying that his categories such as 'B-1' and 'B-2' should not be understood as genetic branches.

A brief mention of Lepcha (Rong), which is spoken in far off Sikkim, may be justified at this point. Strangely, Shafer (1955: 106) insisted that Lepcha belongs to the Ao group, and that it is especially like Tengsa, an extraordinarily precise assignment. The only explanation that I can imagine for this assignment is that Shafer classified primarily on the basis of the words for the numbers. Most areas of the Lepcha and Tengsa vocabulary show no special similarity, but they do have rather similar numerals. Numerals much like these all found in many other languages as well, but not usually in the same combination. The parallel selection of numerals in Lepcha and Tengsa does give the languages a superficial resemblance, but the similarity evaporates once the rest of the vocabulary is considered. Forrest (1962), Bodman (1988), and Sun (1993a: 373–9) have mulled over the place of Lepcha within Tibeto-Burman without demonstrating any special similarity to any Northeastern Indian group, let alone a specific similarity to Tengsa. Shafer's peculiar assignment of Lepcha should, I believe, be quietly forgotten. Since Lepcha lies outside of the geographical range of this chapter I will say no more about it.

### **ix Angami-Pochuri group**

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. First, Chokri (20) is so close to Angami (109) that it might even count as a dialect. Kheza (23) to the east and Mao [Sopvoma] (81), to the south in Manipur stand just a bit further from Angami, but Mao is still close enough that it is sometimes referred to as 'southern Angami'. The second nucleus includes Pochuri [southern Sangtam, eastern Rengma] (13), spoken in the far southeast corner of Nagaland, along with Marrison's Ntenyi (13) and Meluri. The earlier classifications did not include Pochuri, but the small dictionary by Kumar and Pochuri (1972) shows it to be so close to Marrison's Meluri as to be, at most, a slightly different dialect. Marrison's Ntenyi, also called 'Rengma-North', is spoken considerably further west than Pochuri, but it is not much more distinct. Finally, Simi/Sema (132) and Rengma proper (21) (i.e. the Rengma that is neither Pochuri/Meluri/Eastern Rengma nor Ntenyi/Rengma-N) stand somewhat apart, both from these two nuclei and from each other. Marrison classified most languages of the Angami-Pochuri group as C-1 but he placed Rengma with the members of the Zeme group as C-2 and Ntenyi and Meluri with the Ao group as B-2. These assignments can be justified by his typological criteria but the languages are closer lexically to Angami and Pochuri.

### **x Zeme group**

The languages of the Zeme group are even more closely unified than those of the Ao group and some may be close enough to count as dialects. The unity of this group has been recognized at least since Shafer called them the 'Western' Branch of 'Naga' (1955). Zeme [Empeo, Kachcha] (15), Mzieme (29), and Liangmai [Kwoireng] (18) are particularly close. Nruanghmei [Rongmei, Kabui] (59), Puiron, Khoirao (20) and Maram (11) are a bit more divergent. Some speakers of these languages recognize their unity by calling themselves Zeliang or Zeliangrong, names constructed from the first syllables of Zeme, Liangmai and Rongmei, the latter a variant of Nruanghmei. Shafer did not include all of these languages in his classification and he called several by older names, but I believe he grouped those for which he had data correctly. They are all 'C-2' for Marrison.

The LSI classifies these languages very differently from all the later classifications. It divides them into two categories, 'Naga-Bodo' and 'Naga-Kuki' but these appear simply to be geographical designations. The LSI's 'Naga-Bodo' languages included not only the members of the Zeme group that are known today as Zeme, Nruanghmei, and Khoirao, but also Karbi. These are all spoken further west than most 'Naga' languages. They are thus closer geographically to the languages of the Bodo-Koch group, and this seems to be the basis of the name 'Naga-Bodo'. The LSI's 'Naga-Kuki' languages included Maram and Liangmai of my Zeme group, and the very different languages of the Tangkhul group, again mostly called by older names. These languages are all spoken in the southwestern part of the 'Naga' region, close to the Kukis, hence 'Naga-Kuki'. Neither Naga-Bodo nor Naga-Kuki forms any sort of linguistically unified group, and it would be best to forget the LSI's classification. Shafer, I believe, got them right.

### **xi Tangkhul group**

Tangkhul (110) and Maring (16), both spoken in Northeastern Manipur State, form the final group of languages spoken by people who count as 'Nagas'. Tangkhul is said to have several dialects so different from one another that they might better be called separate languages. Maring seems more similar to Tangkhul than to any of the 'Naga' languages spoken further north. All earlier authorities agree on grouping Tangkhul and Maring together, though the LSI unjustifiably adds Maram and Liangmai [Kwoireng].

### **xii Karbi [Mikir]**

Karbi (341), also known as 'Arleng', has generally been grouped with the 'Naga' and Mizo-Kuki-Chin languages although it is spoken considerably further west than any of them. The homeland of the Karbi is a hilly region, known as the 'Karbi Anlong' (formerly 'Mikir Hills') that remains a part of the state of Assam. Karbi shows some resemblance to Mizo and Kuki but not enough to count as a Mizo-Kuki-Chin language, and it does not fit clearly with any particular group of eastern border languages. It seems to differ from the various groups of 'Naga' and Kuki languages about as much as they differ from each other. Like Shafer, I place it in a category by itself.

### **xiii Meitei**

The state of Manipur differs from the other 'hill' states of the northeast in having a large central plateau. This has 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 in the present hill states. The Manipuris (1240) who populate this plateau speak a Tibeto-Burman language, but they are Hindus, and several centuries ago they developed an Indic based script for their language. They are the only Northeastern Tibeto-Burman speaking people with a literate tradition that predates the colonial period. Their language, known as 'Meitei' or 'Meiteilon' (or, often, as 'Manipuri') shows some lexical resemblances to Kuki and some to Tangkhul. These could be the result of borrowing from the language of the politically and culturally dominant Meitei, and the resemblances are not great enough to make the assignment of Meitei to one of the groups obvious. None of the earlier classifications include Meitei 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 Meitei by itself.

#### **xiv Mizo-Kuki-Chin (Kukish)**

Mizo [Lushai] (529) is the majority language of Mizoram. The Chin languages are spoken in the adjacent Chin State of Myanmar. The Kuki languages (or dialects) with a population of at least 200 (thousand) in all, perhaps considerably more, are scattered from Mizoram northward through Manipur and beyond. A score or more Kuki languages are traditionally recognized. The largest number are in Manipur, but Kuki speakers are found in all the surrounding states as well. Their languages are all quite similar to one another and some of them must be mutually intelligible, but it is not at all clear which Mizo and Chin, are also closely related to Kuki and, indeed, the only reason why the languages in Myanmar are called 'Chin' while those in Northeastern India are called 'Kuki' is that the British approached them from opposite directions and learnt the names from different neighbours. Some Mizos, Kukis, and Chin seek to recognize their unity by uniting under the common name 'Zo' but the name has not gained full acceptance. The following are only a few of the names by which groups of the Kuki speakers in Northeastern India have been known: Aimol, Anal (15), Chiru, Chothe, Gangte (11), Hmar (5), Kom (14), Lakher/Mara (22), Paite (45), Pawi (16), Ralte, Thado (79), Vaibe (21).

That the Mizo-Kuki-Chin dialects and languages form a branch within Tibeto-Burman is abundantly clear. Indeed, they are far more unified than the languages of, for example, the Angami-Pochuri group, and they may be no more diverse than Tangkhul, which is usually described as a single 'language', even though it has several mutually unintelligible forms. In view of the manifest similarities among the Mizo-Kuki-Chin languages, it was simply bizarre for Shafer to divide them into six distinct 'Branches'. This is more Branches than he assigned to the vastly more diverse 'Naga' languages. In spite of Shafer's great proliferation of Branches, Units, and 'Languages', the subgrouping within the Mizo-Kuki-Chin group remains murky.

A new complication has recently been added to the classification of these languages. A few groups, whose language a linguist would, without hesitation, classify as 'Kuki' have declared themselves to be 'Nagas'. Everyone agrees that Nagas and Kukis are sharply distinct ethnically. Indeed, they have been killing each other from time to time. Nevertheless, exactly which groups are ethnic Nagas and which are ethnic Kukis can be disputed. Unless the people actually shift languages, which of course, they sometimes do, a shift in ethnic and political affiliation should not change a group's language classification, but when people talk in new ways about ethnic affiliation, they often talk in new (and confused) ways about language affiliation as well.

The languages of the Ao, Angami-Pochuri, Zeme, Tangkhul, and Mizo-Kuki-Chin groups, along with Meitei and Karbi have often been considered to form a genetic subgroup of Tibeto-Burman, more closely related to each other than to other languages. This grouping is not implausible, but it is far from obvious. The Bodo-Konyak-Jinghpaw languages show a good many lexical innovations that set them apart from other Tibeto-Burman languages (Burling 1983). It is more difficult to find lexical items, let alone morphological or syntactic features, that appear to be common innovations of the entire eastern border area. Nevertheless, adjacent languages and adjacent groups frequently share more lexical items than more distant languages. Thus, for example, it is easy to point to lexical items shared by the Ao and Angami groups or by the Angami and Zeme groups or even by adjacent languages of the Ao and Konyak groups. One can get the impression that Nagaland, Manipur, and Mizoram, along with the neighbouring strip of Myanmar, constitute something like a large dialect area, in which

adjacent languages share some characteristics but where distant languages are no more alike than any two randomly chosen Tibeto-Burman languages. The languages are far more diverse than we expect to find in a dialect area, however, for the region is crisscrossed by sharp language boundaries.

Much more work is needed before we will have a definitive classification of the languages of the eastern border. The similarities among these languages are tantalizing but difficult to pin down, and until more solid work is done, it seems best to withhold judgement. The individual groups seem sound, but how they relate to each other and whether they fall into one or several branches of Tibeto-Burman awaits further data and further analysis.

## 7 CONCLUSIONS

Chinese troops occupied a portion of Arunachal Pradesh in 1962 and, briefly, they seemed to threaten the entire northeast. In 1971, Indian troops were deployed in massive numbers in Northeastern India, waiting to give help to the independence movement that soon converted East Pakistan into Bangladesh. Between the mid-1970s and the mid-1990s violent separatist movements, sometimes degenerating to banditry, engulfed several hill states and parts of Assam. All this political turmoil made it almost impossible for foreign scholars to visit, let alone do research, anywhere in the northeast and it restricted the work of Indian linguists as well. Only in the mid-1990s did the political situation finally begin to cool. Foreigners were once again welcomed to Northeastern India, though, as the millennium begins, some areas remain off limits. At the same time, the number of Indian scholars who have a commitment to the area and to Tibeto-Burman linguistics, including linguists who, themselves, have a North-eastern Indian tribal background, has been growing. We can expect the next decades to see new cooperation between Indian and foreign linguists and a burst of productive research on this, the least well-researched portion of the Tibeto-Burman world.

## REFERENCES

- Anonymous (1991) *Zangmianyu Yuyin He Cihui* (Tibeto-Burman Phonology and Lexicon), Beijing: Zhongguo Shehui Kexue Chubanshe.
- Anonymous (1997) *Languages of India*, India Missions Association.
- Benedict, Paul K. (1972) *Sino-Tibetan: a conspectus*, Cambridge: Cambridge University Press.
- Benedict, Paul K. (1976) 'Sino-Tibetan: another look', *Journal of the American Oriental Society* 96.2: 167–97.
- Bernot, Lucien (1966) 'Éléments de vocabulaire Cak recueilli dans le Pakistan Oriental', in Ba Shin, Jean Boisselier, and A.B. Griswold (eds) *Essays Offered to G.H. Luce by his colleagues and friends in honour of his seventy-fifth Birthday*. Artibus Asiae: Supplementum XXIII, Ascona, Switzerland: Artibus Asiae 1: 67–91.
- Bodman, Nicholas C. (1988) 'On the place of Lepcha in Sino-Tibetan: a lexical comparison', *Linguistics of the Tibeto-Burman Area* 11.1: 1–26.
- Brown, R. Grant (1920) 'The Kadus of upper Burma', *Bulletin of the School of Oriental Studies* 1.3: 1–28.
- Burling, Robbins (1959) 'Proto Bodo', *Language* 35.3: 433–53.
- Burling, Robbins (1971) 'The historical place of Jinghpaw in Tibeto-Burman', in F.K. Lehman (ed.) *Papers on the Tibeto-Burman historical and comparative linguistics from the second annual meeting on Sino-Tibetan reconstruction*, Occasional papers of the Wolfenden Society on Tibeto-Burman Linguistics II: 1–54.
- Burling, Robbins (1983) 'The Sal Languages', *Linguistics of the Tibeto-Burman Area* 7.2: 1–32.
- DeLancey, Scott C. (1991) 'Sino-Tibetan languages', in William Bright (ed.) *International Encyclopedia of Linguistics* 4: 445–9, New York: Oxford University Press.

- Dondrup, Rinchin (1988) *A Hand Book on Sherdukpen Language*, Itanagar: Directorate of Research, Government of Arunachal Pradesh.
- Dondrup, Rinchin (1990) *Bugun Language Guide*, Itanagar: Directorate of Research, Government of Arunachal Pradesh.
- Downs, Frederick S. (1994) *Essays on Christianity in North-East India*, New Delhi: Indus Publishing Company.
- Driem, George van (1992) *The Grammar of Dzongkha*, Dzongkha Development Commission, Royal Government of Bhutan.
- Forrest, R.A.D. (1962) 'The linguistic position of Rong (Lepcha)', *Journal of the American Oriental Society* 82.3: 331–5.
- French, Walter (1983) 'Northern Naga: a Tibeto-Burman mesolanguage', unpublished PhD dissertation, City University of New York.
- Grierson, G.A. (1903–1928) *Linguistic Survey of India*. Vol. 1, pt. II, *Comparative Vocabulary* (1928), Vol. III *Tibeto-Burman Family*: Pt. I, *Specimens of the Tibetan Dialects, the Himalayan Dialects, and the North Assam Group*, 1909, Pt. II, *Specimens of the Bodo, Naga, and Kachin Groups*, 1903, Pt III, *Specimens of the Kuki-Chin and Burma Groups*, 1904. Government of India, Central Publication Calcutta Branch.
- Grierson, G.A. (1921) 'Kadu and its relatives', *Bulletin of the School of Oriental Studies* 2.I: 39–42.
- Grimes, Barbara F. (ed.) (1996) *Ethnologue: Languages of the World*, 13th edn, Dallas: Summer Institute of Linguistics.
- Huffman, Franklin E. (1986) *Bibliography and Index of Mainland Southeast Asian Languages and Linguistics*, New Haven and London: Yale University Press.
- Karlsson, B.G. (1997) *Contested Belonging: An Indigenous People's Struggle for Forest and Identity in Sub-Himalayan Bengal*, Lund, Sweden: Lund Monographs in Social Anthropology 4.
- Kumar, B.B. and Pochuri, T. (1972) *Hindi Pochuri English Dictionary*, Kohima: Nagaland Bhasha Parishad.
- Kumar, B.B., Meya, T.P., and Jon, S. n.d. *Hindi Khamngan English Vocabulary*, Kohima: Nagaland Bhasha Parishad.
- Lorrain, J. Herbert and Savage, Fred. W. (1898) *A Grammar and Dictionary of the Lushai Language (Dulien Dialect)*, Shillong: Assam Secretariat Printing Office.
- Lorrain, J. Herbert (1907) *A Dictionary of the Abor-Miri Language with Illustrative Sentences and Notes*, Shillong: Eastern Bengal and Assam Secretariat Press.
- Matisoff, James A. (1991) 'Sino-Tibetan linguistics: present state and future prospects', *Annual Review of Anthropology* 20: 469–504.
- Marrison, Geoffrey E. (1967) 'The classification of the Naga languages of North-East India', unpublished PhD dissertation, University of London, School of Oriental and African Studies.
- Mason, Marcus C. *et al.* (1905) *English–Garo Dictionary*, Shillong: Assam Secretariat Printing Office.
- McCulloch, W. (1859) 'Account of the valley of Munnipore and of the hill tribes, with a comparative vocabulary of the Munnipore and other languages', Selection from the records of the Government of India (Foreign Dept. No. XXVII). Calcutta: Bengal Printing Co.
- Michailovsky, Boyd (1994) 'Bhutan: language situation', in R.E. Asher (ed.) *The Encyclopedia of Language and Linguistics*, 1: 339–40.
- Michailovsky, Boyd and Mazaudon, Martine (1994) 'Preliminary notes on the languages of the Bumthang group', *Tibetan Studies, Proceedings of the 6th Seminar of the International Association for Tibetan Studies*, Oslo: The Institute for Comparative Research in Human Culture.
- Shafer, Robert (1947) 'Hruso', *Bulletin of the School of Oriental and African Studies*, 12: 184–96.
- 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.
- Simon, I.M. (1970) *Aka Language Guide*, Shillong: North-East Frontier Agency.
- Simon, I.M. (1976) 'The Khoa Language', *Resarun* 2.3: 8–10.
- Simon, I.M. (1979) *Miji Language Guide*, Shillong: Directorate of Research, Government of Arunachal Pradesh.

- Sun, Tianshin Jackson (1993a) 'A historical-comparative study of the Tani (Mirish) branch in Tibeto-Burman', PhD dissertation, University of California, Berkeley.
- Sun, Jackson, T.-S (Sun, Tianshin Jackson) (1993b) 'The linguistic position of Tani (Mirish) in Tibeto-Burman: a lexical assessment', *Linguistics of the Tibeto-Burman Area* 16.2: 143–88.
- Tayang, Aduk (1990) *Sulung Language Guide*, Itanagar: Directorate of Research, Government of Arunachal Pradesh.
- 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/Philadelphia: John Benjamins.



**PART 4**

---

**LOLO-BURMESE**

---

**LANGUAGES**

---





# BURMESE

*Julian K. Wheatley*

## 1 BACKGROUND

### 1.1 Affiliations

Standard Burmese is the official language of Burma, or Myanmar (the latter, a cognate name officially adopted in 1989). Burmese is spoken as a first language by the dominant Burman ethnic community and as a first or second language by the vast majority of the many minority groups living within or close to the borders of the country – nearly 50 million people in all. The standard has evolved from a ‘central’ dialect spoken in the lower valleys of the Irrawaddy and Chindwin rivers. As the formal means of exchange, it has penetrated to most parts of the country, but 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, an official language of the Arakan region in the southwest of Burma, formerly home to a littoral state that remained independent until the late eighteenth century. Arakanese is also spoken by groups of people across the western border of Burma, notably the Marma, who settled in the Chittagong Hills several centuries ago (Bernot 1958). Other dialects (or dialect groups) include Tavoyan, spoken mostly in the region around the town of Tavoy in the southeastern ‘tail’ of Burma, Intha, spoken in the Inle Lake region of the Shan highlands, and Yaw, which is found in a central region, west of the Irrawaddy. The dialects differ from the standard mostly noticeably in pronunciation, but also in some lexical and grammatical vocabulary (see Okell 1995).

None of the Burmese dialects preserves the full sets of final consonants that are attested by the orthography. Such evidence is, however, to be found in languages such as Atsi (‘Zaiwa’ in China) and Achang, spoken in the far western part of Yunnan and neighbouring parts of Burma. Both are closely related to Burmese, and with other languages, are grouped in a Burmish branch, which in turn is closely related to the extensive Yiish group of southwest China and bordering regions. Thus Burmese belongs to the Burmish-Yiish sub-branch of Tibeto-Burman (TB).

### 1.2 History

The distribution of the Burmese-Yiish languages suggests that people speaking a Burmese prototype spread westward from what is now southwestern China and passed into a region occupied by speakers of other languages and profoundly influenced by Indian political and religious traditions. Precisely when Burmese speaking people appeared in the central plains of Burma is still very uncertain. The language is not attested until early in the twelfth century, when it begins to appear on stone inscriptions in the temples of Pagan. But the transformation

of a highland TB people into rulers of a powerful and sophisticated state must have involved a complex interplay of influences over a number of centuries.

At first, Burmese competed as the language of inscriptions with Pali, Mon and very occasionally, Pyu – the last, a language (thought to be TB) with probably no more than ceremonial function at the time. Mon, on the other hand, was an important language of the first ‘Burmese’ kingdom, and remained a powerful presence in Lower Burma until the Mon ‘state’ was absorbed into the Burmese in the eighteenth century. Mon, genetically-related to Cambodian, continues to be spoken in the Moulmein region.

The early appearance and relative frequency of Mon inscriptions in Pagan even after Burmese began to be used, plus the presence of a contemporary Mon state in Lower Burma and a much earlier one in Lower Thailand, has led some historians of Burma to assign a dominant position to Mon culture in the early Burmese state. Until the evidence is clearer, however, it may be better to withhold a decision on the amount, and even the direction of influence, and simply observe that the writing systems used for Burmese and Mon are virtually identical in form. Both derive ultimately from a Brahmi prototype. In any case, by the thirteenth century, inscriptions in Old Burmese predominate.

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 AD) and beyond. Manuscripts make their appearance later, with the writing scratched on specially prepared palm leaves, either bound flat with cord (*peṣa*) or folded in zigzag fashion (*pəṛəbai?*). Much of the post-inscriptional literature is verse, typically with four-syllable lines, 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 Ayutthia 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 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 appear in the 1870s. The first novel appeared in 1904, a work said to be inspired by *The Count of Monte Cristo*.

### 1.3 Contact

From earliest times, Burmese has been in contact (*kun+te?*) with speech or writings in other languages, many of which have left, and continue to leave, their mark. Since earliest times Pali, accorded high prestige as the language of the Buddhist scriptures, has been the source of specialized words. 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. the first three ordinal numbers, spelt *PAHTAMA*, *DUTIYA*, *TATIYA* in both Pali and written Burmese. 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: *sei?* ‘mood; thought’, written *CIT*, from Pali *CITTA*; *θa?pe* ‘elegant’, written *SAPPAY*, from Pali *SAPPĀYA*. Pali (and less commonly, 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 foreign stock.

While some Pali material found in Burmese shows the effects of transmission via Mon texts, much Mon material is likely to have been transmitted through speech. Mon influence on Burmese seems to be confined to loanwords having to do with the natural and artificial environment (Hla Pe 1973), though Bradley (1980) detects some structural borrowing as well.

In more recent times, English, first as the language of the colonial power and more recently as the language of worldwide popular culture, has been the source of specialized vocabulary either in the form of loanwords (*dimokərəsi*) or neologisms (*lú-ək'wínəyè* 'human rights'). But as with loanwords from other languages, those derived from English almost always appear as nouns: *dizàin s'wè-* '(design hang) to design'.

#### 1.4 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 by an orthography which stabilized by the eighteenth century or sooner, and 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 (inferred, in part, from traditional Indian values of the letters) differ rather drastically from spoken: *LAKPHAK-RE KRAM*: 'plain tea' (for transliterations of Burmese, see Okell 1971) is read *ləp'ɛʔyeçàn*. However, sound change has been consistent enough that, with the exception of weak syllables and juncture voicing (seen in the first and last syllable of our example), pronunciation can be read off from spelling fairly predictably.

Major sound changes include the reduction of four points of articulation of final consonants to a single feature for nasals (represented as -n) and another for oral stops (-ʔ); and the splitting (sometimes with merger) of vowels according to the syllable type, open vs closed, so that the spellings, *I, IN, IT* are read *i, ein, eiʔ*, and *U, UN, UT* read *u, oun, ouʔ*, etc. Another important change involves an ordered shift of initial consonants, e.g. *\*s > θ; \*c > s; \*ky, kr > c*. Two of these are revealed by comparison of literal and spoken values of Burma's best known name, spelt *ʔOÑ HCAN*: *CU KRAÑ*, pronounced *Aun S'àn Sú Ci* (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 *Càun-ko θwà-tɛ* '(school-GOAL go-RLS) [He]'s going to school' would be rendered *Càun-θó θwà-θi*. LBs forms are not always homologous with colloquial. The LBs particle pronounced -*ywé*, for example, subsumes functions of two colloquial particles, the causal -*ló* and the conjunctive -*pì*; while colloquial -*ko* 'GOAL' subsumes the functions of LBs -*θó* 'towards', -*à* 'DAT', and -*ko* 'D.OBJ'.

## 2 PHONOLOGY

It is necessary to distinguish two types of syllable in Burmese: full, or major syllables and reduced, or minor. In the latter, phonological distinctions are confined to the initial consonant.

The vowel is realized as shwa, and there are no tonal contrasts. Words may contain one, or less commonly, two minor syllables prefixed to a major: *kə̀lə̀* ‘Indian; foreign’, *kə̀lə̀tʰain* (foreign sit) ‘chair’. In some cases, minor syllables clearly derive from full, e.g. *ŋə̀pʰi* ‘fish paste’, the first syllable from *ŋà* ‘fish’.

## 2.1 Burmese sounds

Major syllables contain five syntagmatic positions: initial ( $C_i$ ), medial (M), vowel (V), final ( $C_f$ ) and tone (T), with vowel, tone, and initial (including zero, which represents glottal onset) present in all syllables.

Burmese vowel values vary according to whether the syllable is open (-V) or closed (-n or -ɔ̃). While it is possible to align open and closed values along the columns shown in Figure 12.1, and to follow the writing system by identifying closed /au/ with open /ɔ/, what is gained in consistency is lost in transparency. Here, as indicated in Figure 12.1, we use a broad transcription rather than a strictly phonemic one.

The four tonal distinctions of major syllables can be termed ‘creaky’, ‘low’, ‘heavy’ and ‘checked’, the last symbolized by -ʔ. As the names suggest, tone in Burmese has a complex realization, which may include features of phonation, pitch and contour, length, and vowel quality. The salience of these features varies with contexts, notably phrase final (citation) vs non-final.

The creaky tone is far less common than the others, which reflects its secondary origin (Thurgood 1981). Some compensation is provided by grammatically induced shifts of low tone, and less often, heavy, to creaky. With vocatives, for example, creaky tone can indicate impatience (Okell 1969: 20): *θə̀mì-yé* ‘hey, daughter’ vs normal *θə̀mì-ye*. With pronouns and other personal referents, the shift to creaky can, with or without  $P_o$  marking, indicate grammatical relationships such as possession (or attribution) or object: *cə̀nɔ́(-kə̀)* ‘(first person OBJ [-OBJ]) me’ (but *cə̀nɔ́*); *ŋá tu-ye* ‘(first person GEN nephew-VOC) ... my boy’ (but *ŋa* ‘first person’).

Initials and medials: Figure 12.3 shows the  $C_i$  arranged in three series, based on morphological considerations. The voiced series is far rarer than the aspirate or plain. They 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): *jeiʔ* ‘hook’ and *cʰeiʔ* ‘to hook’. The aspirate and plain series are associated with a vestigial morphological process that links, in regular pattern, about fifty pairs of verbs in which the causative member has an initial from the aspirate series, and the simplex, one from the plain: *cʰò* ‘to break’, defined in the dictionary as *cò-se-θi* ‘cause to be broken’; *sʰouʔ* ‘to tear’, defined as *souʔ-se-θi* ‘cause to be torn’. Pairs such as *eiʔ* ‘sleep’ and *θeiʔ* ‘put to bed’

### Syllable type

-vowel	i	e	ɛ	a	ɔ		o	u
-n	ɪ	eɪ		a		aɪ	aʊ	oʊ
-ʔ	ɪ	eɪ	ɛ	a		aɪ	aʊ	oʊ
Transcribed as	i	e	ɛ	a	ɔ		o	u
		ei				ai	au	ou

FIGURE 12.1 VOWELS OF MAJOR SYLLABLES

creaky (˘):	starts high and falls; tense voicing becoming creaky; not long
low (ø):	persisting low, but rising in phrase final position; clear voicing; long
heavy (˘):	persisting high, but falling sharply in phrase final position; clear, sometimes slightly lax voicing; long
checked (-ʔ):	mid-high with slight fall; clear voicing with occlusion variously realized; very short

FIGURE 12.2 TONAL CONTRASTS

## SERIES

Aspirate	p' t' c' k' s'	hm hn hɲ hŋ	hl ʃ (hw)	h
Plain	p t c k s θ	m n ɲ ŋ	l y w (r)	ø[=?]
Voiced	b d j g z (ð)			

FIGURE 12.3 INITIAL CONTRASTS

reveal a likely origin in *s*-prefixation ( $*s > \theta$ ); pairs like *yʃ* 'be slack' and *ʃʃ* 'slacken' support alignment of /ʃ/ and /y/ in the table above (and the convention followed by some transcriptions of writing *ʃʃ* as *hyʃ*).

Items in parentheses, such as /hw/ are rare; /t/ is found in Pali and other loanwords; and but for a very few cases (e.g. phrases involving the conjunction *ðó*), *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'-/ (more consistently distinguished from plain /s/ in the speech of older speakers), as well as a set of pre-aspirated nasals and voiceless resonants.

The only medials are /-y-/ and /-w-/; the /-r-/ attested by the orthography (like the /-l-/ of Old Burmese) is reflected either in modern /-y-/ or in the palatal affricate series ( $c < *ky$ ,  $*kr$  etc.). Medial /-y-/ patterns with  $C_1$ , appearing only with the labial stops (oral and nasal) and with the lateral /l/. Medial /-w-/ , on the other hand, patterns with the rhymes, appearing only with non-back vowels in open syllables (-*wi*, -*we*, -*wɛ*, -*wa*) and a few others in closed (e.g. -*win*, -*wan* -*wain*, -*wɛʔ*). Historically, *w*-rhymes have been the source of the series of /u/-rhymes in closed syllables (*HPWAT*, read *p'uʔ* 'monitor lizard'), but there remains some variation in the modern reflexes, e.g. *KWAM*: 'betel nut', read *kùn* and *kwàn*.

## 2.2 Juncture

The realization of  $C_2$ s 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 our transcription, non-close juncture is indicated by + (word) or space (phrase), close juncture, by contiguous syllables (word) or hyphens (phrase). Underlined consonants show close juncture voicing in standard speech.

In citation form, or before a syllable in open juncture, the two  $C_2$ s are realized as nasalization on the vowel and final glottal stop (plus associated tonal features) respectively. With following close juncture, however, the final nasal tends to assimilate to the position of the following  $C_1$  (*pàṅwín* 'flower', pronounced [*pàmbwín*], with a labial final, along with close

juncture voicing of the following initial); while the final 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  $C_i$  (*θauʔ-tɛ* ‘drink-RLS’ [*θauttɛ*], with plain dental stop, and *myɛʔsɪ* ‘eye’, [*myɛssɪ*], with plain fricative).

Close juncture also affects the form of the following  $C_i$ . Following a checked syllable (-ʔ), the  $C_i$  retains its original manner, but following a smooth syllable (-v, -n), if the  $C_i$  is voiceable (i.e. if it is aspirate or plain and there is voiced series  $C_i$  corresponding to it), then it is realized voiced (manner varying with speed): contrast *daʔs* ‘petrol’ with *ðunz* ‘coconut oil’ (*s* ‘oil’) or *eiʔk’an* ‘bedroom’ with *ék’an* [*égàn*] ‘(guest room) living room’ (*ək’an* ‘room’).

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, [*zəbwè*] ‘table’ (< *sà* ‘eat’ + *pwè* ‘festival’), with voicing, but *θəp’ðun* ‘gums’ (*θwà* ‘teeth’ + *p’ðun* ‘cover’), without. Sometimes, as the example for ‘table’ shows, where the major syllable has a voiced stop, a process of consonant harmony may induce voicing on the initial of the weak syllable as well *zəpwè* [*zəbwè*].

McDavid (1945: 6) posits an intermediate degree of juncture, characterized by the anticipatory assimilation of  $C_f$ , but not the perseveratory assimilation (voicing) of the following  $C_i$ . Loanwords and verb compounds, which have features of close juncture but not, in most cases, word-internal voicing, may be support for this view. Close juncture (with voicing) is the characteristic of nominal compounds, rare in verbal compounds, and a regular feature of certain phrasal combinations, such as ‘noun plus (unprefixed) stative verb’; most particles follow in close juncture.

### 3 WORD CLASSES

#### 3.1 Verbs and such

Verbs and nouns show virtually no class overlap in Burmese. Nouns are frequently derived from verbs, but the verb repertoire expands through compounding, and not from derivation from other word classes or from borrowing. Disyllabic verbs can usually be analysed as lexical compounds (two roots) or pleonastic compounds in which a stylistic effect is achieved by combining the root with the (usually rarer) synonym, or with a rhyming or alliterating syllable (‘rhyme’ or ‘chime’): *pyð*, *s’o* and *pyð+s’o* ‘speak’; *cwà* and *cwà+wa* ‘boast’; *twè* and *twè+tð* ‘ponder’. Verbs may also draw noun complements into lexical configurations of ‘verbs with tied nouns’ (Okell 1969: 36): *à na-* ‘(strength hurt) be embarrassed’; *θəpð kàun-* ‘(disposition kind) be kind’.

Stative verbs (SVs) can be distinguished from functive (class overlap between the two is rare) by their behaviour as modifiers. While both may precede a  $N_h$  if mediated by a relative marker, SVs more usually follow their  $N_h$ , either as derived ə-prefixed nouns, or without prefix and in close juncture: *ləp’ɛʔye-əc’o-pð* ‘lightly sweetened tea’ (*c’o* ‘sweet’, *pð* ‘light’); contrast *c’o-tɛləp’ɛʔye* ‘tea that is sweet’.

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 by processes of repetition, prefixation, rhyme or chime, or combinations thereof: *sðð t’á-* ‘to get up early’ (*sð* ‘be early’); *əluʔ cɛʔ-* ‘learn by heart’ (*luʔ* ‘be free’); *byàunpyan pyð-* ‘to say s/t back to front’ (*pyàun* ‘change’ [unpredictably voiced in the compound], *pyan* ‘return’); *səniʔ təcá louʔ-* ‘to do s/t systematically’ (*səniʔ cá-* ‘be systematic’). The last example, in which the 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. Derived adverbials may become institutionalized enough to be considered true adverbs, but those that can retain complements, at least, suggest a grammatical process of subordination or embedding. Adverbs formed by repetition may also retain complements: *Bəlau? caḡa ne-ḡun-mə-lè?* ‘(How long stay-more-IRL-Q) How long are [you] staying?’ (*ca* ‘be long in time’).

### 3.2 Nouns and such

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, repetition, compounding (lexical and pleonastic), and combinations thereof, e.g. the elaborate pleonastic noun *əlou? + əkain* ‘work’ with roots *lou?* ‘do’ and *kain* ‘handle’. Nominal compounds have rather more 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: *k’əpai?hnai?* ‘(pouch pick) pickpocket’; *θE?θa?lu?* ‘(life killing free) vegetarian’.

Except for the first person and second person pronouns, *ŋa* and *nin*, which derive from ST pronominal prototypes, pronouns in Burmese are derived from nouns. Pronominal functions are also frequently performed by kin terms, names, titles or other nouns (e.g. *mei?s’we* ‘friend’ or *s’əya* ‘teacher’), which can substitute for a second person or other pronoun.

Standard pronouns with typical usage are shown in Figure 12.4 (cf. Cooke 1968 for greater detail; also Okell 1969: 100–1). The pronouns shown are ordered roughly along the dimension of social distance, with deferential (distant) first and familiar (intimate) last; likely pairings of first person plural and second person plural cannot be safely inferred from the chart. As indicated, in addition to registering social distance, pronouns may also be sensitive to the sex of the speaker. There are also pronouns for showing deference to people of special status, e.g. *dəbélɔ* ‘(disciple royal) first person, to monks’ or *dəbélɔ-má* (female speaker) and *əshinp’əyà* ‘(lord) second person, to monks’.

1p	<i>cunɔ/cəno</i>	(honourable slave) very polite, male speaker
	<i>cunmá/cəmə</i>	(female slave) very polite, female speaker
	<i>ko</i>	(body) polite, mostly among same sex (?)
	<i>cənou?/cou?</i>	(worthless slave) neutral
	<i>ŋa</i>	familiar, mostly male to male (?)
2p	<i>shin</i>	(owner; master) very polite, female speaker
	<i>k’inbyà/k’əmyà</i>	(k’in ‘be close’) polite, male speaker
	<i>tə</i>	familiar, female speaker
	<i>mìn</i>	(king) familiar
	<i>nī</i>	very familiar, female to female, rural
	<i>nin</i>	familiar (crude?)
3p	<i>θu</i>	

FIGURE 12.4 STANDARD BURMESE PRONOUNS



The subject of reflexive and emphatic pronouns needs more investigation, but we can note the following forms built on partially repetitive structures such as *kó* + [ ] + *ko* ‘self.GEN + ( ) + self’ or *thú* + [ ] + *thu* ‘third person. GEN + ( ) + third person. Insertion of *ha* ‘thing’ (or *p’a* or *p’a a* ‘on one’s own’) forms emphatic pronouns: *Kó + ha + ko-p’è louʔ-pa-yá-ge* ‘(myself-only do-POL-must-CAUS) May I do it myself?’ Insertion of body terms creates reflexive pronouns, as for example *pà* ‘cheek’ in the following: *Kó + pà + ko θwà c’á-pa-là* ‘(self + cheek + self go strike-POL-Q) Why don’t you go and slap yourself [on the cheek]!’

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: *Houʔ-pa-tɛ shin* ‘(be.the.case-POL-RLS sir) That’s right, sir [woman speaker]’; *Ci-ta-pa kwa* ‘(look-RLS. NOMZR-POL guy) Just looking, man!’

A subclass of pronouns are deictics, which include the proximate *di* ‘here’; ‘this’ and the distal *ho* ‘there’; ‘that’, plus *da* ‘this one’; ‘that one’. The three also occur with the ejaculation *hò* ‘oh’; ‘aha’ prefixed (*hòdi* ‘this’ etc.) or the ejaculation *è* ‘yes’; ‘yeah’ (*èdi* ‘this’ etc.). The former are often exophoric (pointing), the latter anaphoric (textual).

Burmese, like so many other languages in the region, makes use of numeral classifiers to count nouns or to substitute for them: e.g. *pà* for monks, *yauʔ* for (ordinary) people, *kaun* for animals, *t’ɛ* for clothes, *s’u* for sacred objects (e.g. pagodas), *lòun* for round things. For extralinguistic reasons, some nouns can be classified in several ways: *lounji tət’ɛ* ‘(sarong one-cloth)’ or *təkwin* ‘(one-circle)’. Some nouns are self-classifying: *ein tə-ein* ‘house one-house’. Nouns may also select quantifying measures: *ηəpyəθi tə-p’i* ‘a (small) bunch of bananas’. Nominal forms of classifiers appear in a special construction for counting items in multiples of ten (10, 100, 1000 etc.): *p’òunçì kò-pà* ‘(monks 9-CL) 9 monks’ but *p’òunçì əpà hnə-s’ɛ* ‘(monks əCL 2–10) 20 monks’.

As with verbs, there are subclasses of nouns that show a range of grammatical and semantic specialization. These include plural markers. Pronouns are regularly made plural by the addition of *-tò*: *cənɔ-tò* ‘we’; ‘us’. The same word appended to nouns, however, means ‘and the like; and the others’: *s’imì-tò* ‘oil lamps and so on’; *má sò sò-tò* ‘Ma So So and family’. The usual plural for countable nouns is *-te* (or *-twe*): *mipòun-tɛ* ‘lanterns’; with mass nouns, *-te* signifies ‘a large amount of’: *sɛʔku-tɛ-né* ‘with lots of paper’. In more formal contexts, *-myà*, 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 nouns worth noting are the ‘female’ suffix, *-má* (*əmá* ‘female’; ‘main’), the augmentative suffix, *-cì* ‘big’, the diminutive, *-k’əlè*, *-lè* ‘child’: *s’əya-má* ‘female teacher’; *s’əya-çì* ‘eminent teacher’; *s’əya-k’əlè* ‘(teacher-DIM) addressing a nun’. Another class of nouns specifies location: *ein-fé-hma* ‘(house-front-LOC) in front of the house’ (contrast *ein-əfé* ‘the front of the house’, in which *əfé* is the  $N_p$ ). One of these location nouns, *-t’ɛʔ* ‘above’, has become specialized as a comparative marker: *Lu bəwa-ha k’wè bəwa-t’ɛʔ kàun-θə-là?* ‘(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: *yauʔ-k’a* ‘when [we] arrived’ (*ək’a* ‘time’); *yauʔ-sá* ‘since arriving’ (*əsá* ‘start’). These nouns may also act as heads for RCs: *Məθi-θə-han s’aun ne-tɛ* ‘(not know-RLS-appearance adopt stay-RLS) [They] are pretending not to know’.

## 4 SYNTAX (CONSTRUCTIONS AND PARTICLES)

### 4.1 The verb phrase

Minimally, the VP consists of a head followed (in most cases) by one of a disparate set of main (cf. 4.3) or subordinate clause particles (4.5). In most cases, other elements are also

present. Verbs may appear before the  $v_h$  to form concatenations ('serialized verbs'): *θwà k'aʔ-pa* '(go ladle-POL) Go and fetch [some water]'; *yò sà-yin* '(mix eat-if) if [you] eat [them] mixed [with other fruit]'; *P'əyàuntain t'ɛ t'ün-tɛ* '(candles insert light-RLS) [They] stick candles [in the ground] and light [them].' Concatenations can be resolved into temporally consecutive clauses by the insertion of the subordinating particle *-pi* 'and' (e.g. *yò-pi sà-yin*).

Auxiliary verbs appear in unmediated strings directly after the main verb. They show a range of grammatical properties and semantic specialization that suggests 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. For example, in the phrase *Mə pyin pè nain-p'ù* '(not repair give prevail-NEG) [I] can't fix [it] for [you]', only *pè* allows direct negation, not *nain*: *Pyin mə pè nain-p'ù* (same meaning).

Auxiliaries include verbs like *ne* 'stay', as in *La ne-pi* '(come stay-PUNC) Here he comes!'; or *t'à* 'put' and *pè* 'give' in *K'əná cí t'à pè-pa* '(a.while watch put give-POL) Watch [it] for [me] for a while, would you mind?' Some auxiliaries, such as *pè* 'give' (benefactive) in the previous example and *k'àin* 'ask'; 'tell' (causative) in the next, add arguments to the clause: *əP'wà-çì-ko òunθi c'ìʔ pyá k'àin-yá-tɛ* '(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 *p'ù* 'have ever' (homophonous with the negative clause marker), as in *ηəpí-çə mə sà-p'ù-p'ù-làʔ* '(shrimp.paste-fried not eat-ever-NEG-Q) Haven't [you] ever eaten fried shrimp paste?' Another particle, *-k'ɛ*, conveys displacement in space or time: *Lɛʔs'aun-ko yu la-k'ɛ-mɛ* '(present-OBJ carry come-back-IRL) [I]'ll bring [you] a present [from there]'. It also appears in subordinate clauses conveying unfulfilled conditions: *aunmyin-k'ɛ-yin* '(succeed-k'ɛ-if) if [she] had succeeded'. *-K'ɛ* is paradigmatically paired with *-laiʔ* ('follow') by Allott (1965: 296), a particle whose prototypical meaning is 'action away', but which may also suggest abrupt or effective action: *Yè pè-laiʔ-pa* '(write give-laiʔ-POL) Would [you] mind just writing [it] down [for him]'. Another paradigmatically related set includes the aspectual particles, *θè* and *òun*, both meaning 'still'; 'not yet (no change)', the first used only in realis clauses and the second, only in irrealis; and *tó* 'about to'; 'anymore (change of state)' (realis or irrealis). The last is difficult to distinguish in function from the 'punctative' main clause marker, *-pi* (cf. 4.3). The following five short examples suggest the differences between these four aspectual markers: *sà ne-θè-tɛ* 'still eating (RLS)', *sà-òun-mɛ* 'will eat (IRL)', *sà-tó-tɛ* 'finally ate (RLS)'; *sà-tó-mɛ* 'about to eat (IRL)', and *sà-pi* 'am eating (now); have eaten (PUNC)'.

## 4.2 The noun phrase

The following sentence (taken from a recorded text) is cited to illustrate some of the features discussed in the sections that follow:

<i>èdi sa-ko əlun + myàmyà pyan nain-tɛ</i>	these texts-OBJ very + many recite can-REL
<i>p'òun-tə-çì-myà-ko</i>	monks-great-PL-OBJ
<i>lu-te-ká-lè-p'è</i>	people-PL-s-also-EMPH
<i>innətan c'ì + cù-çá-pa-tɛ</i>	very praise-COLL-POL-RLS

'Monks who can recite back large parts of these texts are highly praised by the people/get high praise from the people.'

In the NP, modifiers, for the most part, precede the  $N_h$ ; these include demonstratives (*èdi*) and RCs (marked by *-té* above); SVs follow (*ci*) as do classifier phrases, noun suffixes

ha	subject
ká	initial point; contrasted subject; agent; source; topic; at (past time); attributive
ko	end point; goal; object; extent; degree; at (future time)
né	manner; instrument; accompaniment; and
hma, t[w]in	locative
yé, ké	genitive; attributive (yé after non-checked syllables, ké after checked)
mó, cáun	because; on account of
p'ó	for the purpose of
áp'yin	as for; in view of

FIGURE 12.5 MAIN 'CASE' MARKING POSTPOSITIONS (COLLOQUIAL)

(-myà, -te), and classes of grammaticalized nouns. Case marking  $P_oS$  (-ká) and general  $P_oS$  (-lè- $\underline{p}$ 'è) occupy final position in the NP. Several of the most common case marking  $P_oS$  show 'local' and 'non-local' functions: -ká, 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 -ká requires positing homophonous particles with several senses. The  $P_o$  -ko, with meanings that cohere around the notion 'goal' (object, goal of motion, extent, etc.) requires the same type of analysis. Figure 12.5 lists the main case marking postpositions and their functions.

### 4.3 The clause

Burmese is consistently verb-final, though positive categorial sentences such as *U Myín Maun cáun s'əya* 'U Myint Maung's 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 subject complements, object complements, inner locatives and direct objects, and all those elements before adverbials and interrogative phrases, the order of nominal elements is, in principle, free. Thus, the long sentence cited above ('Monks who can . . .), in which the object precedes the subject, is not marked, and can be quite naturally translated by the English passive.

The VP is the only obligatory element of clause structure (other than in noun clauses), and although a sentence such as *Pyin pè-lai?mε* '(repair give-follow-IRL) [I]'ll fix [it] for [you]' 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. *Məne?p'yan* 'Tomorrow', given as a one-word answer to a question 'When?' Repetition of the VP 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': *θwà-mε-nɔ* '(go-IRL-okay?) [I]'ll be off then'. Main clause particles include the following (Allott 1965: 288): imperative (affirmative:  $\emptyset$ , negative: -nε) and indicative (affirmative: -tε or -θə 'realis (RLS)', -mε or -mə 'irrealis (IRL)', and -pi or -pə 'punctative (PUNC)';

negative: *-p'ù*). 'Punctative' (which might also be labelled 'change of state') expresses the realization of a state or initiation of an action: *Yau?-pi* '[We]'ve arrived'; *Lou?-pi* 'Now [she]'s doing it'. The distinction between realis, irrealis, and punctative is neutralized in negative clauses.

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: *Mò ywa-tɛ* '(Sky rain-RLS) It's raining'; and that 'complex transitive' clauses may be built around verbs of 'shaping' and 'processing', as well as 'naming' and 'selecting': *S'ànan ye-ko s'à c'ɛ? ne-tɛ* '(Salt water-OBJ salt cook ing-RLS) They are refining salt out of salt water'. Double-subject (topic-comment) sentences are a particularly common type: *Cənɔ pəs'ò ənan cì-tɛ* (first person.GEN sarong width big-RLS) My sarong is (extra) wide'. 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): *Bə̀lù-cì-ha əlun əyou? s'ò-tɛ* '(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 *shí*: *dì-hma shí-tɛ* '(here-LOC have-RLS) It's here'; *θu-hma shí-tɛ* '(third person-LOC have-RLS) He has [one]'.

#### 4.4 Processes

Questions in Burmese require no special structural reorganization. Yes-no questions are formed by the addition of the sentence particle *-là* to the declarative. Content questions are based on the interrogative pronoun *bɛ*, usually placed directly before the VP, in conjunction with another final interrogative sentence particle, *-lè*. *Bɛ* is also an indefinite: *bɛ-tɔ̄ (... lè)* 'when', *bɛ-tɔ̄-hmá* [+neg] 'never'. Negation is indicated by a preverbal enclitic, *mə* (which in main clauses is supported by the markers *p'ù*, *né*, etc.): *Mə pau?-p'ù-là* '(NEG grow-NEG-Q) Don't [they] grow [here]?' Alternative questions pattern like coordinate clauses, with redundant material are often omitted.

#### 4.5 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 VP, may be omitted (gapped). Verbs of locution, as well as certain verbs of cognition, take clauses as complements marked with the 'quotative' particle, *-lò*; in some cases the verb may be 'understood': *θwà-mə ló [can t'á-tɛ]* '(go-IRL-QUOT [intend put-RLS]) [I] intend to go'. A more general 'evidential' particle, *-tɛ*, appearing in final position as if meaning 'it was said', reports on any locution: *θwà-mə ló-tɛ* '[He] said [he] 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 *-ta* (occasionally *-t'a*) or *hma* (parallel in function to the main clause markers *-tɛ* and *-mɛ*): *Cənɔ pəs'ò cu? cá-hma sòyein-tɛ* '(first person sarong undo fall-IRL.NOMZR fear-RLS) I was afraid my sarong would fall off'. Embedding is especially common with verbs of elapsed time: *Yau? ne-ta θòun-lá ca-pi* '(arrive state-RLS.NZR 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: *Luçoun-ká yu la pè-ta* '(courier-S carry come give RLS.NOMZR) 'It was brought by courier' (with *-ta* replacing *-tɛ*). Cleft sentences perform a similar, 'identifying' function by recasting a verb sentence as a noun sentence: *Təyou? ɔ̀ə̀kà pyò ta?-ta-ká ko Nandá-pa* '(Chinese language

speak can-RLS.NOMZR-S brother [Name]-POL) It's Ko Nanda 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<sub>0</sub>'s), then reiterating it in verbal, a process known as 'exposure' or 'verb focus': *θe-ko θe-ðun-hma* 'die-extent die-FUT-IRL.NOMZR, [You]'ll get [yourself] killed (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 *k'an* 'suffer', for example, where it occurs in a construction that is as close to a passive as can be found in Burmese: *θu-ká baʔsəkà ətaiʔ k'an-yá-tɛ* '(third person 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<sub>0</sub>'s). Included in this set is the weakly subordinating particle, *-pì* 'and', that connects temporally consecutive as well as simultaneous clauses: T'ì s'əun-pì seʔbèin-né la-tɛ '(umbrella hold-ing bicycle-by come-RLS) [He] rode up holding an umbrella.' Our final example, a proverb for people who wear sarongs, illustrates conditional clauses, and gives us a glimpse of Burmese wit and poetry:

əpyð mə taʔ-yin s'è-θə-lo	speaking not able-if swear-RLS-like
ət'ain mə taʔ-yin p'yè-θə-lo	sitting not able-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

IRL	irrealis
POL	polite
PUNC	punctuative
RLS	realis

## REFERENCES

- Allott, Anna (1965) 'Categories for the description of the verbal syntagma in Burmese', in G.B. Milner and E.J.A. Henderson (eds) *Indo-Pacific Linguistic Studies II (=Lingua 15)*: 283–309.
- Bernot, Denise (1958) 'Rapports phonétiques entre le dialecte Marma et le Birman', *Bulletin de la Société Linguistique de Paris* 53.1: 273–94.
- Bradley, David (1980) 'Phonological convergence between languages in contact: Mon-Khmer structural borrowing in Burmese', *Proceedings of the Sixth Annual Meeting of the Berkeley Linguistic Society*: 259–67.
- 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.
- McDavid, Raven I. Jr (1945) 'Burmese phonemics', *Studies in Linguistics* 3.1: 6–18.
- 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 (1995) 'Three Burmese dialects', in David Bradley (ed.) *Papers in South-East Asian Linguistics No. 13: Studies in Burmese Languages (Pacific Linguistics, A-83)*: 1–138.
- Pe, Hla (1973) 'A tentative list of Mon loan words in Burmese', *Journal of the Burma Research Society* 50.1: 71–94.
- Thurgood, Graham (1981) *Notes on the Origins of Burmese Creaky Tone, Monumenta Serindica* 9 (Tokyo).

---

**FURTHER READING**

- Allott, Anna J. (1967) 'Grammatical tone in modern spoken Burmese', *Wissenschaftliche Zeitschrift der Karl-Marx Universität, Leipzig, Gesellschafts- und Sprachwissenschaftliche Reihe XVI*, 1/2: 157–62.
- 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)*: 131–54.
- Becker, Alton L. (1995) 'The figure a classifier makes: describing a particular Burmese Classifier', in A.L. Becker (ed.) *Beyond Translation: Essays toward a Modern Philology*, Ann Arbor, MI: University of Michigan Press, 211–30.
- Bernot, Denise (1980) 'Le prédicat en Birman parlé', *Langues et civilisations de l'Asie du Sud-Est et du monde Insulinde*, 8, Paris: SELAF.
- 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.
- Sawada, Hideo (1995) 'On the usages and functions of particles *-kou/-ka*. in colloquial Burmese', *Senri Ethnological Studies* 41: 153–87.
- Sprigg, R.K. (1957) 'Junction in spoken Burmese', in W.S. Allen (ed.) *Studies in Linguistic Analysis*, Oxford: Philological Society: 104–38.
- Wheatley, Julian K. (1995) 'Burmese writing', in Peter T. Daniels and William Bright (eds) *The World's Writing Systems*, Oxford: Oxford University Press: 450–6.
- Wheatley, Julian K. with San San Hnin Tun (1999) 'Languages in contact: the case of English and Burmese', *The Journal of Burma Studies* 4: 61–99.

## LAHU\*

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’).<sup>1</sup> 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 *Coxing* (also spelt Co Sung, Co Xung, or Khu Xung).<sup>2</sup>

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 north 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-nâʔ*) 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.<sup>3</sup> In Thailand, the biggest Lahu groups are known as Red Lahu (*Musəð Dεεη*, *Lahu Nyi*, *Lâhū-ni*) and Lahu Shehleh (*Lâhū šε-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 Yellow Lahu (*Lâhū-ši-bān-kew* and *Lâhū-ši-bān-lán*), as well as two interesting-sounding varieties called White Lahu (*Lâhū-phu*) and *Lâhū A-phū-be-le*. Next to nothing is known about the Lahu dialects spoken in North Vietnam. BL is the dialect described in my grammar and dictionary (Matisoff 1973/82, 1988).

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 8000 to 10,000 persons, while the even smaller Lahu presence in Vietnam seems to number about 1500. The tiny but growing Lahu

---

\* This chapter is a condensed and revised version of Matisoff 1992.

- 1 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.
- 2 See Vuong Hoang Tuyen 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).
- 3 See Nishida 1969, Bradley 1979. YL lacks the post-velar stops /q qh/ (BL *qô*, YL *kɔ*<sup>55</sup> ‘nine’; BL *qhâ*, YL *kha*<sup>55</sup> ‘bitter’), and does not have the central vowels /ɨ/ or /ɜ/, e.g. BL *phî*, YL *phi*<sup>55</sup> ‘dog’; BL *kə*, YL *ke*<sup>33</sup> ‘put in’.

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 hundred 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 fifty-five officially recognized minorities. Since 1953, they have played a major role in the local administration of the Lancang Lahu Autonomous County in the far southwest of Yunnan, where Chinese and Lahu are the joint official languages, and in the affairs of the adjacent Menglian County (officially the 'Tai-Lahu-Lawa Autonomous County').

## 2 PHONOLOGY

Black Lahu has seven tones. Five of these tones are smooth and open, pronounced without constriction; the other two are checked by a glottal stop.<sup>4</sup> See Figure 13.1.

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.<sup>5</sup> Even in syllable-initial position, only a single consonant may

<i>Name of tone</i>	<i>Pitch</i>	<i>Symbol</i>	<i>Example</i>
Mid	33	unmarked	<i>ca</i> look for
High-rising	35	/ʔ/	<i>cá</i> boil
High-falling	53	/ʔ/	<i>câ</i> eat
Low-falling	21	/ʔ/	<i>cà</i> ferocious
Very low	11	/ʔ/	<i>cā</i> feed
High-checked	54ʔ	/ʔʔ/	<i>cáʔ</i> rope
Low-checked	21ʔ	/ʔʔ/	<i>càʔ</i> push; machine

**FIGURE 13.1** THE SEVEN TONES OF BLACK LAHU

4 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).

5 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.



Consonants (24)					Vowels (9)		
p	t	c	k	q	i	ɨ	u
ph	th	ch	kh	qh	e	ə	o
b	d	j	g		ɛ	a	ɔ
m	n		ŋ				
f		ʃ	h				
v	l	y	ŋ̃ (ɣ)				

FIGURE 13.2 BLACK LAHU CONSONANTS AND VOWELS

occur; there are no consonant clusters.<sup>6</sup> BL has twenty-four initial consonants and nine vowels, as shown in Figure 13.2.

The five palatals /c ch j ʃ y/ have special dental pronunciations before /ɨ/: /cɨ chɨ jɨ ʃɨ yɨ/ → [tsɿ tshɿ dzɿ sɿ zɿ]. 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/ → [pfu pʰfu bvuu mvuu]. In these situations, /ɨ/ and /u/ have special allophones, [ɿ] and [uu], 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’.

### 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-phâ* ‘masculine’/ *-šē-ma* ‘feminine’ (*qhâʔ-šē-phâ* ‘headman’, *qhâʔ-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-šì-jɔ* ‘rainbow’; *pā-pā-qú-tí-ní* ‘dragonfly’; *a-gò-a-lí-pē* or *na-gù-na-gá-pē* ‘spider’; *khâʔ-pâ-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-ghâ* ‘ragweed’, *á-ghɔ* ‘home’; *ð-* ‘noun-forming prefix’, e.g. *ð-cè* ‘a plant’, *ð-šì* ‘a ball; a sphere’, *ð-u* ‘an egg’ (<*u* ‘to lay an egg’), *ð-qðʔ* ‘a curved thing’ (<*qðʔ* ‘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-cè ð-cè* ‘pair by pair’; (c) ‘Indefinite’ reduplication indicates uncertainty about the exact scope of reference of the noun: *tê chí kilô-lô* ‘about ten kilos’; (d) ‘Emphatic’ reduplication heightens or intensifies the meaning of the noun: *ð-lɛ-lɛ* ‘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

6 Medial *-w-*, in words like *bwé* (~*bô*) ‘room’, *chɨ-pí-qwèʔ* ‘barking deer’, is to be considered part of the vocalic nucleus. See Matisoff 1973/82: 15–21.

A-B-C-B, e.g.: *khô-mu-khô-nè* [word-high-word-low] ‘high and low pitched words; tones (of a tone-language)’, *ha-lè-ha-qa* [spirit-warm-spirit-qa] ‘happy and relaxed’, *dô-sa-gâ-sa* [think-easy-gâ-easy] ‘be serene and easy in one’s mind’.

Besides its basic stock of vocabulary inherited from Proto-Lolo-Burmese (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-lî* ‘custom’, *šɔ-cɛ* ‘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 twenty-five 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. *lɔ̄lɔ̄* ‘lorry’; ‘truck’), or because of missionary activity (e.g. *kɔ̄mɔ̄tɔ̄* ‘committee’, *nòʔ* ‘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 *šɔ̄* ‘round object’ as their final element. Sometimes these are new combinations of native Lahu morphemes: *à-mɔ̄-šɔ̄* ‘hand grenade’ (fire ball); *ð-fɔ̄-qɔ̄-šɔ̄* ‘stomach ulcer’ (stomach ball); *cɔ̄-gɔ̄-šɔ̄* ‘tonsils’ (spit-spheres). Sometimes they are blends of *šɔ̄* with Shan or Thai roots, e.g. *màʔ-pāw-šɔ̄* ‘coconut’ (1st element < Thai *məphráaw*); *šɛ(n)-šɔ̄(a)* ‘jewel’; ‘precious stone’ (b) ‘lightbulb’ (1st element < Shan *s’ɛŋ* ‘precious’); or blends with Burmese morphemes (via Shan), e.g. *dáʔ-šɔ̄* ‘lightbulb’ (1st element < Bs, ult. < Pali/Skt *dhātu* ‘element’; ‘primitive matter’); *yā-cū-šɔ̄* ‘weight’ (< Bs *ra-ju* ‘scales’). Most recent are blends with English syllables: *bɔ̄-šɔ̄* ‘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:

*mɔ̄-chwɛ chɛ̀ tù ve yò*  
 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-hɛ̀ tɛ̀ qhàʔɔ̄</i>	<i>šā</i>	<i>qhɔ̄-qhe-phɔ̄</i>	<i>gàʔ</i>	<i>câ</i>	<i>e</i>	<i>tù</i>	<i>le</i>
today	our village	game	where	hunt	eat	PRT	PRT	PRT
NP1	NP2	NP3	NP4		VP			
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 *câ* '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 *mâ*. 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 PUs may be further subclassified according to the types of clause in which they appear: (a) Non-final unrestricted particles (Punf) occur only in non-final clauses: e.g. *qo* 'if'; *thô* 'even', 'also'; *le* '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. *ɔ̄ câ ve* 'to eat rice'; *nâ? tâ? ve* 'to carry a gun'.<sup>7</sup> Lahu does have an object-marking particle, *thà?*, 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 (thà?) phu tân ve* 'offer money to the priest'. The particle *thà?* 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. *í-kâ? gò ve* 'water is cold'.
- d instrument + verb, e.g. *í-kâ? chî ve* 'wash with water'.
- e location + verb, e.g. *á-ghɔ chê ve* 'stay at home', *í-kâ? pā-tí? ve* 'sink into water'.

<sup>7</sup> 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.

- 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? hê ve* 'bathe' ('water-bathe'), *i-kâ? ší ve* 'be thirsty' ('water-thirsty'). The verbs *hê* 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'.<sup>8</sup>
- h physical characteristic + te, e.g. *ð-qhè? te ve* 'have ridges' ('do ridges'), *ð-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, ɔ, lɔ*. *ɔ* and *lɔ* show more emotional involvement than the neutral (and most common) *yò*, e.g. *Lâhū-yâ yò* 'He's a Lahu'. *Lâhū-yâ ɔ* 'He's a Lahu, of course'. *Lâhū-yâ lɔ* 'Gee, he's a Lahu!'
- 2 Dubitative: *hê, nè-ɔ*, e.g. *Lâhū-yâ hê* 'I guess he's a Lahu'. *câ ò nè-ɔ* 'I suppose he's eaten by now'.
- 3 Persuasive: *mē* (sense of urging, often requesting assent): *qha-bû? câ 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. *ɔ câ ò lâ* 'Have you eaten already?' Yes-no questions may be disjunctive: *nð qay lâ mâ qay lâ* 'Are you going or not?' (*lit.* 'You go? You not go?').
  - b *lê* 'request for assent; tag questions': e.g. *ɔ câ ò lê* 'You've eaten, haven't you?'
  - c *nā* 'rhetorical or indirect questions': e.g. *yð kâ? qay tù 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-šū le* 'Who is it?'; *à-thò? ma le* 'What is it?'
- 5 Quotative: *cê* 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-šwè thâ ð-chð nî gâ cò ve cê* 'Once upon a time there were two friends (it is said)'. *cê* may be ordered differently with respect to other final particles, with change of meaning. When *cê* appears in the same clause as an interrogative particle (*lâ, le*), it marks a quoted question; but according to whether *cê* follows or precedes the interrogative particle the sentence is interpreted as a direct or an indirect quoted question:
 

*nð ve lâ cê* He said, 'Is it yours?' / *nð ve cê lâ* Did he say it was yours?
- 6 Interjectory (to show emphasis or emotion): *è?, ma, và, nē, yâ, lè?, qò? ma, qò?-lè?*

In rapid colloquial Lahu these particles appear in practically every sentence, sometimes several in a row in a single clause: *yà?-tɔ-phð kà? mâ ší ò yâ ma nē* 'She's absolutely shameless about it now!' (*lit.* 'She does not even know the way of shame any more').

<sup>8</sup> This could also have the object + verb interpretation: *ú-gê te ve* 'make a pillow'.

<sup>9</sup> 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-ni* 'how many?', *qhà-ma* 'how much?', *qhà-hi* 'how big?', *qhà-ši* 'how long?', *qhà-ft* 'how far?', *qhò* 'where?', *qhò-phð* 'what direction?'

## 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-gà yàʔ-qɔ jû le## hâ jâ ve yò*<sup>10</sup>  
 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-gà yàʔ-qɔ jû ve pa-tɔ## hâ jâ ve yò*  
 ‘Since he has been walking all day on the road, he’s very tired.’

c *thô* ‘concessive’

*yô tê ni qha-gà yàʔ-qɔ jû thô## mâ hâ šē*  
 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-gà yàʔ-qɔ jû thâ## hâ jâ ve yò*  
 ‘When he’s been walking all day on the road, he’s very tired.’

e *qo* ‘conditional’

*yô tê ni qha-gà yàʔ-qɔ jû qo## hâ jâ ve yò*  
 ‘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 *-hɛ*, or dual suffixes like *-hí-mà* or *-hí-nè*. There is also an impersonal third person pronoun which cannot be pluralized: *šú* ‘remote or contrastive third person’; ‘they’; ‘others’.<sup>11</sup>

<i>ŋà</i>	I	<i>ŋà-hɛ</i>	we	<i>ŋà-hí-mà</i>	both of us
<i>nò</i>	you	<i>nò-hɛ</i>	you all	<i>nò-hí-mà</i>	both of you
<i>yô</i>	he/she	<i>yô-hɛ</i>	they	<i>yô-hí-mà</i>	both of them

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.

10 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.

11 The same morpheme occurs in the interrogative pronoun *a-šú* ‘who?’.

## 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 tē gā* ‘a/one person’; *pā-vī nī khē* ‘two civet cats’. There are eleven numerals, the numbers from 1 to 9, the interrogative ‘how many?’, and a Shan-derived word for ‘several’ (10, 100, 1000 and other ‘round numbers’ are classifiers, not numerals):

*tē* ‘one’; *nī* ‘two’; *šē* ≠ *šēʔ* ‘three’; *ṣ* ‘four’; *ṇā* ‘five’; *khḁʔ* ‘six’; *šī* ‘seven’;  
*hī* ‘eight’; *qḁ* ‘nine’; *qhà-nī* ‘how many?’; *láy* [<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 auto-classifiers: some nouns take themselves as classifiers, e.g. *yè tē yè* ‘a house’ (‘house one house’); *qhāʔ nī qhāʔ* ‘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; *khē* for animals; *kà* for places; *cè* for plants; *péʔ* for fields; *qōʔ* for books or papers; *qhḁ* for elongated objects; *cāʔ* for stringlike objects; *šī* for round objects.
- c measure classifiers: e.g. *í-kāʔ tē líʔ* ‘a/one liter of water’; *là tē khē* ‘a/one cup of tea’.
- d time classifiers: these usually occur without a head noun, e.g. *tē ha-pa* ‘one month’; *ṣ nī* ‘four days’; *hī qhḁʔ* ‘eight years’.
- e group classifiers: these appear only with the numeral *tē* ‘one’; ‘whole’, e.g. *Lāhū tē phā* ‘all the Lahu’; *ṣ-qā tē gī* ‘a bunch of buffalo’; *Héʔ-pā tē 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 *cè* ‘kind’: *yè nī mà* ‘two houses’; *yàʔ-qḁ tē mà* ‘a/one road’; *nāʔ-chī tē cè* ‘a/one kind of medicine’.
- g round-number classifiers: all round numbers are classifiers. The way to say ‘ten’ is *tē chī*, i.e. ‘one ten-unit’, just as ‘500’ is *ṇā ha* ‘five hundred-units’. The higher round numbers are borrowings from Shan: *hé(n)* ‘thousand’; *mā(n)* ‘ten thousand’; *šé(n)* ‘hundred thousand’; *lā(n)* ‘million’.<sup>12</sup>

## 7.3 The object noun particle *thàʔ*

As an independent noun prefixed by *ṣ-*, or a constituent in noun compounds, Lahu *thàʔ* (<PTB *\*l-tak* ≠ *\*g-tak* ‘ascend’; ‘above’) means ‘upper surface’; ‘top part’. As a noun particle, *thàʔ* (often reduced to *hàʔ* or even *àʔ*) 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, *thàʔ* 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 *thàʔ*, the following sentence could be interpreted as ‘I have given (someone) that book’:

*lìʔ chī ṇà thàʔ pī tā ve yò*  
 book this 1sg OBJ give PV Pu Puf  
 ‘(Someone) has given me that book.’

12 These higher numerals all have final nasal consonants in Shan, and many Lahu speakers still pronounce them with nasalized vowels.

In complex sentences, *thàʔ* may also mark a whole clause which functions as the object of another clause:

{*yô qòʔ la ve*} *thàʔ* *ɲà dɔ-lɔ 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): *ɲà ve ð-chɔ* 'my friend'; *Lâhū-yâ ve mû-mì* 'the country of the Lahu'; *phî ve mē-tu* 'a dog's tail'. Many possessive constructions may be shortened by omitting *ve*, becoming more like unitary noun compounds: *vàʔ ve ð-šā* 'the flesh of a pig' vs *vàʔ-šā* 'pig flesh'; 'pork'. The *ð-* 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 ð-pa*, *nò ð-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', *ò* 'there', *cò* 'way over there'; 'yonder', *nò* 'up there', *mò* '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, *yè ð-qhɔ* 'in(side) the house', *yè ð-hɔ* 'under (neath) the house', *qhâʔ ð-pâ-nê* 'near the village', *he ð-qɔ-ji* 'the middle of the swidden'. Other spatial nouns are *ð-qhò* 'above'; 'over'; 'on top of' (but not touching); *ð-qhòʔ-nɔ* 'in back of; behind'; *ð-thàʔ* 'on top of; above and touching'; *ð-na* 'in front of and above'; *ð-bà* 'outside'; *ð-ǵū-šī* 'in front of'; 'before'.
- c Locative noun particles: Lahu has several noun particles of general locative meaning (*kàʔ*, *ɔ̄*, *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:

*há-qō lo mɛ chê ve*  
 cave sit PROG Pu  
 'He's sitting in the cave.'

*há-qō lo lòʔ e ò*  
 cave enter PV PV  
 'He has already gone *into* the cave.'

*há-qō lo tɔʔ e ò*  
 cave emerge PV PV  
 'He has already come *out of* 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





### 8.1 Verb particles (Pv)

There are more than twenty 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: *dàʔ* ‘reciprocal’, ‘mutual’: *dəʔ dàʔ* ‘strike one another’; *və* ‘transportatory motion’: *fá və* ‘take something and hide it’; *e* ‘motion away’: *qəʔ e* ‘go back to’; *la* ‘motion towards’: *qəʔ la* ‘come back to’; and *lâ* ‘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: *pə̄ la* ‘be almost finished’; *pə̄ e* ‘be finished’, ‘all used up’.
- 2 Experiential: this group of Pvs express subjective attitudes towards the nature of one’s experience. They include: *qhɛ* ‘excessive repetition’; *gâ* ‘desiderative’; *jɔ* ‘experiential’; and *à* ‘asseverative’.
- 3 Aspectual: these Pvs include: *tā* ‘durative’; ‘perfective’; *tù* ‘future’; ‘hypothetical’; ‘purposeful’; *sā* ‘durative’; *sē* ‘inchoative’; and *ò* ‘completed action’; ‘change of state’.
- 4 Imperative: e.g. *a* ‘mild imperative or suggestion’; *sā* ‘intended action of the first person’; *yà* ‘brusque imperative’; *vì* ‘hortatory’; *lò* ‘urge someone with insistence’; *-ʔ* ‘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>lâ</i>	<i>pəʔ</i>	<i>chəʔ</i>	<i>câ</i>	<i>pə̄</i>	<i>sē</i>	<i>ve</i>	<i>cê</i>
tiger	jump	bite	eat	finish	Pv	Pu	Puf
	v <sup>1</sup>	v <sup>2</sup>	v <sup>3</sup>	v <sup>4</sup>			

‘The tiger jumped (out) and bit (into them) and ate (them) all up!’

The fourth verb, *pə̄* ‘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 versatile as Vv. In the following example of a five-verb concatenation, all four of the versatile verbs follow the verb-head:

<i>cì</i>	<i>gə̄</i>	<i>təʔ</i>	<i>mā</i>	<i>pî</i>	<i>cə̄</i>
	vh	Vv	Vv	Vv	Vv
tooth	pull	emerge	teach	give	correct

‘(They) ought to show them how to pull out teeth.’

The verbs *təʔ*, *mā*, *pî*, and *cə̄*, which as main verbs mean ‘emerge’, ‘teach’, ‘give’, and ‘be correct’, have much more abstract meanings as versatile verbs, translated here by ‘out’ (*təʔ*), ‘show how’ (*mā*), ‘them’ (*pî*), and ‘ought’ (*cə̄*).

In the following concatenation, one *vV* precedes the *Vh*, while three *Vvs* follow it:

<i>ŋà-hí</i>	<i>ga</i>	<i>yù</i>	<i>qay</i>	<i>cí</i>	<i>cɔ̄</i>	<i>ve</i>	<i>lâ</i>
	vV	Vh	Vv	Vv	Vv	Pu	Puf
we	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>dɔ̄</i>	drink	<i>tɔ̄</i>	give to drink
<i>mɔ̄</i>	see	<i>mɔ̄</i>	show
<i>cá</i>	eat	<i>cā</i>	feed someone
<i>vî</i>	be far	<i>fî</i>	separate; demarcate
<i>vàʔ</i>	hide oneself	<i>fá</i>	hide something
<i>tòʔ</i>	be burning	<i>tú</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', *pî* 'give', and especially *cɛ* 'send on an errand', e.g. *yù khá* (vV + Vh) 'make blocked up'; *te qè* (vV + Vh) 'make wide', 'widen'; *ce pî* (Vh + vV) 'cause to fall, drop'; *qay cɛ* (Vh + vV) 'cause to go', *chu cɛ* (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 *pî* 'give' and the Pv *lâ* (< *là* 'come'). The outer-directed *pî* is used to indicate that the action affects a third person, while the inner-directed *lâ* shows that the action affects a non-third person, e.g. *chɔ̄ lâ* (Vh + Pv) 'chop for me/us/you'; *chɔ̄ pî* (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 *í-kâʔhé* 'bathe':

- |                    |   |
|--------------------|---|
| a. <i>pā</i>       | agentive nominalizer; one who V's; a V'er   |
| <i>í-kâʔhé pā</i>  | bather; one who bathes                      |
| b. <i>kî</i>       | locative nominalizer; the place where V     |
| <i>í-kâʔhé kî</i>  | bathing place; bathroom                     |
| c. <i>thā</i>      | temporal nominalizer; the time that one V's |
| <i>í-kâʔhé thā</i> | the time for bathing                        |
| d. <i>tù</i>       | purposive nominalizer; something for V'ing  |
| <i>í-kâʔhé tù</i>  | 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: *í-kâʔhé ve* 'bathing; to bathe'. Frequently *ve* marks a clause that is embedded as the topic of a larger sentence:

*ŋà má qòʔte ve tē chi qhòʔgà ò*  
 '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:

à-šwè-thâ chɔ-ghô nî gâ cò 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):

[yà ?-qɔ jû qay ve] a-pi-qu chi a-šu le  
 Nrh  
 ‘Who is this old lady *who is walking along the road*?’

[nò ð-mî-ma cê tã ve] và ?-ó-qō câ pè ò lâ  
 Nrh  
 ‘Is the pig’s head *that your wife boiled* all eaten up?’

In some cases, the relative clause may be shifted to the position directly *after* the Nrh:

và ?-ó-qō [cê tã ve] mè jâ  
 Nrh  
 ‘A *boiled* pig’s head is very tasty.’

There is occasionally ambiguity between nominalizing and relativizing *ve*:

{tê-ghâ ?-tê-lò šī ve} a-pi-qu chi šī e ve yò  
 ‘What the whole village knows is that the old woman has died.’

[tê-ghâ ?-tê-lò šī ve] a-pi-qu chi šī e ve yò  
 Nrh  
 ‘The old woman whom the whole village knew has died.’

## REFERENCES

- Bradley, David (1979) *Lahu Dialects*, Canberra: Australian National University Press.
- Fei Xiaotong (1990) ‘Ethnic identification in China’, in *Thai-Yunnan Project Newsletter*, No. 11, 11–24. Canberra: Australian National University, Research School of Pacific Studies, Department of Anthropology.
- 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, xxv + 1436 pp, 80 plates.
- Matisoff, James A. (1989) ‘The bulging monosyllable, or the mora the merrier: echo-vowel adverbialization in Lahu’, in J. Davidson (ed.) *South-East Asian Linguistics*, London: School of Oriental and African Studies, 163–97.

- Matisoff, James A. (1992) 'The Lahu people and their language', in Judy Lewis (ed.) *Minority Cultures of Laos*, Southeast Asia Community Resource Center, Folsom Cordova Unified School District (Rancho Cordova, CA), 125–247.
- 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 Tuyền (1973) 'De quelques groupes ethniques arrachés de justesse à l'extinction aux confins de la zone du Nord-Ouest', *Etudes Vietnamiennes* 36, Données ethnographiques (II), 149–200. Section entitled 'Les Co Sung ou La Hu', 176–85.

# LISU<sup>1</sup>

*David Bradley*

## 1 INTRODUCTION

Lisu has about a million speakers: over 650,000 in northwestern Yunnan and southwestern Sichuan provinces, China; nearly 300,000 in Northeastern Burma; 35,000 in northern Thailand, a thousand in Northeastern India, and a couple of hundred in Laos. The very closely related Liphong language is spoken by a further 200,000 in north central Yunnan in China; about a quarter of these are also officially classified as Lisu nationality, while the rest are included in the Yi nationality. In some of the literature the Liphong are referred to as Eastern Lisu. In some of the literature the Liphong are referred to as Eastern Lisu. Over 300,000 further members of the Yi nationality in the same area of north central Yunnan and into southeastern Sichuan call themselves Lolopho, and speak another language extremely similar to Liphong. Thus Lisu in a broader sense, including Liphong and Lolopho, has over 1.5 million speakers. Speakers suggest high intelligibility between Liphong and Lolopho, and considerable intelligibility between both and Lisu; but Lisu speakers cannot understand Liphong and Lolopho. Some other languages, such as Micha and Lamu, are also close to this complex.

The documented migrations of the Lisu took them up the upper Salween and upper Lancang or Mekhong rivers from about 1750, thence into extreme northern Burma from about 1900 and further into Tirap District of Arunachal Pradesh, India from the early 1940s. The Lolopho and Liphong are to their east, mainly along and south of the Jinsha or upper Yangtze river. Southward migration took them into the Shan State of Burma in the nineteenth century and especially after 1950, and further into Thailand starting about 1920. The largest concentration is now in the Nujiang Lisu Autonomous Prefecture and Weixi Lisu Autonomous County of northwestern Yunnan – both areas where they arrived after 1750.

Lisu is divided by its speakers into three subgroups. The northernmost /lo<sup>35</sup> nɛ<sup>44</sup>/ ‘black Lo’ (called /lo<sup>35</sup> wu<sup>55</sup>/ ‘northern Lo’ by other Lisu, Hei Lisu in Chinese and Black Lisu in English) are in the northwest of Yunnan and into extreme northern Burma and India. The central /ɕɿ<sup>44</sup> ɕɿ<sup>44</sup>/ (sometimes called Hua Lisu in Chinese and Flowery Lisu in English) are in western Yunnan and adjacent areas of Northeastern Burma. The southern /lo<sup>35</sup> ʃɿ<sup>33</sup>/ ‘yellow Lo’ are in parts of the Shan State in eastern Burma, the extreme southwest of Yunnan and in Thailand; they are sometimes called Lisaw 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.

---

1 The assistance of many Lisu and other colleagues and friends over the years and the support of the Australian Research Council (A59701122) and of the International Council of Philosophy and Humanistic Sciences of UNESCO is gratefully acknowledged. Thanks for comments from Lisu colleagues, the editors and Takashi Kato; of course all errors are solely the author’s responsibility. Examples are from Flowery Lisu unless otherwise stated.

There are several current Lisu orthographies. The first was for Liphō, and is a Pollard<sup>2</sup> script which is still used by the numerous Liphō 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.<sup>3</sup> This is based on the Flowery dialect, but is universally used by Lisu Christians of whatever dialect background. The third is a syllabary using Chinese characters and other symbols devised by Wang Renpo, a Black Lisu 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*.<sup>4</sup> This was used to some extent in China from the late 1950s, but is now rapidly losing ground to the Fraser script. Hope (1976) provides a Thai script for Southern Lisu, but no one uses this.

## 2 LISU PHONOLOGY

Lisu has the following inventory of consonants

p	t	ts	tɕ	(tʂ)	k	
p <sup>h</sup>	t <sup>h</sup>	ts <sup>h</sup>	tɕ <sup>h</sup>	(tʂ <sup>h</sup> )	k <sup>h</sup>	
b	d	dz	dʒ	(dz <sub>ɿ</sub> )	g	
m	n		ɲ		ŋ	ĩ
f		s	ɕ	(ʂ)	x	
		z	j	(z <sub>ɿ</sub> )	ɣ	
w	l					

There is one Flowery subvariety, the first which Fraser encountered, that distinguishes a complete retroflex series /tʂ tʂ<sup>h</sup> dz<sub>ɿ</sub> ʂ z<sub>ɿ</sub>/ from the alveopalatals /tɕ tɕ<sup>h</sup> dʒ ɕ j/, but only before /a/. Retroflexes and alveopalatals are also distinguished in Liphō and in Lolophō, 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’. In most varieties, the retroflexes are in complementary distribution with the alveopalatals: the retroflexes occur before the fricative syllabic [ɺ], before back vowels /u/ and /o/, and

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 Liphō 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) and Bradley and Bradley (1999). This orthography is a generalized Flowery Lisu standard not really based on any one local speech variety.

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<sup>h</sup>/. Tones are written with a final consonant, such as *-l* for /<sup>55</sup>/ and *-t* for /<sup>21</sup>/.

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 ts<sup>h</sup> dz s z/; in Black Lisu they usually have a following medial /w/ plus /a/. Curiously, some varieties which completely lack other retroflex vs palatal 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 most kinds of Lisu sometimes palatalize to initial [ɲ] in some varieties, but not the reverse: many words are /ɲi/ in all dialects. A non-nasal [h] allophone of /h/ occurs only in the final imperative marker [hɑ<sup>55</sup>]; in /h/-initial words with other tones or vowels the entire syllable including the initial is nasal. The Fraser orthography redundantly provides a separate letter, inverted ‘G’, to represent oral [h].

The /w/ is allophonically [v] or [w]; in most varieties [v] occurs initially before front vowels and [w] occurs elsewhere. The /f/ is also marginal; it occurs exclusively before /u/ or /y/ in native Lisu words, as discussed below. There are also a few areas with all /f/ and most /w/ replaced by alveolar or velar fricatives; but all dialects do have /w/ before /u/ and /o/.

Lisu has the following vowel system in most areas.

i	y	u	u
e	ø	ɤ	o
ɛ		a	

There are also fricative syllabics [ɲ] and [ɲ] which occur after homorganic affricates and fricatives: [tsɲ ts<sup>h</sup>ɲ dzɲ sɲ zɲ] and [tʃɲ ts<sup>h</sup>ɲ dzɲ ʃɲ zɲ] respectively. In a few dialects, there are also sequences of [tsɲi tsɲ<sup>h</sup>i dzɲi sɲi zɲi], so these could be treated as an additional set of initial consonant clusters, unless the /ɲ/ is recognized as an additional vowel and [tsɲi] and so on are treated as /tsi/, with [tsɲ] and so on then treated as /tsɲ/. The varieties on which the various Lisu orthographies are based do not have sequences like [tsɲi], so the orthographies treat the [ɲ] 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 a distinction must be made; either a vowel /ɲ/ is needed, occurring only after retroflexes; or five distinct retroflexes are needed. The Fraser orthography recognizes five retroflex initials in any case, since it also distinguishes these initials before /a/; it treats [tʃɲ] and so on as sequences of retroflex plus /y/, and of course [tʃi] and so on as /tʃi/. The Chinese orthography also distinguishes retroflexes, partly in order to represent the many Chinese loans, and treats [ɲ] as another allophone of /i/, again following *pinyin*. But Black Lisu, which it represents, does not have retroflexes before /a/ in native words, except with medial /w/.

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 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 /t<sup>h</sup>ø<sup>33</sup>/, but usually /t<sup>h</sup>e<sup>33</sup>/ in Southern and Black Lisu; ‘child’ is Flowery, Southern, and sometimes Black Lisu /zɑ<sup>21</sup> nø<sup>33</sup>/, but also sometimes Black /zɑ<sup>21</sup> ne<sup>33</sup>/. Thus the Chinese script is somewhat underdifferentiated in failing to distinguish the front rounded vowels. In most kinds of Lisu, /y/ occurs only after bilabial, alveolar stop, nasal and lateral and /h/ initials and /w/, and never after any affricate, fricative, or velar initials.

There is some dialectal fluctuation, especially in Black, fronting /a/ to /ɛ/ after alveopalatals and to a lesser extent after dental affricate and fricative initials; thus, for example, BY: /bjɑ<sup>21</sup>/ ‘bee’ is also /bjɛ<sup>21</sup>/ in some areas.

The high and mid-back unrounded vowels /uɤ/ do not occur syllable-initially; they must be preceded by a voiced velar fricative [ɣ] if no other consonant precedes. The velar fricative is otherwise only found before /a/, and some dialects including Southern replace some or all of these [ɣa] sequences with [ja], thus eliminating the need for postulating a /ɣ/ at all. The Fraser script writes both [uɤ] after a consonant and the syllable [ɣuɤ] with inverted 'L', and [ɤ] after a consonant and the syllable [ɣɤ] with inverted 'D'. The distinction between /u/ and /ɤ/ is highly marginal, but those literate in the Fraser script are now accustomed to it. The Chinese orthography does not distinguish the two; both are written with *e*.

In Southern Lisu and in some other areas the vowel /u/ conditions a preceding labiodental fricative after a consonant, [f] after voiceless initials and [v] after voiced ones; for example, /tu/ is [tɸu] and /du/ is [dɸu]. The vowel /e/ conditions a preceding [j]; for example /de/ is [dje]. All the mid vowels, /e ø ɤ o/, are variably raised to [i y u u]. Often the raised /e/ retains its glide, becoming [ji] rather than [i]. In the case of underlying /u/ vs /o/ these are also usually still distinguished by the labiodental: /u/ is [ɸu] or [vɸu], while /o/ is [o] or [u]. A further related complication is that 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 /ge<sup>33</sup>/, but is consistently written as *ggi* /gi<sup>33</sup>/ in the Chinese *pinyin* for Black and as JE., /dze<sup>33</sup>/ in written Flowery Lisu, reflecting the most frequent pronunciations.

It could be argued that some initials are in complementary distribution with initial glottal stop. There are no native syllables with initial [f] except before /u/ or /y/, and [ɸu] is the allophone of /u/ after a voiceless consonant in some dialects; so we could say that initial /u/ and /y/ (or /u/ and /y/ preceded by a glottal stop) are [ɸu] and [ɸy]. Similarly, initial [ɣ] before /u/ and /ɤ/ and initial [j] before /i/ and /e/ could be treated as automatic allophonic initials, since these vowels also do not occur initially, other than in some baby talk, onomatopoeic, or loanwords. This would completely eliminate the need for an /f/ phoneme in native words and greatly reduce the distribution of /ɣ/ and /j/. As we have seen, the Fraser orthography does treat [ɣ] as an allophone preceding an initial vowel with /u/ and /ɤ/, but /f/ and /j/ are written as such. The Chinese orthography writes /f/ as *f*, but writes initial /ɣ/ before /a/ and /j/ before /i/ both as initial *e*; thus 'he' /ji<sup>55</sup>/ is written *eil* and 'rock' [ɣa<sup>55</sup>] is written *eal*. Initial [ɣuɤ] and [ɣɤ] are both written with *e* representing both initial and vowel, as in [ɣu<sup>55</sup>] '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 /wɨ/. Flowery Lisu has [y] with lip projection rounding, but the degree of rounding differs from area to area. In Black Lisu it is central [ɥ] in some varieties, but completely merged into /u/ in others.

Lisu has six tones. They are high level /<sup>55</sup>/, rising /<sup>35</sup>/, higher-mid creaky /<sup>44</sup>/, mid non-creaky /<sup>33</sup>/, low falling /<sup>21</sup>/, and low falling with final glottal stop /<sup>21</sup>/. The six tones are numbered 1 to 6 in this order in Fraser (1922), and this is also their alphabetical order in the Flowery 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 /<sup>55</sup>/ 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 /<sup>21</sup>+/<sup>35</sup>/ tones, such as /mɑ<sup>21</sup> mɑ<sup>35</sup>/ 'rice', and also occurs after voiced stop initials in a few very frequent words, especially post-head modal elements such as /dɑ<sup>35</sup>/ '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 dialectal 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 /h/ (but not labiodental) plus medial /j/ before /a/; /tʃa/ also occurs in Chinese loans. Some Black dialects have retroflex affricate and voiceless fricative plus medial /w/ clusters before /a/; like the retroflex allophones before [ɿ], these of course 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/ 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 written TY., since the first page of the very first 1915 catechism. It is pronounced /tʃa<sup>33</sup>/ by non-literate Lisu, but has come to be pronounced /tja<sup>33</sup>/ by Lisu Christians in reading or careful speech.<sup>5</sup> 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.

Conversely, speakers of dialects with quite different canonical sequences are happy to follow the now traditional orthography. For example, all bilabial plus /ja/ sequences are replaced by palatoalveolars in most kinds of Black Lisu. Thus /bja<sup>21</sup>/ ‘bee’ becomes /dʒa<sup>21</sup>/ or /dʒɛ<sup>21</sup>/, and /mja<sup>21</sup>/ ‘many’ becomes /ɲa<sup>21</sup>/ or /ɲɛ<sup>21</sup>/, but Black Lisu Christians still write these as BY: /bja<sup>21</sup>/ and MY: /mja<sup>21</sup>/ . The same convention applies in the Chinese orthography: ‘bee’ is written *bbiat* but never as *bbiait*, *jjat* or *jjait* and likewise ‘many’ is *miat* and never *miait*, *niat* or *niat*.

### 3 LISU MORPHOSYNTAX

#### 3.1 Nominals

Like most other Tibeto-Burman languages, Lisu is verb-final. The order of noun phrases (NPs) before the verb is to some extent pragmatically determined, as Hope (1973b) has pointed out; however the most frequent 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 noun phrase may simply be omitted if it is clear from the context. Overt pronouns for syntactic core NPs are relatively infrequent in running text. 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 /tɛ<sup>55</sup>/ in Black and Flowery or /le<sup>55</sup>/ 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 least frequent on an ACC. For most speakers it may only occur on one NP in a sentence.

5 This is usually written as TY, fused with the following DECL /q̄<sup>44</sup>/.

ji<sup>55</sup>    ŋwa<sup>33</sup> tɛ<sup>55</sup>    ni<sup>33</sup> (ʔtɛ<sup>55</sup>)    t<sup>h</sup>o<sup>21</sup>ɣu<sup>21</sup> (\*tɛ<sup>55</sup>)    gu<sup>21</sup>    tsɿ<sup>33</sup>  
 he/she    I-OBJ    you-OBJ    book-(OBJ)    give    CAUS  
 'He makes me give you a book.'

Much less frequent is the subject marker /ne<sup>33</sup>/, also pronounced [le<sup>33</sup>] in some places, which derives from a form homophonous with COMIT and INST, and has grammaticalized into a non-obligatory ergative marker in Southern Lisu. Postpositions on non-core NPs are much less optional. These include the locative/allative /kwɔ̄<sup>44</sup>/ 'at' and /k<sup>h</sup>u<sup>21</sup>/ 'inside'. In some varieties these have alternative forms with voiced initials; in others there are shortened forms such as Southern /wɔ̄<sup>44</sup>/ or /ɔ̄<sup>44</sup>/ for /kwɔ̄<sup>44</sup>/.<sup>6</sup> The 'from' ABL postposition is /tɕo<sup>55</sup>/. There are also some dialect forms which are completely different, such as Southern /to<sup>55</sup>/ for INST, which is thus distinct from COMIT/ERG /ne<sup>33</sup>/ in that variety.

Possessor NPs precede the NP possessed. The genitive postposition /ɣu<sup>21</sup>/ 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, especially if the possessor has more than one constituent. It is obligatory if there is no overt possessed NP. Another way of expressing possession is with /ji<sup>55</sup>/. 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 below.

/ŋwa<sup>33</sup>    ba<sup>35</sup> ba<sup>21</sup>/    /a<sup>55</sup> mo<sup>21</sup>    (ji<sup>55</sup>)    mu<sup>55</sup>/    /ji<sup>55</sup>    mu<sup>55</sup>/  
 I            father            horse    he/she    tail            PREF    tail  
 'my father'            'horse's tail'                            'tail' or 'his/her tail'

/t<sup>h</sup>ø<sup>33</sup>    mɔ̄<sup>44</sup>    ŋwa<sup>33</sup>    ɣu<sup>21</sup>    ŋa<sup>33</sup>/  
 this    CL    I            GEN    be  
 'This one is mine.'

There is also a topic marker /ŋɔ̄<sup>44</sup>/. It is very frequent on subjects, especially transitive ones, but also occurs on NPs in other case roles. Interestingly, in most varieties of Lisu 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 subjects in /na<sup>21</sup>/.

The demonstrative and quantifier phrase if any must follow the head noun in this order; a demonstrative must be followed by a quantifier phrase, minimally the general classifier /mɔ̄<sup>44</sup>/. The quantifier phrase consists of numbers(s) plus a classifier. When a demonstrative is present, the number 'one' may be omitted. As in many other TB languages, NPs with only a demonstrative or numeral plus classifier, without a head NP, are frequent.

The demonstrative system marks relative height as well as distance: /t<sup>h</sup>ø<sup>33</sup>/ 'this', /gø<sup>33</sup>/ 'that (same level)', /nø<sup>33</sup>/ for 'that (higher level)', and /dzø<sup>33</sup>/ for 'that (lower level)'. Some varieties also have a form /a<sup>55</sup> t<sup>h</sup>ø<sup>33</sup>/ 'that (near you)'; all five are within visible distance. There are dialect differences in the demonstratives for greater distance or beyond vision, but they include modified forms of the three 'that' forms with high tone as well as additional forms including /tɕo<sup>55</sup>/, /dzø<sup>55</sup>/ or /ko<sup>33</sup>/ for 'that (extremely far and invisible)'. All demonstratives also occur fused with the LOC /kwɔ̄<sup>44</sup>/: /t<sup>h</sup>a<sup>33</sup>/ 'here', /gwa<sup>33</sup>/ 'there', /nwa<sup>33</sup>/ 'up there' and so on which are used as independent NPs.

6 The 'inside' form is clearly related to the verbal locative nominalizer /gu<sup>33</sup>/. This locative postposition has cognates elsewhere, including Lahu /kɪ<sup>21</sup>/, Nasu /gu<sup>33</sup>/ and so on; in other languages it is usually just a verbal locative nominalizer, not also a nominal locative postposition as in Lisu.

The numbers follow the usual TB pattern, with higher round numbers effectively functioning as classifiers of a preceding lower numbers.

/a <sup>55</sup> na <sup>21</sup>	sa <sup>21</sup>	tu <sup>33</sup>	tʰi <sup>21</sup>	ḥja <sup>33</sup>	li <sup>55</sup>	ts <sup>h</sup> ɿ <sup>33</sup>	ku <sup>44</sup>	mq <sup>44</sup> /
dog	3	1000	1	100	4	10	9	CL
'3149 dogs'								

In Lisu, even in counting, numbers are usually given with a classifier; if nothing else, the general classifier [mq<sup>44</sup>]. There are tone sandhi forms of the numbers /sq<sup>44</sup>/ '3', /li<sup>44</sup>/ '4' and /ku<sup>44</sup>/ '9'; when the following classifier has tones /<sup>33</sup>/ or /<sup>44</sup>/, a sandhi tone [<sup>55</sup>] occurs for '4' and '9', and usually [<sup>21</sup>] or sometimes [<sub>21</sub>] for '3'. However this sandhi is variable before the general classifier /mq<sup>44</sup>/. '1' has an alternative unaspirated form /ti<sup>55</sup>/ which occurs in '11' /ts<sup>h</sup>ɿ<sup>33</sup> ti<sup>55</sup>/ and as a noun suffix 'sole, only'. '10' has an unaspirated form in '20' /pi<sup>21</sup> tsɿ<sup>33</sup>/.

The general classifier /mq<sup>44</sup>/ is homophonous with the nominalizer/relativizer /mq<sup>44</sup>/ discussed below. It classifies many types of nouns including most animals. Humans are classified with /z<sub>o</sub><sup>44</sup>/ in Flowery, the corresponding form /jo<sup>44</sup>/ in Black, and /zo<sup>33</sup>/ or /wa<sup>21</sup>/ 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 /sɿ<sup>21</sup>/ 'fruit' and /dzɿ<sup>33</sup>/ 'tree'. Some non-bound nouns like /ḥi<sup>33</sup>/ 'house' are also used to classify themselves; in such a case the head noun is often omitted.

/ḥi <sup>33</sup>	li <sup>55</sup>	ḥi <sup>33</sup> /
(house)	4	CL
'four houses'		

One interesting set of two-syllable classifiers, which is also found in closely related languages such as Lalo (see Björverud 1998), is for groups of a parent or grandparent plus one or more children. The forms are /ma<sup>55</sup> la<sup>21</sup>/ for mother and children, /pa<sup>55</sup> la<sup>21</sup>/ for father and children, and /pi<sup>55</sup> li<sup>21</sup>/ for grandparent or grandfather and children. A fourth form /mi<sup>55</sup> li<sup>21</sup>/ for grandmother and grandchildren is used by some. These occur without a head noun or can follow a pronoun.

/sq <sup>44</sup>	pa <sup>55</sup>	la <sup>21</sup> /
3	CL	
'father and two children' (or in some dialects also 'father, mother and one child')		

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 /(a<sup>55</sup>) mq<sup>44</sup>/; this is prehead or more frequently posthead and then precedes the demonstrative plus quantifier in the NP.<sup>7</sup> The /a<sup>55</sup>/ is often omitted. This also occurs as a general complementizer for other verb types. Unlike the following constructions, an /(a<sup>55</sup>) mq<sup>44</sup>/ construction with a verb usually has an overt nominal head, though such clauses can also occur headless. A sequence of two /mq<sup>44</sup>/, 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 /a<sup>21</sup> k<sup>h</sup>u<sup>55</sup>/ 'very', there is no marking of embedding.

7 There is no obvious meaning difference between preceding and following REL; speakers often reject preceding REL out of context, but they do occur, especially if the REL itself is complex.

/hĩ<sup>33</sup> p<sup>h</sup>u<sup>33</sup> (ɑ<sup>55</sup>)mǣ<sup>44</sup> t<sup>h</sup>ø<sup>33</sup> (t<sup>h</sup>i<sup>21</sup>) hĩ<sup>33</sup>/  
 house white REL this one CL  
 ‘this (one) house which is white’

/p<sup>h</sup>u<sup>33</sup> (ɑ<sup>55</sup>) mǣ<sup>44</sup> (hĩ<sup>33</sup>) t<sup>h</sup>ø<sup>33</sup> (t<sup>h</sup>i<sup>21</sup>) hĩ<sup>33</sup>/  
 white REL house this one CL  
 ‘this (one) house which is white’

/p<sup>h</sup>u<sup>33</sup> (ɑ<sup>55</sup>) mǣ<sup>44</sup> (t<sup>h</sup>i<sup>21</sup>) mǣ<sup>44</sup>/  
 white REL one CL  
 ‘this white one’

/(hĩ<sup>33</sup>) ɑ<sup>21</sup> k<sup>h</sup>u<sup>55</sup> p<sup>h</sup>u<sup>33</sup> t<sup>h</sup>ø<sup>33</sup> (t<sup>h</sup>i<sup>21</sup>) hĩ<sup>33</sup>/  
 house very white this one CL  
 ‘this (one) house which is very white’

Stative extensive 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 /ǣ<sup>44</sup>/ parallel to the substance question words; diminutive forms with a following /e<sup>55</sup>/; and a variety of reduplicated augmentative forms. Six to nine such verbs pattern this way in different varieties of Lisu; Liphò have eight to ten.

/wu<sup>21</sup>/ /wu<sup>33</sup>/ /ǣ<sup>44</sup> wu<sup>33</sup>/ /ǣ<sup>44</sup> wu<sup>33</sup> e<sup>55</sup>/ /ǣ<sup>44</sup> wu<sup>55</sup> wu<sup>33</sup> be<sup>33</sup>/ /ǣ<sup>44</sup> wu<sup>55</sup> wu<sup>33</sup> be<sup>33</sup>/  
 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 /sǣ<sup>44</sup>/ ‘(animate) subject’, /sɿ<sup>21</sup>/ ‘(non-human, usually) object’, /gu<sup>33</sup>/ ‘locative’, /du<sup>33</sup>/ ‘instrument’ and /t<sup>h</sup>ε<sup>21</sup>/ ‘temporal’ as well as the general /mǣ<sup>44</sup>/. 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. These nominalizations are frequent without a head NP and thus look like complements; but sometimes occur with a head NP, which usually follows; if it precedes, it is hard to distinguish from an internally-headed nominalization unless there is a pause. What seems not to 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.

/(ŋwa<sup>33</sup>) lɑ<sup>33</sup> sǣ<sup>44</sup> tε<sup>55</sup> t<sup>h</sup>o<sup>21</sup>ɣu<sup>21</sup> gu<sup>21</sup> ǣ<sup>44</sup>/  
 I come NOMZR OBJ book give PFV  
 ‘I gave the person who came/is coming 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.

/hĩ<sup>33</sup> p<sup>h</sup>u<sup>33</sup>/  
 house white  
 ‘white house’

/mi<sup>33</sup> nǣ<sup>44</sup>/  
 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.

/dza <sup>21</sup>	du <sup>33</sup> /	/dza <sup>21</sup>	du <sup>33</sup>	do <sup>33</sup>	du <sup>33</sup> /
eat	NOMZR	eat	NOMZR	drink	NOMZR
'food'		'food (and drink)'			

A further kind of nominal form is with the prefix /ji<sup>55</sup>/. It 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' as we have seen above. It is also frequently used as an abstract nominaliser of a stative non-extensive verb.

/ji <sup>55</sup>	mjə <sup>44</sup> /	/ji <sup>55</sup>	mjə <sup>44</sup> /	/ji <sup>55</sup>	nɛ <sup>44</sup> /
he/she	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 a medial /be<sup>33</sup>/. A string of conjoined nouns may be followed by a single /be<sup>33</sup>/; this is most likely if there are three or more. In principle this /be<sup>33</sup>/ may also occur between the nouns and after them, but this is infrequent and rejected by many speakers. Conjunction of animates can also use COMIT/INT /le<sup>33</sup>/ 'with', medially or finally. A third nominal conjunction found in some areas is /tɕe<sup>33</sup>/.

The pronoun forms /ŋwa<sup>33</sup>/ 'I' /nu<sup>33</sup>/ 'you' /ji<sup>55</sup>/ 'he/she' are used to refer to humans. The system includes a third person remote 'other' category, for which the form is /sɯ<sup>44</sup>/; it is of course not an accident that this is homophonous with the animate nominalizer suffix. For second and third person forms there is an optional plural suffix /wa<sup>21</sup>/; these also sometimes fuse to single-syllable forms /na<sup>21</sup>/ from /nu<sup>33</sup> wa<sup>21</sup>/ 'you (pl)' and /ja<sup>55</sup>/ from /ji<sup>55</sup> wa<sup>21</sup>/ 'they'. This suffix does not occur after /sɯ<sup>44</sup>/ or /ŋwa<sup>33</sup>/. There is an inclusive first plural form /zə<sup>21</sup>/ or sometimes /zo<sup>21</sup>/ in Flowery, /jo<sup>21</sup>/ in Black, and /a<sup>21</sup> zo<sup>21</sup>/ in Southern Lisu. The form /ŋwa<sup>33</sup> nu<sup>21</sup>/ or its song form equivalent /ɿ<sup>44</sup> nu<sup>21</sup>/ is the first plural exclusive; note the tone difference from the 'you' pronoun.

A final pronoun is the reflexive. There are dialect differences, but the most widespread construction is an invariant reflexive pronoun /tɕi<sup>55</sup> tɕ<sup>h</sup>a<sup>21</sup>/ 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 Liphon, which has a construction more frequently encountered outside Lisu: pronoun + reflexive marker + pronoun. Lisu also has such a construction, but with a non-reflexive meaning: pronoun + /da<sup>33</sup>/ + pronoun 'by oneself'. Another reflexive possibility is pronoun + /go<sup>33</sup> dɔ<sup>21</sup>/ 'body' or other body part noun + pronoun, but only where the action physically affects one's body. The invariant /tɕi<sup>55</sup> tɕ<sup>h</sup>a<sup>21</sup>/ 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: /mɿ<sup>44</sup>/ 'female' and /p<sup>h</sup>a<sup>21</sup>/ 'male', as in /ma<sup>55</sup> mɿ<sup>44</sup>/ 'female teacher' and /ma<sup>55</sup> p<sup>h</sup>a<sup>21</sup>/ 'male teacher', both from the verb /ma<sup>55</sup>/ 'teach'. These are alternatives to the non-gender-specific /sɯ<sup>44</sup>/ discussed above, but unlike it they can also be added to nominals, even some of those which already contain /sɯ<sup>44</sup>/, as in /li<sup>44</sup> sɯ<sup>44</sup> mɿ<sup>44</sup>/ 'Lisu woman', /li<sup>44</sup> sɯ<sup>44</sup> p<sup>h</sup>a<sup>21</sup>/ 'Lisu man'; but not in Liphon where these are /li<sup>33</sup> ma<sup>44</sup>/ and /li<sup>33</sup> p<sup>h</sup>a<sup>21</sup>/. Note that the female suffix is homophonous with both the general classifier and the nominalizer.

### 3.2 Verbals

There is no agreement with nominals of any kind marked on Lisu verbals. There is also no switch reference or other clause sequencing marking. Reflexive is also marked entirely by the nominals. A morphosyntactically distinct passive does not exist, though the order of NPs is pragmatically determined. There are some clause-final markers with evidential-like meanings, such as the quotative /dzo<sup>21</sup>/ ‘someone said’.

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 /wɑ<sup>44</sup>/ is a verb meaning ‘get’, a pre-head meaning ‘get to’, ‘have the opportunity to’ and a post-head meaning ‘must’.

Post-head elements are quite numerous. They include a wide range of modal meanings: /da<sup>33</sup>/ ‘can’, /wɑ<sup>44</sup>/ ‘must’, /tʃo<sup>35</sup>/ ‘need to’, /tʃo<sup>44</sup>/ ‘experiential’, /da<sup>35</sup>/ ‘durative’, /tʃa<sup>33</sup>/ ‘continuous’, /ɲi<sup>44</sup>/ ‘try to’ and so on. The post-head modal ‘want to’ has two syllables and also varies: /ni<sup>35</sup> ʂɿ<sup>33</sup>/ or /ɲi<sup>35</sup> ʂɿ<sup>44</sup>/; the first syllable is the noun ‘heart’. Some do have full verb uses, such as /wɑ<sup>44</sup>/ ‘get’, /da<sup>35</sup>/ ‘be at’, /tʃa<sup>33</sup>/ ‘stay’ and /ɲi<sup>44</sup>/ ‘look at’; others do not.<sup>8</sup> 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 modal in a non-final clause and with another modal 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, including post-heads which have a homophonous main verb.

As in many TB languages, there are lexical causatives: pairs of etymologically related intransitive and transitive or transitive and ditransitive forms. In Lisu, the causative have voiceless initials and the non-causative forms have voiced initials, as in /dzɔ<sup>44</sup>/ ‘fear’, /tʃo<sup>35</sup>/ ‘scare’ or /dza<sup>21</sup>/ ‘eat’, /tsɑ<sup>55</sup>/ ‘feed (an animal)’. There are often also tonal and semantic differences. In addition to these, there is a post-head causative /tsɿ<sup>44</sup>/ with various alternative dialect forms. Furthermore, /gu<sup>21</sup>/ from the verb ‘give’ is also used post-head in a benefactive meaning,<sup>9</sup> sometimes with a causative result. As a pre-head it is a permissive causative. A lexical causative can also be followed by the postverbal causative. 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 this order.

/gu <sup>21</sup>	dza <sup>21</sup> /	/dza <sup>21</sup>	tsɿ <sup>44</sup> /	/tsɑ <sup>55</sup> /	/tsɑ <sup>55</sup>	tsɿ <sup>44</sup> /	/gu <sup>21</sup>	dza <sup>21</sup>	tsɿ <sup>44</sup> /
P-CAUS	eat	eat	CAUS	feed	feed	CAUS	P-CAUS	eat	CAUS
‘give to eat’		‘cause to eat’	‘feed’	‘cause to feed’		‘cause to give to eat’			

8 For example, the main verb ‘want’ is /nu<sup>33</sup>/ in Black, /ny<sup>33</sup>/ in Flowery, and /nɔ<sup>33</sup>/ in Southern Lisu.

9 For the benefit of any human of whatever person.

/mɑ <sup>55</sup>	ɡw <sup>21</sup> /	/mɑ <sup>55</sup>	ts <sub>1</sub> <sup>44</sup> /	/mɑ <sup>55</sup>	ts <sub>1</sub> <sup>44</sup>	ɡw <sup>21</sup> /
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 markers, and often follow other postverbal elements. The reciprocal, Flowery /lɛ<sup>21</sup> k<sup>h</sup>o<sup>33</sup>/, Black and Southern Lisu /lɛ<sup>21</sup> xo<sup>33</sup>/ forms part of the directional system, the rest of which is derived from motion verbs. These include /la<sup>33</sup>/ ‘motion towards/come’, /ge<sup>33</sup>/ (with various raised and /or palatalized dialectal forms including Flowery and Southern [dʒɛ<sup>33</sup>]) ‘motion away’/‘go’ and the slightly less transparent forms /je<sup>33</sup>/ ‘motion away’ and /le<sup>33</sup>/ ‘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 /ɑ<sup>44</sup>/; perfective /ɑ<sup>44</sup>/, with a transitive perfective alternative /kɣ<sup>44</sup> ɑ<sup>44</sup>/; or habitual/generic /lɑ<sup>44</sup>/, In Southern and some kinds of Flowery Lisu, the intransitive perfective is very often preceded by a directional (/le<sup>33</sup> ɑ<sup>44</sup>/, /je<sup>33</sup> ɑ<sup>44</sup>/), but in other varieties such sequences are not preferred or even ungrammatical.

Some copula sentences have no overt copula; just a sequence of two nominals. Alternatively, there can be an overt final copula /ŋɑ<sup>33</sup>/ which becomes obligatory when negated; when non-negated, this can be followed by the HAB /lɑ<sup>44</sup>/, or with these two fused to /ŋɑ<sup>33</sup>/. These copula forms cannot be combined with other pre- or post-head verbal elements. Another copula-like form is /p<sup>h</sup>jø<sup>21</sup>/ ‘become’, but unlike /ŋɑ<sup>33</sup>/ it is a normal verb and can be combined with modal, directional, and non-habitual aspect postverbal elements.

/ŋwɑ <sup>33</sup>	mɑ <sup>55</sup> mɑ <sup>44</sup>	(ŋɑ <sup>33</sup> )/	/ŋwɑ <sup>33</sup>	mɑ <sup>55</sup> mɑ <sup>44</sup>	mɑ <sup>21</sup>	ŋɑ <sup>33</sup> /
I	teacher-female	be	I	teacher-female	NEG	be
‘I am a (female) teacher.’			‘I am not a (female) teacher.’			

/ji <sup>55</sup>	mɑ <sup>55</sup> mɑ <sup>44</sup>	wɑ <sup>44</sup>	p <sup>h</sup> jø <sup>21</sup>	tɕɑ <sup>33</sup>	ɑ <sup>44</sup> /
he/she	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 /mɑ<sup>44</sup>/, which is often used as a complementizer without a head noun. Other complement types include non-final clauses embedded by a final /ɑ<sup>44</sup> be<sup>33</sup>/, and with certain final verbs involving human action a non-final clause with a verb and then /bɛ<sup>44</sup>/ (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 /ji<sup>33</sup>/ ‘day’ seen in /ji<sup>55</sup> me<sup>33</sup>/ ‘today’ or verbal elements nominalized by the clause-final /t<sup>h</sup>ɛ<sup>21</sup>/ as we have seen. There is one very frequent intensifier adverb /ɑ<sup>21</sup> k<sup>h</sup>w<sup>55</sup>/ ‘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 candidate for adverb status in Lisu is a small but open class of reduplicated two-syllable onomatopoeic forms. A similar set of manner adverbials has a prefix /tɕeŋ<sup>33</sup>/ (varying between [tɕeŋ<sup>33</sup> tɕe<sup>33</sup> tɕiŋ<sup>33</sup> tɕi<sup>33</sup>]) preceding a partially or fully reduplicated two-syllable form with at least the same vowel and tone in both syllables. This prefix may be connected with the nominal conjunction /tɕe<sup>33</sup>/. Dialect differences in adverbial forms are extreme; they usually occur immediately before the verb. Apart from 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 verbals.

### 3.4 Clause types

The final clause has the widest range of possibilities for postverbal elements, as discussed above. Among final clauses, most frequent are declaratives, which most often end in the imperfective final / $\underline{a}^{44}$ / or the perfective final / $\underline{a}^{44}$ /. These aspect markers are usually not found in negatives and negative imperatives, which have the negative / $m\alpha^{21}$ / or negative imperative / $t^h\alpha^{21}$ / immediately preceding the verb or modal to be negated. A positive imperative is most frequently just a bare verb without following marker, but strong imperative markers such as / $h\alpha^{21}$ / exist. In a negative, this verbal element is likely to be the last element in the sentence, though some final particles and directionals may occur after a negated verbal, most frequently / $\tilde{h}e^{21}$ / ‘still’/‘yet’.

/ $dza^{33}$   $dza^{21}$   $\underline{a}^{44}$ /  
 rice eat PFV  
 ‘(I) have eaten.’

/ $dza^{33}$   $m\alpha^{21}$   $dza^{21}$  ( $\tilde{h}e^{21}$ )/  
 rice NEG eat (still/yet)  
 ‘(I) have not eaten (yet).’

/ $dz\gamma^{33}$   $p^h\omega^{21}$   $do^{33}$  ( $h\alpha^{21}$ )/  
 liquor dring (IMP)  
 ‘Drink liquor.’

/ $dz\gamma^{33}$   $p^h\omega^{21}$   $t^h\alpha^{21}$   $do^{21}$ /  
 liquor NEGIMP drink  
 ‘Don’t drink liquor.’

Questions containing a question nominal (‘WH word’) most often have a parallel structure to declaratives, including a final aspect marker if non-negative. Fronting of the question nominal is not necessary, though it is frequent as this is the most topical element in a question. These question elements all start with the prefix / $a$ /, with various tones: / $a^{21}$   $m\alpha^{33}$ / ‘who?’, / $a^{55}$   $\xi^{21}$ / ‘what?’, / $\underline{a}^{44}$   $t^h\varepsilon^{21}$ / ‘when?’, / $\underline{a}^{44}$   $to^{55}$ / ‘where?’ and so on. The question word / $\underline{a}^{44}$   $\underline{e}^{44}$ / ‘which?’ /‘how?’ can occur alone, as in the frequent greeting given below, or with a following nominal element, especially a classifier. ‘Where?’ can also be expressed by ‘which?’ plus LOC / $k\omega\underline{a}^{44}$ / or a contracted form / $\underline{a}^{44}$   $\underline{a}^{44}$ /. It can be noted that some of the second syllables of question words are homophonous with clause nominalizers of similar meaning. Such questions may also have the sentence-final question element / $\underline{a}^{21}$ / fused onto the preceding syllable; this is realized as a falling tone if that syllable ends in / $a$ / with another tone already.

/ $\underline{a}^{44}$   $\underline{e}^{44}$   $\eta\alpha^{33}$  ( $\underline{a}^{21}$ )/ [ $\eta\underline{a}^{31}$ ]  
 how? be Q  
 ‘How are (you/things)?’ (a greeting)

/ $\underline{a}^{44}$   $\underline{e}^{44}$  ( $t^h i^{21}$ )  $m\alpha^{44}$   $\underline{a}^{33}$   $\underline{a}^{44}$  ( $\underline{a}^{21}$ )/ [ $\underline{a}^{41}$ ]  
 which? one CL come IMPFV Q  
 ‘Which one is coming?’



Yes-no questions have a number of alternative structures which differ in frequency and grammaticality between varieties of Lisu. Very frequent is the alternative question: verb+negative+verb, or for some modals verb+modal+negative+modal. As for other negatives, these usually lack most final markers such as directional and aspect. Another widespread possibility is verb+ $\underline{q}^{21}$ , the final question marker we have just seen as an option for substance questions; this is obligatory if it is the sole marker of a Yes-no question, and is most frequent in Southern Lisu. Final  $/la^{21}/$ , which is an earlier form of this final with cognates outside Lisu, is also seen in some Flowery varieties; this can be used finally or combined with the alternative question, in which case it comes after both instances of the verb or modal. There are some other sentence-final question markers for specific situations, such as the first person rhetorical  $/ne^{35}/$ . Some varieties of eastern Black Lisu use a reduplicated verb to form a Yes-no question, but speakers of most other varieties find this ungrammatical.<sup>10</sup>

$/dza^{33} dza^{21} ma^{21} dza^{21}/$   
rice eat NEG eat

$/dza^{33} dza^{21} \underline{q}^{44} \underline{q}^{21}/ [dza^{33} dza^{21} \underline{q}^{41}]$   
rice eat IMPFV Q

$/dza^{33} dza^{21} la^{21}/$   
rice eat Q

$/dza^{33} dza^{21} la^{21} ma^{21} dza^{21} la^{21}/$   
rice eat Q NEG eat Q

$(*)/dza^{33} dza^{21} dza^{21}/$   
rice eat eat  
'Have you eaten?'

$/dza^{33} dza^{21} ne^{35}/$   
rice eat Q  
'Have I eaten?'

Comparisons are formed with  $/s_1^{55} ma^{21} ts_1^{33}/$  'above/more' following the compared NP; so this is structurally like an NP plus postposition.

$/(\underline{j}i^{55}) \eta wa^{33} s_1^{55} ma^{21} ts_1^{33} mo^{21} \underline{q}^{44}/$   
he/she I 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  $/(\underline{q}^{44}) be^{33}/$  type, whose marker is homophonous with a very frequent nominal conjunction, and the  $/be^{44}/$  type, whose marker is the verb 'speak'. As we have seen, there is 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

10 These are the speakers who are in contact with Nosu, which also forms Yes-no questions in this way.

'because' clauses are followed by nominalizer /mǫ<sup>44</sup>/ + noun /pu<sup>55</sup> do<sup>33</sup>/ 'cause' and then the main clause with its final verb.<sup>11</sup>

/mu <sup>21</sup> ĥa <sup>33</sup>	le <sup>44</sup>	la <sup>33</sup>	ǫ <sup>44</sup>	mǫ <sup>44</sup>	pu <sup>55</sup> do <sup>33</sup>
rain	fall	DIR	IMPFV	NOMZR	cause
ng <sup>44</sup> ji <sup>33</sup>	dz <sup>21</sup>	la <sup>33</sup>	ǫ <sup>44</sup> /		
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; after each line there is another line repeating the content with alternative lexical material. Interestingly, the aspect markers /ǫ<sup>44</sup>/ and /ǫ<sup>44</sup>/ do not count as a syllable here. Like most oral literature, songs are very repetitive. Another genre with even more different syntax is proverbs, which are two lines having an equal number of three or more syllables and without postverbal markers.

## REFERENCES

- Bradley, David (1979) *Proto-Loloish*, London and Malmö: Curzon Press.
- Bradley, David (1994) *A Dictionary of the Northern Dialect of Lisu (China and Southeast Asia)*, *Pacific Linguistics* C126.
- Bradley, David (1995) 'Grammaticalisation of extent in Mran-Ni', *Linguistics of the Tibeto-Burman Area* 18/1: 1–28.
- Bradley, David and Bradley, Maya (1999) 'Standardisation of transnational minority languages: Lisu and Lahu', *Bulletin Suisse de Linguistique Appliquée* 69/1: 75–93.
- Bradley, David, Fish, James and Hope, Edward Reginald (2002) *Southern Lisu Dictionary*, Berkeley, CA: Linguistics of the Tibeto-Burman Area.
- Bradley, David and Kane, Daniel (1981) 'Lisu orthographies', *Working Papers in Linguistics, University of Melbourne* 7: 23–38.
- Bradley, David, Lewis, Paul, Jarkey, Nerida, and Court, Christopher (1991/1999) *Hill Tribes Phrasebook: Hill Tribes of South-East Asia*, Melbourne: Lonely Planet.
- Björverud, Susanna (1998) *A Grammar of Lalo*, Lund: Department of East Asian Languages, Lund University.
- 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 (1974) *The Deep Syntax of Lisu Sentences, A Transformational Case Grammar*, *Pacific Linguistics* B34.
- Hope, Edward Reginald (1976) 'Lisu', in William A. Smalley (ed.) *Phonemes and Orthography: Language Planning in Ten Minority Languages of Thailand*, *Pacific Linguistics* C43, 125–48.

11 This is exactly like the structure of various closely related languages including Lahu which has clause + NOMZR /ve<sup>33</sup>/ + /pa<sup>33</sup> tɔ<sup>33</sup>/ 'cause'; the noun form are also cognate.

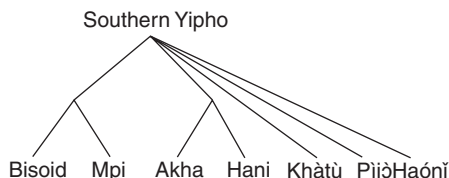
# AKHA

*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 still most of the Akhas 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 lowland 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 7000 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 15.1** RELATIONS AMONG THE SOUTHERN YIPHO LANGUAGES

This scheme does not include Gǎkhý, 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. The texts belong to the offices of *phíma* ‘priest’, *dzòma* ‘village leader’, and *njìphà* ‘shaman’. The texts are recited at 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 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:

				front		mid	back	
				unrounded	rounded	unrounded	unrounded	rounded
p	ph	b	m					
pj	pjh	bj	m					
t	th	d	n	l				
tj	tjh	dj	nj	i	ü		y	u
ts	tsh	dz		e	ø		ə	o
s	sh	z		ɛ				ɔ
sj	sjh	j		m		a		aŋ
k	kh	g	ŋ					
x	xh	ɣ						
h								

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 *-aŋ*, 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>maq</i>	dream
<i>mà</i>	not	<i>màq</i>	group

In Akha, there is both an initial *m-* and a final, syllabic *-m*, which may also be laryngealized:

*í-nm̥q* brain (cf. WB *nok*)

*n̄m-shi* 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 *-ü*, 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:

*míɛ* < *mí nɛ* because

*xə̀qì* < *xə̀q ni* although

*j̄ɔ̀ yà naqù* < *j̄ɔ̀ yà na lùq* 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

*náa* when

*mía* sentence particle, non-past

*mìa* sentence particle, past

*ŋáa* sentence particle, non-past

*ŋàa* sentence particle, past

(For discussion of sentence particles, see below.)

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:

*làq* movement towards first person

*lèq* movement towards non-first person

*nèq* action towards first person

*nè* action towards non-first person

First and second person pronouns change phonation type and/or tone when followed by certain noun particles:

*ŋá* I      *ŋà nɛ* by me      *ŋà áŋ* to me      *ŋàq ə* mine

*nó* you      *nò nɛ* by you      *nò áŋ* to you      *nòq ə* 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      *lá mɛ* > *lá mɛ*      he is coming

Low + Low > Low + Mid      *dzà mè* > *dzà mɛ*      he ate

Depending on the syntactic structure, one or more in a row of syllables on the same tone change.

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 > tʃh	ts > tʃ
dz > dʒ	dz > dʒ
z > ʒ	z > ʒ

but the other way round in:

sjh > sh	sj > s
----------	--------

where there is a loss of palatalization.

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 *-aŋ* 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 go 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>dɔ́</i>	drink
<i>sjhí</i>	die
<i>jùq</i>	sleep
<i>kha</i>	plant
<i>dù</i>	dig
<i>dì</i>	beat
<i>nə̌</i>	think
<i>m̄</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>hɔ̀</i>	cooked rice
<i>m̄</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>ádu</i>	corn
<i>ábjà̀q</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  
*àbjm* head of a plant, *ca bjm-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>khÿ-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>dε-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, *ηà-* for fishes (<*ηà-sjhà* ‘fish’) and *xhà-* for some birds and beasts:

<i>xhà-hm</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>í-luq</i>	billow (< <i>í-tjùq</i> water)
<i>mí-luq</i>	earthquake (< <i>mí-tshà</i> earth)
<i>dzaŋ-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>tjhí-hà</i>	barking deer
<i>bó-hà</i>	bat

*bɔ* – no clear meaning

<i>àbɔ</i>	grandfather
<i>ábɔ</i>	tree
<i>nà-bɔ</i>	ear

A common suffix in the ritual language, *-ba*, is in the modern language lexically restricted to a few spatial nouns, as *xhɔ-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>bì-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>dzɔ-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:

*àjɔq tshɔ-mɔ jɔ-mɔ xhɔ nji ɣà áŋ áshì jɔ-khý jɔ nɛŋ jɔ íshàŋ i bìq í mɔq ɔ nja djé*  
 He old people good those two CLF fruit tasty every day very give go want SP FP  
 ‘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:

*tshɔ-hà jɔ-mɔ xhɔ nji ɣà*  
 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>nɛ</i>	agent, instrument, place from, time from
<i>áŋ</i>	direct object, indirect object, place at or towards, time at
<i>à</i>	topic (subject, object, time, place, clauses)
<i>lé</i>	object (often after verbs like ‘call, name’, enumeration of objects)
<i>ə̀</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:

<i>jàq-the-jà</i>	<i>hə</i>	<i>hy</i>	<i>ɣɛ</i>	<i>ə̀</i>	<i>njèq</i>	<i>ə̀ nɛ</i>	<i>tʃhu</i>	<i>hỳ</i>	<i>ə̀ nɛ</i>	<i>kuq</i>	<i>áŋ</i>
young wild boar	this	big	be	VP	catch	VP	raise	big	VP	pen	NP
<i>mà</i>	<i>lány</i>	<i>ni ə̀</i>	<i>bi</i>	<i>dʒə̀</i>	<i>hə</i>	<i>shá ə̀</i>	<i>tʃɛq</i>	<i>bjoq</i>	<i>í</i>	<i>má</i>	
not	keep	P	VP	let	stay	try	please	run	disappear	away	SP

‘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 *nɛ* in the past tense (marked with low tone on the sentence particle). It is not marked with an intransitive verb:

<i>ə̀jə̀q</i>	<i>lá</i>	<i>mɛ</i>
He	come	SP (SP <i>mɛ</i> : mid tone as sandhi of high)

‘He is coming.’

<i>ə̀jə̀q</i>	<i>lá</i>	<i>mè</i>
He	come	SP

‘He came.’

<i>ə̀jə̀q</i>	<i>àzàq</i>	<i>dì</i>	<i>sèq</i>	<i>mé</i>
He	pig	beat	death	SP

‘He will kill a pig.’

<i>ə̀jə̀q</i>	<i>nɛ</i>	<i>àzàq</i>	<i>dì</i>	<i>sèq</i>	<i>mè</i>
He	NP	pig	beat	death	SP

‘He has killed a pig.’

With inanimate nouns the NP *nɛ* marks an instrument:

<i>ə̀jə̀q</i>	<i>nɛ</i>	<i>mì-tʃhɛ</i>	<i>nɛ</i>	<i>àzàq</i>	<i>dì</i>	<i>sèq</i>	<i>mè</i>
He	NP	knife	NP	pig	beat	death	SP

‘He has killed a pig with a knife.’

An animate object may be marked with *áŋ* if clarity requires it, but usually it is unmarked:

<i>ja-tʃiq</i>	<i>dì</i>	<i>sèq</i>	<i>ŋa</i>	<i>djé</i>
chicken	beat	kill	SP	FP

‘The chicken was killed.’

An indirect object precedes a direct object and is marked with *áŋ*:

*àjðq áŋ áshì thì shì bìq má*  
 he NP fruit one CL give SP  
 '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 *nɛ* and *áŋ*, and to low laryngealized when followed by the NP *ə*:

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
1st	<i>ŋá -ŋà -ŋàq</i>	<i>ŋá njàq</i> (exclusive) <i>àdy njàq</i> (inclusive)	<i>ŋá màq</i> (exclusive) <i>àdy màq</i> (inclusive)
2nd	<i>nó -nò -nòq</i>	<i>nó njàq</i>	<i>nó màq</i>
3rd	<i>àjðq</i>	<i>àjɔq njàq</i>	<i>àjɔq màq</i>

There is also for the third person, a 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 àjðq áŋ áshì thì shì bìq má*  
 I he NP fruit one CLF give SP  
 'I will give him one fruit.'

In the past tense, the subject pronoun marked with *nɛ* and change of tone, the object pronoun marked with *áŋ*, sentence particle on low tone (mid tone as sandhi for low after another low tone):

*ŋà nɛ àjðq áŋ áshì thì shì bìq ma*  
 I NP he NP fruit one CLF give SP  
 'I gave him one fruit.'

Possession is marked with the NP *ə* (mid tone as sandhi when preceded by a low tone) after the possessor:

*xhà-là ə dð-mì*  
 tiger NP tail  
 'tiger's tail'

*ŋàq ə àma*  
 I NP 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 unprefixing kinship term:

*ŋà ma* my mother (< *àma*)  
*nð da* your father (< *àda*)

The interrogative pronouns are:

*àshú ~ àshú ʃà* who  
*àdjè* what  
*àgə+CLF* which

Demonstrative pronouns are followed by a classifier and have a three-way distinction for distance:

<i>hə</i>	here (close to me)
<i>thə</i>	there (close to you)
<i>xhə</i>	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 *ə* for non-past and *ə* for past. The noun may be deleted and the classifier acts as the head:

*mí-naŋ lá ə xhə nji ɣà*  
yesterday come those two CLF  
'Those two persons who came yesterday.'

*tjhé kha kha lá ə tshə-hà*  
rice plant come person (*ə* with mid tone as sandhi of high)  
'Those persons who are coming to plant rice.'

Classifiers are compulsory in Akha between a noun and a demonstrative pronoun and/or a numeral. A number of disyllabic nouns have a classifier as one of its syllables, e.g.:

<i>shà-áŋ thì áŋ</i>	one spindle
<i>lɔ-gɔŋ thì gɔŋ</i>	one stream

Most classifiers are used with a number of nouns, e.g.:

<i>ɣà</i>	for human beings (= strength)
<i>mɔ</i>	for animals (> <i>mɔ-do</i> = body)
<i>shì</i>	for small round things, fruits
<i>bjín</i>	for heaps

Nouns may be conjoined with the coordinate conjunction *hɔə* between two nouns, and with the comitative conjunction *nəqé* after the second noun:

<i>àma hɔə àda</i>	mother and father
<i>àma àda nəqé</i>	mother with father

An adjective may modify a noun in two ways:

- 1 Noun + *jə*  
*ŋà-sjhà jə-né* red fish  
*phi-dì jə-pjhú* white skirt
- 2 Noun + one syllable of the noun repeated + adjective  
*ŋà-sjhà ŋà né* redfish  
*phi-dì phi pjhú* white skirt

There is a difference in meaning between the two structures. A noun modified with the *jə* 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>nâq kha i</i>	very deep
<i>pjhú dâ i</i>	very white
<i>phý dù i</i>	very blue
<i>yâ lø 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>ná hý dzè</i>	I am bigger/the biggest
<i>àshú yà hý dzè</i>	who is bigger
<i>àma hɔɔ àda àshú yà hý dzè</i>	who is bigger, mother or father

#### 4.2 The verb phrase

The verb can be defined as being marked by the VP  $\varnothing$  in its citation form and by being able to be preceded by the negation *mà*. When asking: how do you say ‘eat’ in Akha, you always get as answer *dzâ 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-hý* ‘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 – preverb head auxiliaries (auxV) or after – postverb 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	1st pers.	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>tjhø</i>
Neg	auxV	auxV	Vh	Vr	Vaux	Vaux	Vv	Vaux
not	further	cause	call	back	for/towards	1st pers.	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>	<i>ð</i>		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-ŋɛ</i>	<i>í</i>		go to the fields
	Obj	Vh		
	<i>sjhý</i>	<i>í</i>		go yellow, wither
	Vh	Vv		
	<i>phá</i>	<i>lɔ</i>	<i>í</i>	further topples over
	auxV	Vh	Vv	
	<i>mà</i>	<i>hɔ</i>	<i>xhəm</i>	not allowed to look
	Neg	Vh	Vaux	
	<i>mà</i>	<i>kha</i>	<i>hɔ</i>	have never planted
	Neg	Vh	Vv	
	<i>àzàq</i>	<i>sèq</i>	<i>ə</i>	to kill a pig
	O	Vh	VP	
	<i>dì</i>	<i>sèq</i>	<i>ə</i>	beat to death
	Vh	Vv	VP	
	<i>ná-hà</i>	<i>sèq</i>	<i>ə</i>	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>làq</i>	towards first person
	<i>lèq</i>	towards non-first person
Attitudinal	<i>mòq</i>	want
	<i>náŋ</i>	dare
Potential	<i>tjhø</i>	able
	<i>xhəm</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 unprefixes in concatenations, but usually only as verb head or in the first post-head position

<i>sjhý í</i>	go yellow, wither (< <i>jɔ</i> - <i>sjhý</i> yellow)
Vh Vv	
<i>dɛ à</i>	beaten wet (by rain) (< <i>jɔ</i> - <i>à</i> wet)
Vh Vv	

Adverbs, preceding the verb phrase and the negation, are either always marked with *-i* or *-é*, as e.g. *íshàŋ i* 'very', *átjyq tètqé* 'a little bit', or constructed by repeating a *jɔ*-prefixed adjective or a verb phrase marked with *-é*:

<i>jɔ khó khó</i>	<i>é</i>	<i>í</i>
fast	P	go
'go fast'		

<i>hɔ hɔ</i>	<i>é</i>	<i>djɔ</i>
look	P	stay
'stay looking'		

<i>ləq</i>	<i>thà</i>	<i>ləq</i>	<i>thà</i>	<i>é</i>	<i>òq</i>	<i>lé</i>
peel	leave	peel	leave	P	return	go
'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 *ɔ* (longer form: *ɔ nɛ*) after each verb or concatenation:

<i>àjəq</i>	<i>já-ŋɛ</i>	<i>í</i>	<i>ɔ</i>	<i>tjhé</i>	<i>kha</i>	<i>ɔ</i>	<i>òq</i>	<i>lá</i>	<i>ɔ</i>	<i>hɔ</i>	<i>dzà dzà</i>	<i>ɔ...</i>
Sub	Place	Vh	VP	Obj	Vh	VP	Vh	Vv	VP	Obj	Vh	VP
he	field	go		rice	plant		return	come		rice	eat	
'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>ŋɛ</i>	while
<i>lɔ-nín -à</i>	if
<i>míɛ</i>	as, because
<i>xòqì</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*, *mɛ*) or infer from seeing (*ɲa*, *ɲaa*), 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). To these comes four particles of prediction, only with high tone, expressing fear (*á*) or threat (*ɲá*), doubt (*m̄*) or certainty (*má*) (for more details, see Egerod and Hansson 1974; and Hansson 1996). 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 mɛ</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 m̄ ló</i>	Has he come?
<i>lá m̄</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): *ɲá* 'feeling' and plain statement, *ɲá* 'seeing' and plain statement, *mía* 'hearing'/'feeling and reaction', *ɲáa* seeing and reaction. Some contrasting examples:

<i>hò m̄ɛq mía</i>	I am hungry (it's dinner time).
<i>hò m̄ɛq 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 ɲá</i>	He is an Akha (I can see it).
<i>àkhà ɲáa</i>	Yes, he is an Akha (I can see it).

<i>àkhà míá</i>	He is an Akha (I can hear him speaking).
<i>àkhà á nja</i>	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>áj</i>	<i>dì</i>	<i>ə</i>	<i>àshú yà</i>	<i>mà</i>	<i>há</i>
he	NP	beat	VP	who	not	SP
'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 and 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 1996.) 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 bjm bjm</i>	divide meat into piles
<i>sjhà bjεq bjεq</i>	chop raw meat
<i>sjhà the the</i>	hunt
<i>sjhà tjhé tjhé</i>	pickle

From *ùdù* head

<i>ùdzàη dzàη</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 thì kháŋ dzàŋ</i>	wear one turban
<i>làq bè thì 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 clearly most prevalent and most versatile is *tjhé* ‘rice’, ‘paddy’ as, for example:

<i>tjhé boq boq</i>	pop rice
<i>tjhé dī dī</i>	thresh paddy
<i>tjhé dyq dyq</i>	soak rice
<i>tjhé ləq ləq</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 *hə* ‘cooked rice’:

<i>m̄ djè djè</i>	to thunder
<i>m̄ mjəq mjəq</i>	to lighten
<i>m̄ tjìq tjìq</i>	get towards evening
<i>hə thàŋ thàŋ</i>	make rice cakes
<i>hə tjhu tjhu</i>	wrap rice in leaves
<i>hə tsuq tsuq</i>	make rice balls
<i>hə 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>ádzɣ dzɣ</i>	itch	<i>ájɛq jɛq</i>	to flower
<i>álm̄ 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>áəq əq</i>	to burp		

Sometimes the second syllable occurs both with prefix *a-* and in other combinations:

<i>ábòq bòq</i>	make a pattern
<i>shàŋ bòq bòq</i>	write (wood +)

where *ábòq bòq* has a more general meaning and *shàŋ 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>bɔ 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>h̀ dzà dzà</i>	eat (action of eating)
<i>h̀ dzà mà dzà</i>	not eat
<i>h̀ dzà</i>	eat rice (have a meal)
<i>h̀ 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: 225–84.
- Hansson, Inga-Lill (1982) 'A phonological comparison of Akha and Hani', *Linguistics of the Tibeto-Burman Area* 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 Pije', *Linguistics of the Tibeto-Burman Area* 12.1: 6–91.
- Hansson, Inga-Lill (1991) 'The language of Akha ritual texts', *Linguistics of the Tibeto-Burman Area* 14.2: 155–67.
- Hansson, Inga-Lill (1996) 'The interplay between the verb particle 'ə' and the sentence particles in Akha', *Linguistics of the Tibeto-Burman Area* 19.1: 65–76.
- Hansson, Inga-Lill (1996) 'Object-verb in Akha: the ABB Structure', *Linguistics of the Tibeto-Burman Area* 19.1: 77–95.
- 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.



**PART 5**

---

# **BODISH LANGUAGES**

---

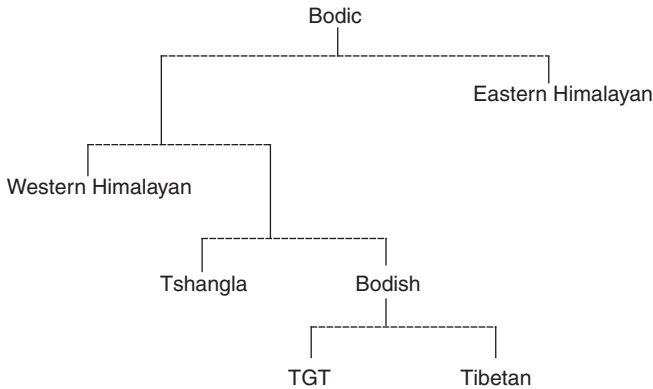


# CLASSICAL TIBETAN

*Scott DeLancey*

## 1 THE TIBETAN LANGUAGE

Tibetan consists of a number of dialects, not all mutually intelligible; for an outline of dialects see the chapter on Lhasa Tibetan. Tibetan belongs to the Bodic branch of Tibeto-Burman, within which it appears to be most closely related to the Tamang-Gurung-Thakali (TGT) nucleus, then to the West Himalayan group:



**FIGURE 16.1** THE BODIC BRANCH

Documents exist in Tibetan script representing two ancient Bodic but non-Tibetan languages, Nam (Thomas 1948) and Zhang-zhung (Hoffman 1967; Thomas 1967; Haahr 1968; Hummel 1995; Martin 1997).

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 attributed to the legendary minister Thon-mi Sambhota (Bacot 1928; Schubert 1937; cf. Miller 1976) in the reign of Srong= btsan sgam-po; the traditional date is 632 CE. Several texts and inscriptions are preserved from the eighth and ninth centuries; including a number of texts recovered from the caves at Dun-huang (see Thomas 1935–63, 1957). The orthography preserves the phonology and morphology of this period, though scribal errors in texts show that the simplification of consonant clusters which distinguishes most modern dialects from the written form had already begun by the tenth century.

‘Classical Tibetan’ is an idealization, referring both to over a millennium of written history and to a tradition of prescriptive grammar which many of the authors of the texts, in some cases down to the present, made greater or lesser efforts to conform to. (Elements of traditional grammatical analysis are explicated in modern terms in Gyurme 1992.) In some texts we can trace evidence of the historical development of the language. The term ‘Old

Tibetan' is used to refer to written material from before about 1000 CE, primarily inscriptions and documents found in the Dun-huang caves (for published examples see Thomas 1935–1963, 1957; Richardson 1985). Already in the vernacular *Mi=la Ras-pa'i rNam=thar* (see De Jong 1959), the biography of Milarepa, which probably dates from the fourteenth century, we find a mixture of classical and more modern features.

## 2 PHONOLOGY

There is 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 of this chapter I have used hyphens 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.

### 2.1 Segmental inventory

There are several transliteration systems for Tibetan in current use; a summary of the major systems can be found in Hoffmann 1975. 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 (seventh century). There are thirty 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:

k	kh	g	ng
c	ch	j	ny
t	th	d	n
p	ph	b	m
ts	tsh	dz	
w	zh	z	ʼ
y	r	l	
sh	s	h	ʔ
	i	u	
	e	o	
	a		

The phonetic interpretation of most of the letters is straightforward, the most problematic being those transliterated as ʼ and ʔ (Clauson and Yoshitake 1929; Li 1933). The ʔ apparently represented a glottal stop, as reflected in its reflexes in modern dialects. The interpretation of ʼ, called 'a-chung' ('little a') is problematic. The writing system requires a place-holder in root-initial position, and ʼ occurs there when there is no other consonant: 'o ma 'milk', etc. In the modern dialects these words have initial breathy [h], and low tone. But ʼ is also a prefix, occurring only before voiced and voiceless aspirated stops and affricates: 'byung 'emerge', etc.

In many modern dialects the reflexes of orthographic *a-chung* + obstruent onsets are prenasalized obstruents. Whether these two functions of the letter represent the same phonological unit remains a matter of controversy. The third orthographic use of *a chung* is to indicate the position of the vowel in an otherwise ambiguous sequence of consonants. For example, the sequence *dg* will be read as the onset and coda of a syllable with the default vowel: /dag/. But *d-g* is also a possible onset, so the syllable /dga/ must be written *dg'*, 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 below). Thus their phonemic status in Classical Tibetan is marginal, and it is clear that at an earlier stage of the language they were not phonemically distinct.

## 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:

*r*  
*bgd*  
*y*

The *r* written above the *g* immediately precedes it; the *b* then precedes the *r*. *y* written below the *g* follows it. Some Western Tibetanists use an upper-case letter in transliteration to indicate the root initial.

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 PTB medial inventory. Combinations of root initials with one of these are probably the only original tautomorphic 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*, or root-initial *g* with medial *y*. Capitalization of the root initial eliminates this ambiguity: *Gyon pa* 'wear', *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 (Section 3.3.2). Thus the Old Tibetan syllable canon is (C<sub>p</sub>)C(G)V(C)-(s/t), where C<sub>p</sub> can be a voiced stop *b d g*, a liquid *r l s, m*, or *ʿ*. Possible root-final codas are the PTB 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 WORD CLASSES AND INFLECTIONS

### 3.1 Nouns

#### 3.1.1 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 Section 2.3):

genitive *kyi/gi/gyil'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 Hahn 1974: 81–116; Beyer 1992: 267–70.)

#### 3.1.2 Alignment

Classical Tibetan is relatively consistently ergative, i.e. the subject of a transitive verb has overt case marking (in Tibetan identical to the instrumental case) contrasting with the zero marking of intransitive subjects:

*dbyig-pa-can de 'bros pa-r brtsams te*  
 dByig-pa-can DEM flee NOM-TERM begin NF  
 'dByig-pa-can tried to flee.'

*dbyig-pa-can gyis bdag gi bu bsad do*  
 dByig-pa-can **ERG** I GEN child killed FINAL  
 'dByig-pa-can killed my child.'

Some texts show an aspectual split pattern, with ergative marking only in perfective clauses (Regamey 1954), but most do not (see Andersen 1987; Tillemans and Herforth 1989; Saxena 1989; Dempsey 1993 for further discussion). Some texts show ergative marking on agentive intransitive subjects, as in Modern Central Tibetan (cf. Saxena 1989: 37–8):

*de-s rtsig-pa zhig gi steng nas mchongs pa*  
 that-ERG wall a GEN upper.surface ABL leapt NOM  
 'He jumped off a wall.'

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 'principal 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:

*shing-la sta-re gzhus-pa*  
 tree-LOC axe hit-NOM  
 'hit the tree with an axe'

*sta-re-s shing 'chad-pa*  
 axe-INSTR tree cut-NOM  
 '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.

### 3.2 Pronouns

Personal pronouns found in the texts include a range of forms, differing probably in dialectal as well as honorific status. 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 the first person. Plural marking is not obligatory on nouns or pronouns, but all pronominal stems can be pluralized using *cag*, *tsho*, or *mams* (Section 3.1). The stem *u* occurs only in first person pluralized forms: '*u cag* '*u bu cag* 'we'.

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*.

All pronominal forms with personal reference frequently occur with one of the reflexive suffixes *nyid* or *rang*: *nga rang*, *nga nyid* 'I'; *de nyid* 'that one', 's/he'. *Rang* occurs alone with first person singular reference; *nyid* for second or third person.

### 3.3 Verbs

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. In the examples below, I follow the standard practice of presenting the stems in this order. 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 1974; Beyer 1992; 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; in addition to the above sources see Conrady 1896; von Koerber 1935 (which combines some solid analysis with some rather wild speculation), Wolfenden 1929; Li 1933, 1959; Uray 1953.

### 3.3.1 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 Kharto n.d.):

see *lta bltas blta ltos*

This phenomenon must reflect a pre-Tibetan imperative suffix \**o*, which is still attested in some other Bodic languages. The other patterns, all less common, are:

***o a a o***: overthrow *rlog brlags brlag rlogs*

***e a a o***: fill up *gengs bkang dgang khong*

***i u u u***: remove *'byin phyung dbyung phyung(s)*

### 3.3.2 Suffixation

The only productive inflectional suffix is *-s*. This is most characteristically found distinguishing the past and imperative stems from the present and future:

finish *sgrub bsgrubs bsgrub sgrubs*

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). 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 Section 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 the 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 *snyegs bsnyegs bsnyeg snyegs*

And in others it occurs only in the present stem (see 'fill up', above). 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' above), and final *-d* in the present stem only:

do *byed byas bya byos*

### 3.3.3 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* *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 particular 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 unprefixated, though it may also show the *g/d* prefix, and generally takes the *-s* suffix. The regular patterns are

‘*b g*: hit ‘*debs* *btab* *gtab* *thob*  
‘*b b*: make ‘*chos* *bzos* *bzo* *chos*  
*g b g*: cut *good* *bcad* *gcad* *chod*  
*g b b*: mince *gtsab* *btsabs* *btsab* *gtsabs*

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.

### 3.3.4 Root-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*

There are fewer than half-a-dozen verbs with unprefixated initial unaspirates; the unaspirated series otherwise occur in the verbal system only following the *b d g r l s* prefixes in alternation with aspirates.

### 3.3.5 Voice and diathesis

Tibetan has no morphologically or syntactically explicit voice alternation, although some scholars suggest that certain patterns of zero anaphora (Section 3.2) and/or NP-fronting (Section 6) may be interpretable as voice constructions (see Andersen 1987; Dempsey 1993). Regardless of how one may choose to label the various constructions, the functional load of an active/passive alternation is divided among zero anaphora, NP-fronting, and the alternation of related intransitive/transitive verb pairs.

There is a large body of such verb pairs, related by semi-regular morphological derivations (Conrady 1896; Wolfenden 1929; Chang 1971). The primary pattern, of PST provenience, is a causative *s-* prefix: *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; see Section 3.3.4), *’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’.

### 3.4 Sentence particles

A CT sentence ends with a final particle: declarative ‘*o*, interrogative ‘*am*, imperative *cig* (see Section 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-NOM-INSTR tongue cut IMP  
 ‘Since [he] didn’t speak, cut out [his] tongue!’

The *a-chung* initial copies a previous final consonant:

*rab-tu dbul-phongs so*  
 very poor-poor FINAL  
 ‘[He was] very poor.’

*mi de dgra yin nam*  
 person DEM enemy be Q  
 ‘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 Section 6.3.)

## 4 WORD FORMATION

Tibetan, like other Sino-Tibetan languages, is very fond of both derived and compound nouns. There is also a large inventory of Noun + Verb constructions which must be considered fixed lexical items, but we can see from the modern dialects that these do not represent formal compounds in the same sense as the compound nouns. (These are more frequent in the modern dialects; see the Lhasa chapter.)

### 4.1 Derived nouns

Many nouns consist of only the simple noun stem: *mi* ‘person’, *bod* ‘Tibet’, *rta* ‘horse’, *ja* ‘tea’, *mgo* ‘door’. As many, if not 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 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*.

## 4.2 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.

## 5 THE NOUN PHRASE

### 5.1 Number words

Number is not an obligatory category, but several ‘number words’ can be used to specify plurality. The commonest are *tsho*, *rnams*, and *cag*: ‘*di tsho* ‘these’, *nga cag* ‘we’, etc. In the modern dialects 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 lexical noun phrase they are not bound to the head:

<i>mi</i>	<i>chen-po</i>	<i>de</i>	<i>rnams</i>
person	large	DEM	PL
‘those big men’			

### 5.2 Demonstratives and articles

The demonstratives are *de* ‘that’, ‘*di* ‘this’. Both occur both 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 Section 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* (Section 4.1) is unclear. This form is still used in the Western dialects, but not in the Central region.

### 5.3 Other noun modifiers

Genitive modifiers precede the head:

<i>khyim</i>	=	<i>bdag</i>	<i>gi</i>	<i>bu</i>
householder		GEN		son
‘the householder’s son’				

A peculiar feature of the Tibetan noun phrase is an appositive-like construction:

*nga ma rgan=mo ni bsam=mno btang*  
 I mother old.woman TOP 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)

Adjectives may precede the head, in which case they are marked as genitive modifiers:

*gzhan zhig gi yul*  
 other a GEN country  
 ‘another country’

Or they may follow, in which case they are followed by one of the nominalizing particles:

*mtshan gzhan pa*  
 sign other NOM  
 ‘other signs’

This construction may be related to the appositive construction.

Relative clauses are nominalized clauses which likewise may precede the head as genitive modifiers, or follow as appositives:

*sgas pa-i zas rnams*  
 gather NOM-GEN food PL  
 ‘the food which had been collected’

*shing=sdong che ru ‘gro ba de*  
 tree great TERMINATIVE go NOM DEM  
 ‘the tree which grows great’

#### 5.4 Relator nouns and postpositions

The only spatial postpositions are those listed in the section on case forms (Section 3.1.1). 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.’

*de-s rtsig-pa zhig gi steng nas mchongs pa*  
 that-ERG wall a GEN upper.surface ABL leapt NOM  
 ‘He jumped off a wall.’

Relator nouns may also have more abstract senses:

*lta ba-‘i ched du*  
 see NOM-GEN purpose TERMINATIVE  
 ‘in order to see’

## 6 CLAUSE AND SENTENCE

Word order is 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.

### 6.1 Complementation

The basic complement construction is simple nominalization with *pa*, but a range of other constructions is found. The most common are clauses marked with the terminative case, and complements of *verba dicendi* marked by *ces* (*zhes/shes*, see Section 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:

*dbyig-pa-can de 'bros pa-r brtsams te*  
 dByig-pa-can DEM flee NOM-TERM begin NF  
 'dByig-pa-can tried to flee.'

*yul pha=rol=po-r ltad=mo lta-r song nas*  
 country other.side-TERM sights see-TERM went NF  
 '[They] went to another country to see the sights.'

Quotations are marked by *ces*:

*nga-s khyod kyī khyo ji-ltar sbyin zhes smras nas*  
 I-ERG you GEN husband how give QUOTE spoke NON-FINAL  
 '“How can I give you [back] your husband?” [he] said.'

The use of *ces* is very 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 QUOTE ask FINAL  
 'The householder's son made this request . . .'

### 6.2 Pragmatic particles

Two particles with pragmatic force are quite common in text. *Ni* marks a continuative or resumptive topic:

*khyim-bdag-gi bu de ni ba-she-cir yin no*  
 householder-GEN son that TOP P.N. 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 Section 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 YANG DEM like TERM do(HON) ABL

'The householder's son made this request . . . The king, in turn, assented to that.'

### 6.3 Subordination and clause chaining

There are two morphemes which function only as subordinating particles: *(s)te*, the generic non-final, and *cing*, which indicates temporal overlap between the clause where it occurs and the next. Since it contrasts with *cing*, *(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 Section 2.3.)

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:

ablative	<i>nas</i>	>	temporal succession, then
elative	<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 CT 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 DEM boy DAT give ABL  
'The householder gave the baby bird to his son.'

*cher-skyes pa dang gcig la gcig chags te*  
grow.up NOM 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 TERM 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 DEM-GEN back LOC sit ABL bird ERG carry CONT  
'sitting on the bird's back, the bird carrying [him]'.

*nam-mkha' la 'phur te 'gro 'o*  
sky LOC fly NON-FINAL go FINAL  
'flying in the sky, [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’.

#### 6.4 Questions and negation

Yes-no questions are marked by the final particle ‘*am*’ (Section 3.4):

*khyod kyis glang brnyas sam*  
 you ERG ox borrow Q  
 ‘Did you borrow [his] ox?’

Wh-words occur in their syntactically appropriate position; Wh-questions have no final particle:

*nga-s khyod kyi khyo ci-ltar sbyin*  
 I-ERG you 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-NOM food and clothes  
*med-pa zhid go*  
 not.have-NOM a FINAL  
 ‘[He was] a very poor one, one without food or clothing.’

#### REFERENCES

Of the various grammars listed here, the most complete and reliable are those of Hahn and Beyer. (I have not been able to directly consult Shoji’s grammar, but it is also very highly regarded.) Linguists may find Beyer the most accessible source.

Andersen, Paul Kent (1987) ‘Zero-anaphora and related phenomena in Classical Tibetan’, *Studies in Language* 11.2: 279–312.

Bacot, Jacques (1928) *Les slokas grammaticaux de Thonmi Sombhota*, Paris: Paul Geuthner.

Bacot, Jacques (1946) *Grammaire du tibétain littéraire*, Paris: Librairie d’Amérique et d’Orient.

Beyer, Stephan (1992) *The Classical Tibetan Language*, Albany, NY: State University of New York Press.

Chang, Betty Shefts (1971) ‘The Tibetan causative: phonology’, *BIHP* 42.4: 623–766.

Clauson, G.L.M. and Yoshitake, S. (1929) ‘On the phonetic value of the Tibetan characters ཨ and ཨ and the equivalent characters in the ḥphags.pa alphabet’, *JRAS* 1929: 843–63.

Coblin, Weldon South (1976) ‘Notes on Tibetan verbal morphology’, *T’oung Pao* 62: 45–70.

Conrady, August (1896) *Eine indochinesische Causative-Denominative-Bildung und ihr Zusammenhang mit den Tonaccenten*, Leipzig: Otto Harrassowitz.

Csoma de Korös, Alexander (1834) *A Grammar of the Tibetan Language in English*, Calcutta: Asiatic Society of Bengal (repr. 1983, New Delhi: Nagawang Topgyal).

- Das, Sarat Chandra (1902) *Tibetan-English Dictionary, with Sanskrit Synonyms*, Calcutta: Bengal Secretariat Book Depot (repr. 1976, Delhi: Motilal Banarsidass).
- De Jong, J.W. (ed.) (1959) *Mi la Ras Pa'i rNam Thar: Texte tibétain de la vie de Milarépa*, (Indo-Iranian Monographs IV), s'Gravenhage: Mouton.
- Dempsy, Jakob (1993) 'Ergativity in Milarepa's Rnam-thar': another viewpoint, *LTBA* 16.2: 113–20.
- Durr, Jacques (1950) *Morphologie du verb tibétain*, Heidelberg: Carl Winter.
- Gyurme, Kesang [sKal-bZang 'Gyur-Med] (1992) *Le clair miroir: enseignement de la grammaire tibétaine*, Paris: Éditions Prajna.
- Haarh, Erik (1968) 'The Zhang-zhung language: A grammar and dictionary of the unexplored language of the Tibetan Bonpos', *Acta Jutlandica* 40.1: 7–43.
- Hahn, Michael (1974) *Lehrbuch der klassischen tibetischen Schriftsprache*, Bonn: Michael Hahn.
- Hannah, Herbert B. (1912) *A Grammar of the Tibetan Language, Literary and Colloquial*, Calcutta: Baptist Mission Press, (repr. 1978: Delhi: Motilal Banarsidass).
- Hoffmann, Helmut (1967) 'Zaṅ-zuṅ: the holy language of the Tibetan Bon-po', *ZDMG* 117.2: 376–81.
- Hoffmann, Helmut (1975) *Tibet: A Handbook*, Bloomington, IN: Research Center for the Language Sciences, Indiana University.
- Hummel, Siegbert (1995) 'Neues Material zur Sprache von Zhang-zhung', *AO* 59: 162–68.
- Jäschke, H.A. (1881) *A Tibetan-English Dictionary*, London: Routledge and Kegan Paul.
- Jäschke, H.A. (1929) *Tibetan Grammar*. Berlin: Walter de Gruyter (contains very useful addenda by A.H. Francke).
- Jäschke, H.A. (1954) *Tibetan Grammar*, New York: Frederick Ungar (The 1929 edition with added reading material and vocabulary, but without Francke's addenda).
- Khartho, Dorje Wangchuk n.d. *Thumi'i dGongs gTer* (The complete Tibetan verb forms), Delhi: C.T. Khartho.
- Koerber, Hans Nordewin von (1935) *Morphology of the Tibetan Language*, Los Angeles: Sutton-house.
- Lalou, Marcelle (1950) *Manuel élémentaire de tibétain classique (méthode empirique)*, Paris: Librairie d'Amérique et d'Orient.
- Li, Fang-kuei (1933) 'Certain phonetic influences of the Tibetan prefixes upon the root initials', *BIHP* 4: 135–57.
- Li, Fang-kuei (1959) 'Tibetan *glo-ba-'dring*', in Søren Egerod (ed.) *Studia Serica Bernhard Karlgren Dedicata*, Copenhagen: E. Munksgaard, 55–9.
- Lyovin, Anatole, (1997) 'Sketch of Classical Tibetan', in Lyovin A. (ed.) 'An introduction to the languages of the world', Oxford: Oxford University Press.
- Martin, Dan (1997) *Zhang-zhung dictionary*. <http://www.comet.net/ligmincha/html/zdzdict1.html>
- Miller, Roy A. (1970) 'A grammatical sketch of Classical Tibetan', *JAOS* 90: 74–96.
- Miller, Roy A. (1976) *Studies in the Grammatical Tradition in Tibet* (Amsterdam Studies in the Theory and History of Linguistic Science III: Studies in the History of Linguistics, vol. 6), Amsterdam: John Benjamins.
- Miller, Roy A. (1992) 'Indic models in Tibetan grammars', *JAOS* 112: 103–9.
- Przyłuski, Jean, and Lalou, Marcelle (1933) 'Le da-drag tibétain', *BSOS* 7: 87–9.
- Regamey, Constantin (1954) 'A propos de la "construction ergative" en indo-aryen moderne', in *Sprachgeschichte und Wortbedeutung: Festschrift Albert Debrunner*, Bern: Francke.
- Richardson, H.E. (1985) *A Corpus of early Tibetan Inscriptions*, London: Royal Asiatic Society.
- Saxena, Anju (1989) 'Ergative in Mi = la = ras = pa'i rnam thar', *LTBA* 12.2: 35–9.
- Schubert, Johannes (1937) *Tibetische Nationalgrammatik: das Sum-cu-pa und Rtags-kyi-'ajug-pa des Grosslamas von Peking Rol-pai-rdo-rje: ein Kommentar zu den gleichnamigen Schriften Thon-mi Sambhota's auf Grund der Erklärung des Lamas Chos-skyon-bzan-po, Lo-tsa-ba von Zha-lung, mit Übersetzung und Anmerkungen versehen von Johannes Schubert* (Artibus Asiae suppl. 1), Leipzig: Offizin Richard Hadl.
- Shafer, Robert (1951) 'Studies in the morphology of Bodic verbs', *BSOAS* 13: 702–24.
- Shoju, Inaba (1954) *Chibettoگو koten bunpogaku*, Kyoto: Hozokan.

- Thomas, Frederick (1935–63) *Tibetan Literary Texts and Documents Concerning Chinese Turkestan*, 4 volumes (Oriental Translation Fund, new series, v. 32, 37, 40, 41), London: Royal Asiatic Society.
- Thomas, Frederick (1948) *Nam: An Ancient Language of the Sino-Tibetan Borderland*, (Publications of the Philological Society 14), London: Oxford University Press.
- Thomas, Frederick (1957) *Ancient Folk-Literature from North-Eastern Tibet*, (Abhandlungen der Deutschen Akademie der Wissenschaften zu Berlin, Klasse für Sprachen, Literatur und Kunst; Jahrg. 1952, Nr. 3), Berlin, Akademie-Verlag.
- Thomas, Frederick (1967) 'The Zhang-zhung language', *Asia Major* 13: 211–17.
- Tillemans, Tom, and Herforth, Derek (1989) *Agents and Actions in Classical Tibetan*, (Wiener Studien zur Tibetologie und Buddhismuskunde, Heft 21), Vienna: Arbeitskreis für Tibetische und Buddhistische Studien Universität Wien.
- Uray, Géza (1953) 'Some problems of the Ancient Tibetan verbal morphology: methodological observations on recent studies', *Acta Linguistica Hungarica* 3: 37–62.
- Wolfenden, Stuart N. (1929) *Outlines of Tibeto-Burman Linguistic Morphology*, London: Royal Asiatic Society.

# LHASA TIBETAN

*Scott DeLancey*

Various dialects of Tibetan are spoken in a broad area reaching from northern Pakistan to Qinghai, Sichuan, and Yunnan provinces in China. The place of Tibetan within Tibeto-Burman is discussed in Chapter 16. A great deal of work remains to be done on Tibetan dialectology. The most comprehensive current classification is that of Nishi (1986), who distinguishes six major groups: Central or U-Tsang (Lhasa, Shigatse, Sherpa, Kagate, etc.), Western Archaic (Balti, Ladakhi, Purik), Western Innovative (Lahul, Spiti), Southern (dialects of Sikkim and Bhutan), Khams, and Amdo. Lhasa, the best-known contemporary Tibetan dialect, belongs to the Central group. It serves as a *lingua franca*, and to some extent as a standard, in both the Tibetan Autonomous Region and the diaspora communities.

## 1 PHONOLOGY

### 1.1 Transcription and transliteration

Forms and examples in this chapter will be presented in orthographic transliteration, as well as phonemic transcription where relevant. With the exception of a handful of irregular words, phonemic representation can be recovered directly from orthographic form, as will be described in the course of this section (where ‘original’ can be read as ‘orthographic’). Orthographic forms are transliterated in a slightly modified version of the Wylie system (Wylie 1959). There is little consistency in the literature in the system of phonemic transcription; there is disagreement among many authors as to analysis (though the analysis of Chang and Chang (Chang and Shefts 1965; Chang and Chang 1978) is accepted by a number of American scholars), and even different works by the same author may show different typographic choices (compare e.g. Goldstein 1975, 1984). The system used in this chapter is close to that of Goldstein 1984, with several modifications: /k<sup>y</sup> kh<sup>y</sup>/ instead of /k kh/, /k kh/ instead of /q qh/, /ä/ instead of /ε/, and a substantively different representation of tone, which will be explained below.

### 1.2 Segmental inventory

The consonant phonemes of Lhasa are:

<i>Labial</i>	<i>Dental</i>	<i>Retroflex palatal</i>	<i>Palatalized velar</i>	<i>Velar</i>	<i>Glottal</i>
p	t	ʈ	k <sup>y</sup>	k	
ph	th	ʈh	kh <sup>y</sup>	kh	
	ts				c
	tsh				ch
m	n	ṅ	ŋ		

(mh)			
	s	š	h
	l	r	
	lh	rh	
w			y

The retroflexes /ʈ ʈh/ reflect original stop + r clusters; /kʷ khʷ/ 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 16). In my data voiceless [mh] occurs only in the allomorph of the negative prefix preceding aspirated stops, and only for some speakers. Some authors (Chang and Chang; 1978–81; Goldstein and Nornang 1970) report a few words with initial voiceless nasals /ñh ŋh/, but these do not occur in the speech of all speakers.

The vowel phonemes are:

i	ü	u
e	ö	o
ä	a	(ɔ)

The five original vowels, /i e a o u/ occur both short and long, in open and closed syllables. Fronted /ü ö ä/ occur only in long rhymes (oral and nasalized) and in glottal/falling tone syllables (see below); they always reflect a lost coronal coda *dn s* or *l*.

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* /ph<sup>h</sup>ö/ ‘the Bon religion’, *bon-po* /ph<sup>h</sup>mpo/ ‘an adherent of Bon’. Nasalization will be represented as raised /<sup>h</sup>/.

The phonemic status of [ɔ] is debatable. It reflects original *og* and *or* rhymes. 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/ is always [khɔɔ]). /a/ is realized as [ə] before tautosyllabic /p, m/, when unstressed, and under vowel harmony.

### 1.3 Vowel harmony

Lhasa Tibetan shows dominant/recessive vowel harmony – all vowels are raised in any word which contains an intrinsically high vowel. (For more detailed descriptions, including exceptions to this generalization, see Sprigg 1961; Chang and Chang 1968, 1978; Dawson 1980). 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:

*îgro* /tɔ/ go  
*îgro* = *gyi yin* /t<sup>u</sup>kiyii<sup>h</sup>/ go (CONJUNCT FUTURE)  
*dkar* = *po* /kaapo/ white  
*zhim* = *po* /šimpu/ 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/; see Chang and Chang 1978: vii–xix for discussion.

### 1.4 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 [ʔ], 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 ŋ/ occur as codas. If the palatalized velars /kʲ/ and /khʲ/ and the prenasalized stops (in those idiolects where they occur) are analysed as unitary segments, there are no consonant clusters. The inventory of syllable patterns is:

CV	short open rhyme
CVV	long open rhyme
CVN	nasal-final rhyme
CVV <sup>n</sup>	long nasalized rhyme
CVC	obstruent-final rhyme
CVʔ	glottalized rhyme ~ CVV <sup>^</sup> long rhyme with falling tone

CVV rhymes reflect original liquid codas *l r*. CVV<sup>n</sup> rhymes reflect original nasal, and CVʔ and CVV<sup>^</sup> original obstruent finals. Labial codas are retained in word-final position. Velar codas are variably present: *chang* ‘beer’ [chaa<sup>n</sup>] or [chaŋ], *lug* ‘sheep’ [lɯk], [lɯʔ] or [lɯŋ]. All original coronal codas are lost in Lhasa, but are reflected in fronting of original non-front vowels.

### 1.5 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 16 for an explanation of ‘root initial’) or sonorant initial with prefix; low tone syllables reflect original voiced obstruent or unprefixated voiced sonorant initials. In this chapter low tone is indicated by underlining of the vowel; high tone is unmarked.

Lhasa 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 a H or mid tone if the first syllable is intrinsically low. In this chapter the morpheme boundary in words showing this tonal pattern 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 rhyme, and always behave as closed syllables with respect to all phonological rules. These will be consistently represented in this chapter as glottal-final, i.e. *stag* /taʔ/ rather than /tã/ as in much modern work (e.g. Chang and Chang 1978; Goldstein and Nornang 1970; Goldstein and Narkyid 1984). Both the final glottal and falling tone can occur only in a final syllable. Non-finally these rhymes are usually realized as long vowels: *bod* /phōʔ/ ‘Tibet’, *bod* = *skad* /phōökāʔ/ ‘Tibetan language’; irregularly in some words as short rhymes: *bod-pa* /phōpa/ ‘Tibetan person’.

While the high/low distinction is universally recognized, other aspects of Tibetan tonal phonology are analysed quite differently by various scholars. For summaries of some areas of disagreement see Mazaudon 1977, 1984; Hu 1982.

## 1.6 Fossilized clusters

Even for the majority of speakers who do not pronounce prenasalized stops, words with orthographic *ʼ* or *m* preceding a root-initial obstruent have an underlying initial nasal which surfaces following a vowel within a word: *gro* /tʰ/ ‘go’, *ma* ‘gro /manto/ ‘not go’; *mda* /ta/ ‘arrow’, *me* = *mda* /mɛnta/ ‘gun’ (Chang and Shefts 1965). This process is very productive. A limited set of morphemes with an orthographic *b* prefix, and a handful of others, manifest a /p/ within compounds under the same circumstances: *bzhi* /ʃi/ ‘four’, *bcu* /cu/ ‘ten’, *bzhi* = *bcu* /ʃipcu/ ‘forty’. This phenomenon shows limited productivity, being mostly limited to specific lexical items (see Shefts and Chang 1967).

## 2 NOUNS, ADJECTIVES, AND NOMINAL MORPHOLOGY

### 2.1 Elements of the noun phrase

Although constituent order in Tibetan is otherwise consistent with the SOV stereotype, order within the NP is unusual. Genitive modifiers (including one variety of relative clause) precede the head noun; all other NP elements follow. Nouns and adjectives will be discussed in separate sections.

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/, historically a reduced form of *gcig* ‘one’:

*rnam* = *lha*    *zer* = *mkhan-gyi*    *ʼbrog-pa*    *cig*  
 Namla        call = NOM-GEN    nomad    a  
 ‘a nomad named Namla’

*lo*    *kha* = *shes*    *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’.

### 2.2 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* /ŋaŋi/, *nga* = *tsho* /ŋatsho/, *khyed* = *rang* = *gnyis* /kh<sup>y</sup>eraŋi/, etc.

### 2.3 Case

Lhasa has four postclitics 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 [kiʔ];  
 ‘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/locative *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 /näʔ/

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: /thubtää<sup>n</sup> khi/.

The case role marked by the ergative has been extensively discussed in recent literature (Chang and Chang 1980; DeLancey 1984, 1985, 1990; Tournadre 1996a; Agha 1993). 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:

*blo = bzang-gis nga-r mthong-byung*  
Lobsang-ERG I-LOC see-PERF/CONJUNCT-GOAL  
'Lobsang saw me.'

It also marks instruments:

*blo = bzang-gis me = mda-s stag bsad-pa red*  
Lobsang-ERG gun-INSTR tiger killed-PERF/DISJUNCT  
'Lobsang killed a tiger with a gun.'

As well as oblique locatives, the *la* case marks recipients in trivalent constructions:

*nga-s blo = bzang-la ngul-tsam sprad-pa yin*  
I-ERG Lobsang-LOC money-some give-PERF/CONJUNCT  
'I gave some money to Lobsang.'

'dative subjects' of predicates such as 'like' and 'need':

*blo = bzang-la ngul dgos = kyi*  
Lobsang-LOC money need-IMPF/DISJUNCT  
'Lobsang needs money.'

and possessors in 'have' constructions:

*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:

*thub = bstan-gyis blo = bzang-la gzhus-song*  
Thubten-ERG Lobsang-LOC hit-PERF  
'Thubten hit Lobsang.'

*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: change of state verbs like *gsod* ‘kill’, *gcod* ‘cut’, *bzo* ‘cook’, etc., always require an unmarked argument; ‘surface contact’ verbs like *gzhus* ‘hit’ require a *la*-marked argument. There can be no alternation between the two constructions with the same verb.

## 2.4 Relict case forms

Remnants of the somewhat more more elaborated Classical Tibetan case system are found in restricted contexts in Lhasa. The ablative *las* no longer occurs in a spatial sense, but is preserved in the comparative construction (Section 2.5). The terminative *du* is still found in a number of frozen constructions, including relator noun constructions such as *N-gi ched-du* ‘for the benefit of N’, and adverbs such as *sger-du* ‘privately, personally’, *lhag=par-du* ‘especially’.

## 2.5 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:

*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-la* ‘in’, *sgang* ‘on’, and ‘og ‘under’, do not allow genitive marking on the lexical noun:

*zim=chung (\*-gi) nang-la*  
 bedroom (\*-GEN) in-LOC  
 ‘in the bedroom’

These represent an intermediate category of erstwhile nouns grammaticalizing into post-positions (DeLancey 1997).

## 2.6 Adjectives

Adjectives in Lhasa 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:

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. (Francke (1929) presents a short list of adjectives which he says still choose a masculine or feminine suffix in concord with the head noun, but this pattern seems to be lost in contemporary spoken Lhasa.)

## 2.7 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 Mazaudon 1978; 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:

*mog = mog zim = po bzo = mkhan*  
 momo delicious cook = NOM  
 'one who makes good momos'

*mog = mog zim = po bzo = mkhan bu = mo de*  
 momo delicious cook = NOM girl that  
 'the girl who makes good momos'

*mo-s bzos-pa*  
 'what she cooked'

*mo-s bzos-pa-'i mog = mog*  
 she-ERG cooked-NOM-GEN momo  
 'the momos which she made'

*mo-s bzo-yag-gi mog = mog*  
 she-ERG cook-NOM-GEN momo  
 'the momos which she makes'

*mo-s mog = mog bzo = sa*  
 'the place where she makes momos'

*mo-s mog = mog bzo = sa(-'i) za = khang*  
 'the food shop 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 *=mkhan* in colloquial speech.

Relative clauses may also follow the head noun, with no genitive marking possible:

*mog = mog kho-s bzos-pa*  
 'the momos which she made'

## 3 THE VERB

### 3.1 The stem

The Classical Tibetan verb is marked for certain derivational (i.e. transitivity) and inflectional (tense/aspect) categories by a complex and only broadly regular system of prefixation, suffix-

ation, and ablaut. (See Chapter 16.) The merger of unprefixated 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 Lhasa can be found in Goldstein and Norngang (1970: 53–6) or Goldstein and Narkyid 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 *t/a* form based on the nominalizer *-pa* requires the past stem, and any form based on *=ki* requires the present stem:

*bsad-pa yin* /sǎʔpayii<sup>n</sup>/ killed (PERF CONJUNCT)  
*gsod=gi yin* /süʔkiyii<sup>n</sup>/ will kill (FUT CONJUNCT)

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:

<i>Intransitive</i>	<i>Transitive</i>
<i>khol</i> /khöʔ/	<i>bkol</i> /köö/ boil
<i>chag</i> /chaʔ/	<i>bcag</i> /caʔ/ break
<i>ril</i> /rii/ fall	<i>dbril</i> /rii/ knock someone down

### 3.2 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 *t/a/e* system consists of (at least etymologically) nominalized verb stems in construction with one of the copular verbs *yin* /yii<sup>n</sup>/ and *red* /reʔ/ or the existentials *yod* /yöʔ/ and *'dug* /tuʔ/. This core paradigm, with rough glosses, is (for further discussion of the use and meaning of several of these forms, see Chang and Chang 1981; DeLancey 1990; Jin 1979, 1983a, b):

<i>-pa yin</i>	/pa-yii <sup>n</sup> /	perfective conjunct
<i>-pa red</i>	/pa-reʔ/	perfective disjunct
<i>-ki yin</i>	/ki-yii <sup>n</sup> /	future conjunct
<i>-ki red</i>	/ki-reʔ/	future disjunct

<i>-ki yod</i>	/ki-yöʔ/	imperfective conjunct
<i>-ki 'dug</i>	/ki(-tuʔ)/	imperfective disjunct

(The /-tuʔ/ of the imperfective non-volitional form is regularly omitted in spoken Lhasa.)

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 Lhasa by *phyin*; the other source verbs are still in use as main verbs. The other suffixes are:

<i>-song</i>	/-soo <sup>n</sup> /	direct-evidential perfective from an obsolete suppletive past stem of 'go'
<i>-bzhang</i>	/-šaʔ/	indirect-evidential perfective from <i>bzhang</i> 'put'
<i>-byung</i>	/-chuu <sup>n</sup> /	perfective, speaker as Goal from 'byung 'appear', 'come to pass'
<i>-yong</i>	/-yoo <sup>n</sup> /	gnomic, future grounded in inherent tendency from <i>yong</i> 'come'

### 3.3 Conjunct/disjunct marking and evidentiality

Modern Tibetan and several other Bodic languages show a peculiar evidential pattern which Hale (1980) has labelled 'conjunct' vs 'disjunct' in Newari. In the copular system, conjunct forms occur with first person subject in statements and second person subjects in questions, disjunct forms elsewhere (for additional discussion, see Chang and Chang 1984; DeLancey 1990, 1992; Agha 1993; Tournadre 1992, 1996b):

*nga bod=pa yin*  
 I Tibetan (person) be/CONJUNCT  
 'I am a Tibetan.'

*kho bod=pa red*  
 he Tibetan be/DISJUNCT  
 'He is a Tibetan.'

*khyed=rang bod=pa yin pas*  
 you Tibetan be INTERR  
 'Are you a Tibetan?'

*nga rgya=mi red pas*  
 I Chinese=person be INTERR  
 '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:

*khos kho bod=pa yin zer=gyis*  
 he (ERG) he Tibetan be/CONJUNCT say-IMPF/DISJUNCT  
 'He<sub>i</sub> says that he<sub>i</sub> is a Tibetan.'

*khos kho bod=pa red zer=gyis*  
 he (ERG) he Tibetan be/DISJUNCT say-IMPF/DISJUNCT  
 'He<sub>i</sub> says that he<sub>j</sub> is a Tibetan.'

*khos nga bod=pa red zer=gyis*  
 he (ERG) I Tibetan be/DISJUNCT say-IMPF/DISJUNCT  
 'He says that I am a Tibetan.'

The existentials likewise show a distinction between conjunct *yod* and disjunct 'dug.

The distribution of conjunct and disjunct forms is not strictly based on person; both can occur with first person, with an evidential distinction:

*nga-r dngul tog=tsam yod*  
 I-DAT money some exist/CONJUNCT  
 'I have some money.'

*ngar dngul tog=tsam 'dug*  
 I (DAT) money some exist/DISJUNCT  
 'I have some money.'

Use of disjunct 'dug here 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 Goldstein (1973: 21–2), Chang and Chang 1984; DeLancey 1990, 1992.

The verbal endings which include one of the copulas have the same conjunct/disjunct value as the copula, with an additional restriction: the conjunct forms can occur only with volitional verbs:

*nga (s) thang=ka 'gel-gyi yod*  
 I (ERG) thangka hang-IMPF/ CONJUNCT  
 'I am hanging up thangkass (religious paintings).'

*kho (s) thang=ka 'gel-gyis*  
 s/he(ERG) thangka hang-IMPF/DISJUNCT  
 'S/he is hanging up thangkass.' (based on the speaker's direct perception)

*nga mthong- gis (\*gi yod)*  
 I see IMPF/DISJUNCT (\*CONJUNCT)  
 'I see [it].'

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:

*khos dkaryol bcag-song*  
 s/he (ERG) cup broke-PERF/EVIDENTIAL  
 'S/he broke the cup.'

*khos dkaryol bcag-zhag*  
 s/he (ERG) cup broke-PERFECT/INFERENCE  
 'S/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 (Jin 1979; DeLancey 1985):

- khong pbebs-byung*  
 he (HON) went (HON) -BYUNG  
 'He came (to where I was).'
- kho-s nga-r yige cig 'bris-byung*  
 he-ERG I-DAT letter a wrote-BYUNG  
 'He wrote me a letter.'
- kho-s nga-r gzugs-byung*  
 he-ERG I-DAT hit-BYUNG  
 'He hit me.'
- nga-s kho = tsho mthong-byung*  
 I-ERG they see-BYUNG  
 'I saw them.'

### 3.4 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 Lhasa. Many of these also occur as main verbs. A given auxiliary requires either the perfective or the imperfective stem.

Auxiliaries requiring the present stem:

- 'gro* /tɔ/ distal motion (from 'go')  
 (perf. stem *phyin* /chi<sup>n</sup>/)
- yong* /yɔ<sup>n</sup>/ proximal motion (from 'come')
- bsdad* /tã<sup>?</sup>/ progressive (from 'sit', 'stay')
- ran* /rã<sup>n</sup>/ be time to, almost (from 'appropriate')
- thub* /thup/ can
- nus* /nũ<sup>?</sup>/ dare
- shes* /ʃe<sup>n</sup>/ know how to (from 'know')
- 'dod* /tö<sup>?</sup>/ want to

Auxiliaries requiring the past stem:

- bzhag* /ʃã<sup>?</sup>/ do with deleterious effect (from 'put')
- tshar* /tshaa/ completive perfect (from 'finish')
- myong* /m̃ɔ<sup>n</sup>/ experiential perfect (from 'taste')
- chog* /chok/ may

### 3.5 Reduplication

Lhasa 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':

- nga-'i yi = ge 'di bris = bris-pa yin*  
 I-GEN letter DEM wrote = wrote-PERF/DISJUNCT  
 'This letter of mine was written long since.'

*kho phyin=phyin-pa red*  
 he went = went-PERF/DISJUNCT  
 'He comes and goes; keeps going back and forth.'

Tense specification in this construction is provided by the alternation between present and past stem:

*kho-s phyin=phyin-pa red*  
 he-ERG went = went-PERF/DISJUNCT  
 'He used to go regularly, went for a while.'

*kho 'gro='gro-ba red*  
 he go = go-PERF/DISJUNCT  
 'He goes regularly.'

For other uses of reduplication, see Wang 1988.

### 3.6 Nominal t/a/m constructions

Some verb suffixes require a copular or existential construction, for example *rtsis* /tsi?/ 'plan to', as in:

*nga 'gro=rtsis yod*  
 I go = plan exist  
 'I'm planning to go.'

Phonologically *rtsis* shows compound phonology. (Note also that *rtsis* governs vowel harmony in its 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* /loo<sup>n</sup>/, 'have time to' and *grab(s)* /tap/ 'about to', 'ready to'.

## 4 WORD FORMATION

### 4.1 Nouns

A substantial number of common nouns are monomorphemic and hence monosyllabic: *mi* /mi/ 'person', *khyi* /kh<sup>n</sup>i/ 'dog', *gri* /thi/ 'knife', *bod* /phö?/ '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* /lap<sup>n</sup>ã/ '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* = *zam* 'bamboo bridge' (*smyug* = *ma* 'bamboo', *zam-pa* 'bridge').

#### 4.2 'Light' verbs

There are no true compound verbs, but Tibetan, like many other South and West Asian languages, has a very large set of 'light' verbs, which consist of a noun in construction with a semantically fairly empty verb. There is a small set of extremely common verbs which account for most of these. For example:

<i>skad gtong</i>	call, invite	( <i>skad</i> speech)
<i>gru gtong</i>	sail	( <i>gru</i> boat)
<i>skad</i> = <i>dpar gtong</i>	play records	( <i>skad</i> = <i>dpar</i> phonograph)
<i>skad rgyag</i>	shout	( <i>skad</i> speech)
<i>gom-pa rgyag</i>	walk	
<i>rgyugs</i> = <i>šad rgyag</i>	comb	( <i>rgyugs</i> = <i>šad</i> a comb)
<i>chams-pa rgyag</i>	catch cold	( <i>chams-pa</i> a cold)

#### 4.3 Honorifics

Central Tibetan, and other dialects such as Ladakhi and Dzongkha which have historically been spoken in a monarchical context, have a system of honorific vocabulary called *zhe = sa*. (Kitamura 1975; Agha 1993; DeLancey 1998). Honorific speech is used with and among people of relatively high social standing, reflecting the absolute social status of the addressee, rather than relative status of speaker and addressee.

Besides the honorific pronouns (Section 2.2), there is an honorific suffix *lags* /laâ/ suffixed to names and terms of address, in either 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<sup>n</sup>/ 'do':

*kho-s yi = ge bris-song*  
 he-ERG letter write-PERF/DIRECT  
 'He wrote a letter (plain).'

*khong-gis phyag = bris bris = gnang-song*  
 3rdHON-ERG letter (HON) write = HON-PERF/DIRECT  
 'S/he wrote a letter (honorific).'

## 5 SYNTAX

### 5.1 Clause and sentence

Unmarked constituent order within the clause is SOV, but any preverbal constituent can be fronted:

*phru = gu-s mog = mog sgan = ga zas-zhag*  
 child-ERG momo all ate-PERF/INFERENTIAL  
 'The children at all the momos.'

*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 Lhasa is defined by the occurrence in final position (or directly preceding a final question particle) of a verb with marking for a complex tense/aspect/evidentiality category. An inflected Lhasa 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. In narrative or expository discourse 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:

*stag-gis gyag-la so brgyab-byas bsad-pa red*  
 tiger-ERG yak-DAT bit NF kill PERF/DISJUNCT  
 'The tiger bit the yak and killed it.'

*kho-s kha=lags zas-byas phyin-pa red*  
 he-ERG meal ate NF went PERF/DISJUNCT  
 'He ate and left.'

The non-final marker (glossed NF) is obligatory in such constructions, and thus serves as a clear diagnostic to distinguish fortuitous concatenation from grammaticalized serial verbs. In this chapter the latter are discussed in Section 3.4.

## 5.2 Adverbial clauses and subordinators

The simplest subordinating construction consists of the bare verb stem with a clitic subordinating particle: *na* 'if', *tsang* 'because', *nas* /näʔ/ and *byas* /cäʔ/ 'non-final', etc. 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* /khonla/ with negated verb 'before':

*kho rgya=gar-la ma-gro khong-la dngul mang-po bsags-pa red*  
 kho kʰäkaala ma<sup>n</sup>t̚o khonla ŋüü mango saʔpareʔ  
 he India-to NEG-go before money much save-PERF  
 'Before he went to India, he saved a lot of money.'

The nominalized past stem with the locative *-r* indicates 'in order to' or, with a negated verb, 'without':

*kho-s nga ma-bsgugs-pa-r phyin-zhag*  
 khöʔ ŋä maküʔpar chi<sup>n</sup>sa  
 he-ERG I NEG-wait-NOM-LOC went-PERF/INFERENTIAL  
 'He left without waiting for me.'

### 5.3 Complementation

Most complement clauses are nominalizations. The nominalizers most commonly used as complementizers are *-pa*, *=yag* (Section 2.6), *=rgyu*, and *=gag*. The complement of a verb of perception or cognition can be a finite clause nominalized with *-pa*:

*nga-s kho las=ka byed=kyi yod-pa mthong-byung*  
 I-ERG he work do=IMPFV exist-NOM see-PERF/CONJUNCT GOAL  
 ‘I see he is working.’

Otherwise complement constructions involve an uninflected verb stem plus a nominalizer. The *=gag* construction generally indicates some sort of purpose or intention:

*nga-s ‘di-r khyed=rang-la skad=cha bshad=gag yongs-pa yin*  
 I-ERG DEM-LOC you-LOC speech talk=NOM came-PERF/CONJUNCT  
 ‘I came here to talk to you.’

The general-purpose complementizers are *-pa*, *=yag*, and *=rgyu*.

*nga-s kho-r bshad=rgyu khas=len byas-pa yin*  
 I-ERG he-LOC tell-NOM promise did-PERF/CONJUNCT  
 ‘I promised to tell him.’

*kho-s las=ka byed=yag ‘go btsugs-song*  
 he-ERG work do=NOM start-PERF/DISJUNCT  
 ‘He started to work.’

*nga bod-la phyin-pa kho-s shes=kyi*  
 I Tibet-LOC went-NOM he-ERG know-IMPV/DISJUNCT  
 ‘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:

*‘di-‘i skad=cha dris=rgyu gus=zhabz med-pa red*  
 this-GEN question ask=NOM polite not.be-PERF/DISJUNCT  
 ‘It’s not polite to ask about this.’

*‘di-îi skad=cha dris-pa gus=zhabz med-pa red*  
 this-GEN question ask=NOM polite not.be-PERF/DISJUNCT  
 ‘It’s not polite to have asked about this.’

Finite complements occur with verbs of speech. These may be unmarked or marked with the complementizer *ze*, usually pronounced as a semi-syllabic [s]:

*kho-s nga-s bcag-pa red zer=gysis*  
 he-ERG I-ERG broke-PERF/DISJUNCT say=IMPF/DISJUNCT  
 ‘He says I broke [it].’

*kho-s nga-s bcag-pa red-ze zer=gysis*  
 /khö? ŋä? caâpareês seeki/  
 ‘idem’.

This *ze* is a grammaticalized form of *zer* ‘say’, but now freely co-occurs with it, as in the above example.

## 5.4 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:

*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:

*nga gnyid 'khugs ma-byung*  
 I sleep fall NEG-PERF/1ST  
 'I couldn't get to sleep.'

In tense/aspect forms based on copulas, the copula carries the negation:

*nga = gnyis thug = gi ma-red*  
 I = two meet = NONPERF NEG-BE  
 'We two won't [be able to] meet.'

*kho za = gi mi-'dug*  
 he eat = NONPERF NEG-EXIST  
 'He isn't eating.'

The conjunct copulas *yod* and *yin* have special negative forms, *med* /meʔ/ and *min* /miʔ/.

Yes-no questions are marked by the final particle *pas* /-pää/, which assimilates to a preceding (underlying) velar:

*kho bod-pa red-pas*  
 /kho phȫpa rɛpää/  
 he Tibetan be-Q  
 'Is he a Tibetan?'

*khyed = rang bod-la phebs myong-ngas*  
 /kh<sup>y</sup>eerəŋ phȫʔla pheb ñöŋää/  
 you Tibet-LOC go (HON) EXPERIENTIAL-Q  
 'Have you ever been to Tibet?'

*khyed = rang chams = pa brgyab 'dug-gas*  
 /kh<sup>y</sup>eerəŋ champa k<sup>y</sup>ap tɯkää/  
 you cold (have) EXIST-Q  
 'Do you have a cold?'

Information questions have the interrogative word in its expected sentence position. The interrogative words are *su* /su/ 'who' and a series based on *ga*: *ga = re* /khare/ 'what', *ga = pa-r* /khapaal/ 'where', *ga = dus* /khatüʔ/ 'when', *ga = 'dra* /khanta/ 'how'. These questions may optionally be marked with the final particle *ga* /kha/:

*slob = gra tshar-na khyed = rang ga = re gnang-ga*  
 /lapʂa tshaana kh<sup>y</sup>eerəŋ khare naŋkha/  
 school finish-when you what do(HON)-Q  
 'What will you do when you finish school?'

## REFERENCES

- Agha, Asif (1993) *Structural Form and Utterance Context in Lhasa Tibetan: Grammar and Indexicality in a Non-Configurational Language*, New York: Peter Lang.
- Chang, Betty Shefts and Chang, Kun (1980) 'Ergativity in spoken Tibetan', *BIHP* 51.1: 15–32.
- Chang, Betty Shefts and Chang, Kun (1981) 'Perfective and imperfective in spoken Tibetan', *BIHP* 52.2: 303–21.
- Chang, Betty Shefts and Chang, Kun (1983) 'Tense and aspect in spoken Tibetan', in Ernst Steinkellner and Helmut Tauscher, (eds) *Contributions on Tibetan language, History, and Culture, proceedings of the Csoma de Körös Symposium held at Velm-Vienna, Austria, 13–19 September 1981*, Vol. 1, 329–37, (Wiener Studien zur Tibetologie und Buddhismuskunde, Heft 10), Vienna: Arbeitskreis für tibetische und buddhistische Studien, Universität Wien.
- Chang, Betty Shefts and Chang, Kun (1984) 'The certainty hierarchy among spoken Tibetan verbs of being', *BIHP* 55.4: 603–34 (on evidential values of Lhasa copulas).
- Chang, Kun and Chang, Betty Shefts (1968) 'Vowel harmony in spoken Lhasa Tibetan', *BIHP* 40: 53–124.
- Chang, Kun and Chang, Betty Shefts (1978–81) *Spoken Tibetan Texts*, Taipei: Inst. of History and Philology, Academia Sinica (In four volumes; volume 1 includes the phonological analysis, with considerable phonetic detail, which has become standard in much work in the US, including Goldstein's pedagogical and reference works).
- Chang, Kun and Shefts, Betty (1964) *A Manual of Spoken Tibetan (Lhasa Dialect)*, Seattle: University of Washington Press.
- Chang, Kun and Shefts, Betty (1965) 'A morphophonemic problem in the spoken Tibetan of Lhasa', *JAOS* 85: 34–9.
- Dawson, Willa (1980) 'Tibetan phonology' unpublished PhD dissertation, University of Washington (a generative treatment of vowel harmony and tone, for the most part following Chang and Chang's descriptive analysis).
- DeLancey, Scott (1984) 'Transitivity and ergative case in Lhasa Tibetan', *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, 131–40.
- DeLancey, Scott (1985) 'On active typology and the nature of agentivity', in F. Plank, (ed.) *Relational Typology*, pp. 47–60, Berlin and New York: Mouton.
- 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 (1997) 'Grammaticalization and the gradience of categories: Relator nouns and postpositions in Tibetan and Burmese', in J. Bybee, J. Haiman, and S.A. Thompson (eds) *Essays on Language Function and Language Type*, pp. 51–69, Benjamins.
- DeLancey, Scott (1998) 'Semantic categorization in Tibetan honorific nouns', *Anthropological Linguistics* 40: 1–15.
- DeLancey, Scott (1999) 'Relativization in Tibetan', in Yogendra Yadava and Warren Glover (eds) *Studies in Nepalese Linguistics*, pp. 231–49, Kathmandu: Royal Nepal Academy.
- Francke, A.H. (1929) 'Addenda', in H.A. Jäschke (ed.) *Tibetan Grammar*, Berlin: de Gruyter.
- Goldstein, Melvyn (1973) *Modern Literary Tibetan*, Occ. Papers of the Wolfenden Society on Tibeto-Burman Linguistics 5, Urbana, IL: Center for Asian Studies, University of Illinois (a detailed exposition of the modern written standard, with detailed discussion of the semantics of verb forms, subordinators, and other constructions, including evidential and conjunct/disjunct forms).
- Goldstein, Melvyn (1975) *Tibetan–English Dictionary of Modern Tibetan*, Kathmandu: Ratna Pustak Bhandar (the only thorough and reliable work of its kind).
- Goldstein, Melvyn and Ngawangthondup Narkyid (1984) *English–Tibetan Dictionary of Modern Tibetan*, Berkeley: University of California Press. (the only thorough and reliable work of its kind).
- Goldstein, Melvyn and Normang Nawang, (1970) *Modern Spoken Tibetan: Lhasa Dialect*, Seattle: University of Washington Press, 2nd edn, Kathmandu: Ratna Pustak Bhandar, 1978 (includes

- useful explanation of many points of phonology, syntax, and semantics, as well as tabular presentations of verb stem alternations and tense/aspect forms).
- Hale, Austin (1980) 'Person Markers: Finite Conjunct and Disjunct Verb Forms in Newari', *Papers in South-East Asian Linguistics*, no. 7, ed. by R. Trail *et al.*, Canberra: Australian National University, 95–106 (not concerned with Tibetan, but the first cohesive analysis of the conjunct/disjunct phenomenon).
- Hari, Anna (1979) *An Investigation of the Tones of Lhasa Tibetan*, Huntington Beach, CA: Summer Institute of Linguistics (very useful phonological data, especially for phonological processes in word formation. The author reports a hitherto unknown tonal contrast which, however, other investigators have been unable to replicate).
- Hu, Tan (1982) 'Recherches sur les tons du Tibétain (dialecte de Lhasa)', *Cahiers de Linguistique, Asie Orientale* 9.1: 11–46 (includes a summary of various competing analyses of the Lhasa tonal system).
- Jäschke, H.A. (1881) *A Tibetan–English Dictionary*, London: Routledge & Kegan Paul (more oriented to the classical than the modern language, but a standard reference work in western research on all varieties of Tibetan).
- Jin Peng (1958) *Zangyu Lasa Rigidze Changdou hua-de bijiao yanjiu* (comparative study of Lhasa, Shigatse, and Chamdo Tibetan), Beijing: Lexue Chubanshe: (a wealth of data, including a detailed description of Shigatse, another major dialect of the Central group).
- Jin Peng (1979) 'Lun Zangyu Lasa-koyu 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), *MZYW* 1979.3: 173–81 (an analysis of the conjunct/disjunct system. Some of the data presented here I have not been able to replicate with my Lhasa consultants).
- Jin Peng (1983a) 'Zangyu Lasa-hua dongci-de shi ji qi biaoda fangfa' (the moods of Tibetan (Lhasa dialect) verbs and the ways to represent them), *MZYW* 1983.1: 9–18.
- Jin Peng (1983b) 'Zangyu jianzhi' (outline of Tibetan), Beijing: Minzu Chubanshe.
- Kitamura, Hajime (1975) 'The Honorifics in Tibetan', *Acta Asiatica* 29: 56–74.
- Mazaudon, Martine (1977) 'Tibeto-Burman tonogenetics', *LTBA* 3.2.
- Mazaudon, Martine (1978) 'La formation des propositions relatives en tibétain', *Bulletin de la Société de Linguistique de Paris* 73: 401–14.
- Mazaudon, Martine (1984) 'Review of Hari', *LTBA* 8.1: 91–9.
- Nishi, Yoshio (1986) 'Gendai Chibetto-go hogen-no bunrai' (a classification of Tibetan dialects), *Bulletin of the National Museum of Ethnology* 11: 837–901.
- Shefts, Betty, and Chang, Kun (1967) 'Spoken Tibetan morphophonemics: p', *Language* 43: 512–25.
- Sprigg, R.K. (1954) 'Verbal phrases in Lhasa Tibetan', *BSOAS* 16: 134–56 (primarily concerned with the phonology of the verbal piece, presented in a Firthian prosodic framework).
- Sprigg, R.K. (1955) 'The tonal system of Tibetan (Lhasa dialect) and the nominal phrase', *BSOAS* 17: 133–53.
- Sprigg, R.K. (1961) 'Vowel harmony in Lhasa Tibetan', *BSOAS* 24: 116–38 (a detailed analysis, presented in a Firthian prosodic framework).
- Sprigg, R.K. (1981) 'The Chang-Shefts tonal analysis, and the pitch variation of the Lhasa-Tibetan tones', *LTBA* 6.1: 49–59 (very useful discussion of tone in polysyllabic constructions).
- Tournadre, Nicolas (1992) 'La deixis en tibétain: quelques faits remarquables', in M-A. Morel and L. Danon-Boileau (eds) *La Deixis : colloque en Sorbonne, 8–9 juin 1990*, 197–207. Paris: Presses universitaires de France.
- Tournadre, Nicolas (1996a) *L'Ergativité en tibétain: approche morphosyntaxique de la langue parlée*. Louvain: Peeters.
- Tournadre, Nicolas (1996b) 'Comparaison des systèmes médiatifs de quatre dialects tibétains', in Zlatka, Guentchéva (ed.) *L'Énonciation médiatisée*; Paris: Peeters 195–211, (on evidentiality in four Tibetan dialects).
- Wang Huiyin (1985) 'Liaolun Zang-yu Lasa-hua qingsheng-de xingzhi – jiantan qingsheng yu jufa jiegou-de guanxi' (the characteristics of light tone of Lhasa dialect – the relations between light tone and syntax), *ZMXX* 1985, no. 4: 86–91 (on the tonal phonology of grammatical elements).

- Wang, Huiyin (1988) 'Zangyu Lasa hua dungcide zhongdie xingshi' (reduplicated forms of verbs in Lhasa Tibetan), *MZYW* 1988.3: 36–43.
- Wylie, Turrel (1959) 'A standard system of Tibetan transcription', *Harvard Journal of Asiatic Studies* 22: 261–7.
- Zimmerman, Heinz (1979) *Wortart und Sprachstruktur im Tibetischen*, Wiesbaden: Harrasowitz (the only available attempt at an extended syntactic analysis of Tibetan. Includes very useful descriptions of word composition).

**PART 6**

---

# **TGTM LANGUAGES**

---





# TAMANG

*Martine Mazaudon*

## 1 INTRODUCTION

Tamang is spoken by about a million speakers in central Nepal. The particular dialect described here is spoken in the village of Risiangku, Sindhu Palchok 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), Manangpa, the Nar-Phu dialects, Chantyal and the dialect of Tangbe, in Mustang zone. Together they form what Robert Shafer (1955) has identified as the Gurung Branch of the Bodish Section of the Bodic Division of Tibeto-Burman.<sup>1</sup>

Tamang was included under the name of Murmi in Grierson's *Linguistic Survey of India*. A few studies have appeared on its phonology, but very little on its grammar. On Western Tamang see Taylor 1973 and Everitt 1973. Some texts in Western Tamang have been published in Hale and Pike 1970 and Hoefler 1981–97. On Eastern Tamang see Mazaudon 1978b and 1988, and Yoncan 1997.

On other languages of the TGTM group the reader can consult Glover 1974 for Gurung, Georg 1996 for Marphali, and Noonan in this volume for Chantyal and Nar-Phu. The next closest language on which grammatical information is available is Tibetan.

## 2 PHONOLOGY

### 2.1 Initial consonants

	<i>labial</i>	<i>dental</i>	<i>sibilant affricate</i>	<i>trilled retroflex</i>	<i>velar/ glottal</i>
aspirated	p <sup>h</sup>	t <sup>h</sup>	c <sup>h</sup>	ʈ <sup>h</sup>	k <sup>h</sup>
unaspirated	p	t	c	ʈ	k
nasal	m	n			ŋ
continuant		l	s	r	h
semi-vowel	(j, w)				

---

1 M. Noonan 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). I refer to this group by the initials of its main ethnic groups, TGTM (Tamang-Gurung-Thakali-Manang).

## 2.2 Initial consonant clusters

	p <sup>h</sup>	p	m	t <sup>h</sup>	t	n	l	c <sup>h</sup>	c	s	t <sup>h</sup>	t	r	k <sup>h</sup>	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 [~].

## 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: <sup>1</sup>a, high falling; <sup>2</sup>a, mid-high level; <sup>3</sup>a, mid-low level; <sup>4</sup>a, 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, <sup>1</sup> and <sup>2</sup>, have clear voice quality, words in the low tones have a breathy voice quality.

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, and the archiphonemic series in that position is pronounced unaspirated with a slight optional voicing. This correlation derives from the origin of the tonal system. The devoicing of an old voiced series in Proto-Tamang,

which merged with an old voiceless unaspirated series, gave rise to the two low tones through a process of merger and compensatory split of the old two-tone system (Mazaudon 1977, 1978a).

## 2.7 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 monosyllabic.

## 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 aspects 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 (± PERIPHERAL ARGUMENTS) V

*Active*

- (1) <sup>2</sup>*ai-Ø*    <sup>3</sup>*naŋkar*    <sup>1</sup>*ni-pa ?*  
 you-ABS tomorrow go-IMPV  
 'Are you going tomorrow?'

*Stative*

Stative verbs carry verbal suffixes. They can be used in the perfective with an inchoative meaning.

- (2) *o,*    <sup>3</sup>*cer-ci*    <sup>4</sup>*tamo*  
 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).

4.1.2 *Non-verbal predicate***S-ABS PREDICATE (± COPULA)***Adjective predicate*

- (3) <sup>2</sup>cu <sup>3</sup>mento <sup>3</sup>caca  
 this flower small  
 'This flower is small.'

The absence of a copula is very frequent. But there exist two main copulas and some others. <sup>1</sup>mu-la is mostly used for existence and attribution, and <sup>3</sup>hin-pa for identification.<sup>2</sup>

Several other verbs can be used as copulas :

- (4) <sup>1</sup>p<sup>h</sup>jukpa-la <sup>3</sup>caca <sup>2</sup>kola-m <sup>2</sup>os-pa <sup>3</sup>a <sup>1</sup>k<sup>h</sup>a-pa-<sup>1</sup>ka  
 rich-GEN small child-TOP thus-NER not come-IMPV-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 <sup>1</sup>ta-ci expresses excess.

- (5) <sup>3</sup>caca <sup>1</sup>ta-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.

*Noun predicate*

The meaning difference of the basic copulas with an adjectival predicate is clear, but it does not seem to hold when the predicate is a noun.

- (6) <sup>1</sup>pa-la <sup>1</sup>apa <sup>1</sup>c<sup>h</sup>oŋpa {∅} /<sup>3</sup>hin-na /<sup>1</sup>mu-la}  
 I-GEN father merchant {∅} /be- NONPST /exist-NONPST}  
 'My father is a merchant.'

*Locative predicate*

- (7) <sup>4</sup>kle <sup>1</sup>it <sup>3</sup>naŋ-ri <sup>1</sup>mu-la  
 king this inside-LOC be- NONPST  
 'The king is inside this [bag].'

*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) <sup>1</sup>pa-la <sup>4</sup>came <sup>4</sup>ni: <sup>1</sup>mu-la  
 I-GEN daughter two be- NONPST  
 'I have two daughters.'

2 In Risiangku Tamang the copulas <sup>1</sup>mu-la and <sup>3</sup>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. <sup>1</sup>mu-la cannot be negated; a substitutive root <sup>3</sup>are is used instead.

- (9) <sup>1</sup>*ŋa-la* <sup>1</sup>*ka*  
 I-GEN FOC  
 ‘It is mine.’

The genitive is the most commonly used case for possession, but the dative is also used:

- (10) <sup>4</sup>*tap* <sup>1</sup>*mu-la* <sup>2</sup>*ani-ta?*  
 needle be-NONPST 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) <sup>1</sup>*ŋa-ta* <sup>3</sup>*pa-u*  
 I-DAT give-IMP  
 ‘Give it to me!’
- (12) <sup>4</sup>*came* <sup>4</sup>*ca* <sup>1</sup>*ʰe-la* <sup>1</sup>*am-la-n* <sup>1</sup>*pin-ci*  
 daughter TOP she-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) <sup>2</sup>*nam* <sup>1</sup>*kʰa-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 absolutive 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) <sup>2</sup>*naka-se* <sup>3</sup>*tap-∅* <sup>1</sup>*ca-ci*  
 chicken-ERG vegetable-ABS eat-PFV  
 ‘The chicken is eating the vegetable.’

Perception predicates follow the direct transitive pattern.

- (15) <sup>1</sup>*ŋa-i* <sup>4</sup>*jo:-∅* <sup>1</sup>*mraŋ-ci*  
 I-ERG thief-ABS see-PFV  
 ‘I saw the thief.’

### s-ERG o-DAT v

- (16) <sup>1</sup>*mam-se* <sup>2</sup>*kol'-kat'-ta* <sup>3</sup>*paŋ-pa*  
 Grandma-ERG children- DAT scold-IMPV  
 ‘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 absolutive. 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) <sup>1</sup>am'-se    <sup>2</sup>kol'-ta    <sup>1</sup>kan-Ø    <sup>2</sup>k<sup>h</sup>waɾ-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) <sup>2</sup>aŋa-ta    <sup>1</sup>pet-pa?  
 sister-DAT shy-IMPV  
 'Do you feel shy?' (said to a young girl addressed as 'sister')

With some other verbs, the experiencer is in the absolutive 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) <sup>1</sup>ŋa-ta    <sup>1</sup>t<sup>h</sup>e    <sup>1</sup>nai-pa  
 I-DAT he disgust-IMPV  
 'I disgust him.'

Inanimate objects of feelings require a subordinate clause:

- (20) <sup>1</sup>kli    <sup>1</sup>mraŋ-si    <sup>1</sup>ŋa    <sup>1</sup>nai-ci  
 faeces see-having I disgust-PFV  
 'Excrement disgusts me.' (*lit.* 'Seeing faeces I feel disgust'.)

a construction which can also be used with animate objects as in:

- (21) <sup>1</sup>ŋa    <sup>1</sup>mraŋ-si    <sup>2</sup>ai    <sup>1</sup>pet-pa?  
 I see-having you shy-IMPV  
 '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) <sup>1</sup>t<sup>h</sup>e-ta    <sup>3</sup>ro    <sup>1</sup>kik    <sup>1</sup>por-pa    <sup>1</sup>ce-n  
 he-DAT friend one take-IMPV 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,<sup>3</sup> as opposed to the expected ergative (23), although topicalization does not necessarily imply such a shift in case marking: compare (89).

(23) *inam bakas k<sup>h</sup>up <sup>1</sup>jaŋ-ci <sup>4</sup>jaŋ-la p<sup>h</sup>alana 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) *<sup>2</sup>os <sup>1</sup>ta-m’ <sup>2</sup>ai-ta inam bakas <sup>3</sup>a <sup>1</sup>jaŋ*  
 thus happen-when you-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, Adjective<sub>1</sub>, Adjective<sub>2</sub>, Noun, Numeral, {Case + Particle/  
 Particle + Case}<sup>4</sup>

### 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. 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: <sup>1</sup>ŋa ‘I’, <sup>2</sup>ai ‘you’, <sup>1</sup>h<sup>e</sup> ‘he’, <sup>4</sup>jaŋ ‘us (inclusive)’, <sup>1</sup>in ‘us (exclusive)’, <sup>2</sup>ai-ni ‘you (plural)’, <sup>1</sup>t<sup>h</sup>e-n ‘they’. The two plural pronouns of the first

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 The order of these elements depends on the tonal or non-tonal status of the particle.



person are often used instead of the singular. ‘You’ is often replaced, for politeness, by a kinship term (10, 18), or by the intensifier <sup>3</sup>*raŋ* ‘oneself’.

The demonstratives <sup>2</sup>*cu* ‘this’ and <sup>1</sup>*oca* ‘that’, and the nouns <sup>3</sup>*mi* ‘man’, and <sup>3</sup>*ro* ‘companion’ are also often used as pronouns.

- (25) <sup>3</sup>*mi-la*       $\emptyset$       <sup>4</sup>*tamom*    <sup>3</sup>*rap-si*      <sup>3</sup>*cin-ci*  
 person-GEN [thing] now      play-having    finish-PFV  
 ‘Her [tape-recorder] has now stopped playing.’

Personal or demonstrative pronouns take the same case markers as nouns, except for <sup>1</sup>*ŋa* ‘I’ with which a variant of the ergative is found <sup>1</sup>*ŋa-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 pronoun <sup>3</sup>*raŋ* ‘in person’, ‘one’s own’ can be used for the reflexive.

- (26) <sup>3</sup>*ponpo-se*    <sup>3</sup>*raŋ-ta-n*    <sup>3</sup>*kjaŋ-si*    <sup>1</sup>*cjaɹ-pa*  
 shaman-ERG self-DAT-INT hit-PART see-IMPV  
 ‘The shaman checks by hitting himself [with the hot broom that it is not too hot for the patient].’

Reciprocity is expressed either by a verb with a conjoined subject:

- (27) <sup>2</sup>*o-te-ma*      <sup>4</sup>*niz-nun*    <sup>2</sup>*c<sup>h</sup>at-cim*,    <sup>3</sup>*rempa-then*    <sup>1</sup>*mriŋ*  
 that-much-when 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) <sup>3</sup>*rempa-ta*    <sup>1</sup>*mriŋ*    <sup>1</sup>*mriŋ-ta*    <sup>3</sup>*rempa*    <sup>1</sup>*nai-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 (<sup>1</sup>*tar* ‘white’, <sup>3</sup>*caca* ‘small’) although some of them end in a non alternating final syllable *-pa* (<sup>1</sup>*k<sup>h</sup>eppa* ‘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) <sup>4</sup>*kle*    <sup>4</sup>*kik*    ***ra:ni***    <sup>4</sup>*kik*    <sup>1</sup>*mu-pa*    <sup>2</sup>*cim*  
 king one queen one be- PAST EVID

<sup>1</sup>*oca* <sup>4</sup>*kle-h'en* *ra:ni-mi* <sup>1</sup>*p<sup>h</sup>jukpo-<sup>1</sup>p<sup>h</sup>jukpo* <sup>1</sup>*mu-pa* <sup>2</sup>*cim*  
 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) <sup>3</sup>*mi* <sup>1</sup>*k<sup>h</sup>a-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, 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: <sup>2</sup>*ai-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 multiplier: <sup>4</sup>*pokal* <sup>4</sup>*ni:* [twenty.two] 'forty'. Addition is indicated by the ablative *-se*: <sup>4</sup>*pokal* <sup>4</sup>*ni:-se* <sup>4</sup>*ni:* [twenty.two-from two] 'forty-two'.

The suffix *-te*, added to a number or measure (91) gives it an imprecise meaning.

Quantifiers like 'all', 'many', etc. are either used as full nouns (31) or as adverbs (see Section 6.3). They are not adjectives.

- (31) <sup>3</sup>*mokko-se-n* <sup>2</sup>*se:-pa*  
 all-ERG-INT know-IMPV  
 '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), *-h'en* 'sociative'. Temporal *-ma* is bound to some words (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, 95).

The genitive is also used to indicate price but expressions of amounts using a measure word do not use an overt genitive marker.

### 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. There is no relative pronoun, and the relativized argument is simply deleted inside the relative clause: <sup>4</sup>*tot* <sup>2</sup>*pwi-pa* <sup>3</sup>*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) <sup>1</sup>*k<sup>h</sup>a-pa-ta*    <sup>1</sup>*pin*    <sup>1</sup>*toɽ-pa*,    <sup>3</sup>*a*    <sup>1</sup>*k<sup>h</sup>a-pa-ta*    <sup>1</sup>*pin*    <sup>3</sup>*a*    <sup>1</sup>*toɽ*  
 come-NER-DAT    give    must-IMPV    not    come-NER-DAT    give    not    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 : 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) <sup>1</sup>*tu-kjuŋ-ma*    <sup>2</sup>*maca-ri*    <sup>2</sup>*ai-ni-kate*    <sup>3</sup>*toɽ-pa*  
 last year    down-LOC    you-PL-COLL    reach-NER  
<sup>1</sup>*oca*    <sup>1</sup>*mam-<sup>1</sup>k<sup>h</sup>wi*  
 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) <sup>1</sup>*sun*    <sup>4</sup>*karte*    <sup>1</sup>*toɽ-pa-cim*    <sup>2</sup>*ote*    <sup>1</sup>*tat-si*    <sup>2</sup>*toɽ-ko*  
 rice    how much    need-IMPV-EVID    that much    take-ing    pound-IMP  
 ‘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) <sup>1</sup>*sun*    <sup>4</sup>*karte*    <sup>1</sup>*ni-pa-cim*    <sup>2</sup>*to*    <sup>1</sup>*toɽ-ci*  
 rice    how much    go-IMPV-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: <sup>1</sup>*ni-ke* ‘let’s go!’ <sup>1</sup>*pin-sai* ‘let us give (alms)’ <sup>4</sup>*tik* <sup>1</sup>*la-sai* ‘what can we do?’

### 6.2 Word order

The verb phrase comprises the following elements, in order: adverb, negation, verb, directional, modal {tense-aspect 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’. 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 adverbs: between NP arguments and the verb (23, 36) or before a complex predicate (37).

- (36) <sup>1</sup>*sun* <sup>4</sup>*ca* <sup>4</sup>*la:mo* <sup>3</sup>*a* <sup>1</sup>*k<sup>h</sup>a* <sup>1</sup>*mu-la*  
 rice TOP much not come be-NonPst  
 ‘Rice does not grow well (*lit.* “much”) [here].’

- (37) *apai*, <sup>4</sup>*la:man* <sup>1</sup>*suŋ* <sup>1</sup>*k<sup>h</sup>eppa*  
 EXCL much mouth big  
 ‘Oh<sup>1</sup> He is very much a big-mouther (= he is very greedy).’

### 6.4 Negation

Negation precedes the verb. It has two forms, <sup>4</sup>*ta* for imperatives and optatives, and <sup>3</sup>*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’ (<sup>1</sup>*k<sup>h</sup>a-pa*): <sup>2</sup>*pit-ka-pa* ‘to send hither’.

Compounds are exceptional (<sup>2</sup>*pit-hur-pa* ‘to throw to someone’ from <sup>2</sup>*pit-pa* ‘to send’ and <sup>1</sup>*hur-pa* ‘to throw’).

A verbal expression consisting of a noun and weak verb sequence (<sup>1</sup>*sem* <sup>1</sup>*ta-pa* [feelings happen] ‘to like’, <sup>3</sup>*prot* <sup>1</sup>*k<sup>h</sup>a-pa* [taste come] ‘to be tasty’) can replace the verb and be constructed like a simple verb (38, 41).

- (38) <sup>1</sup>*ŋa-ta* <sup>3</sup>*twa-la* <sup>1</sup>*sja* <sup>3</sup>*prot* <sup>1</sup>*k<sup>h</sup>a* <sup>1</sup>*mu-la*  
 I-DAT pork-GEN meat taste come be-NonPst  
 ‘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 main verb + modal, or as a sequence 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 below under ‘complex sentences’ (see Section 7.3).

## 6.7 Tense-aspect and sentence final particles

Basic tense-aspects 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), future/irrealis, *-ul-ko* imperative, *-kel-kai* optative, *-sai* conditional. Only main clause or complete 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 past is expressed with the irrealis marker *-la*, followed by the intensifier *-i*: <sup>1</sup>*k<sup>h</sup>a-ci* [come-PFV] ‘he came’; <sup>3</sup>*a-<sup>1</sup>k<sup>h</sup>a-la-i* [NEG-come-IRR-INT] ‘he did not come’.

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<sup>1</sup>mula*, progressive; *V-si<sup>2</sup>cipa*, durative; etc.

The reported speech particle *-ro* is added to the tensed form of the verb (55, 91).

## 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*, conditional; *-si*, temporal succession; *-ma*, temporal simultaneity; *-na*, manner; *-pa-ri*, purpose; *-sil/-pa-sel/-pa-se<sup>1</sup>la-ma*, cause.

# 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) <sup>1</sup>*ni* <sup>4</sup>*ta* <sup>2</sup>*puŋ-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.

- (40) <sup>2</sup>*nam* {<sup>1</sup>*k<sup>h</sup>a-sam*/<sup>1</sup>*k<sup>h</sup>a-ci* <sup>3</sup>*pi-sam*} <sup>1</sup>*ŋa* <sup>3</sup>*a* <sup>1</sup>*k<sup>h</sup>a*  
 rain {*come-if/come- PFV* *say-if*} I not 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) <sup>2</sup>*ai-ni-kate* <sup>4</sup>*pu* <sup>2</sup>*mra* <sup>2</sup>*sjot-pa-ri* <sup>1</sup>*ni-ci*  
 you-PL-COLL field weeds pull out-NER-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) <sup>1</sup>*ni-la* <sup>1</sup>*ta-la*  
 go-FUT happen-FUT  
 'He might go.'

## 7.2 Sentential subjects and completives

### 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) <sup>1</sup>*tir-la* <sup>4</sup>*ni-ra* <sup>2</sup>*som-ra* <sup>1</sup>*ta-ci* <sup>1</sup>*ŋa* <sup>1</sup>*c<sup>h</sup>jai-pa-m*  
 yesterday two-days three-days happen-PFV I ready-IMP-FV-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 Completives: 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* <sup>3</sup>*naŋkar* <sup>3</sup>*pa-u* <sup>3</sup>*pi-si* <sup>3</sup>*pi-m*  
 money tomorrow bring-IMP say-ing say-EMPH  
 'He said that you should bring the money tomorrow.' (*lit.*: "'Bring the money tomorrow" having said he said'.)

Verbs of saying, intention, or desire are often constructed without any overt mark of subordination by two juxtaposed finite clauses.

- (45) <sup>1</sup>*ni-la* <sup>4</sup>*man-pa*  
 go-FUT like-IMPV  
 '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) <sup>1</sup>*ŋa-ta-m* <sup>1</sup>*tamo* <sup>2</sup>*aru-la* <sup>2</sup>*p<sup>h</sup>a:* <sup>1</sup>*to:ci*  
 I-DAT-TOP now aunt-GEN repay need-PFV  
 'Now I have to repay my aunt's [loan].'

The three main modals of ability are: <sup>2</sup>*k<sup>h</sup>am-pa* 'be (physically) able to' <sup>2</sup>*se:pa*, 'know how to', <sup>2</sup>*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) <sup>2</sup>*ai-se* {<sup>4</sup>*pwi* / <sup>4</sup>*pwi-pa-ri*} <sup>4</sup>*kham-la?*  
 you-ERG {carry / carry-NER-LOC} can-FUT  
 'Will you be able to carry it?'

Immediate future is expressed either by a V-V construction with the modal <sup>4</sup>*tam-pa* or, more commonly, by a purpose clause with the verb <sup>1</sup>*chjai-pa* 'be ready to' (61).

- (48) <sup>1</sup>*ŋa-ta* *sarpa-se* <sup>1</sup>*sat* <sup>4</sup>*tam-pa*  
 I-DAT snake-ERG kill be about to-IMPV  
 '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) \*<sup>3</sup>*a* <sup>1</sup>*ca* <sup>2</sup>*k<sup>h</sup>am-la?*  
 not eat can-FUT  
 '\*Can you refrain from eating?'
- (50) <sup>3</sup>*a* <sup>1</sup>*ca-na* <sup>2</sup>*ci* <sup>2</sup>*k<sup>h</sup>am-la?*  
 not eat-ing stay able-FUT  
 'Will you be able to stay without eating?'

### 7.4 Causatives

#### 7.4.1 V<sup>2</sup>*puŋ-pa*

One of two ways of expressing causation is a V-V construction with the verb <sup>2</sup>*puŋ-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 absolutive if it is intransitive (52). In this particular construction the embedded verb has to be voluntary or at least active.

- (51) <sup>1</sup>*t<sup>h</sup>en-kat'-ta-n* <sup>4</sup>*klu* <sup>2</sup>*puŋ* <sup>1</sup>*to:~nam*  
 they-PL-DAT-INT buy make must-INFR  
 'We have to get them to buy [it] [for us].'

- (52) *sikar- se* <sup>1</sup>*nak<sup>h</sup>i* <sup>4</sup>*jar* <sup>2</sup>*puŋ-pa*  
 hunter-ERG dog-ABS run let-IMPV  
 'The hunter lets his dog run/makes his dog run.'

With negation, <sup>4</sup>*jar* <sup>3</sup>*a-<sup>2</sup>puŋ*, 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* <sup>1</sup>*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) <sup>2</sup>*ani-kat'-se-nun* <sup>1</sup>*t<sup>h</sup>ai-na* <sup>1</sup>*la-u*  
 nun-PL-ERG-INT hear-that do-IMP  
 'Tell the nuns themselves.' (*lit.* 'Do so that they hear'.)

This construction can always replace the *v-<sup>2</sup>puŋ* construction, except for the meaning 'let', 'allow'.

Negation of causation (as opposed to denial of permission) uses the *v-na* <sup>1</sup>*la-pa* construction:

- (54) <sup>3</sup>*ca-ta-mi* <sup>3</sup>*mi-la* <sup>1</sup>*ŋon-ti* *icat* <sup>1</sup>*ni-na* <sup>4</sup>*ta* <sup>1</sup>*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) <sup>3</sup>*car* <sup>1</sup>*cjaŋpa* <sup>4</sup>*ca-∅* <sup>3</sup>*kruŋ-∅* <sup>2</sup>*k<sup>h</sup>ru-pa-ri* <sup>2</sup>*pit-cim-ro*  
 son youngest TOP-ABS guts-ABS wash-NER-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 <sup>3</sup>*pi-si* 'having said' used as a conjunction is employed.

- (56) <sup>3</sup>*keŋ* <sup>1</sup>*me-ri* <sup>4</sup>*po-kai* <sup>3</sup>*pi-si* <sup>4</sup>*caŋ-pa*  
 bread fire-LOC rise-OPT say-ing insert-IMPV  
 'I put the bread in the fire so that it may rise.'

- (57) \*<sup>1</sup>*me-ri* <sup>4</sup>*po-pa-ri* <sup>4</sup>*caŋ-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) <sup>1</sup>*cjan-se* <sup>1</sup>*ŋa-ta* {<sup>1</sup>*ca-la*/<sup>3</sup>*a-<sup>1</sup>ca-kai*} <sup>3</sup>*pi-si* <sup>4</sup>*jar-pa*  
 tiger-ERG I-DAT {eat-FUT/not-eat-OPT} say-ing escape-PFV  
 'I ran away for fear the tiger would eat me/so the tiger would not eat me.'

The participle <sup>3</sup>*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) <sup>4</sup>*pu* <sup>3</sup>*su-pa-ri* <sup>3</sup>*pi-si* <sup>1</sup>*apa* <sup>1</sup>*k<sup>h</sup>a-ma* . . .  
 field plant-NER-LOC say-ing father come-when  
 'When the father came thinking that he would plant the field. . .'

Using a conditional form <sup>4</sup>*pu* <sup>3</sup>*su-sai* <sup>3</sup>*pi-si* |field plant-COND say-ing| 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) <sup>2</sup>*nam* <sup>1</sup>*ka* <sup>1</sup>*k<sup>h</sup>a-pa-ri* <sup>4</sup>*mai-ci*  
 rain FOC come-NER-LOC search-PFV  
 'It's going to rain.' (*lit.* 'It's rain that is trying to come.')
- (61) <sup>1</sup>*ni-pa-ri* <sup>1</sup>*c<sup>h</sup>jai-ci*  
 go-NER-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'.

- (62) <sup>1</sup>*sjoŋ* <sup>3</sup>*kuŋ-ri* <sup>3</sup>*toɾ-ka-ma* <sup>1</sup>*sjoŋ-se* <sup>4</sup>*lap* <sup>4</sup>*por-ci*  
 river middle-LOC reach-DIR-when 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) <sup>3</sup>*a-<sup>1</sup>k<sup>h</sup>a-ma-m*, <sup>4</sup>*mai* <sup>1</sup>*ni-ci*  
 NEG-come-when-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) <sup>3</sup>*rit-si* <sup>4</sup>*pra-pa* <sup>1</sup>*paɾci* <sup>2</sup>*mren-na* <sup>1</sup>*ca-ci*  
 beg-ing walk-ner beggar satiated-so that eat-PFV  
 'He, a beggar, 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 <sup>4</sup>*to:na*, or the Nepali loan equivalent *samma*, following a negative form of the verb bearing the suffix *-te*, which indicates extent.

- (65) <sup>1</sup>*jaŋ* <sup>3</sup>*maima-m* <sup>3</sup>*a* <sup>1</sup>*si-te* *samma-m*  
 us women-TOP NEG die-as long as until-TOP  
<sup>3</sup>*rit-sai* <sup>1</sup>*ca-sai*  
 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 <sup>4</sup>*ŋjanchem* ‘from’.

- (66) <sup>1</sup>*oca* <sup>1</sup>*kan* <sup>3</sup>*a* <sup>2</sup>*min-ma* <sup>4</sup>*ŋjanchem*...  
 that rice NEG cook-during 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 anteriority is expressed by a participial clause in *v-si*.

- (67) <sup>1</sup>*nana* <sup>1</sup>*k<sup>h</sup>a-si* <sup>1</sup>*capasai* <sup>3</sup>*ŋjo-si* <sup>1</sup>*ca-ci*  
 sister come-having food cook-having eat-PFV  
 ‘Nana came, we cooked food, and we ate.’

This is also the most common way of expressing a cause.

- (68) <sup>2</sup>*nam* <sup>1</sup>*k<sup>h</sup>a-si* <sup>4</sup>*pu-ri* <sup>1</sup>*ni* <sup>3</sup>*a* <sup>2</sup>*mjaŋ-lai*  
 rain come-ing field-LOC go not 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) <sup>1</sup>*sja-si* <sup>1</sup>*sja-si* <sup>1</sup>*k<sup>h</sup>a-ci*  
 dance-ing dance-ing 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) <sup>2</sup>*nam* <sup>3</sup>*a* <sup>1</sup>*k<sup>h</sup>a-na* <sup>4</sup>*pu* <sup>3</sup>*su* <sup>3</sup>*a* <sup>2</sup>*mjaŋ-lai*  
 rain not come-ing field plant not 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) <sup>1</sup>*am'-se*      <sup>1</sup>*kwan-si*      <sup>1</sup>*pin-pa*  
 mother-ERG dress-having give- IMPFV  
 ‘His mother dresses [him].’
- (72) <sup>1</sup>*t<sup>h</sup>e-jen*    <sup>1</sup>*t<sup>h</sup>ai-si*      <sup>3</sup>*cin-ci*  
 they-too hear-having finish-PFV  
 ‘They too have already heard [the news].’
- (73) <sup>3</sup>*mi*      <sup>4</sup>*pra-si-n*      <sup>2</sup>*ci-pa*  
 people walk-ing-INT stay-IMPFV  
 ‘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 conditionals.

*7.8.1 Simple conditionals*

Simple conditionals are formed with the addition of the topic marker *-m(i)*.

- (74) <sup>2</sup>*nam*    <sup>1</sup>*k<sup>h</sup>a-sa-m*    <sup>1</sup>*ŋa*    <sup>3</sup>*a*    <sup>1</sup>*k<sup>h</sup>a*  
 rain come-if-TOP I 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, expressing greater doubt as to the certainty of the event.

- (75) <sup>1</sup>*iampu-ri*    <sup>1</sup>*mar*    <sup>1</sup>*ni-pa*    <sup>3</sup>*mi*    <sup>1</sup>*mu-sam*,  
 Kathm.-LOC down go-NER man be-if  
<sup>1</sup>*ŋa-m*    <sup>3</sup>*cakir*    <sup>1</sup>*ca-pa-r'*    <sup>1</sup>*ni-sai*  
 I-TOP salary eat-NER- LOC go-COND  
 ‘If someone was to go down to Kathmandu, I would go [with him] to earn a salary.’

A necessary condition ‘only if’ is expressed by a disjunction of the opposite conditions.

- (76) <sup>1</sup>*t<sup>h</sup>unta*    <sup>2</sup>*pit-ka-sam*    <sup>1</sup>*ŋa*    <sup>2</sup>*so-la*,  
 food send-DIR-if I live-FUT  
<sup>1</sup>*t<sup>h</sup>unta*    <sup>3</sup>*a-<sup>2</sup>pit-ka-sam*    <sup>1</sup>*ŋa*    <sup>1</sup>*si-la*  
 food NEG-send-DIR- if I 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, the meaning becomes counterfactual.

- (77) <sup>4</sup>*kjarca* <sup>2</sup>*som* <sup>1</sup>*la-sai* <sup>1</sup>*mu-pa*, <sup>2</sup>*arku-se* <sup>3</sup>*hin-sam*  
 hundred three make-COND be-Past other-ERG be-if  
 ‘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) <sup>1</sup>*sar* <sup>1</sup>*ju:-sa-ka* <sup>4</sup>*mrai-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: <sup>2</sup>*ser-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) <sup>2</sup>*chjai* <sup>1</sup>*ta-sa-i*, <sup>2</sup>*ai-se* <sup>2</sup>*phar* <sup>1</sup>*tor-pa* <sup>3</sup>*ahin*  
 debt occur-if-INT you-ERG pay must-IMPFV isn’t it?  
 ‘And if ever he incurred a debt, you would have to repay it, right?’

The emphatic conditional *-sa-i* is used to express a concessive condition ‘even if’. In this meaning it is often strengthened by a second intensifier *-nun*.

- (80) <sup>1</sup>*he* <sup>3</sup>*a-<sup>1</sup>k<sup>h</sup>a-sai(-nun)* <sup>1</sup>*ŋa* <sup>1</sup>*ni-la*  
 he not-come-if + INT(-INT) I 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) <sup>4</sup>*kjat* <sup>4</sup>*kate-n* <sup>1</sup>*la-sai* <sup>3</sup>*a* <sup>3</sup>*cin-pa*  
 work how much-even do-COND not finish-IMPFV  
 ‘However much work I may do, it is n [ever] finished.’

## 8 COMPARISON

Word order: standard + marker + predicate.

- (82) <sup>1</sup>*ŋa* ***banda*** <sup>2</sup>*tir-tiŋ-la* <sup>3</sup>*purij*  
 I than one-year- GEN younger sister  
 ‘She is one year younger than me. (she is not necessarily my sister).’

The morpheme *bhanda* is a Nepali loan meaning ‘saying’. A Tamang calque <sup>3</sup>*pi-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) <sup>4</sup>me <sup>1</sup>mama-i <sup>1</sup>ni-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 <sup>1</sup>tik 'what' after *ki*.

- (84) <sup>3</sup>plor-si-wa <sup>1</sup>t<sup>h</sup>an-pa ki <sup>1</sup>tik?  
 boil-having-Q put-IMPV 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) <sup>2</sup>c<sup>h</sup>at-wa <sup>2</sup>c<sup>h</sup>at-pa-ki <sup>2</sup>to-ca-kate-m?  
 fight-Q fight-IMPV-Q up-those-COLL-TOP  
 'Are they fighting for good, the people upstairs?'

Alternative questions are frequently used.

- (86) <sup>2</sup>na <sup>4</sup>ŋac<sup>h</sup>a <sup>3</sup>tor-ci <sup>2</sup>na <sup>1</sup>lic<sup>h</sup>a <sup>2</sup>ci-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) <sup>2</sup>osem <sup>2</sup>k<sup>h</sup>acjupa guru <sup>2</sup>sjap-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, indefinites with reduplicated 'wh-' words.

- (88) <sup>2</sup>k<sup>h</sup>aima <sup>2</sup>k<sup>h</sup>aima <sup>1</sup>klaŋ-pa  
 when when play-IMPV  
 '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 (96).

Negative polarity indefinites can also be used in questions.

## 10 INFORMATION STRUCTURE

Tamang possesses two topic markers, *-mil-m* for simple topic and <sup>4</sup>*ca* for contrastive topic, one focus marker *-ka<sup>l</sup>ka*; and a set of intensifiers *-i/-e/-ja* ‘even’, ‘also’ and *-ni/-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. Previously known elements are absent from the sentence.

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 (89).

- (89) <sup>1</sup>*oca-m*    <sup>1</sup>*ḡa-i-mi*    <sup>l</sup>*ʰa:*                    <sup>3</sup>*are, mai*  
 that-TOP    I-ERG-TOP    knowledge    is not    Mai  
 ‘That, as far as I am concerned, I don’t know, Mai.’

A contrastive topic is marked by the tonal particle <sup>4</sup>*ca*, possibly a loanword from Nepali *cahĩ*. (12, 55).

Clauses also can be topicalized with either topic marker. See (43).

- (90) <sup>1</sup>*si-pa*    <sup>4</sup>*ca*    <sup>1</sup>*ḡa-i*    *patta*                    <sup>3</sup>*are*  
 die-NER    TOP    I-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 (91) it prepares thus a dramatic effect expressed in the main clause. Compare with the unemotional reporting in (62).

- (91) <sup>3</sup>*kuḡ-te-ri*                    <sup>3</sup>*tor-ka-ma-m*                    <sup>1</sup>*apa*    <sup>1</sup>*sjog-se*    <sup>4</sup>*lap*    <sup>4</sup>*por-ci-ro*  
 middle-about-LOC    reach-DIR-when-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, 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.

### Argument focus

- (92) <sup>2</sup>*ai-la* <sup>4</sup>*mar-ka* <sup>4</sup>*niz-nun* <sup>2</sup>*cuŋ-o* <sup>3</sup>*pi-pa* <sup>1</sup>*ŋa-i-mi*  
 you-GEN gold-FOC two-INT sell-IMP say-IMPV I-ERG-TOP  
 'It is your gold [earrings] that I said to sell both of (I did).'

### Predicate and sentence focus

- (93) <sup>1</sup>*ŋa-ka*  
 I-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.

- (94) <sup>1</sup>*ŋa-i-ia* <sup>1</sup>*nipa-ka*, <sup>1</sup>*apa*  
 I-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 (95) or to its subject in case of sentence focus (60).

- (95) <sup>2</sup>*ai-se-mi* <sup>1</sup>*ŋa-la* <sup>1</sup>*t<sup>h</sup>o-ri* <sup>1</sup>*c<sup>h</sup>armo-ka* <sup>3</sup>*pa-ci*  
 you-ERG-TOP I-GEN top-LOC cowife-FOC bring-PFV  
 'You brought a CO-WIFE on me.'

### 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

- (96) <sup>1</sup>*ti:la* <sup>1</sup>*ŋa* <sup>1</sup>*oraŋ* <sup>3</sup>*ŋjot-la* <sup>3</sup>*pi-si-ka* <sup>4</sup>*paŋ*  
 yesterday I that way drunk-FUT say-ing-FOC strength  
<sup>3</sup>*are-pa-ri* <sup>2</sup>*k<sup>h</sup>ana-i* <sup>3</sup>*a-<sup>2</sup>waŋ-pa*  
 not be-NER-LOC where-also NEG-enter-IMPV  
 'Yesterday it was for fear of getting drunk, like that, in my weak condition, that I did not enter anywhere.'

As we have seen, it is almost systematically used on counterfactual conditionals (78).

## ADDITIONAL ABBREVIATIONS

Words in bold *italic* are loans from Nepali. Abbreviations specific to this section are:

Excl	exclamation
INT	intensifier
IRR	irrealis
<i>lit.</i>	literally
Ner	nominalizer (infinitive or participle)
NonPst	non past
RS	reported speech
∅	zero marker
*	ungrammatical structure
[ ]	word added in translation
( )	word deleted from translation

## REFERENCES

- Everitt, Fay (1973) 'Sentence patterns in Tamang', in Ronald Trail (ed.) *Patterns in Clause, Sentence, and Discourse in Selected Languages of India and Nepal*, SIL, University of Oklahoma, Vol. 1, 197–234.
- Georg, Stephan (1996) *Marphatan Thakali*, Munich: Lincom Europa.
- Glover, Warren (1974) *Sememic and Grammatical Structures in Gurung (Nepal)*, SIL, University of Oklahoma, 232 pp.
- Grierson, George A. (1908) *Linguistic Survey of India*, Vol. 3, Tibeto-Burman family, repr. Motilal Banarsidass 1967.
- Greenberg, Joseph (1966) *Language Universals*, The Hague, Paris: Mouton.
- Haiman, John (1978) 'Conditionals are topics', *Language* 54: 565–89.
- 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.
- Haspelmath, Martin and Ekkehard König (1998) 'Concessive conditionals in the languages of Europe', in J. Van der Auwera *et al.* (eds) *Adverbial Constructions in the Languages of Europe*, Berlin: Mouton de Gruyter, 563–640.
- Hoefer, Andras (1981, 1997) *Tamang Ritual Texts*, Beiträge zur Südasiensforschung 65, Wiesbaden, Franz Steiner Verlag, Vol. 1, 1981, Vol. 2, 1997.
- Keenan, Edward L. and Comrie, B. (1977) 'Noun phrase accessibility and universal grammar', *Linguistic Inquiry* 8: 63–99.
- Lambrecht, Knud (1994) *Information structure and Sentence form*, Cambridge Studies in Linguistics 71, Cambridge University Press.
- Mazaudon, M. (1973) *Phonologie du Tamang*, Paris: SELAF.
- Mazaudon, M. (1977) 'Tibeto-Burman Tonogenetics', *Linguistics of the Tibeto-Burman Area* 3:2, 123 pp., Berkeley, CA.
- Mazaudon, M. (1978a) 'Consonantal mutation and tonal split in the Tamang sub-family of Tibeto-Burman', *Kailash* 6/3, Kathmandu, 157–79.
- Mazaudon, M. (1978b) 'La formation des propositions relatives en tibétain', *Bulletin de la Société de Linguistique de Paris*, 73/1, Paris, 401–14.
- Mazaudon, M. (1988) 'Temps aspect et négation en tamang', in N. Tersis and A. Kihm (eds) *Temps et Aspects*, Paris Péeters/SELAF.
- Mazaudon, M. (2000– ) Tamang recorded texts, with synchronized transcription and translation. <http://lacito.archivage.vjf.cnrs.fr>



- Taylor, Doreen (1973) 'Clause patterns in Tamang', in Austin Hale and David Watters (eds) *Clause, Sentence, and Discourse Patterns in Selected Languages of Nepal*, SIL, University of Oklahoma, Vol. 2, 81–174.
- Shafer, Robert (1955) 'Classification of the Sino-Tibetan languages', *Word* 11:1, 94–111.
- Yoncan, Amrit (2054 B.S. [1997]) 'Tamang vyakaran' (Tamang grammar), Royal Nepal Academy, Kathmandu (in Nepali).

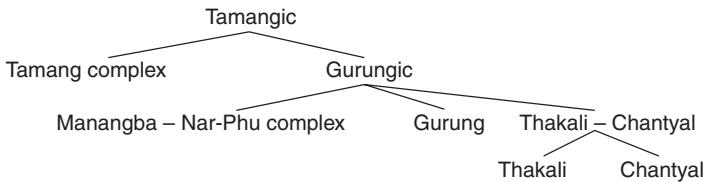
# CHANTYAL<sup>1</sup>

*Michael Noonan*

## 1 INTRODUCTION

The Chantyal language is spoken by approximately 2000 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, Ghyās Kharkā, Caura Khāni, Kuine Khāni, Thārā Khāni, Pātīle 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 and Tamang, the last two 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 as much as shared genetic background. At the moment, the most likely classification is as in Figure 19.1:



**FIGURE 19.1** RELATIONS AMONG THE TAMANGIC LANGUAGES

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

---

1 Work on Chantyal has been supported by the National Science Foundation, grant No. DBC-9121114.

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 [ɐ], 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].

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/=[tʃi]) 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.<sup>2</sup> Nasals, approximants, fricatives, and glides also contrast murmured and non-murmured phonemes. In the transcription used here, <h> indicates aspiration, <fi> murmur.

Unaspirated stop	p	t	c	k
Aspirated stop	ph	th	ch	kh
Voiced stop	b	d	j	g

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 text reflects the intuitions of Chantyls, though this ‘intuition’ has doubtless been influenced by Nepali orthographic tradition. The transcriptional system employed here, with {fi} representing murmur, is compatible with either view.

Murmured stop	bfi	dfi	jfi	gfi
Murmured stop with voiceless onset	pfi	tfi	cfi	kfi
Murmured stop with voiceless aspirated onset		thfi		khfi
Voiced nasal stop	m	n	ny	ŋ
Murmured nasal stop	mfi	nfi	nfiy	
Voiced lateral approximant		l		
Murmured lateral approximant		lfi		
Voiced tap approximant			r	
Murmured tap approximant			rfi	
Voiceless fricatives			s	
Murmured fricatives with voiceless onsets			sfi	fi
Glides			y	w
Murmured Glides			fiy	fiw

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 intervocalically 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 (/bfi/, /dfi/, /jfi/, and /gfi/) occur word internally in words of Nepali origin and /bfi/ occurs finally in words of Nepali and English origin: /bfi/ 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 /t/ and /d/ 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	GK	<i>bāwso</i>	hoe
KL	<i>bakhra</i>	goat	LK	<i>arko</i>	next
KK	<i>bəkselu</i>	larva	KK	<i>tuktuk</i>	hacking
KN	<i>thutno</i>	snout	NK	<i>əncəl</i>	province
NG	<i>kaŋŋyo</i>	comb [N]	GN	<i>bfiyawni</i>	ghost
NL	<i>aŋla</i>	joint	LN	<i>jurni</i>	joint

Prior to the massive influx of Nepali and, recently, English words, allowable word-final segments 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 /bfi/) 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/, /cʰi/] 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’, *fiə-* ‘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 3 dogs’

The plural suffix is often found with the first person plural pronoun, e.g. *nfi-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*  
 I-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:

<i>na</i> I	<i>nə-ye</i> ‘my’ [I-GENITIVE]
<i>tɕim</i> house	<i>tɕim-əŋ</i> ‘homeward’ [house-LOCATIVE]

The case clitics are listed below:

Absolute	-∅
Ergative/Instrumental	-sə
Ablative	-gəm-sə
Elicative	-nɕi-ri-gəm-sə
Abessive-proximal	-nasa-w
Dative	-ra
Allative	-nas
Locative	-ri
	-əfi [with a few common nouns, deictics, case clitics]
	-chə [with the noun mɕiun ‘night’]
	-ra [in a few special expressions]
Inclusive	-muwa
Adessive	-muwa-ri
Inessive	-nɕi-ri
Genitive	-ye, -i
Allative/Comitative	-siŋ, -səŋ, -səŋə, -səŋgə (N, except? -siŋ)
Comitative	-ru
Circumlative	-mar
Subessive	-pɕiri
Sublative	-pɕiri-ŋ-sə
Supersive	-phyara-ŋ-sə
Temporal	-ma (N)
Comparative/Temporal	-bɕinda (N)
Comparative/Temporal	-dɕin, -dɕikɕin (N)
Comparative/Temporal	-rasə
Essive/Comparative	-thōy

(The forms followed by ‘(N)’ are borrowings from Nepali.) As shown in the chart, 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:

*tɕim-nɕi-ri-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 absolutive), but does not demonstrate syntactic ergativity; that is, grammatical processes that refer to subjects refer to ergative and absolutive marked subjects equally and no grammatical processes also refer to absolutives regardless of

grammatical role. Ergative subjects may be agents, but they may also represent other semantic roles, for example experiencers, as in the following:

*khi-sə*            *uttər*    *thāya-m*  
 s/he-ERGATIVE    answer    know-NON.PAST  
 'He knows the answer.'

Animate direct objects are typically marked as datives, i.e. Chantyal evidences 'anti-dative' marking. The use of the genitive (as opposed to simple juxtaposition of an absolutive 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:

*yek gilās cfiā*  
 one glass tea  
 'one glass of tea'

The locative case is used with both static (locational) and dynamic (allative) senses: *fiyunda-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:

*nə-ye*            *naku khi-ye*    *naku- bfiənda*    *thyo mu*  
 I-GENITIVE dog    s/he-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:

*cəŋ*    *nə*    *ram-sə*            *nə*    *bənnu*    *chij-ji*  
 then    TOPIC    Ram-ERGATIVE    FOCUS    gun    hold-PERF  
 '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>Dasē-ye ligəm</i>	after Dasain
<i>Ram-ye wən</i>	before Ram	<i>Ram-ye wənwən</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 s/he 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:

1s	<i>na</i>
1d	<i>nagi</i> [ <i>&lt; na kfi</i> ‘I you’; inclusive sense only]
1p	<i>nfi, nfi-ma</i> [we-PLURAL]
2s	<i>kfi</i> [INFORMAL], <i>nūwə</i> [FORMAL]
2p	<i>nūwə</i>



- 3s *khi* [humans only], *cə* ['that': humans and non-humans]  
 3p *thā wə*, *thā wə-ma* [they-PLURAL], *cə-ma* [that-plural]

The second and third person plural forms derive from forms with overt plural marking, though as *thūwə-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 the *khi* form 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, *kfiya-ra* and *khyā-ra*: the regular forms *kfi-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 *fiə-* '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>fiə-sər</i>	that way	<i>fiə-jəgər</i>	that much
<i>fiə-tuŋ</i>	up there	<i>yi-ddyo</i>	this big
<i>wu-tuŋ</i>	up there yonder	<i>fiə-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>fiə-cu</i>	this (near speaker, further from speaker of two alternatives)
<i>yi-cə</i>	that (distant from speaker, nearer of two alternatives)
<i>fiə-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>kfiəre</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-ē, -ē</i>
Perfective	<i>-ji, -i</i>
Perfective Interrogative	<i>-la</i>
Imperfective	<i>-ma, -wā, -ō</i>
Negative Anterior Imperfective	<i>-s-are, -s-ere</i>
Past Anterior	<i>-sēē</i>
Emphatic, mirative	<i>-wa</i>

#### *Mood:*

Hypothetical	<i>-∅</i>
Suppositional	<i>-ndə, -nnə, -n</i>
Suppositional	<i>-t</i>
Hypothetical	<i>-rə, -r</i>
Desiderative	<i>-to</i>
Imperative	<i>-o</i>
Polite Imperative	<i>-ne</i>
Hortative	<i>-ye</i>
Optative	<i>-kəy, -gəy, -ge</i>

#### *Subordinate:*

Anterior	<i>-si-</i>
Nominalizer	<i>-wa</i>
Infinitive	<i>-nu (N?)<sup>3</sup></i>
Cotemporal	<i>-khiri, -khir, -khi (N)</i>

3 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-gəy</i>	<i>a-thū</i>
drink-INFINITIVE	drink-INFINITIVE	do-PROGRESSIVE	NEGATIVE-drink
‘She was about to drink, but didn’t.’			

Progressive converb	<i>-kəy, -gəy, -ge</i>
Sequential converb	<i>-si-rə, -si-r, -si</i>
Resultative	<i>-nə</i>
Conditional	<i>-la</i>
Negative conditional	<i>-kfiəre, -gfiəre</i>
Remote Conditional	<i>-la-i</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, *fya-* ‘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>fīn</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>-gəy</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 19.2.

	Past	Present	Immediate future	Future
Perfective	<i>-ji</i>		<i>-ji</i>	
Imperfective	<i>-ma</i>		<i>-m</i>	

FIGURE 19.2 THE TENSE-ASPECT SYSTEM

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 periphrastically.

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.

*na-sə*      *na-ra*      *jfi-i*  
 I-ERGATIVE I-DATIVE bite-PERFECTIVE  
 'I bit myself.'

There is, however, a special reciprocal form:

*Ram rə Pirəm-sə*      *lfi-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>
<i>pətk-</i> explode, burst	<i>pətk-i-</i> explode, burst (intr)	<i>pətk-ə</i> explode, burst (tr); fire a gun	
<i>dfiəl-</i> fall	<i>dfiəl-i-</i> fall over	<i>dfiəl-ə</i> knock over	
<i>phutk-</i> loose	<i>phutk-i-</i> escape, untangle, untie (intr)	<i>phutk-ə</i> let loose, untie (tr)	
<i>tər-</i> cross		<i>tər-i-</i> cross	<i>tər-ə</i> take across
<i>buj-</i> understand		<i>buj-i-</i> understand	<i>buj-ə</i> explain
<i>səmji-</i> remember		<i>səmji-i-</i> remember	<i>səmji-ə</i> 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, *bfiər-i-* 'fill' takes substances as objects, whereas *bfiər-ə* 'fill' takes containers as objects. (Note that *bfiər-i-* also has an intransitive sense: 'be full'.)

The case marking associated with verbs having these derivational affixes is entirely predicable and expected: transitive subjects are marked with the ergative; intransitive subjects are absolutive. There is also a periphrastic causative, illustrated below:

*ram-sə*      *pirəm-ra*      *nfiaka*      *rfiə-nə*      *la-i*  
 Ram-ERGATIVE piram-DATIVE chicken steal-RESULTATIVE do-PERF  
 '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 *fīn* 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>fīn</i>	<i>nfiē</i>	<i>fīayn</i>		
<i>mu</i>	<i>mu-ē</i>	<i>fīare</i>	<i>mu-ma</i>	<i>fīare-ma</i>

*mu* is regular except for the suppletive root [*rfiē*] 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.) *fīn* is defective in not having past forms: the past forms of *mu* are used instead. *fīayn* is analysable simply as the negative suffix *a-* plus *fīn*; *nfiē* is simply a contraction of *fīn* and the non-past interrogative suffix *-ē*.

In general, *fīn* 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 *fīn*. 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 5000 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>rfiəŋ rfiəŋ</i>	sound of something dry and crunchy
<i>cwaktə cwaktə</i>	sound of cutting up meat
<i>ʃfiəm ʃfiəm</i>	burning, tingling sensation (as produced by alcohol, nettle, Szechuan pepper)
<i>kfiwak kfiwak</i>	sound of choking, gagging
<i>burrək burruk</i>	action of jumping

<i>phwak phwak</i>	coming off in chunks
<i>phwakkə</i>	coming off in chunks
<i>cakal cikal</i>	eating sloppily, noisily
<i>phappə phuppə</i>	scattered carelessly, easily picked up and/or stolen
<i>khəllə bəllə</i>	action of quarrelling
<i>ləpəpə</i>	appearance of scattered things
<i>khwassə</i>	inserting quickly and neatly
<i>lwarrə</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>gfadi 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 s/he 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>nfi</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 above, 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 below. An important set of derivational morphemes are the converbal 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 Chantyls 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>huri əŋŋula</i>	old woman digit	thumb
<i>dɦopini mimi</i>	washerwoman bird	brown dipper
<i>ɦiar khor</i>	bone shelter	skeleton
<i>bɦiũy 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>nɦie tɦiara</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>pɦiəkən</i>	<i>pɦiara</i> flour + <i>kan</i> ‘food’	porridge
<i>kɦele</i>	<i>kɦə</i> mouth + <i>le</i> tongue	tongue
<i>nagi</i>	<i>na</i> I + <i>kɦi</i> you	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 absolutive 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 lɦi-</i>	gun hit	shoot
<i>ɦutti 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 five thousand 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:

*bfiuluŋ-sə*      *ca-i*      *bura-ra*  
 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 REFERENTS  
 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.

*na-ra*      *bfiuluŋ-sə*      *ca*  
 I-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:

*cəŋ*      *gāw-ri*      *na-sə*      *bənnu* *bfiar-ri*      *kar-ji*  
 that-LOCATIVE    village-LOCATIVE    I-ERGATIVE    gun    rack- LOCATIVE    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:

*cəŋ nə*,    *pirəm nə*,    *lata*    *mənchi*,    *dula*    *kho-i*  
 then TOPIC    Piram TOPIC    stupid person    hole    dig-perf  
 ‘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:

*khi-sə*            *na-ra*      *cini-m*  
 s/he-ERGATIVE    I-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 absolutive 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

*khi-sə*            *nfiaka*    *tha-i*  
 s/he-ERGATIVE    chicken    cut-PERFECTIVE  
 'He sacrificed/killed the chicken.'

with:

*khi-sə*            *nfiaka-ra*      *tha-i*  
 s/he-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*  
 I-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

*nfi bidyarthi fin*  
 we student be + NON.PAST  
 'We're students.'

PREDICATE ADJECTIVE

*na khusi mu*  
 I happy be + NON.PAST  
 'I'm happy.'

OBLIQUE CASE-MARKED NP FUNCTIONING AS PREDICATE

*cə kitab tebəl-phyaraŋ 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.<sup>4</sup> 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, *kfiəmməy* 'never', though this form requires the negative prefix on the verb:

*khi-sə kan kfiəmməy a-ca-m*  
 s/he-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

*nfi-sə su-i-ra pəni a-pin*  
 we-ERGATIVE who-even-DATIVE also NEGATIVE-give-PERFECTIVE  
 'We didn't give it to anyone.'

and

*na-sə təy a-yā*  
 I-ERGATIVE anything NEGATIVE-find-PERFECTIVE  
 'I didn't find anything/I found nothing.'

show. (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'.)

<sup>4</sup> It may also be prefixed onto adjectives, though generally not when they are used predicatively.

### 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 *-ē*, which is added to the non-past suffix *-m*.

*kfi-sə*            *bfialu*    *sar-la*  
 you-ERGATIVE    bear        kill-PERFECTIVE.INTERROGATIVE  
 ‘Did you kill a bear?’

*kfi-sə*            *bfialu*    *sar-m-ē*  
 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.

*kfi-sə*            *bfialu*    *sar-m*            *nfi*  
 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.

*kfi-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*.

*na-ra*            *syaw*    *ca-wa*            *mən*    *kha-m*  
 I-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 *bfi-* ‘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 *bfi-* is thus taking on many of the characteristics of a complementizer, though it is not yet fully grammaticalized in this role.

*khi tisuj kadmandu-ri fya-i bfi-wa*  
 s/he last.year Kathmandu-LOCATIVE go-PERFECTIVE say-NOMINALIZATION  
*na-sə səmjfi-i*  
 I-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.

*khyo-ma-ru duli-si-wa kyata*  
 friend-PLURAL-COMITATIVE wander-ANTERIOR-NOMINALIZER boy  
*tfiem-əŋ fya-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.

*səŋlal-ma məə tara-wa-ri fya-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.

*pirəm kha-lanə a-kha-m*  
 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:

*Ram-sə gāw-ri fya-si-rə tfiim-nfiari*  
 Ram-ERGATIVE village-LOCATIVE go-ANTERIOR-SEQUENTIAL house-INESSIVE  
*wō-si-rə nfiaka-ye sya ca-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:

*na-ra*      *pir-si*                      *pin-o*  
 I-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:

*mənchi-sə*              *thar-ra*                      *bənnu*   *lfi-si-rə*  
 person-ERGATIVE   mountain.goat-DATIVE   gun   hit-ANTERIOR-SEQUENTIA  
*sar-ji*                      *chala*   *tar-si-rə*,                      *ənnarabfiuri*  
 kill-PERFECTIVE   skin   extract-ANTERIOR-SEQUENTIAL   viscera  
*thuti-si-rə*,                      *tfiem-əŋ*                      *kfiə-si- ə*,  
 pull.out-ANTERIOR-SEQUENTIAL   house-LOCATIVE   bring-ANTERIOR-SEQUENTIAL  
*dewri-ri*                      *pəkəysi-rə*,                      *ca-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'.

### REFERENCES

- Very little has yet been published on Chantyal. Articles that focus exclusively or primarily on Chantyal include the following, all of which are written by the author, Michael Noonan:
- 'The fall and rise and fall of the Chantyal language', *Southwest Journal of Linguistics* 15/1–2: 121–36 (1996), (Also: *Milwaukee Studies on Language* 9, 1995.) (Describes the social and historical setting for the decline of the Chantyal language.)
- '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 (1997). (Describes the range of functions of nominalizations in Chantyal and other languages.)
- 'Converbal constructions in Chantyal', in Yogendra Yadava, P. (ed.) *Topics in Nepalese Linguistics*, Kathmandu: Royal Nepal Academy (1999). (Describes the syntax and semantics of the sequential and progressive converbs.)
- Chantyal Dictionary and Texts*, with Ram Prasad Bhulanja, Jag Man Chhantyal, and William Pagliuca. Berlin: Mouton de Gruyter (1999).
- 'Chantyal texts', in George van Driem (ed.) *Himalayan Linguistics*, Berlin: Mouton de Gruyter (forthcoming).

'The "double demonstratives" of Chantyal', in *Linguistics of the Tibeto-Burman Area* 24.2: 173–88 (2001). (Describes the spatial and discourse uses of the combinations of prefixal and root demonstratives.)

'Posture verbs in two languages of Nepal', in John Newman (ed.) *The Linguistics of Sitting, Standing, and Lying*, Amsterdam/Philadelphia: John Benjamins. With Karen Grunow-Härsta (2002). (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.)

An extensive grammar is in preparation.

# NAR-PHU<sup>1</sup>

*Michael Noonan*

## 1 INTRODUCTION

The Nar-Phu language is spoken by the 800 people of the villages of Nar and Phu, located in the Valley of the Nar Khola in the Manang District of Nepal. The territory they inhabit is very high (the lowest point in their territory is approximately 3500 metres in altitude) and their main occupations are yak herding and small-scale farming.

By the standards of other languages in Nepal, the influence of Nepali on the Nar-Phu language has been relatively small. 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 twelve 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, Manangba, 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 Manangba 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 I 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 Manangba designation) was made in consultation with native speakers.

---

<sup>1</sup> Work on Nar-Phu has been supported by the National Science Foundation, grant no. SBR-9600717.

## 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 can be displayed as follows:

i	u
e	o
ɛ	œ
æ	a

The status of [œ] as an independent vowel is not clear. In what follows, it is always written /wɛ/, a transcription that reflects, in part, its pronunciation – it is typically pronounced [wœ] or [yœ] – and the fact that the sequence [wɛ] otherwise does not occur.

With /ɛ/ a preceding glide [y] is often heard in open, stressed syllables. Word final /ɛ/, especially in the suffix /-pɛ/, sounds similar to /a/ in very slow speech. /e/ and /ɛ/ contrast in open syllables, but are neutralized to [ɛ] in closed syllables: <ɛ> 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 in /ɛ/ in closed syllables derives historically from \*/e/ whereas /ɛ/ in open unstressed syllables derives historically from \*/a/.)

/æ/ is a very low front-central vowel.

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 [hrāre] ‘millet’ and [khɛro] ‘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/, /mfɪnse/ ‘all night’, has been recorded.)

Long vowels come about as a result of phonological or morphological processes: the loss (always [ʔ] restorable in slow pronunciation) of syllable final /k/, /p/, or /t/ in Tibetan borrowings, and the coalescence of a final front vowel in nouns with the genitive: *lamɛ-ɛ* ‘lama’s’, *ŋɛ-ɛ* ‘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 /y/ and /w/ (recall the special status of /wɛ/, discussed above):

ae  
ay  
ow  
aw  
uy  
wey

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) /ay/, wherein the first element is a low back vowel and the glide a trajector toward the high front position: this diphthong seems to occur only in Nepali borrowings and onomatopoeic words. /wɛy/ represents [œθ], [œy], [ɣœθ], or [ɣœy]. The conditions under which [ɣ] appears in /wɛy/ and /wɛ/ 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 below), though there is with liquids: [l] and [r] contrast with [l̥] and [r̥]. Murmur is distinctive, but is best considered part of the tone system (see below). Five points of articulation are attested: bilabial, lamino-dental, retroflex, alveolo-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.

<i>Phonetically</i>					<i>In transcription</i>					
p	t	t̥		k	ʔ	p	t	t̥	k	ʔ
p <sup>h</sup>	t <sup>h</sup>	t̥ <sup>h</sup>		k <sup>h</sup>		ph	th	t̥h	k	
	ts		tʃ				c		č	
	ts <sup>h</sup>		tʃ <sup>h</sup>				ch		čh	
	s		ʃ				s		š	
m	n		ɲ	ŋ		m	n		ny	ŋ
	l						l			
	l̥						hl			
	r						r			
	r̥						hr			
w		j	ɥ			w		y		[see below]

The rhotic consonants are sometimes produced as taps [ɾ], sometimes as glides [ɹ]. [l̥] has a not-too-common variant [ɬ], an alveolar lateral fricative.

[ɥ] replaces [w] in some speech styles for some speakers. [ɥ] or [ɣ] also occurs in the diphthong /ae/, e.g. in 'load' [tʰɣɛ̃] /tʰiâe/.

[ʔ] 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 *kfiêpɛ* eighth, a [ʔ] has been recorded – though not consistently: [kfiêʔpɛ]. This does not alternate with /k/, /p/, or /t/ (though the last is what occurs in the Written Tibetan), but 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: [bâsi] 'bamboo'. In words with the murmured tones 3 and 4, the murmured voice extends, typically, to the initial consonant, so that /kfiê/ 'work' is often [gfiê]. In this respect, Nar-Phu differs from its relative Chantyal, in which initial voiceless stops and affricates remain voiceless in murmured syllables.

## 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 character <fi>. Two of the 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.

<i>Tone number</i>	<i>Pitch contour</i>	<i>Example</i>
1	53	<i>nâŋ</i> reciprocal obligation
2	44	<i>naŋ</i> full
3	12	<i>nfiãŋ</i> planted in rows
4	21 or 31	<i>nfiãŋ</i> in

Tones 1 and 4 are falling; tones 3 and 4 are murmured. Tone 2 is distinguished by its clear, high quality. There is a certain amount of variation from speaker to speaker in the pitch contours associated with these tones, but the distinctive elements – plain vs murmured voice and falling vs level or rising pitch – appear to remain constant from speaker to speaker.

Murmured tones may occur with any sort of initial consonant except the aspirated stops and affricates and the voiceless liquids /h/ and /hr/.

The pitch contours are ordinarily distributed over entire words; that is, they distribute over a root and any affixes it may take. 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

All consonants may occur in initial position. There are initial clusters with /r/ and /l/; the following are attested:

pr	tr	kr
phr	thr	khr
mr		
pl		kl
phl		
ml		

Initial clusters with the glides /w/ and /y/ are also attested:

pw	tw	cw	čw	
	thw	chw	čhw	khw
mw	nw			ŋw
			šw	
	rw			
py				ky
phy				khy
my				ŋy

In word final position, nasals, liquids, and unaspirated stops may occur. 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/.

Medial clusters of consonants occur in compounds: in such cases, any combination of final and initial consonant (including consonant clusters) is possible. Words synchronically analysable 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 intervocalically (including internally within compounds), and, intervocalically, 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, it becomes variously /a/ or /o/ when it is in the first element (*nfiê* ‘ear’, *nfiaci tigiŋ* ‘area in front of the ear’, *nfiokli* ‘earwax’); when it is the final element, it usually becomes /ɛ/ (*tæ* ‘horse’, *phôrte* ‘gelding’, *môrte* ‘mare’), though there are apparent exceptions. In inflection, /æ/ becomes /a/ (*ŋâ* ‘I’, *ŋâ-se* ‘I (ERGATIVE)’). 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’, *phôrte* ‘gelding’, *môrte* ‘mare’ (Written Tibetan has *rtā* ‘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* ‘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:

*nôkyu-cuke* dogs  
*nôkyu mflaŋ-cuke* black dogs

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:

*nôkyu-cuke-re* to the dogs

The case clitics are listed below:

Absolutive	-∅
Ergative, instrumental, ablative	-se
Dative, locative	-re
Genitive	-(y)e, -i
Independent genitive	-nê
Comitative	-ten

The absolutive 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 below) 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.

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:

*čfiæ 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 absolutive, 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 follow:

<i>Ungendered noun</i>	<i>Male</i>	<i>Castrated male</i>	<i>Female</i>
<i>rfiô</i> bond friend	<i>rfiopε</i>		<i>rfiomε</i>
<i>nôkyu</i> dog	<i>nôkyupho</i>		<i>nôkyumo</i>
<i>rêgoat</i>	<i>râpho</i>	<i>râpε</i>	<i>râmo</i>
<i>tæ</i> horse	<i>photyεn</i>	<i>phôrtε</i>	<i>môrtε</i>

### 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 honoured 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: *šfiε* (food items), *čha* (body parts), *šfiεp* (clothing). The resulting honorific may still display considerable phonological idiosyncrasy. Some examples follow:

	<i>Ordinary</i>	<i>Honorific</i>
beard	<i>kyôw</i>	<i>šalcham</i>
body	<i>čfiû</i>	<i>kûsuk</i>
boot	<i>kfiyô</i>	<i>šfiεpkfiyô</i>
butter	<i>mfiar</i>	<i>šfiεmfiar</i>
chang (Tibetan beer)	<i>pfiow</i>	<i>čhwεčhâη</i>
eye	<i>mi</i>	<i>čên</i>
face	<i>ηotoη</i>	<i>šfiâl</i>

Honorific verbs are illustrated below:

	<i>Ordinary</i>	<i>Honorific</i>	<i>Humble</i>
be sick	<i>na-</i>	<i>nyûŋ-</i>	
buy, take	<i>kîn-</i>	<i>sfi-</i>	
get up, stand	<i>re-</i>	<i>šfiâŋ-</i>	
give	<i>pîŋ-</i>	<i>nâŋ-, nâŋ kê-</i>	<i>phûl-</i>
smell (tr)	<i>naŋ-</i>	<i>naŋ-</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.

#### *Cardinal numbers*

1	<i>kfirî</i>	11	<i>cûkfirî</i>		100	<i>kfiyâ</i>	
2	<i>ŋfi</i>	12	<i>cûŋfi</i>	20	<i>ŋfiyûšu</i>	200	<i>ŋfi kfiyâ</i>
3	som	13	<i>cwêysôm</i>	30	<i>sômčû</i>	300	<i>som kfiyâ</i>
4	<i>pfilî</i>	14	<i>cûlte, cûlpfilî</i>	40	<i>pfilîčû</i>	400	<i>šipkye</i>
5	<i>ŋfiâ</i>	15	<i>cûŋfiâ</i>	50	<i>ŋfiâčû</i>	500	<i>ŋfiâpkye</i>
6	<i>tfiûk</i>	16	<i>cûtfiûk</i>	60	<i>tfiûkčû</i>	600	<i>tfiûpkye</i> (sic)
7	<i>ŋi</i>	17	<i>cwêyŋi</i>	70	<i>ŋičû</i>	700	<i>tfiûŋkye</i>
8	<i>pfire</i>	18	<i>côpfire</i>	80	<i>pfirečû</i>	800	<i>pfirekye</i>
9	<i>ku</i>	19	<i>cûrku</i>	90	<i>kûčû</i>	900	<i>kupkye</i>
10	<i>ču</i>					1000	<i>toŋta</i>
100,000	<i>lak, pfium</i>						
10,000,000	<i>tfiûŋčur, tfiûŋkur</i>						

Internal to the number phrase itself, numbers follow the order of larger-smaller when additive (forty + seven = 47), smaller-larger when multiplicative (four + hundred = 400), except after *toŋta* '1000', 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:

47	<i>pfilîčû ŋi</i>
439	<i>šipkye-se sômčû ku</i>
1996	<i>toŋta kfirî-se kupkye-se kûčû tfiûk</i>
67,735	<i>toŋta tfiûkčû ŋi-se tfiûŋkye-se sômčû ŋfiâ</i>
347 dogs	<i>nôkyu som kfiyâ-se pfilîčû ŋi</i>

The ordinal numbers are entirely borrowed from Central Tibetan. They evidence numerous complexities and only the first three decades are given below:

#### *Ordinal numbers*

	<i>Tenth</i>	<i>čupe</i>	<i>Twentieth</i>	<i>ŋfiyûšupe</i>
1st	<i>tfiâŋpe</i>	<i>čupcikpe</i>		<i>ŋfiyûšu cackikpe</i>
2nd	<i>ŋfiŋpe</i>	<i>čuŋikpe</i>		<i>ŋfiyûšu caŋfiŋkpe</i>
3rd	<i>sûmppe</i>	<i>čuksumpe</i>		<i>ŋfiyûšu caksûmppe</i>
4th	<i>šfiŋpe</i>	<i>čupšfiŋpe</i>		<i>ŋfiyûšu capšfiŋpe</i>
5th	<i>ŋfiâpe</i>	<i>ceŋfiâkpe</i>		<i>ŋfiyûšu ceŋfiâkpe</i>
6th	<i>tfiûkpe</i>	<i>čutfiûkpe</i>		<i>ŋfiyûšu ce tfiûkpe</i>

7th	<i>tfûnpɛ</i>	<i>čuptɛnpɛ</i>	<i>ŋfɪyûšu captfûnpɛ</i>
8th	<i>kfiêpɛ</i>	<i>cepkfîêpɛ</i>	<i>ŋfɪyûšu capkfîêpɛ</i>
9th	<i>kfiûpɛ</i>	<i>curkupɛ</i>	<i>ŋfɪyûšu capkupɛ</i>

Nar-Phu does not have a system of classifiers.

### 3.5 Pronouns, demonstratives and articles

The paradigms for the personal pronouns are displayed below:

	1st SG	1st EMPH	1st PL INCL	1st PL INCL	1st PL EXCL
ABS	<i>ŋɛ̃</i>	<i>ŋfiɛ̃</i>	<i>ŋfĩ</i>	<i>ŋfĩ-cuke</i>	<i>ŋfiyâŋ</i>
GEN	<i>ŋɛ̃-e</i>	<i>ŋfiɛ̃-ɛ</i>	<i>ŋfĩ-i</i>	<i>ŋfĩ-cuke-e</i>	<i>ŋfiyâŋ-e</i>
IGEN	<i>ŋɛ̃-e-nê</i>	<i>ŋfiâ-nê</i>	<i>ŋfĩ-i-nê</i>	<i>ŋfĩ-cuke-nê</i>	<i>ŋfiyâŋ-nê</i>
E/I/A	<i>ŋâ-se</i>	<i>ŋfiâ-se</i>	<i>ŋfĩ-se</i>	<i>ŋfĩ-cuke-se</i>	<i>ŋfiyâŋ-se</i>
D/L	<i>ŋâ-re</i>	<i>ŋfiâ-re</i>	<i>ŋfĩ-re</i>	<i>ŋfĩ-cuke-re</i>	<i>ŋfiyâŋ-re</i>
COM	<i>ŋâ-tɛn</i>	<i>ŋfiâ-tɛn</i>	<i>ŋfĩ-tɛn</i>	<i>ŋfĩ-cuke-tɛn</i>	<i>ŋfiyâŋ-tɛn</i>

	1st PL EXCL	1st PL PARTN
ABS	<i>ŋfiyâŋ-cuke</i>	<i>kuŋ</i>
GEN	<i>ŋfiyâŋ-cuke-e</i>	<i>kuŋ-e</i>
IGEN	<i>ŋfiyâŋ-cuke-nê</i>	<i>kuŋ-nê</i>
E/I/A	<i>ŋfiyâŋ-cuke-se</i>	<i>kuŋ-se</i>
D/L	<i>ŋfiyâŋ-cuke-re</i>	<i>kuŋ-re</i>
COM	<i>ŋfiyâŋ-cuke-tɛn</i>	<i>kuŋ-tɛn</i>

	2nd SG FAM	2nd PL FAM	2nd SG	2nd PL
ABS	<i>nuŋaŋ</i>	<i>nuŋaŋ-cuke</i>	<i>kfiyaŋ</i>	<i>kfiin-cuke</i>
GEN	<i>nuŋaŋ-e</i>	<i>nuŋaŋ-cuke-e</i>	<i>kfĩ-i/ye</i>	<i>kfiin-cuke-e</i>
IGEN	<i>nuŋaŋ-nê</i>	<i>nuŋaŋ-cuke-nê</i>	<i>kfiyaŋ-nê, kfĩ-i-nê</i>	<i>kfiin-cuke-nê</i>
E/I/A	<i>nuŋaŋ-se</i>	<i>nuŋaŋ-cuke-se</i>	<i>kfiyaŋ-se</i>	<i>kfiin-cuke-se</i>
D/L	<i>nuŋaŋ-re</i>	<i>nuŋaŋ-cuke-re</i>	<i>kfiyaŋ-re</i>	<i>n-cuke-re</i>
COM	<i>nuŋaŋ-tɛn</i>	<i>nuŋaŋ-cuke-tɛn</i>	<i>kfiyaŋ-tɛn</i>	<i>kfiin-cuke-tɛn</i>

	3rd SG FAM	3rd SG	3rd RESP	2nd PL FAM	3rd PL	3rd GEN
ABS	<i>nuŋ</i>	<i>cû</i>	<i>fîota</i>	<i>nuŋyɛ</i>	<i>cû-cuke</i>	<i>cɛ</i>
GEN	<i>nwɛ-ye</i>	<i>cû-i</i>	<i>fîota-e</i>	<i>nuŋyɛ-ye</i>	<i>cû-cuke-e</i>	<i>cɛ-ye</i>
IGEN	<i>nuŋ-nê</i>	<i>cû-i-nê, cû-nê</i>	<i>fîota-nê</i>	<i>nuŋyɛ-nê</i>	<i>cû-cuke-nê</i>	<i>cɛ-nê</i>
E/I/A	<i>nuŋ-se</i>	<i>cû-se</i>	<i>fîota-se</i>	<i>nuŋyɛ-se</i>	<i>cû-cuke-se</i>	<i>cɛ-se</i>
D/L	<i>nuŋ-re</i>	<i>cû-re</i>	<i>fîota-re</i>	<i>nuŋyɛ-re</i>	<i>cû-cuke-re</i>	<i>cɛ-re</i>
COM	<i>nuŋ-tɛn</i>	<i>cû-tɛn</i>	<i>fîota-tɛn</i>	<i>nuŋyɛ-tɛn</i>	<i>cû-cuke-tɛn</i>	<i>cɛ-tɛn</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 \**fota-cuke*. The other third person pronouns have plurals formed with the regular plural morpheme. The generic third person pronoun translates ‘one’: it can be used to refer to both human and non-human referents.

The interrogative pronouns are listed below:

what	<i>tê</i>
who	<i>sû</i>
where	<i>khana</i>
which	<i>khoncu</i>
how much/many	<i>kfiate</i>

These forms are used with the interrogative forms of the verb:

<i>čfæ</i>	<i>kfiate</i>	<i>mu-pɛ</i>
tea	how.much	be-INTERROGATIVE
‘How much tea is this?’		

Interrogative pronouns are not fronted within their clause, being placed ordinarily in the slot appropriate to their grammatical role:

<i>lakpɛ-sɛ</i>	<i>sû-re</i>	<i>mraŋ-pɛ</i>
Lakpa-ERGATIVE	who-DAT/LOC	see-INTERROGATIVE
‘Who did Lakpa see?’		

Determiners constitute a simple set from an areal perspective: proximal *cû* contrasts with distal *têta* (fast speech) and *tfieta* (slow, careful speech). These forms are used attributively and pronominally. The generic third person pronoun *cɛ* also has a definite article-like function and lacks any sense of spatial deixis; it follows the nouns it determines: *pfiâlpe cɛ* ‘the toad’, *thɛ-pɛ cɛ-se* ‘the big one’.

There is an indefinite singular article *-ri*, derived historically from the numeral *kfiŋi* ‘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:

<i>ŋê</i>	<i>mfi-ri</i>	<i>mraŋ-cin</i>
I	person-INDEFINITE	see-PAST
‘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. 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 *-pɛ*. Periphrastic verbal constructions are common: the auxiliary verb is always a copular verb. Verbs may be nominalized (becoming verbal nouns) 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 (but see the next paragraph). Verbs are not morphologically marked for transitivity.

Central to the organization of the verbal paradigm is the direct/indirect 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.

The main tense-aspect forms are given below:

	<i>Direct</i>		<i>Indirect</i>	
	<i>Positive</i>	<i>Negative</i>	<i>Positive</i>	<i>Negative</i>
<i>Past</i>	<i>v-čin</i>	<i>a-v-i</i>	<i>v-či(n) mû</i>	<i>a-v-i mû</i>
<i>Aorist</i>	<i>v-se</i>	<i>a-v-i</i>	<i>v-se mû</i>	<i>a-v-i mû</i>
<i>Present</i>	<i>v(mû)</i>	<i>a-v mû</i>	<i>v(-pɛ/te) mûmu</i>	<i>a-v(-pɛ/te) mûmu,</i> <i>v(-pɛ/te) fîarmu</i>
<i>Durative</i>	<i>v-pɛ/te mû</i>	<i>a-v-i</i>	<i>v-pɛ/te mûmu</i>	<i>a-v-i mûmu</i>

(In the table above, ‘v’ is used for any verbal root.) In general, the indirect forms are identical to the direct forms with the addition of a copula (see below 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. The ‘aorist’ refers to the immediate past and the future; it also has a perfective sense. The aorist suffix is likely derived from the converbal suffix *-se*. The durative is a present tense form and is used with progressive senses and in cases where the ongoing nature of habitual actions or states is emphasized.

Main clause modal affixes include:

<i>Imperative: singular subject</i>	<i>-(t)aw</i>
<i>Imperative: plural subject</i>	<i>-(t)ow</i>
<i>Hortative</i>	<i>-šo</i>
<i>Potential</i>	<i>-kî</i>

Subordinate clause affixes include the following:

<i>Sequential converb</i>	<i>-se</i>
<i>Conditional</i>	<i>-reme, -teme</i>
<i>Indeterminate nominalizer</i>	<i>-pɛ</i>
<i>Determinate nominalizer</i>	<i>-te</i>
<i>Potential nominalizer</i>	<i>-ne</i>
<i>Past relative</i>	<i>-pi</i>
<i>Non-past relative</i>	<i>-pɛ</i>



The distinction between the indeterminate and determinate nominalizers is one of aspect: the determinate nominalizer has a more completive sense, whereas the indeterminate nominalizer has a more progressive sense. This distinction involves a number of complexities, but can be illustrated simply with the following pair of sentences:

*ŋâ-se lakpɛ-re fîkɛ pŋiri-pe mraŋ-čîn*  
 I-ERGATIVE Lakpa-DAT/LOC letter write-NOMINALIZER see-PAST  
 'I saw Lakpa writing the letter.'

*ŋâ-se lakpɛ-re fîkɛ pŋiri-te mraŋ-čîn*  
 I-ERGATIVE Lakpa-DAT/LOC letter write-NOMINALIZER see-PAST  
 'I saw Lakpa write the letter.'

The first sentence, with the indeterminate nominalizer, makes no claim that the writing of the letter was ever completed. The second sentence, which involves the determinate nominalizer, implies that the writing was completed.

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. A few examples follow:

*pŋira ni-w*  
 walk go-IMPERATIVE  
 'Go for a walk!'

*kyâŋ lâ-w*  
 reach do-IMPERATIVE  
 'Reach for it!'

*târ kyû pi ni-w*  
 be.in.order run go.fast go-IMPERATIVE  
 'Run in order!'

*nôkyu cɛ-se tŋiŋ nâpraŋ čhâŋ cɛ pi tê læ-čîn*  
 dog DEF-ERGATIVE bee nest DEF go.fast fall do-PAST  
 'The dog unwittingly knocked down the beehive.'

### 3.7 Copular verbs

There are two copular verbs, one of which, *nfiɛ*, is defective in lacking past forms and in lacking a contrast between direct and indirect forms. The two copular verbs are listed below:

	<i>mû</i>	<i>nfiɛ</i>
	<i>Direct</i>	<i>Indirect</i>
Declarative	<i>mû</i>	<i>mûmu</i>
Negative	<i>fiare</i>	<i>fiârmu</i>
Interrogative	<i>mu-pe</i>	<i>fiin-pe</i> (< <i>fiimpe</i> > <i>mɛ</i> )
Negative interrogative	<i>fiar-pe</i>	<i>fi-in-pe</i>
Declarative past	<i>mû</i>	<i>mû-i</i>
Negative past	<i>fiare</i>	<i>fiare</i>

In terms of use, *mû* is the unmarked form and is the form used in verbal periphrasis. It's basic use is in clauses asserting identity, but it has encroached on the territory of *nfiê*, whose basic use is the expression of location. In the past tense, *mû* is used exclusively.

### 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 *-pɛ* 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 onomatopoeically), 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:

<i>cfiarcar</i>	falling down
<i>tfioktfiok</i>	rapping, pounding
<i>khusti</i>	in a joking manner

An example is provided showing the use of the expressive vocabulary in a sentence:

<i>thô</i>	<i>cfiarcar</i>	<i>tɛ-pɛ</i>	<i>mû</i>
saliva	falling.down	become-NOMINALIZER	be-PRESENT
'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 two most common are *kâ*, which has a contrastive emphatic sense, 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 above. There is also the suffix *-ten*, which means 'one from': *ču-pruŋ-ten* 'one from Nar', *pfiâlpe-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:

*mfiḷaŋkhuŋ* throat [*< mfiḷaŋ* black + *khuŋ* hole]  
*mêphra* ash [*< mê* fire + *pîrâ* flour]  
*kfirêmsô* molar [*< kfirê*m cheek + *sô* tooth]  
*hlačhwê-* pray, worship [*<hla* god + *čhwê-* 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. *pâtlu mfi* ‘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:

*cû nâwar čfi-pi*                      *cû nôkyu aŋye thepɛ*  
 this cat    bite-RELATIVE.PAST    this dog    very big  
 ‘this very big dog that bit this cat’

*lamɛ thepɛ-cuke-e*                      *čfi-cuke*  
 lama    big-PLURAL-GENITIVE    child-PLURAL  
 ‘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) comes first, followed by the rest of the constituents arranged according to their rank on the empathy hierarchy, which is reproduced below:

Speech act 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 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; still, most transitive subjects are so marked. The lack of ergative marking is not conditioned in any straightforward manner by tense-aspect or volitionality of the subject: the matter requires further study. Direct objects are, by default, in the absolutive case, but direct objects high on the empathy hierarchy are marked as datives. This sort of dative marking (*aka* ‘anti-dative’ shift) is obligatory for personal pronouns, virtually so for proper nouns, but is very unusual for NPs representing inanimates. Two examples follow:

*ŋâ-se lakpɛ-re mraŋ-čîn*  
I-ERGATIVE Lakpa-DAT/LOC see-past  
‘I saw Lakpa.’

*ŋâ-se fîikε mraŋ-čîn*  
I-ERGATIVE letter see-past  
‘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:

*ŋê-e mîn lakpɛ*  
I-GENITIVE name Lakpa  
‘My name is Lakpa.’

Ordinarily, though, a copular verb is present and comes last in the clause:

PREDICATE NOMINAL  
*kfiyaŋ tfiâpε mûmu*  
you novice.monk be + PRESENT  
‘You’re a novice monk [I’ve been told].’

PREDICATE ADJECTIVE  
*ŋê khaŋ mû*  
I cold be + PRESENT  
‘I’m cold.’

OBLIQUE CASE-MARKED NP FUNCTIONING AS PREDICATE  
*ŋê tfiôsor pfiâlpε-re mû*  
I now Kathmandu-DAT/LOC be + present  
‘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*:

*lama a-kha-kî*  
*lama* NEGATIVE-come-POTENTIAL  
‘The lama may not come.’

No negative indefinite pronouns or adverbs [i.e. words like ‘never’, ‘nothing’] have yet been found.

### 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 *-pɛ* to the verb. This suffix usually replaces finite indicative verbal morphology,

*tile lakpɛ ʧfæ̃ thuŋ-pɛ*  
 yesterday Lakpa tea drink-INTERROGATIVE  
 'Did Lakpa drink tea yesterday?'

but sometimes *ni-se-pɛ* is used in addition to these suffixes:

*ni-se-pɛ*  
 go-AORIST-INTERROGATIVE  
 'Did he go?'

A construction involving positive and negative interrogative verbs is commonly encountered:

*cû kyû thuŋ-pɛ a-thuŋ-pɛ*  
 this water drink-INTERROGATIVE NEGATIVE-drink-INTERROGATIVE  
 'Did he drink water?'

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. The interrogative suffix is found in such sentences:

*ʧfæ̃ sû-se thuŋ-pɛ*  
 tea who-ERGATIVE drink-INTERROGATIVE  
 'Who drank the tea?'

### 5.6 Complement [nominal] clauses

The two nominalizer suffixes, the 'indeterminate' *-pɛ* and the 'determinate' *-te*, were described briefly above. Both suffixes can be used to nominalize clauses so that they can fill nominal slots, either subject

*hleke hlô-pɛ cû-i kŋeyɑŋ kâ*  
 book read-NOMINALIZER this-GENITIVE work EMPHATIC  
 'Reading books is his profession.'

or direct object:

*phô sê-te mraŋ-čîn*  
 game kill-NOMINALIZER see-PAST  
 'I saw the game shot.'

In addition, there is also the suffix *-ne*, referred to as the 'potential' nominalizer. Verbals marked with this suffix also fill nominal slots and have an interpretation that marks them as coding unfulfilled, potential events or states:

*ŋâ ŋhûn-ne tên mû*  
 I sleep-NOMINALIZER desire be + PRESENT  
 'I want to sleep.'

All three nominalizers are used extensively in verbal periphrasis.

There are no finite subordinate clauses except as direct quote complements of *pfii-* ‘say’.

### 5.7 Adjectival (relative) clauses

Adjectival clauses, clauses that modify nouns, are clearly related to nominalizations. There are two relative suffixes, the present *-pɛ* and the past *-pi*: the present is no doubt identical to the indeterminate nominalizer *-pɛ*, the past is probably *-pɛ* 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 nominalizers 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:

*ɣâ-se*      *šî-pi*                      *tfulthun*      *mraŋ-čîn*  
I-ERGATIVE    die-PAST.RELATIVE    snake.body    see-PAST  
‘I saw a dead snake.’

*mîn*      *te-ne*                              *a-tâ-pɛ*                                      *pfuluŋ*  
name    call-NOMINALIZER    NEGATIVE-become-    PRESENT.RELATIVE    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 are used for a number of adverbial functions, including the expression of purpose. The nominalizers *-pɛ* and *-te* can occur with case suffixes.

*ɣâ-se*      *iŋliš*      *hlô-te-re,*                                      *ni-čîn*  
I-ERGATIVE    English    teach-NOMINALIZER-DAT/LOC    go-PAST  
‘In order to teach English, I went.’

Converbial clauses are used to code a variety of senses, including condition

*kfi-i*              *laki-re*              *nfiô*      *câ-reme*              *tano*      *mû*  
you-GENITIVE    sake-DAT/LOC    garlic    eat-CONDITIONAL    good    be + PRESENT  
‘If you eat garlic, it’s good for you.’

and temporal sequence:

*mfiyê*      *câ-pɛ-re*                                      *ni-se,*              *mfiyê*      *a-yâŋ-se,*  
cow    eat-NOMINALIZER-DAT/LOC    go-CONVERB    cow    NEGATIVE-find-CONVERB  
*fiyul-re*              *yê*      *ni-čîn*  
village-DAT/LOC    return    go-PAST

‘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 verb, are used.

**REFERENCES**

There is very little published on the Nar-Phu language. Martine Mazaudon has published an article on aspects of the historical phonology of the Nar-Phu dialects: she is the only other linguist (to my knowledge) who has investigated Nar-Phu. Besides her article, there is only one other publication to provide information about the Nar-Phu language, a book by the geographer and ethno-botanist Perdita Pohle, in which some botanical vocabulary from the Nar and Phu dialects is given, together with botanical descriptions of the plants and discussions of their use in local medicine. Tones are not marked. The transcription is only approximate, and some of the alleged Nar and Phu terms are, in fact, Tibetan. Still the book is a useful introduction to the regional ethno-botany and contains valuable data. References to these two works are given below:

Mazaudon, Martine (1997) 'An outline of the historical phonology of the dialects of Nar-Phu (Nepal)'. *Linguistics of the Tibeto-Burman Area* 19.1: 103–14.

Pohle, Perdita (1990) *Useful Plants of Manang District: A Contribution to the Ethnobotany of the Nepal-Himalaya*, Stuttgart: Franz Steiner Verlag.

A grammatical description of Nar-Phu is in preparation by the author.

**PART 7**

---

**NEWAR DIALECTS**

---





# DOLAKHĀ NEWĀR

*Carol Genetti*

## 1 INTRODUCTION

Dolakhā Newār is a Tibeto-Burman language spoken in the village of Dolakhā, approximately 145 kilometres to the east of Kathmandu, Nepal.<sup>1</sup> The language has been referred to as Dolakhāli, Dolakhā (Newār), and Dolakhāe. The latter form contains the genitive morpheme =e, thus means ‘of Dolakhā’, and is the form used by the speakers to refer to themselves and their language. I will use the terms Dolakhāe and Dolakhā Newār interchangeably.

Dolakhāe is clearly in the Newār family and is closely related to the dialects of Newār spoken in the Kathmandu Valley and other locations throughout Nepal. I prefer to refer to Dolakhāe as a ‘language’ as opposed to a ‘dialect’ of Newār, as it is mutually unintelligible with the Newār 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 Dolakhāe 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 Dolakhā Newārs have migrated to other parts of Nepal and beyond to seek employment. The number is likely to be between 5000 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 Newār community. Consequently, many of the current children of Dolakhā Newārs are not learning to speak the language. On the other hand, the people of Dolakhā in the outlying areas have a strong civic commitment to the village, and it is hoped that this interest in maintaining both the village and the community will play a positive role in furthering the life of the language.

It is not known if there are any true dialectal divisions within the Dolakhāe language. The village of Dolakhā is relatively small and homogenous, 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 have 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

---

<sup>1</sup> Funds for this study were provided by the National Science Foundation, grant BNS-8811773.

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 Dolakhāe. 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 larger works, see Genetti (1994) and Shrestha (1996).

A list of abbreviations specific to this chapter is given in an appendix.

## 2 PHONOLOGY

The consonant phonemes of Dolakhāe are given in (1):

(1)	p, ph, b	t, th, d	ʈ, ʈh, ɖ	c, ch, j	k, kh, g
	m	n			ŋ
		s			h
		l	r		
	w			y	

Of particular interest are the presence of a retroflex series of consonants (not attested in other Newār dialects), and the absence of breathy voice. The former is attested by the minimal pairs *tar-* ‘hear’ ‘hear’ and *tār-* ‘fix’, *thon-* ‘hide’ and *ʈhon-* ‘wake’, and *don-* ‘finish’, *ḍon-* ‘stand’. The dental/retroflex distinction was probably a feature of Proto-Newār, and has been lost in other dialects. Note that the symbols /c, ch, j/ represent alveo-palatal affricates.

There are a few words in Dolakhāe with breathy-voiced articulation. Most represent old borrowings from Sanskrit or Nepali. Only a few words appear to be of Newār origin. One of these is *dhū* ‘tiger’, which both my consultants produced with breathy voice when repeating it to me slowly. In rapid pronunciation no breathiness was heard. Thus breathiness is marginal in Dolakhāe, at best.

One other note on the consonant inventory is that two of these ‘phonemes’ are actually in complementary distribution. The retroflex stop /ɖ/ 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 very 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 Dolakhāe 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 Dolakhāe are given in (2)

(2)	i, ī	u, ū
	e, ē	o, ō
	ā, ā̃	a, ā̃

The vowel /ā̃/ is low and central, while /a/ ranges from the low, back [ɑ] to [ʌ] and [ə]. The vowel /o/ has a wide range of phonetic variants, some of them predictable, including [wa], [wo], [o] and [ʔo].

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.ɛu] ‘come-3sFUT’. When a suffix beginning with /e/ follows another vowel, disyllabic sequences result: *tā-eu* [tā.ɛu] ‘hear-3sFUT’. If the preceding vowel is high, then a glide is inserted: *si-eu* [si.yɛu] ‘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)’.

Dolakhāe has stress but no tone. Stress falls on the initial syllable of a root, with a secondary stress on third syllables in longer words.

### 3 MORPHOLOGY

#### 3.1 Nouns and pronouns

Dolakhāe 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 <lines 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 Dolakhāe. 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 Dolakhāe 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 Dolakhāe 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 Dolakhāe 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

The basic Dolakhāe numerals are given in (5):

(5)	<i>thi</i>	one	<i>ji(m)</i>	ten
	<i>nis</i>	two	<i>nī(n)</i>	twenty
	<i>sō</i>	three	<i>swi</i>	thirty
	<i>pe</i>	four	<i>pi</i>	forty
	<i>ḡā</i>	five	<i>ḡe/ḡai</i>	fifty
	<i>khū</i>	six	<i>khwi</i>	sixty
	<i>nas</i>	seven	<i>nai</i>	seventy
	<i>cyā</i>	eight	<i>cyai</i>	eighty
	<i>gu</i>	nine	<i>gwi</i>	ninety
	<i>sar</i>	hundred	<i>dol</i>	thousand

In forming the numbers from 11 through 19, *jim* is used to indicate ten: *jimthi* ‘eleven’, *jimnis* ‘twelve’, etc. The form *nin* is only found in *ninthi* ‘twenty-one’, otherwise *ni* (sometimes with the vowel lengthened) is found, e.g. *ninis* ‘twenty-two’, *niipe* ‘twenty-four’. The numeral ‘forty’ *pi* similarly shows lengthening: *pīsō* ‘forty-three’. When constructing multiples of one hundred, the smaller number comes first: *nassar* ‘700’.

Dolakhāe numerals never occur as bare stems, but are always followed by a bound numeral classifier. The numeral-classifier combination generally precedes the modified noun. There are twenty-two classifiers that I know of. Classifiers occur with numerals only; 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. The majority of members of a class are usually amenable to semantic classification. A full description of the classifier system is beyond the scope of this chapter. Of particular interest and centrality, however, are *mā* for animates, *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. 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 Dolakhāe. 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 *gāḡ-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. *tuḡ-gu guābā* 'ripe guava', 'guava that has ripened'.

Although the majority of property concepts in Dolakhāe 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 a few others. When used attributively, these adjectives precede the noun, although *lyās-mā* and *lyāsi-misā* may be used as nominal compounds, 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 Discourse markers of topic and emphasis

There are two morphological clitics which adhere to the noun phrase, and which are used to indicate different degrees of prominence of the referent to which the NP refers. One of these is  $=(u)ḡ$ , which has a number of specific interpretations in particular lexical environments. Compare *nīs-mā* 'two' and *nīs-mā=ḡ* 'both', *lita* 'next' and *lita=ḡ* 'again', *rabi=n* 'Rabi' (as subject of transitive verb), *rabi=n=uḡ* 'Rabi also (did something)', *ānṯi* 'like that' with *ānṯi=ḡ* 'just like that' <33>.

The other clitic is  $=(u)ri$ , which usually 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.

## 4 VERBS

### 4.1 Copulas

There are two copular verbs in Dolakhā Newār, *dam(u)* and *khyayḡ*. These verbs are independent of the four inflectional verb classes. The existential copula *dam* has a number of both finite and non-finite inflectional forms. Especially common are *da-u*, the past stative form, and *ma-da*, the negative. The equational *khyayḡ* may also be pronounced *khyaiḡ*; 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 subject of the location is animate, the verb *con-* 'stay' is used <17>. The copula 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’, *twāṅsona sona khyaj* ‘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 copulas, all Dolakhāe 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* ‘return’) 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. *ṭār yet-* ‘to cross’ (from Nepali *ṭār-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 21.1. Only one affirmative finite verb form 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 affixes to transitive verbs.

The Dolakhāe finite verb inflects for four tenses, the past imperfective, simple past, present and future. Tense markers follow the verb directly. The past imperfective 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 Dolakhāe 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-* respect-

TABLE 21.1 AFFIRMATIVE PARADIGM OF NAR-‘EAT’

	<i>Past imperfective</i>	<i>Past</i>	<i>Present</i>	<i>Future</i>
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>

ively. The verbs in these moods and the imperative also inflect for number, each category demonstrating slightly different inflectional patterns.

There are four non-finite forms of the Dolakhāe 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 may be roughly labelled a perfect aspect, and *bir-* 'give', which adds the presence of a benefactive object.

### 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.

Dolakhā Newār 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>, although the opposite order is also frequently attested; presumably these orderings are exploited for discourse purposes, although no thorough study has yet been conducted on this issue. Adverbial modifiers always precede the verb they modify. Manner adverbs are often derived from participial verbs.

#### 5.2 Grammatical relations

There is clear evidence for subject and object as grammatical categories in Dolakhāe. 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 Dolakhāe 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 form of the verb 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, cf. 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 Dolakhāe grammar is the manner in which clauses are woven together to form long sentences of intricate structure. Sentence breaks in narrative discourse may be identified by the presence of a non-embedded finite verb. The verb may be followed by particles, and occasionally other postposed elements. The final clause is usually produced with falling intonation. Preceding clauses are generally produced with rising or high level intonation. The following sentence, from Genetti and Slater (to appear) is illustrative of sentence complexity:

- (6) *dwā-ku=ri*      *iri*                      *oho: |*    *ji*    *yā*    *githi*    *yer-i*,  
 senior-NR1 = TOP    daughter.in.law    EXCL    1s    EMPH    how    do-1FUT  
*lās*    *cār-a-gi*    *haŋ-an*,    *ithi:*    *yeŋ-an*,    *mikhā*    *tī-en*  
 shy    feel-PR-1s    say-PART    like.this    do-PART    eyes    close-PART  
*on-ŋasin*;    *dhirtarāstra*    *janm-ai*    *jur-a*.  
 go-when    Dhirtarāstra    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 (see, e.g. Johanson 1995). 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/aspect information. 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. sentences 1 and 4>. Note that the first of the chaining verbs, *haŋ-an* in (6) above, 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 to appear), thus entire sentences may be embedded within sentences in this construction.

In (6) the verb in bold, *on-ŋasin*, marks an adverbial subordinate clause. These are distinct from the participial construction because the set of adverbial subordinators contain specific information about the temporal or logical relationship which holds between the subordinate clause and the main clause. More adverbial clauses may be found in <15, 19, 20, 38, 55, etc.>

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-ɲasin*, the final verb in the series and the only intransitive one. The final and finite clause of the sentence has a new subject, Dhirtarastra.

In addition to chaining, adverbial subordination, and the embedding of quoted material as syntactic object, Dolakhāe also 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’, *sagkā* ‘suspicion’). The complement clause verb takes the infinitive plus the emphatic = *uri* if the clause is irrealis, 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/irrealis 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 is *-gu*, *-ku*, or *-u*, while NR2 is *-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 Dolakhāe the two structures are syntactically distinct (see Genetti 1994: 154–6). The two suffixes NR1 and NR2 are in a paradigmatic relationship, and the distribution of these suffixes *vis-à-vis* each other is extremely complicated and varies by speaker. 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 judgements 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 Dolakhāe is a bit unusual. The object of comparison appears in the absolutive (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 khyaj* ‘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 (A or S) of the clause. It is most commonly used with a possessive sense, e.g. *āpen thau chē oŋ-an...* ‘they went to their own house’. If appropriate, the reflexive pronoun may be repeated to give a distributive sense, 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 Dolakhā, in January 1989. A small portion has been ellipped, 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, taken from Genetti and Slater (forthcoming), are marked by the following conventions:

- (7) . 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 underlined 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 disguise 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 dalkyāṅ-an;*  
 sari wear-PART
- 6 *gāñjal=na uṅ-an;*  
 eyeliner=INST draw-PART
- 7 *sā thu-en;*  
 hair braid-PART
- 8 *bakundo u-en;*  
 ball dance-PART
- 9 *harararararaṅ,*  
 ONOM

- 10 *bisnu bāgabān 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, *hararararaṅ*, 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.*  
Kailāś direction=LOC
- 14 *baikuntha=lān.*  
Baikhuntha=ABL

‘Taking the form of a young maiden, he went to Śiva’s garden. In Kailāś. From Baikuntha.’

## 6.3 Sentence 3

- 15 *āle āku on-ḡasin wā;*  
then there go-when TOP
- 16 *nis-mā siba pārbati,*  
two-CL Śiva Pārbati
- 17 *atāli=ku coḡ-an;*  
balcony=LOC stay-PART
- 18 *musukka ḡil-en coḡ-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 caḡṭa mikhā tar-ḡasin wā;*  
then EXP eye put-when TOP
- 20 *gulpa=ḡ lāhā=ku bokunda metha-e-thē,*  
when=EMPH hand=LOC ball play-NR2-as.if,
- 21 *gulpa=ḡ bokunda metha-e-thē,*  
when=EMPH ball play-NR2-as.if
- 22 *sārpuli halyāḡ-an;*  
ribbons swing-PART

- 23 *yārliḡ dalkyāḡ-an;*  
earring wear-PART
- 24 *mikhā tarkyāḡ-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-ḡasin;*  
EXP=INST come-when
- 28 *caḡṭa sibaji=n khoḡ-a ju-en con-a.*  
EXP Śiva=ERG see-NR2 be-PART stay-3sPAST

‘Then when he put his eyes on her, sometimes as if she were playing with a ball in her hand, sometimes 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 doḡ-an li;*  
ground=LOC come-INF finish-PART after
- 33 *āle āmu bisnu bāḡabān=ta lyāsmisā=ta=ḡ;*  
then that Viśnu god=DAT young.woman=DAT=EMPH
- 34 *samj-ai ju-en,*  
remember-BVS be-PART
- 35 *tin lok;*  
three realms
- 36 *caudha bhuban;*  
fourteen worlds
- 37 *dulyeḡ-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 *ḡil-en*;  
 smile-PART
- 40 *phisikka ḡil-en thi-khana*,  
 EXP      smile-PART    one-side
- 41 *phisikka ḡil-en thi-khana*,  
 EXP      smile-PART    one-side
- 42 *phisikka ḡil-en on-ḡa on-ḡa on-ḡa*;  
 EXP      smile-PART    go-when    go-when    go-when
- 43 *āme jati sibaji = e jati*,  
 3sGEN    possessions    Śiva = GEN    possessions
- 44 *swāl hā-u*,  
 EXP      fall-NR1
- 45 *bijan*,  
 seed
- 46 *lū*,  
 gold
- 47 *oho*,  
 silver
- 48 *liuri*,  
 bronze
- 49 *pitāl*,  
 copper
- 50 *ānāgu ānāgu*,  
 like.that    like.that
- 51 =*pen ju-en dhātu-pen khas-ai ju-en ḡḡ-gu ju-en con-a* |  
 =PL    be-PART    metal = PL    fallBVS    be-PART    go-NR1    be-PART    stay-3SPAST

‘Then after he fell to the ground, he 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 *āle tin lok*,  
 then    three    realms
- 53 *caudha bhuban ḡuler-i doḡ-ani*,  
 fourteen    world    lead-INF    finish-PART

- 54 *guli* *u* *sibaji = ta* *ḍulet-ki* *ḍukha* *bi*  
 how.much this Śiva = DAT lead-1SPAST trouble give(INF)
- 55 *'chin* *māyā = ku* *jit-ai* *ju-i* *phar-a-gu* *rā'* *hat-ḡasin.*  
 2SERG love = LOC win-BVS be-INF able-PR-2h Q say-when
- 56 *'ma-pha-gi'*  
 NEG-able-1s
- 57 *haḡ-an* *ha = ku.*  
 say-PART say-2hPAST
- 58 *'phar-a-gi'* *haḡ-an* *ha-ku.*  
 able-PR-1s say-PART say-2hPAST
- 59 *āu* *ākhirī = kuota* *haḡ-ane,*  
 now end = LOC this say-PART
- 60 *phyāḡḡa* *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 *lyāsīmīsā* *haḡ-a* *jon-ḡasin;*  
 young.woman say-NR2 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 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-eni;*  
then embarrassment feel-PART

69 *jīn thaeta pahilā tuṅ ha-ku-ī le.*  
1sERG 2hdAT before TOP say-PI-1s PRCL

70 *'thamu māyā = ku jīt-ai ju-i phar-a-gu rā|*  
2HON love = LOC win-BVS happen-INF able-PR-2h Q

71 *ma-phar-a-gu rā' haṅ-an 'phar-a-gi' ha-kupe.*  
NEG-able-PH-2h Q say-PART able-PR-1s say-PI-2h

72 *gwā le hat-cu|*  
where PRCL 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 *āle ānṭhi har-i-ho;*  
then like.that say-INF-when

74 *phisikka pārbati = ṅ ṅil-en,*  
EXP Parbati = TOP smile-PART

75 *jī ulistule bāla-ku pārbati thae sāt-ku coṅ-an con-a-gi.*  
1s like.this beautiful-NR1 Parbati 2hGEN side = LOC stay-PART stay-PR-1s

76 *thamu = ri ānṭhi = ṅ ṅil-a-gu rā|*  
2h = TOP like.that = EMPH smile-PR-2s Q

'Then when he spoke like that, smiling Parbati smiled and said "I, Parbati, who is beautiful like this, sit at your side. You smile like that (at some other girl)?"

(Small portion elipted here)

## 6.12 Sentence 12

77 *sibaji summake lās mā-cā-en –*  
Śiba silent embarrassment NEG-feel-PART

78 *lās cā-en con-a.*  
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
NR1~2	nominalizer/relativizer 1 or 2
ONOM	onomatopoeia
PH	past habitual/remote

## REFERENCES

- Benedict, Paul K. (1972) *Sino-Tibetan: A Conspectus*, Cambridge: Cambridge University Press.
- Genetti, C. (1994) *A Descriptive and Historical Account of the Dolakha Newāri Dialect* (Monumenta Serindica 25), Tokyo: Tokyo University of Foreign Studies.
- (1987) 'Object relations and dative case in Dolakha Newāri', *Studies in Language* 21.1: 37–68.
- Genetti, C. and Slater, K. (forthcoming) 'Syntax/prosody interactions in a Dolakhāe Nepāl Bhāsā rendition of the Mahābhārata', in G. van Driem (ed.) *Himalayan Linguistics*, The Hague: Mouton de Gruyter.
- Johanson, Lars (1995) 'On Turkic converb clauses', in M. Haspelmath and E. König (eds) *Converbs in Cross-Linguistic Perspective*, Berlin: Mouton de Gruyter.
- Shrestha, R.L. (1996) *Dolakhā Nepāl Bhāsāyā Varnātmak Adhyana*, (in Kathmandu Newār: 'A descriptive study of Dolakhā Nepāl Bhāsā') unpublished doctoral dissertation, Tribhuvan University.

# KATHMANDU NEWAR (NEPĀL BHĀŚĀ)

*David Hargreaves*

## 1 INTRODUCTION

Nepāl Bhāśā (Kathmandu Newar), spoken by approximately 690,000 speakers, is the language of the Newars of the Kathmandu Valley of Nepal (Kansakar 1997). 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 many villages surrounding these urban areas. With two notable exceptions (Dolakhāe 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 the conjunct/disjunct verb conjugations (see Section 3.2.1), while a second grouping, consisting of Badikhel Pahari and the Dolakhāe (Genetti 1994; this volume), is characterized by the presence of subject/verb agreement morphology. While the conjunct/disjunct inflectional morphology among the Kathmandu, Bhaktapur, and Patan related dialects is clearly cognate, the person/number agreement morphology in the Badikhel Pahari dialect is not cognate with the person/number agreement system in Dolakhāe, and neither appears cognate with the conjunct/disjunct morphology. Thus, the historical processes leading to different verb conjugation patterns remain opaque. Moreover, the subgrouping of the Newar languages within the Tibeto-Burman family remains problematic. Based on his analysis of verb morphology in Dolakhāe, 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. 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; Kansakar 1997). In modern Nepal, increasing Newar/Nepali bilingualism continues to influence the lexicon and grammar, and the economic and social status of Nepali, as a national language, exerts a powerful force towards language shift away from Nepāl Bhāśā (Shrestha 1990).

## 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 *khwaũ* ‘cold’. Syllables without onsets occur in a small number of monosyllabic lexical items, *i*: ‘time’, *uĩ*: ‘madwoman’, and occur commonly as the first syllable in bisyllabic words: e.g. *ɔ.nɔ* ‘there’, *i.ku* ‘dizzy’.

Given both the long history of contact with Indic languages and the increasing Newar/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 *mɔca-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.

i	u
ye	wɔ
ɛ:	
æ:	ɔ
a	

There are four simple vowels: /i/, /a/, /u/, /ɔ/. Of these four, the vowel /ɔ/ exhibits the most allophonic variation, occurring as [ɔ], [ɔ̃], and [ɑ]. The mid back rounded vowel, represented as /wɔ/, 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 /wɔ/ (cf. Kölver and Shresthacarya 1994: x–xi). The long vowels [ɛ:] and [æ:] are represented as [ɔe] and [æe] following conventions in Devanagari transliteration, but are not true diphthongs. Diphthongs end in high vowels, the most common of which are /ai/, /au/, /ɔi/, /ɔu/, less frequently /ei/, and /eu/.

### 3 INFLECTIONAL MORPHOLOGY

Word classes in Newari 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 (*hyaū*: ‘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 animate nouns via the suffix *-tɔ*, i.e. *khica-tɔ* ‘dogs’; honoured plural nouns are marked with *-pī*: i.e. *pasa-pī*: ‘friends’. Inanimate nouns do not normally take plural suffixes (Hale 1986: xxxiii).

##### 3.1.2 Case

Case is marked by suffixes on noun phrases (Kölver 1976; Hale and Manandhar 1980). Nouns with a final open syllable and long vowel in their absolutive 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’. The choice of which final consonants occur is lexically idiosyncratic.

Case forms include: unmarked absolutive, 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 *-sā:*, *-s̄:* and *-syā:* 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>pasapī:</i>	<i>mikha</i>
Ergative	<i>pasā:</i>	<i>pasapisā:</i>	

Dative	<i>pasayatɔ</i>	<i>pasapī:tɔ</i>	
Associative	<i>pasayake</i>	<i>pasapī:ke</i>	
Genitive	<i>pasaya</i>	<i>pasapini</i>	<i>mikhaya</i>
Locative			<i>mikhae</i>
Instrumental/Ablative			<i>mikhā:</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:pī:</i>	<i>jipī:</i>
Ergative	<i>jī:</i>	<i>jhi:sɔ:</i>	<i>jimisɔ:</i>
Dative	<i>jitɔ</i>	<i>jhi:tɔ</i>	<i>jimitɔ</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:pī:</i>	<i>chi</i>	<i>chikpī:</i>
Ergative	<i>chɔ:</i>	<i>chimisɔ:</i>	<i>chī:</i>	<i>chikpisɔ:</i>
Dative	<i>chɔntɔ</i>	<i>chimitɔ</i>	<i>chitɔ</i>	<i>chikpintɔ</i>
Associative	<i>chɔnke</i>	<i>chimike</i>	<i>chike</i>	<i>chikpinke</i>
Genitive	<i>chɔ:(-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 (1986) and Josī, Ş.K. (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>ipī:</i>	<i>wɔekɔ:</i>	<i>wɔekɔ:pī:</i>
Ergative	<i>wɔ:</i>	<i>imisɔ:</i>	<i>wɔekɔlɔ:</i>	<i>wɔekɔ:pisɔ:</i>
Dative	<i>wɔyatɔ</i>	<i>imitɔ</i>	<i>wɔekɔ:yatɔ</i>	<i>wɔekɔ:pintɔ</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’, *gɔnɔ* ‘where’, *gɔbɔle* ‘when’, *gɔthe* ‘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 absolutive ‘subject’, as well as in emphatic adnominal position.

### 3.1.3 Numeral classifiers

Numeral 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: *-hmɔ* ‘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: *-mɔna* ‘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/evidentiality – but does not directly encode person/subject. The most important morphophonemic contrasts occur with stem final consonant alternations whose distributions define the verb classes given below (Hale 1973, 1986; Shrestacharya 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>	cry out
Class 5	p,t,k class	<i>bhalɔp-</i>	think
		<i>sɔ:t-</i>	call, invite
		<i>penk-</i>	kick

The verbal template is ROOT-stem 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.

Class	Past	Perfective	Imperfective	Non-past	Non-past
	<i>conjunct</i>	<i>disjunct</i>	<i>disjunct</i>	<i>conjunct</i>	<i>disjunct</i>
1	CVn-a	CVn-ɔ	C $\bar{V}$ :	CVn-e	CVn-i
2	CVn-a	CVt-ɔ	CV:	CV-e	CV-i
3	CVy-a	CVl-ɔ	CV:	CV-e	CV-i
4	CVl-a	CVl-ɔ	CV:	CVl-e	CVl-i
5	CVp-a	CVp-ɔɓ	CVp-yu:	CVp-e	CVp-i
	CVt-a	CVt-ɔɓ	CVt-u:	CVt-e	CVt-i
	CVk-a	CVk-ɔɓ	CVk-u:	CVk-e	CVk-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.

The most commonly used 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 (cf. Hale 1973, 1986; Malla 1985). However, the temporal semantics in finite clauses shows a more complex relationship between the 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: (past) perfective /-ɔ/, (past) imperfective /-V:/, and non-past /-i/.

### 3.2.1 Intentionality/evidentiality (conjunct/disjunct)

In addition to encoding tense/aspect contrasts, the verbal morphology exhibits a set of oppositions that index intentionality/evidential contrasts in finite clauses. Following Hale (1980), the terms most widely used in English language scholarship are the terms Conjunct (CJ) and Disjunct (DJ), though Newar language scholarship often uses the terms *ātmā* ‘self’ and *para* ‘other’, respectively (Josī 1992: 83 [NS 1112]).

Conjunct suffixes will occur whenever the action is construed as intentional, and the actor/agent is also the evidential source reporting the action (Bendix 1974, 1992; Hargreaves 1991, forthcoming). A 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.

Disjunct suffixes occur in all other finite environments except those outlined above.

The construal of intentionality is, for the most part, lexically governed. The normal distribution of conjunct suffixes 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 indexed with adverbial markers of non-volitional or evidential markers of non-awareness.
- b Non-control verbs, such as *gya-* ‘be afraid’, or *si(t)-* ‘die’, entail no potential for intentional action and disallow conjunct inflection.
- c Fluid verbs, such as *sɔn-* ‘move’ or *thi-* ‘touch’, allow for alternative interpretations of intentionality. With this small set of verbs, the occurrence of conjunct or disjunct forms may function as the sole index of intentional or non-intentional action.

### 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’ > *khya-* ‘frighten’.

A second causative process, fully productive, is marked by the causative morpheme *-k-* / *-kɔl-* suffixed to the verb stem. The causative stem plus suffix is then inflected via the tense/aspect and conjunct/disjunct system. The paradigm with the verb *ya-* ‘do’ is given below:

<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>
<i>ya-k-a</i>	<i>ya-kɔl-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-verb like verb concatenations (see Section 5.6) that follows the same stem final patterns as the past conjunct inflection, and indeed, on

morphophonological grounds appears identical with the past conjunct inflection. Nevertheless, morphosyntactically, it is non-finite, and does not index any of the tense/aspect/person/intention/evidentiality contrasts characteristic of the finite morphosyntactic environment. Similarly, the infinitive form /-e/ patterns morphophonologically exactly like the non-past conjunct, although again it is morphosyntactically a non-finite inflection, and does not index any contrasts characteristic of a finite morphosyntactic environment. A core group of sensory adjectival verbs takes the suffix /-se/ in a non-finite construction with the auxiliary *cw̃*.

## 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-sɔlɔ* ‘fly-horse’ > ‘type of mythical animal’.

### 4.2 Suffixation

Productive suffixes include the feminine marker, e.g. *iyapu* ‘labourer’/‘peasant’ > ‘*iyapu-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 is also a variety of derivational-like suffixes, although many are not fully productive (Shresthacarya 1981; Malla 1985). They include: *gyan-pu* ‘fearful’, ‘dangerous’ < *gya-* ‘be afraid’; *nɔ-sa* ‘food’ < *nɔ-* ‘eat’; *nhi-su* ‘cheerfulness’, ‘laughter’ < *nhi-* ‘to laugh’; *ɔela-gulu* ‘drunkard’ < *ɔela* ‘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-hā*: ‘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 *swɔ-* ‘into’, ‘inside’, have direction meanings but do not occur as free direction morphemes: *svɔ-ca(l)-* ‘get thrust into’, be ‘stuck into’; *svɔ-tu(t)-* ‘be bogged down’, ‘stuck in’; *svɔ-thɔn-* ‘put in safe keeping’, ‘thrust in pocket’, ‘stuff, pack’. There is a wide range of other non-productive verbal prefixes, many of which are synchronically opaque (Malla 1985: 23).



#### 4.4 Reduplication

Reduplication occurs with a variety of stems and is part productive and part lexically idiosyncratic (Shresthacharya 1976; Malla 1985: 26). Reduplication of nominal roots indicates a plural meaning elusive in its actual semantics but roughly translated as ‘and similar other things’. 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-ji*; *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’ > *jhyatu-kyatu*. 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 a manner or aspectual colouring, meaning roughly ‘doing/having done in a prototypical manner indicated by the verb root’ (cf. Malla 1985: 27). Reduplication of verbal stems accompanied by nasalization marks a progressive aspect commonly used in narration: *ya-* ‘to do’ > *ya-yā* ‘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 above, 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 *-pī*: ‘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 5.1.3). Adverbial modifiers precede adjectival modifiers. An expanded noun phrase is given below:

- (1) *thwɔji-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-ya pasa chɔ-mhɔe-syā*:  
 3-GEN friend one-CL-ERG  
 ‘By his friend ...’

## 5.2 Relative clauses

Relative clauses are a category of nominalization proposed as modifiers to a head noun and marked with one of the three attributive suffixes: *-gu* ‘inanimate’, *mhɔ* ‘animate singular’, and *-pĩ:* ‘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) | <i>ji-gu</i>     | <i>chẽ</i>                         |  | <i>ji-gu</i>              |
|     | 1st-ATR          | house                              |  | 1st-ATR                   |
|     |                  | ‘my house’                         |  | ‘mine’ (inanimate)        |
| (4) | <i>hyaũ: -gu</i> | <i>chẽ</i>                         |  | <i>hyaũ: -gu</i>          |
|     | red/IMPF-ATR     | house                              |  | red/IMPF-ATR              |
|     |                  | ‘red house’                        |  | ‘red (one)’               |
| (5) | <i>wɔ̃:-gu</i>   | <i>chẽ</i>                         |  | <i>wɔ̃:-gu</i>            |
|     | go/IMPF-ATR      | house                              |  | go/IMPF-ATR               |
|     |                  | ‘The house (s/he) has gone to ...’ |  | ‘(the place) gone to ...’ |
| (6) | <i>wɔ̃:-mhɔ</i>  | <i>manu</i>                        |  | <i>wɔ̃:-mhɔ</i>           |
|     | go/IMPF-ATR      | person                             |  | go/IMPF-ATR               |
|     |                  | ‘The person who went ...’          |  | ‘(The one) who went ...’  |
| (7) | <i>wɔ̃:-pĩ:</i>  | <i>manu-tɔ</i>                     |  | <i>wɔ̃:-pĩ:</i>           |
|     | 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’ (NOM) will be reserved for clauses occurring as arguments of copular or perception verbs (see Section 5.7) below. Note the formal similarity between the headless relativization *wɔ̃:-gu* in (5) above and the nominalization *wɔ̃:-gu* in (8) below:

- |     |             |             |            |                |                          |
|-----|-------------|-------------|------------|----------------|--------------------------|
| (8) | <i>wɔ̃:</i> | <i>sita</i> | <i>chẽ</i> | <i>wɔ̃:-gu</i> | <i>khɔn-tɔ</i>           |
|     | 3/ERG       | sita        | house      | go/IMPF-NOM    | see-PRF.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)  | <i>Manoj</i> | <i>guru</i>   | <i>(khɔ:)</i>         |
|      | Manoj        | teacher       | be/IMPF               |
|      |              |               | ‘Manoj is a teacher.’ |
| (10) | <i>Sita</i>  | <i>wɔl-tɔ</i> |                       |
|      | Sita         | come-PRF.DJ   |                       |
|      |              | ‘Sita came.’  |                       |

- (11) *jī: ja nɔy-a*  
1/ERG rice eat-PST.CJ  
'I ate rice.'
- (12) *imi-sṣ dyɔ-yatɔ me syat-ɔ*  
3/PLUR-ERG god-DAT buffalo kill-PRF.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 verbs are marked with 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) *jī 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-PRF.DJ  
'S/he went and ate/ went to eat.'

#### 5.5 Verb concatenation

Verb concatenations are serial verb-like constructions consisting of a non-finite verb form, marked with */-a/*, followed by verbs which carry directional and aspectual meanings in concatenations but have independent main verb functions in other contexts (Hargreaves 1986).

- (16) *Sita wɔy-a cwɔn-ɔ*  
come-CM stay-PRF.DJ  
'Sita is coming.'
- (17) *pwa syan-a wɔl-ɔ*  
stomach ache-CM come-PRF.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 above). Non-embedded or bare nominalized clauses without copula verbs occur frequently in conversation.

- (18) *sita wɔy-a cwɔ:-gu du*  
Sita come-CM stayIMPF-NOM be/IMPF  
'Sita has come.'

- (19) *wḥ: sita wḥy-a cwḥ:-gu khḥn-ḥ*  
 3/ERG sita come-CM stayIMPF-NOM see-PRF.DJ  
 ‘S/he saw Sita coming.’
- (20) *chḥ gḥnḥ wḥn-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) *wḥ: ja chḥ-mḥna thuy-a: nḥy-a: pihā: won-ḥ*  
 3/ERG rice one-CL cook-NF eat-NF out go-PRF.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 *dhḥka:*, or occur as main clauses marked with the quotative (QT) evidential particle *hḥ*. 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 swḥ wḥn-a hḥ*  
 outside look/PURP go-PST.CJ QT  
 ‘(S/he<sub>i</sub>) said “(I<sub>i</sub>) went outside to look”.’
- (23) *pine swḥ wḥn-ḥ hḥ*  
 outside look/PURP go-PRF.DJ qt  
 ‘(S/he<sub>i</sub>) said “(s/he<sub>i</sub>) went outside to look”.’
- (24) *jī: ḥ:pwḥ twḥn-ḥ hḥ*  
 1/ERG much drink-PRF.DJ QT  
 ‘(S/he<sub>i</sub>) said I<sub>j</sub> drank too much.’

The complementizer *dhḥka* also occurs with mental states.

- (25) *pine swḥ wḥn-e jiu: dhḥka pihā won-ḥ*  
 outside look/PURP go-INF ok/IMPRF.DJ COMP out go-PRF.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’, *-bḥle* ‘when’, *ki* ‘but’, ‘either’/‘or’, ‘if’, ‘when’, *-gulī* ‘because/since’.

- (26) *ya:-sā jiu*  
 do-IMPRF-COND be.OK/IMPERF.DJ  
 ‘(Even) if (s/he) does (it), (that’s) OK.’

The verb *dha-* ‘to say’ may function as a coordination marker: *dhae-wɔ* ‘as soon as’, *dha-sa* ‘if’, and *chae dha-sa* literally ‘why say-if’, i.e. because:

- (27) *dyɔ̃*            *won-ɔ*        *dhae-wɔ* *nhyɔ* *wɔl-ɔ*  
 sleep/PURP go-PRF.DJ say-AND sleep come-PRF.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’, *tɔɔ* ‘but’.

- (28) *dyɔ̃*            *wɔn-ɔ*        *tɔɔ* *nhyɔ* *wɔ-e*        *mɔ-phɔ-ɔ*  
 lie/PURP go-PRF.DJ but sleep come-INF NEG-able-PRF.DJ  
 ‘(S/he) went to lie down, but didn’t fall asleep.’

### 5.10 Comparatives

Comparative clauses are formed with a postpositional morpheme *swɔyaɪ*, which may be derived from the verbal root *swɔ-* ‘watch’, ‘look’ and the non-final inflection *-aɪ*.

- (29) *chɔ*        *swɔyaɪ*.        *wɔ*        *tɔ: dhi*  
 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-PRF.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-PRF.DJ PRT  
 ‘Who has come?’

- (34) *sunā:*     *yat-ɔ*  
       who/ERG come-PRF.DJ  
       ‘Who did (it)?’

### 5.13 Particles

Speech act particles appear clause finally and include: *la* ‘Yes-no interrogative’, *le* ‘presuppositional’, ‘often co-occurs with wh- questions’, *re/le* ‘persuasive’, *h̄s* ‘reported speech’, *ka* ‘assertive’. Other discourse particles appear clause internally with local scope, including: *la* ‘topic’, *n̄s* ‘also’, *ɟɔkɔ* ‘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*, Cambridge: Cambridge University Press.
- Driem, G. van (1993) ‘The Newar verb in Tibeto-Burman perspective’, *Acta Linguistica Hafniensia* 26: 23–43.
- Genetti, C. (1988) ‘A syntactic correlate of topicality in Newari narrative’, in S. Thompson and J. Haiman (eds) *Clause Combining in Grammar and Discourse*, Philadelphia: John Benjamins.
- (1994) ‘A descriptive and historical account of the Dolakha Newari dialect’, *Monumenta Serindica* 24, Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo: Tokyo University of Foreign Studies.
- Hale, A. (1973) ‘On the form of the verbal basis in Newari’, in Braj Kachru *et al.* (eds) *Issues in Linguistics: Papers in Honor of Henry and Renee Kahane*, 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), 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*, Berlin: Mouton de Gruyter.
- (1986) ‘Users’ guide to the Newari dictionary’, in T. Manandhar (ed.) *Newari–English Dictionary*, 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 Mahandhar, T. (1980) ‘Case and role in Newari’, in R. Trail (ed.) *Papers in Southeast Asian Linguistics* 7 (Pacific Linguistics Series A, no. 53), Canberra: Australian National University.
- Hargreaves, D. (1986) ‘Independent verbs and auxiliary functions in Newari’ *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society* 12: 401–12.
- (1991) ‘The conceptual structure of intentional action: data from Kathmandu Newari’, *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society* 17: 379–89.
- (1996) ‘From interrogation to topicalization: PTB \*la in Kathmandu Newar’, *Linguistics of the Tibeto-Burman Area* 19.2: 31–44.
- (forthcoming) ‘Intentional action in Kathmandu Newar’, in G. van Driem and W. Winter (eds) *Himalayan Linguistics*, Trends in Linguistics Series, Berlin: Mouton de Gruyter.
- Jørgenson, H. (1931) ‘A dictionary of the Classical Newari’, *Det. Kgl. Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser* 23.1.
- (1941) ‘A grammar of the Classical Newari’, *Det. Kgl. Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser* 27.3.

- Josī, Ṣ.K. (1992) [NS 1112] 'Nepāl bhāṣāyā bhāṣāvaijñānika vyakaraṇa' (A linguistic grammar of nepāl bhāṣa (Newar)), Kathmandu: Lacoul Publications.
- Kansakar, T.R. (1982) 'Morphophonemics of the Newari verb', in T.R. Kansakar (ed.) *Occasional Papers in Nepalese Linguistics* 12–29. Linguistic Society of Nepal Publication No.1, Lalitpur, Nepal.
- (1997) 'The Newar language: a profile', *Newāh Vijñāna: Journal of Newar Studies* 1.1: 11–28.
- Kölver, U. (1976) 'Satztypen und verbsubcategorisierung der Newari', *Structura* 10, Munich: Fink Verlag.
- (1977) 'Nominalization and lexicalization in Newari', *Arbeiten des Kölner Universalen-Projekts* 30.
- 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 Pvt. Ltd.
- (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.
- Shakya, D.R. (1992) 'Nominal and verbal morphology in six dialects of Newari', unpublished masters thesis, University of Oregon.
- Shrestha, Uma (1990) 'Social networks and code-switching in the Newar community of Kathmandu City', unpublished PhD dissertation, Ball State University.
- Shresthacharya, I. (1976) 'Some types of reduplication in the Newari verb phrase', *Contributions to Nepalese Studies* 3.1: 117–27.
- (1981) 'Newari root verbs', *Bibliotheca Himalayica* 2.1, Kathmandu: Ratna Pustak Bhandar.

**PART 8**

---

**NORTHEASTERN INDIA**

---





# GARO<sup>1</sup>

*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 Northeastern 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 11 on Northeast Indian Languages, this volume.) 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.

---

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 forty 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.

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, but 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 Northeastern 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 (Burling 1961) and hope to complete a fuller description of the language in the near future.

## 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 *č* and *ŋ* 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 has no part in what Garos regard as the 'English' alphabet, and so it does not seem like a real letter. I will deviate from convention by indicating syllable boundaries with a hyphen, 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 per cent of syllable boundaries probably correspond to morpheme boundaries. A single syllable virtually never includes 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 23.1, and finals in Table 23.2. 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

**TABLE 23.1** 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 23.2** CODAS: SYLLABLE FINAL CONSONANTS AND CLUSTERS

p	t	k	'
m	n	ng	
m'	n'	ng'	
l	l'	(r)	(s)

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. 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 complimentary 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.

*A-gan-chak -tai -ja -wa-kon.*  
 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 actor 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 an SOV 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 -not I -EMPH  
'[I] haven't eaten yet, not me.'

*Bi-a gip-in song -o-na kat -ang -a-ha, bi-a.*  
he/she different village-to run-away-PAST he/she  
'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 postposition. Demonstratives and genitives always come first, while case makers 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

TABLE 23.3 PRONOUNS

	<i>Free/Nominative</i>	<i>Combining</i>
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-si-mang-</i>

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.

#### 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 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.*  
 I 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.*  
 I -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-cha 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.*  
 I 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.*  
 I 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.*  
 my 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 centre 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* 'woodpecker', *do'-po* 'owl', etc. The second part of such words gives the word its specific meaning but many of these second parts are never be 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 post-positions: *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 of comparison is mentioned, it is, like all nouns, placed before the verb. It takes the dative case marker, *-na*, and is followed by *bat-e*, which anticipates the later echoing *-bat-* of the verb.

*U-a me’-chik bi-ni se-gip-a -na bat-e chang-ro -bat -a*  
 That woman her husband -DAT than tall -more -PRES  
 ‘That woman is taller than her husband.’

## 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.*

I house-to 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*

I-POS 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*

ten o'clock-LOC I village-to arrive-there-past arrive-there-SUB I

*mi cha'-e tu-si -a-ha.*

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 that *-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*

he come-here-not-if I 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.*

we (inclusive) 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: main, nominalized, or subordinate.

Verbs can also be made subordinating 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-ness because I 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-ing after -LOC we eat -ables -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-have-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, shop keepers, 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.

## 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.
- Duanmu, San (1964) '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. 1794: 17–37 (reprinted at New Delhi: Cosmo Publications, 1979).
- Hamilton, Francis (1940) *An Account of Assam. With Some Notices Concerning the Neighbouring Territories*, edited by S.K. Bhuyan, Gauhati, Assam: Government of Assam (first published in 1820).
- Holbrook, L.M. (1998) *Ku'rongdik: A'chikku into English*, The Garo Literature Society: Tura, Meghalaya (first publication from a 1940 manuscript).
- Keith, T.J. (1874) *Outline Grammar of the Garo Language*, Sibsagar, Assam.
- 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, E.G. (1904) *Outline Grammar of the Garo Language*, Shillong: Assam Secretariat Press.

# JINGHPO

## *Dai Qingxia and Lon Diehl*

The ethnic Kachin are distributed across northern Myanmar, spilling westward into north-eastern India and eastward into southwestern China. In China they are called *Jinghpo*; 1990 census figures number those living in China at 119,276. The Jinghpo speak two main languages: Jinghpo and Zaiwa; they also speak languages such as Maru, Lachi, and Bola, which are relatively similar to Zaiwa (also known as Atsi).

The Jinghpo language belongs to the Jinghpo sub-branch of the Tibeto-Burman branch of the Sino-Tibetan language family. It is in common use in the Dehong Dai and Jinghpo Autonomous Prefecture of Yunnan Province, mainly in the following counties: Yingjiang, Ruili, Lianghe, Longchuan, and Luxi. Some speakers live in small compact communities; others live interspersed with speakers of Zaiwa.

The Jinghpo writing system was devised at the end of the nineteenth century; it is based on the Roman alphabet. Jinghpo dictionaries, newspapers, textbooks, and other reading materials have been published.

In this chapter, Jinghpo pronunciation and grammar are represented by the speech of Tongbi Guan, Yingjiang County.

## 1 PHONOLOGY

Jinghpo phonology is characterized by voiceless plosives and affricates without voiced correlates, by palatalized and retroflex bilabial and velar onsets, by vowels distinguished by the presence and absence of laryngealization, by falling diphthongs and by consonantal codas. There are seven consonantal codas: [-*m*, -*n*, -*ŋ*, -*p*, -*t*, -*k*, -*ʔ*]. There are four tones: high level, mid level, low falling, and falling (from high to low); of these four the falling tone occurs infrequently. In many disyllabic words the initial syllable is reduced: [*mǎ*-, *kǎ*-, *ʃǎ*-] and so on.

### 1.1 Syllable onsets

Jinghpo has been analysed as having thirty-one syllable onsets; they are displayed in Table 24.1.

Three onsets [*tsh*, *tʃh*, *f*] occur only in loanwords. Retroflex onsets [*pʒ*, *phʒ*, *kʒ*, *khʒ*] have the tongue tip curled back slightly, with slight friction during release of the plosives. Nasal segments [*m*, *n*, *ŋ*] can constitute full syllables, as in [*m*<sup>33</sup>*puŋ*<sup>33</sup>] the wind, [*n*<sup>31</sup>*tan*<sup>33</sup>] a bow, [*ŋ*<sup>33</sup>*kau*<sup>33</sup>] part [of some whole]. Lexical contrasts among onsets are illustrated in Table 24.2.

### 1.2 Rhymes

There are eighty-eight rhymes, including ten simple vowels, eight diphthongs, and seventy rhymes with consonantal codas. They are displayed in Table 24.3.

Codas [-*p*, -*t*, -*k*, -*ʔ*] are unreleased. The neutral vowel [ǎ] in reduced syllables assimilates to certain onsets. After [*w*], the vowel in a reduced syllable is [*u*] or [ǔ]; after [*ts*] or [*s*], it is



TABLE 24.1 SYLLABLE ONSETS

		Bilabial	Labio-dental	Front apcal	Apical	Laminal	Front dorsal	Dorsal
Plosives	unaspirated	<i>p pj pʒ</i>			<i>t</i>			<i>k kj kʒ</i>
	aspirated	<i>ph phj phʒ</i>			<i>th</i>			<i>kh khj khʒ</i>
Affricates	unaspirated			<i>ts</i>		<i>tʃ</i>		
	aspirated			<i>tsh</i>		<i>tʃh</i>		
Fricatives	voiceless		<i>f</i>	<i>s</i>		<i>ʃ</i>		<i>x</i>
	voiced					<i>ʒ</i>		
Nasals		<i>m mj</i>			<i>n</i>			<i>ŋ ŋj</i>
Liquids					<i>l</i>			
Semi-vowels		<i>w</i>					<i>j</i>	

TABLE 24.2 LEXICAL CONTRASTS AMONG ONSETS

<i>p</i>	<i>pa</i> <sup>31</sup>	elder brother! (term of address)		<i>l</i>	<i>la</i> <sup>33</sup>	man
<i>ph</i>	<i>pha</i> <sup>31</sup>	thin (not thick)		<i>tʃ</i>	<i>tʃaŋ</i> <sup>33</sup>	if, given that
<i>m</i>	<i>ma</i> <sup>31</sup>	child		<i>tʃh</i>	<i>kuŋ</i> <sup>31</sup> <i>tʃhan</i> <sup>33</sup> <i>taŋ</i> <sup>33</sup>	communist party (< Chinese <i>gòngchǎndǎng</i> )
<i>pj</i>	<i>pja</i> <sup>ʔ55</sup>	to collapse		<i>ʃ</i>	<i>ʃa</i> <sup>31</sup>	only, just
<i>phj</i>	<i>phja</i> <sup>ʔ55</sup>	to tear down, raze		<i>ʒ</i>	<i>ʒa</i> <sup>31</sup>	to lack; to need
<i>mj</i>	<i>mju</i> <sup>55</sup>	kind, class, sort		<i>j</i>	<i>ja</i> <sup>33</sup>	to give
<i>pʒ</i>	<i>pʒi</i> <sup>31</sup>	Whiz! (onomatopoeic)		<i>k</i>	<i>ka</i> <sup>31</sup>	speech; talk; language
<i>phʒ</i>	<i>phʒi</i> <sup>31</sup>	iron (the substance)		<i>kh</i>	<i>kha</i> <sup>ʔ31</sup>	water (as it occurs in nature); river
<i>f</i>	<i>fa</i> <sup>31</sup>	to issue; distribute [< Chinese <i>fā</i> ]		<i>ŋ</i>	<i>ŋa</i> <sup>31</sup>	to be (at a place)
<i>w</i>	<i>wa</i> <sup>33</sup>	tooth		<i>kj</i>	<i>kje</i> <sup>ʔ31</sup>	diligent
<i>ts</i>	<i>tso</i> <sup>ʔ31</sup>	a lock		<i>khj</i>	<i>khjet</i> <sup>31</sup>	strip of land between hills
<i>tsh</i>	<i>tshan</i> <sup>33</sup>	to visit (an institution)		<i>ŋj</i>	<i>ŋje</i> <sup>ʔ55</sup>	my
	<i>kon</i> <sup>33</sup>	(< Chinese <i>cānguān</i> )				
<i>s</i>	<i>so</i> <sup>ʔ31</sup>	to invite (in advance)		<i>kʒ</i>	<i>kʒai</i> <sup>31</sup>	very
<i>t</i>	<i>tu</i> <sup>31</sup>	to arrive		<i>khʒ</i>	<i>khʒai</i> <sup>33</sup>	completely, without exception
<i>th</i>	<i>thu</i> <sup>31</sup>	to pound (something in a mortar)		<i>x</i>	<i>xit</i> <sup>31</sup>	to catch one's breath
<i>n</i>	<i>na</i> <sup>33</sup>	to be drunk				

TABLE 24.3 RHYMES

<i>i</i>	<i>e</i>	<i>a</i>	<i>o</i>	<i>u</i>	<i>ĩ</i>	<i>ɛ</i>	<i>ɑ</i>	<i>o</i>	<i>u</i>
		<i>ai</i>	<i>oi</i>	<i>ui</i>		<i>ai</i>	<i>oi</i>	<i>ui</i>	
		<i>au</i>				<i>au</i>			
<i>im</i>	<i>em</i>	<i>am</i>	<i>om</i>	<i>um</i>	<i>ĩm</i>	<i>ɛm</i>	<i>ɑm</i>	<i>om</i>	<i>um</i>
<i>in</i>	<i>en</i>	<i>an</i>	<i>on</i>	<i>un</i>	<i>ĩn</i>	<i>ɛn</i>	<i>ɑn</i>	<i>on</i>	<i>un</i>
<i>iŋ</i>	<i>eŋ</i>	<i>aŋ</i>	<i>oŋ</i>	<i>uŋ</i>	<i>ĩŋ</i>	<i>ɛŋ</i>	<i>ɑŋ</i>	<i>oŋ</i>	<i>uŋ</i>
<i>ip</i>	<i>ep</i>	<i>ap</i>	<i>op</i>	<i>up</i>	<i>ĩp</i>	<i>ɛp</i>	<i>ɑp</i>	<i>op</i>	<i>up</i>
<i>it</i>	<i>et</i>	<i>at</i>	<i>ot</i>	<i>ut</i>	<i>ĩt</i>	<i>ɛt</i>	<i>ɑt</i>	<i>ot</i>	<i>ut</i>
<i>ik</i>	<i>ek</i>	<i>ak</i>	<i>ok</i>	<i>uk</i>	<i>ĩk</i>	<i>ɛk</i>	<i>ɑk</i>	<i>ok</i>	<i>uk</i>
<i>iʔ</i>	<i>eʔ</i>	<i>aʔ</i>	<i>oʔ</i>	<i>uʔ</i>	<i>ĩʔ</i>	<i>ɛʔ</i>	<i>ɑʔ</i>	<i>oʔ</i>	<i>uʔ</i>

[ɣ] or [ʔ]; after [tʃ], [f] or [ʒ], it is [ɣ] or [ʔ]. Otherwise it is [ə] or [ɛ]. In the high level tone some laryngealized vowels alternate with nonlaryngealized variants. For example, [kɑ̌<sup>55</sup>] alternates with [ka<sup>55</sup>] as earth, place. Lexical contrasts among these rhymes are illustrated in Table 24.4.

### 1.3 Tones

There are four tones altogether. Tone 1 is marked as a high level tone, e.g. [mu<sup>55</sup>] ‘work’, [kɑ̌<sup>55</sup>wa<sup>55</sup>] ‘bamboo’, but only in laryngealized syllables is it actually pronounced as a high level tone ([pǎi<sup>55</sup>] left (as opposed to right)); in non-laryngealized syllables it is pronounced as a mid-to-high rising tone ([pai<sup>35</sup>] again). Tone 2 is a mid level tone: [mu<sup>33</sup>] ‘delicious’, [wa<sup>33</sup>] ‘tooth’. Tone 3 is a low falling tone: [mu<sup>31</sup>] ‘to see’, [wa<sup>31</sup>] ‘to return’. Tone 4 falls from a high pitch across the full pitch range to a low pitch: [nu<sup>51</sup>] ‘mother’ (term of direct address), [wa<sup>51</sup>] ‘father’ (term of direct address). All four tones occur in unchecked syllables, while checked syllables are either the high level tone or the low falling tone. The great preponderance of high-to-low falling tones occur as a result of sandhi. For example, the expression combining [n<sup>55</sup>] ‘not with’ [wa<sup>31</sup>] ‘to return’ is [n<sup>55</sup>wa<sup>51</sup>] ‘not to return’.

### 1.4 Syllable structure

Seven types of syllable structures occur. Type 1 is V+T (vowel with tone): [aʔ<sup>31</sup>] [x]’s. Type 2 is V+V+T: [ai<sup>31</sup>] ‘very’. Type 3 is V+C+T (vowel plus consonant with tone): [aŋ<sup>31</sup>] appropriate. Type 4 is C+V+T: [na<sup>33</sup>] ‘ear’. Type 5 is C+V+V+T: [nai<sup>31</sup>] ‘taro’. Type 6 is C+V+C+T: [naŋ<sup>33</sup>] you (singular). Type 7, which involves syllabic nasals, is C+T: [n<sup>55</sup>] ‘not’.

## 2 GRAMMAR

### 2.1 Word formation

Of monomorphemic words two kinds occur: monosyllabic and polysyllabic. Monosyllables are illustrated by [ɣai<sup>33</sup>] first person singular pronoun, [fa<sup>31</sup>] ‘food’, [sa<sup>33</sup>] ‘to go’, and [phɣo<sup>31</sup>] ‘white’. Most polysyllabic words are disyllabic; these disyllabic monomorphemic words typically involve phenomena such as alliteration and rhyme reduplication, as in [koŋ<sup>31</sup>kaŋ<sup>31</sup>] ‘celery’ and [tʃiŋ<sup>31</sup>liŋ<sup>55</sup>] ‘spirit bird’.

Polymorphemic words involve compounding in some cases and affixation in others. Four kinds of compounds occur. In the first kind the morphemes are coordinate: [phu<sup>31</sup>] ‘elder brother(s)’ and [nɑ̌u<sup>33</sup>] ‘younger brother(s)’ are coordinate in [phu<sup>31</sup>nɑ̌u<sup>33</sup>]; so are [lɑ̌<sup>31</sup>mu<sup>31</sup>] ‘heaven’ and [ka<sup>55</sup>] ‘earth’ [lɑ̌<sup>31</sup>mu<sup>31</sup>ka<sup>55</sup>] ‘in universe’; [kun<sup>33</sup>] ‘to carry’ and [pha<sup>33</sup>] ‘to lift’ or ‘raise’ in [kun<sup>33</sup>pha<sup>33</sup>] ‘to bear’ or ‘shoulder (a burden)’; [dip<sup>31</sup>] ‘to press’ and [sep<sup>31</sup>] ‘to peel’ in [dip<sup>31</sup>sep<sup>31</sup>] ‘to exploit’.

In the second kind of compound formation, one morpheme modifies the other. If the modifying element is nominal, it precedes the nominal head; if it is adjectival, it follows. For example nominal [sai<sup>31</sup>] ‘blood’ precedes [lam<sup>33</sup>] ‘road’, ‘path’ in [sai<sup>31</sup>lam<sup>33</sup>] ‘blood vessel’, and nominal [mji<sup>31</sup>] ‘eye’ precedes [mun<sup>33</sup>] ‘body hair’ in [mji<sup>31</sup>mun<sup>33</sup>] ‘eyelashes’. Adjectival [tui<sup>31</sup>] ‘sweet’ follows [tʃum<sup>31</sup>] salt in [tʃum<sup>31</sup>tui<sup>31</sup>] ‘sugar’ and [tʃɑ̌ŋ<sup>33</sup>] black follows [pjen<sup>33</sup>] ‘board’ in [pjen<sup>33</sup>tʃɑ̌ŋ<sup>33</sup>] ‘blackboard’. When verbal morphemes modify nominal ones, some precede the head but the majority follow. The modifying verbal morpheme [ɣam<sup>31</sup>] ‘to be left over’ follows the nominal head [fat<sup>31</sup>] ‘food’ in [fat<sup>31</sup>ɣam<sup>31</sup>] ‘leftover’.

TABLE 24.4 LEXICAL CONTRASTS AMONG RHYMES

<i>i</i>	<i>pʒi</i> <sup>33</sup>	clear (of speech)	<i>i</i>	<i>pʒi</i> <sup>33</sup>	slippery
<i>e</i>	<i>tʃe</i> <sup>31</sup>	even more so	<i>e</i>	<i>tʃe</i> <sup>33</sup>	to know
<i>a</i>	<i>ka</i> <sup>31</sup>	speech, talk	<i>a</i>	<i>ka</i> <sup>31</sup>	to dance
<i>o</i>	<i>po</i> <sup>33</sup>	head	<i>o</i>	<i>pa</i> <sup>33</sup>	to come into being
<i>u</i>	<i>pu</i> <sup>31</sup>	to put on (a lower garment)	<i>u</i>	<i>pu</i> <sup>31</sup>	intestines
<i>ai</i>	<i>taɪ</i> <sup>33</sup>	that	<i>ai</i>	<i>taɪ</i> <sup>33</sup>	to become
<i>au</i>	<i>kʒau</i> <sup>33</sup>	more so, most so	<i>au</i>	<i>kʒau</i> <sup>33</sup>	withered, dried out
<i>oi</i>	<i>koi</i> <sup>31</sup>	(for a bird or animal to make its characteristic cry); to screech; caw; chirp; hiss	<i>oi</i>	<i>koi</i> <sup>31</sup>	to avoid; stay clear of
<i>ui</i>	<i>tui</i> <sup>31</sup>	sweet	<i>ui</i>	<i>tui</i> <sup>33</sup>	to fester
<i>im</i>	<i>tsim</i> <sup>31</sup>	quiet	<i>im</i>	<i>tsim</i> <sup>33</sup>	to dive; go under (water)
<i>in</i>	<i>tin</i> <sup>33</sup>	to put on; wear (shoes)	<i>in</i>	<i>tin</i> <sup>33</sup>	anxious; impatient; stirred up
<i>iŋ</i>	<i>tiŋ</i> <sup>33</sup>	straight	<i>iŋ</i>	<i>tiŋ</i> <sup>33</sup>	to plant (seed) by dibbling
<i>ip</i>	<i>kjip</i> <sup>55</sup>	narrow	<i>ip</i>	<i>kjip</i> <sup>31</sup>	to shrivel up; droop
<i>it</i>	<i>tʃit</i> <sup>31</sup>	urine	<i>it</i>	<i>tʃit</i> <sup>31</sup>	develop (scabies)
<i>ik</i>	<i>tik</i> <sup>31</sup>	to the extreme; all the way	<i>ik</i>	<i>tik</i> <sup>31</sup>	to be about to (do something)
<i>iʔ</i>	<i>tiʔ</i> <sup>31</sup>	to pick, pluck	<i>iʔ</i>	<i>tiʔ</i> <sup>31</sup> <i>tiʔ</i> <sup>31</sup>	Splat! Splat! (onomatopoeic)
<i>em</i>	<i>tem</i> <sup>31</sup>	to photograph	<i>em</i>	<i>tem</i> <sup>31</sup>	clear-cut, straightforward
<i>en</i>	<i>tʃen</i> <sup>33</sup>	to stretch (something) out taut	<i>en</i>	<i>tʃen</i> <sup>33</sup>	slice; petal; section; segment
<i>eŋ</i>	<i>teŋ</i> <sup>31</sup>	to blame	<i>eŋ</i>	<i>teŋ</i> <sup>31</sup>	true; accurate
<i>ep</i>	<i>tep</i> <sup>31</sup>	catch up with; keep abreast of	<i>ep</i>	<i>tep</i> <sup>55</sup>	in a clasping; wrapping manner
<i>et</i>	<i>tset</i> <sup>31</sup>	diligent	<i>et</i>	<i>tset</i> <sup>31</sup>	brushing past
<i>ek</i>	<i>tek</i> <sup>31</sup>	each	<i>ek</i>	<i>tek</i> <sup>31</sup>	Tick! Tock! or Drip! Drip! (onomatopoeic)
<i>eʔ</i>	<i>keʔ</i> <sup>55</sup>	to congeal	<i>eʔ</i>	<i>keʔ</i> <sup>55</sup>	concave; dented; sunken
<i>am</i>	<i>kam</i> <sup>31</sup>	luck	<i>am</i>	<i>kam</i> <sup>31</sup>	to believe
<i>an</i>	<i>kan</i> <sup>33</sup>	acting at once, losing no time	<i>an</i>	<i>kan</i> <sup>33</sup>	stomach
<i>aŋ</i>	<i>taŋ</i> <sup>31</sup>	to win; gain victory	<i>aŋ</i>	<i>taŋ</i> <sup>31</sup>	in a crosswise manner
<i>ap</i>	<i>tap</i> <sup>31</sup>	fire pit; the ashes in a fire pit	<i>ap</i>	<i>tap</i> <sup>31</sup>	(to speak) in an unconvinced manner
<i>at</i>	<i>kat</i> <sup>31</sup>	to run	<i>at</i>	<i>kat</i> <sup>31</sup>	full; plump
<i>ak</i>	<i>pak</i> <sup>31</sup>	with a thump; thud; or a splash (onomatopoeic)	<i>ak</i>	<i>pak</i> <sup>31</sup>	to smoke (tobacco)
<i>aʔ</i>	<i>kaʔ</i> <sup>31</sup>	to crack; split; tear	<i>aʔ</i>	<i>kaʔ</i> <sup>55</sup>	bamboo basket
<i>om</i>	<i>pom</i> <sup>31</sup>	bomb (< English)	<i>om</i>	<i>pom</i> <sup>31</sup>	to bank up (earth); to hill up (something being planted)
<i>on</i>	<i>ton</i> <sup>33</sup>	to negotiate (a price)	<i>on</i>	<i>ton</i> <sup>33</sup>	dull; blunt
<i>oŋ</i>	<i>toŋ</i> <sup>31</sup>	a cubit	<i>oŋ</i>	<i>toŋ</i> <sup>31</sup>	tide over; pull through (a difficult time)
<i>op</i>	<i>kʒop</i> <sup>31</sup>	to shatter; to collapse	<i>op</i>	<i>kʒop</i> <sup>31</sup>	(for the voice) to grow hoarse
<i>ot</i>	<i>poʔ</i> <sup>31</sup>	to feed (somebody)	<i>ot</i>	<i>poʔ</i> <sup>31</sup>	to grow angry; lose one's temper
<i>ok</i>	<i>kok</i> <sup>31</sup>	room (in a house)	<i>ok</i>	<i>kok</i> <sup>55</sup>	tin, bottle
<i>oʔ</i>	<i>toʔ</i> <sup>31</sup>	to break or at least bend (into two lengths)	<i>oʔ</i>	<i>toʔ</i> <sup>55</sup>	to lack, be short of
<i>um</i>	<i>tum</i> <sup>33</sup>	storehouse	<i>um</i>	<i>tum</i> <sup>33</sup>	pit, stone (in a piece of fruit)
<i>un</i>	<i>tun</i> <sup>33</sup>	to lead (an animal); to tow (a disabled vehicle)	<i>un</i>	<i>tun</i> <sup>33</sup>	to dissolve
<i>uŋ</i>	<i>tuŋ</i> <sup>33</sup>	to sit; to be seated	<i>uŋ</i>	<i>tuŋ</i> <sup>33</sup>	the wife of one's father's elder brother; aunt
<i>up</i>	<i>tup</i> <sup>55</sup>	blunt (point)	<i>up</i>	<i>tup</i> <sup>55</sup>	even in length, of equal length
<i>ut</i>	<i>tut</i> <sup>31</sup>	to sell	<i>ut</i>	<i>tut</i> <sup>31</sup>	head-on; in one's face; face-to-face
<i>uk</i>	<i>puk</i> <sup>31</sup>	book (< English)	<i>uk</i>	<i>puk</i> <sup>55</sup>	to shout; to shout to somebody
<i>uʔ</i>	<i>puʔ</i> <sup>31</sup>	to come down with a case of (malaria)	<i>uʔ</i>	<i>puʔ</i> <sup>31</sup>	pop! (onomatopoeic)

food', 'leftovers'; likewise, [*lu*<sup>31</sup>] 'to drink' follows [*kha*<sup>31</sup>] water in [*lu*<sup>31</sup>*kha*<sup>31</sup>] 'drinking water'. In [*jam*<sup>33</sup>*ŋa*<sup>33</sup>] 'slave', however, verbal [*jam*<sup>33</sup>] 'to keep in bondage' precedes [*ŋa*<sup>33</sup>] 'livestock animal', and in [*pjen*<sup>33</sup>*li*<sup>33</sup>] 'aeroplane' [*pjen*<sup>33</sup>] 'to fly' precedes [*li*<sup>33</sup>] 'boat'.

The third kind of compound formation involves the object-verb relationship. In [*tsɿ*<sup>31</sup>*ku*<sup>31</sup>] 'double-edged fine-tooth comb' [*tsɿ*<sup>31</sup>] 'lice' is the object of [*ku*<sup>31</sup>] 'to get rid of'; in [*man*<sup>33</sup>*ju*<sup>33</sup>] 'mirror' [*man*<sup>33</sup>] 'face' is the object of the verb [*ju*<sup>33</sup>] 'to look at'.

The fourth kind of composition involves the subject-predicate relationship. In [*mji*<sup>31</sup>*khɜ*<sup>55</sup>] 'to be united', [*mji*<sup>31</sup>] 'thought' is the subject of the predicate [*khɜ*<sup>55</sup>] 'to meet up with'; in [*ka*<sup>31</sup>*lo*<sup>55</sup>] 'to quarrel', [*ka*<sup>31</sup>] 'speech', 'talk' is the subject of the predicate [*lo*<sup>55</sup>] 'to be much'.

Polymorphemic words formed by affixation involve prefixation, infixation, or suffixation, with prefixation predominating. The derivational prefix [*tʃ*<sup>55</sup>-] attaches to a monosyllabic adjectival morpheme to form a noun. For example, [*tʃ*<sup>55</sup>-] prefixed to [*kha*<sup>55</sup>] '(food) burnt' derives [*tʃ*<sup>55</sup>*kha*<sup>55</sup>] 'substance that has been cooked to burning'; prefixed to [*khɜ*<sup>33</sup>] 'sour it' derives [*tʃ*<sup>55</sup>*khɜ*<sup>33</sup>] 'sour substance'. The prefix [*m*<sup>55</sup>-] derives nouns from verbal and adjectival monosyllables. For example, with verbal [*tin*<sup>31</sup>] 'to separate', 'partition' it derives [*m*<sup>55</sup>*tin*<sup>31</sup>] 'partition board'; 'diaphragm'; with [*khɜ*<sup>33</sup>] 'sour' it derives [*m*<sup>55</sup>*khɜ*<sup>33</sup>] 'sour bamboo shoot'.

The infix [-*m*<sup>55</sup>-] or [-*mi*-] occurs between two instances of verbal or adjectival monosyllables to derive a noun meaning all that which... or all those who... For example, with [*ŋa*<sup>31</sup>] 'to be (at a place)' it forms [*ŋa*<sup>31</sup>*m*<sup>55</sup>*ŋa*<sup>31</sup>] 'all who are present', 'all those present'; with [*laŋ*<sup>33</sup>] 'to use' it forms [*laŋ*<sup>33</sup>*m*<sup>55</sup>*laŋ*<sup>33</sup>] 'everything being used' or 'everything of use'.

Suffixed to pronouns or nouns, [-*the*<sup>33</sup>] yields a plural form. For example, with [*n*<sup>33</sup>*tai*<sup>33</sup>] 'this'; 'this one' it yields the [*n*<sup>33</sup>*tai*<sup>33</sup>*the*<sup>33</sup>] 'these'; with [*n*<sup>33</sup>*au*<sup>33</sup>] 'younger brother' it yields [*n*<sup>33</sup>*au*<sup>33</sup>*the*<sup>33</sup>] 'younger brothers'. The suffix [-*la*<sup>31</sup>] marks animal nouns as male, [-*ji*<sup>31</sup>] as female: with [*u*<sup>31</sup>] 'chicken', 'fowl' they form [*u*<sup>31</sup>*la*<sup>31</sup>] 'cockrel', 'rooster' and [*u*<sup>31</sup>*ji*<sup>31</sup>] 'hen' respectively.

## 2.2 Word classes

The words of Jinghpo can be classified into twelve categories: nouns, pronouns, verbs, auxiliary verbs, adjectives, numerals, measure words, adverbs, manner adverbs (MAS), post-positions, conjunctions, and sentence-final words (SFWs). We briefly point out below some special characteristics.

Personal pronouns distinguish three values for number: singular, dual, and plural; these values are marked by internal alternation and the plural suffix [-*the*<sup>33</sup>]. They are given in Table 24.5.

The singular personal pronouns have special possessive forms: [*ŋje*<sup>55</sup>] 'my', [*na*<sup>55</sup>] 'your', [*khji*<sup>55</sup>] and [*fi*<sup>55</sup>] 'his', 'her'.

Among verbs, causatives are distinguished from non-causatives; this distinction can be marked with morphological prefixes or a syntactic construction involving a causative verb.

TABLE 24.5 PERSONAL PRONOUNS

	Singular	Dual	Plural
First person	<i>ŋai</i> <sup>33</sup>	<i>an</i> <sup>55</sup>	<i>an</i> <sup>55</sup> <i>the</i> <sup>33</sup>
Second person	<i>naŋ</i> <sup>33</sup>	<i>nan</i> <sup>55</sup>	<i>nan</i> <sup>55</sup> <i>the</i> <sup>33</sup>
Third person (oral usage)	<i>khji</i> <sup>33</sup>	<i>khan</i> <sup>55</sup>	<i>khan</i> <sup>55</sup> <i>the</i> <sup>33</sup>
Third person (written usage)	<i>fi</i> <sup>33</sup>	<i>fan</i> <sup>55</sup>	<i>fan</i> <sup>55</sup> <i>the</i> <sup>33</sup>

The great majority of the former are distinguished with a phonologically conditioned prefix that takes three forms: [tʃä<sup>31</sup>-], [ʃä<sup>31</sup>-] and [sä<sup>31</sup>-]. This prefix distinguishes [tʃä<sup>31</sup>phai<sup>33</sup>] ‘to make (somebody) lift (something) up’ from [pha<sup>33</sup>] ‘to lift (something) up’, ‘tʃä<sup>31</sup>khɹat<sup>31</sup>’ ‘to drop (something) from (khɹat<sup>31</sup>)’ (‘for something) to drop’, [ʃä<sup>31</sup>pa<sup>55</sup>] ‘to tire (somebody)’ from [pa<sup>55</sup>] ‘to be tired’, [ʃä<sup>31</sup>tʃɛ<sup>33</sup>] ‘to let (somebody) know’ from [tʃɛ<sup>33</sup>] ‘to know’ and [sä<sup>31</sup>tsam<sup>33</sup>] ‘to make (something) rot’, ‘decay’ from [tsam<sup>33</sup>] ‘to rot, decay’ and so on.

A number of verbs are causativized with the prefix [a<sup>31</sup>-]: [kɹop<sup>31</sup>] ‘(for something) to shatter’ and [a<sup>31</sup>kɹop<sup>31</sup>] ‘to shatter (something)’. A few verbs distinguish the causative by suppletive alternation: [ɹop<sup>33</sup>] ‘(for something) to be inside (something else)’ and [ɹop<sup>55</sup>] ‘to place (something) inside (something else)’; [tʃuŋ<sup>31</sup>] ‘(for something) to be loaded with (something else)’ and [tʃuŋ<sup>55</sup>] ‘to load (something) with (something else)’.

The syntactic causative construction simply places the verb [ʃä<sup>31</sup>ŋun<sup>55</sup>] ‘to cause’ after the non-causative verb: [sa<sup>33</sup>] ‘to go’ and [sa<sup>33</sup>ʃä<sup>31</sup>ŋun<sup>55</sup>] ‘to cause to go’; [kä<sup>31</sup>lo<sup>33</sup>] ‘to do’ and [kä<sup>31</sup>lo<sup>33</sup>ʃä<sup>31</sup>ŋun<sup>55</sup>] ‘to cause to do’.

Verbs can be reduplicated with the added meaning of ‘constantly, regularly’ or ‘slightly’. Compare [sa<sup>33</sup>] ‘to go’ and [sa<sup>33</sup>sa<sup>33</sup>] ‘to go on a regular basis’, or [tʃɛ<sup>33</sup>] ‘to know’ and [tʃɛ<sup>33</sup>tʃɛ<sup>33</sup>] ‘to know to a limited extent’.

Auxiliary verbs are relatively numerous; they follow verbs to express the direction, progress, or nature of the action. For example, [khat<sup>55</sup>] expresses that the action is reciprocal: [kä<sup>31</sup>ɹum<sup>33</sup>] means ‘to help’ and [kä<sup>31</sup>ɹum<sup>33</sup>khat<sup>55</sup>] ‘to help each other’. The auxiliary [khɹat<sup>31</sup>] specifies that the action involves downward motion: [jon<sup>33</sup>] means ‘to trickle, drip’ and [jon<sup>33</sup>khɹat<sup>31</sup>] ‘to trickle down’, ‘to drip down’. The great preponderance of auxiliary verbs has developed from verbs, with the meaning of the verb being relatively vivid and concrete and that of the auxiliary more abstract. For example, the verb [mat<sup>31</sup>] means ‘to come to be lost’; the auxiliary of the same form specifies conclusion of the action. Auxiliary verbs cannot be modified by adverbs. Like verbs, most auxiliaries can be reduplicated with the meaning of ‘constantly’, ‘regularly’: compare [kä<sup>31</sup>ɹum<sup>33</sup>khat<sup>55</sup>] ‘to help each other’ and [kä<sup>31</sup>ɹum<sup>33</sup>khat<sup>55</sup>khat<sup>55</sup>] ‘to help each other on a regular basis’.

Measure words are few. When what is measured or counted is the individual instance of what the noun represents, most nouns do not use measure words: the numeral applies directly to the noun. For example, the expression [mä<sup>31</sup>ʃa<sup>31</sup>lä<sup>55</sup>ŋai<sup>51</sup>] ‘one person’ uses only the noun [mä<sup>31</sup>ʃa<sup>31</sup>] ‘person’ and the numeral [lä<sup>55</sup>ŋai<sup>51</sup>] ‘one’; [nam<sup>31</sup>si<sup>31</sup>lä<sup>55</sup>khon<sup>51</sup>] ‘two pieces of fruit’ uses only the noun [nam<sup>31</sup>si<sup>31</sup>] ‘fruit’ and the numeral [lä<sup>55</sup>khon<sup>51</sup>] ‘two’. On the other hand, when what is being counted or measured is in terms of extent, capacity, weight, collective units, and so on, then a measure word is used. For example, the phrase [tʃum<sup>31</sup>ɹop<sup>31</sup>mji<sup>33</sup>] ‘a liang (50 grams) of salt’ uses not only the noun [tʃum<sup>31</sup>] ‘salt’ and the numeral [mji<sup>33</sup>] ‘one’, a but also the measure word [ɹop<sup>31</sup>] ‘liang’. Likewise the expression [ŋa<sup>33</sup>wä<sup>33</sup>noŋ<sup>33</sup>mji<sup>33</sup>] ‘a herd of cattle’ uses not only the noun [ŋa<sup>33</sup>] ‘cow’ and the numeral [mji<sup>33</sup>] ‘one’, a but also the measure word [wä<sup>33</sup>noŋ<sup>33</sup>] ‘herd (of some kind of animal)’. And [ʃar<sup>31</sup>wan<sup>33</sup>mji<sup>33</sup>] ‘a bowl of rice’ uses the measure word [wan<sup>33</sup>] ‘bowl (of something)’ together with the noun [ʃar<sup>31</sup>] ‘rice’, ‘food’ and the numeral [mji<sup>33</sup>] ‘one’, ‘a’.

Most measure words come from nouns and verbs, with some being loans from other languages. When a measure word is applied to a verb, the common choice is [läŋ<sup>31</sup>] ‘time(s)’, which follows the numeral instead of preceding it: [lä<sup>55</sup>khon<sup>51</sup>läŋ<sup>31</sup>] ‘twice’, ‘two times’.

Manner adverbs (MAS) are extremely numerous; they specify details about the fashion or manner in which verb-designated actions unfold. While the meanings of other adverbs are relatively abstract, the meanings of manner adverbs are relatively vivid, synthesizing and condensing perception-amenable details of state, condition, shape and so on. MAS cannot

modify ordinary verbs; they modify only certain generalised verbs (GVs) such as [*ti*<sup>33</sup>], [*ʒai*<sup>31</sup>], [*ŋa*<sup>33</sup>], [*ŋu*<sup>55</sup>] and [*ʒe*<sup>33</sup>].

- (1) *fɪ*<sup>33</sup> *ŋjan*<sup>31</sup> *ʒai*<sup>31</sup> *wa*<sup>31</sup> *mat*<sup>31</sup> *sai*<sup>33</sup>  
 3SG dilly-dally GV return completive SFW  
 'He took his time going back.'
- (2) *n*<sup>31</sup>*luŋ*<sup>31</sup> *the* *ɽ*<sup>31</sup> *mjet*<sup>55</sup> *ti*<sup>33</sup> *kǎ*<sup>31</sup>*mjet*<sup>55</sup> *tɔn*<sup>31</sup> *u* *ɽ*<sup>31</sup>  
 stone INST press-press GV press leave (in a place) SFW  
 'Press it with the stone really hard!'

Postpositions are structural particles such as [*e*ɽ<sup>55</sup>], [*a*ɽ<sup>31</sup>] and [*e*<sup>31</sup>] used following objects, adnominal modifiers, and other constituents to mark relationships among sentence constituents, as in the following examples.

- (3) *ŋai*<sup>33</sup> *fɪ*<sup>33</sup> *e* *ɽ*<sup>55</sup> *ja*<sup>33</sup> *sǎ*<sup>33</sup>*ŋai*<sup>33</sup>  
 1SG 3SG OBJ give SFW  
 'I gave it to him.'
- (4) *ŋai*<sup>33</sup>-*nau*<sup>33</sup> *a* *ɽ*<sup>31</sup> *lai*<sup>31</sup>*kǎ*<sup>33</sup>  
 1SGPOSS-younger.brother GEN book  
 'My younger brother's book.'
- (5) *fǎ*<sup>31</sup>*wa*<sup>31</sup>*mǎ*<sup>31</sup>*fǎ*<sup>31</sup> *e*<sup>31</sup> *ʒim*<sup>31</sup> *tɔn*<sup>31</sup>  
 the.masses AGT capture set down and leave [in a fixed place]  
 'To be seized by the masses.'

The SFW has numerous inflected forms. Through a process of phonological change this kind of word has come to synthesize and embody such grammatical meanings of the sentential predicate as mood, subject person and number, direction, and aspect. Mood has six values: declarative, interrogative, imperative, consultative, inferential and inexpectative. The imperative and consultative moods each subdivide into ordinary vs emphatic manners of speaking. The declarative, interrogative, inferential and inexpectative moods each subdivide on the basis of speaker attitude and intent into two subcategories: static and dynamic. Person has three values: first person, second person, and third person. Number has two values: singular and plural. Since the sentence-final word can reflect person and number, one can determine subject person and number even when the subject constituent is omitted. We display a subset of the declarative SFWs in Table 24.6. These SFWs displayed are illustrated in the following sentences.

- (6) *ŋai*<sup>33</sup> *ko*<sup>31</sup> *tfoŋ*<sup>31</sup>*ma*<sup>31</sup> *ʒai*<sup>55</sup> *ŋ*<sup>31</sup>*ŋai*<sup>33</sup>  
 1sg TOP student be SFW (1SG subject, static, declarative)  
 'I am a student.'
- (7) *naŋ*<sup>33</sup> *kʒai*<sup>31</sup> *tʃe*<sup>33</sup> *n*<sup>31</sup>*tai*<sup>33</sup>  
 2SG very know SFW (2SG subject, static, declarative)  
 'You know how to do a lot.'
- (8) *khji*<sup>33</sup> *kʒai*<sup>31</sup> *kǎ*<sup>31</sup>*pu*<sup>33</sup> *ai*<sup>33</sup>  
 3SG very glad SFW (3SG subject, static, declarative)  
 'He is very happy.'

TABLE 24.6 A SAMPLE SUBSET OF SFWs IN THE DECLARATIVE MOOD

	Static		Dynamic	
	Singular	Plural	Singular	Plural
First person	$\eta^{31}\eta ai^{33}$	$ka\tau^{31}ai^{33}$	$s\check{a}^{33}\eta ai^{33}$	$s\check{a}^{55}ka\tau^{55}ai^{33}$
Second person	$n^{31}tai^{33}$	$m\check{a}^{31}tai^{33}$	$sin^{33}tai^{33}$	$m\check{a}^{33}sin^{33}tai^{33}$
Third person	$ai^{33}$	$ma\tau^{31}ai^{33}$	$sai^{33}$	$m\check{a}^{33}sai^{33}$

- (9)  $\eta ai^{33}$   $mu^{31}$   $ju^{33}$   $\eta^{31}\eta ai^{33}$   
 1SG see have [done something] SFW (1SG subject, static, declarative)  
 'I have seen it.'
- (10)  $sa^{33}$   $wa^{31}$   $m\check{a}^{33}sai^{33}$   
 go auxiliary verb SFW (3PL subject, static, declarative)  
 'They have gone back.'

### 2.3 Constituent order

As major mechanisms for expressing grammatical meaning, Jinghpo uses sentence-final word inflection, structural particles and constituent order. Basic orders include the following: subject, object, predicate; subject, indirect object, direct object, predicate; adnominal modifying noun, head noun; head noun, adnominal modifying adjective; head noun, numeral adnominal; numeral adverbial, verbal head; adverbial, head verb or adjective. These constituent-order patterns are illustrated below.

- (11)  $khji^{33}$   $wa^{31}$   $sai^{33}$   
 3SG return SFW (3SG subject, static, declarative)  
 'He has come back.'
- (12)  $\eta ai^{33}$   $lai^{31}ka^{33}$   $thi^{55}$   $\eta a^{31}$   $\eta^{31}\eta ai^{33}$   
 1SG book read be.doing SFW (1SG subject, static, declarative)  
 'I am reading.'
- (13)  $\eta ai^{33}$   $khji^{33}$   $e\tau^{55}$   $lai^{31}ka^{33}$   $ja^{33}$   $s\check{a}^{33}\eta ai^{33}$   
 1SG 3SG OBJ book give SFW (1SG subject, static, declarative)  
 'I gave him a book.'
- (14) a.  $phun^{55}$   $ka^{31}pa^{31}$  b.  $ka^{31}pa^{31}$   $ai^{33}$   $phun^{55}$   
 tree big big SFW tree  
 'a big tree' 'a tree that is big'
- (15) a.  $n^{55}$   $fa^{55}$  b.  $m\check{a}^{31}fa^{31}$   $la^{55}\eta ai^{51}$   
 not eat person one  
 'not eat' 'one person'
- (16)  $n^{33}tsin^{33}$   $pa^{33}lin^{33}$   $mji^{33}$   
 water bottle one  
 'a bottle of water'
- (17) a.  $la^{55}kho\eta^{51}$   $la\eta^{31}$   $sa^{33}$  b.  $k\check{a}ai^{31}$   $ka^{31}fa^{33}$   
 two times go very good  
 'to go twice' 'very good'

# HAKHA LAI\*

*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 about 100,000 people, Lai is also used extensively as a second language by speakers of other Chin languages in the Chin Hills. It 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 aspects of the language's phonology. Certain characteristics 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; spoken in the Falam area), Bawm (Reichle 1981; spoken mostly in Bangladesh), and Mizo (Chhange 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.

## 2 PHONOLOGY

### 2.1 Segmental phonology

Table 25.1 gives the segmental phoneme inventory. Where the orthography used does not reflect standard phonetic values, a more 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 unnoted in previous descriptions. Nonetheless, tonal distinctions are important. 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

---

\* Many thanks to Ken VanBik for commenting on an earlier draft of this sketch and for providing and discussing Lai data on demand. Almost all of the insights in this chapter are a result of my own work with him, or his work and work with others over the last several years. Thank you also to Bernard Comrie for helpful criticisms and suggestions.



**TABLE 25.1** SEGMENTAL INVENTORY

<i>Consonants</i>					<i>Vowels</i>	
<b>Stops</b>					<b>Simplex vowels</b>	
p	t	t̚	k	ʔ	i [i], ii [i:] (y)	u [u], uu [u:] (w)
ph [pʰ]	th [tʰ]	th̥ [tʰ̥]	kh [kʰ]			
b	d		g		e [ɛ], ee [ɛ:]	o [o], oo [o:]
<b>Affricates</b>					a [ə], aa [a:]	
	c/ts				<b>Diphthongs</b>	
	ch/tsh [tʃʰ]				uy, oy, ooy	iw, ew, eew,
	tl [tʃ]				ay [ɛy], aay	aw [ow], aaw
	thl [tʃʰ]					
<b>Fricatives and laterals</b>					ia, iaa, ua, uaa	
f	s [~ʃ]			h		
v	z				<b>Triphthongs</b>	
	r	l			uay	iaw
	hr [r]	hl [ɬ]				
<b>Nasals</b>						
m	n		ŋ			
hm [m̩]	hn [n̩]		hŋ [ŋ̩]			

following forms: falling *ʔa-dìŋ* ‘he drinks’ vs level *ʔa-diŋ* ‘he is honest/it is straight’; level *ʔan-ʔii-law* ‘they’re similar’ vs rising *ʔan-láw* ‘they disappear’; and falling *ʔa-màn* ‘it sells’ vs rising *ʔa-mán* ‘its price’.

Glottalization is a suprasegmental feature involved in marking distinctions in particular verbal ablaut classes and in a transitivizing derivation (see below). 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 voiceless sonorants, fricatives, affricates, and voiced and voiceless aspirate stops do not occur 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 below.

#### 3.1.1 Pronominals

Independent pronouns are listed on the left-hand side of Table 25.2.

Independent (non-bound) pronouns distinguish three persons and two numbers and consist of the generic demonstrative element *maʔ* combined with a pronominal element proper. The use of *-niʔ* with plural pronouns as opposed to *-maʔ* (i.e. *kan-niʔ*, *nan-niʔ*, 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 (*kaymaʔ-lee kaymaʔ* ‘myself’, *kanmaʔ-lee kanmaʔ* ‘ourselves/each other’).

#### 3.1.2 Demonstratives

Nominals may be modified by demonstrative elements. The generic demonstrative element is *maʔ*. 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 (forthcoming) 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. 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 25.2. These prefixes express person and number of the possessor and are identical in form to the verbal subject agreement 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).

#### 3.1.4 Case and spatio-temporal relationships

With the exception of absolutes, which are unmarked, nominal phrases bear one of a set of clitic case particles. Subjects of most transitive verbs bear the ergative case clitic *=niʔ*.

TABLE 25.2 PRONOMINALS

	<i>Independent</i>		<i>Possessive prefixes</i>	
	<i>sg</i>	<i>pl</i>	<i>sg</i>	<i>pl</i>
1	key-maʔ	kan-maʔ	ka-	ka-n-
2	naŋ-maʔ	nan-maʔ	na-	na-n-
3	?a-maʔ	?an-maʔ	?a-	?a-n-

The locative case clitic =*ʔaʔ* marks static locations, locations towards which a figure moves, or temporal location. The instrumental-ablative case clitic =*ʔin* 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-ʔin*, and standards in equative constructions are marked by =*thluk-ʔin*, ('be.equal-instrumental/ablative'). In subordinate clauses, the ergative, locative, and instrumental-ablative case particles have an optional allomorph, =*ʔii*.

Most spatial notions are encoded through the use of abstract nouns which refer to locations, e.g. *tshuŋ* 'area inside', *tsuŋ* 'area on top', *taŋ* '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. *ʔin tshuŋ*=*ʔaʔ* 'inside the house', *ʔin taŋ*=*ʔaʔ* 'underneath the house', etc.).

A few elements are categorially intermediate between true relational nouns and case particles in that they may occur with or without oblique case markers. The element *tiaŋ* occurs with nouns which express the extent (spatial or temporal) to which an action occurs. *tshuŋ*, 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. *koŋ* 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), =*hmaa* (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, but their semantic effect is quite complex. Like the corresponding 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.

### 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', *faŋ-* 'units of money', container classifiers (e.g. *kheeŋ-* 'plate', *dur-* 'small container'), and group classifiers (e.g. *buu-* 'animal group', *tuaʔ-* '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 particles.

#### 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 character of 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 strategy occur if the verb is transitive (2). However, there are also notionally transitive clauses in which form 1 of the verb is used (3).

- (1) *paalaw* *?a-thii*  
 Paalaw 3SS-die<sub>1</sub>  
 'Paalaw died.'
- (2) *paalaw=ni?* *thil* (*khaa*) *?a-ba?*  
 Paalaw=ERG clothes DEIC 3SS-hang.up<sub>2</sub>  
 'Paalaw hung up the clothes.'
- (3) *paalaw* (*khaa*) *thil* *?a-bat*  
 Paalaw DEIC clothes 3SS-hang.up<sub>1</sub>  
 'Paalaw hangs up/hung up the 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 Falam Lai described by Osborne, 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 absolutive 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 omissible, 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 3SS-hang.up<sub>1</sub>-INTERR  
 'Did Paalaw hang up the clothes?'

- (5) *paalaw(=niʔ) thil ʔa-bat-law*  
 paalaw(=ERG) clothes 3SS-hang.up<sub>1</sub>-NEG  
 ‘Paalaw did not hang up the clothes.’

On the other hand, subordinate clauses frequently require the form 2 grade, even if the subordinate clause involves the negative marker (6):

- (6) *ka-paa = niʔ tsa ʔuk ʔa-ha ʔw-law tik = ʔaʔ ka-nuu = niʔ*  
 1SS-father=ERG book 3SS-need<sub>2</sub>-NEG time=LOC 1SS-mother=ERG  
*ʔa-zuar*  
 3SS-sell<sub>2</sub>  
 ‘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, and sometimes one suffix, which exhibit a nominative-accusative alignment. Table 25.3 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 *O* 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 *O* in the first person are the same as the markers for *A* and *S*. Third person *O* is zero-marked, but in the third plural, a postverbal particle *-hnaa* indicates plurality of the object. Second person *O* agreement exhibits allomorphy between *ʔin-*, which occurs after a consonant-final (i.e. plural) *A/S* marker, and *ni-*, 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-O* 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 *O* 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 25.4.

### 3.2.3 Directionals

Verbal complexes may contain one of a class of directional markers occurring between the subject and object agreement markers, though reportedly the semantics of a number of these

TABLE 25.3 VERBAL AGREEMENT MARKERS

	<i>A/S</i>	<i>O</i>	Reflexive object <i>A<sub>i</sub> O<sub>i</sub></i>
1s	ka-	-ka-	-a-
2s	na-	-ni-~ʔin-	-a-
3s	ʔa-	-∅-	-a-
1p	ka-n-	-ka-n-	-ʔii-
2p	na-n-	-ni- ... -hnaa~ʔin- ... -hnaa	-ʔii-
3p	ʔa-n-	-∅- ... -hnaa	-ʔii-

TABLE 25.4 JUSSIVE AGREEMENT AND NEGATION

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
cohortative			
1 affirmative	-niŋ	-ʔu-siʔ	-hnaa-ʔu-siʔ
negative	-hlaʔ-niŋ	-hlaʔ-ʔu-siʔ	-hnaa-hlaʔ-ʔu-siʔ
imperative			
2 affirmative	-∅		-ʔu-∅
negative	-hlaʔ-∅		-hlaʔ-ʔuʔ-∅
exhortative			
3 affirmative	-seʔ		-hnaa-seʔ
negative	-hlaʔ-seʔ		-hnaa-hlaʔ-seʔ

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 tiaʔ ʔa-tii*  
 tiger-COLL 1SS-DIREC-shoot<sub>1</sub>-PL.OBJ-FUT QUOT 3SS-say<sub>i</sub>  
 ‘“... 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. *ruŋ-* marks motion from a point above the speaker towards the speaker. *huŋ-* indicates motion upwards towards the speaker and *vuŋ-* motion downwards away from the speaker, respectively. At this point, however, the latter two particles are used pretty much interchangeably.

### 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 morpho-syntax, 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~khoʔ*. There are also modal elements which are probably best regarded as independent verbs taking a bare verb complement which themselves bear agreement morphology, usually to the exclusion of the complement verb, such as the desiderative *duʔ*, the potential *thiam*, the permissive causative *sian~siaŋ*, and obligative *haaw*. 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), *-tsaŋ* (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 (*-deeq* (*maaq*)); and a number focus on the event's continuity: continuative involving effort (*-leen*), continuous but ineffective activity (*-seek*), neutral continuous (*-peŋ*), and negative neutral continuous/superfluous (*-hley*). Temporary activities are marked by *-taa* and *-tshuŋ*. Other categories include habitual (*-toon*), perseverative (*-riʔ*), experiential perfect (*-bal*), instantaneous (*-tsoʔl* – the action occurs instantaneously), instantaneous unexpected (*-duak*), iterative (*-leŋmaq*), permanent (*-beʔ*), exhaustive (*-diʔ*), and for older speakers, also *-thluu* – all of an absolutive entity is effected by the action), repetitive (*-ṭhaan* – the action is performed again; *-hoy* – the action is unfortunately performed again), additional (*-vee* – another subject performs the action), associative (*-tʃi* – the action is performed jointly by a plural subject), accidental (*-sual*), and unpremeditated (*-tshom*).

### 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 Evidential and subjective evaluation markers

There are a few markers which indicate the speaker's evaluation of the accuracy of the proposition, or an emotional response of the speaker to 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 *-tuŋ* 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 post-verbal elements clearly related to demonstratives/discourse deictics (*hiʔ*, *khiʔ*, *tsuʔ*, *khaʔ*) which also conveys quite intricate spatial and evidential information. *ʔaay* indicates regret on the part of speaker or subject and *ʔee* 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) *ʔuytsaw ʔa-baw-duapmap/ʔa-baw-diapmap*  
 dog 3SS-bark<sub>1</sub>-IDEO/3SS-bark-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 ?a-nuu=naak?in ?a-saaŋ-de ?w*  
 Paalaw 3S.POSS-mother=STAND 3SS-tall<sub>1</sub>-COMP  
 'Paalaw is taller than his mother.'
- (10) *paalaw ?a-saaŋ-biik*  
 Paalaw 3SS-tall<sub>1</sub>-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 above in Table 25.4, negation is marked instead by *-hla?*.

### 3.2.4.8 Mood

Indicative mood is morphologically unmarked. We saw above (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 *-hŋaa*, 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=?a? khan ka-rak-tlum-vee-kaw-hŋaa-law*  
 1S.POSS-mother vicinity=LOC DEIC 1SS-PAST-return<sub>2</sub>-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', *sii-vaay* 'medicine-wander=potion', *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, also functions as a deverbal nominalizer which productively creates locative, instrumental, and action nominals: *sam-me ?-naak* 'hair-cut-nomlZR=barbershop', *thil-tsook-naak* 'thing-buy-nomlZR=shop', *?ay-din-naak* 'eat-drink-nomlZR=restaurant', *hmu ?-naak* 'see-nomlZR=seeing', *peek-naak* 'give-nomlZR=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', *hni ?-naak* 'second', *pa-hli* 'four', *hli-naak* 'fourth'.

### 3.3.4 Valence-affecting morphology

Besides the effect that ablaut has 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. *hmu?sak* ‘to show’). Next, causatives which involve devoicing (in the case of sonorants) or aspiration (in the case of stops) of the initial consonant of the root (e.g. *thlaak* ‘to fall’, *thlaak* ‘to fell’) reflect the widely recognized Tibeto-Burman *\*s-* causative prefix. *\*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.

The newer system of causativization, on the other hand, effected by the addition of a particle *-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 addition of one of seven postverbal elements to the verbal complex, depending on the semantics of the applicative object. (12) to (18) give examples of the applicative morphology.

- (12) benefactive/malefactive applicative  
 law *?a-ka-thlo ?-piak*  
 field 3SS-1SO-hoe<sub>2</sub>-BEN.APP  
 ‘He hoed the field for me/in my place.’
- (13) additional benefactive applicative  
 law *?a-ka-thlo ?-tse ?m*  
 field 3SS-1SO-hoe<sub>2</sub>-ADD.BEN.APP  
 ‘He hoed the field for my benefit (in addition to his own benefit).’
- (14) comitative applicative  
 law *?a-ka-thlo ?-pii*  
 field 3SS-1SO-hoe<sub>2</sub>-COM.APP  
 ‘He hoed the field along with me.’
- (15) allative/malefactive applicative  
 law *?a-ka-thlo ?-hno ?*  
 field 3SS-1SO-hoe<sub>2</sub>-MAL.APP  
 ‘He hoed the field to my deriment.’
- (16) prioritive applicative  
 law *?a-ka-thlo ?-ka ?n*  
 field 3SS-1SO-hoe<sub>2</sub>-PRIOR.APP  
 ‘He hoed the field ahead of/before me.’

- (17) relinquitive applicative  
*law* *ʔa-ka-thlo ʔ-taak*  
 field 3SS-ISO-hoe<sub>2</sub>-RELINQ.APP  
 ‘He left me and hoed the field.’
- (18) instrumental applicative  
*tuuhmuy law ʔa-thlo ʔ-naak*  
 hoe field 3SS-hoe<sub>2</sub>-INST APP  
 ‘He hoed the field with a hoe.’

Only in the case of *-hnoʔ*, which is generally malefactive, but which with some roots expressing motion (e.g. *kal* ‘go’ and *kaay* ‘climb’) may have an allative meaning, 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 1999 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.

## 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

A nominal phrase consists minimally of 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-absolute noun phrases bear a 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. This is less often the case for other obliquely marked nouns, however.

## 4.2 Clausal syntax

### 4.2.1 Word order

Arguments are often not instantiated by full NPs, their instantiation being assumed by a pronominal interpretation of verbal agreement marking. When full NPs are present, the basic word order is subject-object-verb, with variable placement of subject and object depending on their case marking and semantic factors. There is a left dislocation position, which attracts heavy constituents and is also used extensively for topicalization. In addition, constituents may sometimes appear in a postverbal afterthought position.

### 4.2.2 Grammatical relations

Marking of the basic S, A and O 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/absolutive alignment, and verbal ablaut achieves the effect of an antipassive construction, in which both arguments of a transitive verb are absolutive, 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 fundamentally 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-luŋ*            *kan-rook*  
 1S.POSS-heart    1PS-break.down<sub>1</sub>  
 '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.

### 4.2.4 Subordination

Subordination involves 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 the object of the oblique case particles =*ʔaʔ* and =*ʔin* to yield adverbial subordinate clauses expressing the circumstance under which an event described in the main clause occurs, as in (20).

- (20) *ka-ʔiin = ʔaʔ*                      (*khan*) *ʔa-kal*  
 1SS-come.home<sub>2</sub>=LOC    DEIC    3SS-go<sub>1</sub>  
 ‘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-ʔiin = ʔaʔ*                      *tsun* *ʔa-kal-laay*  
 1SS-come.home<sub>2</sub>=LOC    DEIC    3SS-go<sub>1</sub>-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 = ʔaʔ* ‘back (=after) v’, *v hlaan = ʔaʔ* ‘front (=before) v’, *v tshuŋ = ʔaʔ* ‘inside (=while) v’, *v laay = ʔaʔ* ‘towards (=when about to) v’, *v tsaʔ = ʔaʔ* ‘sake (=because) v’). In other cases, the relational elements only occur in these constructions (e.g. *v tik = ʔaʔ* ‘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 = ʔin* and purposive clauses are marked by *-diŋ = ʔaʔ*. Besides the *v tsaʔ = ʔaʔ* construction, 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 = ʔin* and *-paʔ = ʔin* 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-tsaʔ-diŋ = ʔaʔ*.

#### 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) Intransitive subject target.  
*in = ʔii*                      *ʔa-it-mii*                      *lawthlawpaa*    *ka-hmuʔ*  
 house=LOC    3SS-sleep<sub>1</sub>-REL    farmer                      1SS-see<sub>2</sub>  
 ‘I saw the farmer who slept in the house.’

- (23) Transitive subject target.  
*thil ʔa-bat-tuu lawthlawpaa ka-hmuʔ*  
 thing 3SS-hang<sub>1</sub>-REL farmer 1SS-see<sub>2</sub>  
 'I saw the farmer who hung up the clothes (*lit.* the things).'
- (24) Transitive subject target.  
*thil ʔa-bat-mii lawthlawpaa ka-hmuʔ*  
 thing 3SS-hang<sub>1</sub>-REL farmer 1SS-see<sub>2</sub>  
 'I saw the farmer who hung up the clothes.'
- (25) Transitive object target.  
*lawthlawpaa = niʔ ʔa-baʔ-mii thil ka-hmuʔ*  
 farmer = ERG 3SS-hang<sub>2</sub>-REL thing 1SS-see<sub>2</sub>  
 'I saw the clothes the farmer hung up.'
- (26) Locative target.  
*lawthlawpaa = niʔ thil ʔa-baʔ-naak thiŋkuŋ ka-hmuʔ*  
 farmer = ERG thing 3SS-hang<sub>2</sub>-REL tree 1SS-see<sub>2</sub>  
 'I saw the tree the farmer hung the clothes up on.'
- (27) Instrumental target.  
*lawthlawpaa = niʔ ŋa ʔa-tan-naak naam ka-hmuʔ*  
 farmer = ERG fish 3SS-cut<sub>2</sub>-REL knife 1SS-see<sub>2</sub>  
 '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 25.5 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 intransitive subject, transitive subject, and object targets.

- (28) *ʔin = ʔii lawthlawpaa ʔa-it-mii ka-hmuʔ*  
 house = LOC farmer 3SS-sleep<sub>1</sub>-REL 1SS-see<sub>2</sub>  
 'I saw the farmer who slept in the house.'
- (29) *nikum = ʔii lawthlawpaa thil ʔa-bat-mii ka-hmuʔ*  
 last.year = LOC farmer thing 3SS-hang<sub>1</sub>-REL 1SS-see<sub>2</sub>  
 'I saw the farmer who hung up the clothes last year.'
- (30) *nikum = ʔii lawthlawpaa = niʔ thil ʔa-baʔ-mii ka-hmuʔ*  
 last.year = LOC farmer = ERG thing 3SS-hang<sub>2</sub>-REL 1SS-see<sub>2</sub>  
 'I saw the clothes the farmer hung up last year.'

**TABLE 25.5** PRIMARY RELATIVIZATION STRATEGIES

Role of target	Intransitive subject	Agent	Patient	Instrument, locative
Ablaut grade	form 1	form 1	form 2	form 2
Relativizer	-mii	-mii/-tuu	-mii	-naak

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 =*ʔii* rather than =*ʔaʔ*. As discussed earlier in the section on case marking, =*ʔii*, 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, because if it is, 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 was 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 = niʔ ʔa-tsoo ʔa-zuar-mii khaa ʔa-thaʔy*  
 Paalaw = ERG 3S.POSS-cow 3SS-sell<sub>2</sub>-COMP DEIC 3SS-know<sub>2</sub>  
 ‘He knows that Paalaw sold his cow.’

There is no distinction between direct and indirect speech, and verbs of speaking and other verbs of cognition require a (quotative) complementizer based on the proverb *-tii* ‘do/say’, *tiaʔ*, as in (32).

- (32) *paalaw = niʔ ʔa-tsoo ʔa-zuar tiaʔ ʔa-tii/ʔa-zuʔm*  
 Paalaw = ERG 3S.POSS-cow 3SS-sell<sub>2</sub> QUOT 3SS-say<sub>1</sub>/3SS-believe<sub>2</sub>  
 ‘He said/believes Paalaw sold his cow.’

Finally, *diŋ = ʔaʔ*, which acts as a purposive clause marker in other contexts, acts as a complementizer:

- (33) *lawthlawpaa = niʔ ʔa-tsoo ʔa-zuar diŋ = ʔaʔ paalaw = niʔ*  
 farmer = ERG 3S.POSS-cow 3SS-sell<sub>2</sub> COMP = LOC Paalaw = ERG  
*ʔa-leem/ʔa-nool*  
 3SS-persuade<sub>2</sub>/3SS-request<sub>2</sub>  
 ‘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, *ʔum*.

### 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 *-tuaʔ* 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 thoɔŋsew ʔa-sii=maa*  
 2S-name Thawng Ceu 3S-be<sub>1</sub>=INTERR  
 'Is your name Thawng Ceu?'

Content questions are characterized by (optionally left-dislocated) dedicated question words (*zay* 'what', *ʔahaw* 'who', *khoy/khoykaa=ʔaʔ* 'where', *zay tii=ʔin* 'how', *zay ruaŋ=ʔaʔ* 'for what reason', etc.) In most cases, the particle *=daʔ* is also added to the question word, as in (35).

- (35) *faalaam=ʔaʔ ʔahaw=daʔ na-thaʔy*  
 Falam=LOC who=QUES 2SS-know<sub>2</sub>  
 'Who do you know in Falam?'

## REFERENCES

- Barnes, Jonathan (1998) 'Tsuu khaa tii hla?: deixis, demonstratives and discourse particles in Lai Chin', *LTBA* 21.1: 53–86.
- Bedell, George (1995) 'Agreement in Lai', papers from the Fifth Annual meeting of the Southeast Asian Linguistics Society, Tempe: Program for Southeast Asian Studies, Arizona State University, 21–32.
- (1996a) *Clitic Climbing in Lai. Pan-Asiatic Linguistics*, proceedings of the Fourth International Symposium on Languages and Linguistics, Vol. 2, Nakhon Pathom: Mahidol University, 405–15.
- (1996b) 'Passives and clefts in Lai', to appear in papers from the Sixth Annual Meeting of the Southeast Asian Linguistics Society.
- (1997a) 'Causatives and clause union in Lai (Chin)', *Mon-Khmer Studies* 27: 219–32.
- (1997b) 'Benefactives and clause union in Lai', to appear in papers from the Seventh Annual Meeting of the Southeast Asian Linguistics Society.
- (1998a) 'Nominal auxiliaries in Lai', to appear in papers from the Eighth Annual Meeting of the Southeast Asian Linguistics Society.
- (1998b) Describing and explaining Lai, ICU ms.
- (1999) 'Word combination in Lai', to appear in papers from the Ninth Annual Meeting of the Southeast Asian Linguistics Society.

- (forthcoming) 'The syntax of deixis in Lai', *LTBA*.
- (2000) 'Postpositions and relational nouns in Lai', paper presented at Pan-Asiatic Linguistics, Ho Chi Minh City.
- Bedell, George and Kenneth VanBik (2000) 'Lexical and syntactic causatives in Lai', to appear in papers from the Tenth Annual Meeting of the Southeast Asian Linguistics Society.
- Bickel, Balthasar (2000) 'On the syntax of agreement in Tibeto-Burman', *Studies in Language* 24.3: 583–609.
- Chhangte, Lalnunthangi (1993) 'Mizo syntax', unpublished University of Oregon PhD dissertation.
- Haye-Neave, D.R. (1948) *Lai Chin grammar and dictionary*, Rangoon: Superintendent of Government Printing and Stationery, Burma.
- Kathol, Andreas (2000) 'The morphosyntax of Lai relative clauses', in Ronnie Cann, Claire Grover, and Phillip Miller (eds) *A Collection of Papers on Head-Driven Phrase Structure Grammar*, Stanford: Stanford University Press.
- Kathol, Andreas and Kenneth VanBik (1999) 'Morphology – syntax interface in Lai relative clauses', Pius Tamanji, Masako Hirotani and Nancy Hall (eds) *NELS* 29: 427–41.
- (2000) 'Lexical constraints and constructional overrides: on the syntax of verbal stem alternations in Lai', unpublished manuscript, UC Berkeley.
- (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* 20.2.
- 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: The University of Illinois Press.
- (1996) 'Relative clauses in Lai Chin, with special reference to verb stem alternation and the extension of control theory', *LTBA* 19.1: 43–58.
- Lehman, F.K. and VanBik Kenneth (1997) 'Notes on Lai Chin personal pronouns and overt case marking', *Studies in the Linguistic Sciences* 27.2: 81–86.
- Macnabb, D.J.C. (1891) *Hand-book of the Haka or Baungshe Dialect of the Chin Language*, Rangoon.
- Melnik, Nurit (1997a) 'The sound system of Lai' *LTBA* 20.2: 9–19.
- (1997b) 'Verbal alternations in Lai', *LTBA* 20.2: 163–72.
- Newland, A.G.E. (1897) *A Practical hand-book of the Language of the Lais*, Rangoon.
- Olawsky, Knut J. and VanBik, Kenneth (2000) 'Introduction to Lai tonology', unpublished manuscript, UC Berkeley.
- Osburne, Andrea Gail (1975) 'A transformational analysis of tone in the verb system of Zahao (Laizo) Chin', unpublished PhD dissertation, Cornell University.
- Patent, Jason D. (1998) 'A willy-nilly look at Lai ideophones', *LTBA* 21.1: 155–200.
- Peterson, David A. (1998) 'The morphosyntax of transitivization in Lai (Haka Chin)', *LTBA* 21.1: 87–153.
- (1999) 'Discourse-functional, historical, and typological aspects of applicative constructions', unpublished PhD dissertation, UC Berkeley ms.
- Peterson, David A. and VanBik, Kenneth (forthcoming) *A Reference Grammar of Lai*, Max Planck Institute for Evolutionary Anthropology and UC Berkeley.
- Plauché, Madelaine C., de Azcona, Rosemary Beam Roengpitya, Rungpat and Weigel, William F. (1998) 'Glottalized Sonorants: A Phonetic Universal?' *BLS* 24: 381–90.
- Reichle, Verena (1981) *Bawm Language and Lore: Tibeto-Burman Area*, Bern: Peter Lang.
- Roengpitya, Rungpat (1996) 'Classifiers in Lai', unpublished manuscript, UC Berkeley.
- (1997) 'Glottal stop and glottalization in Lai (connected speech)', *LTBA* 20.2: 21–56.
- VanBik, David (1986) *English–Chin (Haka) Dictionary*, Haka.
- VanBik, Kenneth (1997a) 'Lai adverb classification', unpublished manuscript, UC Berkeley.
- (1997b) 'Relative clause in Lai', unpublished manuscript, UC Berkeley.
- (1998) 'Lai psycho-collocation', *LTBA* 21.1: 201–33.
- (2000) 'Three types of causative constructions in Lai', unpublished manuscript, UC Berkeley.



- (forthcoming) ‘Causatives in Lai’, to appear in Graham Thurgood (ed.) papers from the Eighth Annual Meeting of the Southeast Asian Linguistic Society (SEALS VIII), Tempe, Arizona: Department for Southeast Asian Studies, Arizona State University.
- VanBik, Kenneth and VanBik, David (2000) Lai–English dictionary, unpublished manuscript, UC Berkeley.
- Yamashita Smith, Tomoko (1998) ‘The middle voice in Lai’, *LTBA* 21.1: 1–52.

### 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 last century by British military personnel (Macnabb 1891 and 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 all produced independent studies on a wide variety of topics (see references); although they so far do not constitute a complete reference grammar, these studies nonetheless cover a wider range of topics than is covered in many sources available for any other language in this part of Tibeto-Burman. Besides the extensive literature which has appeared on Lai in recent years, a full reference grammar and collection of narrative texts is under preparation by the author and Kenneth VanBik. 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 complete it in the near future. Lorenz Löffler has also been working independently on dictionary materials with a collaborator from Hakha.

# MEITHEI

*Shobhana L. Chelliah*

## 1 LOCATION AND CLASSIFICATION

Meithei, also known as *Manipuri* and *Meitheirón*, 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, 1813 square kilometres of which are level country approximately 750 metres above sea level. This level area is populated mainly by the Meithei and 120,000 Muslims who are the progeny of the intermarriage of Muslim traders and labourers with Meithei 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 Meithei, although the number of actual speakers is higher since Meithei is used as a *lingua franca* in the state (*Encyclopedia Britannica Online* 1999). English and Meithei are the state languages.

In Matisoff's (1991) heuristic model, consisting of seven geographical groups, Meithei 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 Meithei 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 Meithei languages do not fit readily into this or other sub-groups of the area. DeLancey (1987: 800) postulates a distinct Mikir-Meithei sub-branch. Pending the collection of more data on other languages in the group, however, the genetic position of Meithei within Tibeto-Burman remains uncertain.

The description in this entry is of the Imphal dialect of Meithei, which is considered to be the standard. Other documented dialects include Sekmai (H.S. Singh 1988) and Pheyeng (P.R. 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 Meithei is given in Table 26.1.

The aspirated and unaspirated voiced stop and affricate series can be attributed to large scale borrowing of Indo-Aryan words. The aspirated affricate /tʃ<sup>h</sup>/ is phonetically realized as [s], [s<sup>h</sup>], [ʃ] or [ʃ<sup>h</sup>] in native words. Additionally, /tʃ/ is realized as [ts] before /i/, as in [tsim] '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 Meithei is given in Table 26.2.

**TABLE 26.1** CHART OF CONSONANT PHONEMES

	<i>Labial</i>	<i>Alveolar</i>	<i>Palatal</i>	<i>Velar</i>	<i>Laryngeal</i>
Stops	p	t		k	
	p <sup>h</sup>	t <sup>h</sup>		k <sup>h</sup>	
	b	d		g	
	b <sup>h</sup>	d <sup>h</sup>		g <sup>h</sup>	
Affricates			č č <sup>h</sup> j j <sup>h</sup>		
Fricatives		s			h
Nasals	m	n		ŋ	
Lateral/Flap		l			
Trill		r			
Semivowels	w		y		

**TABLE 26.2** CHART OF VOWEL PHONEMES

	<i>Front</i>	<i>Central</i>	<i>Back</i>
high	i		u
mid	e	ə	o
low		a	

There are no indigenous words beginning with /a/. The vowel /e/ occurs in initial position in a few words.

## 2.1 Syllable structure

The Meithei 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, ph, t, th, č, čh, k, kh, 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 [ʔə-ʔibə] ‘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 unaspirated stop, fricative, or voiceless aspirated stop and /l/, /w/ or /y/. Onsets of borrowed words may consist of voiced unaspirated 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, and l. There are no complex codas. This restriction on codas is upheld for borrowed words also.

## 2.2 Tone

Meithei 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.

Diphthongization and gemination rules work to insure that syllables in Meithei must have an onset. When a root is concatenated with a suffix that begins with a vowel, a diphthong is formed, e.g. *u-* ‘see’ + *-í* ‘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 *puʔo* ‘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. *thəm-* ‘keep’ + *-u* ‘imperative’ results in *thəmmu* ‘Keep!’.

In sequences of identical oral stops the second stop may be dissimilated. Thus in *čət-* ‘go’ + *-u* ‘imperative’, gemination results in *čəttu* which is followed by the dissimilation of the second consonant to *l*, resulting in *čətlu* ‘Go!’.

## 3 MORPHOLOGY

A noun 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* ‘male’ can indicate gender, e.g. *núpi* ‘woman’, ‘wife’ and *núpa* ‘man’, ‘husband’ where the stem *nú* means ‘human’. The feminine and masculine suffixes also appear in traditional proper names which are based on adjectives, e.g. the youngest male child is named *Tomba* and the youngest female child is named *Tombi*, from *tóm* ‘last’ and one of the gender suffixes. Plurality is indicated with *-siŋ* (e.g. *əŋáŋsiŋ* ‘children’), but this suffix cannot occur with pronouns or proper names.

The sole nominal inflectional category is semantic role marking through cliticization. An example is the inflectional paradigm of *nupá* ‘male child’ which is as follows: *nupanə* ‘by the boy’, where *-nə* is agentive/instrumental; *nupabu* ‘the boy’ (patient), where *-pu* is accusative; *nupadə* ‘at/to the boy’, where *-tə* is locative; *nupadəgi* ‘from the boy’, where *-təgi* is ablative; *nupagi* ‘of the boy’, where *-ki* is genitive; and *nupagə* ‘with the boy’, where *-kə* is associative.

Non-category changing derivational prefixes are *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, *məpa* ‘his grandfather’ where *pa* is ‘grandfather’, *imít* ‘my eye’ where *mít* is ‘eye’ and *nəkhóŋ* ‘your foot’ where *khóŋ* is ‘foot’. The scope of inalienable possession does not include objects or dwellings that are momentarily possessed, but is metaphorically extended to include the ancestral home, e.g. *məyím* ‘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)	<i>phúgayəmlə</i>				
	phú	-khay	-ləm	-lə	-e
	beat	-totally affect	-indirect evidence	-perfect	-assertive
	Root	1st level	2nd level	3rd level	Inflection
		derivation	derivation	derivation	
	'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 below 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 derivational suffixes are grammaticalized roots. Examples of the first level derivational suffixes are *-khay* 'totally affect' from *kháy-* 'cut with a knife'; *-thət* 'partially affect' from *thət-* '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 *pí-* 'give'; *-čə* 'V for sake of self' from *sá* 'body'; *-məŋ* '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'. Meithei verb morphology does not mark number, person, gender or pronominal agreement between the verb and its arguments.

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 16).

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', *khóŋləm* 'footpath' from *khóŋ* 'foot' and *ləm* 'way', and *mítna* 'organs' from *mít* 'eye' and *na* 'nose'. There are no right- or left-headed verbal compounds; the second stem in left-headed compounds has grammaticalized to form the first and second level derivational suffixes described above. The few verbal compounds that exist are exocentric, e.g. *čáthək* 'dine' from *čá-* 'eat' and *-thək* 'drink'.

Lexical collocations are formed through a variety of processes of reduplication 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. *čáiphəm lakphəm* ‘place of much activity’, literally ‘go place come place’ or *həway čəjway* ‘things to eat’, literally ‘lentils and rice, and such’.

#### 4 MAJOR LEXICAL CATEGORIES

Nouns and verbs can be distinguished on formal grounds: nouns are free roots whereas verbs are bound roots 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 *əčəwbə* ‘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 *mətúŋ* ‘behind’ from *túŋ-* ‘be in back’. Temporal adverbs are frozen compounds: *ŋəsi* ‘today’, *ŋəraŋ* ‘yesterday’ and *həyəŋ* ‘tomorrow’ are most likely bimorphemic although the meaning of the individual roots is not clear.

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 *həw-* ‘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 *əy* ‘I’, *nəŋ* ‘you’ and *má* ‘he/she’. Possessive pronouns are formed through the suffixation of *-ki* ‘genitive’ to these personal pronouns: *əygi/nəŋgi/mági yum* ‘my/your/his or her house’. The plural is indicated by *khoj* ‘this and others like this’. The plural forms are not built on the regular personal pronouns, but on the pronominal possessor prefixes, e.g. *əykhoy* 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 strategy for creating relative clauses in Meithei is to place the relativized noun 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 contrastive *-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 Meithei 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; *khitəŋ* ‘ever so little’, ‘a particle’ (composed of *khit* ‘a little’ and *təŋ* ‘exclusive’) of some tangible material; and *khəjiktə* which indicates a short amount of time.

Enclitics in Meithei 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’, *-tə* ‘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

Evidential values are signalled through derivational verb morphology, in particular the indirect evidence, aspect, and direction markers. Additionally, as illustrated in (2), Bhat and Ningomba (1997: 270) note that there is an implication with *-həw* ‘inceptive’ that the speaker is a witness to the initiation of an action.

- (2) *túren*     *pahəwwí*  
 túlen     pa-həw-í  
 river     overflow-start-declarative  
 river     began to overflow  
 ‘The river began overflowing.’

Evidential values are also indicated through the attitude markers. 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) *Kakčìŋdənine*  
 Kakčìŋ-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 *háynə* 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 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) | <i>məsi</i>                                | <i>phúrəbəjatni</i>             |
|     | mə-si                                      | phú-ləbə-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

There is no evidence in Meithei for a verb phrase constituent; thus, the Meithei clause minimally consists of a verb and the arguments (i.e. noun phrases) the verb subcategorizes for. As reflected in phrase structure rule (5), the asterisk which follows the noun phrase indicates, following the convention used in Hale (1983), that the verb may occur with any number of noun phrases.

- (5) sentence → noun phrase\* verb

In the most common word order, the verb occurs sentence finally, with the agent or actor appearing sentence initially. The order of oblique arguments is free and determined by pragmatics; constituents may occur after the verb as afterthoughts and focused constituents occur sentence initially.

The maximum number of noun phrases that may occur with a verb is restricted by the subcategorization frame of that verb. Since Meithei allows for the omission of arguments, sometimes a verb may occur without any accompanying noun phrases.

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 Meithei.

### 6.1 Semantic role marking and information packaging

The grammatical status of the arguments that a predicate subcategorizes for is indicated through semantic role markers. A verb may subcategorize for an agent (marked by *-nə*), actor (marked by *-ə*), patient (marked by *-pu*), experiencer/goal (marked by *-tə*), path (marked by *-ə*), and theme (marked by *-ə*). The agent, which occurs only with verbs that are marked by the causative suffix *-hən*, is defined as the volitional and intentional instigator of an action. The actor, which occurs with verbs that are not marked by the causative, is the effector of action. The patient is an argument that is in or changes states; the theme is an argument that is in or changes location. The experiencer and goal arguments, which occur with verbs of cognition and perception, are those towards which an action is directed.

In the default case the correct interpretation of the status of arguments in a sentence can be read off of semantic role markers and the observance of real-world semantics. Thus in (6) even though neither *əy* or *part* are marked, it is understood that *əy* is the actor.

- |     |                                 |             |  |
|-----|---------------------------------|-------------|--|
| (6) | <i>əy</i>                       | <i>part</i> | <i>ləyrukhini</i>                        |
|     | <i>əy</i>                       | part        | ləy-lu-khi-ni                            |
|     | I                               | spare parts | buy-action away from speech-still-copula |
|     | I                               | spare parts | will buy there                           |
|     | ‘I will buy spare parts there.’ |             |  |



Morphologically encoded grammatical information is often obscured through the overlay of a system of pragmatic marking which may delete or replace the semantic role marker with one of these enclitics: *-tə* 'exclusive', *-ti* 'delimitative', *-tu* 'distal', *-nə* 'contrastive', *-pu* 'adversative', *-si* 'proximate' or *-su* 'inclusive'. Pragmatic information can also be signalled by adding one of these enclitics to a semantic role marker and/or manipulating word order so that topics occur sentence initially.

In some instances, the pragmatic system makes recovery of grammatical relations difficult so that sentences may often have more than one interpretation. In these cases the larger discourse context must be used to recover the intended meaning. Examples are (7) and (8).

- (7) *əydi*                      *Ramnə*                      *phúniŋŋí*  
 əy-ti                      Ram-nə                      phú-niŋ-í  
 I-delimitative      Ram-contrastive      beat-wish-declarative  
 I                      Ram                      wish to beat  
 '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.'
- (8) *əŋáŋdi*                      *Tombəsinə*                      *phúy*  
 əŋáŋ-ti                      Tombə-si-nə                      phú-í  
 child-delimitative      Tomba-determiner-contrastive      beat-declarative  
 child                      Tomba                      beat  
 'This Tomba (out of all the others) beats children.' or  
 'This child beats this Tomba (and no one else).'

The contrastive *-nə* and the agentive *-nə* are distinguished on the basis of distribution; while *-nə* agentive obligatorily marks the subject of a causative verb, *-nə* contrastive can mark any sort of argument, e.g. (9) where it marks the patient.

- (9) *əybunə*                      *Ramnə*                      *nuŋsirəbədi*  
 əy-pu-nə                      Ram-nə                      nuŋsi-lə-pə-ti  
 I-patient-contrastive      Ram-contrastive      love-prospective-nominalizer-delimitative  
 I                      Ram                      if love  
*phəgədəwni*  
 phə-kə-təw-ni  
 good-potential-obligation-copula  
 would be good  
 'If Ram (and no one else) loved me (and no one else), 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 (10). Verbs do not form interrogatives unless they are first nominalized as in (11) or are in the potential mood, i.e. the derivational suffixes *-kə* 'potential', *-loy* 'nonpotential', *-təw* 'certain future', and *-tə* 'strong possibility/obligation' act as nominalizers.

- (10) *Tomba u kəkəpə*                      *míra*  
 Tomba u      kək-pə                      mí-lə  
 Tomba u      cut-nominalizer      man-interrogative  
 Tomba tree to cut                      is it man  
 'Is Tomba the man who is a wood cutter?'

- (11) *čətʔəbra*  
 čət-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 (12).

- |      |                               |                         |                                |
|------|-------------------------------|-------------------------|--------------------------------|
| (12) | <i>učék paybə</i>             | <i>ŋəmdə</i>            | <i>ŋəmbra</i>                  |
|      | učék pay-pə                   | ŋəm-tə-e                | ŋəm-pə-lə                      |
|      | bird fly-nominalizer          | able-negative-assertive | able-nominalizer-interrogative |
|      | bird to fly                   | not able                | are they able                  |
|      | ‘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 verbal 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 (13), a question word can occur with an inflected verb. Second, as in (14), the question word occurs in a sentence that ends with *-no* ‘inquisitive’. Third, *-no* is affixed directly on the question word as in (15).

- |      |                                      |                |
|------|--------------------------------------|----------------|
| (13) | <i>kəna kəyθheldə</i>                | <i>čətli</i>   |
|      | kəna kəy-θel-tə                      | čət-li         |
|      | who grain-display-locative           | go-progressive |
|      | who to the market                    | going          |
|      | ‘Who goes to the market?’            |                |
| (14) | <i>kənagi yénawno</i>                |                |
|      | kəna-ki yén-naw-no                   |                |
|      | who-genitive chicken-new-inquisitive |                |
|      | whose is this chicken, tell me       |                |
|      | ‘Whose chicken is this?’             |                |
| (15) | <i>kənagino</i>                      |                |
|      | kəna-ki-no                           |                |
|      | who-genitive-inquisitive             |                |
|      | whose is it                          |                |
|      | ‘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 a noninflected verb as in (16). It is used in relative clause formation where the relativized argument occurs to the right of a nominalized verb as in (17). The determiner complement is a nominalized clause followed by the determiner *-si* ‘this’ or *-tu* ‘that’ as in (18).

The quotative complement uses forms of the quotative as the complementizer, as shown in (19).

- |      |                                     |                              |                  |                         |
|------|-------------------------------------|------------------------------|------------------|-------------------------|
| (16) | <i>ʒy</i>                           | <i>čak</i>                   | <i>čábə</i>      | <i>həwre</i>            |
|      | <i>ʒy</i>                           | <i>čak</i>                   | <i>čá-pə</i>     | <i>həw-lə-e</i>         |
|      | I                                   | cooked rice                  | eat-nominalizer  | start-perfect-assertive |
|      | I                                   | rice                         | to eat           | started                 |
|      | 'I have started eating rice.'       |                              |                  |                         |
|      |                                     |                              |                  |                         |
| (17) | <i>kolom</i>                        | <i>páyrəbə</i>               | <i>nipa</i>      |                         |
|      | <i>kolom</i>                        | <i>páy-lə-pə</i>             | <i>ni-pa</i>     |                         |
|      | pen                                 | hold-perfect-nominalizer     | person-male      |                         |
|      | pen                                 | one who held                 | boy              |                         |
|      | 'the boy who held the pen'          |                              |                  |                         |
|      |                                     |                              |                  |                         |
| (18) | <i>isiŋ</i>                         | <i>thək-pə-tu</i>            |                  |                         |
|      | <i>isiŋ</i>                         | <i>thək-pə-tu</i>            |                  |                         |
|      | water                               | drink-nominalizer-determiner |                  |                         |
|      | water                               | that drinking                |                  |                         |
|      | 'from that drinking water'          |                              |                  |                         |
|      |                                     |                              |                  |                         |
| (19) | <i>məhák-nə</i>                     | <i>thoyre</i>                |                  |                         |
|      | <i>mə-hák-nə</i>                    | <i>thoy-lə-e</i>             |                  |                         |
|      | she-present participant-contrastive | win-perfect-assertive        |                  |                         |
|      | she                                 | had won                      |                  |                         |
|      |                                     |                              |                  |                         |
|      | <i>háybəsi</i>                      | <i>Tombinə</i>               | <i>khəŋŋí</i>    |                         |
|      | <i>háy-pə-si</i>                    | <i>Tombi-nə</i>              | <i>khəŋ-í</i>    |                         |
|      | say-nominalizer-determiner          | Tombi-contrastive            | know-declarative |                         |
|      | that                                | Tombi                        | knew             |                         |
|      | 'Tombi knew that she had won.'      |                              |                  |                         |

Adverbial subordinators are derived from the associative, genitive, locative, ablative, and instrumental markers, following a pattern that is common in Tibeto-Burman languages, as discussed in detail for twenty-six languages of the family by Genetti (1988). Thus *-kə* 'associative' is used to mean 'at the same time as Ving', *-ki* 'genitive' is used to mean 'for the purpose of Ving', *-tə* 'locative' is used to mean 'after Ving', *-təgi* 'ablative' is used to mean 'resulting from Ving', *-nə* 'instrumental' is used to mean 'by Ving, because of Ving'. An example is given in (20).

- |      |                             |                         |
|------|-----------------------------|-------------------------|
| (20) | <i>ʒykhoydə</i>             | <i>lakpədə</i>          |
|      | <i>ʒy-khoy-tə</i>           | <i>lak-pə-tə</i>        |
|      | I-plural-dative             | come-nominalizer-dative |
|      | to our place                | after coming            |
|      | 'after coming to our place' |                         |

Adverbial participials are suffixed to non-nominalized verbs. The participializers are morphologically complex, lexicalized units. The participial marker *-tínə* 'Ving' is composed of the determiner and instrumental markers, *-tínə* 'by Ving' is composed of the locative and instrumental markers, *-nəbə* 'for Ving' is composed of the instrumental marker and the nominalizer *-pə*, *-ləgə* 'after Ving' is composed of the perfect and associative markers, and

-*lǎbə* ‘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 (21) and (22).

- (21) *áy mági mətúŋ indúnə laʔi*  
*áy má-ki mət-túŋ in-túnə lak-í*  
 I he-genitive nominalizer-behind follow-ing come-declarative  
 I his behind following came  
 ‘I followed him.’

- (22) *ləyrəbədi*  
*ləy-lə-pə-ti*  
 be-pro 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 (23) where several adverbial clauses are strung together. The ampersand ‘&’ marks the beginning of each clause.

- (23) & *čithi ədo parubədudə* ‘on reading that letter’  
 & *senpanəbə nupa ədunə čithi ɛ purəkʔibə nupa ədudə sen píkho háybəninə*  
 ‘since the treasurer was told to give money to the person who brought the letter’  
 & *məraybəkphəbə dolaypəbədo ədudə* ‘and so to that fortunate gate keeper’  
 & *sen píkħəre* ‘(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 gate keeper, the money.’

## REFERENCES

- Benedict, P.K. (1984) ‘Proto-Sino-Tibetan interrogative \*ga(ng)~ \*ka’, *Linguistics of the Tibeto-Burman Area* 8.1: 1–10.  
 Bhat, D.N.S. and Ningomba, M.S. (1997) *Manipuri Grammar*, Munich: Lincom Europa.  
 ‘Manipur’ *Encyclopedia Britannica Online*, Internet, 28 February 1999, available at <http://www.eb.com:180/bol/topic?eu=121199&sctn=1>.  
 DeLancey, S. (1987) ‘Sino-Tibetan languages’, in B. Comrie (ed.) *The World’s Major Languages*, London: Croom Helm.  
 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.  
 Hale, K.L. (1983) ‘Warlpiri and the grammar of non-configurational languages’, *Natural Language and Linguistic Theory* 1: 5–47.  
 Grierson, G.A. (1919) *Linguistic Survey of India and the Census of 1911*, Calcutta: Superintendent Government Printing, India (reprinted 1967 by Motilal Banarsidass: Delhi, Varanasi, Patna).  
 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, W.R. (1989) 'The Kwatha dialect of Meitei', *Linguistics of the Tibeto-Burman Area* 12.2: 101–22.
- Thoudam, P.C. (1980) 'Grammatical sketch of meiteiron', unpublished 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.

### FURTHER READING

- Chelliah, S.L. (1997) *A Grammar of Meithei*, Berlin: Mouton.
- 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.
- Singh, C.Y. (1984) 'Some aspects of Meiteilon (Manipuri) syntax', unpublished PhD dissertation, Jawaharlal Nehru University, New Delhi.
- Singh, C.Y. (1991) 'Causativization in Meiteilon', *Linguistics of the Tibeto-Burman Area* 14.2: 133–44.
- 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', *Linguistics of the Tibeto-Burman Area* 12.2: 93–100.

# TSHANGLA

*Erik Andvik*

## 1 INTRODUCTION

Tshangla is a Tibeto-Burman language of the Bodic subgroup, close to but just outside of the inner nucleus of so-called Tibetan ‘dialects’. 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 *Sharchhop*, or *Sharchhokpa*, ‘Easterners’. Most Bhutanese have 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 roughly 6000 people just across the border from eastern Bhutan in Arunachal Pradesh, 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 19th century (Aris 1980: 9).

## 2 PHONOLOGY

Table 27.1 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 ( $[\phi]$  and  $[x]$  or  $[h]$ ) intervocally.

There are five monophthong vowels in Tshangla, as shown in Table 27.2.

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 consonant cluster, the initial Cr (preserved in Written Tibetan (WT)) for labial initials (cf. *brangtong* ‘chest’). The WT velar-initial Cr clusters are pronounced with a retroflex coronal in Trashigang but retained in some dialects (cf. for example *krame* ‘to distribute’; Trashigang: *tame*). The combination  $/pɕi/$  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

TABLE 27.1 TSHANGLA CONSONANTS

	<i>Labial</i>	<i>Alveolar</i>	<i>Retroflex</i>	<i>Palatal</i>	<i>Velar</i>	<i>Glottal</i>
vl. stop	p*	t*	ʈ {tr}		k*	
+ asp.	p <sup>h</sup> {ph}	t <sup>h</sup> {th}	ʈ <sup>h</sup> {thr}		k <sup>h</sup> {kh}	
vd. stop	b	d	ɖ {dr}		g	
vl. aff.		ts		tɕ {ch}		
+ asp.		ts <sup>h</sup> {tsh}		tɕ <sup>h</sup> {chh}		
vd. aff.		(dz)		dʒ {j}		
vl. fric.		s*		ɕ {sh}		
vd. fric.		z		(ʒ) {zh}		
nas.	m*	n*		ɲ {ny}	ŋ* {ng}	
lat.		l				
lat. fric.		(ʎ) {lh}				
flap		r*				
approx.	w			y		h

TABLE 27.2 TSHANGLA VOWELS

	<i>Front</i>	<i>Central</i>	<i>Back</i>
high	i		u
mid	e		o
low		a	

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 27.3. 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 (cf. pha-n-a).

TABLE 27.3 TSHANGLA VERB CLASSES

<i>Form of root</i>	<i>Root</i>	<i>Copula ča</i>	<i>Copula la</i>	<i>Non-final clause</i>
obstruent-final	<i>yek</i> 'speak'	<i>yekča</i>	<i>yekla</i>	<i>yek</i>
nasal-final	<i>laŋ</i> 'sit'	<i>laŋča</i>	<i>laŋla/laŋna</i>	<i>laŋ</i>
liquid-final	<i>šor</i> 'lose'	<i>šorča</i>	<i>šorla</i>	<i>šor</i>
vowel-final	<i>pha</i> 'bring'	<i>phanča</i>	<i>phana</i>	<i>phan</i>
<i>Form of root</i>	<i>Nominalized</i>	<i>Infinitive</i>	<i>Imperative</i>	<i>Dubitive</i>
obstruent-final	<i>yekpa</i>	<i>yekpe</i>	<i>yekčo</i>	<i>yektu</i>
nasal-final	<i>laŋma</i>	<i>laŋme</i>	<i>laŋšo</i>	<i>laŋdu</i>
liquid-final	<i>šorba</i>	<i>šorbe</i>	<i>šoršo</i>	<i>šordu</i>
vowel-final	<i>phawa</i>	<i>phale</i>	<i>phayo/phai</i>	<i>phadu</i>

Two groups of verb roots do not pattern according to the generalizations presented in Table 27.3. First, certain nasal-final and liquid-final verb roots pattern together with the obstruent-final roots, as shown in Table 27.4 below for exceptional roots *laŋ* ‘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, contrasting with *laŋ* ‘sit’ from Table 27.3. The exceptional liquid-final root alternates between final /r/ and glottalized /rʔ/.

**TABLE 27.4** EXCEPTIONAL NASAL/LIQUID-FINAL VERB ROOTS

<i>Form of root</i>	<i>Root</i>	<i>Copula ča</i>	<i>Copula la</i>	<i>Non-final clause</i>
excep. nas-final	<i>laŋ</i> mount	<i>laŋča</i>	<i>lakla</i>	<i>lak</i>
excep. liquid-final	<i>sor</i> exchange	<i>sorča</i>	<i>sorʔla</i>	<i>sorʔ</i>
<i>Form of root</i>	<i>Nominalized</i>	<i>Infinitive</i>	<i>Imperative</i>	<i>Dubitive</i>
excep. nas-final	<i>laŋpa</i>	<i>laŋpe</i>	<i>laŋčo</i>	<i>laŋtu</i>
excep. liquid-final	<i>sorpa</i>	<i>sorpe</i>	<i>sorčo</i>	<i>sortu</i>

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 27.5 for exceptional *ke* ‘send’ and *qi* ‘write’. There is evidence to suggest that these are historically derived from a syllable-final /l/ (cf. Andvik 1999 for a complete account):

**TABLE 27.5** EXCEPTIONAL VOWEL-FINAL VERB ROOTS

<i>Form of root</i>	<i>Root</i>	<i>Copula ča</i>	<i>Copula la</i>	<i>Non-final clause</i>
excep. e-final	<i>ke</i> send	<i>keča</i>	<i>kela</i>	<i>ke</i>
excep. i-final	<i>qi</i> ‘write’	<i>qiča</i>	<i>qila</i>	<i>qi</i>
<i>Form of root</i>	<i>Nominalized</i>	<i>Infinitive</i>	<i>Imperative</i>	<i>Dubitive</i>
excep. e-final	<i>keba</i>	<i>kebe</i>	<i>kešo</i>	<i>kedu</i>
excep. i-final	<i>qiba</i>	<i>qibe</i>	<i>qišo</i>	<i>qidu</i>

#### 4 NOUN PHRASE

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.

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:

- (1) *dukpu waksta khepa*  
 poor child TOP  
 ‘the poor child’ (identification)



- (2) *waktsa dukpu khepa*  
 child poor TOP  
 ‘the child, who is poor’ (description)

Relative clauses, like adjectives, occur both before and after the NP-head, the pre-nominal position being restrictive (3) and the post-nominal descriptive (4):

- (3) *Onya phai chot-khan songo ja-ga chharo gila.*  
 DEM house make-REL person 1s-LOC friend COP  
 ‘That person building the house is my friend.’
- (4) *Onya songo phai chot-khan khepa ja-ga chharo gila.*  
 that person house build-REL TOP 1s-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 27.6 DEMONSTRATIVES

	<i>Non-locative</i>	<i>Locative</i>
Proximate	<i>uthu</i> ‘this’	<i>otha</i> ‘this/here’
Distal	<i>unyu</i> ‘that’	<i>onya</i> ‘that/there’

A possessive NP precedes the determiner as well as the NP head. Possessive pronouns are transparently analysable as a personal pronoun plus a locative case particle:

- (5) *ro-ka uthu pechha*  
 3-LOC DEM book  
 ‘this book of his’

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 *tshebang* ‘some’, *mangpu* ‘many’, *nyungpu* ‘few’, or the definite plural clitic *-ba*. There is no other number marking on nominals:

- (6) *ngam zum*  
 day seven  
 ‘seven days’
- (7) *pechha chhilu-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 (8) indefinite *thur* occurs with *mangpu* ‘many’:

- (8) *Jang songo mangpu thur-gi-rang kha + tang-nyi . . .*  
 1s 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.):

- (9) *Ji-gi pura hang tshat-pa thur nan-ga bi-wa.*  
 1s-AGT completely what need-NOM one 2s-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:

- (10) *Unyu waktsa khepa ji-gi she-le khe-le dang.*  
 DEM child TOP 1s-AGT kill-INF must-INF PRT  
 '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:

- (11) *Jang nangpa chho khepa-ga ten-nyi...*  
 1s Buddhist faith TOP-LOC adhere-NF  
 'I adhered to the Buddhist faith...'

A diverse class of focus particles follow the case-marker. Examples are the contrastive emphatic marker *-bu*, an emphatic marker *-rang*, and marked topic particles *-chho*, and *-la*.

- (12) *Tshebang-gi waktsa-ba-ka-bu ngalong-ga chhas-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 (13), 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 (14), the semantic patient subject of the non-final clause controls the final clause subject:

- (13) *Gopen-gi tsonpa kong-nyi, Ø phiska di-wa.*  
 chief-AGT prisoner strike-NF Ø outside go-NOM  
 'The chief beat the prisoner and went outside.'  
 \*'The chief beat the prisoner and he (the prisoner) went outside.'
- (14) *Tsonpa gopen-gi brak-nyi, Ø phiska di-wa.*  
 prisoner chief-AGT scold-NF Ø 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 multi-valent verb may be omitted when it is regarded as unimportant, as in example (15) below:

- (15) *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 + go-le* ‘heart + put – INF’ = ‘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*) 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, (cf. Andvik 1999 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:

- (16) *Ji-gi nan chhowang-gi za-me!*  
 1s-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:

- (17) *Ro don-gi nyos-pa.*  
 3 demon-AGT crazy-NOM  
 ‘He became 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 (cf. Section 14 below).

### 6.2 Locative/dative

The locative/dative case-marker *-ga* (*-ka*) marks the recipient or goal of a di-transitive 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’:

- (18) *Ro-ki ja-ga brang-pa.*  
 3-AGT 1s-LOC scold-NOM  
 ‘He scolded me.’

Oblique locative or temporal adverbials, which often precede the core subject, may also be marked with *-ga*:

- (19) *Thimpu-ga phom khe-na.*  
 Thimpju-LOC snow fall-COP  
 ‘It’s snowing in Thimpu.’

Locative/dative case-marking occurs on the subject to encode possession in a copular clause.

- (20) *Jelpo-ga waktsa sam chho-wa.*  
 king-LOC child three stay-NOM  
 ‘The king had three children.’

- (21) *Ja-ga gari man-cha.*  
 1s-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.

- (22) *Na-ga mi gapthang lekpu ma-la.*  
 2s-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*) marks oblique locative or temporal arguments as the source or reference point away from which the action of the proposition is directed:

- (23) *Ro-ki yek-khan nyan-pe-ga songo-ba throm-gai zo-ma-la.*  
 3-AGT speak-REL listen-INF-LOC person-PL town-ABL gather-NOM-COP  
 ‘The people gathered from the town to listen to him speak.’
- (24) *Chhutse gu-gai ja-ga sungjapa a-le kor chho-wa.*  
 hour nine-ABL 1s-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:

- (25) *A-chhing nyiktsing ung nang-kai leng jong-khe.*  
 1p-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 (26):

- (26) *Jang uthu taphur di-wa-gai kukhaila-kap di-wa drik-pe.*  
 1s 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 27.7 shows the entire final verb paradigm for the verb *dile* ‘to go’ (non-mirative forms only, see section 9 below):

TABLE 27.7 AFFIRMATIVE VERB PARADIGM

	<i>Past</i>	<i>Present</i>	<i>Future</i>
Simple pftv.	<i>di-wa</i> ‘went’		<i>di-le</i> ‘will go’
Simple impf.	<i>din-chho-wa</i> ‘was going’	<i>din-cha</i> ‘is going’	<i>din-chho-le</i> ‘will be going’
Perfect pftv.	<i>di-wa-chho-wa</i> ‘had gone’	<i>dī-wa-cha</i> ‘has gone’	<i>dī-wa-u-phe</i> ‘will have gone’ ‘would have gone’
Perfect impf.	<i>din-chho-wa-chho-wa</i> ‘had been going’	<i>din-chho-wa-cha</i> ‘has been going’	<i>din-chho-wa-u-phe</i> ‘will have been going’ ‘would have been going’
Prospective pftv.	<i>di-le-chho-wa</i> ‘was going to go’ ‘would have gone’	<i>di-le-cha</i> ‘is going to go’	<i>di-le-u-phe</i> ‘will be about to go’ ‘may go’

The nominalized or infinitive participle (cf. Table 27.3 above), functions as the main verb, though it is structurally embedded under the auxiliaries *chhole* ‘to stay’, *uphe* ‘to come’, and the existential/descriptive copula *cha*. The perfective verbal phrase is formally analogous to a complement clause structure (cf. section 13), while the imperfective verbal phrase is formally analogous to a serial verb construction (cf. section 16). Note that the 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*, *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 contrafact, permission, and obligation (e.g. *di-le chho-wa* ‘would have gone’, *din chho-wa u-phe* ‘would have been going’).

## 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 27.8. With the exception of the prospective forms, the negative prefix is attached to the verb root in perfective constructions, to the auxiliary in imperfective. In the negative constructions, the place of the existential copula is occupied by the auxiliary *chhi*. The verb suffix *-lu* occurs only in the simple imperfective negative past and present forms, where the negative prefix occurs on the auxiliary.

TABLE 27.8 NEGATIVE VERB PARADIGM

	<i>Past</i>	<i>Present</i>	<i>Future</i>
Simple pftv.	<i>ma-din-chhi</i> 'did not go'		<i>ma-di-la</i> 'will not go'
Simple impf.	<i>di-lu-man-chhi</i> 'was not going'	<i>di-lu-ma-n-cha</i> 'is not going'	<i>din-ma-chho-la</i> 'will not be going'
Perfect pftv.	<i>ma-di-wa-chho-wa</i> 'had not gone'	<i>ma-di-wa-cha</i> 'has not gone'	<i>ma-di-wa-u-phe</i> 'will not have gone' 'would not have gone'
Perfect impf.	<i>di-wa-man-chhi</i> 'had not been going'	<i>di-wa-ma-n-cha</i> 'has not been going'	<i>di-wa-ma-(u)-pha</i> 'will not have been going' 'would not have been going'
Prospective pftv.	<i>ma-di-le-chho-wa</i> 'was not going to go'  'would not have gone'	<i>ma-di-le-cha</i> 'is not going to go'	<i>ma-di-le-u-phe</i> 'will not be about to go' 'may not go' alt: <i>di-le-ma-(u)-pha</i> 'should not, ought not go'

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 mang-gi*).

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 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 *cha* is *la*:

- (27) *Ama khamung zik-cha.*  
mother clothes wash-COP  
'Mother is washing the clothes.'
- (28) *Ama khamung zik-la.*  
mother clothes wash-COP  
'Mother is evidently washing the clothes.'

The non-mirative utterance in (27) would be spoken by a person with prior knowledge of the proposition. The mirative utterance (28) 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  $\emptyset$ -marking, rather than with the non-mirative *cha*.

## 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 *cha* (or its mirative counterpart *la*) and those containing the equative copula *gila* (or its mirative counterpart *giwala*).

- (29) *Otha phai chhilu cha.*  
 DEM house great COP  
 'That house is big.'
- (30) *Ja-ga dung Wamrong gila.*  
 1s-LOC village Wamrong COP  
 'My village is Wamrong.'

Copular clauses built on *cha* 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:

- (31) *Unyu to zhimpu gila!*  
 DEM food sweet COP  
 'This food is delicious!'

In non-final clauses, for descriptive and other predications, the grammaticalized verb *chhole* 'to stay' takes the place of the existential copula *cha*. 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, and 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 *gi-du*. It presents a proposition as a potentiality:

- (32) *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':

- (33) *Ser nyong-tu dak-pa rewa-gi rok-tsing thur thur-gi*  
 gold receive-SUB say-NOM hope-AGT 3-DUAL one one-AGT  
*kholong phi-wa.*  
 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 both in an equative predicate, or as a grammatical auxiliary in a verbal predicate. There is no corresponding subjunctive final verb suffix *-sa*. (However cf. section 12 for a homophonous relativizer suffix *-sa*.)

- (34) *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 . . .’:

- (35) *A-chhing ja-ga ajang-ga brang-ka din-than chhas a-khe.*  
 1p-DUAL 1s-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/-cho/-i(yo)* encodes the imperative mood:

- (36) *Ja-ga tam thur shik-cho!*  
 1s-LOC story one explain-HOR  
 ‘Tell me a story!’

The third person hortative suffix *-chhen* expresses an injunction by the speaker upon a third party, as in English ‘Let him . . .’ ‘He should . . .’, or ‘He must . . .’:

- (37) *Songo mar-khan-gi man zan-chhen*  
 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 (38), the addressee of the second-person imperative *-sho* is the first-person subject of the interrogative sentence:

- (38) *Jang thar-sho mo?*  
 1s 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 the content question with the sentence-final particle *ya*, as in (40). 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 same syntactic position as would the constituent it replaces in a declarative sentence:

- (39) *Unyu wa daza brang-ma mo?*  
 DEM cow young bear-NOM QUES  
 ‘Has this cow born a calf?’
- (40) *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 *chhole* ‘to stay’ takes the place of the existential copula *cha*. The equative copula does not occur in the hortative mood (*\*gikhe*, *\*gi-sho*, *\*gi-chhen*).



## 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 external-headed, i.e. the head noun occurring outside the relative clause itself. The most common relativizing suffix is *-khan*:

- (41) *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.’
- (42) *Ro-ki gadang-gi tsung-khan songo thamchen-rang rolong ri-le.*  
 3-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’:

- (43) *Nyi lok-nyi jang chho-sa phai-ga shek-pa.*  
 PRT return-NF 1s stay-REL house-LOC arrive-NOM  
 ‘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:

- (44) *Nangka shek-pa(-ga) songo-ba shadar phi-nyi cha giwala.*  
 inside arrive-NOM (-LOC) person-PL shout do-NF cop COP  
 ‘The people who had arrived inside were shouting.’
- (45) *Dangpo shing khuk-pe (-ga) zogo khepa thur chho-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* (46), *-sa* (47), or a nominalized or infinitive verb, (48) and (49):

- (46) *Ngang se-khan-gi ngang jang-cha.*  
 song know-REL-AGT song sing-COP  
 ‘Those who know the songs are singing.’
- (47) *Nan-gi yek-sa chho-le gi-du.*  
 2s-AGT speak-REL stay-INF COP-SUB  
 ‘I suppose we will live where you say (‘... live at your saying’).’
- (48) *Nyi rang-gi re-ba (-ga) thur a-n-chho-wa.*  
 PRT RFLX-AGT can-NOM-LOC one do-NF-stay-NOM  
 ‘They were doing whatever they could.’
- (49) *Na-shi zemu-ba za-le ja-me (-ga) bi-le khe-le.*  
 2s-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:

- (50) *Jang yu ja-me (-ga) thup tha-wa.*  
 1s wine drink-INF (-LOC) thow leave-NOM  
 'I left off drinking wine.' (*thup thale* = 'abandon, leave')
- (51) *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:

- (52) *Ro-ki jelpo zi-me bi-wa.*  
 3s-AGT king lie-INF give-NOM  
 'He had the king lie down.'

Case-marking evidence 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 (53), 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*:

- (53) *Ibi u-na mo got-cho!*  
 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 (54) and objects of postpositions like *saken/sakpu* 'until', *korgai* 'about', and *dawa* 'as', 'like' (55). The locative/dative case-marker occurs optionally on the embedded noun-complement, but not on the complement of a postposition:

- (54) *Nyi dan shi-wa-ga tshang khepa songo thamchen-gi*  
 PRT 3 die-NOM-LOC news TOP person all-AGT  
*na tha-wa.*  
 ear leave-NOM  
 'All the people heard the news that he had died.'
- (55) *Uthu chho jang kawa chhilu chat-pa dawa*  
 DEM TOP 1s hardship great suffer-NOM like  
*thong-mu man-chhi.*  
 see-PRT 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:

- (56) *Shama di-nyi-bu, jang dangsanken yitka u-na.*  
 long.time go-NF-PRT 1s clearly memory come-COP  
 ‘Even though a long time has passed, I still remember clearly.’

- (57) *Chhoga phi-le-ga, ro tormu katang chos-pa.*  
 ritual do-NP-LOC 3 TORMU big make-NOM  
 ‘In order to do the ritual, he built a large *tormu*.’

The rhetorical relationship between adverbial and matrix clause 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 ...’)
<i>-nyi-bu</i>	concessive (‘even though ...’)
<i>-nyi</i>	unspecified non-final verb
<i>-la</i>	concurrent (‘while ...’)
<i>-la-gai</i> (ABL)	temporal anteriority (‘until ...’, ‘before ...’)
<i>-le(-ga)</i> (LOC)	purpose/goal (‘in order to ...’)
<i>-wa(-ga)</i> (LOC)	attendant circumstance (‘with/without ...’)
<i>-wa-gi</i> (AGT)	cause (‘because ...’)
<i>-wa-gai</i> (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:

- (58) *Nyi gisa na gisa dak-nyi, bozong zong-nyi,*  
 PRT maybe PRT maybe say-NF cassava boil-NF  
*khoptang khop-nyi, laga-gi chhom-nyi, nyi sa nangka*  
 skin peel-NF leaf-AGT wrap -NF PRT ground in  
*chhe-nyi, onya wang thur tsuk-nyi tha-wa.*  
 plant-NF thus hole one put-NF leave-NOM.

‘Thinking, “Well, maybe”, boiling the cassava, peeling it, wrapping it with a leaf, 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 (1999) shows that clause chains, ‘serial predicates’, and ‘serial verbs’ represent three points on a ‘concatenation continuum’, representing the degree of syntactic and semantic merging.

- (59) *Dorji tiru chum-nyi u-pha.*  
 Dorji money finish-NF come-NOM  
 ‘Dorji used up the money and came.’
- (60) *Ro ro-ten jap a-nyi chhas mangpu zhu-wa-la.*  
 3s 3s-RFLX save do-NF talk much offer-NOM-COP  
 ‘He said many things in his own defence (lit. ‘Saving himself, he offered much talk.’)
- (61) *Kha thur phai nangka phur-nyi u-pha.*  
 bird one house in fly-NF come-NOM  
 ‘A bird flew into the house.’

Example (59) 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 (60) 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 (61) 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 (62).

- (62) *Nyi ro thong-ma-kap-nyi songo thur-gi ngar-nyi bi-wa-la.*  
 PRT 3 see-PTC-with-NF person one-AGT laugh-NF give-NOM-COP  
 ‘And upon seeing him, one person laughed at him.’

As seen in Section 7 above, the serial verb construction with *chhole* ‘to stay’ is unique in having grammaticalized to the point where *chho* 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’:

- (63) *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 (63), 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 (64) or cognition/intention (65):

- (64) *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!”’
- (65) *Lhangpochhe mang-Ø-chhen dak-nyi, ro shu a-nyi lhak-chho-wa.*  
 elephant NEG-come-HOR say-NF 3 strong do-NF read-stay-NOM  
 ‘Thinking, ‘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:

- (66) *Phama-gi waktsa-ba thamchen dolo a-nyi phang-cha.*  
 parent-AGT child-PL all equal do-NF love-COP  
 ‘Parents love their children equally.’

An instrumental argument may be embedded under *anyi*:

- (67) *Ro-ki nai-ba ri-gi a-nyi thrisor phi-wa.*  
 3-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’:

- (68) *Songo mangki ma-se-wa a-nyi ro thup tha-le khe-le-la.*  
 person public NEG-know-NOM do-NF 3s thow 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*:

- (69) *Jang thola to hang-rang ma-ga-la-n, ri thur-rang*  
 1s up.there food what-EMPH NEG-give-PRT-NF water one-EMPH  
*ma-ga-la-n, otha nangka dap tha-nyi...*  
 NEG-give-PRT-NF DEM in insert keep-NF  
 ‘They kept 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 *chhole* ‘to stay’:

- (70) *Ro ma-shi-la-n chho-la-kap, khamung tshok-pa chho-wa.*  
 3 NEG-die-PRT-NF stay-PTC-with clothing sew-NOM stay-NOM  
 ‘Before she died, she had made clothing.’ (lit. ‘While she was staying  
 not dying, she had made clothing’.)

## REFERENCES

- Andvik, E. (1999) *Tshangla Grammar*, University of Oregon Ph.D. dissertation.  
 Aris, M. (1980) ‘Notes on the history of the Mon-Yul Corridor’, in M. Aris and Aung San Suu Kyi (eds) *Tibetan Studies in Honour of Hugh Richardson*, New Delhi: Vikas.  
 Das Gupta, K. (1968) *An Introduction to Central Monpa*, Shillong: North-East Frontier Agency.  
 Egli-Roduner, S. (1987) *Handbook of the ‘Sharchhokpa-lo/Tshangla’*, Thimpu: Helvetas.  
 Sun H., Luy S., Zhang J., and Ouyang J. (1980) *The Languages of the Monpa, Luoba, and Teng Peoples*, Beijing: Chinese Academy of Social Sciences.  
 van Driem, G. (1998) *Dzongkha*, Leiden: Research School of Asian, African and Amerindian Studies.  
 Zhang, J. (1986) *Sketch Grammar of Cangluo Monpa*, Beijing: Chinese Academy of Social Sciences.

## FURTHER READING

- Andvik, E. (1999) *Tshangla Grammar*, University of Oregon Ph.D. dissertation. (A comprehensive reference grammar with illustrative examples.)  
 Egli-Roduner, S. (1987) *Handbook of the ‘Sharchhokpa-lo/Tshangla’*, Thimpu: Helvetas. (A wordlist containing hundreds of words with short glosses, especially medical terms.)

# TANI LANGUAGES

*Jackson T.-S. Sun*

## 1 BACKGROUND

Tani refers to a compact cluster of Tibeto-Burman languages situated at the eastern end of the Himalayas, in an area skirted on four sides by Tibet, Assam, Bhutan, and Burma. The Tani languages are spoken by about 600,000 indigenous people of Arunachal Pradesh and northern Assam, including the Adi (many subtribes), Nishi(ng)-Bengni, Hill Miri, Tagin, and Apatani tribes of East Kameng, Lower Subansiri, Upper Subansiri, West Siang, East Siang, and Dibang Valley districts of Arunachal Pradesh, as well as the Mising people of Assam. In Arunachal Pradesh alone the Tani-speaking area covers some 40,000 square kilometres, or roughly half the size of the state (Simon 1978: 8). 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 (Ouyang 1985: 76), where they are lumped with certain linguistically non-Tani peoples (e.g. the Idu, Sulung, and Bangru) to form the Luoba nationality.

Tani languages constitute a distinct branch in Tibeto-Burman, as argued in Sun 1993a. Their closest relatives appear to be their eastern neighbours the Digarish (Taraon and Idu) languages. They also share some deep-rooted similarities with the Hrusish (Hruso, Dhammai, and Levai or Bangru) languages to the west, which however may also be a result of prolonged contact. By both phonological and lexical criteria, Tani languages fall into two divisions: Eastern Tani consisting of Mising plus many of the Adi dialects, and Western Tani consisting of dialects spoken by the Nishi(ng)-Bengni, Tagin, Gallong, and Hill Miri tribes. Bokar and Damu represent transitional types and are more difficult to subclassify. In addition, the two aberrant languages Apatani and Milang show overall affinities with Western and Eastern Tani, respectively. Unless noted otherwise, the (Na) Bengni and Bokar data in this chapter are from my personal research conducted in Tibet.

## 2 PHONOLOGY

The syllable canon in Tani is (C)(C)V(C). The overwhelming majority of syllables show either the CV or CVC syllable structure. Tani languages diverge greatly with respect to initial consonant clusters. The eastern languages tend to have no initial consonant clusters at all, or at most a few labial plus *-j* clusters; Proto-Tani consonant clusters are better preserved in Western Tani.

Segmental inventories are relatively simple. The most striking features of the modern initial-consonant system are the absence of distinctive aspiration in the stops, and the paucity of spirants and affricates. Many languages have only one single supraglottal spirant, represented in the sources as either *s* or *ʃ*, and a glottal spirant *h* (or *ɦ*); some languages have developed an additional velar spirant *x*, while others have the labiodental spirants *f* and *ν* instead. Most languages have only palatoalveolar affricates, *tʃ* and *dʒ*. Dental/alveolar and retroflexed affricates, if any, were introduced in loanwords. Most languages have

only four contrasting places of articulation for stops/affricates and nasals: bilabial, dental/alveolar, palato-alveolar, and velar. The other three common initial consonants are the liquids *l*, *r*, and the semivowel *j*. The phonemic status of *j* in languages such as Bokar is established by minimal pairs like *i:* ‘bow (weapon)’, *ji:* ‘to bury’. Typically, a seven-vowel system is found, with two central unrounded vowels *ə* and *u* in addition to *a*, *i*, *u*, *e*, and *o*. Languages with fewer vowels (e.g. Apatani, with six vowels *a*, *i*, *u*, *e*, *o*, *u*; Abraham 1985: 7) or more (e.g. Damu, with as many as ten vowels *a*, *i*, *u*, *e*, *o*, *ə*, *u*, *y*, *ɸ*, *ɣ*; Ouyang 1985: 77) are also reported. All Tani languages seem to distinguish vowel length, the lexical function of which differs from language to language. In Bokar and Bengni, vowels contrast in length regardless of position in a word, whereas vocalic length seems to be distinctive only in non-final open syllables in Gallong and Apatani (Weidert 1987: 215–23), and only in the first syllable of polysyllabic words in certain Nishi dialects (Chhangte 1992: 13). Most instances of long vowels in modern Tani occur in open syllables and are clearly secondary; yet there is some evidence of length contrast in closed syllables as well, suggesting that distinctive vowel length could be an old feature from Proto-Tani. True diphthongs did not exist originally; the diphthongs in the modern languages are from neighbouring languages or are the result of syllable contraction. Tani languages vary greatly in their inventories of syllable codas. The language that preserves the largest number of the original proto-codas is Padam Adi, with *-p*, *-t*, *-k*, *-m*, *-n*, *-ŋ*, *-r*, and *-l*. Milang is the only other language documented that preserves the Proto-Tani *\*-l* coda. The velar nasal *\*-ŋ* and the dental/alveolar stop *\*-t* codas are the most prone to loss, especially in Western Tani. The tendency towards coda attrition is epitomized in Apatani, where only two of the original Proto-Tani codas remain, i.e. *-ʔ* (from the original stop codas) and *-r*. New final consonants, including *-s*, *-l*, and *-m*, have been imported in loans from Indo-Aryan and English. In many varieties of Nishi(ng), Gallong, and Hill Miri, non-original codas are created via the clipping of short final vowels from original disyllables, leading to an innovative voicing contrast in the stop codas, e.g. Nishi *ab* ‘father’, *tap* ‘pumpkin’ (Chhangte 1992: 4); cf. Bengni *a-bu:* ‘father’, *ta-pu:* ‘pumpkin’.

Phonemic tone is not a prevalent feature. While completely non-existent in such eastern languages as Bokar, lexically significant pitch occurs to varying degrees in many Western languages. In one type of Na Bengni, for example, high-low pitches need to be lexically marked for certain words, e.g. *ta<sup>H</sup>-nu:<sup>L</sup>* ‘hemp’, *ta<sup>L</sup>-nu:<sup>H</sup>* ‘snail’. In another Bengni dialect, interestingly, the same words are distinguished by vowel qualities: *ta-no:* ‘hemp’, *ta-nu:* ‘snail’. Further afield, Weidert (1987: 219–21) reports a system of word tone for Gallong, with three tonal melodies operating on the domain of the phonological word. Lastly, Apatani appears to have a true syllable-tone system in which each syllable potentially bears a two-way (Weidert 1987: 216–19) or three-way (Abraham 1985: 5–6) tonal contrast.

Phonological alternations are quite extensive, particularly syllable-structure adjustments or assimilation. The rich allomorphy of the nominal prefix *\*a-* in Bokar provides a typical example of the attested sandhi variations: *a-ŋa:* ‘child’; *i-ki:* ‘dog’; *u-puk* ‘arrow’; *e-tfe* ‘clothes’; *o-ŋo:* ‘fish’; *ə-jək* ‘pig’; *u-luŋ* ‘stone’. Morpheme identification can be tricky without knowing the sandhi processes at work. For example, given *luŋin*, the phonetic form of the Bokar word ‘fingernail’, four morphophonemic rules have to be undone to retrieve the original shapes of the component morphemes *lok* ‘hand’ and *jɪn* ‘nail’, namely initial *j*-simplification (→ *lok-in*), obstruent voicing assimilation (→ *log-in*), spirantization (→ *luŋ-in*), and vowel assimilation (→ *luŋ-in*), compare Bokar *a-lok* ‘hand’, *lə-jɪn* ‘bird’s claw’ (compare also Bengni *lak-sin* ‘fingernail’, from Proto-Tani *\*lak-zin*).



### 3 MORPHOLOGY

#### 3.1 Lexical categories

The major lexical categories are nominals, verbs, adverbials, and particles. Adjectives are a subclass of stative verbs. Like other verbs predicate adjectives co-occur with existential verbs in the declarative, as in (1); in their attributive use they are optionally nominalized, especially in Eastern Tani (2):

- (1) Apatani (Abraham 1985: 72)  
*mju si dʒaŋtu do*  
 person this fat exist  
 ‘This person is fat.’
- (2) Bokar (Ouyang 1985: 45)  
*etfe luŋkaŋ-na po da*  
 clothes red-NOMZR good exist  
 ‘The red clothes are good.’

Adverbials of manner and result take distinct adverb markers, such as *-bo* in Bokar, e.g. *po-bo lop* (good-ADV learn) ‘to learn well’; *jarro-bo tuktfin* (long-ADV stretch) ‘to stretch until long’. Temporal adverbials take the object case instead, e.g. *lako-me* ‘sometimes’. Notably, morphemes that form compounds with the main verb root are used to express many common adverbial meanings, e.g. Bokar *ten* ‘again’, *in-ten* ‘to go again’. Particles are a broad category, expressing a host of modal and illocutionary-force meanings. Illocutionary-force particles are of particular syntactic importance, as they are the main coding device for major sentence types (e.g. interrogative, imperative, prohibitive, see Section 4.3 below). The minor category of classifiers is grammaticalized to varying degrees. In Bokar, for example, classifiers do not occur with numerals higher than one, and even there their use is optional. Classifiers are fully functional in Western Tani where they are required on quantified count nouns (except nouns denoting humans). In many languages, moreover, a few common statives such as ‘to be big’ and ‘to be small’ must co-occur with the same classifiers as the head nouns they modify, as illustrated in Bengni:

- (3) *afa: so:-tuur-but so:-pi-gu*  
 rope CL-big-NOMZR CL-two-PART  
 ‘two big ropes’

#### 3.2 Morphological processes

Words are typically disyllabic (or quadrisyllabic in some reduplication constructions), consisting commonly of an affixed root or a compound.

Unaffixed monosyllabic words are scarce; this is particularly true of nominals, most of which take nominal prefixes. Unlike in many neighbouring Tibeto-Burman groups, the small inventory of Tani prefixes applies to nominal and stative-verb roots, but rarely to other types of verbs. The four most commonly attested prefixes are, in reconstructed form, *\*a-*, *\*sa-*, *\*ta-*, and *\*pa-*. The distribution of these prefixes is grounded in semantics: *\*sa-* is attached to roots referring to higher animals; *\*ta-* is used for lower animals, insects, diseases, trees, and small objects in general; *\*pa-* is confined mainly to bird names and related nouns (e.g. Bokar *pu-tuur* ‘coop’). As for *\*a-*, it is found not only on kinship terms, body parts, and nominal roots from assorted semantic areas, but also on many adjectives. The prefixes *\*sa-* and *\*pa-* appear to have evolved from first components of earlier compounds, reflecting perhaps Proto-Tibeto-Burman

\**sya* ‘flesh/meat/animal’ and \**bya* ‘bird’ (Benedict 1972: 46, 177). Other interesting examples of lexical reanalysis as an ongoing process are the ‘weather prefix’ stemming from Proto-Tani \**doŋ* ‘sky’, ‘rain’ and appearing in such forms as Bengni *doŋ-muk* ‘cloud’, *doŋ-ri* ‘wind’, *doŋ-pi huŋ* ‘rain’, *doŋ-gum* ‘thunder’, and the ‘eating prefix’ *du-* in Bengni, etymologically a reduced form of the root *duŋ* (Proto-Tani \**do*) ‘to eat’, occurring not only in compounds related semantically to eating: *du-rom* ‘feast’, *du-gam* ‘food’, *du-muŋ* ‘curry’, *du-duk* ‘poison’, but also by extension to actions involving the mouth: *du-rjak* ‘to lick’, *du-bjuŋ* ‘to suck’, *du-tfiŋ* ‘to gnaw’, *du-mit* ‘to swallow’.

Compounds abound in the Tani lexicon. Attested are noun-noun (e.g. Bengni *siŋ-lak* ‘branch’ < ‘tree’ + ‘hand/arm’), noun-stative verb (e.g. Gallong *dʒe-ku* ‘old cloth, rag’ < ‘cloth’ + ‘to be old’; Das Gupta 1963: 16), verb-adverb (e.g. Bokar *ben-pjoŋ* ‘to speak first’ < ‘to speak’ + ‘first’), as well as compound verbs. Compound verbs, formed by combining verbs with a whole array of *auxiliary verbs*, are of primary importance. The auxiliaries in this construction complement the main verb by adding various meanings, e.g. directionality (e.g. Bokar *puk-tfaŋ* ‘to sell uphill’ < ‘to sell’ + ‘to ascend’), result (e.g. Bokar *jit-keŋ* ‘to beat to death’ < ‘to beat’ + ‘to die’), and causativity (e.g. Bokar *koŋ-mo(:)* ‘to cause to see’ < ‘to see’ + ‘to do’).

Stem modification occurs sporadically, e.g. modification by means of vowel length in Bokar: *i-si* ‘water’, *i-si:* ‘urine’; *a-gu* ‘to be hot’, *gu:* ‘to burn v.i.’. The familiar Tibeto-Burman voicing alternation in the formation of simplex-causative pairs is preserved in some measure by certain languages, such as Padam-Mising (Lorrain 1907):

- (4) *dir* (as of sticks) to break  
*tir* to break (sticks)  
*ben~bet* (as of ropes) to break  
*pen~pet* to break (ropes)

Reduplication is an important morphological device favoured by Tani languages to increase lexical bulk. Words can also contain four syllables, often consisting of two disyllabic ‘couplets’ with some material reduplicated. In (5), some Bengni quadrisyllabic words with a superordinate-hyponym construction are offered:

- (5) a. *su-kir* *kiŋ-puŋ* stallion  
 PREF-horse horse-male  
 b. *a-ki* *ki-tfiŋ* small intestines  
 PREF-intestines intestines-small  
 c. *u-si* *si-lap* wave n.  
 PREF-water water-wave

Another important subtype is expressives, illustrated by these Bengni examples: *a-tuŋ a-juŋ* ‘garbage’, *a-bak a-jak* ‘thick (liquid)’, *dʒi-sit dʒi-mit* ‘to pass something around’, *jip-miŋ jip-maŋ* ‘sleepy’, *muŋ-ruk muŋ-rak* ‘to sleepwalk’. Productive reduplication also occurs in such Apatani formal idioms as *v-si v-ha* ‘difficult to v’, e.g. *lu-si lu-ha* ‘difficult to tell’ (Abraham 1985: 90).

### 3.3 Derivational morphology

Relatively little derivation is exploited in word formation. A good number of nouns function directly as verbs with no denominalizing marking. Shown below is the cognate-verb construction in Bengni:

- (6) to lay egg *pu-pu pu* (cf. *pu-pu* egg)  
 to spit *ta-tur tfiur* (cf. *ta-tfiur* spittle)

to dream      *jip-maː maː*      (cf. *jip-maː* dream)  
 to wear shoes *lu-kjam kjam*      (cf. *lu-kjam* shoe)

Nominalization produces action or state nominals as well as participant nouns. The following are the various nominalizers in Apatani, where nominalizing devices are more fully developed (Abraham 1985: 117–20):

- |   |   |
|---|---|
| (7) Action nominalizer <i>-nũ</i>         | ( <i>mu</i> to do; <i>mu-nũ</i> doing)                                |
| Actor nominalizer <i>-nu</i>              | ( <i>puru</i> to read; <i>puru-nu</i> student)                        |
| Undergoer-locative nominalizer <i>-ko</i> | ( <i>du</i> to eat; <i>du-ko</i> place of eating;<br>something eaten) |
| Instrument nominalizer <i>-na-nu</i>      | (e.g. <i>upu</i> to sweep → <i>upu-na-nu</i> broom)                   |

### 3.4 Inflectional morphology

#### 3.4.1 Nominal inflection

Nominals are inflected for case, definiteness, and deixis. The major cases, marked by suffixes, are genitive, locative, instrumental, and the ‘object’ cases. The allative cases are often identical to, or built on, the locative; the ablative and comitative are likewise often derived from the locative, e.g. Padam-Mising *lok* ‘from’; *kə-lo* ‘along with’ < *lo* ‘in/at’; Apatani *ko-ku* ‘from’; *agi-ho* ‘along with’ < *ko-ho* ‘in/at’ (Abraham 1985: 50–1). The multifunctional object case marks various undergoer (patient, recipient, benefactive) arguments, standards in comparative structures (8), complement clauses (9), as well as temporal adverbials (see (29) below). In some languages, both undergoer arguments of ditransitive verbs are case marked (10), whereas other languages allow only definite nominals to take the object case (11a–b):

- (8) Padam-Mising (Lorrain 1907: 527)  
*ŋo-m pənamə bu bottə ja dak*  
 1SG-OBJ than 3SG big more exist  
 ‘She/he is bigger than I am.’
- (9) Padam-Mising (Lorrain 1907: 511)  
*ŋo-k jokfik lə-nam-əm no muɸa du je*  
 1SG-GEN knife put-NOMZR-OBJ 2SG remember exist Q  
 ‘Do you remember where I put my knife?’
- (10) Padam-Mising (Lorrain 1907: 434)  
*no-k jokfik də-m ŋo-m bipoŋ ika*  
 2SG-GEN knife that-OBJ 1SG-OBJ lend IMP  
 ‘Lend me your knife.’
- (11) Bokar  
 a. *ŋoː pətəŋ ako(\*-m) koŋpoŋ-pa*  
    1SG bird INDEF-OBJ see-PFV  
    ‘I saw a bird.’  
 b. *ŋoː pətəŋ-hə-m koŋpoŋ-pa*  
    1SG bird-DEF-OBJ see-PFV  
    ‘I saw the bird.’

Agents are not case marked, except for demoted agents marked by the instrumental case in the passive construction, as in (12):

- (12) Apatani (Abraham 1985: 140)  
*njummu-huu aki-mi tade-lo bi-ko-ne*  
 woman-DEF dog-OBJ Tadde-INST give-PASS-PFV  
 ‘The woman was given a dog by Tadde.’

Definiteness marking, distinct from demonstratives and third person pronouns, has arisen in some languages (e.g. Padam-Mising *də*, Bokar *fə*; see (11b) above for illustration). Deictic pronouns and adverbials are well-developed in Tani, with a distance-based (proximal, distal, far-distal) dimension often in combination with a vertical (‘up’ and ‘down’) dimension, as shown in the following Bengni expressions (*niː* = ‘person’; *buː* = ‘down’):

- (13) *buː siː niː buː* this person down here  
*buː niː buː* that person down there  
*buː loː niː buː* that person down yonder

Deixis figures in nominal inflection to the extent that deictically differentiated demonstrative pronouns also serve as third person pronouns; consider the following distinct ways of saying ‘him/her’ in Bengni and Bokar:

- (14) Bokar Bengni Position of referent  
*su-m a-so-m* nearby  
*a-m a-lo-m* at some distance  
*tə-m tuː-lo-m* at some distance uphill  
*bə-m buː-lo-m* at some distance downhill

### 3.4.2 Verb inflection

Tani is one of the few major Tibeto-Burman branches with no attestation of argument cross-referencing. Instead, the relevant inflectional categories of the verb are tense-aspect and mode, much as in Tibetan.

The tense-aspect system consists of a set of tense-aspect marking suffixes. The meagre descriptive sources on Tani languages often mention the so-called ‘past’ tenses. This is highly suspect, for one often finds sentences with unmistakable past-time reference without any of the ‘past-tense’ markers, such as the following:

- (15) a. Bokar (Ouyang 1985: 69)  
*ko-m ŋoː kajum tʃen moŋ*  
 3SG-OBJ 1SG formerly know be not  
 ‘I did not know him/her before.’  
 b. Padam-Mising (Lorrain 1985: 548)  
*sijo ŋoː juː maŋ*  
 last night 1SG sleep be not  
 ‘I did not sleep last night.’

Thus, the relevant contrasts appear to be aspectual ones between the imperfective, perfective, and perfect. The imperfective markers seem to be built on the existential copula (< Proto-Tani \**duŋdak*), on which the progressive is often based (e.g. Bokar *də-na*). For examples of perfective marking, see further on in this chapter.

Morphological passives involving demoted agents are also attested, exemplified earlier in (12) where the passive is explicitly indicated by a suffix *-ko* in Apatani.

Modal distinctions are made which have to do with the epistemological status of reported events. Special verb marking is required in statements about other people's inner sensations and mental activities not objectively observable, as in Bengni:

- (16) a. *ɲuː kanoː-pa*  
 1SG hungry-PFV  
 'I have become hungry.'  
 b. *tatuuk kanoː-puu-diː (\*kanoː-pa)*  
 Tatek hungry-PFV-QUOT  
 'Tatek has become hungry.'

The Tani verb also carries a person-based contrast comparable to the 'conjunct–disjunct' distinction operating in Tibetan and Kathmandu Newari, where one set occurs with first person subjects in statements, second person subject in questions, and in complement clauses of *verba dicendi* when the complement and main clause subjects are coreferential, and another set in all other contexts (DeLancey 1992: 39), which is interpreted as a 'self-person vs other-person' opposition in Sun 1993b: 955–6. This is illustrated in (17–18):

- (17) Nishing (Das Gupta 1969: 20–1)  
 a. *ɲo tɔla-m te-təne*  
 1SG 3S-OBJ ask-PFV:self person  
 'I asked him.'  
 b. *tani tɔla-m te-pəne*  
 Tani 3SG-OBJ ask-PFV:other person  
 'Tani asked him.'

- (18) Nishi (Chhangte 1990: 4)
- |             |               |              |
|-------------|---------------|--------------|
|             | self person   | other person |
| Progressive | <i>du-den</i> | <i>du-do</i> |
| Perfective  | <i>ten</i>    | <i>num</i>   |

Some languages also exhibit more typical evidential phenomena. Thus, Bokar differentiates direct–indirect evidentials through distinct perfective markers *-gəɲe* and *-pa(na)*:

- (19) a. *tamə səku-ko rəː-gəɲe*  
 Tame horse-INDEF buy-PFV:direct  
 'Tame bought a horse (I saw it).'  
 b. *tamə səku-ko rəː-pana*  
 Tame horse-INDEF buy-PFV:indirect  
 'Tame bought a horse.'

The perfective markers in other languages may reflect the same evidential contrast, e.g. Padam-Minyong Adi *duŋ* (direct), *tuŋ* (indirect) (Roy 1960: 47–8).

## 4 SYNTAX

### 4.1 Syntactic relations

Tani nominal case marking follows a nominative–accusative pattern. Thus, the unmarked subject relation encompasses the single intransitive argument and the actor of a transitive verb, while the (definite) transitive undergoer is explicitly marked by the object case (e.g. Bokar *-me~-fiəm*).

## 4.2 Constituent order

Morpheme order in compounds is predictable: noun–noun and verb–verb compounds are head-final; whereas noun-adjective compounds are modifier-final. Inside a noun phrase, genitive nominals, demonstrative adjectives, as well as relative clauses precede the head noun, whereas definiteness markers and numeral phrases follow the head noun. A striking characteristic of demonstrative adjectives is they can occur on both flanks of the head (e.g. Bengni *si: əki si:* ‘this dog’). In the western languages the structure of the numeral phrase is more complex, as in Bengni: head+classifier+numeral+numerical particle. The classifier is repeated even before attributive adjectives, see (5) above. The preferred order of attributive adjectives varies from language to language; some languages (e.g. Mising, Padam) allow both orderings.

Of the major clausal constituents, the verb always occurs in the clause-final position. While the unmarked order of the core arguments is agent > recipient > patient, pragmatically conditioned ordering alternatives are possible. Adverbs, especially monosyllabic ones, tend to follow the verb.

## 4.3 Major sentence types

Indicative sentences are optionally marked by declarative illocutionary-force particles, such as *jeter* in the following Bokar sentence (Ouyang 1985: 58):

- (20) *topo: lo: hum lə:me: jeter*  
 buckwheat day three plant-finish DECL  
 ‘To finish planting the buckwheat takes three days.’

Copulas are often left unexpressed in sentences with nominal predicates (21a–b):

- (21) a. Bokar (Ouyang 1985: 44)  
*məni megam (hə:)*  
 3DU hunter be  
 ‘They two are hunters.’
- b. Apatani (Abraham 1985: 104)  
*nuhu nu-ka ude*  
 which 2SG-GEN house  
 ‘Which one is your house?’

Existential copulas are firmly established in the grammar. They occur not only in existential sentences but also in sentences denoting possession, in which the possessor argument often takes the genitive case. In certain languages, existential copulas exhibit functional specialization. Thus Bengni has *do:* and *ka:*, the latter is a negative polarity item collocating only with *ma:* ‘to be not’. Furthermore, Apatani existential copulas have innovated a kind of nominal-classifying function, with a three-way lexical differentiation based on the ‘posture’ of the objects whose existence is predicated. Observe the following examples (Abraham 1985: 70–1):

- (22) a. *so: mju da*  
 here person exist:standing posture  
 ‘Here stands a man.’
- b. *so: mju ako du*  
 here person INDEF exist:sitting posture  
 ‘Here sits a man.’

- c. *so: mju ako do*  
 here person INDEF exist:lying posture  
 'Here lies a man.'

As in many neighbouring Tibeto-Burman groups, the negator (reflecting Proto-Tani *\*maj*) is structurally a postposed auxiliary verb (see Section 3.2). There are no separate existential negators, nor is there differentiation along aspectual lines.

Various non-declarative illocutionary forces are also signalled by means of particles. The Bokar interrogative, imperative, and exhortative sentences in (23a–e) are typical (Ouyang 1985):

- (23) a. *no: o: tuŋ nuŋ-da hə:*  
 2SG liquor drink want-IMPV Q  
 'Do you want to drink liquor?'  
 b. *no: ake: do: dəbo je ifi tuŋ-dəbo je*  
 2SG rice eat-FUT Q water drink-FUT Q  
 'Will you eat rice, or will you drink water?'  
 c. *əmə mo:mit to*  
 fire CAUS-go out IMP  
 'Put out the fire!'  
 d. *no: muku tuŋ joka*  
 2SG smoke drink PROH  
 'Don't smoke!'  
 e. *ŋolu: laku in ɲe:dʒu*  
 1PL together go EXH  
 'Let's go together!'

#### 4.4 Clause coordination

Complex events are most commonly expressed by a chain of loosely organized clauses linked by sequence linkers. In the Bengni clause chain below, the sequence of clauses mirrors the temporal ordering of the individual events; notably, aspect marking is not expressed in the medial clauses:

- (24) *ŋu: dugam du-ɲa: dula udʒi: ku:luk-dula arjap*  
 1SG food eat-finish-LINK clothes put on-LINK door  
*nuŋ-kor-dula ni:ko lin-punu:*  
 push-open-LINK outside exit-PFV:EVID  
 'I finished eating, put on (some) clothes, pushed the door open, and went out.'

#### 4.5 Clause subordination

Grammatically dependent clauses include nominalized, relative, and adverbial clauses. Nominalized clauses occupy a central place in Tani syntax, on which relative, cleft, and purposive clauses are built. In many languages the subject in the nominalized clause is marked with the genitive, shown by this Padam-Mising example (Lorrain 1907: 513):

- (25) *bu-k galuk gə-nam Ø də-m la toka*  
 3SG-GEN coat wear-NOMZR:OBJ that-OBJ take IMP  
 'Take the coat which he is wearing.'

The nominalized clause in (25) functions as an internally-headed relative clause. The basic relativization strategy, however, seems to be gapping inside the relative clause, illustrated below by another Padam-Mising example (Lorrain 1907: 513):

- (26) *no-k*       $\emptyset$  *ge-nam*                      *ami*   *də*   *ai*      *maŋ*  
 2SG-GEN      scold-NOMZR:OBJ   person   that   good   be not  
 ‘That man whom you scolded is bad.’

It is in Apatani that the most elaborate marking on relative clauses is found. Importantly, the genitive marking has become an obligatory feature, even where the deleted noun in the relative clause is the subject. Observe the ‘stranded’ genitive marker below (Abraham 1985: 131):

- (27) *ŋo*      *su-mi*       $\emptyset$ *ka*      *pa-nubo*                      *mju-mi*      *kapa-to*  
 1SG      cattle-OBJ    $\emptyset$ -GEN   kill-NOMZR:A   person-OBJ   see-PFV  
 ‘I saw the person who killed the cattle.’

Purposive clauses are marked by the nominalizer *-na* in Bokar (28a); in some other languages, however, special purposive markers are used (28b–c):

- (28) a. Bokar (Ouyang 1985: 51)  
*məlu*   *topo:*      *lə:-na*                      *in-dəbo*   *je*  
 3PL   buckwheat   plant-NOMZR   go-FUT   Q  
 ‘Will they go and plant buckwheat?’
- b. Apatani (Abraham 1985: 134)  
*mólu*   *sinima*   *ka-tepa*      *i-ne*  
 3PL   movie   see-PURP   go-PFV  
 ‘They went to see the movie.’
- c. Padam-Mising (Lorrain 1907: 471)  
*ati:firi*      *re-kapə*      *gi-duŋ*  
 something   buy-PURP   go-IMPV  
 ‘I am going in order to buy something.’

As a notable Tani trait, temporal and circumstantial adverbial clauses are marked with the object case:

- (29) Bokar (Ouyang 1985: 71)  
*ifi*      *tuŋ-ja-me*                      *aruŋ*   *du-nam*                      *mir-ha-m*                      *mitpen*  
 water   drink-when-OBJ   well   dig-NOMZR:OBJ   person-DEF-OBJ   forget  
*moŋ-bo*  
 be not-FUT  
 ‘When drinking water, we will not forget those who dug the well.’

Some other types of adverbial clauses are marked by specific markers, such as these conditional markers: Apatani *koda*, Bokar *banuŋ*, Bengni *lukulur...lu:*.

## ADDITIONAL ABBREVIATIONS

EXH	exhortative
PART	particle
QUOT	quotative



## REFERENCES

- Abraham, P.T. (1985) *Apatani Grammar*, Mysore: Central Institute of Indian Languages.
- Benedict, P.K. (James A Matisoff, contributing editor) (1972) *Sino-Tibetan: A Conspectus*, Cambridge: Cambridge University Press.
- Chhangte, T. (1990) 'Nishi grammar sampler', paper presented at the 23rd International Conference on Sino-Tibetan Languages and Linguistics, University of Texas Arlington, Texas, October 5–7.
- (1992) 'Phonology of some Nishi (Dafla) dialects', paper presented at the 25th International Conference on Sino-Tibetan Languages and Linguistics, University of California, Berkeley, California, October 14–18.
- Das Gupta, K. (1963) *An Introduction to the Gallong Language*, Shillong: North-East Frontier Agency.
- (1969) *Dafla Language Guide*, Shillong: Research Department, North-East Frontier Agency.
- DeLancey, Scott (1992) 'The historical status of the conjunct/disjunct pattern in Tibeto-Burman', *Acta Linguistica Hafniensia* 25: 39–62.
- Lorrain, J. Herbert (1907) *A Dictionary of the Abor-Miri Language*, Shillong: Eastern Bengal and Assam Secretariat Printing Office.
- Ouyang, Jueya (1985) *Luobazu Yuyan Jianzhi: Bengni-Boga'eryu (Brief Description of a Language of the Luoba Nationality: the Bengni-Bokar Language)*, Beijing: Nationality Press.
- Roy, S. (1960) *Aspects of Padam-Minyong Culture*, Shillong: North-East Frontier Agency.
- Simon, I.M. (1978) 'The largest language group of Arunachal', *Resarun* 4.1: 7–13.
- Sun, Jackson T.-S. (1993a) *A Historical Comparative Study of the Tani (Mirish) Branch of Tibeto-Burman*, unpublished PhD dissertation, University of California, Berkeley.
- (1993b) 'Evidentials in Amdo Tibetan', *Bulletin of the Institute of History and Philology, Academia Sinica* 63.4: 945–1001.
- Weidert, Alfons (1987) *Tibeto-Burman Tonology*, Amsterdam and Philadelphia: John Benjamins.

**PART 9**

---

**GYALRONG LANGUAGES**

---



# COGTSE GYARONG

*Yasuhiko Nagano*

## 1 INTRODUCTION

Gyarong (*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 morphology.

It is true that Gyarong has a lot of words very similar or sometimes identical to WT, but almost all of those are cultural words and are likely to be loanwords. On the other hand, the basic lexical items such as body parts, adjectives, and verb roots are more comparable to Benedict's PTB forms. It is likely, then, that Gyarong should not be sub-classified with the Tibetan branch.

As for the affixation system, its whole structure is quite reminiscent of that of Jinghpaw, Rawang, and PTB, while its pronominalization system is somewhat parallel to certain Himalayish languages. However, we have no clear-cut claim about its genetic relationship to specific TB branches yet. As far as the very sophisticated system of pronominalization is concerned, it is apparently not a direct reflex of the PTB system but is probably a later development.

### 1.1 Distribution

The majority of Gyarong people inhabit the Aba (=阿坝: WT *rnga ba*) Tibetan and Qiang Autonomous State and the Gantse (=甘孜: WT *dkar mdzes*) Tibetan Autonomous State of Sichuan Province. The exact number is unknown since they are not officially recognized as an independent nationality in China but are categorized as of Tibetan nationality. But, the number of native speakers of Gyarong is estimated at least at 150 thousand. Unlike Tibetan, their native language is not used in education in this region, and consequently, bilingualism with Gyarong at home and Chinese in public places has become rather dominant.

### 1.2 Dialects

It is traditionally said that there are eighteen dialects. This classification came from the historical division of the area into eighteen administrative zones in the twelfth century, with a separate dialect spoken in each zone. On the basis of modern data, however, the language seems to have three dialects; northern, eastern, and western. Datsang (WT *da tshang*; 大藏) is in the centre of the northern area where about ten thousand 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 fifty thousand people. The eastern area covers a relatively vast region including Cogtse (WT *lcog rtse*; 卓克基), Barkham (WT *'bar khams*; 马尔康), Suomo (WT *so mang*; 梭磨), Tshakunao (=Tshako, WT *bkra shis gling*; 雜谷脑), Lishan

(WT *lis rdzong*; 理県), 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*, 麻窩). There are around eighty thousand Gyarong people in this area.

These dialects are classified according to the initial consonant clusters and the behaviour of pronominals in VPs. However, not all the descriptions of dialects are complete, so this classification remains tentative. 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

p	py	t	ṭ	ky	k	ʔ
ph	phy	th	ṭh	khy	kh	
b	by	d	ḍ	gy	g	
		ts		č		
		tsh		čh		
		dz		ǰ		
		s		š		h
		z		ž		fi
m		n		ñ	ṇ	
		l	r			
		l̥				
w			y			

Phonetic values of phonemes are:

Py-	palatalized P
ṭ-	retroflex
Ky-	palatalized K-. Palatal stops [c-, ɟ-] may appear as the allophones. Chinese linguists often describe these as [cç-, cç'-, ɟj-]
č, ǰ	palato-alveolar affricates
š, ž	alveolo-palatal fricatives
fi	voiced glottal fricative
ñ	palatal nasal
ṇ	velar nasal
l̥	voiceless alveolar lateral fricative.

In addition to the above, we have /N/ as a phoneme of nasalization, which occurs before the stops and affricates to nasalize them, as well as at the C<sub>f</sub> position to nasalize the preceding vowel.

### 2.2 Vowels

The vowels are /a, i, u, e, o, ə/. /i/ usually appears as [ɪ] and /e/ as [ɛ]. [u] is an allophone of /u/ [u]. Some scholars also have /a/ and /e/, said to distinguish verb stems.

### 2.3 Tone

Tone is not distinctive. Some say that there exist several minimal pairs, but they are in fact homonymous when they were checked with informants.

## 2.4 Syllable structure

The syllable structure is (C)<sub>1</sub>(G)V(C<sub>1</sub>)(s), where the parenthesized portions are optional. (C) can be occupied by p-, t-, k-, r-, l-, s-, š-, m- or n-. Among these, p-, t-, k-, s- and š- get voiced when followed by a voiced C<sub>1</sub>. All the consonants shown in 2.1 except for fi- can occur at C<sub>1</sub>. (G) stands for glides, -r-, -l-, -w- and -y-. C<sub>1</sub> is either -p-, -t-, -k-, -ʔ-, -č-, -s-, -fi-, -m-, -n-, -ñ-, ñ-, -l-, -r-, -w-, -y or -N.

There are syllables that could be interpreted as vowel-initial, but these are interpreted a ʔV here.

## 3 MORPHOLOGY AND MORPHOSYNTAX

The most characteristic features of Gyarong morphology and morphosyntax are the productive affixes, which specify a variety of grammatical functions. The following description is, therefore, focused on them.

### 3.1 Nouns

The majority of nouns are marked by the prefix *tə-*, for example, *tə-rmi* ‘man’, *tə-yak* ‘hand’. This prefix may appear 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 *pak tə-rgi* ‘one pig’, while *tə-pak* ‘a pig’ is a normal expression. For example, ‘three pigs’ is *tə-pak kə-sam*, where *sam* ‘three’ is prefixed by *kə-*.

#### 3.1.1 Syntactic order

The syntactic order of noun and noun-qualifier(s) is as follows:

this umbrella	<i>štə wə-dek</i> (this of-umbrella)
two rooms	<i>kho kə-ñes</i> (room two)
these four pen	<i>štə wə-sñekə kə-wdi</i> (this of-pen four)
a big tent	<i>sgar ku-de</i> (tent big)
this tasty gruel	<i>štə wə-pepe kə-mem</i> (this of-gruel tasty)
these pretty girls	<i>tə-mi kə-snaña kə-sam</i> (woman pretty three)
these three black pencils	<i>štə wə-žasñu kə-na kə-sam</i> (this of-pencil black three)

#### 3.1.2 Nominalizers

*ʔa-* is prefixed to the direction markers to nominalize them; *ʔa-ta* ‘(the) above’, ‘the upper place’, against *ta-* ‘upward’, *ʔa-na* ‘down(N), the lower place’ against *na-* ‘downward’. *ʔa-* also implies ‘the place near the speaker’. In contrast to *ʔa-*, *ha-* nominalizes direction markers to stand for ‘the place (and direction) over there’. Thus, *ha-ta* ‘the upper place over there’.

If *to-* is prefixed to the infinitive of a verb, it stands for the semantic patient of the verb. Thus, *to-ka-žu* means ‘accused person’ against *ka-žu* ‘to accuse’.

A similar nominalizer is *sa-*, which indicates ‘a place/utensil to do something’. Examples are *sa-gyup* ‘bedroom’ vs *ka-gyup* ‘to sleep’, *sa-top* ‘hammer’ vs *ka-top* ‘to hit’.

#### 3.1.3 The gender/number markers

The gender markings on nouns are *-pho* (male) and *-mo* (female). The number markings on nouns are *-ñjas* (dual) and *-ñe* (plural).

### 3.1.4 Relative clause

The general structure of relative clauses is VP non-final + *wə*-noun, where the VP non-final consists of (aspect marker/directive)-*V*<sub>inf.</sub> -pronominal suffix. As for the aspect markers and directives, see Sections 3.4.2 and 3.4.3. Pronominal affixes are described under 3.4.5. *V*<sub>inf.</sub> is a verb root prefixed by *ka-*. Thus,

*ta-pu θ-ka-ndzafi ma nə-rga-w wə-za*  
child INF-eat NEG like-3SG 3-food  
'The food that a child does not like to take.'

*mi-šer phe ndzokhan wu-nguy to-ka-nə-č'a-ra-ñ wə-tha*  
yesterday library of-in PFT(up)-INF-read-1SG 3-book  
'The book that I read in the library yesterday.'

*ka-* sometimes disappears when an aspect marker occurs. For instance,

*mi-šer ta-θ-ki-ñ wə-tha tə na-pši-ñ.*  
yesterday PFT(up)-θ-buy-1SG 3-book NP boundary marker PFT (down)-lose-1sg.  
'I lost the book that I bought yesterday.'

## 3.2 Pronouns

### 3.2.1 Personal pronouns

Independent personal pronouns are as listed below. Pronominal affixes in verb phrases (see Section 3.4.2) and possessives (see below) are the reflexes of the pronouns.

	Singular	Dual	Plural
<b>First</b>	<i>ña</i>	<i>čhi-gyo</i> <i>yi-Njo</i> (exclusive)	<i>yi-gyo</i> <i>yi-ño</i> (exclusive) <i>yo</i>
<b>Second</b>	<i>na-gyo</i> <i>nə-yo</i> (honorifics)	<i>ji-gyo</i>	<i>ñi-gyo</i> <i>ño</i>
<b>Third</b>	<i>wu-yo</i> <i>ñi-yo-ñe</i> (honorifics) <i>mə</i>	<i>wu-yo-jis</i>	<i>wu-gyo-ñe</i> <i>wu-yo-ñe</i> <i>ñi-yo-ñe</i>

FIGURE 29.1 INDEPENDENT PERSONAL PRONOUNS

Among these, only older people use *yo* for first person plural, *ño* for second plural and *mə* for third singular.

*nə-ña* is sometimes used for first person dual (inclusive), which seems to be a newer formation, 'you and I'.

#### 3.2.1.1 Possessives are:

*tə-mo* mother  
*ñə-mo* my mother  
*nə-mo* your mother  
*wə-mo* his/her mother  
*yə-mo* our mother (exc. and inc.)

<i>ñə-mo</i>	your/their mother
<i>ŋjə-mo</i>	the mother of ours/yours/theirs (duals)

### 3.2.1.2 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 those mentioned above. Thus,

<i>tə-pa</i>	father
<i>ña ñə-pa</i>	my father
<i>na nə-pa</i>	your father
<i>wu-yo wə-pa</i>	his/her father
<i>yì-gyo yə-pa, yo yə-pa</i>	our father
<i>ñi-gyo yə-pa, yo yə-pa</i>	your father
<i>wu-yo-ñe ñə-pa</i>	their father
<i>čhi-gyo ŋjə-pa</i>	the father of us two
<i>ji-gyo ŋjə-pa</i>	the father of you two
<i>wu-yo-jis ŋjə-pa</i>	the father of them two

### 3.2.2 Demonstrative pronouns

Principally, *štə* ‘this’ and *wətə* ‘that (over there)’ are distinguished. *štə* seems to originate from \**šətə*, where \**šə-* means ‘near’. The component *-tə* in the both pronouns is treated as a referential indefinite demonstrative pronoun.

### 3.2.3 Interrogatives

The main interrogatives are as follows;

<i>sə</i>	who
<i>thə</i>	what
<i>kə-r̄ti</i>	when
<i>kə-ce</i>	where
<i>thə-ni</i>	how
<i>thə-ste</i>	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-čem* ‘small’, *kə-pram* ‘white’. Those borrowed from WT do not bear *kə-*. Examples: *ljan-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-štak* ‘cold’, *kə-sə-mo* ‘wicked’, and *kə-mə-skru* ‘pregnant’. These affixes seem to behave the same as those in the verb phrases. See Section 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-šta-štak* ‘very



cold' against *kə-mə-štak* 'cold'. If the root ends with a consonant, it disappears in the first component of the reduplication.

### 3.3.3 Comparative and superlative

*ndzok* prefixes adjective roots to indicate the comparative and *stui-* indicates the superlative. For instance, *kə-skren* 'long', *ndzok-kə-skren* 'longer', and *stui-kə-skren* 'longest'.

### 3.3.4 Modification

When an adjective modifies an NP, the modifier is placed after it; *me-tok kə-wə-rne* '(flower-red) red flower'.

### 3.3.5 Composite adjectives

Composite adjectives are formed in two ways: NP + adjective and verb root + adjective. *tə-lto kə-mo* 'to be hungry', where *tə-lto* means 'stomach' and *kə-mo* 'empty'.

### 3.3.6 Adjectives in the predicate

When an adjective is the predicate, the adjective behaves just as verbs do in terms of pronominalization and aspect. Instances of pronominals with adjectives are shown below. The shape in the braces shows the underlying form of the preceding word.

<i>kə-mšor</i>	beautiful
<i>na mšor̄n̄r {θ-mšor-n̄}</i>	I am beautiful.
<i>na-gyo tə-mšor {tə-mšor-θ}</i>	You are beautiful.
<i>ji-gyo tə-mšor-Nč̄ {tə-mšor-Nč̄}</i>	You two are beautiful.

## 3.4 Verb phrases

Gyarong sentences are either simple or compound. The former includes one VP that is necessarily the VPfinal, while the latter has any number of VPnon-finals and a VPfinal. The structures is illustrated schematically as

$[(NP) + VPnon-final]^n(\text{particle})[(NP) + VPfinal] (\text{aux})$ ,  
where *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:

VPfinal → (ka)-(ke)-P1-P2-(P3)-ROOT-(s)-S1  
where the parenthesized parts are optional.

(ka) 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 in this language, and actually this prefix does not point to a specific time but refers to a relatively remote stage.

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. Manners include causative, progressive, transitivizer and some others.

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 verb is in 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:

	Morphological component	Function	Semantic class
S y n t a c t i c o r d e r	<i>ka</i>	signals VP	accompanist
	<i>ke</i>	tensifies aspect	aspectuals
	P1	tells if it is done or tells direction of act	perfectives  locationals
	P2	tells who to whom	pronominals
	P3	tells manner of act	specifics

FIGURE 29.2 MORPHOSYNTAX OF PREFIXES

3.4.1.2 Layers of prefixation

Thus, the prefixes occur as follows:

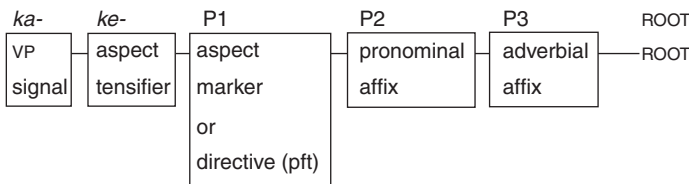


FIGURE 29.3 LAYERS OF PREFIXES

On the semantic level, the chart can be interpreted as follows:

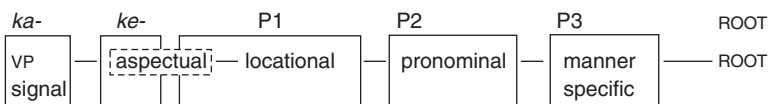


FIGURE 29.4 SEMANTIC LAYERS OF PREFIXES

It can be claimed, therefore, that the more remote from the root, the less concrete 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  $\emptyset$ - and perfective by  $nə$ -. Thus,

- (1) *na diñ { $\emptyset$ -dit-ñ}*.  
1SG  $\emptyset$  -give-1SG  
'I am going to give (it).'
- (2) *na nə-diñ {nə-dit-ñ}*.  
1SG PFT-give-1SG  
'I have given (it).'
- (3) *na nə-mñak ro { $\emptyset$ -ro}*.  
1SG 1SG-eye  $\emptyset$  -wake  
'I wake up.'
- (4) *na nə-mñak nə-ros {nə-ro-s}*.  
1SG 1SG-eye PFT-wake-PFT.  
'I have awakened.'
- (5) *ñi-gyo tə-rgyap tə-sarñ { $\emptyset$ -tə-sar-ñ} mo ños*.  
2PL (HON) marriage 2PL-marry-2PL INTERR AUX  
'Are you going to marry?'
- (6) *ñi-gyo tə-rgyap nət-sarñ {nə-tə-sar-ñ} mo ños*.  
2PL (HON) marriage PFT-2PL-marry-2PL INTERR AUX  
'Have you got married?'

The perfect marker  $nə$ - is often replaced by a direction marker. But, the following verbs conventionally require  $nə$ -: *khyop* 'to cut', *khak* 'to peel', *ki* 'to borrow', 'to buy', *krok* 'to scratch', *krot* 'to cut', *kya* 'to untie', *mčhi lat* 'to bite', *mžit* 'to fall', *phot* 'to break', *physis* 'to wipe', *pšit* 'to drop', *pya* 'to take', *sat* 'to kill', *skyo* 'to write', *ta* 'to take off', *yo* 'to rob'.

#### 3.4.2.1 The aspect tensifier *ke*-

The *ke*- 'tensifies' the aspect, moving the time of action to a more remote stage. Compare the following four:

- (7) *na pyañ { $\emptyset$ -pya-ñ}*.  
1SG  $\emptyset$ -take-1SG  
'I am going to take (it).'
- (8) *na ke-pyañ {ke-pya-ñ}*.  
1SG TSF-take-1SG  
'I will take (it).'
- (9) *na nə-pyañ {nə-pya-ñ}*.  
1SG PFT-take-1SG  
'I have taken (it).'



- (14) *ña mi-šthis to-pšin {to-pšit-ñ}*.  
1SG saliva up-spit-1SG  
'I spat upward.'

(13) is a normal statement, whereas (14) is a proverb that means 'The wheel has come full circle'.

- (15) *ña ñə-ŋgla to-khyeñ {to-khye-ñ}*.  
1SG my-step up-walk-1SG  
'I walked.'

- (16) *ña ñə-ŋgla no-khyeñ {no-khye-ñ}*.  
1SG my-step down-walk-1SG  
'I walked step by step.'

(15) is a normal utterance, while in (16), 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'. Thus,

- (17) *wu-yo-ŋis ña-ñə-mki kow-ptsirč {ko-wu-ptsir-č}*.  
3DL my-neck coiling up-inv.-wring-3DL  
'They (two) wrung up my neck.'

- (18) *čhi-gyo tə-tak ko-pač {ko-pa-č}*.  
1DL weaving coiling up-do-1DL  
'We (two) have woven.'

- (19) *ña ti-gi ko-wa-stsheñ {ko-wa-stshe-ñ}*.  
1SG hot 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:

- (20) *wu-yo-ñe ña-ñə-rpak rew-ñthen {re-wu-ñthen-ñ}*.  
3PL my shoulder back-INV-pull-1SG  
'They have pulled my shoulder.'

- (21) *wu-yo-ñe ña-ñə-rpak row-ñthen {ro-wu-ñthen-ñ}*.  
3PL my shoulder front-INV-pull-1SG  
'They have pulled my shoulder.'

In (21), the agents and the speaker are in a face-to-face position and the speaker's shoulder was pulled towards the agents' noses. In (20), 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.

- (22) *ʃtə wu-mi-tə re-diŋ̃ {re-dit-ŋ̃}*.  
 this red-NOMINALIZER 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

Gyarong people are so sensitive to the social order 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) across a hearth from the seat of honour. This socio-cultural distinction is not always observed in Gyarong in general. Lin (1993) describes still another system of direction markers.

#### 3.4.3.3 Other markers

There are two more markers, *ne-* and *yi-*. *ne-* implies the movement of ‘to go and return’, ‘to get back’. For instance, *ne-ya* ‘to return’ is a compound of *ne-* ‘to get back’ and *ya* ‘to go home’. *yi-* stands for a general movement. ‘to go’ and ‘to come’ require *yi-* unless a specific direction of going and coming has to be indicated. Thus,

- (23) *wu-yo yik-thal {yi-kə-thal-θ} nos*.  
 3SG general movement-3SG-go AUX  
 ‘He has gone.’

A similar, but a slightly extended, usage of *yi-* is observed in an elegant expression for ‘to die’.

- (24) *no-ʃis {no-ʃi-s}*.  
 down-die-PFT  
 ‘He/She died.’
- (25) *ŋi-ʃis {nə-yi-ʃi-s}*.  
 PFT-general movement-die-PFT  
 ‘He/She passed away.’

### 3.4.4 Adverbial affixes

The P3 position is occupied by an adverbial affix that specifies the manner of verb. 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. Gyarong 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 the P3, *sə-*, *ʃə-*, *rə-* and *wa-*, will be described. As for the old vestiges, see Section 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 [E]; if the root has a low/back/rounded vowel, it becomes [U]; otherwise, it remains -ə-.

The following pairs typically illustrate the function of *sə-*.

- (26) *mi-šer tə-rmi ke-ta-key-dzu {ke-ta-kə-yi-dzu}*.  
yesterday men TSF-up(pft.)-3PL-general movement-gather  
'People gathered yesterday.'
- (26a) *ña mi-šer tə-rmi ke-to-sey-dzuñ {ke-to-sə-yi-dzu-ñ}*.  
1SG yesterday men TSF-up(PFT)-CAUS-gather-1SG  
'I assembled people yesterday.'
- (27) *štə wu-ṭha wu-Nguý {wu-Ngu-y} tə-dok ta-ña-kyo-lo {ta-ña-kyo-lo} no-to*.  
this tea of-in-LOC poison up(PFT)-mutual ACT-mix AUX  
'Poison has been mixed in this tea.'
- (27a) *štə wu-smán tə-gi wu-Nguý {wu-Ngu-y} tə-sə-kyo-low {tə-sə-kyo-lo-w}*.  
this of-drug water of-in-LOC 2SG-CAUS-MIX-2sg  
'Mix this drug in the water.'

Examples of the conversion of transitive verbs to causative are as follows:

- (28) *ña ñə-Nga ke-nə-tañ {ke-nə-ta-ñ}*.  
1SG my-cloth TSF-PFT-take off-1SG  
'I took off my clothes.'
- (28a) *ña wu-Nga ke-nə-sə-tañ {ke-nə-sə-ta-ñ}*.  
1SG his-cloth TSF-PFT-CAUS-take off-1SG  
'I undressed him.'
- (28b) *ña wu-Nga ñə-Nḡi nə-sə-tañ {nə-sə-ta-ñ}*.  
1SG his-cloth my-servant PFT-CAUS-take off-1SG  
'I made my servant undress him.'

Besides this productive *sə-*, Gyarong has a transitive/intransitive contrast at the initial position of verb roots. Thus,

<i>ka-Ngyop</i>	to burn (vi)	<i>ka-Nkyop</i>	to burn (vt)
<i>ka-Nbak</i>	to split	<i>ka-phak</i>	to tear
<i>ka-Nglak</i>	to fade	<i>ka-klak</i>	to scour off

The affix *šə-* serves not only to make verbs causative but also to add the meaning of 'to help'. Compare the following three sentences:

- (29) *ña ke-rwas {ke-rwas-ñ}*.  
1SG TSF-rise-1SG  
'I will rise.'
- (29a) *ña wu-yo ke-sə-rwas {ke-sə-rwas-ñ}*.  
1SG 3SG TSF-CAUS-rise-1SG  
'I will raise him.'

- (29b) *na wu-yo ke-šə-rwas {ke-šə-rwas-ñ}*.  
 1SG 3SG TSF-CAUS-rise-1SG  
 ‘I will help him rise.’

The affix *rə-* is the third causative marker. *ka-kšut* means ‘to get out’ while *ka-rə-kšut* means ‘to expel’. Another example is *ka-rə-čhak* ‘to decrease’ from *ka-čhak* ‘few’.

The last causative marker *wa-* functions to transitive adjectives and nouns. Examples are:

- (30) *ñi-gyo ti-gi ke-wa-stsheñ {ke-wa-stshe-ñ} mo nos*.  
 2PL water TSF-CAUS-hot-2PL INTERR AUX  
 ‘Will you boil water?’
- (31) *wa-rgyap gya-roñ na-čhe na-wa-rmow {na-wa-rmo-w}*.  
 his-wife Gyarong PFT-go PFT-CAUS-dream-3SG  
 ‘He dreamt that his wife went to Gyarong.’

#### 3.4.4.2 Mutual act marker

When it occurs before the root, *ñə-* serves to indicate the act is mutual. Thus,

- (32) *wu-yo-jis kew-top {ke-wu-top}*.  
 3DL TSF-3DL-hit  
 ‘They two will hit (somebody).’
- (32a) *wu-yo-jis kew-ñə-top {ke-wu-ñə-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-*. Kin P’eng *et al.* (1957/8) says that *na-* is followed by reduplicated roots, but, in my data, it is not necessarily so.

- (33) *na nə-ra-kroñ {nə-ra-kro-ñ}*.  
 1SG PFT-repetitive act-scratch-1SG  
 ‘I have scratched and scratched.’
- (34) *štə wa-key ko-ho-ke mə-ma ra-skyoñ {ra-skyo-ñ}*.  
 this than nice-adverbializer polite demand repetitive act-write-2pl.  
 ‘Would you please write more nicely than this?’
- (35) *štə wu-rmi-yo ke-kə-na-riñ {ke-kə-na-ri-ñ}*.  
 this 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. Thus,

- (36) *na to-mə-mphañ {mphañ-ñ}*.  
 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-*:

(37) *to-mə-mphat!*  
up-automatic act-vomit  
‘Vomit!’

(37a) *to-mphat!*  
up-Ø-vomit  
‘Vomit intentionally.’

Sentence (37) is of a neutral sense, where the addressee feels sick and the speaker tells him not to counteract his natural physiology. In (37a) 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

*sa-* serves to establish psychological distance from a mental action of the agent. For instance,

(38) *ña-rmo ke-no-sa-pañ {ke-no-sa-pa-ñ}*.  
your-dream TSF-down-OBJECTIVIZER-make-2PL  
‘Please dream.’

(39) *ña wu-mi ke-no-sa-nə-ñan {ke-no-sa-nə-ña-ñ}*.  
1SG his daughter TSF-PFT-OBJECTIVIZER-love(stative)-1SG  
‘I loved his daughter.’

The verb root in (39) is *nə-ña*, which consists of a progressive 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 maker, but little ambiguity occurs because of its position.

(40) *wu-gyo-ñe ña-mñok wu-dza {wu-dza}*.  
3PL my-grain 3PL-eat  
‘They are going to eat my grain.’

(41) *wu-gyo-ñe ña-mñok wu-nə-dza {wu-nə-dza}*.  
3PL my-grain 3PL-PROG-eat  
‘They are eating my grain.’

(42) *wu-gyo-ñe ña-mñok tu-dza {to-wu-dza}*.  
3PL my-grain PFT-3PL-eat  
‘They have eaten my grain.’

(43) *yi-ño ñi-gyo nə-mñok no-nə-dzey {no-nə-dza-y}*.  
1PL(EXC) 2PL your-grain PFT-PROG-eat-1PL  
‘We were eating your grain.’

#### 3.4.4.7 Reflexive marker

*nə-*, identical in shape to the progressive marker, 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’ (Jin Peng *et al.* 1958: 81). Another example is *ka-nə-ña* ‘to love’ (Nagano 1984: 55). This now behaves as a root, but analysis suggests that *nə-* was originally an adverbial affix and has been lexicalized. The exact translation for *ka-nə-ña* will be ‘to love from inside/irresistibly’. According to Jin Peng, *ka-ždar* ‘to be afraid’ may take *-nə-*, which functions to emphasize intransitiveness.

#### 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 participants in the VP. 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 Gyarong, Rawang, Lushai, Qiang and some Himalayish languages, in which pronominal components are indispensable constituents of VPs. ‘Pronominalization’ in this chapter specifies, in most cases, the morphological affixes in the VP which reflect agent and patient/goal/beneficiary agreement. Pronominal affixes appear at 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 intransitive paradigm are

	P2	S1
1SG	( <i>kə-</i> )	<i>-ñi</i>
1DL	( <i>kə-</i> )	<i>-č</i>
1PL	( <i>kə-</i> )	<i>-y</i>
2SG	<i>tə-</i>	<i>-n</i>
2DL	<i>tə-</i>	<i>-Nč</i>
2PL	<i>tə-</i>	<i>-ñi</i>
3SG	( <i>kə-</i> )	<i>-θ</i>
3DL	<i>kə-</i>	<i>-θ</i> (or <i>-Nč</i> )
3PL	<i>kə-</i>	<i>-θ</i> (or <i>-ñi</i> )

The affixes at S1 are recognized to be the remnants of independent personal pronouns (Section 3.2.1). Those are further analysed as *-ñi* being for the first person, *-ñi* for the second person, *-č* for dual and *-y* 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 the third persons is unclear, but this phenomenon may be parallel to that observed in some TB languages in Assam where the first and third person pronominals merge partly.

## 3.4.5.2 The affixes of the transitive paradigm

When there is a set of a person agent and a person patient (or a goal or a beneficiary), the following sets of pronominal affixes (P2-R[oot]-S1) appear:

PTT/BNF/GOA	AGT		
	1	2	3
1SG		<i>kəw-R-ñ</i>	<i>wu-R-ñ</i>
1DL		<i>kəw-R-č</i>	<i>wu-R-č</i>
1PL	<i>ka-R-y</i>	<i>kəw-R-y</i>	<i>wu-R-y</i>
2SG	<i>ta-R-n</i>		<i>təw-R-n</i>
2DL	<i>ta-R-Nč</i>		<i>təw-R-Nč</i>
2PL	<i>ta-R-ñ</i>		<i>təw-R-ñ</i>

PTT	AGT								
	1SG	1DL	1PL	2SG	2DL	2PL	3SG	3DL	3PL
3	<i>θ-R-ñ</i>	<i>θ-R-č</i>	<i>θ-R-y</i>	<i>tə-R-n</i>	<i>tə-R-Nč</i>	<i>tə-R-ñ</i>	<i>θ-R-w</i>	<i>wu-R-θ</i>	<i>wu-R-θ</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>3), *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 *θ-wu-* respectively, where *kə-* and *tə-* are the affixes for the first and second person categories of non-pronominal origin, *-wu-* an inverse marker, and *-a-* a direct marker. Therefore, P2 also shows a patient agreement. *-wu-* occurs only in the 2>1, 3>1 and 3>2 agreements, and *-a-* only in the 1>1pl and 1>2 agreements. This *-a-* seems related to *ʔa-*, a nominalizer of direction markers (Section 3.1.1), that always implies a near deixis.

Unlike the first and second person patient series, the third person patient series show an agent agreement. *wu-*, an inverse prefix, occurs for the 3dl/3pl>3 agreements. *-w* for the 3sg>3 agreement is not the inverse affix but the third person marker. *-w* or *-u* is a widespread marker for the third person category in the TB that appears primarily in the transitive structure.

If a person patient is absent, the affixing system is as follows:

AGT	P2	S1
1SG	<i>θ-</i>	<i>-ñ</i>
1DL	<i>θ-</i>	<i>-č</i>
1PL	<i>θ-</i>	<i>-y</i>
2SG	<i>tə-</i>	<i>-w(u)</i>
2SL	<i>tə-</i>	<i>-Nč</i>
2PL	<i>tə-</i>	<i>-ñ</i>
3SG	<i>θ</i>	<i>-w</i>
3DL	<i>wu-</i>	<i>-θ</i>
3PL	<i>wu-</i>	<i>-θ</i>

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 agt 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 WT where *-s*, which is in complementary

distribution with *-d*, is generally employed in the perfective roots, the *-s* in Gyarong marks the perfective of intransitive process verbs.

- (44) *wu-yo la-sa-s no-kə-skyes {no-kə-skye-s}*.  
3SG Lhasa-ABL PFT-3SG-be born-PFT  
'He was born in Lhasa.'
- (45) *wu-yo-ñe gya-gar-s no-kšis {no-kə-ši-s}*.  
3PL India-LOC PFT-3PL-die-PFT  
'They (went to India and) died.'

If the subject is 1sg in (44), the VP is of the shape *no-skye-ñ*, and if it is 1pl in (45), *nəp-ši-y* 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-khya* 'to be able' and *kə-sa-kha* 'to be difficult', may take *-s* for their perfective. For instance, *ña ka-che no-ra* 'I need to go' vs *ña ka-che no-ra-s* 'I needed to go'.

### 3.4.7 Lexicalization of prefixes

Gyarong 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. Gyarong had several waves of this kind of lexicalization. The following examples are the result of that phenomenon:

to change	<i>s-gyur</i> (vt)		<i>N-gyur</i> (vi)
to turn around	<i>s-kor</i> (vt)		<i>N-kor</i> (vi)
to wind	<i>s-kru</i> (vt)		<i>N-kru</i> (vi)
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>
to see	<i>m-to</i>	to meet	<i>r-to</i>

On the other hand, however, we know in some other dialects that 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 Gyarong seems quite reminiscent of what happened in the remote stages of Tibeto-Burman.

## 3.5 Auxiliary verbs

The auxiliary verbs frequently used are: *ka-khya* 'can', *ka-špa* 'can', 'to be able', *khut* 'may', 'to be ready', *ra* 'need', *ka-yok* 'may', 'to be allowed', *ka-sə-yok* 'to finish', *ndo* 'to exist'. These require the infinitive form of verbs (normally *ka-ROOT*) before them.

- (46) *ña juŋjak ka-pa khyañ {khyā-ñ}*.  
1SG swimming to-do can-1SG  
'I can swim.'

- (47) *ña ku-ru-skat ka-pa špañ {špa-ñ}*.  
1SG Tibetan to-do can-1SG  
'I can speak Tibetan.'
- (48) *tə-gyim ka-ñi ma nə-khut*.  
the-house to-live NEG ready  
'The house is not ready to be lived in.'
- (49) *ña tə-gyim wu-Nguy ka-Ngo məyok*.  
1SG the-house of-in to-enter INTER may  
'May I enter the house?'
- (50) *semḍə ka-pa ma ra*.  
worry to-do NEG need  
'You don't need to worry.'
- (51) *nəgyo ka-nə-ndza mə tə-sə-yok*.  
2SG to-PFT-eat INTER 2SG-finish  
'Have you finished eating?'
- (52) *nəgyo chamdo-y ka-che mə no-ndo-s*.  
2SG Chamdo-LOC to-go INTER PFT-AUX-PFT  
'Have you ever been to Chamdo?'

Besides these auxiliary verbs above, there are auxiliary verbs of statement (attribute markers), *nos* (aff) and *mak* (neg), and existential auxiliary verbs, *ndo* (aff) and *me* (neg). They function as main verbs, but also bear the role of auxiliary verbs. Thus,

- (53) *so-sni tə-mu no-lat je-lat ji, ña ka-che nos*.  
tomorrow rain down-fall not-fall also, 1SG go AUX  
'Whether it rains or not, I will go.'

In (53), the main clause without *nos*, *ña ka-chen {ka-che-ñ}*, is fully grammatical. The combination of aspect marker + *to* may appear in place of *ndo*. It does not take any pronominal affixes.

### 3.6 Numerals

#### 3.6.1 Basic numerals

The basic numerals are:

one	<i>kə-rek</i>	two	<i>kə-ñes</i>
three	<i>kə-sam</i>	four	<i>kə-wdi</i>
five	<i>kə-mño</i>	six	<i>kə-ʈok</i>
seven	<i>kə-šñes</i>	eight	<i>wə-ryat</i>
nine	<i>kə-ŋgu</i>	ten	<i>sgye</i>
eleven	<i>sgye rek</i>	twelve	<i>sgye ñes</i>
twenty	<i>kə-ñes sgye</i>	twenty-two	<i>kə-ñes sgye kə-ñes</i>
hundred	<i>pə-rya</i>	thousand	<i>stoñ-tso</i>

### 3.6.2 Ordinals

The ordinals are loans from WT. Thus,

	<i>Gyarong</i>	<i>WT</i>
first	<i>tañ-bo</i>	dang po
second	<i>ñes-pa</i>	gnyis pa
third	<i>səm-ba</i>	gsum pa
fourth	<i>bžə-ba</i>	bzhi pa
fifth	<i>rña-pa</i>	lnga pa
sixth	<i>tək-pa</i>	drug pa
seventh	<i>bdən-ba</i>	bdun pa
eighth	<i>rjat-pa</i>	brgyad pa
ninth	<i>rgu-ba</i>	dgu pa
tenth	<i>pču-pa</i>	bcu pa

The names of the months require ordinals; for instance, February is *zla-wa ñes-pa* ‘month second’. For comparison, ‘two months’ is *kə-ñes tsə-la*.

### 3.6.3 Frequency

*tə-lok* stands for ‘times’; *kə-sam tə-lok* ‘three times’.

### 3.6.4 Fractions

Fractions are expressed by *tə-šok*; *sgye tə-šok wu-nguy kə-sam tə-šok* ‘three tenths’.

### 3.6.5 Classifiers

Classifiers are fairly abundant. For example, *phyar* ‘a piece of (paper, leather)’, *rgi* ‘a grain of’, ‘a drop of’, *lpek* ‘a piece of (meat, cloth)’, *pyam* ‘a suit of (cloth)’, *rzək* ‘a bundle of’, *nthak* ‘a drop of liquid’, and so on.

## 3.7 Case marking particles

Nominal case markers seem to have been quite alien to Gyarong 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 -y(i)

-y(i) is a locative marker in a most general sense, which means ‘in’, ‘at’, ‘to’, ‘towards’.

- (49) *bi-sni-so pot-pa wu-tha tseng-du-y par wu-nə-lat.*  
 yesterday-tomorrow-day Tibetan of-book Chengdu-LOC photo 3PL-PROG-hit  
 ‘Nowadays Tibetan books are being printed in Chengdu.’

Another function of -y(i) is to link verbs to mean ‘in order to’.

- (50) *ña tə-tha kə-ki-y (kə-)čheñ {kə-čhe-ñ}.*  
 1SG book to-buy-LOC (1SG-)go-1SG  
 ‘I go to buy a book.’

3.7.2 *-s as locative marker*

*-s* is another locative marker, which represents ‘stationary’, instead of ‘shifting or moving’ as implied by *-y(i)*.

- (51) *wu-yo la-sa-s no-kə-skyes {no-kə-skye-s}*.  
 3SG Lhasa-LOC PFT-3SG-be born-PFT  
 ‘He was born in Lhasa.’

3.7.3 *-s as ablative and instrumental marker*

*-s* also functions as an ablative and instrumental case marker. This *-s* makes a complementary distribution with *-ki*, the ergative maker (see Section 3.7.5); when the preceding syllable ends in a vowel, *-s* appears; otherwise, *-ki*.

- (52) *na kyo-mkyo-s nə-paṅ {nə-pa-ṅ}*.  
 1SG Kyomkyo-from PFT-come-1SG  
 ‘I came from Kyomkyo.’
- (53) *kə-zor wə-yi-s*,  
 painful of-reason/cause-by  
 ‘because of pain’

3.7.4 *-i as a genitive marker*

They have *-i* as a genitive marker. This seems to be a loanword from WT *'i-*.

- (54) *wu-tə-ṅ nə-yo-i nos*.  
 that over there-PL 2SG(HON)-GEN is  
 ‘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, the 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.

- (55) *štə wu-rmi-tə-ki štə wu-dzat na-nə-mšor*.  
 this of-man-the-ERG this of-woman PFT-PROG-love  
 ‘The man was loving the woman.’

Just like in WT, this *ki-* marks the instrumental case (see Section 3.7.3).

## 3.8 Negation and question marker

Negation is always expressed by *ma-* and question marker by *mo-*, both of which are followed by VPfinal, VP non-final or auxiliary verb.

## REFERENCES

- Bauman, James J. (1975) *'Pronouns and pronominal morphology in Tibeto-Burman'*, unpublished PhD dissertation, University of California, Berkeley.
- Chang, Kun and Chang, Betty S. (1975) 'Gyarong historical phonology', *BIHP* 46.3: 391–524.
- Delancey, Scott (1981) 'An interpretation of split ergativity and related patterns', *Language* 57.3: 626–57.
- Jun P'eng (金朋鳥) (1949) 'Étude sur le Jyarung', *Han Hiue* 3: 211–310.
- 金朋鳥 *et al.* (1957/58) 「嘉戎語梭磨話的語音和形態」『語言研究』2:123–51, 3: 71–108.
- Jin, Peng, Tan, Kerang, Qu Aitang, and Lin, Xiangrong (1957/58) 'Jianrongu Suomohua de yuyin he xingtai', (The phonology and morphology of the Suomo dialect of Jiarong, parts 1–2), *Yuyan Yanjiu* 2: 71–108.
- LaPolla, Randy (1992) 'On the dating and nature of verb agreement in Tibeto-Burman', *BSOAS* 55.2: 298–315.
- 林向榮 (1993) 『嘉戎語研究』成都：四川民族出版社。
- Lin, Xiangrong (1993) *Jiarongyu yanjiu (Studies on Jiarong)*, Chengdu: Sichuan Minzu Chubanshe.
- Lin, Youjing (2000). Tense, aspect, and modality in the Zhuokeji rGyarong verb, M.A. thesis, National Tsing Hua University, Hsinchu.
- Nagano, Yasuhiko (1984) *A Historical Study of the rGyarong Verb System*, Tokyo: Seishido.
- 瞿靄堂 (1984) 「嘉戎語概況」『民族語文』2: 67–80.
- Qu Aitang (1984) *Jiarongyu gaikuang (A brief description of Jiarong)*, Minzu Yuwen 1984.2: 67–80.
- van Driem, George (1993) 'The Proto-Tibeto-Burman verbal agreement system', *BSOAS* 56.2: 292–334.
- Wolfenden, Stuart N. (1929) *Outlines of Tibeto-Burman Linguistic Morphology*, London: Royal Asiatic Society.
- (1936) 'Notes on the Jyarong dialect of Eastern Tibet', *T'oung Pao* 32: 167–204.



## CAODENG RGYALRONG

*Jackson T.-S. Sun*

## 1 BACKGROUND

Caodeng is a dialect of Sidaba rGyalrong. Sidaba (northwestern), Situ (eastern), and Chabao (northeastern) are the three major dialects of the rGyalrong language. rGyalrong (proper), together with two neighbouring languages, Lavrung and Horpa, comprise in turn the rGyalrongic cluster of languages in the Tibeto-Burman family (Huang 1991, 2001; Sun 2000a,b). Sidaba rGyalrong is composed of two distinct subdialects: Caodeng and Shawu. Most Sidaba speakers live in the three townships Caodeng, Kangshan, and Ribu in Ma'erkang county, Aba prefecture, northwestern Sichuan; small outlier communities (of the Shawu subdialect) are also found at Aba, Rangtang, and Seda counties in the same prefecture. The Caodeng data in this chapter, representing the speech of Gaqiuli village, are from my fieldwork conducted over the past few years.

## 2 PHONOLOGY

The structure of the Caodeng 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 p <sup>h</sup> b <sup>n</sup> b	t t <sup>h</sup> d <sup>n</sup> d	c c <sup>h</sup> ʃ <sup>n</sup> ʃ	k k <sup>h</sup> g <sup>n</sup> g	q q <sup>h</sup>  <sup>n</sup> G
		ts ts <sup>h</sup> (dz) <sup>n</sup> dz	tʂ tʂ <sup>h</sup> (dʒ) <sup>n</sup> dʒ	tʃ tʃ <sup>h</sup> (dʒ) <sup>n</sup> dʒ	
	m  (f) v	n s z r (ʃ) l	ɲ ʃ ʒ	ŋ x ɣ	χ ʁ
	w		j		

An enormous number of initial clusters are attested, including many three-member clusters. Caodeng phonology treats nasal+voiceless stop combinations as consonant clusters, but nasal+voiced stop combinations as unitary prenasalized stops. The consonants permitted as syllable codas are -v, -t, -ɣ, -m, -n, -ŋ, -r, -l, -j, and -s. The coda -r is often devoiced while the -s is usually voiced. The lateral coda -l is distinctly laminodental. Syllables may also take

a glottal-stop coda which may even 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*, *i*, *u*, *e*, *ɛ*, *o*, *ɔ*, *ə*, and *ɐ*. Only two intrinsic complex vocoids *ɛj* and *oj* are noted in the native vocabulary, 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. As shown in (2) below, glottalized verb roots regularly drop the glottal coda in the past and progressive, resulting in a predictable shift of pitch from level (marked by a macron) to falling (marked by a grave accent). Significantly, this alternation of pitch patterns is also extended to verb roots checked by the stop coda *-t*:

- (2) a. *rōʔ*                    to fetch  
       *ɐsɐ-rò*                She/he is fetching.  
       *jɐ-rò*                    She/he fetched.  
   b. *fkōt*                    to carve  
       *ɐsɐ-fkòt*             She/he is carving.  
       *ne-fkòt*                She/he carved.

Unlike in the closely related Shawu dialect where all syllable types show a two-tone contrast (level vs falling), Caodeng syllables with *-t* normally do not carry the falling pitch barring a small number of exceptions (e.g. *ts<sup>h</sup>ət* ‘goat’; *xtət* ‘to be short’). The phenomenon in (2), therefore, represents a grammatically conditioned type of tonality alternation. Caodeng phonology is characterized further 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 the lexical and morphosyntactic levels. The locus of the accented syllable, i.e. the high-pitched syllable immediately preceding the pitch drop, is highly restricted. Stem-final accent (the default situation) 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 turns out to be the penult of the stem. Interestingly, monosyllabic words contrast in accent even though the citation pitch remains identical; the latent accent (marked herein by an apostrophe) materializes only when a preceding morphological element is added, resulting in a pitch drop, e.g. *ʔχser<sup>H</sup>* ‘gold’, *ɐ<sup>H</sup>χser<sup>L</sup>* ‘my gold’; *rŋul<sup>H</sup>* ‘silver’, *ɐ<sup>L</sup>-rŋul<sup>H</sup>* ‘my silver’. A number of segmental phonological processes are found, including vowel assimilation (e.g. *mə-tə-ɔ-ɐ<sup>n</sup>doŋʔ* → *motoɐ<sup>n</sup>doŋʔ* ‘She/he did not beat me.’; *ró<sup>n</sup>di* → *rú<sup>n</sup>di* ‘as soon as she/he rides’), consonant assimilation (e.g. *nɐ-scət-aŋ* → *nɐscənaŋ* ‘I was comfortable’); syncope (e.g. *kətfoʋʔ-sqe* → *kətfosqe* ‘sixty’; *ɐsɐ-ro-aŋ* → *ɐsɐroŋ* ‘I am fetching’), consonant epenthesis (e.g. *n-rewe* → *<sup>n</sup>dʒewe* ‘to hope’; *sə-rem* → *səyrem* ‘to make dry’). For productive phonological alternations at the morphosyntactic level, see further on.

### 3 MORPHOLOGY

#### 3.1 Lexical categories

The major lexical categories in Caodeng are nominals and verbs; 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. *paʔiʔ kə-snojʔ* (crab apple NOMZR-be.deliciously.ripe) ‘deliciously ripe

crab apples'. Few true classifiers (e.g. *-b<sup>h</sup>byeʔ* 'classifier for any long object, such as hairs, sticks, logs, and trees') are used; numerals in most cases directly quantify their head nouns. Many illocutionary-force meanings are coded by various sentence-final particles, compare for example (3a) through (3c) below:

- (3) a. *kəkoʔ kəruʔ ʃteʔ ko*  
 3SG rGyalrong be:EMPH SFP  
 'So (I realize now that) she/he is rGyalrong!'
- b. *kəkoʔ kəruʔ ʃteʔ mu*  
 3SG rGyalrong be:EMPH SFP  
 'She/he is rGyalrong, isn't she/he?'
- c. *kəkoʔ kəruʔ ʃteʔ taʔ*  
 3SG rGyalrong be:EMPH SFP  
 'I am positive that 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)-<sup>n</sup>gə-<sup>n</sup>ge* 'to call each other'; from *<sup>n</sup>ge* 'to call'; *nə-<sup>v</sup>də-<sup>v</sup>də* 'to become better and better', from *vde* 'to be good'. Reduplication also applies to lexical categories other than verbs; examples are certain temporal phrasal idioms (e.g. *k<sup>h</sup>e-s<sup>h</sup>i-s<sup>h</sup>i* 'day after day' and *s<sup>h</sup>i-ku-s<sup>h</sup>i* 'everyday', from *s<sup>h</sup>i* 'day') and complex nominals (e.g. *kə-<sup>n</sup>dʒə(-l<sup>h</sup>ə)-l<sup>h</sup>əʔ* 'parent and child', from *tə-l<sup>h</sup>əʔ* 'child'). Expressives constitute a minor but fascinating lexical category where reduplication is amply exploited, for example *nə-buy-buyʔ* 'in a swarming fashion' and *χ<sup>c</sup>o-<sup>v</sup>ə-χ<sup>c</sup>e ~ χ<sup>c</sup>o-<sup>p</sup>ə-χ<sup>c</sup>o-leʔ* 'in an unshapely or messy way'.

Stem modification via consonantal alternation is utilized in the derivation of a few lexical causatives, e.g. *prət* 'to cause (a rope) to break' vs *<sup>n</sup>brət* '(as of a rope) to break'. A noteworthy example of productive stem alternation in inflection is past stem formation by glottality inversion, e.g. *p<sup>h</sup>jəsʔ* 'to wipe', *p<sup>h</sup>jəs* 'wiped'; *sroŋ* 'to guard', *sroŋʔ* 'guarded'. Equally remarkable is the phenomenon of ablaut. Ablaut plays a role in compound nouns (e.g. *k<sup>h</sup>e* 'house', *k<sup>h</sup>ə-rmiʔ* 'house-name') and compound numerals (e.g. *sqeʔ* 'ten', *sqə-m<sup>h</sup>o* 'fifteen'), but is at full play in verb inflection. One type of ablaut helps form the past stem, e.g. *rp<sup>h</sup>altfem* 'to be in heat (estrus)', *rp<sup>h</sup>altfemʔ* 'was on heat (oestrus)'; *mdεʔ* 'to be level', *mdε* 'was level'. The other ablaut type occurs in certain specific transitive contexts, such as the singular imperative, e.g. *nəji* 'to wait', *nə-nəje* 'You [SG] wait for him/her!'; *pe* 'to 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. *nənəʔ* 'that', *nənəs* 'there'; *tforʔ* 'to be sour', *kətfor* 'sour radish-leaf pickles'. Systematic application of accent modification in Caodeng grammar is shown below in the formation of vocatives (Section 3.4.1) and polar interrogatives (19–20).

Compounds, the majority of which are nominal, may be composed entirely of unaffixed roots, e.g. *qəj-ɣveʔ* 'wheat awn'. As a rule, the nominal prefixes of the initial compound component are retained, resulting in affixed compounds, e.g. *tə-se-roʔ* 'hemp stalk', cf. *tə-seʔ* 'hemp', *tə-roʔ* 'stalk'. Compounds containing more than two components are rare, e.g. *rəp-pyə-ləʔ* 'hen house', literally 'Chinese-fowl-nest'.

### 3.3 Derivational morphology

Generous use is made of derivational processes to form morphologically related words. With nominal roots, suffixation yields gender (male  $-p^ho$ , female  $'-mo$ ) and diminutive ( $-pu$ ) forms, while denominalization is achieved by means of the verbalizing prefixes  $nə-nv-$ , e.g.  $tfape$  'loach',  $nə-tfape$  'to catch loaches';  $tē-ŋv$  'sun',  $nē-ŋv$  'to bask in the sun';  $tə-mja$  'eye',  $nə-mja$  'to aim'. Sometimes actors or instruments are also denominalized in this way, yielding transitive verbs, e.g.  $rtfaχpe$  'robber',  $nə-rtfaχpe$  'to rob'. The other major type of derived words 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.  $kə-n^ngu$  'to be poor',  $kə-n^ngu$  'to become poor',  $kə-xsor?$  'to give birth (as of bovines)';  $kə-sə-sce$  'to give birth (as of humans)'. Participant nominalization (Payne 1997: Section 9.1.2), 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.  $kə-n^ngi?$  'patient',  $kə-tim?$  'rich person', as well as actors of transitive verbs, e.g.  $kə-mətku$  'thief'. The prefix  $kə-$  turns transitive verb roots into undergoer deverbal nouns 'that which is verb-ed', e.g.  $kə-n^dze$  'food';  $v-kə-rge-rv$  'those loved by me'; 'my beloved ones'. There is in addition a manner nominalizer  $tə-$ , e.g.  $o-tə-rtfone?$  'the way she/he dances', and also an oblique (instrument/location/time) nominalizer  $sə-$ , e.g.  $sə-sce?$  'birthplace'; 'birthday';  $sə-n^dzv^hi$  'dining place'; 'instrument with which to eat meals', both of which are attached to past verb stems. Another important type of derivational morphology changes verb valence. Since Caodeng does not have any analytical causative, causativization is achieved predominantly by adding a number of valence-increasing causative prefixes  $sə-$ ,  $sə-$ ,  $wə-$ , or  $f-$ , e.g.  $smi$  'to be cooked',  $wə-smi$  'to cook';  $t^hi$  'to drink',  $f-t^hi$  'to give to drink'; 'to suckle'. On the other hand, a detransitivizing prefix turns a transitive verb into a corresponding intransitive, e.g.  $ntf^he$  'to kill',  $rē-ntf^he$  'to do slaughtering'. Reflexive verb formation (involving the reflexive prefix  $ʃv-$ ) 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 progressive marker as well as inverse marking on the reflexive verb:

- (4)  $oʃi?$   $t^h v-o-ʃv-ʒ^ndu-cə$   
 3SG PROG: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  $-rv$ ), case, and possession. The two basic case forms, the ergative-instrumental  $-kə$  and the locative  $-s$ , are subject to ellipsis where their absence does not cause ambiguity. Nominals bearing the undergoer roles (patient, recipient, theme), furthermore, are usually 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; Sun and Shi 2002). With an abundance of instrument-incorporating verbs such as  $kə-rtov$  'to beat with a rod' and  $kə-ʒrov?$  'to beat with fists', the occurrence of case-marked instruments is

also infrequent. Importantly, the presence of an instrumental argument increases the valence of the verb, marked by the causative prefix:

- (5) *<sup>n</sup>dʒəŋo-kə tə-sə-nə<sup>n</sup>dzɛt<sup>h</sup>i*  
 chopsticks-INST IMP-CAUS-eat.meal  
 ‘Eat your meal with chopsticks!’

The local case marker *-s* expresses a vague local meaning, occurring in locative, ablative, as well as allative contexts. More precise locational senses are conveyed through the construction possessive prefix +relator noun (+ local case), e.g. *jəye ó-ta(-s)* ‘on the book’, literally ‘book its-top-LOC’. The most common relator nouns are *ta* ‘space on/over/above’, *p<sup>h</sup>i* ‘space under/below/beneath’, *naŋ ~ naŋʔ* ‘space inside’, *p<sup>h</sup>iʔ* ‘space outside’, *vzər* ‘space at the corner of’; ‘space beside’, *ʋə(r)* ‘space on the body/vertical side of’, *ʋóri* ‘space in front of’, and *q<sup>h</sup>u* ‘space at the back of’. Several other oblique case markers are possessed relator nouns in form, including *krə* (comitative), *χte/fcə* ‘on behalf of’, *tənkə* ‘for the sake of’, and *p<sup>h</sup>u* (animate goal/source), the last of which is illustrated by these examples:

- (6) a. *sonəm-óp<sup>h</sup>a tɛk<sup>h</sup>u v-nə-tə-k<sup>h</sup>ɛ*  
 Sonam-to cigarette IRR-PFV-2-pass  
 ‘Pass the cigarette to Sonam!’  
 b. *sonəm-óp<sup>h</sup>a tɛk<sup>h</sup>u v-fə-nə-tə-rojʔ*  
 Sonam-from cigarette IRR-go and-PFV-2-fetch  
 ‘Go and get the cigarette from Sonam!’

Standards in comparative constructions take the comparative case markers *fəχtenəʔ* or *fəʋrenə*, as in:

- (7) *vjiʔ krɛfi fəχtenəʔ <sup>n</sup>briʔ-aŋ-cə*  
 1SG Krashi than be.tall-1SG-MED  
 ‘I am taller than Krashi.’

Standards of equative comparison, on the other hand, take the semblative *jərmər*. The vocative case is marked suprasegmentally by accent; e.g. *dzomv* (a woman’s name), *dzómv* (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.

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 possessive prefix:

TABLE 30.1 CAODENG POSSESSIVE PREFIXES

	<i>Sg</i>	<i>Du</i>	<i>Pl</i>
1	<i>v-</i>	<i>nə-</i>	<i>o-</i>
2	<i>tsə-</i>	<i><sup>n</sup>dzə-</i>	<i><sup>n</sup>dzə-</i>
3	<i>jə-</i>	<i>nə-</i>	<i>nə-</i>
Reflexive		<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:

- (8) *tš-lo* milk  
*é-lo* my own (mother's) milk (inalienable)  
*ɐ-tš-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 (21, 23, see below), participant nominals (9a–b), and complex transitive predicates (9c):

- (9) a. *ɐ-kə-j<sup>n</sup>dʒu-cə* *toʔ*  
 1SG:POSS:U-NOMZR:A-accuse-INDEF exist  
 'There is someone who accuses me.'  
 b. *ɐ-ké-j<sup>n</sup>dʒu-cə* *toʔ*  
 1SG:POSS:A-NOMZR:U-accuse-INDEF exist  
 'I have someone to accuse.'  
 c. *nənəʔ vɿɐrme-kə ɐ-mt<sup>h</sup>u t<sup>h</sup>e-lɛt*  
 that monk-ERG 1SG:POSS:U-spell PFV:TR-cast:PAST  
 '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, and evidentiality.

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):

TABLE 30.2 CAODENG PERSON-NUMBER MARKERS ON INTRANSITIVE VERBS

	1	2	3
Sg	<i>v-ɑŋ</i>	<i>tə-v</i>	∅
Du	<i>v-tsə</i>	<i>tə-v<sup>n</sup>dzə</i>	<i>v<sup>n</sup>dzə</i>
Pl	<i>v-jə</i>	<i>tə-v-nə</i>	<i>v-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 happens to be the speaker her/himself, the verb may optionally carry a second index representing the actor. In this construction, more emphasis is put on the actor, as shown by:

- (10) *kəkəʔ-ni-kə ɐjiʔ*  
 3-DU-ERG 1SG  
*t<sup>h</sup>o-sə-wɛʔ-ɑŋ<sup>n</sup>dzə ŋoʔ*  
 PFV:downstream:INV-CAUS-come:PAST-1SG-3DU be  
 'It was the two of them who made me come downstream.'

The category direction marks the relative place of actor and undergoer arguments on the empathy hierarchy (DeLancey 1981). 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 *tʷ-*; 2>1 is marked as inverse as well as by the prefix *kə-*, interchangeable with *tə-*.

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 (upstream–downstream), 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 below:

TABLE 30.3 BASIC CAODENG ORIENTATION PREFIXES

up	down	upstream	downstream	eastward	westward
<i>tə-</i>	<i>nʷ-</i>	<i>lʷ-</i>	<i>tʰʷ-</i>	<i>kə-</i>	<i>nə-</i>

Required on all perfective and imperative verb forms, orientation markers do triple duty, coding aspect, imperativity, as well as spatial orientation *per se*. With non-motion verbs, the selection of collocating orientation prefixes is largely conventionalized.

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 (*ʷ* or *ə*) in the orientation prefixes uniformly to *e*. In direct imperfective sentences with a singular non-first-person actor, moreover, a transitivity marker *-jə* is added to verb stems containing no coda other than the glottal stop (e.g. *jə-tə-qʰʷʷjə* ‘You will dislike it.’); for strong (ablauting) verbs such as *tʰi* ‘to drink’, this transitive marking is optional (e.g. *jə-tə-tʰiɛ(-jə)* ‘You will drink it.’). For yet another transitivity-related distinction, see (12) below.

Caodeng has grammaticalized absolute tense, distinguishing non-past and past stems 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 prefix *jə-*, 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 in this language: an aorist (perfective past) denoting completed actions and processes or resultant states thereof, and an imperfect (imperfective past) denoting ongoing situations existing prior to the moment of speech, marked with *nʷ/ne-*. The contrast is shown in (11a–b):

- (11) a. *forʔ koʔ tə-nʷgi-cə*  
 yesterday 3SG PFV-be.ill:PAST-MED  
 ‘She/he became ill yesterday (AORIST).’  
 b. *forʔ koʔ nʷ-nʷgi-cə*  
 yesterday 3SG IMPFV-be.ill:PAST-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. Several other aspectual categories figure in verb inflection, including two progressives and a habitual. Progressives are of two kinds, high-transitivity and low-transitivity. The high-transitivity progressive prefix *ʋsʋ-* is attached to the progressive stem, a variant of the non-past stem (see (2) above); the low-transitivity progressive prefix *t<sup>h</sup>ʋ-* combines rather with the past stem. Consider:

- (12) a. *stat<sup>h</sup>ʋr lɛju ʋsʋ-pe/\*t<sup>h</sup>ʋ-peʔ*  
 Stathar song PROG:HTR-do  
 ‘Stathar is singing songs.’  
 b. *stat<sup>h</sup>ʋr t<sup>h</sup>ʋ-rlɛjuʔ/\*ʋsʋ-rlɛju*  
 Stathar PROG:LTR-sing:PAST  
 ‘Stathar is singing.’

Events which recur with regularity can be expressed by the bare non-past verb or by adding a habitual prefix *<sup>n</sup>ge-*. This marked habitual form also applies to recurring events in the past:

- (13) *sɣonts<sup>h</sup>es krɛʃi-kə t<sup>h</sup>emgon <sup>n</sup>ge-mtɛʔ nəʔ*  
 formerly Krashi-ERG Tshemgon HAB-see SUB  
*te-rɛmúru-cə*  
 IMPFV-get.angry-MED  
 ‘Krashi used to get angry whenever he saw Tshemgon.’

An experiential meaning is conveyed by the complement-taking verb *rɣiʔ* ‘to taste’; ‘to experience’.

The evidential system consists basically of a mediative *-cə* and a hearsay *-tɛtsə*. 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, and distal imperatives (6; 23), jussives, and optatives, all involving the prefix *ʋ-* (see Sun, forthcoming).

## 4 SYNTAX

### 4.1 Syntactic relations

Neither nominal case marking nor argument cross-referencing, which operate on semantic and pragmatic principles, can give any clue regarding syntactic relations in Caodeng. Participant nominalization (Section 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 *kɛ-*. 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 (Section 4.5.2).

### 4.2 Constituent order

The sequence of morpheme slots in any verb form is: irrealis+negator+go and+low-transitivity progressive+orientation/aspect/imperative+nominalizer+direction+inverse+ reflexive +high-transitivity progressive+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 + subordinator. 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 subordinator at the end of the nominal phrase.

Within a clause, the verb occupies final position. Of the core arguments, recipients normally precede patients, while both kinds of undergoers follow actors except in topicalized or inverse scenarios. Temporal adverbials precede locational ones.

### 4.3 Major sentence types

The indicative sentence is the unmarked sentence type. Fully conjugated equational copulas (positive *ŋoʔ*, negative *maʔ*, and emphatic/contrastive *ʃteʔ*) occur obligatorily in predicate nominal clauses:

- (14) *c<sup>h</sup>eʔ-s*                    *ʃɸniʔ* *vlɸrge* *nɸ-ŋo<sup>n</sup>-dzə*  
 former.time-LOC 3DU monk IMPFV-be:PAST-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 *toʔ* (past stem *to*) and its negative form *mɸ* (past stem *mɸʔ*). In addition to predicating the existence of entities, they express possession via the construction (possessor noun) + possessor prefix – possessed noun + existential verb, as in:

- (15) *nɸjiʔ* *nɸ-lŋaʔ*            *ʃ-toʔ*  
 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 ‘to be present at a certain place’:

- (16) *sonɸm* *pəʃk<sup>h</sup>o* *k<sup>h</sup>ɸnag* *t<sup>h</sup>ɸ-smo-cə*  
 Sonam now home PROG:INTR-be.present-MED  
 ‘Sonam is at home now.’

However, Caodeng does not differentiate its existential verbs further for the purpose of nominal sub-classification, a salient morphosyntactic trait in the neighbouring Qiangic languages.

Topic-comment sentences are extremely popular. The most widely used topicalizing marker in this language is *nəʔ*:

- (17) *nənəʔ* *o-lɸtʃ<sup>h</sup>e*            *nəʔ* *ʃə-kə-ro-nə*  
 that 3SG:POSS-object DET go.and-IMP-fetch-2PL  
 ‘That stuff of his, go and fetch it!’

There are three other topicalizers, *rcənə*, *<sup>n</sup>dɸr*, and *ŋónə*, which are used to express emphasis or contrast.

Negation is expressed by prefixal negators: the imperfective *mɸ-*, the perfective/prohibitive *mə-*, and the high-transitivity progressive, habitual, and non-finite *me-*. The copula *ŋoʔ* and existential verb *toʔ* have distinct negative forms, respectively *maʔ* and *mɸ*. The latter also serves as a propositional negator (i.e. ‘it is not true at all that’):

- (18) *kəkoʔ rɛrɛvʔ ʰbriʔ mɐ-cə*  
 this mountain high exist:NEG-MED  
 ‘This mountain is not high at all.’

The following different 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 accenting the main verb:

- (19) *forʔ sonɐm tɛwaʔ nɛ-tʰi*  
 yesterday Sonam liquor PFV:Q-drink:PAST  
 ‘Did Sonam drink liquor yesterday?’

Polar questions with sentence-focus, also accented, are formed by attaching an interrogative prefix *ə-*.

- (20) *forʔ sonɐm tɛwaʔ nɛ-tʰi ə-ŋo*  
 yesterday Sonam liquor PFV-drink:PAST Q-be  
 ‘Is it true that Sonam drank liquor yesterday?’

Alternative questions show an interrogative particle *soʔ* in 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 with an imperative prefix and, where appropriate, a second person number suffix (dual *-<sup>n</sup>dzə*; plural *-nə*). The negative imperative (prohibitive) takes the imperative negator *mə-*. The prefix *ɐ-* marks imperatives presented for action by a third person, e.g. *ɐ-tə-<sup>n</sup>dzɐ* ‘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 (23) below. The second-person index *tə-*, as well as the 2 > 1 configuration complex 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

Caodeng makes no use of coordinating conjunctions at the clausal level. Instead, one finds in abundance sequences of finite clauses loosely connected by the sequence linker *qʰoʔ* or *qʰónə*:

- (21) *kəkoʔ qa tɛ-<sup>n</sup>di qʰoʔ*  
 3SG hoe PFV:up-carry:PAST LINK  
*komʔ tɛ-cu qʰoʔ*  
 door PFV-open:PAST LINK  
*təjtʃɛ-naŋ kə-rɛmɐ jɐ-fɛʔ*  
 field-in PURP-do.labour PFV-go:PAST  
 ‘He picked up the hoe, opened the door, and went off to do labour in the field.’

Also recorded are verb sequences with no connective morphology in between, indicating simultaneous verbal events (22) or those in immediate temporal sequence (23):

- (22) *pa ɲɛmgon ʧsɐm kɛ-ntʃa kɐ-nlɔʔ cʰɐ*  
 pig river three NOMZR-swim NOMZR-cross be.able  
 ‘Pigs are able to swim across three rivers (at a stretch).’

- (23) *o-kórnɔ tə-zɣi ne-ldɐ-cə*  
 3SG:POSS-brain PFV-spin:PAST PFV-fall.own:PAST-MED  
 'His head spun and he fell down.'

## 4.5 Clause subordination

### 4.5.1 Non-finite subordinate clauses

The two nominalizing prefixes *kɐ-* (dynamic/human) and *kə-* (stative/non-human) yield non-finite nominalized clauses (infinitives) which function as sentential subjects or complements to modal and other complement-taking verbs. The following example exemplifies the purposive or supine use of the nominalized clause in *kə-*, serving as an adjunct to a motion verb:

- (24) *ɐ-kə-q<sup>h</sup>ro ɐ-jɐ-tɐ-wi*  
 1SG-PURP-welcome IRR-PFV-2-come  
 'Come and welcome me.'

There is also the converb (verbal adverb), composed of the prefix *sɐ-* and a reduplicated past stem, which modifies the situation predicated by the main verb:

- (25) *ó-<sup>n</sup>ge sɐ-nɐc<sup>h</sup>i<sup>h</sup>i? noofsɛfsət te-<sup>n</sup>geɽ-cə*  
 3SG:POSS-clothes CVB-wet:REDUP:PAST in.that.way PFV-put.on:PAST-MED  
 'She/he put on her/his clothes while they were still wet.'

### 4.5.2 Finite subordinate clauses

Finite dependent clauses, marked by subordinators, function as arguments (nominalized clauses), argument modifiers (relative clauses), or sentential adjuncts (adverbial clauses).

Propositions are nominalized straightforwardly by adding the generic subordinator *nəɽ*. The nominalized complement clause in (26) is the object of a cognition verb:

- (26) *ɣves-kə k<sup>h</sup>ezeɽ o-fneɽ te-<sup>n</sup>dzeɽ nəɽ ɐji?*  
 hog.badger-ERG dog 3SG:POSS-nose PFV-bite:PAST SUB 1SG  
 'sis-aŋ  
 know-1SG  
 'I know that the hog badger bit the dog on the nose.'

Nominalized clauses do not show evidential marking but are finite in all other respects. Relative clauses are built on nominalized clauses. The most common relativizing strategy is head-marked gapping. In the examples that follow, the relative head *támɛ* 'woman' appears in the relative clause, leaving a gap in the main clause:

- (27) [*for? támɛ nɐ-mti-aŋ*] *∅ nəɽ jɐ-weɽ-cə*  
 yesterday woman PFV-see:PAST-1SG SUB PFV-come:PAST-MED  
 'The woman I saw yesterday has come.'

Dependent-marked gapping is also permitted, as shown in the following syntactic alternative to (27):

- (28) [*for? ∅ nɐ-mti-aŋ*] *támɛ nəɽ jɐ-weɽ-cə*  
 yesterday PFV-see:PAST-1SG woman SUB PFV-come:PAST-MED

The semantic role of the relative-clause pivot is distinctly marked. In (29), the form of the nominalizing prefix *kə-* on the relative verb indicates that the syntactic pivot of the relative clause is the undergoer ‘cup’, rather than some unexpressed actor:

- (29) [tʃóxtse-ta p<sup>h</sup>ərtʃu kə-te]                      Ø nəʔ tə-mjə  
 table-on bowl NOMZR:U-put:PAST SUB IMP:up-take  
 ‘Pick up the bowl that has been put on the table.’  
 \*‘Pick up the bowl that she/he has put on the table.’

Examples (27–29), in fact, reveal a significant generalization in Caodeng syntax; namely, relativization by means of finite nominalization is accessible only to the syntactic pivot (subject) of a clause. This important principle explains the obligatory absence of the undergoer nominalizer *kə-* in (27–28) and, conversely, the relativizability of the undergoer argument in (29), which lacks an overt actor. The principle can be further illustrated with (30), where the undergoer ‘meat’ cannot be relativized with *kə-*:

- (30) qəje-kə je (\*kə-)vəv<sup>n</sup>dze nəʔ vjiʔ-kə  
 tiger-ERG meat (NOMZR:U)-PROG:TR-eat DET 1SG-ERG  
 nə<sup>n</sup>bi-aŋ ŋoʔ  
 PFV-give:PAST-1SG be  
 ‘The meat that the tiger is eating was given by me.’

Transitive scenarios predicated from the viewpoint of an animate undergoer, however, may be expressed as self-caused events, in which the undergoer is promoted to pivot status by becoming a derived agent and thus relativizable by the nominalizer *kə-*. Contrast (31a) with (31b):

- (31) a. o-lŋaʔ kə-kə-rqɔ tɛpe nəʔ  
 3SG:POSS-child PFV-NOMZR:A-embrace:PAST father DET  
 ‘the father who embraced his child’  
 b. ó-pe kə-kə-o-ʃv-sə-rqɔ tɛlŋaʔ  
 3SG:POSS-father PFV-NOMZR:A-INV-REFL-CAUS-embrace:PAST child  
 nəʔ  
 DET  
 ‘the child who got himself embraced by his father’  
 \*ó-pe kə-kə-rqɔ tɛlŋaʔ nəʔ  
 3SG:POSS-father PFV-NOMZR:U-embrace:PAST child DET  
 \*‘the child who was embraced by his father’

Adverbial clauses expressing temporal relations are marked by morphemes showing various temporal meanings in combination with the subordinator *nəʔ*, e.g. *orjagnəʔ(s)* ‘when’, *ʃɔχtenəʔ* ‘before’, *oq<sup>h</sup>unəʔ* ‘after’, *oʒorʔ* ‘while’. The meaning ‘as soon as’ is however coded in a special way by an accented non-past stem plus the prefix *ro-*:

- (32) pje o-vɛr<sup>n</sup>dzem ró-qvɛt t<sup>h</sup>v-mvɛl<sup>n</sup>bjəm  
 bird 3SG: POSS-wing as.soon.as-stretch.open PFV-fly:PAST  
 ‘The bird flew away 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 subordinator *nəʔ*. Alternatively, the conditional form of the copula, *v-nv-ŋo* follows the unmarked verb:

- (33) a. *osto v-nv-tə-vzjaŋ nəʔ kəc<sup>h</sup>iʔ tɐ<sup>n</sup>biʔ*  
 hard IRR-PFV-2-study SUB candy 1>2-give  
 b. *osto tə-vzjaŋ v-nv-ŋo nəʔ kəc<sup>h</sup>iʔ tɐ<sup>n</sup>biʔ*  
 hard 2-study IRR-PFV-be SUB candy 1>2-give  
 ‘I will give you candy if you study hard.’

The instrumental-ergative *-kə* expresses a loose logical connection between two clauses:

- (34) a. *kənm,ŋəmsu-kə pəntjaŋ ó-ta kə-mdzu-kə ne-c<sup>h</sup>ov*  
 fat.person-ERG chair 3SG:POSS-top NOMZR-sit-INST PFV-break:PAST  
 ‘The fat person broke the chair by sitting on it.’  
 b. *mə-nv-fsə-aŋ-kə leʔ tś-tsə*  
 NEG:PFV-PFV-hear:PAST-1SG-INST again IMP-say  
 ‘Say it again, for I did not hear (you).’

Cause and purpose clauses can both take the complex subordinators *otənkə*, while concession clauses are marked with *ntf<sup>h</sup>on(nəʔ)*.

## ADDITIONAL ABBREVIATIONS

- HTR high-transitivity  
 INV inverse  
 IRR irrealis  
 LTR low-transitivity  
 SFP sentence-final particle

## REFERENCES

- DeLancey, Scott (1981) ‘The category of direction in Tibeto-Burman’, *Linguistics of the Tibeto-Burman Area* 6.1: 83–101.  
 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* (Institute of Linguistics, Academia Sinica, Taiwan) 2.1: 69–92.  
 Kuno, Susumo (1976) ‘Subject, theme, and the speaker’s empathy: subject and topic’, in L.N. Li (ed.) New York, San Francisco, London: Academic Press, 417–44.  
 Lin, Xiangrong (1993) *Jiarongyu Yanjiu (A Study of rGyalrong)*, Chengdu: Sichuan Nationality Press.  
 Payne, Thomas E. (1997) *Describing Morphosyntax, A Guide for Field Linguists*, Cambridge: Cambridge University Press.  
 Sun, Jackson T.-S. (2000a) ‘Parallelisms in the verb morphology of Sidaba rGyalrong and Lavrung in rGyalrongic’, *Language and Linguistics* (Institute of Linguistics, Academia Sinica, Taiwan) 1.1: 161–90.  
 — (2000b) ‘Stem Alternations in Puxi Verb Inflection: toward validating the rGyalrongic subgroup in Qiangic’, *Language and Linguistics* (Institute of Linguistics, Academia Sinica, Taiwan) 1.2: 211–32.  
 — (forthcoming) ‘The irrealis category in Caodeng rGyalrong’.  
 Sun, Jackson T.-S. 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* (Institute of Linguistics, Academia Sinica, Taiwan) 3.1: 79–99.

**PART 10**

---

# **KIRANTI LANGUAGES**

---



# KIRANTI LANGUAGES: AN OVERVIEW

*Karen H. Ebert*

## 1 INTRODUCTION

The area where the thirty-two Kiranti languages are spoken stretches over eastern Nepal roughly from the Likhu river in the west into Darjeeling and Sikkim in the east. Hayu and Sunwar are situated a bit further west. The Kiranti people and languages between the rivers Likhu and Arun, including some small groups east of the Arun, are usually referred to as ‘Rai’, which is a somewhat vague geographic grouping rather than a genetic grouping. Most Kiranti languages have less than 10,000 speakers and are threatened by extinction. Some are spoken only by elderly people. Practically all Kiranti speakers are also fluent in Nepali, the language of literacy and education and the national language of Nepal. Kiranti languages have been heavily influenced by Nepali.

Our knowledge of Kiranti languages has increased immensely over the past fifteen years. Up to 1984 the only existing monograph was Allen’s *Sketch of Thulung Grammar*. Since then grammars of Limbu (Weidert and Subba 1985, van Driem 1987), Hayu (Michailovsky 1988), Dumi (van Driem 1993), Athpare (Ebert 1997a), and Yamphu (Rutgers 1998) have been published. Shorter descriptions exist for Khaling (Toba 1984), Bantawa (Rai 1985), and Camling (Ebert 1997b). Ebert (1994) presents a synopsis of grammatical structures of six Kiranti languages. An extensive documentation covering all Kiranti languages on a village to village basis was carried out by the Linguistic Survey of Nepal Project under the direction of Werner Winter. The survey questionnaires consisted of a 250-item wordlist, translations of sentences, and a verbal paradigm. The data collected remain largely unpublished. Thus, for many languages there is still no information available (see Map 31.1; language status based on Survey results, Hansson 1991).

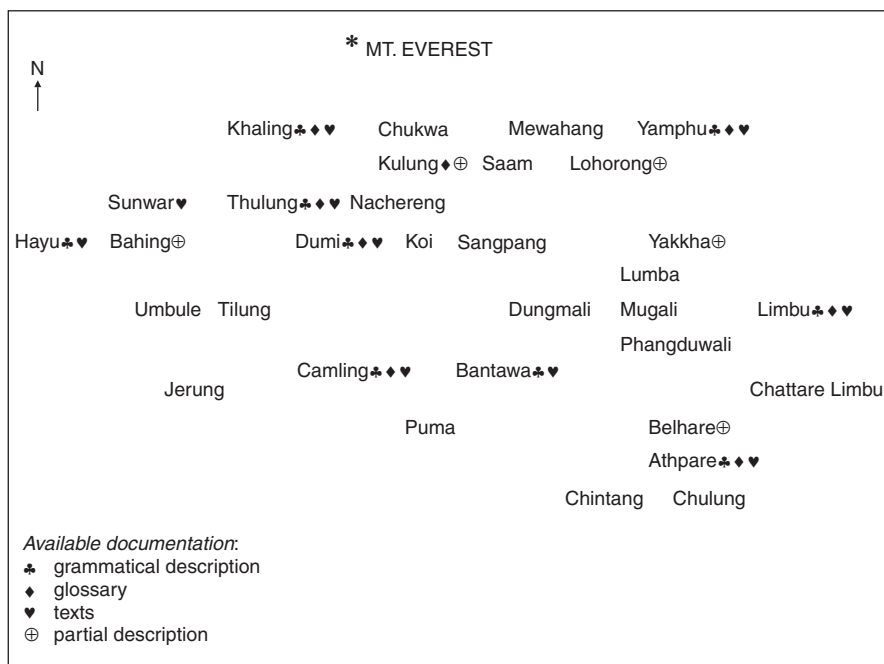
I find it useful to distinguish broadly between southeastern languages, like Limbu, Athpare, Belhare, Bantawa, Camling, and northwestern languages, like Hayu, Sunwar, Thulung, Khaling. This is of course not meant to be a genetically valid grouping (see Section 3), nor is it relevant for all features. Some traits are, for example, found only in eastern languages (including Yamphu and Yakkha in the northeast, but not Camling and Bantawa), others only in central languages (including Camling and Bantawa). As the following attempt to give a general characterization of Kiranti languages is necessarily sketchy, it will be useful to consult also the Chapters on the individual Kiranti languages.

## 2 TYPOLOGICAL CHARACTERISTICS

### 2.1 Phonology

All Kiranti languages have a stop series *p, t, k* and the affricate *ts* (written <c>, if it cannot be split up into *t + s*). NW languages have a regular voiced vs voiceless opposition. Eastern languages have at most a handful of lexemes with initial voiced stops, but in intervocalic and





MAP 31.1 MAP OF KIRANTI LANGUAGES

postnasal position voicing takes place: Athpare *yep-yuk* ‘he will stand up’, *yeb-e* ‘he stood up’. All languages have an opposition between aspirated and nonaspirated stops. Some have breathy voiced stops, with breathiness sometimes optional: Thulung *jham* ‘possible’, *jam* ‘cooked rice’, but *joomu* ~ *jhoomu* ‘plough’. Camling seems to be the only language in the documented set that has phonemic breathy voice with nasals and laterals, e.g. *lhamma* ‘catch’, *lamma* ‘look for’, *lammha* ‘dump’. The only fricative is *s*. Some northern languages have a tonal contrast, which could be related to (loss of) voiceless finals, cf. Thulung *loaana* ‘you see it’ vs *lóaana* (<\*loak-/t/-na see-PT-2s) ‘you saw it’.

The inventory of vowel phonemes ranges from five in Sunwar to nine in Khaling, which has rounded front vowels. Bantawa and Dumī have two high back vowels, /u/ and /u:/ (written <í>): Dumī *nu* ‘nose’, *ní* ‘name’. In NW languages vowel length is phonemic in open syllables: Thulung *romu* ‘come’, *roomu* ‘snatch’; in some central and eastern languages also in closed syllables: Dumī *delme* ‘daughter-in-law’, *deelme* ‘bazaar’, Yamphu *ce?ma* ‘plough’, *cee?ma* ‘spit’.

Initial consonant clusters are restricted to bilabial or velar stop + r/l and occur only in the west, including NW Camling; cf. Thulung *khrapd-*, Umbule *khrams-*, NW Camling *khrap-* and SE Camling, Bantawa, Kulung *khap-* ‘weep’. Nasals, liquids, voiceless stops and *s* occur word-finally. Stops in syllable-final position are unreleased and glottalized in most languages. Occasionally a stop is inserted at juncture before a consonant: Thulung *kutlam* ‘from the water’ (*ku* ‘water’ + *lam* ablative); see also (5b). Some SE languages have a tendency to add stops (mainly *k*) to nominals; Belhare *dabhek* ‘sickle’ (cf. Camling *dabe*); Athpare *nawat* ‘snot’ (*na* ‘nose’ + *wa* ‘water’), *appek* ‘self’, ‘own’ (< Nep. *āphno*). Often the euphonic stop is optional; Bantawa *nabu(k)* ‘nose’, *nicho(k)* ‘younger sibling’, *masi(t)* ‘lentil’.

## 2.2 Nominals

Pronouns distinguish dual and plural as well as inclusive and exclusive; see Table 31.1 for the forms in three languages. NW languages have dual and plural markers with all pronouns; SE languages have only a non-singular marker with third person pronouns (and with nouns). The third person forms are restricted to humans; elsewhere demonstratives are used. The inventory of possessive prefixes is reduced in eastern languages, like Athpare, where the prefixes stand for person independent of number. Western languages, like Thulung, have a complete system of prefixes; Bantawa and Camling are somewhere in between.

Eastern languages mostly use the personal pronouns for non-singular possessors, often together with the possessive prefix. Possessive pronouns are otherwise formed from personal pronouns or prefixes with a genitive (with many idiosyncrasies).

- (1) ATHPARE *u-paŋ* his house (rarely: their house)  
 3POSS-house  
*unci* (*u-*) *tak-ci* their friends  
 they/their 3POSS-friend-ns
- KHALING *i-beŋmä-häm* ~ *in-po* (*i-*)*beŋmä-häm* your sisters  
 2sPOSS-sister-p you<sub>s</sub>-GEN

Eastern languages insert a nasal before some kinship terms or terms for social relations: Limbu *a-m-ba* ‘my father’, *kɛ-n-jum* ‘your friend’. Athpare has reduplication with some kinship terms, preferably with first person plural possessors: *jha-n-jha* ‘our younger sibling’. Possessor nouns are in the genitive case, sometimes followed by a nominalizer/attributivizer. The possessor can be cross-referenced by a prefix on the head noun, though this is usually avoided with words borrowed from Nepali; cf. the following text occurrences with Nepali loans in underlined italics.

- (2) ATHPARE *beula-ŋa-na* *u-tak* the groom’s friend  
 groom-GEN-NOMZR:s 3POSS -friend  
*beuli-ŋa-ga* *natedar-ci* the bride’s relatives  
 bride-GEN-NOMZR:p relative-ns
- KHALING *us-celpa-po* *u-kham* their brother’s house  
 3dPOSS-brother-GEN 3sPOSS -house  
*ut-po* *sanawä* the camel’s conciousness  
 camel-GEN conciousness

TABLE 31.1 PRONOUNS AND POSSESSIVE PREFIXES

	<i>Athpare</i>		<i>Bantawa</i>		<i>Thulung</i>	
1s	<i>aŋa</i>	<i>a-</i>	<i>iŋka</i>	<i>iŋ -</i>	<i>go</i>	<i>a-</i>
1di	<i>anci</i>		<i>iŋkaci</i>	<i>anco-</i>	<i>guci</i>	<i>ici-</i>
1de	<i>anciga</i>		<i>iŋkaca</i>		<i>gucuku</i>	<i>aci-</i>
1pi	<i>ani</i>		<i>iŋkan</i>		<i>guy</i>	<i>iki-</i>
1pe	<i>aniga</i>		<i>iŋkanka</i>		<i>guku</i>	<i>aki-</i>
2s	<i>khana</i>	<i>ka-</i>	<i>khana</i>	<i>am-</i>	<i>gana</i>	<i>i-</i>
2d	<i>khanci</i>		<i>khanaci</i>	<i>amco-</i>	<i>gaci</i>	<i>ici-</i>
2p	<i>khani</i>		<i>khananin</i>	<i>an-</i>	<i>gani</i>	<i>ini-</i>
3s	<i>un</i>	<i>u-</i>	<i>kho(ko)</i>	<i>i-</i>	<i>gu</i>	<i>u-</i>
3d	<i>unci</i>		<i>khoci</i>	<i>unco-</i>	<i>guci</i>	<i>uci-</i>
3p	<i>unci</i>		<i>khoci</i>		<i>gumi</i>	<i>uni-</i>

Demonstrative pronouns consist of a deictic root, usually followed by a nominalizer. The deictic roots combine also with locative case markers.

## (3) BANTAWA

	proximate		<i>distal</i>	
<i>o-ko</i>	PROX-NOMZR	this	<i>mo-ko</i>	that
<i>o-ko-ci</i>	PROX-NOMZR-d	these two	<i>mo-ci</i>	those two
<i>o-da</i>	PROX-LOC	here	<i>mo-da</i>	there
<i>o-yu</i>	PROX-lowLOC	down here	<i>mo-yu</i>	down there

Classifiers distinguish humans and non-humans, seldom shape, e.g. Camling *-li* for round things. The classifiers are not used systematically any more, as mentioned by several authors (Allen 1975: 113, Rai 1985: 166). Natural gender is distinguished in some terms designating humans and larger animals. Female is marked by *-ma*, male by *-pa* or *-cö*: Thulung *recukuma/recukupu* 'orphan', *bəŋme/bəccö* 'guest'. Sometimes natural gender is distinguished with participles or adjectives: Limbu *kəyuyupakəyuyuma* 'a sitting person' (<yuy- 'sit'), *cukpa/cukma* 'small'.

There is usually a small number of lexical adjectives, but most words corresponding to adjectives in English are derived from verbs through nominalization. Comparative and superlative are expressed with the help of ablatives in some northern and NW languages, whereas most SE languages have borrowed the Nepali comparative marker *bhanda*.

(4) THULUNG	<i>a-ma-ḍa-m</i>	<i>ḍokpu</i>	<i>khotle-ḍa-m</i>	<i>ḍokpu</i>
	my-NML-LOC-ABL	big	all-LOC-ABL	big
ATHPARE	<i>aŋa-na-bhanda</i>	<i>thena</i>	<i>sob-bhanda</i>	<i>thena</i>
	my-NOMZR-than	big	all-than	big
		'bigger than mine'		'the biggest'

Case markers include a combined ergative-instrumental, a genitive, and a comitative, which is also used to coordinate nouns: Khaling *soroli-kolo ut-kolo* 'jackal and camel'. Besides a locative and an ablative, some languages have an optional allative marker (5b). Ablative and allative are suffixed to the locative marker, but mediatives derived from *\*lam* 'path' attach directly to a nominal: Limbu *peenibaan-lam* 'in Nepali', Athpare *toba-lam* 'from above', *uthuy-lam* 'over the top'. Case markers for higher, lower, and same altitude constitute a unique trait of Kiranti languages (see also 'Camling' in this volume).

## (5) a. THULUNG

*mö chörcü-la ləs-ta-ma mö-la-ŋa bay-ra.*  
 that basket-highLOC go-PT-SEQ that-highLOC-EMPH be-PT  
 'She went up into the basket and stayed up there.'

## b. BANTAWA

*in̄ka papa-ḍu-t-nin khat-ŋa-ne, nana, khana-nin*  
 I father-highLOC-[t]-ALL go-1s-OPT e.sister you-p/hon  
*mama-yu-t-nin khara-ne.*  
 mother-lowLOC-[t]-ALL go-OPT

'I should go up to father's place, sister, may you go down to mother.'

Altitudinal case markers were not found in the most western and in some eastern languages (Limbu, Athpare), though cognate suffixes do occur with deictic roots; Limbu *kət-thoo* 'up here', *kət-yoo* 'down here', *kət-na* 'over here'.

TABLE 31.2 CASE MARKERS

	<i>Athpare</i>	<i>Bantawa</i>	<i>Camling</i>	<i>Thulung</i>	<i>Khaling</i>	<i>Yamphu</i>
ERG/INST	-ŋa	-a	-wa	-ka	-ä	-æʔ, -tæʔ
GEN	-ŋa	-ko	-mo	-kam	-po	-mi(n)
COM	-lok	-nin		-nuŋ	-kolo	-nu(ŋ)
LOC (general)	-ŋi, -i	-da	-da	-ɖa	-bi	-pe
ALL		-nin	-ni		(-tha)	
ABL	-lam(ma)	-ŋka	-ka	-m	-ka	-pa(ŋ)
MED	-lam(ma)		-la	-lam		-la(n)
highLOC		-du	-dhi, -di	-la	-tü	-tu
lowLOC		-yu	-i	-yu	-yü	-mu
levLOC		-ya	-ya	-no	-yo	-yu

### 2.3 Verbal systems

Most Kiranti languages have two stems for each verb. In SE languages the full stem occurs before vowels, the reduced stem before consonants and word-finally: Athpare *nis-ue* ‘he saw it’, *ni-ne* ‘I saw you’, *ni nepma* ‘show’. Stem alternations are more complex in NW languages. Simple verb stems have the structure C(r/l)V(C), or C(r/l)VCC if an older causative with an augment *t/d/s* is lexicalized. Causative stem formation is only partially productive. Sometimes a secondary causative/applicative can be formed, e.g. Bantawa *i-* ‘laugh’, *is-* ‘make laugh’, *itt-* ‘laugh at’; *par-* ‘shout’, *pays-* ‘make shout’, *patt-* ‘shout at’. Some western languages and Limbu also form causatives by devoicing and/or aspiration; Hayu *bok-* ‘be born’ / *phok-* ‘give birth’, Thulung *get-* ‘come up’ / *khet-* ‘bring up’, Limbu *tæks-* ‘be torn’, *thæks-* ‘tear’. In the more frequent analytic causatives the verb ‘make’ is either attached to the stem in root serialization, or it follows the preconsontal stem as an independent verb: Yamphu *kæʔ-mett-e* (come-make-IMPER) ‘let him come’, Bantawa *khaŋ tɨ-mett-u* (see 2-make-3P) ‘you showed him’.

The Kiranti finite verb is characterized by a complex system of person and number affixes. Some widespread Tibeto-Burman suffixes are well preserved in Kiranti: *-ŋa* first person singular, *-na* second person singular, *-i* first and second person plural (the latter also *-nin*), *-u* third person patient, *-ci*, *-si* dual. SE languages have prefixes with parallels in other TB languages, like *kɛ-*, *ta-*, *a-* for second person, *ɨ-* for inverse, *i-*, *a-* for ‘marked scenarios’ (inverse configurations and second person participation).

In principle both actants of a transitive scenario are marked on the verb, but the two markers cannot always be identified in a straightforward way. SE languages typically use the first and second person affixes in 2 → 1, whereas the old second person marker *-na* serves as a portmanteau for 1 → 2 (also in some central languages).

(6)	<i>2s</i> → <i>1s</i>		<i>1s</i> → <i>2s</i>	
BANTAWA	<i>tɨ-dhat-ŋa</i>	you beat me	<i>dhat-na</i>	I beat you
ATHPARE	<i>a-lems-a-ŋ-e</i>		<i>lem-n-e</i>	
	2-beat-PT-1s(-PT)		beat-1→2(-PT)	
YAKKHA	<i>mokt-a-ŋ-ga-na</i>		<i>mok-nan-na</i>	
	beat-PT-1s-2-NOMZR:s		beat-1→2-NOMZR:s	

Person markers, especially that of second person, are often independent of role:

(7)	<i>2s</i> → <i>3s</i>		<i>3s</i> → <i>2s</i>	
BANTAWA	<i>tɨ-khaŋ-u</i>	you saw him	<i>tɨ-khaŋ-a</i>	he saw you
LIMBU	<i>kɛ-nis-u</i>		<i>kɛ-nis-ɛ</i>	

ATHPARE	<i>a-nis-u-e</i> 2-see-3P(-PT)	<i>m-a-nis-e</i> 3A-2-see-PT
BELHARE	<i>nis-e-ga</i> see-PT-2	<i>n-nis-e-ga</i> 3A-see -PT-2
THULUNG	<i>loaa-na</i> see-2	<i>loaa-na</i> see-2

The dual marker is independent of person and role in all Kiranti languages; in NW languages this is also true for the plural marker (cf. Thulung *-mi*).

TABLE 31.3 THULUNG PERSON-NUMBER MARKERS

A	P											
	<i>1s</i>	<i>1de</i>	<i>1pe</i>	<i>1di</i>	<i>1pi</i>	<i>2s</i>	<i>2d</i>	<i>2p</i>	<i>3s</i>	<i>3d</i>	<i>3p</i>	<i>itr</i>
1s						-ni	-nici	-ni	-(p)u	-(p)uci	-(p)umi	-ŋu
1de						-naci		-nicimi	-cuku			
1pe						-nami	-nacimi	-nimi	-ku		-ki	
1di											-ci	
1pi											-i	
2s	-ŋi								-na			
2d4	-ŋici								-ci			
2p	-ŋimi			-ciki	-kimi						-ni	
3s	-ŋi			-saci	-sa	-na	-naci	-nimi	-ü		-	
3d	-ŋici								-ci			
3p	-ŋimi			-sami	-nami	-nacimi						-mi

TABLE 31.4 BANTAWA PERSON-NUMBER MARKERS

A	P												
	<i>1s</i>	<i>1de</i>	<i>1pe</i>	<i>1di</i>	<i>1pi</i>	<i>2s</i>	<i>2d</i>	<i>2p</i>	<i>3s</i>	<i>3ns</i>	<i>itr.</i>		
1s						-na	-naci	-nanin	-uŋ	-u-ciŋ	-ŋa		
1de						-naca		im- in	-ca		-a		
1pe						im-	im--ci		-umka	-umcumka	-ika		
1di											-cu	-ci	
1pi											-um	-umcim	-in
2s	ti- ŋa	ti- ca	ti- inka						ti- u	ti- uci	ti-		
2d	ti- ŋaciŋ								ti- cu		ti- ci		
2p	ti- ŋaniŋ	ti- nici						ti- um	ti- umcim	ti- i			
3s	i- ŋa	i- c-a	i- inka	-ci	im- in	ti-	ti- ci	ti- in	-u	-uci	-		
3d	im- ŋa	im- c-a	im- inka	im- ci	im- in	im-	im- ci	im- in	im- cu		-ci		
3p											i-	im- uci	im-

In spite of such common principles, the paradigms are arranged quite differently, especially in the 2→1, 3→1 and 3→2 configurations, as Tables 31.3 and 31.4 illustrate. This is partly due to different alignments of non-singular configurations, in which number is sometimes marked for A, sometimes for P. Generally the verb forms of NW languages are shorter than those of SE languages, some of which produce lengthy suffix chains with copied suffixes, e.g.

(8)	ATHPARE	<i>nis-u-m-ci-m-cim-ma</i>		$w_{pe}$	will see them
		see-3P-1pA-3nsP	-[copy]-NPT:[copy]-e		
	YAKKHA	<i>moŋ-me-ŋ-c-u-ŋ-ci-ŋ-ha</i>		$w_{de}$	will beat them
		beat-NPT-1e-d-3P-[copy]-3nsP	[co py]-NOMZR:ns		

In some SE languages traces of direction marking can be found. In the following forms the prefix *i-* indicates inverse and *-u* indicates direct in a system based on a participant hierarchy 1/2 > 3s > 3p.

(9)	BANTAWA				
3s→3s	<i>khaŋ-u</i>	he saw him	3p→3s	<i>i-khaŋ-a</i>	they saw him
	see-DIRECT			INV-see-PT	
1pe→3s	<i>khaŋ-u-m-ka</i>	$w_{pe}$ saw him	3s→1pe	<i>i-khaŋ-in-ka</i>	he saw us <sub>pe</sub>
	see-DIRECT-1/2pA-e			INV-see-1/2pS/P-e	
2p→3ns	<i>tí-khaŋ-u-m-cím</i>	you saw them	3ns→2p	<i>i-m-khaŋ-in</i>	they saw you
	2-see-DIRECT-1/2pA-3nsP:[copy]			INV-3n sA/S-see-1/2pS/P	

The inverse marking system is not very transparent in any of the languages. Thus Bantawa *i-* amalgamates with second person *tí-* and the prefix *im-* occurs also in third person singular and in 1pe→2 forms (Table 31.4).

Some SE languages have developed new paradigms for inverse configurations with a first person patient. In Athpare and in Limbu two patterns exist side-by-side, whereas in Camling they are found in different dialects (see Ebert in this volume). Limbu indicates first person exclusive by preposing *naapmi* (*yapmi* in other dialects) 'person'. Athpare has *yaŋ-*, though some speakers also use *yapmi*.

(10)	LIMBU				
2s→1s	<i>ke-hip-ʔe</i>	you'll beat me	or: 2→1e	<i>naapmi ke-hip</i>	you beat me/us
	2-beat-1s:NPT			1eP (= person) 2-beat	
2s→1ns	<i>a-g e-hip</i>	you'll beat us			
	1nsP-2-beat				
	ATHPARE				
3s→1s	<i>nis-a-ŋ-e</i>	he saw me	or: 3s→1e	<i>yaŋ-nis-e</i>	he saw me/us
	see-PT-1s-PT			1eP-see-PT	
3s→1pe	<i>nis-i-ŋ-e</i>	he saw us	3p→1e	<i>yaŋ-u-nis-e</i>	they saw me/us
	see-1pS/P-e-PT			1eP-3pA/S-see-PT	

These paradigms have been restructured in such a way that A is now marked like an intransitive subject (cf. Ebert 1997a: 38–41).

Most Kiranti languages have two basic tense forms. In SE languages past is mostly marked by /a/ following the stem, in some NW languages by /t/. The simplest situation is found in Limbu and Bantawa, which have past /a/ and leave the non-past unmarked. As /a/ is elided by a following vowel, only half of the verb forms are marked for tense. Most other languages have a marked non-past (with a great variety of markers). Athpare joined a new

past marker *-e* to the old past forms; the analogous formation serves as a non-past (and imperfective) in Camling, where stem + *a* became a finite base.

(11)	PT		NPT	
ATHPARE	<i>khad-e</i>	he went	<i>khat-yuk</i>	he will go
	<i>khad-a-c-e</i>	they <sub>d</sub> went	<i>khat-ci-ci</i>	they <sub>d</sub> will go
	go-PT-d-PT		go-d-NPT:[copy]	
	<i>nis-u-m-e</i>	we <sub>pi</sub> saw it	<i>nis-u-m-t-um</i>	we <sub>pi</sub> will see it
	see-3P-1pA-PT		see-3P-1pA-NPT-[copy]	
CAMLING	<i>khata</i>	he went	<i>khat-e</i>	he will go
	<i>khata-ci</i>	they <sub>d</sub> went	<i>khata-c-e</i>	they <sub>d</sub> will go
	go (:FIN)-d		go (:FIN)-d-NPT	
	<i>tyok-u-m</i>	we <sub>pi</sub> saw it	<i>tyok-u-m-e</i>	we <sub>pi</sub> will see it
	see-3P-1pA		see-3P-1pA-NPT	

A similar flip-flop exists between Dumi nonpast and Thulung past, marked by /t/ after the stem or the person markers.

(12)	PT		NPT	
DUMI	<i>a-yum-ə</i>	he/you hit me	<i>a-yum-t-ə</i>	he/you will hit me
	MS-hit-1s		MS-hit-NPT-1s	
	<i>a-yəm-ki</i>	he hit us <sub>pi</sub>	<i>a-yəm-ki-ti</i>	he will hit us <sub>pi</sub>
	MS-hit-1pi		MS-hit-1pi-NPT	
THULUNG	<i>yal-ŋi-di</i>	he/you hit me	<i>yal-ŋi</i>	he/you will hit me
	hit-1P-PT		hit-1P	
	<i>yal-to-mi</i>	I hit them	<i>yal-u-mi</i>	I will hit them
	hit-PT:1→3-p		hit-1→3-p	

Many Kiranti verbs have ingressive or ingressive-phasal aktionsart, e.g. Athpare *les-* ‘get to know’, *tuk-* ‘hurt’, Limbu *lakt-* ‘(come to the) boil’. We therefore often find a past or perfect form where English has a present; Athpare *lesume* ‘we know’ (i.e. ‘we got to know’), *tuge* ‘it hurts’ (i.e. ‘it came to hurt’). Aspect and aktionsart has been investigated in depth only for Belhare (Bickel 1996).

Compound verbs occur either in root serialization (Hayu, Khaling, Yamphu, Belhare), or both verbs are finite-marked and contracted to some degree (most SE languages). They fulfil similar functions as in other South Asian languages; i.e. the second verbs (‘explicators’, ‘aspectivizers’) are often telicizing, but some are stativizing, and ‘give’ is benefactive/malefactive (see Thulung (18)).

(13)	THULUNG	<i>pü-leađ-d-ü</i>	he ate it up
		eat:3s→3-V2:TEL-PT-3s→3	
	ATHPARE	<i>lept-u-des-u-e</i>	he threw it away
		throw-3P-V2:TEL-3P-PT	

Periphrastic tenses show a great variety of formations. The main verb can have converbal or participial form, or it can be a finite-marked verb followed by a subordinator. In SE languages we also find grammaticalized compound verbs. Periphrastic tenses are used to a greater degree in SE languages. Athpare, Belhare, and Bantawa have fully grammaticalized perfects, progressives and/or imperfectives. Most SE languages moreover have an ambulative, which indicates that something is done while moving around or ‘all over the place’.

- (14) BELHARE (Bickel 1996: 164)  
*kaepma-chi hit-kon-u-chi.*  
 girl-ns look-AMB-3P-3nsP  
 ‘He is going around watching girls.’

Verbs are negated by a negative prefix (mostly *ma(n)-*, *mi-*), in SE languages also by a negative suffix or by both. Often the negative forms lack tense markers and/or direction or 3P markers.

- |         |                          |                    |                                    |
|---------|--------------------------|--------------------|------------------------------------|
| (15)    |                          |                    | <i>NEG</i>                         |
| ATHPARE | <i>a-nis-u-t-u</i>       | you’ll see it      | <i>a-nis-u-n-na</i>                |
|         | 2-see-3P-NPT- [copy]     |                    | 2-see-3P-NEG-NOMZR                 |
|         | <i>a-nis-u-e</i>         | you saw it         | <i>a-nis-u-n-get-u-n-na</i>        |
|         | 2-see-3P-PT              |                    | 2-see-3P-NEG-V2:negPT-3P-NEG-NOMZR |
| LIMBU   | <i>huʔr-u-ŋ-si-ŋ</i>     | I’ll teach them    | <i>mɛ-huʔ-ʔɛ-n-chi-n</i>           |
|         | teach-3P-1sA-3nsP-[copy] |                    | NEG-teach-1s:NPT-NEG-3nsP-[copy]   |
|         | <i>kɛ-dɛs-u</i>          | ‘you sow/sowed it’ | <i>kɛ-n- dɛː s-u-n</i>             |
|         | 2-sow-3P                 |                    | 2-NEG-sow-3P-NEG                   |

There is no straightforward correspondence between positive and negative tenses. One answers to the question ‘Have you eaten?’ in Thulung with *pe-tt-o* (eat-PT-1s→3) ‘I have eaten’, but in the negative with *mi-pe-thiŋa bu-ŋa* (NEG-eat-CONV be-1s), and not with the negative past *pewwa* (Allen 1975: 88). Also in other languages a past is often negated by a ‘negative perfect’; cf. with the Limbu negated form in (15):

- (16) LIMBU (van Driem 1987: 180)  
*khɛnɛʔ maaki kɛ-dees-w-iiʔ – mɛn-dee-ʔtee waa-ʔɛ.*  
 you(s) maize 2-sow-3P-Q NEG-sow-CONV be-1s:NPT  
 ‘Have you sown the maize?’ – ‘No, I haven’t.’

Non-finite forms include besides converbs ((16) (22a,b)) an infinitive, a purposive (19), agentive (18) and patientive participles and (in some languages) a combined locative-instrumental participle, e.g. Yamphu *-tha(m)*:

- (17) a. YAMPHU (Rutgers 1998: 206)  
 waawa *pen-dham yaksa*  
 e.brother stay-LOC PCPL hut  
 ‘the hut in which elder brother stays’
- b. YAMPHU (Rutgers 1998: 494)  
*goru pæk-tha tonʔur-æʔ hiik-yuŋ-nuŋ*  
 ox beat-LOC PCPLstick-INST tie-put-CONV  
*mook-led-u-ŋ-ji-ŋ nip-paŋ-noʔ.*  
 thrash-stop-3P-1s-3ns-[copy] two-CL-FOC  
 ‘I stopped and tied them [i.e the wives] up and gave them both a thrashing with the stick with which I beat the oxen.’

Participles are used as nouns or as attributes, though there are restrictions in some languages. Patientive participles are mainly attested as attributes without an explicit actor; Athpare



*mog-balak cuwa* ‘boiled water’, Dumi *thip-mpo suule* ‘sewn thread’, Thulung *khər-ma maakai* ‘parched maize’.

## 2.4 Syntax

Kiranti languages are head-marking SOV languages with a rather strict order of modifier before head. Participants are marked on the verb and pronouns are not obligatory. At most two participants can be coded, animates having precedence in bitransitive clauses. The pronominal affixes mark participation in the events, often independently of the noun phrases in the sentence (for examples see Chapter 33 by Ebert and Chapter 34 by Bickel in this volume).

Kiranti languages are morphologically ergative. Some languages exhibit a split between speech-act participants and non-participants (Camling, Thulung) or between pronouns and non-pronouns (Limbu), only the latter taking the ergative suffix. Nonhuman undergoers are unmarked for case; human undergoers optionally take the Nepali suffix *-lai* in some languages. The experiencer of emotions is coded by a possessive prefix: Thulung *uci-chokco bək-ta* (their<sub>d</sub>-anger come.up-PT) ‘they<sub>d</sub> got angry’, Limbu *kε-sira dhaŋ-ŋi* (your-pleasure come.up-Q) ‘do you like it?’ Possessive prefixes also indicate the undergoer with agentive participles and with purposives.

- |      |         |                                |                   |                             |
|------|---------|--------------------------------|-------------------|-----------------------------|
| (18) | ATHPARE | <i>a-ka-pik</i>                |                   | ‘the one who speaks to me’  |
|      |         | 1sPOSS-AgPCPL-speak            |                   |                             |
|      | THULUNG | <i>ima i-pep-saŋ-pa</i>        |                   | ‘the one who eats from you’ |
|      |         | your 2sPOSS-eat-V2:give-AgPCPL |                   |                             |
|      |         |                                |                   |                             |
| (19) | CAMLING | <i>kap-tum-si</i>              | <i>t-uŋ-ko.</i>   | ‘I have come to see you.’   |
|      | LIMBU   | <i>kε-dum-se</i>               | <i>ty-aŋ-ba.</i>  |                             |
|      | THULUNG | <i>i-reb-ɖa</i>                | <i>bi-ŋdo-m.</i>  |                             |
|      |         | 2sPOSS-meet-PURP               | come -1s:PT-NOMZR |                             |

With the infinitive only non-singular number of the undergoer can be indicated; Thulung *sii-mu-ci basi* (teach-INF-d must) ‘we must teach those two.’

The typical SOV pattern of subordination with preposed non-finite verbs is followed strictly only by the westernmost language Hayu. Non-finite clauses are infrequent in the southeast. Relative clauses can be expressed both with participles and with nominalized finite-marked verbs in most languages.

- (20) a. DUMI (van Driem 1993: 273)
- miin dumoot hoo-kpɪ hompɪ-bi, baadzarɪm-bi dumoo bus-t-a.*  
 person much come-PCPL place-LOC bazar-LOC much scream-NPT-3s.  
 ‘He screams a lot in places where many people come, in the bazar.’
- b. DUMI (van Driem 1993: 289)
- nacɪrmi-mɪl-ʔa kaand-ini-m sura-mɪl sɪŋ-si-kə*  
 savage-p-ERG throw-3p-NOMZR paddy-p pick.up-d-SE  
*huut-siʔe.*  
 bring-d UNW  
 ‘The two collected and brought home the paddies which the savages had thrown away.’

Internal head relative clauses have been attested for Limbu, Belhare, Khaling and Dumi.

- (21) KHALING (Ebert 1994: 263, from Toba)

*būrā ut samundra-po phār-bi carey ma-si-pä*  
 old camel ocean-GEN edge-LOC graze do-REFL-AgPCPL  
*soroli-ä düm-tä.*  
 jackal-ERG meet-PT

‘The jackal met the old camel grazing by the water.’

In Hayu, adverbial clauses and event sequences are always expressed with converbs. Other languages use both converbs and finite verbs, with the latter again much more common in the southeast. Often case markers function as subordinators.

- (22) a. HAYU (Michailovsky 1988: 205)

*nikai wonoŋ-boŋ buubu-ha cuŋ-noŋ*  
 much high.place-until carry:REDUP-CONV mount-CONV:TEMP  
*suuk-o-m lxtse bilv-ha.*  
 scratch:PT-3P-ASS UNW tiger-ERG

‘When she had carried him up high, tiger scratched her.’ (-ha = ERG, -noŋ = LOC)

- b. SUNWAR (Bieri *et al.* 1973: 447)

*rawa ŋóyk-se-nu, koo-sha ŋóyks-o.*  
 tree climb-2s-TEMP look-CONV climb-IMP

‘When you climb a tree, look first and then climb.’ (-nu = COM)

See also Yamphu (17b): CONV -*nuŋ* = COM; Dumi (20b): SEQ -*kə* = ABL. It is remarkable that the SE languages have no sequential converb (‘conjunctive participle’). Even in most NW languages finite verb forms are more common in sequence clauses (Thulung (5a)).

In most languages the subordinate verb has its full finite form, but in Athpare and Belhare it lacks past and non-past markers. In some forms this brings out the past marker *a*, (elided by past -*e*, see (11)), resulting in a subjunctive. Cf. Athpare with Camling, which retains tense markers in subordination:

- (23) ATHPARE *unci pay-i u-khad-e. u-khad-a-na-m...*  
 their house-LOC 3pA/S-go-PT 3pA/S-go-SUBJ-NOMZR-TOP  
 ‘They went home. When they went . . . .’

CAMLING *kic-khim-da mi-khat-e. mi-khat-e-ko-na...*  
 3pPOSS-house-LOC 3pS-go-NPT 3pS-go-NPT-NOMZR-TOP  
 ‘They will go home. When they go . . . .’

Kiranti languages make ample use of nominalizers, which derive nouns from deictic roots (Camling *u-ko/tyu-ko*, Khaling *tä-m/mä-m* ‘this/that’), adverbials (Yamphu *iskul-i-ha-ji* school-POSS-NOMZR-ns ‘the school children’, Limbu *anchem-ba inghəŋ* yesterday-NOMZR news ‘yesterday’s news’) and from finite-marked verbs (Camling *mi-ta-e-ko* 3pS-come-NPT-NOMZR ‘those who will come’). Any nominalized form can be attributed. Nominalized clauses function as complements and temporal clauses. Moreover, they are focalizers (Bickel 1999; Ebert 1997a: 58–60). Nominalizing the whole utterance, as in (19), is extremely frequent. In some languages there is a sheer inflation of nominalizers; e.g. in Athpare all verbs are nominalized in negation and in questions (cf. (15)); all Yakkha verb forms were given with a nominalizer in the Survey Questionnaires (see (6), (8); cf. also Allen 1975: 85 for Thulung).

### 3 GENETIC AND AREAL GROUPINGS

In the *Linguistic Encyclopedia* (Bright ed. 1992) Magar, Kham, and Chepang are subsumed under Kiranti languages. It is unclear what the criteria for this inclusion are. On the other hand it has never been shown that Kiranti – in the sense in which it is normally understood – is a valid genetic unit. There is also no consensus about its place inside Sino-Tibetan. Recently van Driem (1998) proposed a close relationship between Sinitic and Bodish-Himalayan on the basis of reconstructed Old Chinese and (mainly) present-day Limbu. The time gap of a few thousand years is, of course, problematic. If we were able to reconstruct, say, Proto-SE Kiranti and compare it with Proto Naga (which is placed in a different subfamily), the correspondences would probably be much more numerous than with Old Chinese.

There is no valid genetic subgrouping of Kiranti languages. Glover's (1974) evaluation of the Swadesh hundred wordlist for seven languages yielded only 26 per cent of cognates for the Rai group. Bantawa, Kulung, and Khaling come out as a subgroup with 42 per cent shared cognates. This result is at odds with Winter's comparison of 85 basic verbs in twelve Rai languages (Winter 1991–2). Western languages (Umbule, Bahing, Thulung, Khaling) turned out to share less than 20 cognates with central (Kulung, Nachereng) and SE languages (Bantawa, etc.). It can of course be difficult to recognize cognates if no sound laws have been established. Systematic sound changes have so far been discovered only for a narrow set of data, e.g.  $r > y$  in eastern languages,  $ak > o$  in Camling. Camling *rodon* is thus the same word as Limbu *yakthum-ba* 'Kiranti' (and not a name for the Camling language).

Application of the methods of historical linguistics to data from the Linguistic Survey of Nepal yields no neat bundles of isoglosses for specific phonological, morphological, or lexical features. Winter therefore finds it 'too early to call it the Kiranti family' (1986: 763), and Hansson assumes in an unpublished report of the Survey Project that the cluster of Kiranti languages 'results from several migration waves of Tibeto-Burman groups that have influenced each other for a longer period'.

I have mainly distinguished NW and SE Kiranti languages. Some traits that distinguish SE from NW Kiranti, with isoglosses for the first two features running right through the Camling area, are:

- no initial consonant clusters,
- agentive participles formed with *ka-*,
- prefixes for second person and third person plural on verbs,
- traces of inverse morphology and/or intransitivized 1st patient forms,
- agglutinative morphology with long suffix chains,
- preference for finite-marked forms in subordination.

Such areal features should be compared not only with Tibeto-Burman languages, but also with Indo-Aryan, Munda and Dravidian languages of the broader area. SE Kiranti features like prefixed person markers and the participle formation with *ka-* are found also further east, e.g. in Naga and Chin languages. The inverse marking of Camling and Bantawa has the closest parallel in Gyarong, which is not a neighbour nor supposedly closely related. Reduplication in possessive constructions with kinship terms is known from the North-Dravidian Kurukh. Syllable closing euphonic *k* exists also in eastern Indo-Aryan and in Munda languages. The highly agglutinative morphology of SE Kiranti is characteristic of North Munda languages, and so is the preference for finite-marked subordinate clauses (cf. Ebert 1993, 1999b). All those traits are foreign to western Kiranti languages like Hayu and Sunwar, and I know of no close parallels with other Tibeto-Burman languages of Nepal.

## ADDITIONAL ABBREVIATIONS

AMB	ambulative
INV	inverse
MED	mediative
MS	marked scenario (inverse or second person participation)
NW	northwestern
PCPL	participle
SE	southeastern
SEQ	sequential
TEL	telicizer
v2	second verb ('explicator', 'aspectivizer').

## REFERENCES

- Allen, Nicholas J. (1975) *Sketch of Thulung Grammar*, Ithaca NY: Cornell University.
- Bickel, Balthasar (1996) *Aspect, Mood, and time in Belhare. Studies in the Semantics – Pragmatics Interface of a Himalayan language*, Zürich: ASAS-Verlag.
- (1999) 'Nominalization and focus constructions in some Kiranti languages', in Yogendra Yadava and Warren Glover (eds) *Topics in Nepalese Linguistics*, Kathmandu: Royal Nepal Academy, 271–96.
- Bieri, Dora, Schulze, Marlene and Hale, Austin (1973) 'An approach to discourse in Sunwar', in Austin Hale (ed.) *Clause, Sentence, and Discourse Patterns in Selected Languages of Nepal*, Norman, Oklahoma: Summer Institute of Linguistics, 401–62.
- Bright, William (ed.) (1992) *International Encyclopedia of Linguistics*, Oxford: Oxford University Press.
- Driem, George van (1987) *A Grammar of Limbu*, Berlin: Mouton de Gruyter.
- (1993) *A Grammar of Dumi*, Berlin: Mouton de Gruyter.
- (1998) 'Sino-Bodic', *Bulletin of the School of Oriental and African Studies* 60/3, 455–88.
- Ebert, Karen H. (1993) 'Kiranti subordination in the South Asian areal context', in Karen H. Ebert (ed.) *Studies in Clause Linkage*, Zürich: ASAS-Verlag, 83–110.
- (1994) *The Structure of Kiranti Languages*, Zürich: ASAS-Verlag.
- (1997a) *Athpare Grammar*, Munich: Lincom.
- (1997b) *Camling*, Munich: Lincom.
- (1999a) 'The UP-DOWN dimension in Rai grammar and mythology', in Balthasar Bickel and Martin Gänszle (eds) *Himalayan Space: Cultural Horizons and Practices*, Zürich: Völkerkundemuseum, 107–34.
- (1999b) 'Kiranti non-finite forms – an areal perspective', in Yogendra P. Yadava and Warren W. Glover (eds) *Topics in Nepalese Linguistics*, Kathmandu: Royal Nepalese Academy, 371–400.
- Glover, Warren W. (1974) *Sememic and Grammatical Structure of Gurung (Nepal)*, Oklahoma: SIL.
- Hansson, Gerd (1991) *The Rai of Eastern Nepal: Ethnic and Linguistic Grouping*, Kathmandu: Tribhuvan University.
- Michailovsky, Boyd (1988) *La Langue Hayu (Nepal)*, Paris: CNRS.
- Rai, Novel K. (1985) 'A descriptive study of Bantawa', unpublished PhD dissertation, Poona.
- Rutgers, Roland (1998) *Yamphu: Grammar, Texts, and Lexicon*, Leiden: Research School CNWS.
- Toba, Sueyoshi (1984) *Khaling*, Tokyo: ILCAA.
- Weidert, Alfons and Subba, Bikram (1985) *Concise Limbu Grammar and Dictionary*, Amsterdam: Lobster Publications.
- Winter, Werner (1986) 'Bantawa rV- <? An exercise in internal and comparative reconstruction', in Dieter Kastovsky and Aleksander Szwedek (eds) *Linguistics Across Historical and Geographical Boundaries*, Vol. 1, Berlin: Mouton de Gruyter, 763–72.
- (1991–92) 'Diversity in Rai languages: an inspection of verb stems in selected idioms', *Lingua Posnaniensis* XXXIV: 141–56.

# HAYU

*Boyd Michailovsky*

## 1 INTRODUCTION

The Hayu (Nepali *hāyu*, Hayu *wa:ju*) inhabit an area between 50 and 100 kilometres south east of Kathmandu in the valley of the Sun Kosi and across the Mahabharat range to the south. There are a dozen or so Hayu villages in this area, but the language is currently spoken in only two, Murajor in Ramechhap District, where the data presented here was gathered, and Manedihi in the Marin valley, Sindhuli District. There are probably about 200 speakers, all bilingual in Nepali, the Indo-Aryan national language.

Hayu was first studied by B.H. Hodgson (1857) during his retirement in Darjeeling, and rediscovered and studied by Michailovsky and Mazaudon (1973), Michailovsky (1988) in the village of Murajor. Since 1973 it has stopped being spoken by most children in Murajor, but it is still being learnt in Manedihi, 30 kilometres to the west. There is no contact between the two villages. Dialect differences are slight; Hodgson's material is closer to the dialect of Manedihi in a few respects.

Hayu is closely related to Shafer's (1955) East Himalayish Section (Bodic Division), which includes the languages known as Rai and Limbu in Nepal and extends from the Hayu area eastward to Nepal's eastern border and into Sikkim (India). For the phonological correspondences between these languages (including Hayu), see Michailovsky 1994. Shafer classified Hayu slightly apart with Chepang and Magar (spoken in central Nepal west of Kathmandu) in his West-Central Himalayish Section.

## 2 PHONOLOGY

### 2.1 Syllable structure

The syllable canon is  $(C_1)V(C_2)$ , where  $C_1$  represents either an initial consonant  $C_i$  or a cluster.

### 2.2 Vowels

There are seven vowel timbres, *i, ɪ, e* (front); *a* (central); *o, ʊ, u* (back).

Vowel quantity is distinctive in open first syllables of polysyllabic words – in particular on the stem syllable of verb forms (see below). Closed syllable vowels are realized as short.

Nasality occurs only on open, non-final syllables. Almost all nasal vowels are long and precede a stop  $C_i$  (of the following syllable). A very few words, mainly of a phonaesthetic nature, have short nasal vowels.

A slight glottal stop (not transcribed) separates sequences of two identical vowels as in *mii* 'that one', *maɪŋ* 'NOMINAL NEGATIVE'.

### 2.3 Consonants

There are three series of initial stops, plain (unvoiced), aspirated, and voiced. The syllable initial consonants ( $C_i$ ) are as follows:

dorso-velar: *k, kh, g, ŋ*

lamino-palatal: *c* ([tʃ]), *ʃ* ([dʒ]) (affricates); *x* ([x], [xw]) (fricative)

apico-alveolar: *ts, tsh, dz* (affricates), *s* (fricative)

dental: *t, th, d, n*

bilabial: *p, ph, b, m*

apico-alveolar tap: *r*

apico-dental (bi)laterals: *l* (voiced), *hl* (voiceless).

approximants: *y* (palatal), *w* (labiovelar).

aspirate: *h*.

The opposition between the apico-alveolar and lamino-palatal series is neutralized before front vowels, the realization being apico-alveolar. There is no aspirated lamino-palatal affricate. The fricative *x* is phonetically velar, with the labialized allophone [x<sup>w</sup>] before *a* and *o*.

The approximant *y* appears before all vowels; thus *yi* ‘blood’ is distinct from *i* ‘this’. The approximant *w* appears only before the vowels *a* and *o*.

Initial groups are relatively rare, except in phonaesthetic adverbs, where they appear to be somewhat favoured. The following initial groups occur: *kl, khl, gl; kr, khr, gr; pl, phl, bl*. All words with initial bilabial + *l* clusters have doublets with simple bilabial initials.

The inventory of syllable finals ( $C_f$ ) is: *p, t, k, m, n, ŋ, r, l, x, ʔ*. The finals *x, ʔ* do not occur in word-final position. *x* occurs only before voiceless stops and *ʔ* only before sonorants; it will be seen that they function as allophones of the final stops *p, t, k*.

There is only one series of final stops. Each final stop has two allophones, voiced and voiceless. The voiced allophone occurs before a voiced initial stop either within the same word or in very close juncture (final of a verb root before a following modal). It is realized with glottalization or laryngealization [*b̥, d̥, g̥*], as in [*bebdzere*] ‘bedbug’, [*id bi:to*] ‘allow him to say’. In all other contexts syllable final stops are realized unreleased with simultaneous glottal closure. The final *x* is realized as [ç] after front vowels and [x] after back vowels or *a*.

### 3 MORPHOPHONOLOGY

Alternations are frequent in verbal forms, involving root finals and suffix initials. Those which are strictly phonologically conditioned are treated here; those which concern only particular roots or suffixes are treated with the morphology below.

Verb forms consist of a lexical root of the form ( $C_i$ )V( $C_f$ ), where  $C_f$  represents one of the finals *p, t, k, m, n, ŋ, r, l*, to which markers of tense, agreement, etc., are suffixed. There is no lexical opposition between roots with long and short or nasal and oral vowels. In considering the combinations at the boundary between root and suffix, three classes of ‘homorganic’ consonants are identified, bilabial, velar, and ‘coronal’, a class comprised of the lamino-palatal, apico-alveolar, and dental orders.

Among the regular alternations listed below, the first applies to suffix-initials, and is conditioned by the root final in verbal forms, while the others apply to root finals and are conditioned by the suffix initial.

- 1 After bilabial  $C_f$ , there is no opposition between velar and bilabial initials, the archiphoneme being realized as bilabial. (The resulting homorganic combinations are subject to the rules below.)

- 2 Stop  $C_f$  are realized as  $x$  before homorganic stop  $C_i$ .
- 3 Stop  $C_f$  are realized as  $\mathcal{P}$  before homorganic nasal or liquid  $C_i$ .
- 4 Nasal  $C_f$  are realized as vowel length and nasality before homorganic stop  $C_i$ . (This also applies to the realization of  $n$  before the fricative  $s$ ).
- 5 Before the homorganic nasal  $C_i$ , a nasal  $C_f$  is deleted, leaving an open root syllable with a distinctively short vowel. In contrast, the root syllable of forms of lexically open roots is always realized as long.
- 6 Before  $C_i s$  or  $tsh$ ,  $C_f t$  is deleted, leaving an open root syllable with a distinctively short vowel.

These rules apply whenever their conditions are met. The first can be stated as a neutralization, and the others as allophony, if one allows more than one phoneme to share an allophone. Thus, for example, syllable-final  $x$  or  $\mathcal{P}$  are the allophones of the  $C_f$  stop homorganic with the following initial, etc.

Other junctures between morphemes are looser than that between verb and suffix. Before postpositions, (1) and (4) above do not apply: *top-khen* 'after striking', *puxtay-khata* 'mat-PL'; (2) applies optionally *khok-khen-khox-khen* 'after walking'; (3) and (5) apply generally:  $[lat + noŋ]$  *la $\mathcal{P}$ noŋ* 'after going',  $[lɔn + noŋ]$  *lɔnoŋ* 'after running'.

The rules above have been presented as morphophonemic, as they operate at morpheme boundaries. However, there is reason to believe that they may have operated generally in the native vocabulary.

#### 4 VERB MORPHOLOGY

The suffixes of the verbal morphology are presented in Table 32.1. Roots are classified as 'intransitive' (vi), 'transitive' (vt or vtt or both, see below), 'deponent' (vd) or 'stative' (vs). Finite intransitive verbs show agreement with one argument, here referred to as S; transitive verbs show agreement with two arguments, A (loosely 'agent') and O ('object'). Deponent verbs may be regarded as transitives with impersonal A; they only use the forms in the row corresponding to third person singular A in the transitive paradigm. Stative verbs, essentially adjectives, have no finite forms. The tense distinction corresponds roughly to past vs non-past semantically. Past forms are listed under non-past forms in the table.

Morphological alternations occur in certain verb roots: between the vowels  $a$  and  $o$  in open roots (e.g. *dza-dzo* vt 'to eat'), and between open and t-final roots (e.g. *bi-bit* vtt 'to permit').

There are a few morphophonological alternations that affect particular suffixes. In Table 32.1 these are noted as follows:

The suffix transcribed  $N$  in some past forms nasalizes the preceding root final if it is a stop.

$\eta \sim N \sim s\eta$  is realized as  $\eta$  after an open root,  $N$  after a stop, and  $s\eta$  after a nasal.

¶ (in the  $n$  reflexive suffix element *nts*) has the effect of adding a final  $n$  (realized as nasality before *ts*) to most open roots.

*mi-m*: the assertive suffix *mi* optionally has the form *m* after a vowel.

##### 4.1 Structure of the paradigm

Transitive forms with first person singular O are identical to intransitive first person singular forms, and different from transitive forms with first person singular A. 1s  $\rightarrow$  2 forms have a special mark *no*.

Otherwise, most verbal suffixes mark tense, person, and number, without regard to transitivity or function. This is true of first and second person non-singular markers (e.g. 1de PA *tshoŋ*), and of non-past second person singular and third person markers.

**TABLE 32.1 VERB MORPHOLOGY**

*Transitive*

A/O	<i>1s</i>	<i>Idi</i>	<i>Ipe</i>	<i>Ipi</i>	<i>2s</i>	<i>2d</i>	<i>2p</i>	<i>3s</i>	<i>3d</i>	<i>3p</i>
<i>1s</i>					no Nno	notshe Nnotshe	none Nnone	η~N~soη koη	η~N~soη + tshe koηtshe	η~N~soη + me koηme
<i>Idi</i>		tshok tshoη								
<i>Idi</i>		tshik tshiη								
<i>Ipe</i>					kok kikoη	/tikok /tikkoη				
<i>Ipi</i>					kok kikoη	/tike /tikeη				

<i>2s</i>	ηo soη	tshok tshoη	tshik tshiη	kok (ki)koη	ke (ki)keη			∅ ko	/to	me kome
<i>2d</i>	ηotshe soηtshe							tshik tshe		
<i>2p</i>	ηone soηne							ne Nne		
<i>3s</i>	ηo soη							∅ ko	/to	
<i>3d</i>	ηotshe soηtshe							tshik kotshe	/to /toṣhe	
<i>3p</i>	ηome soηme							me kome	/tome	

*Intransitive*

<i>1s</i>										
<i>o</i> soη	tshok tshoη	tshik tshiη	kok (ki)koη	ke (ki)keη	∅	tshik tshe	ne Nne	∅	tshik tshe	me Nme

*Reflexive*

<i>1s</i>										
<i>nts</i> oη	natshok natshuη	natshik natshij	ntsikok ntsikoη	ntsike ntsikeη	ntse	natshik natshe	ntsine	ntse	natshik natshe	ntsime



1s,2s,3 → 3 past transitive forms have an added mark *k(o)* (in the suffixes *ko* and *kʉŋ*) which does not occur in the intransitive paradigm. The *ki* in 1p A past transitive forms no doubt reflects the same element, although it has intruded into the intransitive 1p forms as an optional variant. These suffixes, which indicate a third person object, will be referred to as ‘transitive k-suffixes’.

#### 4.2 Applicative forms

‘Applicative t-suffixes’ in *to*, *tʉŋ*, *ti* are listed in Table 32.1 in the same positions as the transitive k-suffixes. They are normally unmarked for tense. Transitive verbs which use the k-suffixes are identified as vt; those which use the t-suffixes as vtt. Many transitive verbs use both. For these verbs (identified by vt, vtt), the t-suffixes are used in ditransitive clauses (see below).

Transitive verbs with alternations between open and t-final roots use the latter (as well as t-suffixes) in applicative forms, giving them a full applicative conjugation (e.g. *bu-but* vt, vtt ‘to carry’).

#### 4.3 Negative indicative

The negative of a finite, indicative verb is indicated by *ma* or *makhɪ* placed before the postive form (cf. 9, 26, 31). The negative *makhɪ* ‘not yet’ is used with non-past forms only, but the sense (or at least the translation) is past (cf. 19).

#### 4.4 Imperatives, exhortatives, optative

Almost all positive imperatives are identical to the corresponding past indicative forms: *dza:ko* ‘eat it!’ (= ‘you ate it’), *sxtome* ‘kill them!’ *tesʉŋ* (*|tet + sʉŋ*) ‘let me go!’ *bʉŋme* ‘get up (pl)!’. Exception: the second person singular intransitive imperative is the same as the second person singular non-past indicative: *bʉk* ‘get up!’ (= ‘you get up’).

Negative imperatives are identical to the non-past indicative preceded by the negative imperative marker *tha* ‘don’t!’: *tha bʉk* ‘don’t (sg) get up!’, *tha bʉkne* ‘don’t (pl) get up!’, *tha dzo* ‘don’t eat it!’, *tha stme* ‘don’t kill them!’ *tha tetŋo* ‘don’t let me go!’

Exhortatives are simply non-past 1ps forms (cf. 18). Optatives add the suffix *ju*: *dzo-ju* ‘let him eat!’, *gorɪshik-ju* ‘[long] may we (1di) live!’

#### 4.5 Non-finite forms and verbal nouns

Non-finite forms: gerundive: the verbal root; infinitive: root + *mʉ* genitive; past/passive participle: root + *ta*; non-past/active participle root + *ji*; progressive gerundive: root + *ni* + root; conjunctive participle: reduplicated root + *ha* ‘instrumental/adverbial’. In the conjunctive participle, CV roots are reduplicated as CV:CV; C<sub>i</sub>VC<sub>f</sub> roots are reduplicated variously, most often as C<sub>i</sub>VC<sub>i</sub>VC<sub>f</sub>.

Verbal nouns are formed by adding suffixes to the root: agent nominal *ji*; theme nominal *ta*; patient nominal *tanj*; instrument nominal *canj*; locative nominal *luŋ*; time nominal *siŋ*; manner nominal *si*, *sina*; descriptive nominal (of a person) *tso*. Examples: *topji* ‘blacksmith’ (*top* ‘to strike’); *mexta* ‘carrion’ (*met* ‘to die’); *piptan* cigarette (*pip* ‘to suck’) (cf. 7); *topcan* ‘hammer’ (cf. 7); *tmluŋ* ‘bed’ (*tm* ‘to sleep’); *usiŋ* ‘moment of meeting’ (*ʉt* ‘to meet’); *umsɪŋ* ‘bedtime’; *husina* ‘voice, manner of speaking’ (*huɪ* ‘to speak’); *umtso* ‘sleepy-head’ (cf. 6).

All can also nominalize clauses. The negative (as with other nominals) is *maan̩*, e.g. *maan̩ noxta darbu* (NNEG be:PPT thing) ‘nonsense [“non-existent things”]’ (cf. 6).

#### 4.6 Non-productive derivational morphology

Hayu preserves traces of the well known Tibeto-Burman causative morphology in the form of word families, some thirty-five 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 ‘to tie’ vs *kik* vt, vtt ‘to put on a belt’; *dzol* vi ‘to live’ vs *tsol* vt ‘to save the life of’; *dok* vi ‘to fall’ vs *tok* vt, vtt ‘to drop’, ‘to cause to fall’; *bok* vi ‘to rise, to get up’ vs *pok* vt, vtt ‘to rouse, to raise’; *dat* vi ‘to run out’ vs *that* vtt ‘to use up’; *bek* vi ‘to enter’ vs *phek* vtt ‘to bring/take in’. A few examples do not involve stops: *im* vi ‘to sleep’ vs *hum* vt, vtt ‘to put to sleep’; *ot* vtt ‘to meet’, ‘to find’, ‘to reach’ vs *hU~hot*, vt, vtt ‘to look for’; *re* vs ‘to be broken in pieces’ vs *se* vtt ‘to break’ (dial. **x**); *ram* vi ‘to be afraid’ vs *xwam* vtt ‘to 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, *khata* ‘plural/collective’ (cf. 21).

There are a number of derivational suffixes: *tso* ‘human’, ‘collective human’, ‘male human’, *mi* ‘female human’, *wo* ‘male human’, *st* ‘fruit’. Examples: *thartso* ‘grandson’, *kartso* ‘friend’, *juxkutso* ‘married couple’, *ta:wo* ‘son’, *tha:mi* ‘granddaughter’, *ta:mi* ‘daughter’, *kokst* ‘*Ficus semicordata*’. (See also deverbal nominals, below.)

#### 5.2 Pronouns

The independent absolutive pronouns are: first person: *gu* ~ *guu*; second person: *gon* (2s), *gontshe* (2d), *gone* (2p); third person (human) *komi*. The first and third person pronouns, like nouns, can optionally be followed by *nakpu* ‘two’ or *khata* ‘PL’ to show number. The demonstrative *mi* may also be used as a third person pronoun; it is not limited to human referents.

The ergative postposition *ha* is added to the absolutive form; the first person ergative form is *ga*, and the second person singular form *gona* occurs as a variant of *gon-ha*.

The possessive pronouns are: 1s *an̩*; 1de *ant̩she*; 1di *uñt̩she*; 1pe *ã:ki*; 1pi *ũ:ki*; 2s *uñ*; 2d *uñt̩she*; 2p *uni*; 3s *a*; 3d *atshe*; 3p *ami*. These forms are used before nouns to indicate possession (cf. 1, 12, 21, 25, 30), and before postpositions other than *ha* (cf. 9, 10, 15).

#### 5.3 Demonstratives

The demonstratives are *i~ii* ‘this’, *mi~mii* ‘that’. Only the shorter variants are used before suffixes or postpositions: *mi-ha* ‘he-ERG’, ‘that-INST’; *mi-khen* ‘then’, ‘from there’, *i-noñ* ‘here’, *mi-noñ* ‘there’. Also: *itha* ‘this much’ ‘(towards) here’, *mitha* ‘that much’; *ithara* ‘to this extent (with an adjective)’, *ine* ‘here’, *ibe* ‘here’; *ima* ‘in this manner’ etc.

### 5.4 Interrogative-indefinite words

The indefinite-interrogative words are *su* ‘who, someone’, *mtsɪ* ‘what, something’, and a series of words based on the indefinite-interrogative morpheme *ha* (only interrogative glosses are cited): *hanɔŋ* ‘which?’, *hatha* ‘how much?’, *hani* ‘where?’, *hakhi* ‘when?’, *haŋa* ‘how?’, etc. In indefinite uses these words may carry the indefinite suffix *dum* (cf. *dum* vi ‘to become’): *hatha-dum* ‘however much’. In negative contexts they are often reinforced by the particle *le* ‘also, even’ (cf. 9).

### 5.5 Numbers

Nepali numbers are used for numbers over four, and often for lesser numbers as well. The numbers from 1 to 4 show a rudimentary classifier system: *pu* ‘human’ vs *uŋ* ‘non-human’ ([ʔuŋ] after a vowel). The numbers are: *koŋpu* ‘1:HUM’, *kolu* ‘1’ (human and non-human); *nakpu*, *nauŋ* ‘2’; *tshukpu*, *tshuuŋ* ‘3’; *bliuŋ* ‘4’ (rarely used).

### 5.6 Adjectives, adverbs

There are at most a few lexical adjectives apart from stative verbs.

Colour terms generally carry the suffix *mi* (?<genitive *mɔ*) or are marked by the postposition *ba-baha-baŋa* ‘like’: *dawaŋmi* ‘white’, *jitshiŋ-baha-mɔ khursānī* ‘red chili-pepper’.

There is a very large class of phonaesthetic words which may function as manner adverbs (often marked by the postposition *ha*) or as adjectives (often marked by the postposition *mɔ*), e.g. *kulkul* ‘round’: *kulkul-mɔ mɔrli* ‘round belly’; *kulkul-ha xwārtse?* ‘are you fully [“roundly”] sated?’ Some phonaesthetic manner adverbs are marked by a derivational element *sit*, e.g. *krmsit* ‘in/to the right spot’: *krmsit to:ta* ‘safely put away’.

The basic locative adverbs are *wani-ani* ‘up’, *huti-uti* ‘down’, *wati-ati* ‘across’; the corresponding directionals are *jō:kha* ‘down’, *lō:kha* ‘up’, *dokha* ‘across’.

### 5.7 Postpositions, particles

Postpositions serve to mark the function of an element, which may be an NP or a subordinate clause, in a larger unit, NP or clause. Among the most common: *ha* ‘ergative, instrumental, adverbial’; *khen* ‘from’, ‘after’; *mɔ* ‘genitive’; *thik* ‘like’; *noŋ* ‘at’, ‘with, after’; *noŋmo* ‘from’; *le:si* ‘for’ (*mɔ-le:si* after nouns and, optionally, after oblique pronouns); *he* ‘in, while’; *boŋ* ‘as far as’, ‘as much as’; *tiliŋ* ‘because’; *bhandā* ‘than’ (cf. 5); *sa* subordinator (cf. 26).

Conjunctions: *nom* ‘if’ (cf. 27); *phen* (marks both protasis and apodosis in contrary-to-fact conditionals).

Discourse particles: *ko* ‘TOPIC’, *le* ‘also, even’, *na* ‘precisely, indeed, the very’, *dzi* ‘only’, *pai*, *pi* ‘counter-expectancy focus’.

Sentence particles: *ro* ‘rhetorical question’, *re*, *are* ‘hearsay’, *te* ‘insistence’.

## 6 SYNTAX

A sentence consists essentially of a finite clause, that is, a clause with a conjugated indicative or imperative verb, and, optionally, core and non-core arguments consisting of noun phrases, adverbs, or subordinate clauses. The verb is last in its clause, except when a topical argument is right-dislocated after the verb (cf. 11, 21, 31). There is no requirement that core arguments be present in the clause in lexical or pronominal form. Case marking is ergative: that is,

A appears in the ergative case, marked by the postposition *ha*, while S and O appear in the absolutive case (unmarked). Functors – postpositions, discourse particles, complementizers, etc. – 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* (cf. 2, 3, etc.). This mark is not used in interrogative, imperative, or irrealis clauses. Questions are marked by intonation, or by tags like *ki ma* ‘or not’ (cf. 19, 31).

### 6.2 The intransitive clause

The verb in an intransitive clause shows agreement with an S, in the absolutive case.

- (1) *aba aŋ xwaptso met*  
 now 1s:OBL spouse die:3s:NPA  
 ‘Now my wife will die.’
- (2) *nukun le gu itha belā na phi:ŋomi*  
 tomorrow too I this.much time EMPH come:1s:NPA: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. The verb in the following is morphologically unambiguously intransitive:

- (3) *mi-khata buti-ha benmem*  
 that-PL meal-INST satisfy:3p:PA:ASS  
 ‘They had enough food’. *lit.* ‘They were satisfied by the food.’

Meteorological verbs are intransitive, with 3s morphology: *lo:gaŋ kak* (sun shine:3s:NPA) ‘the sun is shining’; *na?num tŋmi* (rain rain:3s:PA:ASS) ‘it rained’ or ‘it is raining’ (the past used with inchoative sense).

### 6.3 Copular sentences

Nominal or attributive predicates require the copula, the stative copula being *no~not vi* (irregular) ‘to be’. Two types of copular sentence are distinguished below, attributive/identificational, and situative/possessive/existential. The order in the second type depends on information structure; in addition, a topical argument may be omitted or right-dislocated. Inchoative and causative senses are expressed by using verbs like *dum* ‘to become’, *pa~po* vt, vtt ‘to make’, *pon vr* ‘to make oneself’, ‘to pretend to be’, etc., in place of the copula (cf. 5, 6, 10). Non-verbal predicates are negated by *maŋ* (cf. 6).

Attributive: subject – attribute (NP, modifier, quantifier) – copula:

- (4) *e gon su nono? – gu no:ŋom*  
 hey you who be:2s I be:1s:NPA:ASS  
 [In the dark.] ‘Who are you? – It’s me!’
- (5) *pukhŋli-bhandā mi xō:ta dum*  
 king-than that big:PPT become:3s  
 ‘He became richer than the king.’

- (6) *stst-ha gon maṅ stxso pō:tse*  
 know:CJP you NNEG know:VN make:2s:RF  
 ‘Although you know, you play dumb.’

Situative/possessive/existential: subject – situative or possessive complement – copula:

- (7) *mi noktshuṅ-he siktṅ mi ho:caṅ-he no:mi*  
 that ear-in wear:NV:PATIENT that open:NV:INST-in be:3s:ASS  
 ‘The earring is in the box.’
- (8) *tsa:pimi nom are*  
 mother-in-law be:3s:ASS HS  
 ‘Once there was a mother-in-law, they say.’ (beginning of a story)
- (9) *ā:ki-thik-mṀ wa:ju hani le ma no:me*  
 1pe:OBL-like-GEN Hayu where also NEG be:3p  
 ‘There are no Hayu like us anywhere.’
- (10) *a-mṀ ko lom-noṅ ko nakpu xwaptso dum –*  
 3s-GEN TOP road-on TOP two:HUM spouse become:3s  
*kem-he kolu xwaptso no [...] jamma ko tshukpu*  
 house-in one spouse be:3s all TOP three:HUM  
*dume*  
 become:3p  
 ‘Now he had two wives on the road – he had one at home... altogether they were three.’

#### 6.4 Stative verbs

Stative verbs, e.g. *xṀn* vs ‘to be big’, *wol* vs ‘to be withered’, have no finite forms, but appear in intransitive clauses, either in the form of a past/passive participle used as a predicate attribute (cf. 5), or the gerundive used with the modal *la-lat* ‘to go’ (see below): *xṀn laxtse* ‘it’s too big’, *wol lasṀṅ* ‘I have become withered up’.

#### 6.5 The transitive clause

The verb in a transitive clause shows agreement with two core arguments, A, roughly speaking the agent, marked by the ergative postposition, and O, absolutive (unmarked):

- (11) *ga thuṅnomi gon*  
 1s:ERG convey:1s→2s:ASS 2s  
 ‘I will take you there.’
- (12) *kolu buḏhā mānche-ha a bārī-noṅ lalat-ha rā:pi luxtom are*  
 one old man-ERG 3s:OBL field-in go:CJP taro plant:3s:PA:ASS HS  
 ‘An old man went to his field and planted taro, they say.’ (Beginning of a story.)

#### 6.6 Deponent verbs

Transitive verbs may have inanimate A:

- (13) *turī nuṅma-kheri naṅnum-ha thopsṀṅ,*  
 today day-time rain-ERG throw.down:3s→1s:PA

*gu dzū:sa-ha usŋ*  
 1 fever-ERG meet:3s→1s:PA

‘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 in the forms above is homophonous with the 1s intransitive past, the verbs *thop* and *ut* are transitive, as can be seen in other contexts, where they take personal A, and in *dzū:sa-ha ūxtō* (fever.ERG meet.3s→3s) ‘he caught a fever’, where the form is unambiguously transitive. The verb in 14 is somewhat different, since it can never take a personal A, but the form used with a third person O is unambiguously transitive:

(14) *komi ti-ha thimtome*  
 3 water-ERG splash:3s→3p

‘They got sprinkled with water’ (*thim* ‘to be struck from above (by a falling object)’).

Verbs of this type, of which there are about a dozen, are called ‘deponent’ (vd); other examples are *kot* ‘to be pricked’, *dzt* ‘to be affected by a bad odour’. They can be regarded simply as transitives with impersonal A and personal O. However, the impersonal A, unlike other ergatives, can be omitted without implying anaphora: *thimto* means ‘he was splashed’ and not necessarily ‘he was splashed by it’; this is perhaps not a proper syntactic property. (See also relative clauses below.) In addition to using transitive 3→3 forms, all such verbs use the applicative t-suffixes rather than k-suffixes.

## 6.7 The ditransitive clause

Many transitive verbs can appear in a ditransitive clause, with an A and two absolutive arguments, O and O2 (the second object). The verb shows agreement with A and O. Semantically, O2 generally corresponds to a theme or patient, and O to a beneficiary, goal, or to (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 O. The lack of such a form, however, is no obstacle to the use of the ditransitive construction.

The verbs *ha vtt* ‘to give’, *mum vtt* ‘to give (food to eat)’, *thun vtt* ‘to give (to drink)’ are inherently ditransitive, always agreeing with the recipient as O.

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 *rŋk vt, vtt* ‘to plough’ and *ke-ke vt, vtt* ‘to peel’: *ga rŋkŋmmi* (–APP) ‘I ploughed it’, vs *ga rŋktŋmmem* (+APP) ‘I ploughed it for them’; *ke:ko* (–APP) ‘I peeled it’ vs *kexto* (+APP) ‘I peeled it for him’. The verb *rŋk* in a ditransitive clause can also mean ‘to use (a bullock – O) to plough (a field – O2)’.

With other verbs, 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 O2 is human, or otherwise particularly salient (like the animate, moving goal of *tot vtt* ‘to chase’, which always functions as O), and first or second person O2s are never encountered. Where the ditransitive construction cannot be used, a beneficiary can be introduced marked by the postposition *lesi* ‘for’. As an example, with the verb ‘to kill’ only the most insignificant victim can be displaced from the function O to O2 to make room for a beneficiary. Thus the ditransitive construction is used in 15 but not in 16:

(15) *gu sek stsŋ – stŋno*  
 1 louse kill:2s→1s:IMP:±APP kill:1s→2s:NPA:±APP  
 ‘Kill my lice for me! – I’ll kill them for you.’

- (16) *komi-ha aŋ-le:si kolu xo:co sɪxtom*  
 3-ERG 1s:OBL-for one chicken kill:3s→3s:±APP  
 ‘He killed a chicken for me (i.e. in my honor as a dinner guest).’

In 15, not only are the lice beneath notice, but their host, a salient speech-act participant, is directly implicated (more so than the honoured guest of 16) because the action takes place on his body. The fact that the verb *su* vt ‘to kill’ has no distinctively applicative forms, and that *sɪxʉŋ* which usually means ‘kill me!’, and *sɪʃno* ‘I will kill you’, is no obstacle to the use of the construction.

In some ditransitive clauses, O and O2 may have the same referent. Thus, related to *paxkʉŋ* (–APP) ‘I raised it’, ‘I roused him’, etc., we find *pʉktʉŋ* (+APP) ‘I raised it for him’, ‘I roused him<sub>i</sub> for him<sub>j</sub>’ but also ‘I helped him up’ (i=j). That the latter is ditransitive is clear from the use of a distinctively applicative form.

### 6.8 Reflexive clauses

Reflexive forms agree with a single core argument, which appears in the absolutive case. Most or all transitive verbs may have reflexive forms, indicating that A and O have the same referent. Non-singular reflexive forms have either reflexive or reciprocal sense (but cf. the specialized compound reciprocal form below).

- (17) *mi wolta cāhĩ u wathe muxtsem txtse*  
 that wither:PPT TOP there over.there sit:3s:RF:ASS say:3s:RF  
 ‘That withered one sat way over there, they say.’

- (18) *kar:tso cʉʃnatshik*  
 friend look:2di:RF:NPA  
 ‘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:ʃtse!* ‘eat your meal!’ (verb *dza~dzo* vt ‘to eat’) the reflexive indicates identity of reference between the eater and the beneficiary; *buti* ‘cooked grain’ can only have the function O2. On the other hand, when the disappearance of a particular food item is in question we find transitive forms:

- (19) *e, gona dza:ko ki haŋa? – makht dzaŋmi*  
 hey 2s:ERG eat:2s:PA or how? not eat:1s:NPA:ASS  
 ‘Hey, did you eat it or what? – I didn’t eat it.’

Similarly, in threats – *ga gon dzo:nom* ‘I’ll eat you’ (cf. 27) or when a child eats something other than food or medicine – *ko dzo:mi* ‘he eats earth!’ the reflexive is not used.

### 6.9 Compound predicates

A number of operators such as modals, aspectuals, directionals, etc., appear as coverbs following the bare-root gerundive of either the main clausal verb or another coverb. In such VV concatenations, only the (last) coverb can have a finite form. This coverb often agrees in transitivity with the governed verb and, if the latter is transitive, may agree with its logical object. Thus, in the examples below, *cut* ‘finish’ appears as a reflexive with the intransitive ‘come’, but it appears as a transitive with ‘eat’, showing agreement with its object.

- (20) *kem-he phi cuxtse*  
 house-in come finish:3s:RF  
 ‘He already came home.’

- (21) *ā:ki gāū-mU tso-khata jamma dza cuxtomem bumi*  
 1pe:OBL village-of child-PL all eat finish:3→3p:ASS bumi  
*pixpi-ha*  
 g'mother-ERG  
 'Grandma Bumi had already eaten up all the children of our village.'

The verb *la~lat* vi, vr 'to go' has a number of uses as a coverb: centrifugal directional: *lun la?natshe* (run go:3d:RF) 'they (dual) ran away'; change of state with stative verbs (which themselves have no finite forms): *ima wol lasoŋ* (thus withered go:1s:PA) 'I have become withered up like this'; agentless 'go-passive': *ma jeŋ la* (not see go:3s:NPA) 'it will not be seen'. The latter construction is also used with intransitive verbs: *a thum dzik laxtse* (3s:OBL heart break go:3s:RF) 'her heart broke'.

Other coverbs taking the root gerundive are *dak* 'to be obliged/necessary to', *phat* vi and *on* vtt 'to be able', *ktn* vtt 'to try', *wat* vtt 'to stop', *hok* invariable, impersonal 'is suitable', *piŋ* vt causative (= 'to send so. '), *bi* vtt 'to permit', *dzok* vi 'to know how to', *dan* invariable 'to be about to' (used with the copula when a main verb). The verbs *ha* vtt 'to allow', *bi* vt 'to request permission to', *liŋ* vt 'to get to' take infinitive complements. In the causative construction, the causee appears in the absolutive case:

- (22) *ga gon pheri xwan-xwan dza piŋnom*  
 1:ERG 2s again satisfy-satisfy eat cause:1s→2s:ASS  
 'I will let you eat your fill again.'

Purpose clauses appear with the verb in the gerundive form: *siŋ pa la?natshe* (wood make go:3d:RF:ASS) 'they (dual) went to cut wood'.

The reciprocal is formed with the fully reduplicated gerundive followed by the coverb *pa~po* vi 'to do': *āsik ha-ha po:tshē* (blessing give-give do:3d:PA) 'they exchanged blessings'.

## 6.10 Subordinate clauses, sentential complements

The function of a subordinate clause adjunct in a higher clause is marked by a postposition, exactly as an NP adjunct. Almost all the postpositions used are the same as those used with NPs. The subordinate clause has verb-last order, with the verb in the gerundive form, except with the postposition *ha*, where the verb appears in reduplicated form (conjunctive participle).

The subordinate clause with the conjunctive participle usually shares its S or A argument with the S or A of the higher clause, which determines the case of this argument. The sense is loosely adverbial, either of temporal priority ('after x, y'), or of simultaneity, or of manner. Thus, in 12 above, the S of the subordinate has the same referent as the A of the main clause, and appears in the ergative case. The S-or-A-sharing rule, which might be invoked to justify the identification of a 'subject' pivot, is not without exceptions:

- (23) *syāl hū:pukumi-ha tsitsik-ha men*  
 jackal wasp-ERG bite:CJP die:3s:PA  
 'The jackal was stung by the wasps and died.'

The same construction with the copula is occasionally used as a kind of passive:

- (24) *gu ko xō:ta pōkhōli-ha ima tt-ha no:ŋom*  
 I TOP big chief-ERG thus say:CJP be:1s:NPA:ASS  
 'I have been commanded thus by the great chief.'



Subordinates marked by other postpositions have no particular tendency to share S-or-A with the higher clause, often sharing no argument at all:

- (25) *a umu khok-khen danḍaṅ le dum*  
 3s:OBL mother walk-after bright also become:3s  
 ‘After her mother left, it became light.’

The subordinator *sa* always governs a gerundive negated (oddly, for a non-finite form) by *ma*, as in *kokori ma ok-sa* (cock NEG cry-SUB) ‘before cock-crow’.

- (26) *gon kem-he lalat-ha ma dza-sa ma teʔno*  
 2s house-in take:CJP NEG eat-SUB NEG release:1s→2s:NPA  
 ‘I’ll take you home and I won’t let you go without eating you!’

Conditions are marked by **nom** ‘if’ (<**nom** ‘be:3s’):

- (27) *ot nom dzo:mi txtsime*  
 catch:3s→3s:NPA if eat:3s→3s:NPA say:3p:RF  
 ‘If he catches her he’ll eat her, they say.’

Complements of verbs implying citation are usually marked by the postposed complementizer *paha-pa*, composed of the verb *pa~po* vt, vtt ‘to do’ with the manner-adverbial post-position *ha*. The complementizer may be omitted.

- (28) *su:ŋom paha txtom txtse*  
 scratch:3s→1s:NPA:ASS COMP say:3s→3s:PA:ASS say:3s:RF  
 ‘“It’s clawing me,” she said.’

- (29) *jamma suko ine ine nom paha*  
 all know:3s→3s:PA 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 topic marker *ko* (cf. 10, 17, 31) is not required.

Focal arguments tend to be placed immediately before the verb. The contradiction of a focal element with what is presupposed may be emphasized by the particle *pi~pai*: *jexpi pi khuwāi pozsŋmi* (shit FOC feed do:3s→1s:PA:ASS) ‘he fed me SHIT!’

## 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 *mŋ* (cf. 21), participles, relative clauses, lexical adjectives, etc. Thus: *rā:pi-mŋ lo* (taro-GEN leaf) ‘taro leaf’; *haŋa-mŋ siŋtoŋ* (how-GEN man) ‘what manner of man?’; *xō:ta pokhŋli* (big:PPT chief) ‘big chief’; *ŋlawo phum* ‘huge tree’. Quantifiers may precede or follow; in the latter case any postposition is repeated, e.g. *nonotso-ha nakpu-ha*

(sisters-ERG two:HUM-ERG). Adverbial modifiers precede adjectives: *ithara xō:ta le gotji* (thus big:PPT foot have:APT) ‘having feet big like this’.

Any potential nominal modifier may serve as head. That is, ‘big’ may mean ‘big one’ (cf. *wolta* in 17).

Coordination is usually realized by simple juxtaposition: *ga got le tsikno* (1s:ERG hand foot break:1s→2s:NPA) ‘I’ll break your arms and legs!’ (ditransitive). The postposition *noŋ* ‘after,with’ may also be used: *bhālu-noŋ syāl toxtotshem* (bear-with jackal chase:3s→3d:PA) ‘he chased the bear and the jackal’ (note the dual object 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 antecedent is omitted from the relative. When the antecedent corresponds to the intransitive S in the relative, the form used depends on temporal considerations, the past/passive participle in *ta* being used for states (stative verbs have only this form) or for completed events, and the agentive/active participle in *ji* for non-past time:

- (30) *āki lañkā-khen dzokta deutā*  
 1pe:OBL Ceylon-from come.up:PPT god  
 ‘our goddess who came up from Ceylon’

- (31) *inoŋ jō:kha latji siŋtoŋ nakpu gona jē:ko ki ma?*  
 here below go:APT man two 2s:ERG see:2s→3s:PA or not  
 – *ma jē:kōŋ, thattso, ga ko*  
 – not see:1s→3s:PA 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: *yeso tyeso pa:pa-ha dzo:zi siŋtoŋ* (this.way that.way do:CJP eat:APT man) ‘a man who lives by chicanery’ (*lit.* ‘eats by doing thus and so’); 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 had been wearing’. When the antecedent is the impersonal A of a deponent verb, the active participle is used: *mi kotji tsu* ‘the thorn which is pricking me’.

When a pronoun S or A is expressed in the relative, it appears in the oblique/possessive form: *gon-ha aŋ txa thāū-he thuŋo ta* (2s-ERG 1:OBL say:PPT place-in escort:2s→1s:IMP TOP) ‘you just take me to the place I say [and]’.

When the antecedent corresponds to an adjunct in the relative, the past/passive participle in *ta* can generally be used: *ima-mō lxta thāū-noŋno* (thus-GEN plant:PPT place-from) ‘from the place where it had been planted like this’. More often, however, the appropriate nominal is used: *bumi unluŋ kem-noŋ* (bumi sleep:NV:LOCATIVE house-at) ‘in the house where Bumi slept’. The absolutive (not genitive) case of ‘Bumi’ shows that it is the whole clause ‘Bumi sleeps’ that is nominalized.

### ADDITIONAL ABBREVIATIONS

1s, etc.	1, 2, 3 represent persons, s singular, d dual, p plural, i inclusive, e exclusive
3s → 1d, etc.	3rd person singular A, 1st dual exclusive O.
APP	applicative
APT	active/present participle

ASS	assertive
CJP	conjunctive participle
EMPH	emphatic particle ('the very')
ERG	ergative
HS	hearsay, reported speech
HUM	human classifier
IMP	imperative
NEG	negative
NNEG	nominal negative
NPA	non-past
OBL	oblique/possessive
PA	past
PL	plural
PPT	past/passive participle
RF	reflexive
SUB	subordinator
vd	deponent verb
Vi	intransitive verb
VN	verbal noun
vr	reflexive verb
vs	stative verb
Vt	transitive verb using k-suffixes
vtt	transitive verb using t-suffixes

Bold italic words in the transcriptions are Nepali loanwords; they appear in transliterated Nepali orthography.

## REFERENCES

- Hodgson, B.H. (1857) 'Grammatical analysis of the Hayu language', *JASB* 26: 373–475.
- Michailovsky, Boyd (1988) *La langue hayu*, Paris CNRS.
- (1994) 'Manner vs place of articulation in the Kiranti initial stops', in Kitamura, H. *et al.* *Current issues in Sino-Tibetan Linguistics*, Osaka, 1994.
- (2000– ) Hayu recorded texts, with synchronized transcription and translation. <http://lacito.archivage.vjf.cnrs.fr>.
- Michailovsky, Boyd and Mazaudon, M. (1973) 'Notes on the Hayu language', *Kailash* 1.2: 135–52.
- Shafer, R. (1955) 'Classification of the Sino-Tibetan languages', *Word* 11.1: 94–111.

# CAMLING

*Karen H. Ebert*

## 1 INTRODUCTION

Camling is one of the many endangered languages of the Kiranti cluster (see Chapter 31). 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 *-nA* in NW, *-nA* or *-ki* in SE.

## 2 PHONOLOGY

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>thuŋma</i>	cough	<i>tuŋma</i>	village
		<i>duŋma</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* (*khru*+*ma*) ‘get up’. Five vowel phonemes and four 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
<i>khōma</i> ( <i>khaŋ-</i> )	see		

TABLE 33.1 CONSONANT PHONEMES

	<i>Labial</i>	<i>Dental</i>	<i>Alveolar</i>	<i>Velar</i>	<i>Glottal</i>
stop, –voice	p	t	c [ts]	k	
	ph	th	ch [ts <sup>h</sup> ]	kh	
+voice	b	d	(j) [dz]	(g)	
	bh	dh	(jh) [dz <sup>h</sup> ]	(gh)	
fricative	f	s			h
nasal	m	n		ŋ	
	mh	nh			
continuants	l, lh, r, rh				
glides	w, y				

*khaima* (*khat-*) go      *kheima* (*khet-*) split, cut up (SE)  
*maima* (*maid-*) make      *muima* (*muit-*) ripen  
*məima* (*məid-*) 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 *ɤ* is unclear. As there is no [a] before *r* in verb stems, a pair like *chorsyu* ‘he payed it’ vs *chɤrsyu* ‘he urinated’ 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 *-nɤ*, but both are often pronounced the same ([nɤ] ~ [nɔ]).

The canonical syllable structure is CV(C). The NW dialect has initial consonant clusters, restricted to *p(h)*, *k(h)+r*, *l*: NW *khɤrupma* ‘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 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. 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* (cf. 8a) are also possible.

TABLE 33.2 PERSONAL PRONOUNS, POSSESSIVES AND DEMONSTRATIVES

	<i>pers. pronouns</i>	<i>poss. pronouns</i>	<i>poss. prefixes</i>
1s	<i>kaŋa ~ kã</i>	<i>aŋa</i>	<i>a-</i>
1di	<i>kaici</i>	<i>icmo ~ Acmo</i>	<i>ic- ~ Ac-</i>
1de	<i>kAccka</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 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-tuŋma* ‘our<sub>c</sub> village’, *i-la* ‘our language’ (also when speaking to a non-Camling). Only *aŋa* 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 aŋa</i>	this is mine
<i>aŋa a-khim = a-khim = aŋa 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-dunŋkha* ‘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’): *was*a ‘chicken’, *ngasa* ‘fish’, *sosa* ‘porcupine’, *lusa* (frog species), *bosa* (\*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 class a noun both in the meaning ‘bird’ and ‘water’: *tuwa* ‘pheasant’, *khlaw*a ‘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’, *masuŋma/pasuŋpa* ‘old woman’/‘man’, *wama/wapa* ‘hen’/‘cockrel’. The gender suffix can be replaced by *cha* ‘child’ to express a diminutive: *Khocilipa*, ‘the cultural hero’, is also called *Khocili-cha*. 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’, *hinkha* ‘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/ḡosi</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 are exemplified in: *i-ruḡ* ‘once’, *haka-ruḡ* ‘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 a combined ergative-instrumental case marker *-wa* and a genitive *-mo*. The suffix *-lai* is borrowed from Nepali and optional with human undergoers. All can be demonstrated with the following constructed example:

(1)	<i>Ram-mo</i>	<i>m-nicho-wa</i>	<i>a-woini(-lai)</i>	<i>bhe-wa</i>	<i>ap-u.</i>
	(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 *ra* ‘and’: *kusya (ra) pucho* ‘youngest daughter and the snake’. Comitative meaning can sometimes be expressed with the postposition *iōda* ‘near.’

Camling shares with other Kiranti languages a unique system of case 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’ (cf. 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 (cf. Ebert 1999).

## 4 THE VERB

### 4.1 Stems

The full verb stem occurs only before vowels. Before a consonantal suffix stem-final *t*, *d*, *s* are elided. This results in diphthongization of the stem vowels *a*, *e*:

<i>stem</i>	+ <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 consonantal 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 stem-final *s*.

<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 somebody

Secondary causatives are possible with a few verbs, e.g. *set- + t > seid-* ‘kill for someone’.

- (2) *kap-ghicro-da*    *a-ma*                    *sor*    *bo*    *nyabd-e-ŋas-e*.  
 2sPOSS-neck-LOC    1sPOSS-mother    louse    PART    stick-NPT-V 2:stay-NPT  
*-seid-uŋ-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 also as primary causatives: *sei maima* ‘kill for someone’, *khō 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-uŋa/hupd-āic-uŋa* ‘I washed myself’, *mi-ras-umca/mi-ras-āica* ‘they separated’ (cf. *ras-* ‘divide’).



## 4.2 Person and number affixes

Camling, like the neighboring 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-uŋa</i>	I went
1s→3s	<i>phlod-uŋa</i>	I helped him
3s→1s	<i>pa-phlod-uŋa</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 <sub>p</sub> helped him
	2-help-3P-1/2pA	

TABLE 33.3 CAMLING PERSON-NUMBER MARKERS NW DIALECTS

A	P											
	<i>1s</i>	<i>1de</i>	<i>1pe</i>	<i>1di</i>	<i>1pi</i>	<i>2s</i>	<i>2d</i>	<i>2p</i>	<i>3s</i>	<i>3ns</i>	<i>itr.</i>	
1s						-na	-naci	-nani	-uŋa	-uŋcuŋa	-uŋa	
1de									-cka			
1p									-umka	-umcumka	-(i(m)ka)	
1di									-ci			
1pi									-um	-umcum	-i	
2s	ta- -uŋa	ta- -cka	ta- -(i(m)ka)						ta- -u	ta- -ucyu	ta-	
2d									ta- -ci		-	
2p									ta- -um	ta- -umcum	ta- -i	
3s	pa- -uŋa	pa- -cka	pa- -(i(m)ka)	pa- -ci	pa- -i	ta-	ta- -ci	ta- -i	-u	-ucyu		
3d									pa- -ci		-ci	
3p									pa-	pa- -ucyu	mi-	

SE dialects: 1st person patient scenarios

	<i>1s</i>	<i>1de</i>	<i>1pe</i>	<i>1di</i>	<i>1pi</i>
2s	kha-ta-				
2d	kha-ta- -ci				
2p	kha-ta- -i				
3s	kha-				
3d	kha- -ci				
3p	kha-mi-				

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 /*m*/ with the allomorphs *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
3d→3	<i>loda-ci</i>	they told him/them
=1d→3		we told him/them

The homophonous third non-singular patient marker *-ci* always follows *-u* and the person-number markers *-m*, *-ŋ*, 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>	I will tell them
	tell-3P-1s-3nsP-[copy]-NPT	
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<sub>d</sub> saw him/them’; ‘we<sub>di</sub> saw him/them’.

### 4.3 Inverse configurations

The prefix *pa-* is an inverse marker 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 him/them
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 33.3 demonstrates, the system is disturbed. Second person *ta-* does not combine with inverse *pa-* due to a one-prefix-restriction, *-uŋa* (<u + ŋa) 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 (cf. ‘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 left 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 <sub>d</sub> told me/us
3p→1di	<i>pa-loda-ci</i>	3p→1	<i>kha-mi-loda</i>	they <sub>p</sub> 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 33.3. 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 have 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*.

	PT (unmarked)		NPT	
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→2s	<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. *hija* 'he was' (not: 'he became', as in most neighbouring languages), and in periphrastic tenses  $\emptyset$  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 *nhame* is a statement based on general knowledge.

A resultative-perfect is expressed by a compound verb formation with the verb *ɲas-* 'stay', 'remain, keep'. The SE dialect has a progressive marker *-uɲs* ~ *-ōs* (contracted from *ɲas-?*). Both dialects have a grammaticalized ambulative marked by *-hod*:

<i>mi-khat-e-ɲas-e</i>	they have gone/they are gone
3ps-go-NPT-V2:stay-NPT	
<i>rō m-uɲs-yo</i> (SE)	she is preparing food
rice make-PROG-3P:NPT	

*suŋ khop-hod-yo* 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!*  
*dingu! diŋaci! diŋani! drink it!*

First person patient forms differ according to the dialectal inverse paradigms: NW *iduŋa!* 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* (a possible 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-uŋa</i>	<i>pa-khai-n-uŋa</i>	I did not go
1di	<i>ta-khata-ci</i>	<i>pa-khata-ci-na/-c-āi/-c-aina</i>	we did not go
2d	<i>ta-khata-ci</i>	<i>ta-khata-ci-na/-c-āi/-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 nonpast, with *-ō* in forms that have no personal suffixes and *-mi* in first and second person forms (again with some variation): *pa-los-ō* ‘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 *ŋas-*.

	<i>asserted</i>		<i>negated</i>	
1pi	<i>tip-i-ŋas-i-e</i>		<i>mi-tim ŋas-i-e</i>	we have not met
	meet-1pS/P-V2	:stay-1pS/P-NPT	NEG-meet V2:stay-1pS/P-NPT	

In the negative imperative the stem is negated by *mi-* and followed by the negative auxiliary *da-*:

2s *khata! mi-khai-da!* don’t go!  
 2s-1s *setuŋa! mi-sei-duŋa!* 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 desemantized to some degree, most often they telicize the main verb.

<i>ca-m-pak-u-m-ka</i>	we ate it up
eat-1/2pA-V2:put-3P-/2pA-e	
<i>mi-pera-khata</i>	they flew away
3pS-fly-V2:go	
<i>ta-mobdh-yu-kas-yu</i>	you spilt it
2-dump-3P-V2:throw-3P	

*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<sub>p</sub>-e 2sPOSS-fetch-PURP come-1/2pS/P-e-NPT  
 ‘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 (cf. Ebert, Chapter 3) – 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 SOV. 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 (cf. (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. *kanga khana(-lai) lo-na.*  
 I you(-DAT) tell-1→2  
 ‘I told you.’

- b. *khu-wa kaini pa-phlod-i.*  
 s/he-ERG we<sub>pi</sub> 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’, ‘strike’). 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<sub>pe</sub> [=itr. S]            go-1pS/P-e  
 b. (*kaika*)                    *phlod -u-m-ka*                we helped him  
           we<sub>pe</sub> [=Actor]           help-3P-1/2pA-e  
 c. (*kaika*)                    *pa-phlod-im-ka*                he /they helped us  
           we<sub>pe</sub> [=Patient]        INV-help-1ps/P-e

That participant marking on the verb is not in grammatical agreement and is evident from examples like:

- (6) a. *uko-lai phlakai mu-ma sopa-wa ta-cap-u-m-ne?*  
           this-DAT court do-INF who-ERG 2-can-3P-1/2pA -NPT  
           ‘Who of you can [go and] court him?’  
 b. *khim-da mi-hiη-e-ko-ci-lai rairewa-kukuwa kha-maid-e*  
           house-LOC 3pS-be-NPT-NOMZR-p-DAT (ritual) 1P-make -NPT  
           *nochun-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 *hiηma* ‘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 (cf. (9), (11a)) and are negated by *aina* ‘no’, ‘not’.

- (7) a. *kap-khim khoda hiη-e?*  
           2sPOSS-house where be-NPT  
           ‘Where is your house?’ i.e. ‘Where do you live?’  
 b. *kic-nicho-ci mi -hiη-e? – mihije./paina.*  
           3nsPOSS-y.sibling-ns 3pS-exist-NPT  
           ‘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 SOV 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 diŋ-u-pana 'lo, u-ko-wa diŋ-u!*  
 wood pecker-ERG drink-3P-TEMP INTJ PROX-NOMZR-ERG drink-3P  
*mi-riŋa-ki tyu-ko-pa-mo m-lem*  
 3pS-say-SEQ DIST-NOMZR-NOMZR-GEN 3sPOSS-tongue  
*pa-bhurda-ko raicha.*  
 INV-pull.out-NOMZR REP
- b. *m-lem pa-bhurda-ki kamalapa pa-maida-ko raicha.*  
 [repet.] -SEQ upside.down INV-make-NOMZR REP
- c. *kamalapa pamaida tyudaka Arko thaū mi-khata-ko-na tyu-da*  
 [repet.] then other place 3pS-go-NOMZR-TOP DIST-LOC  
*pAni hiŋa-ko wa-na chirkucipa-wa diŋ-u-ko*  
 FOC be-NOMZR water-TOP grasshopper-ERG drink-3P-NOMZR  
*raicha.[...]*  
 REP
- d. *tyudaka lam-sa lam-sa mi-khata-pana cikalemma*  
 then search-SIM search-SIM 3pS-go-TEMP (name)  
*ruŋma-ko wasa-wa 'tyuda hiŋ-e, tyuda hiŋ-e!'*  
 QUOTE-NOMZR bird-ERG there be-NPT [repet.]  
*lod-yi -ŋas-yi-c-yi-ko raicha-ki tyu-ko-ci-lai*  
 tell-3P-V2:keep-3P-3nsP-[copy]-NOMZR REP-SEQ DIST-NOMZR-ns-DAT  
*khō maid-yi-c-yi-ko...*  
 see make-3P-3nsP-[copy]-NOMZR

(During a drought the ancestor-god Nayima sends out animals to search for water: ...)

- When the woodpecker drank, they said: 'Look, he drank!' and they pulled out his tongue.
- They pulled out his tongue and turned it the wrong way round.
- Then, when they went to some other place, the grasshopper drank some water that was there. [...]
- 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 *ruŋma(pa) ~ ruŋmanA* (<*ruŋma* 'say'), which is optional before a speech-act verb, e.g. SPEECH (*ruŋmanA*) *lod-yu* 'he told him'. Thoughts are rendered in the form of direct speech, i.e. with the deictic elements unshifted.

- (9) *a-kuruṅpa ruṅma pa-chaid-aina, Saphopte-wa paṅi*  
 1sPOSS-mat.uncle quote NEG-know-NEG (name) also  
*a-cyodum-ci ruṅma 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 mihiṅe-ko-ci* those who live in the house (6b)

Nominalized sentences can constitute headless relative clauses, or they can be attributed as relative clauses; cf. *tyuda paṅi hiṅako 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-NPT prox-NOMZR 1sPOSS-rice/food-NO MZR  
 ‘I won’t give it to you; this is my food!’  
 b. *kho-ni ta-khat-e-ko?*  
 which-ALL 2-go-NPT-NOMZR  
 ‘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)

### REFERENCES

- Ebert, Karen H. (1997) *Camling*, Munich: Lincom.  
 — (1999) ‘The UP-DOWN dimension in Rai grammar and mythology’, in Balthasar Bickel and Martin Gänszle (eds) *Himalayan space: Cultural Horizons and Practices*, Zürich: Völkerkunde Museum, 107–34.  
 — (2000) *Camling Texts and Glossary*, Munich: Lincom.



# BELHARE

*Balthasar Bickel*

## 1 INTRODUCTION

Belhare is a Kiranti language spoken by about 2000 people living on the Belhara (written *Belahārā*) hill, one of the southern foothills of the Himalayas situated in Eastern Nepal (Dhankuṭā district; Kośī zone; 87° 18' E and 26° 57' N; c. 1150 m above sea-level). The Belhare are culturally closely related to the neighbouring Athpare community of Dhankuta. 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 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 homogenous; its closest relatives are Chilling and Athpare (on Kiranti subgrouping, see Chapter 31 (Ebert) this volume).

## 2 PHONOLOGY

Table 34.1 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 below). Segments in brackets occur in loanwords only, but the breathy voiced consonants also appear as regular allophones of aspirated voiceless stops between sonorants, e.g., /taŋkhek/: [ʔaŋ<sup>N</sup>gʰɛk-] ‘head’ or /laphe/: [ʔlabʰɛ] ‘caught’, and are represented as such in the practical orthography used here. Coronal is postalveolar; /t/ has a secondary glottal release before laterals and dissimilates to [ʔ] before palatal glides. Apart from this, there is not much allophonic variation beyond low-level assimilation of vowels to consonantal points of articulation.

**TABLE 34.1** 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		
ʔ	h	l	r	(rh)	s	y	w	N	

(NB: <c> stands for [tʃ] and <j> for [dʒ])

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.

## 2.1 Stress, syllable and word structure

The major syllable canon of Belhare is (C)V(C). Onsets are obligatory word-initially; if there is no underlying consonant, a glottal stop is prothesized in this environment (not written in the orthography), cf. e.g. /ʔu.uk.ma/ 'to roast and bring down'. Coda consonants are optional in all word positions. They are either nasals or unreleased versions of /p, t, k, ʔ/. There is in addition a minor syllable type consisting of the syllabic nasal that arises from phrase-initial *N*, e.g. /ŋ.ka/ 'I'.

Apart from a few lexical exceptions, main stress is initial in the prosodic 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. (<sub>ϕ</sub><sup>1</sup>phagi)(<sub>ϕ</sub>,det)(<sub>ϕ</sub>,lem) 'butterfly': [ˈphagɪ,det̚<sup>2</sup>,lɛm] ~ [ˈphagɪ,det̚<sup>2</sup>,lɛm].

The prosodic word starts at stem boundaries, e.g. ʔu(<sub>PrWd</sub>'hopchi) from *u-hop-chi* '3sgPOSS-calebash-nsg', unless this results in end stress, e.g. (<sub>PrWd</sub>'ʔuhop). The prosodic word often fails to align with the morphological word, mostly because of the multifunctional prefix *N-* '3nsgS/A; 3A; NEG; 2sgPOSS', which is realized as a nasality feature on preceding codas (but written as a distinct segment in the practical orthography): e.g. /un.c<sup>h</sup>iŋ k<sup>h</sup>are/ from *un-chik ŋ-khar-e* (3-nsg[ABS] 3nsg[S]-go-PAST) 'they went'. Also note that the morphological word 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'.

## 2.2 General morphophonology

The single most important morphophonological regularity is due to consonant prothesis to avoid morpheme boundaries within feet, e.g. (<sub>ϕ</sub><sup>1</sup>lap-)(<sub>ϕ</sub>,buk-)ma from *lap-uk-ma* 'to catch and bring down' instead of \*(<sub>ϕ</sub><sup>1</sup>la)(<sub>ϕ</sub>,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. (<sub>ϕ</sub><sup>1</sup>la)b-u-k=cha 'catch-3P-2A=ADD' instead of \*(<sub>ϕ</sub><sup>1</sup>lap)(<sub>ϕ</sub>,b-u-k)=cha or \*(<sub>ϕ</sub><sup>1</sup>la)(<sub>ϕ</sub>,b-u-k)=cha. However, lexical roots always receive stress, and as a result of this, some heteromorphemic feet are left unrepaired, cf. e.g. (<sub>ϕ</sub><sup>1</sup>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 affecting suffix boundaries: stops generally voice between sonorants if they are on either side of a suffix boundary, with a few exceptions of invariably unvoiced markers (most notably *-t* 'non-past'). Intervocalic /t/ not only becomes voiced but also continuant, cf. /kar<sup>h</sup>e/ from *kat-he* 'come.up-PAST' vs /k<sup>h</sup>ɛmdahe/ from *khem-ta-he* 'hear-come-PAST'.

There are two deletion rules, one affecting vowels, the other consonants. First, /a/ and /i/ regularly delete before /u/ (cf. e.g. /luisuha/ from *lui-sa-u-hak* 'tell-TR.PERF-3P-PERF') and /u/ deletes after /e/ (e.g. /lur<sup>h</sup>e/ from *lut-he-u* 'tell-PAST-3P'). Second, word-final /k/ is obligatorily deleted 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 affix boundaries, e.g. /unna/ from *un-ŋa* 's/he-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 fitted 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. *yaŋ-* ~ *yaĩ-* ‘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-* ~ *kaiʔ* ‘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. Grammatical morphemes are agglutinative 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’).

### 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 (a category I term ‘environmental space’; see 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 *ibaŋ* ‘one’, interrogatives (*sati = cha* ‘who = ADD’, i.e. ‘whoever’) or by zero (on the latter, see Bickel 1999d).

#### 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; the first person differentiates exclusive from inclusive forms (Table 34.2).

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*).

TABLE 34.2 PRONOUNS AND DEMONSTRATIVES

	<i>Singular</i>	<i>Plural</i>	<i>Dual</i>
1st inclusive	<i>ŋka</i>	<i>ŋke</i>	<i>ŋkechi</i>
1st exclusive		<i>ŋkeŋ</i>	<i>ŋkeŋchiŋ</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>

## 3.1.2 Case

The case forms are summarized in Table 34.3, with a brief description of their use and distribution.

The absolutive is unmarked except for the interrogative pronoun *sa-ti*, where the absolutive is marked by *-ti* in the singular (cf. *sa-a* ‘who-ERG’ and *sa-chi* ‘who-nsg [ABS]’). Occasionally, the Nepali dative *-lāī* is used on high-empathy objects, but not as commonly as in other Kiranti languages (cf. Ebert, Chapter 31 this volume). 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 *-ŋa* 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 *-ŋa* (e.g. *maʔi-a* ~ *maʔi-ŋa* ‘person-ERG’ or *sa-a* ~ *sa-ŋa* ‘who-ERG’); there is a slight preference for *-ŋa* in repetitions. Free allomorphy is also characteristic of the mediative (*-lam* vs *-lamma*) and the allative (*-sAm* vs *-samma*, from Nep. *samma*). The ablative in *-huŋ* is restricted to temporal and spatial concepts, e.g. *hamba-huŋ* ‘from today on’, while *-etnahuŋ* is distributionally free. The underlying C in the locative triggers gemination of a subsequent C but deletes in all other environments (cf. *mi-et=to* ‘fire-LOC=ID’ vs *mi-e* ‘fire-LOC’). The locative allomorph *-pak* is limited to environmental space demonstratives (e.g. *tu-ba* ‘up there’), some temporal roots (e.g. *hale-ba* ‘earlier’) and converbs (*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. *helleŋ*. The form *hena* ‘which’ derives from *he-* suffixed by the 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-leŋ*); dir. \**ina-leŋ* and med. \**ina-lam* contract to *illeŋ* and *illam*, respectively (also cf. abs. pl. *inkha* ~ *ikha*). Case combination is possible only with the environmental case locatives: *na-pmu-leŋ* (DEM-DOWN-DIR) ‘towards down there’. However, it is likely that the ablative allomorph in *-etnahuŋ* derives historically from the locative *-eC* suffixed by an article (= *na*) and the older ablative in *-huŋ*. In complex NPs, there is no case stacking (*Suffixaufnahme*) and neither is there case agreement: Belhare case-marking has phrasal scope.

**TABLE 34.3** CASE MARKERS AND THEIR USE

absolutive	<i>-∅</i> ~ <i>-ti</i>	S, P (i.e. primary objects), secondary objects, experiencers in experiencer-as-goal constructions, predicate nominals
erg./instr.	<i>-ŋa</i> ~ <i>-a</i>	A, instruments, causes, forces
genitive	<i>-ŋahak</i> ~ <i>-hak</i> (after V)	possessors, attributive nouns in NPs
comitative	<i>-lok</i>	accompanying referents or situations
ablative	<i>-etnahuŋ</i> ~ <i>-huŋ</i>	from, after
mediative	<i>-lam</i> ( <i>ma</i> )	via, through, from, in (language X)
allative	<i>-sAm</i> ( <i>ma</i> )	until, up to
directional	<i>-leŋ</i>	towards, in the direction of
locative	<i>-eC</i> ~ <i>-pak</i> ~ <i>-∅</i>	at, in, on, to
UP-locative	<i>-ttaŋ</i>	up at, in, on, to
DOWN-loc.	<i>-pmu</i>	down at, in, on, to
ACROSS-loc.	<i>-ʔyã</i>	across at, in, on, to

TABLE 34.4 POSSESSIVE PREFIXES

	<i>Singular</i>	<i>Plural</i>	<i>Dual</i>
1st inclusive	<i>a-</i>	<i>ŋke-</i>	<i>ŋkechi-</i>
1st exclusive		<i>ŋkeŋ-</i>	<i>ŋkeŋchiŋ-</i>
2nd	<i>N-</i>	<i>hani-</i>	<i>hanchi-</i>
3rd	<i>u-</i>		<i>unchi-</i>

### 3.1.3 Possession

Nouns can inflect for possessors (Table 34.4). In the non-singular, the relevant prefixes are closely related and sometimes identical with free pronouns (cf. Table 34.2). 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 disyllabic kin terms used for address). 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 34.6), the possibilities are severely limited, however, and only involve *-ŋa* ‘[1]ex’ and *-kak* ‘2’; third person is zero-marked:

- (1) a. *ŋka*            *n-tak-ŋa*  
 1sg[ABS]    2sgPOSS-friend-[1]ex  
 ‘I am your friend.’  
 b. *un*            *n-tak*  
 3[sgABS]    2sgPOSS-friend[-3sg]  
 ‘S/he is your friend.’

## 3.2 Adjectives

Belhare has a small set of adjectives which are marked by the nominalizer *-khak* (cf. Section 3.5) or by the enclitic article, which indicates specific reference of the head noun and agrees in number (= *na* ‘s’, = *kha* ‘p’, = *khachi* ‘d’). One small subset 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. *phabeleŋ=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 34.5 (from Bickel 1996, with revisions); Table 34.6 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 to the

TABLE 34.5 FINITE VERB INFLECTION

pf1	<i>maʔi-</i>	first person nonsingular 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 A or S	
	~ <i>mi-</i>	(in pf1 position) in negative forms with third person S or P	
	<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	simultaneously in 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 temporary
sf2	<i>-t</i>	'non-past indicative'	- <i>kone</i> inconsequential
	$\emptyset$	before C (but blocking intervocalic voicing of C)	- <i>ŋe</i> resultative Perfect
	<i>-ʔ</i>	before N	~ <i>-se</i> in transitives
	~ <i>-yuk</i>	in [ <sub>prWd</sub> σ <sub>1</sub> ]	- <i>ŋa</i> perfect (only in combination with sf10 filler <i>-khak</i> ~ <i>-hak</i> )
	<i>-he</i>	past indicative	~ <i>-sa</i> in transitives
	~ <i>-e</i>	after /s/ and in <i>C<sup>h</sup>Vr<sub>1</sub></i>	- <i>att</i> perfect negative
	<i>-a</i>	past subjunctive	
	<i>-att</i>	past negative	
	<i>-a</i>	imperative	
	~ <i>-an</i>	in plural affirm. 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</i> ~ <i>-sa</i> ) or 'nominalizer'	
	~ <i>hak</i>	after V or N	

left and right (e.g. neg. imp. 2du > 3sg:  $\Sigma$ -*n-a-n-ch-u-n* <  $\Sigma$ -*a-ch-u-n*, 2sg > 3nsg:  $\Sigma$ -*u-n-chi-n* <  $\Sigma$ -*u-n-chi*). In non-imperative forms, this is limited to copying around the sf8 filler *-chi* 'nsg' (e.g. neg. 3sg > 3nsg:  $\Sigma$ -*u-n-chi-n* <  $\Sigma$ -*u-n-chi*). The first person markers of sf7 copy around *-chi* 'nsg' in all forms.

TABLE 34.6 AFFIRMATIVE (UPPER FORM IN EACH CELL) AND NEGATIVE (LOWER FORM) NON-PAST PARADIGM (LUMA 'TO TELL' AND KHATMA 'TO GO')

	Primary object						Intransitive		
	1sg	1nsg.ex	1nsg.inc	2sg	2pl	2du		3sg	3nsg
1sg				lui ꞑha nloi ꞑhan	lui ꞑhani nloi ꞑhanin	lui ꞑhachi nloi ꞑhanchin	luituy nloi ꞑting	luituychit nloi ꞑtingchit	khai ꞑha ꞑkhai ꞑting
1pl.ex				lui ꞑhachina nloi ꞑhanchinna			luitumma nloi ꞑhimma	luitumchimma nloi ꞑhimchimma	khai ꞑiga ꞑkhai ꞑinna
1du.ex							luituuga nluitchunna		khai ꞑiga ꞑkhai ꞑinna
1pl.incl							luitum nloi ꞑhim	luitumchim nloi ꞑhimchim	khai ꞑiti ꞑkhai ꞑitin
1du.inc							luitu nluitchun		khai ꞑi ꞑkhai ꞑichin

		Primary object							Intransitive	
		1sg	1sg.ex	1sg.inc	2sg	2pl	2du	3sg	3nsg	
A	2sg	<i>kaluika</i> <i>kaluiʔhiga</i>	<i>maʔluika</i> <i>maʔluuʔhiga</i>					<i>luutuga</i> <i>nluitunga</i>	<i>luituchiga</i> <i>nluitunchinga</i>	<i>khaika</i> <i>ʔkhaiʔhiga</i>
	t	<i>kaluutiga</i> <i>kaluutiga</i>	<i>maʔluutiga</i> <i>maʔluutiga</i>					<i>luutunga</i> <i>nluiʔhiga</i>	<i>luitumchinga</i> <i>nluiʔnumchinga</i>	<i>khaiiga</i> <i>ʔkhaiiga</i>
	2du	<i>kaluichiga</i> <i>kaluichinga</i>	<i>maʔluichiga</i> <i>maʔluichinga</i>					<i>luuchuga</i> <i>nluuchuga</i>		<i>khaichiga</i> <i>ʔkhaichinga</i>
	3sg	<i>mailuyu</i> <i>mailuiʔhi</i>	<i>maʔluuyu</i> <i>maʔluuʔhi</i>	<i>kaluyu</i> <i>kaluiʔhi</i>	<i>nluiika</i> <i>nluiʔhiga</i>	<i>nluitiga</i> <i>nluitiga</i>	<i>nluihiga</i> <i>nluihinga</i>	<i>luutu</i> <i>nluitun</i>	<i>luituchi</i> <i>nluitunchin</i>	<i>khaʔyu</i> <i>ʔkhaiʔhi</i>
	3pl	<i>mapluyu</i> <i>mapluiʔhi</i>		<i>kapluyu</i> <i>kapluiʔhi</i>				<i>nluitu</i> <i>minluitun</i>	<i>nluituchi</i> <i>minluitunchin</i>	<i>ʔkhaʔyu</i> <i>miʔkhaiʔhi</i>
	3du	<i>mapluichi</i> <i>mapluichin</i>	<i>maʔluichi</i> <i>maʔluichin</i>	<i>kapluichi</i> <i>kapluichin</i>				<i>nluichu</i> <i>minluichun</i>		<i>ʔkhaichi</i> <i>miʔkhaichin</i>



### 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. /nluiʔniŋ/ rather than \*/nluituŋ/ from *lui-t-u-n(i)-ŋ* ‘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 *-ŋatt* 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 in *-sa* has a negative counterpart marked by *miN-Σ*. There is no negative supine, a paradigm gap which is compensated for by NEG-transport; cf. (37) in Section 5.6.

### 3.3.3 Tense and mood

The past vs non-past opposition also covers a modal 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]    2sgPOSS-child[-sgABS]    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*<sub>(PrWd)</sub> *lu-yu*, not \**mai*<sub>(PrWd)</sub> *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 modal 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]?’

- (3) b. *iŋa uŋ-ch-u!*  
 beer[ABS] drink-[1]du[incS]-3P  
 'Let's have some beer!'  
 c. *cuŋ-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 (40).

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 34.5) 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. *aŋ-khat-ni!*  
 OPT:NEG[-3sgS]-go-NEG  
 'She shouldn't go!' (i.e. 'I don't want her to go.')
- 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<sup>P</sup> 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 (cognate) Aktionsart modifier *kond-* (cf. Section 4.2) combined with the imperfective, e.g. *wat chi-goŋ-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-ke-t-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 allows specification of the time of reference, i.e. of the resulting situation (5b):

- (5) a. *namniŋŋ-etnahuŋ misen niu-s-u-ŋŋ-ha.*  
 last.year-ABL know-TR.PERF-3[sg]P-1sgA-PERF  
 ‘I have known him since last year.’  
 b. *asen tai-ŋe.*  
 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 money 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) *pī-sa uŋ-gone-ŋa 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* (<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*                      *khuiʔ-yu*.  
 REM-DEM[-sgABS]    chicken[-ABS]    [3sgS-]steal-NPAST  
 ‘This [guy] steals chicken.’
- b. (*i-na-ŋa*)                      *wa*                      *khuiʔ-t-u*.  
 REM-DEM[sg]-ERG    chicken[-sgABS]    [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 (relativization on A is impossible in this construction; cf. (26b) in Section 5.4):

- (9) *tombhira*    *wa*                      *seiʔ-sa-ha*                      *chitt-he-m*.  
 lynx[-sgABS]    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 suggests 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, forthcoming). 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 34.5 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 1999e). 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 maʔi-ŋa 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*, *-pilan*, *-pilak*, *-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[sgABS] tell-PASS.PART-NOMZR hey! this.much time  
*ŋŋ-hon-dai-ʔ-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. *khon-yak-ma* ‘to keep playing’ vs *khon-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-maʔ*  
 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) below.

The three other non-finite forms, the supine (*-si*) and the affirmative (*-sa*) and negative (*miN-*) converb 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*)*ʔwa* ‘like’) and evidentiality/status (e.g. = *phu* ~ = *bu* ‘hearsay’, ‘reportative’, = *phe* ~ = *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 on a noun in (43)), 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* ~ = *naŋa* (cf. Section 5.6; for the historical reasons of this, see Bickel 1999b).

Clause-final particles are the interrogative particle *i*, the ‘obvious’ marker *mu* (signalling clear and obvious evidence), the mirative marker *rAicha* (< Nep.) and also include the conjunctions *ki* ~ *kina* ~ *kinahuŋ*(*go*) ‘and’, *huŋcha* ‘whereas’, ‘but’, and *muhungo* ~ *hundo* ‘since’, ‘because’. Clause-final particles bear stress and are unaffected by the voicing and prothesis rules of suffixes and clitics, cf. *taheŋa ki* (not \**taheŋ = gi*) ‘I came and’ or *kharega i* (\**kharegak = gi*) ‘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* ‘3sgPOSS-mother’ for women. An example appeared in (7) above. 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. *maʔi khim* ‘a foreigner’s house’, *maŋ khim* ‘god-house’), *sa* ‘meat, flesh’ (e.g. *phak sa* ‘pork’, *ŋa sa* ‘fish’) or *taŋ* ‘tree’, ‘plant’ (e.g. *siŋ taŋ* ‘firewood tree’, but *ambiu u-taŋ* ‘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’, *phabeleŋ-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’, *aũbu* ‘three days ago’.

### 4.2 Verbs

Apart from one suffix (*-ap*), which serves to integrate loanwords (e.g. *har-ap lima* ‘to lose’, ‘be defeated’ from Nep. *hārāunu* ‘id.’ and the auxiliary *lima* ‘to be(come)’), the verbal lexicon

can be productively expanded only through compounding, which is often lexicalized or grammaticalized. (The reflexes of Proto-Tibeto-Burman transitivizers are unproductive, although there are many pairs like *pok-* ‘rise’ vs *phok-* ‘raise’ (< \**s-pok*), *ta-* ‘come’ vs *tat-* ‘bring’, *hon-d* ‘appear’ vs *hot-t* ‘drive out’, or *lik-* ‘enter’ vs *liŋs-* ‘insert’.)

Compounds often reflect stereotyped ways of performing activities such as *eʔwa 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. *chanḍahe* ‘it (the snake) came and curled up’ in (43) below. Compound verbs agree in transitivity with regard to the roots chosen:

- (14) *cama*            *ca-apt-he-i-ga*                            *iʔ* (\**ca-ab-he-i-ga*)  
 food[-ABS] eat-bring.ACROSS-PAST-pl-2 Q eat-come.ACROSS-PAST-2pl-2  
 ‘Did you come over here after you had your meal?’ (Nep. *khānā khāera āyauʔ*)

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 V 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. \**maimethe*, *hit*. The most important grammaticalized compounds include the following:

#### 4.2.1 Benefactives

The root *pir-* ‘give’ expresses benefactive or malefactive affection of 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-laŋ*    *tok-cand-he-ŋ*.  
 1sgPOSS-leg[-sgABS] hit.against-S/A.BEN -PAST[-3sgP]-1sA  
 ‘I hit my leg against something.’  
 b. *u-nucha*    *khon-ca-he*.  
 3sgPOSS-younger.sibling[-sgABS] [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 a morpheme in the agreement system, Belhare has apparently reanalysed the cognate of this marker 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* ~ *-nɛ*, etc.); e.g. present *ten-chiŋ-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 N-*lok* (‘COMIT’) *khima* ‘to be angry with N, to quarrel with N’. 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-PT[-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 *hanna* ‘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) with intransitives and *-tend-* with transitives (but note exceptional *khat-lott-* ‘take-TEL-’ and *khat-ca-* ‘go-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-*; cf. Section 3.3), relinquitive (*-ye-* ‘to leave P where the action takes place’), perseverative (*-yuh-* ‘keep doing’), accelerative (*-itt-* ~ *-is-* ‘quickly’, ‘just’), diminutive (*-khet-* ~ *-khat-* ‘only’, ‘just’, ‘a little’, ‘quickly’), errative (*-khet-* ~ *-khat-* ‘mistakenly’), restrictive (*-tis-* ‘do nothing else but V’), 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 a rhematic increase from left to right, ending with the predicate. The pre-predicate position often attracts question words and focus-marked constituents. There is an intonationally marked post-clausal afterthought position (illustrated by (24b) and (43) below), which is highly thematic and does not accommodate, for instance, true question words (Bickel, forthcoming). NPs are all optional (and often absent) in the clause.

The predicate usually consists of a verb (or verb complex). Nominal predicates are either marked by agreement morphology (Section 3.1), a copula (*yuhha-* in the non-past, *yuhhase-* 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 rhematic enough, pronouns can be co-referent with clause-mate antecedents that bear a lower grammatical relation. Rhematicity can be increased by word order and/or focus particles:

- (17) *Mohan-naha<sub>i</sub> u-kitab un-na = ro<sub>i,j</sub> chitt-he.*  
 M.-GEN 3sgPOSS-book[ABS] 3[sg]-ERG = ID [3sgA-]find-PAST[3sgP]  
 ‘HE<sub>\*i,j</sub> found Mohan<sub>i</sub>’s book.’

Apart from the theme–rheme 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 clauses:

- (18) *sa-a*                    *thuu-s-u=na*                    *iŋa*  
 who[sg]-ERG    [3sgA-]cook-TR.PERF-3P=ART    beer[ABS]  
*chept-he-ga?*  
 taste-PAST[-3P]-2[sgA]  
 ‘\*Who did you try the beer that\_\_made?’

Instead of *saa* ‘who’, a focused NP like *a-na-ŋa=ro* ‘my-elder.sister-ERG=ID’ would be equally grammatical here.

## 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:

- (19) a. [<sub>NP</sub> *sik-kira*    *phabeleŋ=ma*    *phuŋ*]  
 two-NHUM    red=COLOUR.ART    flower[ABS]  
*tar-he-ŋ.*  
 bring-PAST[-3P]-1sgA  
 b. [<sub>NP</sub> *phabeleŋ=ma*    *sik-kira*    *phuŋ*]  
 red=COLOUR.ART    two-NHUM    flower[ABS]  
*tar-he-ŋ.*  
 bring-PAST[-3P]-1sgA  
 c. [<sub>NP</sub> [<sub>NP</sub> *sik-kira*    *phuŋ*]            *phabeleŋ=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 are embedded 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:

- (20) a. *maʔi-ha*                    *khim*                    (cf. *\*na*    *maʔi-ha*                    *khim*)  
 person[sg]-GEN    house[ABS]                    dem    person[sg]-GEN    house[ABS]  
 ‘someone (foreign)’s house’  
 b. *maʔi-ha*                    *u-khim*  
 person[sg]-GEN    3sgPOSS-house[ABS]  
 ‘the (known) person’s house’  
 c. *\*maʔi*                    *u-khim*  
 person[sg.ABS]    3sgPOSS-house[ABS]

Articles and the nominalizer are complementary means of attribution (and other Kiranti languages merge them formally; cf. Bickel 1999e): the use of articles indicates specific reference of the head noun, while the nominalizer is unmarked in this respect. Since possessed nouns have inherently specific reference, only article-marked adjectives can be used:

- (21) *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 (cf. (18) above), 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, *uchoūat* 'new', is attributed without further marking (*uchoūat khim* or *khim uchoūat* '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. *ηatlabu u-soη* [banana 3POSS-ripe] 'the/a ripe banana'.

Classifiers are obligatory with numerals and draw a basic distinction between humans (*-paη*) 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:

- (22) *sip-paη maʔi-chi / \*maʔi*  
 two-HUM person-nsg [ABS] person[sgABS]  
 'two people'

With other nouns, number agreement is optional; cf. the examples in (19) above.

The NP head is always optional, whence demonstratives and attributes can function as NPs of 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 (23a) but also allows appositional (23b) and partitional (23c) interpretations of the relationship between the agreement trigger and its target:

- (23) a. *kaepma-chi η-khar-e.*  
 girl-nsg[ABS] 3 nsgS-go-PAST  
 'The girls went.' (*η-* 'they' = *kaepmachi* 'girls')
- b. *masiη=cha siη taηη-e thaū-ʔ-ηa.*  
 old.woman=ADD fire.wood tree-LOC go.UP-NPAST-[1sg]ex  
 'Even as an old woman I climb trees.' (*-ηa* 'I' as *masiη* 'old woman')
- c. *sa-ti khar-e-i-gaʔ*  
 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; as such they support extraction (24), in contrast to non-arguments, which do not (25):

- (24) a. [*a-tak-ηaha u-phu-ηa*]  
 1sgPOSS-friend[sg]-GEN 3sgPOSS-elder.brother [sg]-ERG  
*mai-lur-he.*  
 1sgP-[3sgA-]tell-PAST



- (28) a. *khali* [ $\phi_{*i,j,k}$   $\phi_{i,*j,*k}$  *set-ma*] *ka<sub>i</sub>- $\eta_j$ -ko $\bar{i}$ -yu*.  
 only [ERG]<sub>A</sub> [ABS]<sub>P</sub> kill-INF 1incP-3nsgA-want-NPAST  
 ‘They just want us to get drunk.’ (*lit.* ‘They want us to get killed by [the beer].’)
- b. [ $\phi_{*i,j,k}$   $\phi_{i,*j,*k}$  *ten-ma*] *ma<sub>i</sub>- $\eta_j$ -na-rend-he*.  
 [ERG]<sub>A</sub> [ABS]<sub>P</sub> beat-INF 1sgP-3nsgA-stop-TEL-PAST  
 ‘They stopped beating me.’ or ‘They stopped x from beating me.’

#### 5.4.2 Objects

Belhare is a primary object language (in Dryer’s 1986 terminology). Thus, in ditransitives it is the goal rather than the theme argument that triggers P-agreement, cf, e.g. (32) below. The same pattern is found in the pivot of non-finite complement constructions. Unlike ditransitive goals, themes cannot be controlled by the matrix inflection:

- (29) [ $\phi_{i,*j}$  *kitap-chi<sub>\*i,j</sub>* *pi-ma*]  *$\eta_j$ -khe-yu*.  
 [ABS]<sub>Primary Object</sub> book-nsg[ABS]<sub>Secondary Object</sub> give-INF 3nsgS-must-NPAST  
 ‘They (\*s/he) must be given books.’

Primary and secondary objects are not differentiated, however, by case-marking, where all objects are in the absolutive, or by the pivot of internal-head relatives, where all objects are part of the pivot.

### 5.5 Experiencer constructions

Belhare features both experiencer-as-goal constructions, corresponding to South Asian-style dative subjects, and experiencer-as-possessor constructions, corresponding to South-East Asian-style psychocollocations (Matisoff 1986):

- (30) a.  *$\eta$ ka* *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-ni $\bar{u}$ a* *ta-he*.  
 1sgPOSS-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 (cf. Bickel 1999c, forthcoming), e.g.:

- (31) a. *caleppa* *ka-khik-pa*  
 bread[ABS]<sub>P</sub> ACTIVE.PART-taste.bitter-PART  
 ‘one to whom the bread tastes bitter’
- b. [*( $\eta$ ka<sub>\*i,j</sub>)* *caleppa<sub>i,\*j</sub>* *khik-ma*]  *$\eta$ a-rend-he*  
 1sg[ABS]<sub>A</sub> bread[ABS]<sub>P</sub> taste.bitter-INF [3sgS-]stop-TEL-PAST  
 / \**mai<sub>i</sub>-na-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 in *-khak* ~ *-hak* and, in the past tense, both are limited to subjunctive mood. Compare the following relative construction with a complement clause as in (34) below:

- (32) *dahi ka-khuiʔ-sa-ha maʔi*  
 yoghurt[-sgABS] 1incP-[3sgA-]carry-TR.PERF-NOMZR person[-sgABS]  
*he-lleŋ khar-eʔ*  
 wh-DIR [3sgS-]go-PAST  
 ‘Where did the guy go who brought us the yoghurt?’

Relative constructions are prenominal, as in (32), or, less frequently, internally headed, as in (9) above. Instead of the nominalizer, relative clauses (but not complement clauses) can also be marked by the article; cf. (18) for a prenominal and (26) for internally-headed examples). Participial constructions, as in (11) and (31a), 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.’:

- (33) *‘hena bela tai-s-u-k-kha?’ (ceksa) kisi mai-lapt-he.*  
 which time reach-TR.PERF-3P-2 [sgA]-PERF say-CVB 1sgP-[3sgA-]ask-PAST  
 ‘S/he asked me when I had arrived.’

Indirect quotation involves complementation by means of *-khak* ~ *-hak* and triggers deictic shift:

- (34) *tas khons-a-ŋŋi-ha \*i,j maʔ-yakt-he.*  
 card play-SUBJ.PAST-[1sg]ex-NOMZR [3sgS-]narrate-IMPV-PAST  
 ‘S/he was saying that I played cards.’ (not: “‘I, played cards”, s/he<sub>i</sub> 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 (35a) with the referentially uncontrolled construction in (35b):

- (35) a. [ $\phi_i$   $\phi_j$  *lu-ma*  $i, *i, k$  *koĩ-t-u*  $*i, j, *k$ .  
 [ERG]<sub>A</sub> [ABS]<sub>P</sub> tell-INF [3sgA-]want- NPAST-3[sg]P  
 ‘S/he wants x to tell him/her.’  
 b. [ $\phi_i$   $\phi_j$  *lu-ma*  $i, j, k$  *koĩ-yu.*  
 [ERG]<sub>A</sub> [ABS]<sub>P</sub> 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. *koĩtu*, *luma* as an alternative to (35a). 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 converbs, supines, or comitative-marked verbs are all embedded within the relational periphery of the matrix clause:

- (36) *a-tak-ŋa* [ne-e *yuŋ-sa*] *sappaɪ*  
 1sgPOSS-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.:

- (37) *wa-si* *ŋ-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. (36) above). In repeated form, they are also used for alternating activities, often supported by a generic verb in auxiliary function (e.g. *metma* ‘to do’):

- (38) *kop-sa* *yuŋ-sa* *kop-sa* *yuŋ-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 (37)) 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, e.g. *khar-a-ŋ-lo* (go-SUBJ.PAST-[1sg]ex) ‘when I went’. 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 enclitic topic marker =*naa* ~ =*naŋa* or by one of the phonologically independent, clause-final conjunctions *muhungo* ~ *hudo* ‘since’ and *huŋcha* ‘whereas’, ‘although’, ‘but’, which are all based on the ablative in *-huŋ*. 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* ~ =*naŋa* 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*:

- (39) *ŋka=na ten-he-ŋ=cha pheŋd-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*, *kinahuŋ*, 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:

- (40) *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 chained clauses can be postponed:

- (41) *ca-ma=na ca-yau-t-u tARA he-lleŋ*  
 eat-INF=TOP [3sgA-]eat -IMPFV-NPAST-3[sg] P but where-DIR  
*leŋ kina?*  
 [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:

- (42) *khar-e ki jutta ŋŋ-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:

- (43) *chokt-he ki, n-celi-ŋa,*  
 [3sgA-]point.with.finger-PAST[3sgP] SEQ 2sgPOSS-unmarried.agnatic.girl-ERG  
*doko-ep=phu chaŋ-da-he.*  
 basket-LOC=HS [3sgS-]curl.up-come-PAST  
 ‘When your *celi* pointed with her finger [at him<sub>i</sub>], [Rainbow<sub>i</sub>] came and curled up in [her] basket.’

Independent sentences can be linked by demonstratives in the ablative (*inetnahuŋ* ‘after that’) or the Nepali loans *ani* ‘and then’ and *pheri* ‘again, and then.’ For contrastive linkage, Nep. *tARA* ‘but’ is often used; the native strategy involves either the inconsequential (Section 3.3) or subordination by *huŋcha* ‘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).

## REFERENCES

- Bickel, B. (1993) 'Belhare subordination and the theory of topic', in K.H. Ebert (ed.) *Studies in Clause Linkage*, 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 béharien', *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. Nuyts and E. Pederson (eds) *Language and Conceptualization*, Cambridge: Cambridge University Press.
- (1998) 'Rhythm and feet in Belhare morphology', Working Paper No. 287, Rutgers Optimality Archive, <http://www.rucss.rutgers.edu/roa.html>.
- (1999a) 'Cultural formalism and spatial language in Belhara', in B. Bickel and M. Gaenszle (eds) *Himalayan Space: Cultural Horizons and Practices*, Zürich: Museum of Ethnography Press.
- (1999b) 'From *ergativus absolutus* to topic marking in Kiranti: a typological perspective', *BLS* 25.
- (1999c) 'Grammatical relations, agreement, and genetic stability', University of California Berkeley. (<http://socrates.berkeley.edu/~bickel/papers>.)
- (1999d) 'How important are referents? Syntactic typology and cognitive effects', paper presented at the Conference on Event Conceptualization in Language and Cognition, Max Planck Institute for Psycholinguistics, Nijmegen, December 17–19, 1999 (Ms., <http://socrates.berkeley.edu/~bickel/papers>).
- (1999e) 'Nominalization and focus constructions in some Kiranti languages', in Y.P. Yadava and W.W. Glover (eds) *Topics in Nepalese Linguistics*, Kathmandu: Royal Nepal Academy.
- (2000a) 'Deictic transposition and referential practice in Belhare', *Journal of Linguistic Anthropology* 10: 224–47.
- (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*, Osaka: National Museum of Ethnology.
- (forthcoming) 'Hidden syntax in Belhare', in A. Saxena (ed.) *Himalayan Linguistics*, Berlin: Mouton de Gruyter.



Dryer, M. (1986) 'Primary objects, secondary objects, and antitativity', *Language* 62: 808–45.

Ebert, K.H. (1986) 'Reported speech in some languages of Nepal', in F. Coulmas (ed.) *Direct and Indirect Speech*, 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.

Van Valin, R.D. and LaPolla, R.J. (1997) *Syntax: Structure, Meaning, and Function*, Cambridge: Cambridge University Press.

### **Further reading**

Phonology and especially the prosody-morphology interaction is treated in Bickel (1998). 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 (forthcoming, 1999c) and for clause linkage, Bickel (1993, 1995b, 1999b, 1999e). Spatial deixis is discussed in detail in Bickel (1997, 2000a). Ethnographic information is given in Bickel (1999a, 2000c). A dictionary is available in electronic form from the author (bickel@rz.uni-leipzig.de).

**PART 11**

---

**QIANGIC LANGUAGES**

---



# QIANG\*

*Randy J. LaPolla*

Qiang is spoken in Aba Tibetan and Qiang Autonomous Prefecture in northwest Sichuan Province, China; it belongs to the Qiangic branch of Tibeto-Burman. 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 ) (Sun 1981a: 177–78). The dialect presented here is the Northern Qiang variety spoken in Ronghong Village, Yadu Township, Chibusu District, Mao County.

## 1 THE PHONOLOGICAL SYSTEM

Qiang has thirty-nine consonants at seven points of articulation (Table 35.1), 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 /ʑ/ and /x/, 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, can be finals. All of the original Proto-Tibeto-Burman finals were lost (cf. 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 35.1** THE QIANG CONSONANTS

	<i>Labial</i>	<i>Dental</i>	<i>Retroflex</i>	<i>Palatal</i>	<i>Velar</i>	<i>Uvular</i>	<i>Glottal</i>
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	dʂ	dʑ			
Voiceless fricative	ɸ(f)	s	ʂ	ɕ	x	χ	h
Voiced fricative	(v)	z	ʐ	(ʑ)	(ɣ)	ʁ	ɦ
Nasal	m	n		ɲ	ŋ		
Voiceless lateral		ɬ					
Voiced lateral		l					
Approximant	w			j			

\* I would like to thank R.M.W. Dixon and Alexandra Aikhenvald for valuable comments on a draft of this chapter.

Phonemically, consonant clusters are formed by /s/ and one of the following initials: /p, t, k, q, tɕ, b, d, g, m, dz/, /x/ followed by /k, s, tɕ, ʂ, tʂ, ʧ, l, z, dz, z, dz/, or /χ/ followed by /q, s, ʂ, tʂ, ʧ, l, d, z, n, dz, ŋ, z, dz/. Phonetically /s/ becomes [s] before /t/ and /d/, and becomes [ɕ] before /pi/, /pe/, /bi/, /tɕ/ and /dz/, and the pre-initials all become voiced before voiced initials (e.g. /mi:-xkam/ ‘eyebrow’; /χsu/ ‘living’, ‘to be alive’; /sta/ ‘entrust to’; /zbu/ ‘drum’; /zdu/ ‘deer’; /zdzi/ ‘disease’; /ʂdzəs/ ‘toenail’). Some examples of clusters in final position: /tʂhexʧ/ ‘sip (vt.)’; /dzɑtɕ/ ‘laugh (v)’; /wəχs/ ‘horse dung’; /laxs/ ‘palm’; /əxts/ ‘shade (vt)’.

The Qiang vowels are given in (1):

(1)	i, i: y, y:		u, u:
	e, e:	e, e:	o, o:
		ə	
	a, a:	a, a:	ɑ, ɑ:

There are fourteen 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 future tense marker /-ɑ:/ is added to a root such as /tɕye/ ‘carry’: [tɕya] ‘I carry’, [tɕya:] ‘will 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 /-ʃ/). The r-colouring participates in the vowel harmony scheme (see below) and so is treated as a vowel feature rather than a consonant.

The syllable canon is given in (2):

(2)	(C)	(C <sub>i</sub> )	(V)	V	(V)	(C)	(C <sub>f</sub> )
	[fric]		[glide]		[glide]	[fric]	

The minimum syllable is a single vowel, e.g. /a/ ‘one’, the maximum is CCVVCC, e.g. /çpiexʧ/ ‘scar’. Any of the consonants listed in Table 35.1 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. /p/ > [ɸ/f]: DIR + /phə/ ‘blow’ > [əɸ] ‘blow (imperative)’; /kh/ > [x]: DIR + /khə<sup>t</sup>te/ ‘hit (people)’ > [nəxte] ‘hit (past)’; /dz/ > [l]: /ma/ NEG + /dzə/ ‘able’ > [ma-lə] ‘not able’; /dz/ > [z]: DIR + /dzuə/ ‘sit’ > [əzūə] ‘sit (imperative)’; /k/ > [β]: DIR + /kə/ ‘go’ > [dəβ] ‘go out’; /b/ > [w]: DIR + /bə/ ‘pile’ > [təw] ‘piled’. Comparing Ronghong and Mawo dialect (Sun 1981a) forms, we can see that a similar type of weakening has occurred historically to pre-initial consonants in Ronghong (e.g. Ronghong /xsə/, Mawo /khsi/ ‘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 + /tɕə/ ‘drink’ > [sətɕ] ‘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’ + /ʂpu/ ‘flock’ > [wuʂpu] ‘(wild) pigeon’; /ha/ ‘ten’ + /tʂil/ ‘one’ > [hatʂil] ‘eleven’). 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<sup>ʃ</sup>/ ‘reduce’ > [me<sup>ʃ</sup>-we<sup>ʃ</sup>] ‘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. [ʒdʒi-tʃɒp-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. *phis* ~ *phiχ* ‘white’; *ɲiq* ~ *ɲix* ‘black’; *muxũ* ~ *muɸũ* ‘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. Ex. (4) contains two NPs (bracketed).

- (4) [*the*:tɕə-tɕiɸua ba-the-zɪ] *piɲa-la* [zawa ba-xɔə-zɪ] ʃə  
 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. /ɸua/ ‘help’ + /-m/ > /ɸuam/ ‘servant’. Nouns can also be formed from adjectives by simply adding one of the definite markers, e.g. /ɲiq-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-ɸuatʃa/ [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 Liu 1981 on Tibetan loans). Verbs borrowed into Qiang are treated as nouns, and take the verbalizing suffix /-tha/, if monosyllabic, or the verb /-pə/ ‘to do’, if polysyllabic, e.g. /tuen-tha/ ‘squat’ < dūn (蹲); /ɕunɲian-pə/ ‘train’ < xùnliàn (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* (藍)).

In Qiang only the natural gender of animals is marked: /-mi/ or /miaha/ for females; /zdu/, /zə/, /xə/, /ɕi/, and /pi/ for males. The diminutive has the concrete sense of ‘child’; it is not used for hypocoristic or other abstract uses: /-tʃuə ~ tʃũ/ (< /tʃuə/ ‘child’ (general)); /-ɸl/ (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/ (> /ɣle/ 'few'), used only on words referring to people, e.g. /ipi-le/ 'uncles'. Following a proper name, it 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 35.2.

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 [tsəhə]) and distal /the/ (plural [thəhə]). 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>ɲawu / ɲawe</i>	how much/many?
	<i>tʃa-la ~ tʃa:</i>	where?	<i>ɲi:ke</i>	how?
	<i>ɲiɣi</i>	what? / whatever	<i>tʃho:</i>	when?
	<i>ɲiɣi-χuaɲi</i>	why?		

The numeral system is a simple decimal system, with 'one' to 'ten' being unique forms, 'eleven' to 'nineteen' being 'ten' + 'one' etc., 'twenty' to 'ninety' being 'two' + 'ten', etc., and the numerals in between being 'two' + 'ten' + 'one', etc. There are no ordinal numbers in Qiang; /tʃi-qə<sup>1</sup>-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.

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' < /ʃqu/ 'mouth', /sa/ 'classifier for sections' < /saq/ '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; /xʃe/ for one item of a pair.

TABLE 35.2 THE QIANG PERSONAL PRONOUNS

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>	<i>Reflexive (sg/pl)</i>
1	<i>qa</i>	<i>tʃizzi</i> [tʃi-zi]	<i>tʃi-le</i>	<i>qa-qəi/tʃil-tʃile</i>
2	<i>ʔü</i>	<i>ʔizzi</i> [ʔi-zi]	<i>ʔi-le</i>	<i>ʔi-ʔi:/il-ile</i>
3	<i>the:/qupu</i>	<i>thizzi</i> [the-zi]	<i>them-le</i>	<i>(the:) ɲiɲi/ɲil-ɲile</i>

### 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, perlativ (‘through’, ‘along’) marker	<i>wu</i>
Genitive, recipient marker	<i>tʃə̌</i>
Locative, allative marker	<i>la</i>
Locative, allative, temporal, goal marker	<i>ta</i>
Locative, temporal marker	<i>ʁa</i>
Comitative, conjunction marker	<i>ŋa</i>
Comparative marker	<i>sə̌/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:-tʃə̌ pi:-xsə-lə 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. *moʁu-wu qa da-tuə-zə̌-ʃa.*  
 wind-AGT 1sg DIR-fall.over-CAUS-1sgU<sup>1</sup>  
 ‘The wind knocked me over.’

An instrumental NP is marked by the postposition /wu/, the same form as the agentive and ablative markers.

- (8) a. *qa ʁduəʃ-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 /-tʃə̌/, e.g. /tʃiautʃaŋ-tʃə̌ puŋu/ [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

---

1 U is the gloss for a bound non-actor marker, e.g. ‘1sgU’ means ‘1sg non-actor’.



does not have agentive marking the locative marker /-ta/ can be used after the NP representing the undergoer.

- (9) a. *the: qa-ta dze!*  
 3sg 1sg-LOC hit  
 'He is hitting me!'

Qiang formally distinguishes between goal, the referent at which an action is directed, and recipient, and 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-zə-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 /-tɕə̃/, the genitive postposition:

- (11) *sum the:-tɕə̃/ pi:-xsə-lə 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 /-tɕə̃/. 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) *mə: qa-tɕə̃ fa spetɕ.*  
 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 /χuaŋi/ 'because', 'in order to', 'in place of', 'for':

- (13) *the: qa-χuaŋi sə ge-kə̃.*  
 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 /-ŋa/ 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 /-ŋa/. 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. *mutsitʂu-ŋa-tugantʂu zəpəq-ta fiə-lua-ji.*  
 Mutsitʂu-COM-Tugantʂu earth-LOC DIR-come-CSM  
 'Mutsitʂu and Tugantʂu came to earth.'  
 b. *qa khumtsi-ŋa tiantsə-βa ka.*  
 1sg Khumtsi-COM store-LOC go:1sg:FUT  
 'I will go to the store with Khumtsi.'

The particle /*ŋa*/ is also used for the arguments of certain verbs, such as in (15):

- (15) a. *ʔiʔi-ŋa tɕe-khue.*  
 2sgREFL-COM NEG.IMP-upset  
 ‘Don’t be angry at yourself.’

In comparative clauses, the positive comparative marker is /*sâ*/, while the negative comparative is /*ŋiki*/.

- (16) a. *qa the:-sâ tɕe-fia.*  
 1sg 3sg-COMPAR still-white:1sg  
 ‘I am lighter (in colour) than him.’  
 b. *qa ʔi-ŋiki 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. /*petçin-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. /*ɕdzyta-la ~ ɕdzyta:*/ ‘in/to Chengdu’.

The locative /*ba*/ generally marks containment in some sort of vessel or movement in or out of one, e.g. /*pankoŋsə-ba*/ [office-LOC] ‘in the office’.

The ablative postposition is the same form as the agentive and instrumental postposition, /*wu*/. It can be used alone (/*petçin-wu*/ ‘from Beijing’), or with one of the other locative postpositions introduced above (/*məq-ta-wu la*/ [above-LOC-ABL fly] ‘fly from above’). 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, but if the situation is simply one of temporary possession and not ownership, then the possessor takes the locative/dative marker /*ta*/. If the situation involves ownership of an object or relationship (brother, sister, etc.), then 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.

- (17) a. *khumtsi tutɕə-ɣzə-zɨ zɨ-zə.*  
 Khumtsi younger.brother-four-CL have/exist-CAUS  
 ‘Khumtsi has four younger brothers.’  
 b. *qa dzoqu-ji-tua wa.*  
 1sg leg-two-CL have/exist:1sg  
 ‘I have two legs.’  
 c. *ʔi-dzobu-le qa-ta sə.*  
 2sg-key-DEF 1sg-LOC have/exist  
 ‘I have your key.’

#### 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, future tense 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<sup>2</sup>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 /-zə̌ə/, 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.

- (18) a. *qa tsə tu-χsu-zə̌ə-ja.*  
 1sg water DIR-boil-CAUS-CSM:1sg  
 'I brought the water to a boil.'
- b. *qa the:-wu pəitsə-e-ze zə-pə̌ə-za.*  
 1sg 3sg-AGT cup-one-CL DIR-buy-CAUS:1sg  
 'I made him buy a cup.'
- c. *qa χumtsi-wu ləyz-te-pen khumtsi-ta tə-χua-za.*  
 1sg Xumtsi-AGT book-DEF-CL Khumtsi-LOC DIR-buy-CAUS:1sg:FUT  
 '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. 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>fiə-phe</i> | tear (of clothes)                    |
| <i>da-be</i>  | <i>fiə-qhe</i> | break (of bowls, etc.)               |
| <i>da-bli</i> | <i>fiə-χfi</i> | break (in two) (of tree limbs, etc.) |

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 many 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; *patšpatš* 'some round things'), Au:A (intensification; *patšú:patš* 'very round'), AAu: (plurality plus intensification; *patšpatšú:* 'some very round things'; /u:/ is a stressed syllable added to the reduplicated form).

There are four main existential/locative verbs: /sə/, 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; /zi/, for animate referents; and /we/, for possession of qualities and for immovable referents or referents inalienably connected to a larger entity.

**TABLE 35.3 THE QIANG PERSON MARKING SUFFIXES FOR INTRANSITIVE VERBS**

	1	2	3
Singular	-a	-n	-θ
Plural	- <sup>1</sup>	-i	-tçi

Following are some auxiliary verbs and their meanings/uses: /γzə/ learnt ability; /dzə/ or /qe/ natural (physical) ability; /gu/ ability to fit into something else; /bu/ willingness to perform an action or to allow others to perform an action, or in some cases the possibility of some situation; /xşu/ ‘to dare’; /kze/ ‘ought to’; /se/ permission or lack of it; /çtçəq-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 and the single direct argument of an intransitive clause.

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 35.3 gives the forms of the suffixes.

Another set of suffixes can be used for marking a non-actor human referent.<sup>2</sup> These forms are given in Table 35.4:

**TABLE 35.4 THE QIANG NON-ACTOR PERSON MARKING SUFFIXES**

	1	2	3
Singular	-şa	-san	-wə~-u
Plural	-şa <sup>1</sup>	-sai	-wə~-u

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/ ‘look at’ only takes one prefix. The form of the prefix follows the rules of vowel harmony. Following are the prefixed forms of the verb /lə/, which also means ‘look at’:

- (20) *təl* look upwards      *fiəl* look downwards
- zəl* look towards centre      *dəl* 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’; /tçhə/ ‘eat’ (activity), [sətçə] ‘eat’

<sup>2</sup> This distinction of actor vs non-actor parallels the use of agentive and non-agentive marking on NPS (see LaPolla 1992a, 1995), 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).

(accomplishment); /ɣlu/ 'roll' (activity), [doɣlu] 'roll' (achievement). 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 of the argument of some intransitive verbs, or can be part of the causativization or transitivization of some verbs, e.g. [da-ʔa] 'slip', [fi a-ʔa] 'slide'; [do-tshu] 'fall (e.g. of fruit from tree)', [fi o-tshu] 'pick (fruit from tree)' (see also (19); cf. Huang 1997: 73).

Only future tense is overtly marked in Qiang. It is marked by the suffix /-ɑ:/, which replaces the root vowel of the verb if it is /ə/ or /a/, as in (21), where /məpa/ becomes [m̄əpa:] in the second clause.

- (21) pəs məpa wa, təp-ŋi tsə-s tɕa-m̄əpa: lu.  
 today cold very tomorrow this-COMPAR still-cold:FUT will  
 'Today is very cold, and tomorrow is going to be even colder than this.'

There are several suffixes for marking the different types of aspect. The most common aspect marking is /-ji/, which marks a recent change of state or situation.

- (22) a. meʔ: ɕi-ji. b. meʔ: de-ɕi-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 future tense marking to express the idea 'about to V'.

- (23) meʔ: ɕa:-ji.  
 rain fall:FUT-CSM  
 'It's about to rain.'

The form /-jy/ marks an action that has 'already' begun or been carried out:

- (24) qa a-tian u-z̄ulu-jya.  
 1sg one-hour DIR-wait-ASP:1sg  
 'I have already waited for one hour.'

The prefix [tɕe- ~ tɕi- ~ tɕa- ~ tɕo-] '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 future marking) there usually is no problem of ambiguity.

- (25) meʔ: tɕe-ɕi.  
 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 and the optional polite imperative particle /-na/. The prefix, which may be any one of the directionals, is stressed, unlike non-imperative directional prefixes.<sup>3</sup> In an imperative clause the person-marking is optional, though the imperative sense is stronger if person marking is used (ə-zə-nə! [DIR-eat-IMP] 'Eat!' vs ə-zə-n-nə [DIR-eat-2sg-IMP] 'You eat!').

3 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-tɕ '(S/he drank'. vs ə-tɕ 'Drink!'

The prohibitive is expressed by the prefix /tʃa-/ ([tʃa ~ tʃe ~ tʃə ~ tʃo]) (<PTB \*ta-), which appears in the same position as the negative prefix. For example: *ha-tʃə-βə!* [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 /-ɲua/ (often pronounced [wa]) for all other persons or numbers.

- (26) a. *ʔü zme ɲuə-n-a?*      b. *the: zme ɲuə-ɲua?*  
 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) *ʔü zdzyta: fiə-qə-n-a      fiə-mə-qə-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 /βze/ ‘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: tʃəu-la      lu-s      ɲuə-ɲüə. (< 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 above).

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: zdzyta fiə-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 /tʃa/ ([tʃa ~ tʃi ~ tʃə ~ tʃo]) (also used to mean ‘still’, ‘yet’) is generally used in the comparative construction (*tʃə-wa* [still-big] ‘relatively big’). The superlative of adjectives and some stative verbs is marked by the prefix /tʃi/: *tʃi-wa-la-lə*

[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-mə-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-tɕhə̌*      b. *akha-ŋi ə-tɕhə̌*  
           slow            DIR-eat            slow-ADV    DIR-eat  
           ‘eat slowly’                    ‘eat slowly’

Some adjectives, when acting as manner adverbs, take /-ji/ or /tɕi/ 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 /-tɕ/ is used:

- (33) *pət̪sa-ŋi-tə-bəl-ɟy-tɕ*                    *tɕuatsə̌*  
           just.now-ADV-DIR-make-ASP-GEN    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-zə-s*                    *khuə-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) *qɑ-fia-ɸdʒe-m*                    *khuə-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, generally by /-s/ (e.g. 37) and sometimes by /-m/.

- (36) *the: e-ze stu nə mo-xɕu.*  
           3sg one-CL alone sleep NEG-dare  
           ‘S/he doesn’t dare sleep alone.’

- (37) *tɕile-(ŋuəŋi) pə-s pies ŋuə.*  
 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 /mɑ-tɕi-Verb/ ‘had not yet Verb’ in the first clause and be nominalized by /-tɕ /, as in (38).

- (38) *nəs, qa mɑ-tɕ-kə-tɕ, 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<sup>1</sup>/ ‘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 /ŋiantɕi/ ‘following’ is used at the end of the first clause:

- (39) *qa stuaha sə-tɕhə-ŋike, suə χuəla.*  
 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 /ŋiantɕi/. To mark the purpose of an action, the post-position /χuaŋi/, 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 /χuaŋi/.

A cause–effect relation can also be marked by adding the instrumental postposition or the manner adverbial marker /ŋi/ to the end of the first clause:

- (40) *the:-dzoqu-le dagə-wu, pitɕ sei mɑ-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: dɕiq hɑ-qə-ŋi die-ɕe.*  
 3sg cliff DIR-go-ADV DIR-die  
 ‘S/he died from falling off the cliff.’

To express the concessive, the phrase /hɑ-ŋə-ŋə-lu/ (DIR-COP- ‘come’) is added to the end of the first clause:

- (42) *qa quaha hɑ-χuəla-haŋə-ŋə-lu, χo:ts fie-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) *χsutsqha-le-wu tə-ke<sup>1</sup>-l-kui, ‘ha! ʔū ŋi-χuaŋi kə-zei-n-a?’*  
 stomach-DEF-AGT DIR-ask-HS EXCL 2SG What-because DIR-cry-2sg-Q



*jə-kui*, ‘*niχuaŋi kə-zei-n-a?*’     *jə-kui-tu*,     *kəpətʂ-tou*  
 say-HS     why     DIR-cry-2sg-Q     say-HS-LINK     orphan-DEF:one:CL  
*he-zgue-kəi*, ‘*qɑ-ŋuəŋi ep lə me-zə-i*’  
 DIR-answer-HS     1sg-TOP     father     also     NEG-have/exist-CSM  
*əw lə me-zə-i*,     *gəx-ŋi fiə-xtʂəp ŋiaufu*  
 mother     also     NEG-have/exist-CSM     night-ADV     DIR-dark     as.soon.as  
*ʂkup-le qɑ dzə-la:’ jə-kəpə-ʂə, ...*  
 orangutan-DEF     1sg     eat-come:FUT     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 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.

## REFERENCES

- Graham, David Crockett (1958) *The Customs and Religion of the Ch'iang (Smithsonian Miscellaneous Collection Vol. 135, No. 1)*, Washington, D.C.: 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, Chenglong (1992a) ‘Qiangyu fufuyin de yanbian’ (The changes in the consonant clusters of Qiang), *Qiangzu Yanjiu* 1992.2: 152–7.
- Huang, Chenglong (1992b) ‘Qiang zu’ (The Qiang), *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 Ronghonghua (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: 181–89.
- Huang, Chenglong (1997) ‘Qiangyu dongci de qianzhui’ (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.
- 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. (1995) ‘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. with Huang, Chenglong (to appear) ‘Grammatical sketch of the Qiang language, with texts and annotated glossary’, Berlin: Mouton de Gruyter.
- 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: 50–8.
- 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: 30–6.
- Editorial Board (1985) *Aba Zangzu Zizhi Zhou Gaikuang (Overview of Aba Tibetan Autonomous Prefecture)*, Chengdu, Sichuan: Sichuan Nationalities Press.
- 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), *Minzu Yuwen Yanjiu Wenji*, Minzu Yuwen (ed.) 189–224.
- Zhou, Xiyin and Liu, Zhirong (1993) *Qiang Zu (The Qiang nationality)*, Beijing: Minzu Chubanshe.

# PRINMI: A SKETCH OF NIUWOZI\*

*Picus Sizhi Ding*

## 1 INTRODUCTION

Spoken by the Pumi and Tibetan nationalities in northwestern Yunnan and southwestern Sichuan, Prinmi (or Pumi) is a member of the Qiangic branch of the Tibeto-Burman languages (H. Sun 1982; Bradley 1997: 35–37). With widespread settlement in the mountain areas, speakers of Prinmi live close to many communities of distinct Tibeto-Burman languages. The two major dialectal groups are Northern Prinmi and Southern Prinmi, with some mutual unintelligibility (Lu 1983: 90). In addition, there are also clan-based sociolects.

Niuwozi Prinmi, the variety presented in this chapter, is from central Ninglang Yi Autonomous County, Yunnan, neighbouring Sichuan. It falls within Southern Prinmi under Lu's (1983) scheme, in the borderland between Southern and Northern Prinmi.

## 2 PHONOLOGY

### 2.1 Syllable structure

Unless treated as independent segments, glides are considered as part of the nucleus of the syllable, to avoid overloading the consonant inventory (but see Matisoff 1997 for a different approach). Thus, Prinmi syllables are analysed in Figure 36.1.

### 2.2 Vowels, glides, and consonants

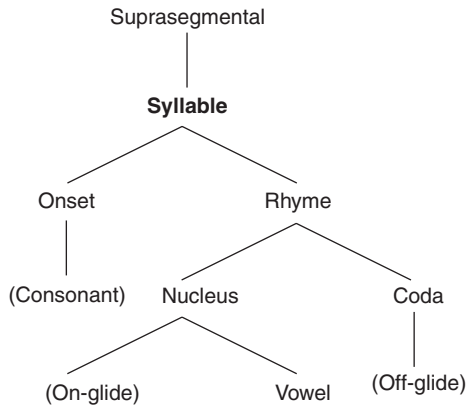
Niuwozi Prinmi has fifteen monophthongs. These are given in Table 36.1 according to their height, frontness, roundedness, and nasality.

The phoneme /i/ is often realized as one of the three fricative vowel allophones, which are homorganic with a preceding sibilant fricative segment (including that found in an affricate), e.g.  $s^R$  [sʃ] 'to die',  $tʃ^h$  [tʃʰ] 'dog', and  $z^F$  [zʃ] 'four' (the raised letters at the end of syllables indicate surface tones: R = rising, F = falling, and H = high level; the low level is left

---

\*The description of Prinmi presented here is based on fieldwork done in Ninglang County, Yunnan. I am indebted to many Pumi people for their invaluable help.

Except for suprasegmental notation, Prinmi data are presented in IPA according to the latest revision of 1996. Abbreviations – *Assr*: assertive; *Dc*: discourse clitic; *ExT*: external topic; *fr.sp*:- away from the speaker; *Inct*: incitee; *InT*: internal topic; *intr*: intensitive (causative); *Inv*: involmental; *LGS*: locative/goal/source; *M*: modificatory; *nInv*: non-involmental; *N<sub>ds</sub>*: desiderative negator; *NS*: nominal sentence marker; *to.sp*:- towards the speaker; and *Vlt*: volitive.



**FIGURE 36.1** THE STRUCTURE OF PRINMI SYLLABLES

unmarked). The non-fricative allophone contrasts with its rounded counterpart, e.g.  $g^F$  ‘nine’ vs  $g^R$  ‘(of things) old’.

Phonetically /y/ is realized as [ɥi], but /y/ is a true monophthong. Apart from /y/, /ỹ/, and /ĩ/, high and mid high vowels are prone to lowering; hence /i/ can be [i] or [ĩ], /o/ can be [o] or [ɔ̃], and so forth. When words with lowered vowels were played back from isolated word lists, some listeners confused /k<sup>h</sup>u<sup>R</sup>/ [k<sup>h</sup>ɥ<sup>R</sup>] ‘head’ with /k<sup>h</sup>o<sup>R</sup>/ ‘needle’. Preliminary cross-dialectal comparison reveals that Niuwozi /ɜ/ is historically related to \*ɛ. (The shift probably responds to the pressure from vowel lowering.) This central vowel contrasts with schwa, e.g.  $mɜ^F$  ‘bamboo’ vs  $mɔ^F$  ‘sky’.

Three glides occur in Niuwozi Prinmi: /j/, /ɥ/, and /w/. The latter two are almost in complementary distribution but contrast in such minimal pairs of native words as  $kɥɐ^H$  ‘melon’ vs  $kwɐ^H$  ‘(of eye) to open’. Treating glides as part of the rhyme in the syllable leads to a number of complex vowels, mostly rising diphthongs. All three glides can function as the on-glide in rising diphthongs, as shown in (1a), but only /j/ and /w/ may serve as the off-glide in falling diphthongs or triphthongs, as shown in (1b).

- (1) a. ju, jɜ, jɐ, jẽ, jõ, ɥe, ɥɛ, ɥɐ; ɥẽ; wɜ, wɐ, wa, wã  
 b. ɜj, ɐw; wɜj

**TABLE 36.1** MONOPHTHONGS OF NIUWOZI PRINMI

		Front		Central		Back
High	oral	i	y	ɨ	ɥ	u
	nasal	ĩ	ỹ			
Mid high	oral	e				o
	nasal					õ
Mid				ə		
Mid low				ɜ		
	oral			ɐ		ɑ
Low	nasal			ẽ		

**TABLE 36.2** CONSONANTS OF NIUWOZI PRINMI (CONSON = CONSONANT)

	<i>Bilabial</i>	<i>Dental</i>	<i>Post-alveolar</i>	<i>Retroflex</i>	<i>Velar</i>	<i>Glottal</i>
Plosive	b p p <sup>h</sup>	d t t <sup>h</sup>			g k k <sup>h</sup>	
Fricative		z s	ʒ ʃ	ʒ ʃ	ɣ	h
Nasal	m ɱ	n ɳ				
Lateral		l ɭ				
Rhotic				ɻ ɻ		
Complex consonant	b <sup>ʰ</sup> p <sup>ʰ</sup> p <sup>hʰ</sup>	dz ts ts <sup>h</sup>	dʒ tʃ tʃ <sup>h</sup>	ɖz tʂ tʂ <sup>h</sup>	g <sup>ʰ</sup> k <sup>ʰ</sup> k <sup>hʰ</sup>	

Table 36.2 displays all forty consonants in Niuwozi Prinmi. Complex consonants are those requiring more than one manner of articulation. Except for the two fricatives /ɣ/ and /h/, all the consonants show voicing contrasts. Further, the plosives and the complex consonants also show distinctive aspiration.

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 of /ʒ/ before /i/ or /j/ (not applicable to /ʃ/); the bilabial fricative [ɸ], an allophone of /h/ before /u/; and the optional palatalization of /n/ to [ɲ] before /i/ or /j/.

### 2.3 Suprasegmentals

In spite of there being a great number of compounds, Prinmi lexical words are predominantly monosyllabic. Monosyllabic words may bear a high tone, a falling tone, or a rising tone; e.g. *bjɛ<sup>H</sup>* ‘busy’, *bjɛ<sup>F</sup>* ‘urine’, *bjɛ<sup>R</sup>* ‘to fly’. The suprasegmental contrast between the high tone and the falling tone is rather difficult to perceive when words occur in isolation, as this environment is not ideal for manifesting the distinction between them. However, when monosyllabic words appear in connected speech, these three tones effectively contrast in terms of two pitches: high vs low. The falling and rising tones are often split into a high-low and a low-high sequence respectively in longer domains. Similarly, the high tone gives rise to high-high (details in Table 36.3 below). The tone partition is obligatory in clitic groups (cf. examples in Section 3.5).

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 combinations of these two settings yields a total of seven categories:

**TABLE 36.3** SUPRASEGMENTAL CATEGORIES IN NIUWOZI PRINMI

<i>Category</i>	<i>Parameters</i>	<i>Quadrisyllabic</i>	<i>Trisyllabic</i>	<i>Disyllabic</i>	<i>Monosyllabic</i>
1	[1] [+spread]	H-H-L-L	H-H-L	H-H	High
2	[1] [-spread]	H-L-L-L	H-L-L	H-L	Falling
3	[2] [+spread]	L-H-H-L	L-H-H	L-H	Rising
4	[2] [-spread]	L-H-L-L	L-H-L		
5	[3] [+spread]	L-L-H-H	L-L-H	L-R	
6	[3] [-spread]	L-L-H-L			
7	[4] [spread]	L-L-L-H	L-L-R		

The bracketed number indicates the location of the H tone on the prosodic domain, while the other signifies the spreading of the H tone. The tonal category, as specified by the parameters, remains unchanged throughout a given row, although the tones on the right are subject to truncation as the domain becomes smaller.

These suprasegmental categories suggest that Prinmi has a pitch-accent system, which is based on the word rather than the syllable, a departure from the better known tone languages of East and Southeast Asia (some word-based tone languages of Nepal are discussed in Mazaudon 1977: 76–84).

## 2.4 Phonological alternations

Phonologically-conditioned segmental variation is relatively limited, but ‘tone sandhi’ is pervasive in Prinmi (cf. numbered examples below where surface tones instead of citation tones are given). Much of the variation ensues from differences in the prosodic domains. The effects of suprasegmental phonological processes are (a) shifting of the H tone to an adjacent syllable (usually to the right), (b) the reversal of spreading, (c) the merger of two prosodic domains, and (d) the split of a prosodic domain into two.

## 3 MORPHOLOGY

### 3.1 Lexical categories

Prinmi lacks derivational morphology that exclusively changes lexical categories. Consequently there is considerable overlap between noun, verb, and adjective. Particularly scarce are mono-functional adjectives – some known instances are *bulp<sup>H</sup>* ‘many’ and *gu<sup>H</sup>ʒi* ‘(adjective) middle’. Incidentally, the latter also has a noun form *guʒi<sup>H</sup>* ‘(noun) middle’; the two constitute the only minimal pair of related lexical items contrasting in category. As a rule of thumb, when a simplex non-grammaticalized word modifies a noun postnominally, it functions as an adjective. Some adjectives may serve as nouns, verbs, or both. Another notable property of adjectives is that they tend to form noun-adjective compounds with their head nouns.

Eleven minor lexical categories can be recognized, including (low tones in the examples here are marked explicitly, as toneless syllables get no marking):

- existentials – *ʃi<sup>R</sup>* ‘(abstract things) to exist’, *ʒe<sup>F</sup>* ‘(animate) to exist’, *dʒi<sup>H</sup>* ‘(animate) to exist among a group’, *djɔ<sup>F</sup>* ‘(inanimate) to exist’, *kɥe<sup>F</sup>* ‘(inanimate) to exist within something’, and *tʒ<sup>H</sup>* ‘(inanimate) to exist on something’;
- auxiliary verbs – e.g. *k<sup>h</sup>u<sup>R</sup>* ‘must; need’, *ʒjɛ<sup>R</sup>* ‘can’;
- demonstratives – e.g. *tɔ<sup>F</sup>* ‘this’, *ɔ<sup>H</sup>tɔ<sup>L</sup>* ‘that’, *tɔ<sup>H</sup>ɔ<sup>L</sup>* ‘these’;
- pronouns – e.g. *ʒ<sup>H</sup>* ‘I’, *ne<sup>R</sup>* ‘you (singular)’, *ne<sup>L</sup>dʒɛ<sup>H</sup>* ‘you (dual)’, *ʒ<sup>L</sup>ɔ<sup>H</sup>* ‘we (inclusive)’, *d<sup>L</sup>ɔ<sup>H</sup>* ‘we (exclusive)’, *ni<sup>F</sup>* ‘she/he/it’;
- numerals – e.g. *ti<sup>R</sup>* ‘one’, *ni<sup>R</sup>* ‘two’, *sɔ<sup>R</sup>* ‘three’;
- counters (a less-grammaticalized kind of classifier) – e.g. *tsi* ‘(for humans)’, *bɔ* ‘(for trees)’;
- ideophones – e.g. *tʃɛtʃɛ* ‘an intensive state of whiteness’, *ʃɔʃɔ* ‘an excessive state of abundance’;
- adverbs – e.g. *ti<sup>L</sup>ti<sup>H</sup>* ‘slowly’, *lʒ<sup>L</sup>ljɛ<sup>R</sup>* ‘very’;
- onomatopes – e.g. *k<sup>h</sup>ɔ<sup>L</sup>ɔ<sup>h</sup>ɔ<sup>L</sup>* ‘sound of weighty objects falling into deep water’;
- interjections – e.g. *wi<sup>F</sup>* ‘expressing an unpleasant surprise’; and
- postpositions – *po* ‘under’, *to* ‘on’, *k<sup>h</sup>u* ‘on top of’, *wu* ‘inside’, *lo* ‘outside’, *tʃe* ‘at one’s place’, *be* ‘at; to; from’ (also marking the semantic roles of locative, goal, and source), *ɔ* ‘with (instrumental)’, *ni* ‘with (comitative)’, *bo* ‘for (beneficiary)’, and *ki* ‘(incitee marker)’.

Many postpositions are grammaticalized from nouns. The ongoing process is changing these words into clitics; the instrumental has already become a clitic, able to form composite variants with other clitics, e.g. *ɬə* (with the plural *ɬə*).

### 3.2 Affixes

Suffixes are more common than other affixes in Prinmi. They impart additional meanings to a word, e.g. *pɔ-tsɿ<sup>H</sup>* ‘small flower’ (flower+suffix); *tsəw<sup>H</sup>-dɿ* ‘beating implement’ (beat+suffix); *t<sup>h</sup>jɛ-ji<sup>H</sup>* ‘beverage’ (drink+suffix); *ɬə-ɜw<sup>H</sup>* ‘to be prone to laugh’ (laugh+suffix).

Antonyms of some descriptive verbs/adjectives are derived with a negative prefix, e.g. *mα-ɜd<sup>H</sup>* ‘unclean’ (prefix+clean).

Direction marking is frequently found on verbs. Most verbs can be affixed with at least one of the six directional prefixes given in Table 36.4 (cf. Fu 1998: 27–72 for a detailed discussion of the directional prefixes in Dayang Prinmi):

TABLE 36.4 THE DIRECTIONAL PREFIXES

Category	Prefix	Meaning	Prefix	Meaning
Person	dɜ-/də-	towards the speaker	t <sup>h</sup> ɜ-	away from the speaker
Space	gə-/k <sup>h</sup> ə-	outwards	ɜ-/hɜ-	inwards
Vertical	nɜ-	downwards	tɜ-	upwards

Other identified nominal affixes are the vocative prefix *ɜ-* and the conjunctive infix *-mə-*, both unproductive, e.g. *ɜ-k əw<sup>H</sup>* ‘maternal uncle’; *ŋi<sup>H</sup>-mə-ŋi* ‘days’.

### 3.3 Inflection

A small number of verbs and auxiliary verbs in Niuwozi Prinmi inflect for ‘subject’ agreement in terms of person and number (but cf. 4.1), as in Table 36.5:<sup>1</sup>

TABLE 36.5 THE CONJUGATION OF THE COPULA AND VERBS

Root	First person singular	Second person singular	1/2 person plural	Third person	Meaning
<i>dzɿ<sup>F</sup></i>	<i>djɛ<sup>F</sup></i>	<i>dəw<sup>F</sup></i>	<i>dɿ<sup>F</sup></i>	<i>dzɿ<sup>F</sup></i>	be
<i>pα<sup>F</sup></i>	<i>pjɛ<sup>F</sup></i>	<i>pu<sup>F</sup></i>	<i>pɿ<sup>F</sup></i>	<i>pα<sup>F</sup></i>	do
<i>t<sup>h</sup>jɛ<sup>R</sup></i>	<i>t<sup>h</sup>ɔ<sup>R</sup></i>	<i>t<sup>h</sup>jɔ<sup>R</sup></i>	<i>t<sup>h</sup>ɿ<sup>R</sup></i>	<i>t<sup>h</sup>qɛ<sup>R</sup></i>	drink
<i>mɛsɿ<sup>H</sup></i>	<i>mɛsjɛ<sup>H</sup></i>	<i>mɛsəw<sup>H</sup></i>	<i>mɛsɿ<sup>H</sup></i>	<i>mɛsɿ<sup>H</sup></i>	know

The inflections are often associated with the perfective aspect but their occurrence is variable.

Reciprocal verbs are derived through reduplication, e.g. *tsɜ<sup>H</sup>tsəw* ‘to fight’ < *tsəw<sup>H</sup>* ‘to beat’; *gjɜge<sup>R</sup>* ‘to help each other’ < *ge<sup>R</sup>* ‘to help’. Notice the modification of the vowel in the first syllable.

1 A rich verb inflection is found in dialects spoken in Lanping County, Yunnan. Cf. Jinghua Prinmi in Lu (1983: 42–45) and Dayang Prinmi in Fu (1998: 104–14).

A few causative verbs are remnants of an archaic derivational prefix of Proto-Tibeto-Burman. They survive in the form of a minimal contrast on the consonants of the base forms, e.g.  $k^h j\bar{e}^H$  ‘to break (transitive)’ vs  $g^h j\bar{e}^H$  ‘to break (intransitive)’;  $p^h e^F$  ‘to destroy’ vs  $be^F$  ‘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 quadrisyllabic; trisyllabic ones often result from compounding a reduplicated formative such as an ideophone, e.g.  $b\mu^H t_j v t_j v$  ‘really thin’. With a binary structure, embedding is permitted in complex compounds. For instance, the structure of ‘fibula’ is built on three levels of compounding:

$$\{[(k^h \bar{a}^H .ju^H) .\mathcal{I}3ka] [(lo t_j^h \eta e^H) .\mathcal{I}3^H ka]\}$$

foot stem bone out side bone

### 3.5 Clitics

Nominal and verbal clitics can be distinguished on the basis of the lexical category of the host. Except for negative and interrogative clitics, all others are enclitics. The major nominal clitics are the plural  $\mathcal{I}2$  and the dual  $dz\bar{e}$ , e.g.  $m\mathcal{I}$ (=plus) $\mathcal{I}2^H$  ‘fire’, ‘flames’ and  $k^h \bar{a}^H dz\bar{e}^H$  ‘two feet’. Other nominal clitics mark topics (see Section 4.9) and modifiers (see Section 4.2). They can cling to verbs occasionally. A minor type of nominal clitic includes *bo* and *ne*, called ‘discourse clitics’. They function much like gap fillers and can be readily omitted (as witnessed in editing recorded spontaneous speech).

Verbal clitics are richer in number and form. The experiential  $t\#$  signifies having an experience of doing something. The durative  $n\bar{o}$  expresses a continual state (cf. (20) and (27)). The hortative *gi* conveys the wish to engage in some activity together. The evidential  $t_j\bar{i}/t_j\bar{v}$  is used for marking hearsay information whose truth is not committed to by the speaker, e.g. (2a) and (2b):

(2a)  $mjv^H bu^H ge bo k^h \bar{a}^H gj\bar{o} to n\mathcal{I} \text{-} t_j swa + si^H t_j \bar{i} + .jju$   
 eyelid InT Dc knee on down-contact + Pfv say + nInv  
 ‘It is said that the eyelids touch down to the knees.’

(2b)  $p\mu^H fi \mathcal{I}^H \mathcal{I}^H m\bar{a} ti g\bar{a} \text{-} t_j^h j\bar{o},$   
 last night 1sg dream one out-appear  
 $\mathcal{I}^H h\bar{e} ti^H be^H t^h \mathcal{I} \text{-} p^h v w^H + si t_j \bar{i} + .jju.$   
 1sg parrot one LGS fr.sp-transform + Pfv say + nInv  
 ‘Last night I dreamt of myself transforming into a parrot.’

(2c) ‘*ne bo^H \mathcal{I}^H gj\bar{a}^H g\bar{a} \text{-} dz\bar{e} v w^H + si.*’  $t_j \bar{i} + .jju.$   
 2sg ExT 1sg M out-eat:2sg + Pfv say + nInv  
 ‘“You’ve eaten mine,” said (the Dragon King).’

(2d)  $t_j v^R t_j \bar{i} d\bar{i}^H t_j \bar{i}^H l\bar{e} t_j^h j\bar{o} + t_j j\bar{e} t_j^h i + .jju.$   
 now deluge appear + say do + nInv  
 ‘Now a deluge is about to take place.’



**TABLE 36.6** ATTACHMENT PATTERNS OF THE INTERROGATIVE AND NEGATIVE CLITICS

Bare monosyllabic	<i>a/ma/me/tjɛ</i> + Verb
Prefixed, no other clitics	Prefix + <i>a/ma/me/tjɛ</i> + Verb
With other verbal clitics	(Prefix) + Verb + <i>a/ma</i> + Asp/Mod/(non)-Involvemental

Notes: ASP = ASPECT CLITIC; MOD = MODAL CLITIC.

In specific discourse contexts, it may mark mirative, quotative, or imminent future, as exemplified respectively in (2b)–(2d) (cf. J. Sun 1993; DeLancey 1997).

Aspect, modality, and evidentiality can be expressed with sets of clitics sensitive to the factor of controllability. The set of perfective clitics consists of *si*, *sjɛ̃*, *sī*, and *me*, the last being exclusively for negation (cf. example (6)). The basic perfective is *si* (cf. (8) and (14)). When an activity is controllable by the agent, *sjɛ̃* is chosen (for first person singular) or *sī* (for non-third person plural). Likewise, the pair of modal clitics – *fo* and *kɔj* – are differentiated in terms of controllability. When the wish to do something is under the agent’s direct control, *fo* is selected; otherwise, *kɔj* is employed (cf. (24)). The set of involvemental clitics – *ɔ̃*, *ju*, and *ɿ*, opposed to the non-involvemental *ɿju* – also embodies controllability in the basic meaning (cf. (9) and (11)). The involvementals convey a direct source of information and the agent’s control of specific situations, whereas the non-involvemental indicates a lack of control over a situation and/or direct access to information. This is comparable with similar (but not identical) evidentials in Amdo Tibetan (J. Sun 1993) and rGyalrong (Lin 1993: 194–5). The variants are used according to the person and number of the agent – *ɔ̃* (1 singular), *ju* (2 singular), and *ɿ* (1/2 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: *gɣɸ* (expressing surprise), *ma* (making a suggestion), *pa* (speculation), and *mə* (making an assumption).

There are three specialized negative clitics: the general negator *ma*, the perfective negator *me*, and the desiderative negator *tjɸ* (occurring mainly in imperative sentences). These negators and the interrogative clitic *a* are the only ones that may precede their hosts. Their overall attachment patterns are summarized in Table 36.6.

## 4 SYNTAX

### 4.1 Grammatical system

Prinmi does not operate on a grammatical system with a clearly definable subject and object. Although there is unstable agreement found on a small number of verbs, no other grammatical phenomena suggest the existence of subjects in the language.

It is noteworthy that Prinmi tends to mark agents with the instrumental clitic *ɔ̃*, especially for a third-person singular agent (cf. examples (13) and (22)). Sometimes this optional marking is also found after non-third-person agents. The inconsistency suggests that the development of ergativity in Prinmi is at an early stage (cf. 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 36.2):

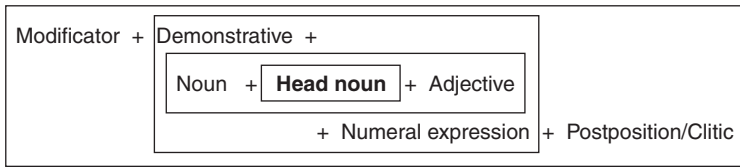


FIGURE 36.2 A LAYERED ANALYSIS OF THE NOUN PHRASE

The ‘modifier’ is a special kind of modifier, typically signified by the modificatory clitic *a* or its complex variants such as *gja*. The term is meant to distinguish these from other more general modifiers such as lexical or compound attributes. Genitivization (cf. (2c)) and relativization (cf. (3)) are effected through the modificatory relation.

The following noun phrase contains all the possible elements except for the postposition/clitic:

- (3) [mjɐ<sup>H</sup>bu tɔj] [tʃɪdɪ<sup>H</sup> di<sup>H</sup>-mi + a] tɔ<sup>H</sup>  
 eyelid big flood cast-er+M this  
 ʃɪ<sup>H</sup>bɑ<sup>H</sup> ɲəgɟɛ<sup>H</sup> ʃɪ<sup>H</sup>bɑ<sup>H</sup> pudi nɔ<sup>H</sup>tsɪ  
 Dragon family old man Dragon family old woman two Counter  
 ‘these two old man and old woman of the Dragon family who have big eyelids  
 and started the flood’

There are two modifiers (placed within brackets), a demonstrative, a pair of coordinate heads (boldfaced), each modified by a noun, 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. Modifiers of the predicate such as temporals and locatives are situated in the periphery. These three together form a clause expandable to a simplex sentence as follows, as shown in Figure 36.3.

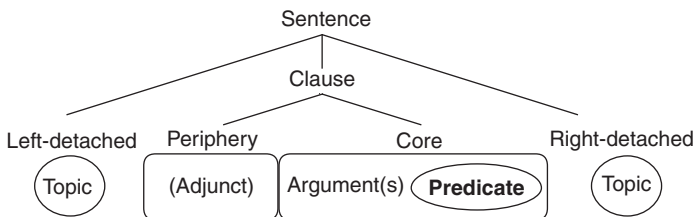


FIGURE 36.3 THE STRUCTURE OF THE SIMPLEX SENTENCE

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) in the core layer and then the verb in the nucleus. 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. Initial topics are dominant while final ones are marginal. For instance (bracketing indicates the layers within a clause; N = nucleus, C = core, and P = periphery),

- (4)  $ne^R$   $P[ɰəbe^H]P$   $C[ʈʂɿʈʂu$   $ge^H$   $N[nɜ-di$   $nɜ-ʃi^H$   $ke^H]N]C$ .  
 2sg first LGS pestle InT down-cast down-go Caus:intt  
 ‘You throw the pestle down first.’

The presence of the temporal modifier  $ɰəbe^H$  ‘first’ in (4) is particularly helpful for recognizing a left-detached argument. For the core argument  $ne^R$  ‘you’ to precede the modifier at the periphery, it must be situated at the left-detached position outside the clause. Otherwise, it could only follow the adverbial modifier and occur within the core layer.

#### 4.4 Major sentence types

Major sentence types are exemplified in (5)–(7): declarative, negative, and interrogative.

- (5)  $pɜ^Hdi^H$   $ɜ^H$   $qz_u$   $dzɿ^H$ .  
 frog 1sg friend Copula  
 ‘The frog is my friend.’
- (6)  $mɜ^Hʈʂe^H$   $ɰi^R$   $me+t^hɜ^H$ .  
 seek get Neg:Pfv + succeed  
 ‘(He) searched (but) couldn’t get (one).’
- (7)  $pɔ̃^Hpɔ̃$   $a^H+ʒe?$   
 uncle Q+exist:anim  
 ‘(Is uncle (here))?’

Prinmi also has tag questions which always imply pragmatic presupposition, e.g.

- (8)  $ne^R$   $dɜ-ji^H+si$   $a?$   
 2sg to.sp-come:2sg + Pfv Q  
 ‘You’ve come over, haven’t you?’

Information-seeking questions are formed with an interrogative pronoun, without any change of word order, e.g. (cf. also (15))

- (9)  $ne^R$   $hɜ^Hgi$   $ʃi+ɰu?$   
 2sg where go + Inv:2sg  
 ‘Where are you going?’

Imperative sentences generally show consistent agreement with the agent, e.g.

- (10a)  $(ne)$   $dzɐw^H$ .  
 2sg eat:2sg  
 ‘(You) eat.’
- (10b)  $(neɰə^H)$   $tjɐ+dzɿ^H$ .  
 2pl  $N_{ds}$  + eat:2pl  
 ‘(You) don’t eat.’

They are negated with the desiderative negator, as in (10b). The presence of the agent in imperative sentences is optional.

#### 4.5 Clause compounding

When two verbs are adjacent in a sentence, the serial verbs may form a unit – the ‘double-verb predicate’ (the term ‘double’ is not to be taken literally). With embedding, it is possible to have more than two verbs in the complex predicate. As with morphological compounds, the conjoining is binary in nature. These verbs are syntactically interdependent without subordinating one to the other. (11) illustrates a typical double-verb predicate in Prinmi.

- (11)  $t\check{\sigma}gy^H$   $ge^H$   $ni^H$   $be$   $ge-j\check{\sigma}t^h\check{\sigma}-k^h$   $j\check{\epsilon}$   $f\check{i}^H$   $a+ .jju?$   
 attire InT 3sg LGS out-fetch fr.sp-give Aux Q+nInv  
 ‘Will (the Golden Pheasant) go bring and give the clothes back to him?’

- (12)  $t\check{\sigma}gy^H$   $ge^H$   $ge-j\check{\sigma}^R$   $ni^H$   $be$   $t^h\check{\sigma}-k^h$   $j\check{\epsilon}$   $f\check{i}^H$   $a+ .jju?$   
 attire InT out-fetch 3sg LGS fr.sp-give Aux Q+nInv  
 ‘Will (the Golden Pheasant) go bring the clothes and give (them) back to him?’

Note that the auxiliary verb is not part of the double-verb predicate; rather, it takes the clause headed by the serial verbs as its complement. Given in (12) is another type of clause compounding. With the shared argument (single-underlined) now intervening between the verbs, the sentence cannot be analysed as comprising a double-verb predicate which is essentially formed with two verbs at the nucleus layer. Notwithstanding its identical meaning to (11), (12) is syntactically a clause-chaining sentence, sharing its structure with the following sentence:

- (13)  $m\check{\sigma}^H$   $g\check{\sigma}$   $ne$   $m\check{\sigma}^H$   $to$   $b\check{u}^H$   $\check{\sigma}^H-dy$   $t\check{i}^H$   $\check{\sigma}^H-dy$   
 sky Inst Dc sky on sun in-cast:3 moon in-cast:3  
 ‘The heavenly god placed a sun and a moon in the sky.’

As clause-chaining sentences involve the conjoining of units at a level higher than the nucleus, they are looser in structure and laxer in prerequisites, and thus occur more commonly in the language.

#### 4.6 Subordinate clauses

Embedded clauses are prevalent in subordination, for example, in relative clauses, in copular periphrastic clauses, as complements to auxiliary verbs, and as complements to verbs of cognition. Prinmi uses nominal sentences to convey a sense of discovery gained from cursory observation of the surroundings. The structure is straightforward, with simply a relative clause and the marker *ti*, e.g.

- (14)  $t\check{\sigma}^H$   $mi^H$   $ge$   $n\check{\sigma}-s\check{i}$   $si^H+a$   $ti$ .  
 this person InT down-die Pfv+M NS  
 ‘This person (is) dead.’ (occurring in a direct speech)

As in other relative clauses, the relative clause in a nominal sentence may contain a modificatory clitic, as in (14), but the clitic is omissible.

When the copula is used in periphrastic constructions, its complement clause is nominalized by the suffix *-ji* in the obligational construction, e.g. (15), or by the suffix *-mi* in the focus-supposition construction, e.g. (16).

- (15) *me<sup>H</sup> t<sup>h</sup>ɔ-tʃ<sup>h</sup>i-ji dzɪ?*  
 what fr.sp-do-nomzr Copula  
 ‘What should (we) do?’
- (16) *ʃʃɛ<sup>H</sup> gō<sup>H</sup> gɔ-tsɐw<sup>H</sup>-mi<sup>H</sup> dzɪ.*  
 iron Inst out-beat-nomzr Copula  
 ‘(The tool,) what I’m saying is: (it) is forged from iron.’

Complement clauses to auxiliary verbs are not marked by grammatical means, frequently leading to a juxtaposition of the complement verb and the auxiliary, e.g.

- (17) *kwɔʒ<sup>H</sup> bo<sup>H</sup> tʃɐ+kwɔʒ<sup>H</sup> ku<sup>H</sup>.*  
 cry ExT N<sub>ds</sub>+cry must  
 ‘As for crying, (you) mustn’t cry (any more).’
- (18) *[[[ɔɪ<sup>H</sup> bo dʒzɔdʒzɔu tʃ<sup>h</sup>ɪ<sup>H</sup>] ʒō<sup>H</sup>ʒī a+ɔju] ɔ-ʃɐʃɐ<sup>H</sup>] ti*  
 1pl:inc ExT friend do want Q+nInv in-test  
*pu.*  
 one do o:2sg  
 ‘(You) give 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 reflected in the negation pattern in (17) and embedded subordination in (18). When a ‘double-verb predicate’ is negated, the negative clitic appears on the final unit, as in (19). However, the negator occurs within the complement clause in (17).

- (19) *ne<sup>R</sup> tɔ<sup>H</sup>-pe<sup>H</sup> tʃ<sup>h</sup>ɪ gɔ-ʒɪ<sup>R</sup> ma<sup>R</sup>+ke bo*  
 2sg up-spew do out-come Neg+Caus:intt ExT  
 ‘If you don’t cough (him) out.’ (*lit.* ‘Cough up – don’t let come out’).

Three orders of subordination are observed in (18). Of these, only the head of the inner most complement is adjacent to a verb. The other instances see intervention of a clitic 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.

- (20) *ni<sup>H</sup>ɔ t<sup>h</sup>ɔ-ʃuʃu<sup>H</sup>+nō tʃ<sup>h</sup>wa tɔʒ<sup>H</sup>pa+nō k<sup>ʰ</sup>ōk<sup>ʰ</sup>t<sup>H</sup>.*  
 3pl fr.sp-joint+Dur voice big do+Dur sing  
 ‘They sing aloud together.’

Both adverbial clauses in (20) happen to contain the durative clitic *nō*, 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, using the following structure:

causer+ affectee/goal/incitee+ complement clause+ *ke/kyɛ*

Irrespective of how the causee is encoded in the construction, it is always shared between the causative verb and the complement head. With the presence of elements in the peripheral layer (the underscored part), (21) illustrates that the causee is not part of the complement clause. If it were, it could be easily rendered after the locative inside the complement clause.

- (21)  $\eta i^H$   $k\alpha ts\alpha j$   $ge^H$   $bo$   $\underline{I\alpha t f^h y\epsilon^H}$   $\alpha\text{-}\alpha^H$   $ky\epsilon^H$ . (affectee)  
 daughter small InT Dc front side in-sleep Caus  
 ‘(She) made the youngest girl sleep at the front.’
- (22)  $fj\alpha^H$   $g\tilde{o}$   $ki^H pu$   $be$   $n\alpha\text{-}gw\alpha^H$   $ke$ . (goal)  
 g. pheasant Inst cuckoo LGS down-dress Caus:intt  
 ‘The Golden Pheasant got the Cuckoo to put (the clothes) on.’
- (23)  $\alpha^H$   $ne$   $ki^H$   $g\alpha^R po$   $n\alpha\text{-}f\alpha^H$   $ke^H$ . (incitee)  
 1sg 2sg Inct hill below down-go Caus:intt  
 ‘I let you go down the mountain.’

The causee is expressed differently according to the degree of its volitionality and/or affectedness. When a powerless causee is inevitably caught in a causation, it is encoded as affectee without any postposition marking, as in (21), although it may host topic clitics. In situations where the causee could have avoided the causation, it is marked by *be* as goal, as in (22). Finally, the causee may be eager to partake in some causation. In such cases, the willing causee is signified as incitee with *ki*, as in (23).

The use of the causative verb is influenced by the speaker’s perspective on the causation (for the notion of perspective, see Caughley 1982: 18–21). That is, the choice is subject to: (i) the speaker’s perspective on a particular causation and (ii) whether the causer factually has a direct control of the causation. When the causation is viewed as under the causer’s direct control, i.e. with an intentive causer, *ke* is used, as in (22) and (23). To express from other than the speaker’s perspective, *kyε* is employed, as in (21).

#### 4.8 The comparative construction

The comparative construction has the following structure:

theme + noun phrase/clause *to* + descriptive verb

The construction, strictly speaking, lacks morphosyntactic marking. The intended comparison is achieved only when the postposition *to* ‘than (literally: on)’ is construed as a sort of complementizer.

In transitive sentences, the comparative construction is embedded as a relative clause modifying the patient argument, as underscored in (24).

- (24)  $\alpha^H$   $n\alpha ge^H$   $\underline{t\alpha^H}$   $\underline{dj\epsilon ba^H}$   $to$   $\underline{t f^h y^H + a}$   $ti$   $bo$   
 1sg again this place than good +M one Dc  
 $m\alpha^H$   $t\alpha e^H$   $p^h\alpha^R$   $ma+k\alpha^H$ .  
 seek meet Neg +VIt  
 ‘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 clitics – *bo* for external topics, which typically lay out the discourse background, and *ge* for internal topics, which are targets of comments. However, these clitics are not reliable for indicating pragmatic topics. As their grammaticalization continues, the morphosyntactic markings do not always correspond to pragmatic topics. Moreover, a pragmatic topic can also appear bare, without any morphosyntactic marking (cf. examples (4) and (5)).

There are three subtypes of topic-comment construction. The ‘double-topic construction’ consists of a ‘scene topic’ (double-underlined below) and a topic in an aboutness relation (single-underlined) with its comment (inside brackets), e.g.

- (25)  $tʃi^H$   $ge^H$   $nj\epsilon^H$   $k^{h\epsilon}$   $\partial^H$   $po$   $tʃ^H-t\alpha^H$   $\lambda\alpha$   $k^h\epsilon$   $bo$ ,  
 water InT 2sg:M foot below up-reach nInv:M time ExT  
 $ne^R$  [ $\lambda\partial$   $be^H$   $tʃi^H$   $tʃu$   $ge^H$   $n\partial$ -di  $n\partial$ - $f^H$   $ke^H$ ].  
 2sg first LGS pestle InT down-cast down-go Caus:inT  
 ‘When the water reaches your feet, you throw the pestle down first.’

Besides temporal settings, scene topics are also used for locational and conditional settings (for the latter, cf. (19)). The ‘chained comment construction’ involves a single topic with several comments in chained clauses, e.g.

- (26)  $l\alpha^H b\partial$   $ge$  [ $sj\epsilon b\partial^H$   $tʃ^H$   $ts^h\partial$ ], [ $bi^H$ ], [ $\lambda\partial$ ] $j\epsilon^R$   $gq\epsilon^H$ ].  
 China fir InT tree one kind radially thick very tall  
 ‘The China fir is a kind of tree, radially thick, and very tall.’

It is also possible for a larger comment to comprise a topic-comment relation, forming an ‘embedded topic-comment construction’ (cf. Ding 1993: Chapter 4). The outer topic and the embedded topic in the construction must hold the semantic relation of set-member, which can easily be fulfilled when the embedded topic is a body part term, as in (27):

- (27)  $j\partial dz^H$   $papa$  [ $z\epsilon$   $ge^H$  [ $qz\partial+n\partial$ ]] [ $tʃo$   $ge^H$  [ $dzi$ ]].  
 bat face InT bad+Dur organ InT wonderful  
 ‘The bat, (its) face is ugly; (yet its) organs are perfect.’

- (28)  $tʃ\partial^H$   $to+a$   $qz\partial qz\partial^H$   $tʃ^h$   $f^H$   $mi$   $bo$   
 earth on+M bad do person ExT  
 $tʃi^H$   $\lambda\partial$   $bo$  [ $wu$   $qz\alpha^H$   $k^h\partial-tʃ^h j\partial$   $j\partial$ ],  
 seed sow ExT harvest poor out-appear Assr  
 $g\bar{i}$   $f\epsilon w^H$   $bo$  [ $qz\partial^R$   $m\alpha^R+j\partial$ ],  
 livestock raise ExT smooth Neg+Assr  
 $ts^h\partial$   $tʃ^h$   $f^H$   $\lambda\alpha^H$  [ $ku^H dz^H$   $\lambda i^R$   $m\alpha^R+j\partial$ ].  
 business do also profit get Neg+Assr

‘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.’

In (28), the set–member relation is realized as agent–activity. The set denotes activities done by the agent. The semantic relation of set–member is necessary, but not sufficient, for topic–comment embedding. Note that the complex structures of (27) and (28) contain not only topic–comment embedding but also chained comments.

## REFERENCES

- Bradley, David (1997) ‘Tibeto-Burman languages and classification’, in D. 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, Sizhi (1993) ‘The ba resultative construction: a comprehensive study of Mandarin ba sentences’, unpublished MA thesis, Simon Fraser University, Vancouver.
- Fu, Ailan (1998) *Pumiyu Dongci de Yufa Fanchou (Grammatical Categories of Pumi Verbs)*, Beijing: Chinese Cultural and Historical Publisher.
- LaPolla, Randy (1995) ‘Ergative marking in Tibeto-Burman’, in Y. Nishi, J. Matisoff, and Y. Nagano (eds) *New Horizons in Tibeto-Burman Morphosyntax*, Osaka: National Museum of Ethnology, 189–228.
- Lin, Xiangrong (1993) *Jiarongyu Yanjiu (An Investigation of rGyarlong)*, Chengdu: Sichuan Nationalities Press.
- Lu, Shaozun (1983) *Pumiyu Jianzhi (An Outline of Pumi)*, Beijing: Nationalities Press.
- Matisoff, James (1997) ‘Dayang Pumi phonology and adumbrations of comparative Qiangic’, in S. Premrirat *et al.* (eds) *Memorial Volume for Paul K. Benedict (Mon-Khmer Studies vol. 27)*, Bangkok: Mahidol University, 171–213.
- Mazaudon, Martin (1977) ‘Tibeto-Burman tonogenetics’, *Linguistics of the Tibeto-Burman Area* 3.2: 1–123.
- Sun, Hongkai (1982) ‘On the subgrouping of the Qiangic languages’ (in Chinese), in *Minzu Yuwen Yanjiu Wenji*, Xining: Qinghai Nationalities Press, 189–224.
- Sun, Jackson (1993) ‘Evidentials in Amdo Tibetan’, *Bulletin of the Institute of History and Philology, Academia Sinica* 63.4: 945–1001.
- van Valin, Robert and LaPolla, Randy (1997) *Syntax: Structure, Meaning, and Function*, Cambridge: Cambridge University Press.



# 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* 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) 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 w/. Tangut has no 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 ninety-seven rhymes, whereas the rising tone comprises eighty-six 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 in twelve rhyme groups, abbreviated to RG in Table 37.1.

**TABLE 37.1 RECONSTRUCTION OF TANGUT FINALS**

RG	Short vowels			Long vowels		
	Lax vowels	Tense vowels	Retroflex vowels	Lax vowels	Retroflex vowels	
I	1	[1.1] u [2.1]	61 [1.58] ʊ [2.51]	80 [1.75] ur [2.69]	5 [1.5] uu [2.5]	
	2	[1.2] ju [2.2]			6 [1.6] juu	
RG	3	[1.3] ju [2.3]	62 [1.59] jʊ [2.52]	81 [1.76] jur [2.70]	7 [1.7] juu [2.6]	
	4	[1.4] u [2.4]				
II	8	[1.8] e [2.7]	68 [1.65] ɛ [2.58]	82 [1.77] er [2.71]	12 [1.12] ee [2.11]	99 [2.84] eer
	9	[1.9] ie [2.8]	69 [1.66] iɛ [2.59]	83 [1.78] ier	13 [1.13] iee	
RG	10	[1.10] ji [2.9]	70 [1.67] jɪ [2.60]	84 [1.79] jir [2.72]	14 [1.14] jii [2.12]	101 [1.93] jiiir [2.86]
	11	[1.11] ji [2.10]				
III	15	[1.15] ē [2.13]				
RG	16	[1.16] jī				
IV	17	[1.17] a [2.14]	66 [1.63] ʌ [2.56]	85 [1.80] ar [2.73]	22 [1.22] aa [2.19]	88 [1.83] aar
	18	[1.18] ia [2.15]		86 [1.81] iar	23 [2.20] iaa	
RG	19	[1.19] ja [2.16]	67 [1.64] jʌ [2.57]	87 [1.82] jar [2.74]	24 [1.23] jaa [2.21]	89 [2.75] jaar
	20	[1.20] ja [2.17]			21 [1.21] jaa [2.18]	
V	25	[1.24] ā [2.22]				
	26	[1.25] iā [2.23]				
RG	27	[1.26] jā [2.24]				
VI	28	[1.27] ə [2.25]	71 [1.68] ʔ [2.61]	90 [1.84] ər [2.76]	32 [1.31] əə	
	29	[1.28] iə [2.26]		91 [1.85] iər		
RG	30	[1.29] jī [2.27]	72 [1.69] jī [2.61]	92 [1.86] jīr [2.77]	33 [1.32] jīī [2.29]	100 [1.92] jīīr [2.85]

TABLE 37.1 (CONTINUED)

RG	<i>Short vowels</i>			<i>Long vowels</i>	
	<i>Lax vowels</i>	<i>Tense vowels</i>	<i>Retroflex vowels</i>	<i>Lax vowels</i>	<i>Retroflex vowels</i>
	31 [1.30] j̥ [2.28]				
	34 [1.33] ej [2.30]	63 [1.60] ej [2.53]	77 [1.73] ejr [2.66]	38 [1.37] eej [2.34]	
VII	35 [1.34] iej [2.31]	63 [1.60] iej [2.53]	78 [2.67] iejr	39 [1.38] ieej	
RG	36 [1.35] j̥j̥ [2.32]	64 [1.61] j̥j̥ [2.54]	79 [1.74] j̥jr [2.68]	40 [1.39] j̥j̥j̥ [2.35]	
	37 [1.36] j̥j̥ [2.33]				
	41 [1.40] əj				
VIII	42 [1.41] iəj [2.36]	76 [2.65] iəj			
RG	43 [1.42] j̥j̥ [2.37]	65 [1.62] j̥j̥ [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]				
RG	46 [1.45] j̥iw [2.40]		94 [1.88] j̥iwr [2.79]	49 [1.47] j̥iww	
	47 [1.46] j̥iw				
	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]	
RG	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
	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	
RG	58 [1.56] jow [2.48]		98 [2.83] jowr	60 [2.50] joow [2.49]	
	104 [1.96] ū				
XII	105 [1.97] jwa				
RG					

## 2.2 Vowels

Tangut has a seven-vowel system consisting of *i i̥ u 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 *ii̥ j̥j̥ uu ee əə*

oo aa. The tense vowels are represented by putting a dot under the vowels, as i̇ ð̇ ɛ̇ ə̇ ɔ̇ ɑ̇ and the retroflex vowels by adding r after the vowels, as ir̄ ɨr̄ ur̄ ər̄ or ar̄. The long retroflex vowels are written as iir̄ ɨir̄ uir̄ ər̄r̄ aar̄. In addition to these vowels there are four nasalized vowels ī ū ē ā, which are found mainly in Chinese loanwords.

### 2.3 Consonants

There are twenty-eight consonants and two semivowels. Stops (except the glottal stop) and affricates contrast in voicing and aspiration. Spirants have a contrast between voicedness and voicelessness. The Tangut consonants are presented in Table 37.2.

TABLE 37.2 INVENTORY OF CONSONANTS

	<i>Bilabial</i>	<i>Dental</i>	<i>Palato-alveolar</i>	<i>Velar</i>	<i>Glottal</i>
Voiceless stop	p	t		k	ʔ (·)
Aspirated stop	ph	th		kh	
Voiced stop	b	d		g	
Voiceless affricate		ts	tʃ (tʃ)		
Aspirated affricate		tsh	tʃh (tʃh)		
Voiced affricate		dz	dʒ (dʒ)		
Nasal	m	n		ŋ	
Voiceless spirant		s	ʃ (ʃ)		h (x)
Voiced spirant		z	ʒ (ʒ)		ɦ (ɣ)
Tap		r			
Lateral		l			
Lateral spirant		ɬ(lh)			
Glide	w		j		

Note:

Symbols given in the parentheses are those used in the reconstruction.

### 2.4 Tones

There are two tones in Tangut, the level tone and the rising tone. The level tone comprises ninety seven rhymes, whereas the rising tone comprises eighty six 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) 𐰇 *bie*<sup>2</sup> to release, to open (vi)  
𐰇 *phie*<sup>2</sup> to release, to open, to untie (vt)
- (2) 𐰇 *bjɑ*<sup>2</sup> end, cut off (vi)  
𐰇 *phja*<sup>1</sup> cut off (vt)
- (3) 𐰇 *dwər*<sup>2</sup> to burn, to blaze (vi)  
𐰇 *thwər*<sup>1</sup> to burn, to cause to blaze (CAUS)
- (4) 𐰇 *dʒjwi*<sup>1</sup> to clean, be pure, melt (vi)  
𐰇 *tshjwi*<sup>1</sup> make clean, to purify, melt (vt)

2.5.2 *Alternation of medials*

The alternation of medials includes the alternation between the presence and absence of medial -j- as well as -w-.

(5)	絳	<i>dzuu<sup>2</sup></i>	to plant, to erect	絳	<i>dzjuu<sup>2</sup></i>	idem
(6)	續	<i>thu<sup>1</sup></i>	to draw (a bow)	續	<i>thju<sup>1</sup></i>	idem
(7)	叢	<i>dzu<sup>1</sup></i>	to love, to like	叢	<i>dzju<sup>1</sup></i>	idem
(8)	癩	<i>lwu<sup>1</sup></i>	to mix	癩	<i>ljwu<sup>1</sup></i>	idem
(9)	綴	<i>jie<sup>1</sup></i>	strength, influence, power	綴	<i>ywie<sup>1</sup></i>	idem
(10)	辦	<i>sji<sup>1</sup></i>	poor, exhausted; to come to an end	辦	<i>sjwi<sup>1</sup></i>	strength, force, power
(11)	覘	<i>sjɿ<sup>1</sup></i>	go, reach, enter	覘	<i>sjwi<sup>1</sup></i>	idem

2.5.3 *Alternation of tones*

As mentioned above, 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)	鞞	<i>ljii<sup>1</sup></i>	trousers (n)	鞞	<i>ljii<sup>2</sup></i>	to put on trousers (v)
(13)	編	<i>tɕiej<sup>1</sup></i>	cap, hat (n)	編	<i>tɕiej<sup>2</sup></i>	wear cap (v)
(14)	鞞	<i>zji<sup>1</sup></i>	leather shoes, boots (n)	鞞	<i>zji<sup>2</sup></i>	to put on shoes (v)
(15)	覘	<i>wɑ<sup>1</sup></i>	the shoulder (n)	覘	<i>wɑ<sup>2</sup></i>	both mean 'to carry on the shoulder' (v)

## b. Level tone representing verbs and rising tone representing nouns

(16)	覘	<i>tji<sup>1</sup></i>	to put, to place (v)	覘	<i>tji<sup>2</sup></i>	a place (n)
(17)	鞞	<i>ɲewr<sup>1</sup></i>	to count (v)	鞞	<i>ɲewr<sup>2</sup></i>	a number (n)
(18)	鞞	<i>dʒjwo<sup>1</sup></i>	to pierce, to bore a hole (v)	鞞	<i>dʒjwo<sup>2</sup></i>	hole (n)
(19)	鞞	<i>sju<sup>1</sup></i>	to hoard, to conceal (v)	鞞	<i>sju<sup>2</sup></i>	a cupboard (n)

## c. Tone alternation not related to syntactic categories

(20)	覘	<i>sjwi<sup>1</sup></i>	who	覘	<i>sjwi<sup>2</sup></i>	idem
(21)	覘	<i>wjij<sup>1</sup></i>	there is/are	覘	<i>wjij<sup>2</sup></i>	idem
(22)	辦	<i>mjiij<sup>1</sup></i>	tail, end, the lower part	辦	<i>mjiij<sup>2</sup></i>	idem
(23)	覘	<i>do<sup>1</sup></i>	poison	覘	<i>do<sup>2</sup></i>	idem
(24)	覘	<i>nɛr<sup>1</sup></i>	yellow	覘	<i>nɛr<sup>2</sup></i>	idem
(25)	覘	<i>ɲwɛr<sup>1</sup></i>	blue	覘	<i>ɲwɛr<sup>2</sup></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 備 *.jij<sup>1</sup>*, instrumental by 覘

*ɲwu*<sup>2</sup>, locative by 悝 *u*<sup>2</sup> ‘in (the garden)’, 漚 *kha*<sup>1</sup> ‘in (the book, the water)’, 翳 *gu*<sup>2</sup> ‘in (the water)’, 翳 *ya*<sup>2</sup> ‘in (the sky, the heart)’, 翳 *tshjaa*<sup>1</sup> ‘on’, 翳 *khju*<sup>1</sup> ‘under’, allative by 翳 *do*<sup>2</sup> ‘to’, comitative by 翳 *rjir*<sup>2</sup> ‘with’, and comparative by 翳 *su*<sup>1</sup>. In addition to this, there is an ergative marker 翳 *džji*<sup>2</sup> *wj1*<sup>1</sup>, indicating the subject of transitive verbs.

### 3.1.2 Pronouns

There is a rich system of pronouns. Plurality can be expressed by the suffix 翳 *nji*<sup>2</sup>. 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 翳 *jij*<sup>1</sup>, three demonstrative pronouns, and several interrogative pronouns.

#### 3.1.2.1 Personal pronouns

person	singular	honorific	plural
	familiar		
1st	翳 <i>ɲa</i> <sup>2</sup>	翳 <i>mjo</i> <sup>2</sup>	翳 翳 <i>ɲa</i> <sup>2</sup> <i>nji</i> <sup>2</sup> 翳 翳 <i>gja</i> <sup>2</sup> <i>mji</i> <sup>2</sup> we [inc] 翳 翳 <i>gji</i> <sup>2</sup> <i>mji</i> <sup>2</sup> we [ex]
2nd	翳 <i>nja</i> <sup>2</sup>	翳 <i>nji</i> <sup>2</sup>	翳 翳 <i>nji</i> <sup>2</sup> <i>nji</i> <sup>2</sup>

#### 3.1.2.2 Reflexive pronouns

翳 *jij*<sup>1</sup> oneself, himself, herself

#### 3.1.2.3 Demonstrative pronouns

翳 <i>thji</i> <sup>2</sup>	this	翳 <i>thja</i> <sup>2</sup>	(third person, irrespective of animate, inanimate, male or female)
翳 <i>thju</i> <sup>2</sup>	here	翳 <i>thja</i> <sup>1</sup>	he, his, him, she, her, it, that
翳 <i>tshju</i> <sup>2</sup>	this, that	翳 <i>thja</i> <sup>2</sup>	there
翳 <i>tšju</i> <sup>1</sup>	that	翳 <i>tji</i> <sup>1</sup>	there
		翳 <i>tshji</i> <sup>1</sup>	that
		翳 <i>tshjiw</i> <sup>1</sup>	that

### 3.1.3 Adjectives

Adjectives occur after the nouns they modify. Many adjectives appear in reduplicated forms.

Examples:

翳 <i>ɲa</i> <sup>2</sup>	翳 <i>ɲa</i> <sup>2</sup>	翳 <i>ljuu</i> <sup>2</sup>	翳 <i>ljuu</i> <sup>2</sup>	翳 <i>ljɿ</i> <sup>1</sup>	翳 <i>mji</i> <sup>1</sup>	翳 <i>džji</i> <sup>1</sup>	翳 <i>lju</i> <sup>1</sup>
good	good	pretty	pretty	though	not	only	flute

翳 <i>məð</i> <sup>1</sup>	翳 <i>rjijr</i> <sup>1</sup>	翳 <i>ljɿ</i> <sup>1</sup>
blow	skilled	AFF

‘(She) is not only very pretty, but also good at playing a flute.’

Comparison is expressed by the postposition 𪛗 *su*<sup>1</sup> 'than' followed by an adjective.

𪛗	𪛗	𪛗	𪛗	𪛗	𪛗	𪛗	𪛗
<i>dzjwo</i> <sup>2</sup>	<i>we</i> <sup>2</sup>	<i>ljwɿ</i> <sup>1</sup>	<i>dzjwow</i> <sup>1</sup>	<i>su</i> <sup>1</sup>	<i>bjuu</i> <sup>1</sup>	<i>tja</i> <sup>1</sup>	<i>tjiɿ</i> <sup>2</sup>
man	be	beast	bird	than	noble	TOP	propriety

𪛗	𪛗	𪛗
<i>dzjiɿ</i> <sup>2</sup>	<i>nioow</i> <sup>1</sup>	<i>ljɿ</i> <sup>1</sup>
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; 1993: 2). The latter are derived from the former by changing the finals into *-jij*.

	<i>Directional markers</i>	<i>Optative markers</i>	
	<i>Perfective markers</i>		
1	𪛗 <i>ja</i> <sup>1</sup>	𪛗 <i>jij</i> <sup>1</sup>	upward
2	𪛗 <i>nja</i> <sup>1</sup>	𪛗 <i>njij</i> <sup>2</sup>	downward
3	𪛗 <i>kji</i> <sup>1</sup>	𪛗 <i>kjij</i> <sup>1</sup>	here, inside
4	𪛗 <i>wji</i> <sup>2</sup>	𪛗 <i>wjij</i> <sup>2</sup>	there, outside
5	𪛗 <i>dji</i> <sup>2</sup>	𪛗 <i>djij</i> <sup>2</sup>	towards the speaker
6	𪛗 <i>dja</i> <sup>2</sup>	𪛗 <i>djij</i> <sup>2</sup>	away from the speaker
7	𪛗 <i>rjiɿ</i> <sup>2</sup>	𪛗 <i>rjijɿ</i> <sup>2</sup>	(direction not found)

Examples:

𪛗	𪛗	𪛗	𪛗	𪛗
<i>mə</i> <sup>1</sup>	<i>dzju</i> <sup>2</sup>	<i>ljiɿ</i> <sup>2</sup>	<i>nja</i> <sup>1</sup>	<i>ljiɿ</i> <sup>2</sup>
sky	rain	big	DIR	come

'It has rained hard.'

𪛗	𪛗	𪛗	𪛗	𪛗
<i>ŋa</i> <sup>2</sup>	<i>jij</i> <sup>1</sup>	<i>dji</i> <sup>2</sup>	<i>khjiɿ</i> <sup>1</sup>	<i>nja</i> <sup>2</sup>
me	DAT	DIR	give	-2sg

'Give (them) to me!'

𪛗	𪛗	𪛗	𪛗	𪛗	𪛗
<i>sjiɿ</i> <sup>1</sup>	<i>ji</i> <sup>2</sup>	<i>zji</i> <sup>1</sup>	<i>mjaa</i> <sup>1</sup>	<i>djiɿ</i>	<i>we</i> <sup>2</sup>
now	again	husband	wife	OPT	become

'Let's be husband and wife again.'

## 3.2.2 Agreement

The Tangut verb shows agreement only with first and second person. Third person is left unmarked (Kepping 1994: 339).

The verb-agreement system also distinguishes between subject agreement and object agreement (Kepping 1976). The subject 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	𐞗 <i>dźij<sup>2</sup></i>	𐞗 <i>dźjo<sup>2</sup></i>	have
27	𐞗 <i>tji<sup>1</sup></i>	𐞗 <i>tjo<sup>1</sup></i>	let drink
28	𐞗 <i>phji<sup>1</sup></i>	𐞗 <i>phjo<sup>2</sup></i>	make, let
29	𐞗 <i>wji<sup>1</sup></i>	𐞗 <i>wjo<sup>2</sup></i>	to do, to make
30	𐞗 <i>rjir<sup>1</sup></i>	𐞗 <i>rjor<sup>1</sup></i>	to get
	𐞗 <i>rjir<sup>2</sup></i>	𐞗 <i>rjor<sup>2</sup></i>	to get
31	𐞗 <i>pju<sup>2</sup></i>	𐞗 <i>pjo<sup>1</sup></i>	to burn, to bake
32	𐞗 <i>lhju<sup>2</sup></i>	𐞗 <i>lhjo<sup>2</sup></i>	to get, to acquire, to obtain

## 3.2.2.1 Third person

When the subject is in the third person, the basic forms of verbs such as 𐞗 *dźij<sup>2</sup>* 'have', 𐞗 *tji<sup>1</sup>* 'let drink', and 𐞗 *rjir<sup>1</sup>* 'get' are used. For example:

𐞗    𐞗    𐞗    𐞗    𐞗  
*jow<sup>1</sup>*   *swē<sup>1</sup>*   *mji<sup>1</sup>*   *gji<sup>2</sup>*   *dźij<sup>2</sup>*  
 Wang Sun daughter a have  
 'Wang Sun has a daughter.'

𐞗    𐞗    𐞗    𐞗    𐞗    𐞗    𐞗  
*buu<sup>2</sup>*   *kow<sup>1</sup>*   *tshji<sup>2</sup>*   *rjar<sup>2</sup>*   *o<sup>2</sup>*   *khjow<sup>1</sup>*   *dja<sup>2</sup>*   *tji<sup>1</sup>*  
 Mu Gong immediately wine bestow PERF let drink  
 'Mu Gong immediately entertained (them) with wine.'

𐞗    𐞗    𐞗    𐞗    𐞗  
*dzji<sup>1</sup>*   *mji<sup>2</sup>*   *dzjw<sup>1</sup>*   *rjir<sup>1</sup>*   *sju<sup>2</sup>*  
 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 subject is in the first or second person, the derived forms of verbs such as 𐞗 *dźjo<sup>2</sup>*, 𐞗 *tjo<sup>1</sup>*, and 𐞗 *rjor<sup>2</sup>* are used.

𐞗    𐞗    𐞗  
*ŋa<sup>2</sup>*   *lj<sup>1</sup>*   *dźjo<sup>2</sup>*   *ŋa<sup>2</sup>*  
 I treasure have -1sg  
 'I have treasure.'

𐞗    𐞗    𐞗    𐞗    𐞗    𐞗    𐞗    𐞗  
*khji<sup>2</sup>*   *dzjwo<sup>2</sup>*   *dja<sup>2</sup>*   *tjo<sup>1</sup>*   *ŋa<sup>2</sup>*   *nji<sup>1</sup>*   *o<sup>1</sup>*   *dja<sup>2</sup>*   *lhji<sup>2</sup>*  
 ten thousand person PERF let drink -1sg near owner PERF lack  
 'I have fed ten thousand people but have neglected my neighbours.'



數	甌	彝	耗	瓊	祥	獲	緇
so <sup>1</sup>	ji <sup>2</sup>	kji <sup>1</sup>	lji <sup>2</sup>	·o <sup>2</sup>	dji <sup>2</sup>	rjo <sup>2</sup>	ŋa <sup>2</sup>
three	hundred	bottle	jar	wine	OPT	get	-1sg

‘Let me get three hundred bottles of wine.’

## 3.2.2.2 Subject agreement

Only when the personal suffixes of the verbs refer to the subject of the sentence, the derived forms of the verbs are used. In the following examples 肥 *nja<sup>2</sup>* refers to the subject of the sentence, therefore the derived forms of the verbs such as 覓 *phjo<sup>2</sup>* and 緝 *wjo<sup>1</sup>* are used.

緇	緝	緝	緝	覓	緝	緝	緝	緝
ljo <sup>2</sup>	tsəj <sup>1</sup>	·jij <sup>1</sup>	dzju <sup>2</sup>	phji <sup>1</sup>	wo <sup>2</sup>	kha <sup>1</sup>	gji <sup>2</sup>	·jij <sup>1</sup>
brother	little	ACC	rule	CAUS	should	while	son	ACC

覓	緝	覓	肥
kji <sup>1</sup>	dzju <sup>2</sup>	phjo <sup>2</sup>	nja <sup>2</sup>
PERF	rule	CAUS	-2sg

‘You have made your son, instead of your younger brother as you should have, a ruler.’

概	發	舖	覓	悞	瓊	榜	覓	緝	肥
nji <sup>2</sup>	śja <sup>1</sup>	bji <sup>2</sup>	·io <sup>1</sup>	·u <sup>2</sup>	dzjo <sup>2</sup>	·ja <sup>2</sup>	dzwə <sup>2</sup>	rji <sup>2</sup>	wjo <sup>1</sup> nja <sup>2</sup>
you[sg,HON]	seven	step	scope	in	poem	one	CL	PERF	make -2sg

‘Write a poem in the time it takes you to make seven steps.’

## 3.2.2.3 Object agreement

In the following examples 緇 *ŋa<sup>2</sup>* and 肥 *nja<sup>2</sup>* both refer to the objects of the sentences, accordingly the basic forms of the verbs such as 覓 *phji<sup>1</sup>* and 緝 *wji<sup>1</sup>* are used.

覓	覓	緝	緝	覓	緇
ko <sup>1</sup>	tshja <sup>1</sup>	wji <sup>2</sup>	dzuu <sup>2</sup>	phji <sup>1</sup>	ŋa <sup>2</sup>
coach	on	PERF	sit	CAUS	-1sg

‘Let me sit on the coach.’

覓	悞	覓	緝	覓	緝	肥
kji <sup>1</sup>	dji <sup>2</sup>	lji <sup>2</sup>	lji <sup>1</sup>	lji <sup>1</sup>	wji <sup>1</sup>	nja <sup>2</sup>
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	覓	nju <sup>2</sup>	to suck a breast (vt)	緝	nju <sup>2</sup>	to suckle (CAUS)
34	緝	thji <sup>1</sup>	to drink (vt)	覓	tji <sup>1</sup>	to give to drink, to feed (CAUS)

35	蕪	<i>gju<sup>2</sup></i>	to wear clothes, to put on clothes (vt)	統	<i>gju<sup>1</sup></i>	to make to wear clothes, to clothe (CAUS)
36	蕪	<i>ɣwej<sup>1</sup></i>	to fight (vi)	繳	<i>ɣwej<sup>1</sup></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	辨	<i>lwu<sup>1</sup></i>	to get mixed (vi)	辨	<i>lwu<sup>1</sup></i>	to mix (vt)
38	蕪	<i>le<sup>2</sup></i>	to boil (vi)	蕪	<i>le<sup>1</sup></i>	to boil (vt)

Nouns and adjectives are changed into verbs or causative verbs by the same mechanism.

39	繳	<i>nju<sup>2</sup></i>	tame, mild (adj)	蕪	<i>nju<sup>2</sup></i>	to tame (vt)
40	繳	<i>bji<sup>1</sup></i>	low, below, down (adj)	蕪	<i>bji<sup>1</sup></i>	to lower, to bend
41	尾	<i>bji<sup>2</sup></i>	low, below, down (adj)	尾	<i>bji<sup>1</sup></i>	to heighten, to elevate, to promote (v)
42	批	<i>bji<sup>2</sup></i>	high (n, adj)	批	<i>bji<sup>1</sup></i>	to heighten, to elevate, to promote (v)
43	批	<i>khwa<sup>1</sup></i>	far, distant (adj, adv)	批	<i>khwa<sup>1</sup></i>	keep away from, keep at distance (vt)
44	批	<i>lɔ<sup>1</sup></i>	a pit, a gully, a ditch (n)	批	<i>lɔ<sup>1</sup></i>	to throw into a pit, to bury
44	蕪	<i>mji<sup>2</sup></i>	traces, footprints (n)	蕪	<i>mji<sup>2</sup></i>	to trace, to trace source (v)

Nouns are derived from verbs, as in:

45	蕪	<i>thji<sup>1</sup></i>	to drink (vt)	蕪	<i>tji<sup>1</sup></i>	food (n)
46	蕪	<i>dʒju<sup>1</sup></i>	to deceive, to cheat (vi,vt)	蕪	<i>dʒju<sup>2</sup></i>	a bait, an allurement (n)
47	蕪	<i>sji<sup>2</sup></i>	to know (v), knowledge (n)	蕪	<i>sji<sup>2</sup></i>	wisdom (n)

## 4.2 Compounding

### 4.2.1 Reduplicative compound

Reduplication is an important device for Tangut word-formation.

致	致	<i>ljur<sup>2</sup> ljur<sup>2</sup></i>	(beautiful beautiful) beautiful
殺	殺	<i>lji<sup>1</sup> lji<sup>1</sup></i>	(one one) one by one, in every detail
廩	廩	<i>ɣowr<sup>2</sup> ɣowr<sup>2</sup></i>	(all all) all, every
處	處	<i>rjur<sup>1</sup> rjur<sup>1</sup></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.

迺	蕪	<i>ɣwə<sup>1</sup> ɣwej<sup>1</sup></i>	(fight fight) to fight
秀	蕪	<i>sə<sup>2</sup> sej<sup>1</sup></i>	(clean clean) clean

投	亂	<i>džiə² dźiej²</i>	(revolve revolve) to revolve; transmigration
報	龍	<i>siə¹ siej¹</i>	(lead lead) to lead
踴	緒	<i>bji¹ bji²</i>	(low low) the lower part
踴	妮	<i>mji² mji²</i>	(silent silent) silent(ly)
踴	尾	<i>bji¹ bji²</i>	(high high) the upper part
踴	隄	<i>sjie² sjij²</i>	(know know) acquaintance
踴	隄	<i>lji¹ ljij²</i>	(midday midday) midday
踴	滌	<i>sjie¹ sjii¹</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.

鞦	鞦	<i>giee¹-gioo¹</i>	silly, foolish
鞦	鞦	<i>šiee¹-sioo¹</i>	to collect
鞦	鞦	<i>wjii¹-wjoo¹</i>	to exchange, to trade
鞦	鞦	<i>djii¹-djoo¹</i>	to divide, to share
鞦	鞦	<i>gjii¹-gjoo¹</i>	to bite, to peck
鞦	鞦	<i>no²-nej²</i>	tranquillity, stability, safety, peace
鞦	鞦	<i>ror²-rejr²</i>	to surround
鞦	鞦	<i>kio²-kie²</i>	hate, dislike
鞦	鞦	<i>sjwo²-sjwij¹</i>	to grind, to whet
鞦	鞦	<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

鞦	鞦	<i>tji¹ ηa²</i>	(food good)	good food
鞦	鞦	<i>lji¹ ljij²</i>	(wind big)	strong wind
鞦	鞦	<i>dji² rji²</i>	(doctor good)	good (medical) doctor
鞦	鞦	<i>we¹ nər²</i>	(bird yellow)	yellow bird

#### b. Noun + noun

鞦	鞦	<i>məə¹ :joow²</i>	(fire stove)	brazier
鞦	鞦	<i>tsur¹ tsiej¹</i>	(winter hat)	winter hat
鞦	鞦	<i>sjie² dzjwo²</i>	(female person)	woman
鞦	鞦	<i>dźji¹ gjwi¹</i>	(leather clothing)	leather jacket

## c. Noun + adjective/noun + numeral

𐰇	𐰇	𐰇	<i>rjijr<sup>1</sup> ηa<sup>2</sup> gjɨ<sup>2</sup></i>	(horse good one) a good horse
𐰇	𐰇	𐰇	<i>sjjɨ<sup>2</sup> dzjwo<sup>2</sup> gjɨ<sup>2</sup></i>	(female person one) a woman

## d. Noun + adjective + noun

𐰇	𐰇	𐰇	<i>rjijr<sup>1</sup> ηa<sup>2</sup> tshjɨ<sup>1</sup></i>	(horse good meat) meat of a good horse
---	---	---	---	--

## e. Noun + adjective + adverb

𐰇	𐰇	𐰇	<i>dji<sup>2</sup> rjijr<sup>1</sup> yiej<sup>1</sup></i>	(doctor good really) really good doctor
---	---	---	---	---

f. Verb + auxiliary (Here the auxiliary verb 𐰇 *lew<sup>2</sup>* functions as a nominalizer.)

𐰇	𐰇	<i>thwə<sup>1</sup> lew<sup>2</sup></i>	(burn should)	fuel
𐰇	𐰇	<i>gjwi<sup>2</sup> lew<sup>2</sup></i>	(wear should)	clothing
𐰇	𐰇	<i>zji<sup>2</sup> lew<sup>2</sup></i>	(wear-shoes should)	shoes
𐰇	𐰇	<i>phji<sup>1</sup> lew<sup>2</sup></i>	(send should)	servant, slave

## g. Noun + numeral + classifier

Classifiers occur after numerals.

𐰇	𐰇	𐰇	<i>dzjo<sup>2</sup> ja<sup>2</sup> dzwə<sup>2</sup></i>	(poem one CL) a poem
𐰇	𐰇	𐰇	<i>sow<sup>1</sup> ja<sup>2</sup> phu<sup>2</sup></i>	(mulberry-tree one CL) a mulberry-tree
𐰇	𐰇	𐰇	<i>da<sup>2</sup> ja<sup>2</sup> gjwi<sup>2</sup></i>	(word one CL) a word
𐰇	𐰇	𐰇	<i>kja<sup>2</sup> ja<sup>2</sup> kho<sup>1</sup></i>	(song one CL) a song

## h. Pronoun + noun

𐰇	𐰇	<i>ηa<sup>2</sup> wə<sup>1</sup></i>	(my husband) my husband
𐰇	𐰇	<i>mjo<sup>2</sup> ηwuu<sup>1</sup></i>	(my word) my words
𐰇	𐰇	<i>thjɨ<sup>2</sup> da<sup>2</sup></i>	(this word) these words
𐰇	𐰇	<i>nja<sup>2</sup> wjij<sup>2</sup></i>	(your[sg] uncle) your (paternal) uncle

## i. Noun/pronoun + genitive marker + noun

𐰇	𐰇	𐰇	𐰇	𐰇	<i>thej<sup>1</sup>-pie<sup>1</sup> ·jij<sup>1</sup> gjɨ<sup>2</sup> lhjɨ<sup>1</sup></i>	(proper noun POSS sons grandsons) 'the descendants of Thej-pie',
𐰇	𐰇	𐰇			<i>ηa<sup>2</sup> ·jij<sup>1</sup> la<sup>1</sup></i>	(I POSS hand) 'my hand'

## 5.2 Structure of the clause

## 5.2.1 Word order

The basic pattern of the Tangut clause is SOV. The verb can be followed by an auxiliary verb.

𐰇	𐰇	𐰇	𐰇	𐰇	𐰇	𐰇	𐰇	𐰇	𐰇	
<i>xu<sup>1</sup></i>	<i>·jow<sup>1</sup></i>	<i>·jɨr<sup>2</sup></i>	<i>da<sup>2</sup></i>	<i>phja<sup>1</sup></i>	<i>dzjɨ<sup>2</sup></i>	<i>rjijr<sup>1</sup></i>	<i>mji<sup>1</sup></i>	<i>dzjij<sup>1</sup></i>	<i>niow<sup>1</sup></i>	<i>mjiij<sup>1</sup></i>
Fu	Rong	learn	thing	cut	judge	good	NEG	only	also	dream
	𐰇	𐰇								
	<i>seew<sup>2</sup></i>	<i>rjijr<sup>1</sup></i>								
	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.

𗇑	𗇒	𗇓	𗇔	𗇕	𗇖	𗇗	𗇘	𗇙
<i>dzjwɿ<sup>1</sup></i>	<i>ljij<sup>2</sup></i>	<i>tshja<sup>1</sup></i>	<i>zjɿ<sup>1</sup></i>	<i>tšier<sup>1</sup></i>	<i>dzjwo<sup>2</sup></i>	<i>zeew<sup>2</sup></i>	<i>pju<sup>2</sup></i>	<i>rjɿr<sup>2</sup></i>
emperor	greatly	angry	left	right	person	send	palace	hall

𗇚	𗇛	𗇜	𗇝	𗇞	𗇟	𗇠	𗇡
<i>khju<sup>1</sup></i>	<i>tshjij<sup>1</sup></i>	<i>phji<sup>1</sup></i>	<i>rjɿr<sup>2</sup></i>	<i>.u<sup>2</sup></i>	<i>njij<sup>2</sup></i>	<i>ɣe<sup>1</sup></i>	<i>ji<sup>2</sup></i>
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: 𗇑 *mji<sup>1</sup>*, 𗇒 *mji<sup>1</sup>*, 𗇓 *mji<sup>1</sup>*, 𗇔 *mji<sup>2</sup>*, 𗇕 *tji<sup>1</sup>* (expressing prohibition), and 𗇖 *njaa<sup>2</sup>*.

The negative markers generally precede the verb they negate. The exceptions are 𗇓 *mji<sup>1</sup>* and 𗇖 *njaa<sup>2</sup>*, which occur in the sentence-final position.

The most general form of negation is 𗇑 *mji<sup>1</sup>*, as in:

𗇑	𗇒	𗇓	𗇔
<i>da<sup>2</sup></i>	<i>mjor<sup>1</sup></i>	<i>mji<sup>1</sup></i>	<i>tshjij<sup>1</sup></i>
word	true	NEG	say

‘(S/He) does not tell the truth.’

𗇕	𗇖	𗇗	𗇘	𗇙	𗇚	𗇛	
<i>dzjij<sup>2</sup></i>	<i>rjur<sup>1</sup></i>	<i>zjɿ<sup>1</sup></i>	<i>lji<sup>2</sup></i>	<i>zji<sup>2</sup></i>	<i>dzjij<sup>1</sup></i>	<i>mji<sup>1</sup></i>	<i>dwewr<sup>2</sup></i>
other	all	children	all	coldness	not	feel	

‘All the other children did not feel cold.’

𗇒 *mji<sup>1</sup>* is used to negate auxiliary verbs, such as 𗇓 *kjir<sup>2</sup>* ‘dare’, 𗇔 *njwi<sup>2</sup>* ‘can’, 𗇕 *ljii<sup>1</sup>* ‘possible’, and 𗇖 *dzjij* ‘willing’ etc. For example:

𗇑	𗇒	𗇓	𗇔	𗇕
<i>tji<sup>2</sup></i>	<i>phji<sup>1</sup></i>	<i>mji<sup>1</sup></i>	<i>kjir<sup>2</sup></i>	<i>ɳa<sup>2</sup></i>
decency	discard	NEG	dare	-1sg

‘I did not dare to offend decency.’

𗇖	𗇗	𗇘	𗇙	𗇚	𗇛	𗇜	𗇝	𗇞		
<i>thjij<sup>2</sup></i>	<i>sjo<sup>2</sup></i>	<i>thjɿ<sup>2</sup></i>	<i>dzjwo<sup>2</sup></i>	<i>tji<sup>1</sup></i>	<i>jij<sup>1</sup></i>	<i>rjur<sup>1</sup></i>	<i>ar<sup>2</sup></i>	<i>mji<sup>1</sup></i>	<i>njwi<sup>2</sup></i>	<i>ji<sup>2</sup></i>
why	this	person	alone	ACC	restrain	NEG	can	COMP		

‘Why can’t you restrain this person alone?’

級 級 級  
*dza*<sup>2</sup> *mji*<sup>1</sup> *lji*<sup>1</sup>  
 estimate not possible

'It is not possible to estimate it.'

難 難 難  
*sj*<sup>1</sup> *mji*<sup>1</sup> *dzij*  
 go NEG willing

'S/He is not willing to go.'

Finely differentiated existential verbs such as 有 *dju*<sup>1</sup>, 無 *tshju*<sup>1</sup>, 度 *wji*<sup>2</sup>, and 難 *dzij*<sup>2</sup> are all negated by 無 *mji*<sup>1</sup>, which normally occurs after nouns.

有 *dju*<sup>1</sup> 'there is' and 無 *mji*<sup>1</sup> 'not exist' are put in contrast as in:

這 那 懼 織 織 韋 慨 辭 鏡 有 難 終  
*thj*<sup>2</sup> *we*<sup>2</sup> *u*<sup>2</sup> *tja*<sup>1</sup> *yu*<sup>1</sup> *khjwi*<sup>1</sup> *mji*<sup>1</sup> *le*<sup>2</sup> *bjuu*<sup>2</sup> *dju*<sup>1</sup> *mji*<sup>1</sup> *dji*<sup>2</sup>  
 this city in TOP head cut NEG fear general there is but  
 殺 鏡 織 無  
*yio*<sup>1</sup> *bjuu*<sup>2</sup> *tja*<sup>1</sup> *mji*<sup>1</sup>  
 surrender general TOP not exist

'In this city there are generals who are not afraid of being beheaded, but no generals who will surrender.'

有 *tshju*<sup>1</sup> 'have' and 無 *mji*<sup>1</sup> 'not have' are put in contrast in:

殺 有 殺 無  
*sjj*<sup>2</sup> *tshju*<sup>1</sup> *sjj*<sup>2</sup> *mji*<sup>1</sup>  
 wisdom have wisdom not have

'No matter whether (he) has wisdom or does not have wisdom.'

The contrast between 度 *wji*<sup>2</sup> 'there is' and 無 *mji*<sup>1</sup> 'not exist' are observed in the following pair of sentences.

賤 難 散 度  
*khjow*<sup>1</sup> *tji*<sup>2</sup> *ljo*<sup>2</sup> *wji*<sup>2</sup>  
 give place where there is  
 'How is it possible to give (her) away.'

有 變 禍 難 無  
*lo* *we*<sup>2</sup> *phji*<sup>1</sup> *tji*<sup>2</sup> *mji*<sup>1</sup>  
 rich become make place not exist  
 'It is not possible to make (him) rich.'

有 *dzij*<sup>2</sup> 'have' and 無 *mji*<sup>1</sup> 'not have' are put in contrast in the following sentences.

殺 無 殺 有 有  
*zji*<sup>1</sup> *mji*<sup>1</sup> *so*<sup>1</sup> *mji*<sup>1</sup> *dzij*<sup>2</sup>  
 son not have three daughter have  
 '(He) has three daughters but no sons.'

褲 到 幾 絹 繼 低 籠 織  
*njwo<sup>2</sup> lew<sup>1</sup> ljii<sup>1</sup> mjij<sup>1</sup> sjij<sup>1</sup> ŋwə<sup>1</sup> lhwu<sup>1</sup> dʒji<sup>2</sup>*  
 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 織 *dʒji<sup>2</sup>* is negated by 慨 *mji<sup>1</sup>*.

繡 賢 翻 銀 鞣 疑 織 慨 織  
*ɣjwə<sup>1</sup> xia<sup>1</sup> lu<sup>2</sup> niow<sup>1</sup> war<sup>2</sup> gjiw<sup>2</sup> lew<sup>2</sup> mji<sup>1</sup> dʒji<sup>2</sup>*  
 Ruan Xian poor therefore property clothes not have

'Ruan Xian is so poor that he has no property and clothes.'

絹 *mji<sup>1</sup>* can also be used as a negative answer to a Yes-no question.

敵 絹 絹 敵 鞣 鞣 肥  
*thji<sup>2</sup> sjo<sup>2</sup> mjij<sup>1</sup> lji<sup>1</sup> rji<sup>2</sup> tshji<sup>2</sup> nja<sup>2</sup>*  
 why not exist AFF PERF say -2sg

'Why did you say "no"?''

鞣 *mji<sup>2</sup>* is used in a negative sentence in reference to future.

鞣 鞣 鞣 鞣 鞣 鞣  
*thja<sup>1</sup> dzwo<sup>2</sup> rjur<sup>1</sup> kha<sup>1</sup> mjij<sup>2</sup> wee<sup>1</sup>*  
 that person world in NEG born

'That person has not been born yet in the world.'

The equational verb 鞣 *ŋwu<sup>2</sup>* is negated by 悞 *njaa<sup>2</sup>*. These two words are put in contrast in the following sentence.

繼 鞣 鞣 鞣 悞 鞣 鞣 鞣 鞣 鞣  
*sjij<sup>1</sup> nja<sup>2</sup> nə<sup>2</sup> dʒiə<sup>1</sup> njaa<sup>2</sup> ku<sup>1</sup> sɪwə<sup>1</sup> wə<sup>1</sup> ŋwu<sup>2</sup> nja<sup>2</sup> lji<sup>1</sup>*  
 now you[sg] fox NEG then rat be -2sg AFF

'You are either a fox or a rat.'

悞 *njaa<sup>2</sup>* can occur alone without a verb as a negative answer to a question. In that case it is equivalent to English 'no!'

鞣 鞣 鞣 悞 鞣 鞣 鞣 鞣 鞣 鞣  
*xjiw<sup>1</sup> tshjow<sup>1</sup> da<sup>2</sup> njaa<sup>2</sup> ŋa<sup>2</sup> ŋa<sup>2</sup> wji<sup>2</sup> sjii<sup>2</sup> nja<sup>2</sup> ji<sup>2</sup>*  
 Fei Zhong say no good good PERF think -2sg COMP

'Fei Zhong said, "No! Please think about it".'

鞣 鞣 鞣 悞  
*jiv<sup>2</sup> sjwo<sup>1</sup> lew<sup>2</sup> njaa<sup>2</sup>*  
 doubt raise should NEG

'Don't have a doubt!'

鞣 *tji<sup>1</sup>* 'don't' is the negative marker used in imperatives.

鞣 鞣 鞣  
*tji<sup>1</sup> khji<sup>1</sup> nja<sup>2</sup>*  
 NEG give -2sg

'Don't give (it to them)!'

玆 織 純 肥  
*tʃi<sup>1</sup> sjwi<sup>1</sup> sjij<sup>2</sup> nja<sup>2</sup>*  
 NEG worry -2sg  
 'Don't worry!'

懼 覓 玆 禿  
*ʃiw<sup>2</sup> ljij<sup>1</sup> tʃi<sup>1</sup> sjwo<sup>1</sup>*  
 疑 惑 勿 生  
 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 覓 *mo<sup>2</sup>* in sentence-final position signals the construction is a question, as in:

覓 覓 覓 覓 覓 覓  
*nʃi<sup>2</sup> kã<sup>1</sup> tʃja<sup>1</sup> wʃi<sup>2</sup> dzjo<sup>1</sup> mo<sup>2</sup>*  
 you[sg,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.

覓	<i>sjwi<sup>1</sup></i>	who
覓	<i>sjwi<sup>2</sup></i>	who
覓	<i>wã<sup>2</sup></i>	what
覓	<i>ljɔ<sup>2</sup></i>	where
覓	<i>ljij<sup>1</sup></i>	which
覓	<i>thij<sup>2</sup> sjo<sup>2</sup></i>	why, how
覓	<i>wã<sup>2</sup> zji<sup>1</sup></i>	how many, how much
覓	<i>zji<sup>1</sup> mɔ<sup>2</sup></i>	how many kinds

In general, question words occur in the same position in the sentence as do non-question words that have the same grammatical function. Examples are:

覓 覓 覓 覓 覓  
*ʃji<sup>1</sup> wɔ<sup>1</sup> gi<sup>2</sup> tja<sup>1</sup> sjwi<sup>1</sup>*  
 former husband TOP who  
 'Who is your former husband?'

覓 覓 覓 覓 覓  
*thji<sup>2</sup> lhwu<sup>1</sup> sjwi<sup>1</sup> jij<sup>1</sup> ŋwu<sup>2</sup>*  
 this clothes who POSS be  
 'Whose clothes are these?'

覓 覓 覓 覓  
*thij<sup>2</sup> sjo<sup>2</sup> nwɔ<sup>1</sup> nja<sup>2</sup>*  
 how know -2sg  
 'How do you know it?'



儼 敗 級 狡 殼 羸  
*ŋwə<sup>1</sup> bji<sup>1</sup> nja<sup>1</sup> nji<sup>2</sup> ljɔ<sup>2</sup> dju<sup>1</sup>*  
 five brightness divine pearl where be  
 ‘Where is the divine pearl of five planets?’

微 猶 數 樞 朧 翫 殼 緝  
*sju<sup>2</sup> po<sup>1</sup> tji<sup>1</sup> nji<sup>2</sup> phji<sup>1</sup> ya<sup>2</sup> thji<sup>2</sup> sjo<sup>2</sup>*  
 Su bhū ti you[sg,HON] mind in how  
 ‘How does it appear to thee, Subhūti?’

媼 孺 緬 猷 羸 樞 稽  
*sjɪ<sup>2</sup> .ji<sup>1</sup> lji<sup>1</sup> tshja<sup>2</sup> dju<sup>1</sup> nji<sup>2</sup> zji<sup>1</sup>*  
 woman GEN four virtue there are you[sg,HON] how many

殼 覆 貳  
*mə<sup>2</sup> dzjo<sup>2</sup> nja<sup>2</sup>*  
 kind have -2sg

‘There are four virtues of women. How many do you have of them?’

### 5.3.3 Subordinate clauses

The adjective clause is placed before the noun it modifies.

緝 韋 慨 憚 篋  
*yu<sup>1</sup> khjwi<sup>1</sup> mji<sup>1</sup> le<sup>2</sup> bjuu<sup>2</sup>*  
 head cut NEG fear general  
 ‘The generals who are not afraid of being beheaded.’

#### 5.3.3.1 Adverb clause

The conditional meaning is expressed by means of the conditional conjunction 羸 *tji<sup>1</sup>*.

樞 羸 慨 蕪 貳 鑄 錐 殼 務 貳  
*nji<sup>2</sup> tji<sup>1</sup> mji<sup>1</sup> sji<sup>1</sup> nja<sup>2</sup> ku<sup>1</sup> ka<sup>1</sup> phja<sup>1</sup> wji<sup>1</sup> nja<sup>2</sup>*  
 you[sg,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 樞 *bju<sup>1</sup>*.

媼 駁 答 駁 駁 羸 殼 訊 緝 憚  
*sjɪ<sup>2</sup> dzjwo<sup>2</sup> ku<sup>2</sup> da<sup>2</sup> sji<sup>1</sup> thji<sup>2</sup> gji<sup>2</sup> kju<sup>1</sup> kjɪr<sup>2</sup> u<sup>2</sup>*  
 woman person answer say now this child carriage room in  
 緝 駁 羸 駁 錐 訊 駁 駁  
*wji<sup>2</sup> wee<sup>1</sup> bju<sup>1</sup> tshjwo<sup>1</sup> mjij<sup>2</sup> kju<sup>1</sup> tsə<sup>1</sup> .ji<sup>2</sup>*  
 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 駁 *lji<sup>1</sup>*.

樞 駁 駁 孺 駁 駁 駁 駁 駁 駁  
*nji<sup>2</sup> nji<sup>2</sup> tsjiw<sup>1</sup> .ji<sup>1</sup> sji<sup>2</sup> mji<sup>1</sup> dzji<sup>1</sup> lji<sup>1</sup> .ji<sup>2</sup> djij<sup>2</sup>*  
 you [pl,HON] Zhou POSS grain NEG eat although COMP but

發	菜	通	逐	發	蔬	蔬	蔬	蔬
tʃjw <sup>1</sup>	ɲər <sup>1</sup>	lwu <sup>2</sup>	dʒjij <sup>1</sup>	tʃjw <sup>1</sup>	naa <sup>2</sup>	dʒjij <sup>1</sup>	dʒji <sup>1</sup>	tja <sup>1</sup>
Zhou	Mountain	hidden	dwel	Zhou	vegetable	only	eat	TOP
豉	蒜	勁						
thjij <sup>2</sup>	ɣiej <sup>1</sup>	ji <sup>2</sup>						
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?'

## REFERENCES

- 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: 935–68.
- (1994) 'A hypothesis of three grades and vowel length distinction in Tangut', *Journal of Asian and African Studies* 46–47: 305–14.
- (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: 265–89.
- (2001) 'Personal agreements and phonological alternations in the Tangut verb', *Language and Linguistics* 2.1: 21–67.
- Kepping, K.B. (1976) 'Subject and object agreement in the Tangut verb', *Linguistics of the Tibeto-Burman Area* 2.2: 219–31.
- (1982) 'Once again on the agreement of the Tangut verb', *Linguistics of the Tibeto-Burman Area* 7.1: 39–54.
- (1983) «Les Kategorij» – Utračennaja 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'.
- (1994) 'The conjugation of the Tangut verb', *Bulletin of the School of Oriental and African Studies* 57.2: 339–46.
- 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.
- Neveskij, N.A. (1960) *Tangutskaja filologija. Issledovanija i slovar'* (in two vols), Moscow: Izdatel'stvo Vostočnoj Literatury.
- 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 & 2, Tōkyō: 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.
- Shi, Jinbo, Huang, Zhenhua, and Nie, Hongyin (1993) *Leilin Yanjiu* (*A study in the Forest of categories*), Yinchuan: Ningxia Renmin Chubanshe.

- Sofronov, M.V. (1968) *Grammatika tanguskogo jazyka* (in two vols), Moscow: Izdatel'stvo 'Nauka'.
- Sun, Hongkai (1981) *Qiangyu jianzhi* (*A concise description of the Qiang language*), Beijing: Nationality Press.
- van Driem, George (1991) 'Tangut verbal agreement and the patient category in Tibeto-Burman', *Bulletin of the School of Oriental and African Studies* 54.3: 520–34.

**PART 12**

---

**KAREN LANGUAGES**

---



# EASTERN KAYAH LI

*David Solnit*

## 1 INTRODUCTION

Eastern Kayah Li is spoken in the area of Kayah State (Karenni) 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 /kəjē li phú/ (for the people; the language is /kəjē li ŋò/). *kəjē* is ‘person’, ‘Kayah’ and *li* is ‘red’. An older term is Karenni, which is anglicized Burmese; Burmese <ni> is also ‘red’. ‘Karenni’ 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. Not enough is known about the Padaung languages to fit them into one of these groupings.

The Kayah Li are numerically the dominant group in the Central Karen area, with perhaps 100,000 speakers, all in Burma except for some 1500 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 (Karenni) 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 alone in Tibeto-Burman, 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.

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 (c and t are separated by 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 twenty 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 e ɤ o ε ʌ ɔ 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 [ɨ].
- 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 *du* 'cut', *du* 'slice', *du* 'at', *du* '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 [ɕ] and [ɔ̃] respectively. /wa/ occurs with all initials except /v/. /wi/ occurs with all simple initials except /v t n h Ø/. /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: 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 h/, 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 [ŋ<sup>j</sup>~ɲ] (fronted velar or palatal nasal) 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 /p/ 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 whose vowel copies the preceding full syllable and whose tone pre-empts that of the full syllable. In the Upper (*kè khu*) dialect of Eastern Kayah this suffix retains its own vowel quality; compare Lower *chwa khri*<sup>52</sup>, Upper *chwa khreu*<sup>52</sup> ‘crab’; Lower *təple*<sup>52</sup>, Upper *təpleu*<sup>52</sup> ‘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 merges completely with 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: ʔithu ‘post’, ʔilò ‘to plant (seeds)’ (conditioned high before low-level and low-falling); ʔikhré ‘to winnow’ (conditioned low before high); ʔilū ‘the Kayah New Year festival’ (lexical low before mid); ʔivī ‘to whistle’ (lexical high before mid).

Something similar can be seen in *khɛ~khɛ* ‘leg’. Low-level *khɛ* occurs before mid tone in compounds like *khɛ mā* ‘knee’ (*mā* ‘joint’) and *khɛ rē* ‘paw’. Mid tone *khɛ* occurs before low-level, as in *khɛ do* ‘lower leg’ and *khɛ le* ‘foot’. Low-level *khɛ* 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 below) can be analysed as made up partly of verbal compounds.



Prefixes are of two types: *ʔ*, 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 *ʔi* in names of instruments: *sé ʔiché* ‘sewing machine’ (*sé* ‘large machine’, *ché* ‘sew’), *ʔithá* a plough (*thá* ‘to’ plough). Examples in which the prefix has no identifiable function are *ʔise* ‘salt’, *ʔikhré* ‘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’). Examples are *kəjē li phú ~ kəjē liú* ‘the (red) Kayah’ (*kəjē* ‘person’, ‘Kayah’, *li* ‘red’). The suffix has the meaning ‘member of a class’, ‘instance of a category’. Examples are *kláú ~ kláphú* ‘soldier’ (*klá* ‘army’) and *tēú* ‘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) *vē ma ʔe ká phé thé ja ja*  
 I 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: *sínɛ 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 ʔo* ‘fruit+exist: useful, meaningful’.

For verb-verb compounds, see 4.2.2 ‘Serial verbs: the verb complex’ below.

## 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 dipɔ du* ‘I+pot+big: my big pot’ or *phremə hA 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, *kəjɛ̄ [pā tha]* 'people to cut sesame (plants), sesame-choppers' consists of a head noun *kəjɛ̄* '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 *dīpɔ tɔthé* 'pot+iron: iron pot' and *bēʔū təphé* 'cloth+cotton: cotton cloth'. Second, many names of plant and animal species begin with a general term for the kind, as *thu tò* 'drongo (*Dicrurus* spp.)' (*thu* 'bird') or *sɔ lehλ* 'teak tree' (*sɔ* 'tree'). Third, ethnic designations follow the head in the meaning 'X-style', as in *hλca phrè* 'clothes+Shan: Shan-style clothes'.

Classifier-headed NPs have the structure (Dem) (s) Clf'. Dem is a Demonstrative, either *ʔā* 'this' or *nλ* 'that'. S is a clause, in this function known as a preposed attributive clause (see below). 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?'.

Localizer-headed NPs consist of a localizer obligatorily modified by a preceding ordinary-headed NP, the resulting NP in turn serving as object of a preposition. An example is *dʂ khrā kū* 'inside the bottle-gourd', consisting of the Preposition *dʂ* followed by an NP made up of the localizer *kū* 'inside-part' modified by *khrā* 'bottle-gourd'.

## 4.2 Structure of the clause

### 4.2.1 Word order and clause structure

Clause structure is as follows:

[NP<sub>1</sub> [[VC NP<sub>2</sub> NP<sub>3</sub> PP<sub>1</sub>]V' PP<sub>2</sub> Clf']VP]s SPTc

VC=verb complex, PP=prepositional phrase, Clf'=quantifier+classifier, SPTc=sentence particle. Grammatical relations are NP<sub>1</sub>=subject, NP<sub>2</sub>=indirect object, and NP<sub>3</sub>=direct object. The remaining phrasal categories represent various oblique relations, with the one represented by the clause-final Clf' termed Extent.

Example of indirect and direct objects:

- (2) *sárá ʔiswá phúcè li*  
 teacher teach child writing  
 'The teacher teaches the children (their letters).'

Example of PP<sub>1</sub> (inside of V') and PP<sub>2</sub> (outside of V'):

- (3) *ʔa phā nλ ʔo [dʂ hi thū]<sub>1</sub> [hú pe ʔo ʔā]<sub>2</sub> nλ*  
 he grandmother PTC exist at house edge like we exist this PTC  
 'His grandmother lived at the edge of the village, as we do here.'

Example of extent expression, with alternate classifiers:

- (4) *phúcè cwá dʂ hóhé 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 ' $V_2$  from having  $V_1$ -ed'. When  $V_1$  is  $\text{ʔo}$  'exist', this construction is the most usual way to express the notion source:

- (5) *ʔo dʂ sɔ̄ kū tā the nɔ̄ 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 will illustrate the major properties of the verb complex (the bracketed string):

- (6) *ʔa [vī jo cwá] chā m̀ bē nA rA*  
 3 throw fly go chicken female yellow Ptc Ptc  
 '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 (*ʔa* 'she/he') and a single (Direct) Object (*chām̀ bē* 'yellow hen').
- 2 The semantic arguments (participant roles, theta-roles) of the VCs constituent verbs map onto the grammatical relations specified by the VC in a complex but partially predictable way. In this example the Subject *ʔa* has the actor role of *vī* 'throw' (it is the one that does the throwing). The Direct Object *chām̀ 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 patient 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.

Example of sequence:

- (7) *ʔa ka dēhā rá ʔaphā*  
 s/he return ask PTC grandmother  
 'He went home and asked his grandmother.'

The actions of *ka* 'go towards home' and *dēhā* 'ask' take place in temporal sequence.

Example of modification:

- (8) *ké r̀ chwa*  
 AMB cold strong  
 'It's very cold.'

The predicate *r̀* '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. Example of a verbal sentence:

- (9) *nō nō hóhé to ɛ ɔ*  
 enter at.all school NEG QUES HUH  
 ‘Aren’t you going to school, hey?’

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, síʔichē ʔa ʔó he*  
 want put dry afraid it mildew  
 ‘I want to put it out to dry; I’m afraid it may mildew.’

*pó~pō* urging, ‘why don’t you’:

- (12) *ne thō tēmē ʔa dā jū vē pō*  
 you drum one-CLF it give easy me  
 ‘Your drum, won’t you give it to me?’

*mē* mild counter-assertion, also used in answers to questions:

- (13) *ʔa me sā á lū mē*  
 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) above. The other is the negative imperative *me*.

## 4.3.3 Questions

Yes-no questions are signalled by the final particle  $\bar{e}$  (this is one of the few Kayah Li lexical items with zero initial). Question-word (Wh-) questions use  $\text{ʔit}\bar{e}$  ‘what’ and a set of morphemes incorporating its abbreviated form  $t\bar{e}$ :  $me\ t\bar{e}$  ‘why’,  $t\bar{o}\ \acute{u}\ t\bar{e}$  ‘where (general)’,  $b\acute{s}\ t\bar{e}$  ‘where (nearby)’,  $b\acute{s}\ k\bar{e}\ t\bar{e}$  ‘when’,  $b\acute{a}$  CLF  $t\bar{e}$  ‘how many’,  $h\acute{u}\ t\bar{e}$  ‘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,  $d\acute{s}\ t\bar{e}\ t\bar{o}$ -CLF...  $\bar{e}$  ‘which one’ and  $\text{ʔ}\bar{u}$ ...  $p\bar{e}$  ‘who’. An example of the former is

- (14)  $d\acute{s}\ t\bar{e}\ t\bar{o}\ c\acute{x}\ \text{ʔ}a\ v\bar{i}\ l\acute{s}\ \bar{e}$   
 at what one-CLF it 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 below), as in (clause in square brackets):

- (15)  $\text{ʔ}a\ c\acute{w}a\ n\bar{i}h\bar{o}\ k\bar{\alpha}\ d\acute{s}\ [k\bar{o}\acute{z}\bar{e}\ k\acute{e}\ t\bar{o}h\bar{\alpha}] n\Lambda$   
 they 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)  $\text{ʔ}a\ d\acute{a}\ l\acute{a}t\acute{e}a\ li\ ph\acute{a}\ h\acute{u}\ [ph\bar{e}\ \text{ʔ}\bar{i}r\bar{o}] h\acute{e}\ n\Lambda$   
 he 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)  $[ne\ ch\acute{a}\ m\bar{\alpha}\ phr\acute{e}\ t\bar{e}\ S\acute{e}th\bar{u}ph\bar{e}] \text{ʔ}a\ h\acute{e}\ n\Lambda$   
 you fight PCT Shan PTC (name) she 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 n $\Lambda$  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)  $kw\bar{a}\ n\Lambda\ ma\ d\acute{s}\ [\text{ʔ}a\ m\bar{u}\ l\bar{i}p\bar{o}na\ k\bar{u}kl\acute{s}] \underline{n\Lambda t\bar{o}be}$   
 axe PTC be.so at he 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 and Object NP but not both, and often consist only of a single verb, as in

*dipɔ 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 *kəjē mō ʔo to* 'person+mother+exist+not: a person with no mother', in which the attributive clause includes the sentence particle *to* 'negative'.

## REFERENCES

- 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).
- Jones, R.B. (1975) 'The question of Karen linguistic affiliation', presented at the 8th International Conference on Sino-Tibetan Languages and Linguistics, Berkeley.
- Karenni Literature Department (1994) *The Modern Western Kayah Li-English Lexicon*, Karenni 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, F.K. (1967a) 'Ethnic categories in Burma and the theory of social systems', in P. 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 J.H. Steward (ed.) *Contemporary Change in Traditional Societies* v. 1, Urbana: University of Illinois Press. (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. (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.

# PWO KAREN

*Atsuhiko Kato*

## 1 INTRODUCTION

Pwo Karen belongs to the Karen branch. Its closest relative is probably Sgaw Karen with more distant relationships to the other languages of the Karen branch such as Pa-O and Karenni (Kayah).

The dialects of Pwo Karen can be divided into two groups: the eastern dialects and the western dialects (see Kato 1995). The eastern dialects are spoken in the Karen State, the Mon State, the Tenasserim Division of Burma and western Thailand. The main towns where the Eastern dialects are spoken are Hpa-an, which is the capital of the Karen State, Hlaingbway, Kawkaeik. The western dialects are spoken widely in the area of the Irrawady delta, and the main towns where they are spoken are Bassein, Myaungmya, Kyonbyaw. The dialect treated in this chapter is one of the Eastern dialects spoken around Hpa-an, referred to here as the Hpa-an dialect.

The western and eastern dialects differ in many aspects and are barely intelligible largely because of phonological and semantic differences. For example, in the eastern dialects /caĩN:/ [tɕ ai<sup>55</sup>] means ‘to walk’, but the cognate word in the western dialects is /saiN\_/ [sai<sup>11</sup>], which means ‘to run’. The dialects are not very different at the syntactic level, although there are some striking differences; for example, the eastern dialects have a causative construction which takes a complement sentence such as /jə- ʔaŋ:məN. lə- ʔəwe. lɪ\_/ (1sg-order-COMP-3sg-go) ‘I ordered him to go’, but the western dialects have no such construction. Instead, in the western dialects, the same thing is expressed by using a causative auxiliary /ʔaŋ\_mə\_/ which is cognate with eastern /ʔaŋ:məN./, i.e. /jə- ʔaŋ\_mə\_ le: ʔəwe\_/ (1sg-CAUS-go-3sg).

The exact number of Pwo Karen speakers is unknown. According to the estimated population statistics published by the Burmese government in 1993, there were 2.86 million Karen in Burma, but it does not say how many of these were Pwo Karens. An estimate would be that over one million Pwo Karens live in Burma and since there are also a number of Pwo Karens 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 neighbours.

Several writing systems have been created for the Pwo Karen dialects. The most widely used in Burma are the Monastic script and the Mission script (Stern 1968). The Monastic script was created for one of the eastern dialects (perhaps for the Hpa-an dialect) and is based mainly on the Mon script. Its history is poorly understood, but surviving records in it date back to the middle of the nineteenth century (U Phon Myint 1975). It is now coming into widespread use throughout the Karen State, since it is taught in many Buddhist monasteries. The Mission script, based on Burmese alphabet, was originally invented for one of the eastern dialects by an American missionary, but it is not in vogue in the eastern areas where Buddhists overwhelmingly outnumber Christians. It is, however, presently popular in the delta, where Christians have been increasing in number. Unfortunately, it does not altogether suit the phonological systems in the Western dialects.

## 2 SOUNDS

The syllable structure of the Hpa-an dialect can be represented as C1(C2)V1(V2)(C3)/T, where C1 is an onset consonant, C2 is the second member of a cluster, V is a vowel, C3 is a coda, and T is a tone. In this structure, -V1(V2)(C3) is referred to as the rhyme.

There are consonant phonemes as shown below:

Stops					
p	θ	t	c	k	ʔ
ph		th	ch	kh	
b		d			
Fricatives					
			ç	x	h
			ʃ	ʁ	
Nasals					
m		N	ɲ	(ŋ)	N
Semivowels					
w			J		
Liquids					
		L	R		

/θ/ is an interdental stop. /c/ and /ch/ are affricates; [tɕ][tɕh]. The /b/[ɸ] and /d/[ɗ] are imploded, but while the /b/ is consistently imploded the /d/ often is not. All of the consonants except /N/ can occur as onsets, but only the consonants /w/, /l/, /r/, /j/ can occur as the second member of the clusters. In the coda, only /N/, which nasalizes preceding vowels, can occur. In some dialects, including the western Kyonbyaw dialect and the eastern Tavoy dialect, /ʔ/ also can occur as a coda (Kato 1995). In the Hpa-an dialect, however, the final /ʔ/ has already disappeared.

The rhymes are as follows:

### (a) Plain rhymes

Simple vowels			Diphthongs	
i	ɨ	u	ai	aʊ
ɪ	ʊ			
e	ə	o		
ɛ	a	ɔ		

### (b) Nasalized rhymes

<i>Simple vowels</i>		<i>Diphthongs</i>		
əN		eiN	əʊN	oʊN
aN	oN	aiN		

(Note: Nasalization of /eiN/, /əʊN/, /oʊN/ is very weak.)

There are four tones:

High-level	/V:/	[55]	Pronounced with normal voice.
Mid-level	/V=/	[22(3)]	Pronounced with breathy voice. It is sometimes accompanied by a rising contour especially in utterance final position.
Low-level	/V_/	[11]	Pronounced with normal voice.
Falling	/V./	[51]	Pronounced with slightly creaky voice.

When two falling tones are juxtaposed morphologically in a single word, the first falling tone often changes to the low-level tone:

(1) *thi*. 'water' + *khlan*. 'hot (of water)' > *thi\_khlan*. '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. When they are followed by a word boundary, they are represented as /Cə-/.

Intonation sometimes distorts the pitch contours of the tones. For instance, tones of verbs before the perfective particle /jəu-/ often have a contour <sup>[223]</sup>, which resembles the contour of the mid-level tone:

- (2) *phla= ba: [223] jəu-*  
 arrow hit PERF  
 'The arrow has hit (the mark)!'

Since the verb /ba:/ has the high-level tone, it may be pronounced as <sup>[55]</sup>, but when it is pronounced as <sup>[223]</sup>, the sentence clearly shows that the speaker has hoped the occurrence of the event. All of the tones may be pronounced <sup>[223]</sup> before the particle /jəu-/. The influence of intonation patterns can also be observed in the verbs of yes-no interrogative sentences and the verbs of sentences answering them:

- (3) *nə- mə- lɿ [223] ʔa.*  
 2sg IRR go QUE  
 'Are you really going?'  
*mə- lɿ [223]*  
 IRR go  
 'Yes, of course!'

This contour shows doubt or suspicion in a question, and a strong belief or confidence in an answer. Instead of an intonational pattern, this pitch pattern might alternately be analysed as a tone change of the verb to the mid-level tone, but this matter requires further investigation.

### 3 PARTS OF SPEECH

Pwo Karen words can be divided into those which can constitute an utterance in isolation (i.e. verbs, nouns, adverbs, and interjections) and those which cannot (i.e. particles).

The words which can constitute an utterance in isolation, other than interjections, can be defined as follows. Verbs can have verb particles attached to them (preverb particles and postverb particles are discussed later). For example, /lɿ-/ 'to go' is a verb, since it can be preceded by the preverb particle /mə-/ (irrealis):

- (4) *mə- lɿ-*  
 IRR go '(I) will go.'

Nouns can be the argument of a verb but cannot have verb particles attached to them, for example, /phloʊN-/ 'person; Karen' below:

- (5) *\*mə- phloʊN-*  
 (6) *phloʊN- ʔɔ- lə- ɣa-*  
 person exist one NC  
 'There is one person.'

Adverbs can neither have verb particles attached to them nor can they be the argument of a verb, for example, /ləpəʊN-/ 'much' below:

- (7) *\*mə- ləpəʊN-*  
 IRR much

- (8) \**ləpəuN\_ ʔɔ:*  
 much exist

In contrast, particles cannot constitute an utterance in isolation. On this point, they are similar to affixes (discussed later). In this chapter, the relatively independent morphemes are referred to as particles, but ultimately it may be impossible to draw a strict dividing line between particles and affixes. Below are the most important particles (each list is not an exhaustive one):

### 3.1 Adpositions

Pwo Karen has prepositions, such as these illustrated below:

*lə-/ (~lə:/~le:/)* ‘in, at’ (location), ‘to’ (goal) or ‘from’ (source).

- (9) *ʔəwe. ʔɔ: lə- ʔə- yeiN:*  
 3sg exist in his house  
 ‘He is in his house.’

*/de=* ‘with’ (instrument or accompanier) or ‘and’ (conjunction)

- (10) *jə- ʔaN: mI\_ de= nɔ:θoUN\_*  
 1sg eat rice with spoon  
 ‘I eat rice with a spoon.’

- (11) *kho\_θa: de= θa\_kwi\_θa:*  
 mango and banana  
 ‘a mango and a banana’

There is also a circumposition:

*/be.~θo\_/* ‘like’, ‘as’.

- (12) *ʔaN: mI\_ be. cəpaN. θo\_*  
 eat rice like Japanese like  
 ‘(He) eats rice as a Japanese.’

### 3.2 Demonstratives

There are demonstratives, such as shown below (they follow nouns):

- (13) *yeiN: jo\_*  
 house this  
 ‘this house’
- (14) *yeiN: nɔ:*  
 house that  
 ‘that house’
- (15) *yeiN: ʔo\_*  
 house that  
 ‘that house’ (very far)

Demonstratives are often used as topic markers (especially */nɔ:/*):

- (16) *ʔə- yeiN: nɔ: jə- lI\_ ʔe:*  
 his house TOP 1sg go NEG  
 ‘To his house, I didn’t go.’

### 3.3 Adverbial particles

Particles function as adverbs, but they cannot constitute an utterance in isolation. For example:

- (17) *ɣl\_ ma=*  
 good very  
 '(It) is very good.'
- (18) *ɣl\_ pəθai\_ ʔe:*  
 good not so NEG  
 '(It) is not very good.'

### 3.4 Preverb particles

Particles before a verb show various meanings including irrealis, negation, obligation, and causation:

/mə-/ Irrealis marker. There is no marker for realis.

- (19) *jə- mə- l\_*  
 1sg IRR go  
 'I will go.'

/lə-/ Negative marker:

- (20) *jə- lə- l\_ ba: ʔəkhə:con\_ ...*  
 1sg NEG go NEG because  
 'Because I didn't go ...'

/ba:/ 'must, have to':

- (21) *jə- ba: l\_ ɣa.*  
 1sg must go QUE  
 'Do I have to go?'

/da\_/ Causative marker:

- (22) *jə- da\_ l\_ ʔəwe.*  
 1sg let go 3sg  
 'I let him go./I made him go.'

### 3.5 Postverb particles

Particles after a verb add various meanings including 'to try', 'to do for a purpose', 'to do in advance', and various directions.

/jʊ=wa=/ 'try to (do)' (< /jʊ=wa=/ 'to look after (someone)')

- (23) *jə- ke\_ jʊ=wa= lai:*  
 1sg write PVP alphabet  
 'I tried to write the alphabet.'

/tha:/ (~/da:/) 'to do something for some purpose'. Its meaning is similar to the Burmese auxiliary *-tha:* or the Japanese morpheme '(-te) oku'.

- (24) *jə- ke\_ tha: lai:*  
 1sg write PVP letter  
 'I wrote a letter (for a certain purpose).'

/wɛ\_/ ‘(to do) in advance’

- (25) *jə- lɔ\_ wɛ\_ ʔəwe.*  
 1sg tell PVP 3sg  
 ‘I told him in advance.’

/thaN:/ ‘upward movement’, ‘increase’ or ‘change to a better condition’, etc. (< /thaN:/ ‘to ascend’ ‘to climb’)

- (26) *khlain\_ thaN:*  
 speak PVP  
 ‘He spoke upward.’

/laN\_/ ‘downward movement’, ‘decrease’ or ‘change to a worse condition’ etc. (< /laN\_/ ‘to descend’)

- (27) *ke\_ laN\_*  
 write PVP  
 ‘to write down’

### 3.6 Subordinate clause markers (see also Section 5.6.3)

For example:

- (28) *jə- ʔaN: mɿ\_ ʔoN\_, ʔaN:lu= thi.*  
 1sg eat rice after bathe water  
 ‘After eating, I took a shower.’

### 3.7 Conjunctions

For example:

- (29) *jə- lɿ\_ pəjan\_ khaN=, de= ʔəwe. lɿ\_ θaiN. khaN=*  
 1sg go Burma country and 3sg go Thai country  
 ‘I went to Burma, and he went to Thailand.’

### 3.8 Sentence-final particles

The sentence-final particles denote various attitudes of the speaker. For example, /nɛ./ is used when the speaker seeks the agreement of the hearer.

- (30) *ʔəjo\_ yɿ\_ nɛ.*  
 this good SFP  
 ‘This is good, isn’t it?’

## 4 MORPHOLOGY

### 4.1 Inflectional morphology

Pwo Karen nouns and verbs do not inflect. Nonetheless, the pronoun paradigm might be termed inflectional. Each pronoun has two forms: form I and form II.

	<i>Form I</i>	<i>Form II</i>
1sg	<i>jə-</i>	<i>jə_</i>
1pl	<i>pə-~hə-</i>	<i>pə_~hə_</i>
2sg	<i>nə-</i>	<i>nə_</i>
2pl	<i>nəθi:</i>	<i>nəθi:</i>
3sg	<i>ʔəwe.</i>	<i>ʔə_~ʔəwe.</i>
	<i>ʔə-</i>	
3pl	<i>ʔəθi:</i>	<i>ʔəθi:</i>

The form I is used before a verb for subjects, or before a noun to denote a possessor. There are two forms for form I 3sg pronoun; /ʔəwe./ is used before verbs and /ʔə-/ before nouns. But for subjects of subordinate clauses, /ʔəθi:/ is sometimes used. The 3pl form /ʔəθi:/ sometimes becomes /ʔəθi:ʔə-/ before nouns. Form II is used after verbs and prepositions, or when pronouns are topicalized. Each pronoun has an emphatic form: *jəwe.* (1sg), *pəwe. ~ həwe.* (1pl), *nəwe.* (2sg), *nəθi:we.* (2pl), *ʔəwe.* (3sg), *ʔəθi:we.* (3pl).

## 4.2 Derivational morphology

As mentioned already, it is difficult to draw a strict dividing line between particles and affixes. In this article, morphemes which cannot constitute an utterance in isolation, those which are less independent morphologically, are termed affixes. The important affixes are shown below.

### 4.2.1 Suffixes

/cha=/ makes nouns which denote owners or experts.

- (31) *yeiN:cha* = owner of a house < *yeiN*: house  
 (32) *toUN=cha* = dancer of Karen dance < *toUN* = ‘Karen “Don” dance’

/phʊ:/ makes nouns which denote members of a group. cf. /phʊ:/ ‘child’

- (33) *təwan.phʊ*: villagers < *təwan*. ‘village’  
 (34) *phloUN\_phʊ*: Pwo Karen people < *phloUN\_* ‘Pwo Karen; person’

### 4.2.2 Prefixes

/pə-/ (pronounced as /hə-/ also) makes nouns denoting certain kinds of people.

- (35) *pəθa\_baN*: young people < *θa\_baN*: ‘young’  
 (36) *pəmwī\_* guest < *mwī\_* guest (*mwī\_* is a bound form.)

/ʔə-/ makes nouns from verbs.

- (37) *ʔəθa*: fruit < *θa*: to bear fruit  
 (38) *ʔədi*: egg < *di*: to lay (an egg)

/ʔe\_/ (also pronounced as /ʔə-/) makes adverbs from verbs.

- (39) *ʔe\_phlɛ*: fast, swiftly < *phlɛ*: to be fast  
 (40) *ʔe\_blɛ\_* as to be satisfied < *blɛ\_* satisfied

/chə-/ makes nouns from verbs. cf. /chə\_/ ‘thing’

- (41) *chədoN*: fence < *doN*: to fence  
 (42) *chəchəN\_rain* < *chəN\_* to rain

### 4.2.3 Reduplication

Reduplication is used for making adverbs from verbs.

- (43) *phlɛ:phlɛ:* fast, swiftly < *phlɛ:* to be fast  
 (44) *ɣl\_ ɣl\_* well < *ɣl\_* to be good

## 4.3 Compounding

As in many of the monosyllabic languages of the region, many compounds are found in Pwo Karen. Some examples of these are as follows (N and V denote nouns and verbs respectively):

### 4.3.1 $N < N + V$

- (45) *thi. ʔɔ\_* drinking water < *thi.* water + *ʔɔ\_* to drink  
 (46) *mɪ:dwai\_* matchstick < *mɪ:* fire + *dwai\_* to light

### 4.3.2 $V < N + V$

- (47) *θa\_θaN:* to be angry < *θa\_* heart + *θaN:* to ascend  
 (48) *na = ɣəN.* to hear < *na =* ear + *ɣəN.* to hear

### 4.3.3 $N < N + N$

- (49) *me:thi. tear* < *me:* eye + *thi.* water  
 (50) *wɔ\_βeiN:nu = thi.* milk < *wɔ\_βeiN:* cow + *nu = thi.* milk  
 (In N+N compounds, the second element is the head.)

### 4.3.4 $V < V + N$

- (51) *chəN.na =* comfortable to listen < *chəN.* sweet + *na =* ear  
 (52) *ke\_pəɾəN.* to write (a letter) < *ke\_* write + *pəɾəN.* news  
 As in: */ke\_pəɾəN. lai:/* (write-letter) to write a letter

### 4.3.5 $V < V + V$

- (53) *cheiN\_xi.* clean < *cheiN.* clean + *xi.* beautiful  
 (54) *ɣəuN = khəN.* stable < *ɣəuN =* stable + *khəN.* hard

Concatenated serial verbs (see Section 5.5.1) might also be considered verb compounds, since criteria have not been found to distinguish between serial and compound verbs.

## 5 SYNTAX

### 5.1 Word order

The basic Pwo Karen word order is Subject-Verb-Object.

- (55) *θa\_ʔwa thɛ\_ thwi:*  
 Thawa kick dog  
 'Thawa kicked the dog.'

The basic word order is retained even in existential sentences. Pwo Karen has only one existential verb: /ʔɔ:/.

- (56) *lai: ʔaʊ\_ lə- bein: ʔɔ: lə- cəpwɛ= ʔəphan.khʊ:*  
 book one NC exist at table top  
 'There is a book on the table.'
- (57) *jə- phʊ: ʔɔ:*  
 my child exist  
 'I have a child.'

Adverbs and adpositional phrases occur after the verb and the object (if there is one).

- (58) *ʔaʊ: mɪ\_ ʔe\_phlɛ:*  
 eat rice fast  
 'Eat rice fast!'

In the comparative sentence, which is made with a postverb particle /da:/, standard noun is put after it. Since /da:/ is a postverb particle, it cannot be separated from the verb.

- (59) *jə- thɔ. da: ʔəwe.*  
 1sg long than 3sg  
 'I am taller than he.'

## 5.2 Negation

There are two particles which denote negation: /ʔe:/ and /lə-/. /ʔe:/ is a sentence-final particle and is used in main clauses:

- (60) *ʔəwe. khɫaiN\_ chəkhɫaiN\_ xɛ\_xɛ\_ ʔe:*  
 3sg speak language slowly NEG  
 'He does not speak slowly.'

/lə-/ is a preverb particle, and is used in subordinate clauses. When /lə-/ is used before the verb, the particle /ba:/ is usually put immediately after the verb or in clause-final position.

- (61) *ʔəwe. ʔe\_ lə- ʔaʊ: mɪ\_ ba: nɔ:, jə\_ mə- ʔaʊ:*  
 3sg if NEG eat rice NEG that 1sg IRR eat  
 (= ʔəwe. ʔe\_ lə- ʔaʊ: ba: mɪ\_ nɔ:, ...)  
 'If he doesn't eat the rice, I will eat it.'

## 5.3 Interrogative sentences

Yes-no interrogative sentences are made with the sentence-final particle /ba:/:

- (62) *nə- mə- thain\_ ba.*  
 2sg IRR return QUE  
 'Are you going back?'

In sentences that include an interrogative word, the sentence-final particle /lɛ:/ occurs:

- (63) *ʔəjo\_ (mwɛ=) chənɔ: lɛ.*  
 this be what QUE  
 'What is this?'

## 5.4 Noun phrases

Stative verbs and demonstratives follow nouns:

- (64) *phloʊN\_ ɣI\_ jo\_*  
 person good this  
 ‘this good man’

Pwo Karen has many numeral classifiers, which follow numerals. The numeral classifier construction occurs after the noun. The origins of the numeral classifiers are unknown.

- (65) *tho\_ nI= du\_*  
 pig two NC (animal)  
 ‘two pigs’

In possessive constructions, the possessor precedes the possessed. The 3sg pronoun (form I) is often attached to the possessed noun.

- (66) *θa\_ʔwa (ʔə) yeiN:*  
 Thawa his house  
 ‘Thawa’s house’

Relative clauses are introduced by the relative marker /lə-/. The relative clause follows the head noun when the head noun is identical to the subject of the relative clause. In such cases, a pronoun which refers to the head noun occurs in the relative clause, but in informal speech, both the relative marker and the following pronoun tend to be omitted. Relative clauses are often followed by demonstratives.

- (67) *phloʊN\_ (lə- ʔə-) lI\_ lə- ʔə- yeiN: nɔ:*  
 person REL 3sg go to his house that  
 ‘the person who went to his house’

When the head noun is not identical with the subject of the relative clause, the relative clause may either precede or follow the head noun. The relative /lə-/ is used when the relative clause follows the head noun but not when it precedes it, and no coreferential pronoun occurs in the relative clause. The head final type seems preferred in informal speech.

- (68) *khan:phai\_ lə- jə- thaʋ: lə- daʊ\_ phən\_ nɔ:*  
 shoe REL 1sg wear in room inside that  
 ‘the shoes that I wear in the room’  
 = *jə- thaʋ: lə- daʊ\_ phən\_ khan:phai\_ nɔ:*  
 1sg wear at room inside shoe that

## 5.5 Verb serialization

Like other Southeast Asian languages, Pwo Karen has verb serialization. Serialization involving two verbs is examined, since 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 or Vietnamese. In Thai, a neighbouring language, it is possible for noun phrases or preposition phrases to intervene between serialized verbs as /pay talàat súfu plaa/ (go-market-buy-fish) ‘(I) went to the market and bought a fish’. In Pwo Karen, however, the corresponding serialization is not acceptable:

- (69) *\*jə- lI\_ phja. xwe\_ ja:*  
 1sg go market buy fish



In some Karen dialects with more contact to Thai, this type of verb serialization seems acceptable to some speakers, but in the Hpa-an dialect it is unacceptable. 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 /phja./ 'market' somewhere other than between the two verbs, making it a non-argument:

- (70) *jə lɪ\_ xwe\_ ja: lə phja.*  
 1sg go buy fish at market  
 'I went and bought a fish at the market.'

Of course, the same notion can be expressed without using verb serialization at all, for example by using combined clauses:

- (71) *jə lɪ\_ phja. yəN\_, xwe\_ ja:*  
 1sg go market after buy fish

In Pwo Karen, the only time a noun phrase can separate the two verbs is when the second verb denotes the result of the first verb or an objective description about the situation denoted by the first verb. In such cases a noun phrase or preposition phrase can occur between V1 and V2 (shown below).

Serialized verbs in Pwo Karen can be divided into the concatenated and separated serialization according to the position of the negative marker /lə-/ when they occur in a subordinate clause. Pwo Karen also has another type of serialization, the appositive type, where /lə-/ occurs before both the first and the second verb (Kato 1998). But only the two important types will be discussed here.

(1) Concatenated type: the negative marker /lə-/ occurs before V1.

- (72) *jə- lə- lɪ\_ xwe\_ ba: ʔəkhʉ:coN\_,...*  
 1sg NEG go buy NEG because  
 'Because I did not go to buy (something), ...'

(2) Separated type: the negative marker /lə-/ occurs before V2.

- (73) *jə khlaiN\_ phloʉN\_ lə θr: ba: ʔəkhʉ:coN\_,...*  
 1sg speak Pwo Karen NEG can NEG because  
 'Because I cannot speak Pwo Karen, ...'

### 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 combinations of V1 and V2 is quite free and productive. V1 and V2 are usually arranged in accordance with the temporal order of events. The semantic relationships denoted by the sequence of V1 and V2 include cause–result, means–purpose, manner–action, and simultaneity. There are four possible combinations of V1 and V2 according to whether the verbs are intransitive or transitive:

(a) Intransitive + intransitive: the whole concatenated verb functions as intransitive.

- (74) *ʔəwe. laN thi.pha= θi.*  
 3sg tumble die  
 'He tumbled and died.'

- (75) *jə chi.naN\_ ko\_ɕa\_*  
 1sg sit cry  
 ‘I sat and cried.’

(b) Intransitive + transitive: the whole concatenated verb functions as transitive.

- (76) *jə chi.naN\_ ʔan:kho\_ ʔəwe.*  
 1sg sit wait 3sg  
 ‘I waited for him while seated.’

- (77) *jə l\_ xwe\_ ja: lə bein:*  
 1sg go buy fish one NC (flat thing)  
 ‘I went and buy a fish.’

(c) Transitive + transitive: the whole concatenated verb functions as transitive.

- (78) *jə θɜ:ja. na:θɪ: ʔəyaiN=ʔəcon\_*  
 1sg know understand reason  
 ‘I know and understand the reasons.’

- (79) *jə xwe\_ ʔaN: kɔ:*  
 1sg buy eat confectionary  
 ‘I bought a cake and ate it.’

(d) Transitive + intransitive: the whole concatenated verb functions as transitive.

- (80) *jə- dɔ: θi. thwi:*  
 1sg strike die dog  
 ‘I struck the dog to kill it.’

- (81) *ʔəwe. thaɔ\_ cheiN. ʔə mɛ:*  
 3sg brush clean his tooth  
 ‘He brushed his teeth to clean them.’

In the patterns (a), (b), and (c), both V1 and V2 are either volitional verbs (agentive verbs) or non-volitional verbs (non-agentive verbs), and the subjects of V1 and V2 are identical. The objects in pattern (c) are usually, but not always, identical for V1 and V2. See the example below:

- (82) *jə chu.laN\_ ʔaN: mɪ\_ de= thi.la:*  
 1sg put in eat rice with salt  
 ‘I ate rice after putting salt in it.’

In such cases, the object of the concatenated verbs is always an argument of V2. In the example above, /mɪ\_/ is an argument of V2. The word /thi.la:/, which is the patient of V1, cannot be put in the object position for V1 + V2:

- (83) *\*jə chu.laN\_ ʔaN: thi.la:*  
 1sg put in eat salt  
 ‘I ate (rice) after putting salt in it.’

Pattern (d) differs from the other patterns on two points. First, the logical subjects of V1 and V2 are not identical. In /jə- dɔ: θi. thwi:/, the logical subject of V1 is ‘I’ which is the subject

of the whole V1 + V2 construction, whereas the logical subject of V2 is the 'dog', which is the object of V1 + V2. The logical subject of V2 is usually identical with the patient of V1, but this is not always the case, although such exceptions are relatively rare:

- (84) *jə- khəuN: laN\_bəN\_ phloʊN\_cʊ:*  
 1sg dig be buried dead body  
 'I dug (a hole) to bury the dead body.'

Second, the volitionality of V1 and V2 differ from each other in this pattern; i.e. V1 is volitional but V2 is not. Volitional verbs cannot occur as the second verb:

- (85) *\*jə- dʊ: mi.naN\_ ʔəwe.*  
 1sg strike lie down 3sg

Note that, although the transitive + transitive pattern may show the same properties as pattern (d), this is quite rare. Consider the example below:

- (86) *jə- lɔ\_ ɣəN. ʔəwe.*  
 1sg tell hear 3sg  
 'I told him (a story).'

In this sentence, the logical subjects of V1 and V2 are different, and the verbs also differ in volitionality. In such cases, the second verb is always very low in transitivity.

Concatenated verbs usually occur in temporal order. However, there are two exceptions. One such case is found below:

- (87) *jə- mi. kəN\_ thədoN\_*  
 1sg sleep wear sarong  
 'I wore my sarong and went to sleep.'

In this example, the order /*kəN\_mi*/ (wear-sleep) is not allowed. It seems 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. Another exception is where one of the consecutive events involves movement:

- (88) *jə- ɣɛ. ʔaN: mI\_*  
 1sg come eat rice  
 'I came after having lunch.'

Here also the order /*ʔaN: ɣɛ*/ (eat-come) is not allowed. Verbs denoting movement always have to occur as V1. Since this sentence also means 'I came to eat lunch', it is semantically ambiguous. The next sentence is another example:

- (89) *ʔəwe. ɣɛ. kli: (\*kli: ɣɛ.)*  
 3sg come run  
 'He came running.'

### 5.5.2 Separated type

In verb serialization with separated verbs, noun phrases, or prepositional phrases may occur between V1 and V2. With these, the co-referentiality of the arguments of V1 and V2 vary from case to case. Below are the examples:

- (90) *jə dʊ: θwi: θi. pəʊN=*  
 1sg strike dog die SFP  
 ‘When I struck the dog, it happened to die.’
- (91) *jə ʔaN: mɿ\_ blɛ\_ jaʊ\_*  
 1sg eat rice full PERF  
 ‘I’ve eaten rice and got full.’
- (92) *ʔəwe. kli: phlɛ:*  
 3sg run fast  
 ‘He runs fast.’
- (93) *jə- khlain\_ phloʊN\_ bəʊN: ʔe:*  
 1sg speak Pwo Karen brave NEG  
 ‘About speaking Pwo Karen, I don’t dare (to do it).’

In separated verb serialization, *v2* denotes the result of *v1*, as in the first two examples above, or denotes an objective description or judgment about the event denoted by *v1*, as in the last two examples. Note that *ʔəwe. kli: phlɛ:/* is an objective description of ‘his running’. Thus, this serialization cannot occur in imperative sentences: *\*/kli: phlɛ:/* ‘Run fast!’. An imperative would have to be given as */kli: phlɛ:phlɛ:/* (run-fast) or */kli: ʔe\_phlɛ:/* (run-fast)). Only verbs denoting non-volitional situations can occur as the second verb. They are usually, but not always, intransitive, as below:

- (94) *jə- ʔaN:xu. khaN:phai\_ da: ʔe:*  
 1sg look for sandal find NEG  
 ‘I looked for (my) sandals but I couldn’t find (them).’

What should be noted is the semantic difference between concatenated serialization and separated serialization. Consider the examples below:

- (95) *jə dʊ: θi. θwi:* (Concatenated type)  
 1sg strike die dog  
 ‘I struck the dog to kill it.’
- (96) *jə dʊ: θwi: θi. pəʊN=* (Separated type)  
 1sg strike dog die SFP  
 ‘When I struck the dog, it happened to die.’

In the first example, the ‘death’ of the dog was expected from the beginning, thus it is purposive. In the second example, however, the ‘death’ occurred unexpectedly or accidentally, thus this serialization denotes a cause-result. Interestingly, the first sentence does not always imply the death of the dog, but the second always implies it. This fact is established by putting a clause meaning ‘but it did not die’ after them:

- (97) *jə dʊ: θi. θwi:, la=naN. θi: θi. ʔe:*  
 1sg strike die dog but also die NEG  
 ‘I struck the dog to kill it, but it did not die.’
- (98) *\*jə- dʊ: θwi: θi., la=naN. θi: θi. ʔe:*  
 1sg strike dog die but also die NEG

The addition of this clause to clauses containing separated serialization yields a contradiction, whereas its addition to cases of concatenated serialization is fine. This is because the separated serial verbs always imply a result but the concatenated ones do not. This is quite similar to verb serialization in Kayah, where concatenated verbs like (97) also do not always imply a result (Solnit 1997: 68).

## 5.6 Sentences with plural clauses

### 5.6.1 Complement sentences

The complementizer /*lə-/* (~/*lə:/~/le:/*) may occur before complement sentences which are embedded as the object of matrix sentences.

- (99) *jə- da: (lə-) ʔəwe. kli:*  
 1sg see COMP 3sg run  
 'I saw him running.'

But, /*lə-/* is not attached to the complement sentences which occur as the subject of matrix sentences.

- (100) [*hə- ʔaN: chədo\_chəla:] ɣl\_ ma=*  
 1pl eat vegetable good very  
 'It is good to eat vegetables.'

### 5.6.2 Coordinated clauses

Conjunctions are used to coordinate clauses. In the next example, /*la=naN/* 'but' is used to mark the coordination of two clauses, but it is used in the next section as a subordinate clause marker. When it is used as a conjunction, a pause may occur before it.

- (101) *jə- l\_ pəjaN\_ khaN=, la=naN. θi: ʔəwe. l\_ θaiN. khaN=*  
 1sg go Burma country but also 3sg go Thai country  
 'I went to Burma, but he went to Thailand.'

### 5.6.3 Adverbial clauses

Adverbial clauses usually precede main clauses and come marked in various ways.

(a) Adverbial clauses may be marked by various kinds of subordinate clause markers. In the examples below, some subordinate markers occur clause-finally:

- (102) *jə- ʔɔ: lə\_ maɔ: ba: ʔəkhɔ:coN\_, ləni\_jo\_ jə\_ ma\_ kəloUN.*  
 1sg exist NEG healthy NEG because today 1sg do job  
*kɛ: ʔe:*  
 take place NEG  
 'Since I'm not fine, I can't do my job today.'

- (103) *paU\_ thaN: pai\_təlaN. ɣoN\_, chəphɔ:xa= naV: laN\_ we. ləpoUN\_ ma= lɔ.*  
 open PVP window after insect enter PVP PVP much very SFP  
 'After (I) opened the window, many insects came in.'

- (104) *lai: ʔaU\_ ʔɔ: ləpouN\_ la=naN. θi:, jə pɔ= maU: ʔe:*  
 book exist much but also 1sg read comfortable NEG  
 ‘Although there are many books, they aren’t fun.’

/la=naN./ of the last example above may occur immediately after the verb:

- (105) *lai: ʔaU\_ ʔɔ: la=naN. ləpouN\_ θi:, jə- pɔ= maU: ʔe:*  
 book exist although much also 1sg read comfortable NEG

In the next examples subordinate clause markers occur clause-initially:

- (106) *kəla\_ lə ʔaN: mI\_ dai\_ ba:, θi.ja\_ khwai: wi= nə- cu:*  
 before NEG eat rice still NEG wash PVP PVP your hand  
 ‘Wash your hands before you eat rice.’

- (107) *jə wai\_ phr: nə- lu:, θoN=, jə- jaw: ʔaU\_*  
 1sg spin PVP your thread till my age run out  
 ‘I will spin your threads for the rest of my life.’

Although adverbial clauses usually precede main clauses, those with /θoN=/ usually follow main clauses. /θoN=/ perhaps originated from /θoN\_/ ‘to reach’. Next, The subordinate clause marker /ʔe\_/ occurs before the verb:

- (108) *maN\_cɔ\_ θa\_bjɔ\_ ʔe\_ ye:, da\_ ʔɔ\_ ləxl\_*  
 uncle Thabyaw if come let drink don’t  
 ‘If uncle Thabyaw comes, don’t let him drink (liquor).’

The subordinate clause marker /be. ~ θo\_/ surrounds clauses:

- (109) *be. jə- mə- θr: θo\_, jə- klɪ\_cɪ\_ cha. ma=*  
 so as to 1sg IRR can so as to 1sg endeavour much very  
 ‘I endeavoured so much so that I could (do it).’

(b) Adverbial clauses may also be marked by special nouns which introduce subordinate clauses. For example:

- (110) *jə- l\_ ʔəkha., chə- chəN\_ cha. ma=*  
 1sg go time thing rain much very  
 ‘When I went (there), it rained hard.’

(c) Finally, adverbial clauses may be marked by topic markers. For example:

- (110) *l\_ jU= nɔ:, chə- ʔɔ: naN= meiN\_ ʔe:*  
 go look TOP thing exist any kind NEG  
 ‘When I went there, I found nothing.’

## ADDITIONAL ABBREVIATIONS

IRR	irrealis marker
NC	numeral classifier
PVP	postverb particle
QUE	question marker
SFP	sentence-final particle
TOP	topic marker

## REFERENCES

- Cooke, Joseph R., Hudspith, Edwin and Morris, James A. (1976) 'Phlong (Pwo Karen of Hot district, Chiang Mai)', in William A. Smally (ed.) *Phonemes and Orthography: Language Planning in Ten Minority Languages of Thailand*, pp. 187–220. Pacific Linguistics Series C-No. 43, The Australian National University.
- Jones, Robert B. Jr. (1961) *Karen Linguistic Studies*. Berkeley and Los Angeles: University of California Press.
- Kato, Atsuhiko (1995) 'The phonological systems of Pwo Karen dialects', *Linguistics of the Tibeto-Burman Area* 18.1: 63–103.
- (1998) 'Poo Karen-go (Toobu Hoogen) no doosirenzoku ni okeru syudoosi ni tsuite (On head verbs of serial verb constructions in Pwo Karen (the eastern dialect))', *Gengo Kenkyu* (Journal of the Linguistic Society of Japan) 113: 31–61. (in Japanese)
- (1999) 'Two types of causative construction in Pwo Karen (the Eastern dialect)', in Shintani Tadahiko (ed.) *Linguistic and Anthropological Study on the Shan Culture Area*, pp. 55–93. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- (2001) 'Poo Karen-go no kankeisetsu (Relative clauses in Pwo Karen)', *Tokyo University Linguistic Papers* 20: 275–300. (in Japanese)
- 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.
- (2000) 'West-central Thailand Pwo Karen phonology', *33rd ICSTLL Papers*: 99–110, Ramkhamhaeng University, Bangkok.
- Purser, W.C.B. (1922) *A Comparative Dictionary of the Pwo-Karen Dialect*. Rangoon: American Baptist Mission Press.
- 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 Bhuddhist Pwo Karens)* Rangoon: Sapreuu:caapetuik. (in Burmese)

**PART 13**

---

**OTHER LANGUAGES**

---





# YUNNAN BAI\*

*Grace Wiersma*

## 1 INTRODUCTION

The Bai language is spoken today by more than 1 million people of the Bai nationality, a scheduled minority ethnic group of the People's Republic of China since the establishment of an autonomous prefecture of the same name in 1956. Speaker strength and density are concentrated in and around communities such as Xiaguan, Dali, Yunlong, Eryuan, Jianchuan, and Heqing, all located within the borders of the Dali Bai Autonomous Prefecture in north-west Yunnan. But Bai speakers are also found in other communities radiating outward in all directions from the autonomous prefecture: to the northwest in Weixi, northwards in Lijiang, to the southwest in Baoshan, southeast in Chuxiong, and in the Kunming area. Many native speakers of Bai are bilingual; in addition to bilingualism involving Chinese, depending on the situation, bilingualism involving Lisu, Yi, or Naxi is also common. Yunnan census data for 1990 report the Bai population within the province as 1.3 million. Bai population statistics should be viewed against the population of Yunnan as a whole, which in 1990 amounted to 36.9 million, of whom 24.6 million were Han Chinese. Compared to other ethnic minorities in Yunnan, the Bai people are second in numbers only to the Yi, of whom there were 4.5 million in 1990.

---

\* This survey is a much-reduced version of a longer work which remains in manuscript, cited here as Wiersma 2001. Citation to sources published in Chinese is given in full in the longer paper, and a fuller sample and discussion of the data are also provided. Although I alone am responsible for the views expressed here, my perspective on Bai has grown from discussion with many experts over the past fifteen years. Most recently, and during the preparation of this chapter, I have enjoyed fruitful exchanges with Fu Jingqi and Xu Lin. My original inspiration was provided through contact with James Matisoff at Berkeley. Jim challenged me to commit to the study of Bai, aided and abetted by the late Paul K. Benedict. My efforts to grasp the substance and context of the language might have foundered without the support of the late Zhao Yansun, who generously shared his views, contacts, and unpublished material with me in China despite considerable inconvenience. My field trips to Dali were supported and facilitated by the late Yang Yingxin, then head of the Yunnan Minority Language Advisory Commission. I regret the passing of these two departed Bai enthusiasts before I was able to thank them publicly for their generosity and encouragement. Li Shaoni was my principal language consultant in China, and I am grateful to him for patiently enduring my interviews and for generously introducing me to his home community during the months of April and May, 1989. Finally I wish to express grateful thanks to the former US Committee on Scholarly Communication with the People's Republic of China through which funding for my research in China from 1987 to 1989 was provided, as well as to the Special Fund for Exchange with China of the Hong Kong Polytechnic for a grant which allowed me to attend the Bai corpus planning symposium in 1993. And thanks also to the Graduate Division at Berkeley for its support under the Humanities Graduate Research Program which allowed me to consult with François Dell at the Centre de Recherches linguistiques en Asie Orientale in 1986.

### 1.1 Dialects, subgrouping, and genetic affiliation

The first comprehensive statement of the Bai dialect situation was made in a 1958 report to the Chinese Academy of Sciences by the Bai Language Working Group, which had surveyed Bai speakers extensively in fifty-nine Yunnan locales (or *dian*<sup>3</sup>) over a ten-month period in 1957. At this time only two varieties were recognized as distinctive dialects: Jianchuan and Dali. The Bai Language Working Group was only one of several groups organized under the larger Language Survey Team No. 3, which conducted fieldwork throughout Yunnan province on many languages in addition to Bai, under the overall direction of Fu Maoji.

Sporadic differences later assigned to the Bijiang dialect were also reported by the Working Group in 1958, notably an atypical repertoire of consonant initials, including a series of retroflex stops and affricates, and a reduced inventory of tonal reflexes: only six as opposed to eight for Dali and Jianchuan. Lexical and syntactic isoglosses were minimized at this point, and the report concentrated on outlining general syntactic patterns where Bai resembles Chinese (e.g. SVO ordering), in contrast to other co-existing patterns where Bai resembles the Yi (Loloish) languages (e.g. SOV as well as {N + NUM + CL} ordering). Isoglosses for the three established dialects are still drawn from synchronic differences and not based on historical comparison, and therefore it cannot be said that the issue of subgrouping for Bai dialects has been publicly addressed to date.

Beginning in the late nineteenth century and continuing throughout the twentieth, the genetic affiliation of Bai has been the subject of continued scholarly speculation, and the ongoing controversy has given rise to a considerable literature. It is interesting that while modern linguists in China consolidated their view of Bai as belonging to Tibeto-Burman either as an isolate or as a member of the Yi (Loloish) group, Bai-speaking literati of the early to mid-twentieth century had relied on their knowledge of Chinese philology to uncover archaic Chinese etymologies for the basic terms of their own language. Western scholars discovered such etymologies independently within the last two decades, and the situation at present is that scholarly opinion in China, as well as in the international community, is sharply divided as to whether Bai is historically closer to Sinitic or to Tibeto-Burman. Meanwhile, the possible roots of intimate borrowing by an autochthonous language from literary Chinese are suggested by recent anthropological work in Dali, which documents the effects of Han Chinese in-migration in local marriage customs, and suggests how autochthonous cultural terms may have been ‘bleached’ by Chinese through the growth of local documentary traditions (Yokoyama 1994 and elsewhere).

### 1.2 Language history, writing, and Romanization

Bai lexical items belonging to the colloquial stratum of Chinese loans, the ‘earliest’ layer of Chinese identifiable in the language, are now routinely compared to Chinese cognates as reconstructed for Han times, or in other words to Old Chinese (Benedict 1981; Starostin 1994). Whether or not such comparisons constitute evidence for genetic affiliation, they do attest to a history of intimate contact with Chinese that involves early material, whatever the date of the contact period. So, despite differences of opinion on the genetic question it may still be said that the outlines of Bai language history cannot be pursued independently of Chinese historical developments. Meanwhile, we know that the Jianchuan dialect area was politically under Tibetan control during the early eighth century (Wiersma 1990: 18–19). It might be tempting to posit that the partial case-marking paradigm affecting Bai personal pronouns has resulted from a period of contact with Tibetan. However, local control of Jianchuan by Tibetans was short-lived, and had been thrown off by the late eighth century.

The subsequent history of the region virtually obscured Tibetan political (though perhaps not cultural) influence, making it more likely that case inflection through vowel and tone shift in Bai actually reflects a pattern of independent development. As such, genitive case marking in Bai and its sporadic extension to other functions invite comparison to the parallel developments recently attributed to historical drift for other TB languages (LaPolla 1994: 64–70, 76–7). Returning to the question of Chinese contact, it is interesting that while Dali's local rulers were finally brought down by Mongol conquest in the thirteenth century, this conquest actually ushered in a period of administration in Yunnan by an Arab Muslim official whose influence throughout the province has extended well beyond the Mongol period (Armijo-Hussein 1996: 150–92). It is not difficult to suppose that prior mastery of written Chinese by elite sectors of the local population would have been a key to their political survival over the long term from this point onwards.

Bai-speaking people have been traditionally known as 'Min Chia', both in China and abroad. This is a Chinese term that came into use in the local context only around 1575 when it was editorially superimposed upon the local ethnic term, whose Chinese representation was then redeployed elsewhere in official historical documents (Wiersma 1990: 12–13). The contemporary ethnonym 'Bai' reflects long-standing common usage within the speech community. Bai speakers typically refer to themselves as /pɛ<sup>4</sup>.xo<sup>2</sup>/, /pɛ<sup>4</sup>.tsi<sup>7</sup>/ or /pɛ<sup>4</sup>.jĩ<sup>5</sup>/, all terms derived through addition of a suffix to the ethnic term 'Bai' {WHITE} (a Chinese 'loan' of the colloquial stratum, for which see Section 2.2.2), using a suffix meaning {MEMBER OF A GROUP}, {HUMAN}, or {PERSON}. The term Min Chia, on the other hand, is open to various explanations. Whether it originally meant 'commoners' households', as glossed in traditional dictionaries, or 'members of the nobility' as interpreted by one local historian (Wiersma 2001: 8), Min Chia is a literary Chinese term whose interpretation requires a scholarly understanding of Chinese historical phonology. It is also worth mentioning that terms currently used to identify various Yi speaking groups (Nosu, Nasu, etc.) are typically analysed as being composed of a native morpheme meaning {BLACK} followed by a suffix meaning {PERSON}.

Given the wealth of archaeological and historical documentation that is available for the social history of the Dali region, it is fair to suppose that historical development of the Bai language as a whole has been influenced – for as long as the current varieties have co-existed – by local traditions of learning in Chinese. Documentary sources attest that considerable levels of attainment were achieved in the official examinations of the Chinese court, beginning from the Mongol dynasty (thirteenth to fourteenth centuries) and increasing steadily thereafter, in several historical locales where today Bai speaker density is known to be the greatest. Authorship of the best-known local inscription in Chinese, dating from the late eighth century (and still on display in Dali), is generally attributed to a sojourning Tang official from the north, but later extant inscriptions and a tradition of lost historical works (appropriated by and preserved in Chinese works of the Ming period) are thought to represent the work of local writers beginning around the rise of the Dali state (tenth century and onward) (Wiersma 1990: 15–17; Wiersma 2001: 6). In addition to local writings in Chinese, artifacts from the Dali region also document simultaneous use of the Brahmi script for inscriptional functions between roughly AD 800 and 1100. Finally, a later (fifteenth century) tradition of inscriptional Chinese instantiated mainly by Dali funerary texts is thought to reflect adaptation of the Chinese script to encoding a local vernacular, and though controversy still surrounds the notion, texts of this type have been interpreted as Bai language documents.

Reasons for early adoption of Jianchuan Bai as the basis for creation of a Roman-based draft orthographic scheme were partly inherited from the Sino-Soviet mentoring relationship which shaped the practice of ethnolinguistics and minority language policy in new China. Linguists assumed that adoption of a standard language was the first step towards a uniform

orthography, and that the actual standardization should be based on a form of speech used for daily communication by a significant population situated in dense proximity and subscribing to a uniform local identity. Adoption of the Dali variety as a standard was not mandated in 1958 because it failed the speaker density test. Compared to the relatively mixed population of the 'southern' or Dali dialect area, in Jianchuan County, over 90 per cent of the population was and remains Bai speaking, while native Chinese speakers are fewer than 5 per cent. In Jianchuan, Bai is the spoken medium of choice in and around local government offices and commercial enterprises, and it is also sometimes used for inter-ethnic communications. The draft Bai orthographic scheme based on the Jianchuan dialect was withheld for political reasons at the time of the Working Group's original report, but was published during the 1980s and introduced in a series of educational programmes mainly located in Jianchuan County.

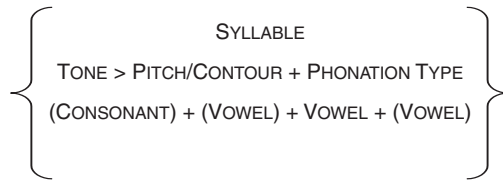
By the early 1990s, however, a large-scale exercise in community consultation was underway, conducted by the provincial language policy and oversight body, culminating in a well-orchestrated Bai language and orthography symposium held in 1993. The symposium entertained diverse proposals for revision of the draft orthographic scheme, and heard new policies articulated bypassing the earlier emphasis on uniformity and standardization. Despite cautionary remarks by senior linguists, the symposium as a whole recommended an inclusive orthography that promised to accommodate both Jianchuan and the newly prestigious Dali variety through systematic or legal variations in spelling. The current situation is that the significant body of primary materials which had appeared based on the 1984 draft orthography (most recently, Zhao and Xu 1996) cannot be interfiled with any locally produced orthographic materials based on the 1993 symposium, because alphabetically salient modifications were then introduced to the system of tone marking by final consonants, thereby licensing widely divergent spellings for the same words. Although the 1993 revised scheme has been recognized at the national level, it was published only in a provincial-level language planning journal, annexed to summary comments on the 1993 symposium (Wiersma 2001: 10).

## 2 PHONOLOGY

The phonology of Bai is comparable to that of northern Yi and certain other LB languages, where historical checked syllable endings have been lost and have given rise to tense/lax phonation type contrasts. However it is also the case that syllable prosodies including register and voice quality contrasts from earlier initial voicing have been found in LB, and linked to historical contact with Mon (Bradley 1982: 127–9). Syllable prosodies from \*-p, \*-t, or \*-k endings in Bai show evidence of 'crosscutting' by register from earlier initial voicing. In this and following sections, illustrations are drawn from the Jianchuan dialect of Bai except as otherwise indicated.

### 2.1 Syllable structure and phonotactics

Syllables in the Bai language are invariably open and every syllable is considered to be tone-bearing. The syllable canon may be represented as in Figure 40.1, where unbracketed sub-syllabic components make up the minimal syllabic unit. No Bai dialect has been described with checked or stopped syllable endings, but there is a rhotic final ending to one class of syllables in the Dali variety. The rhotic ending is not applied to relevant loans from modern Chinese in Dali Bai, nor can the process itself be considered cognate to mandarin because while the conditioning in Mandarin is semantic or lexical, in Dali Bai the conditioning is clearly an aspect of phonological history and the ending has no semantic meaning.



**FIGURE 40.1** JIANCHUAN BAI SYLLABLE STRUCTURE

### 2.1.1 Consonant initials

The Jianchuan Bai consonant initials are as follows:

*p p<sup>h</sup> m f v t t<sup>h</sup> n l ts ts<sup>h</sup> s [z] tʃ tʃ<sup>h</sup> ʃ j k k<sup>h</sup> ŋ x ʂ*

The standard inventory also includes a set of four retroflex (affricate and fricative) sounds to accommodate educated pronunciation of Beijing Mandarin loans, but these sounds have not been included here because they are not required by the local inventory, such loans popularly assuming a pronunciation that fits within the local parameters.

### 2.1.2 Vowel rhymes

The Jianchuan Bai vowel rhymes are as follows:

*i e ε a o u u v ao iε ia io iuu ui uε ua iao*

In Jianchuan Bai, all rhymes except / u /, / ao /, and / iao / have nasalized reflexes, although the distribution of such nasalized rhymes is limited. The feature of nasalization is lexically distinctive in Jianchuan, while in Dali Bai, no nasalized rhymes are found. As observed by Starostin 1994, some Jianchuan nasalization appears to be secondary rather than historical. The Jianchuan syllabic inventory may be represented as in Table 40.1, where the combinatory possibilities of consonant initial and vowel rhyme are specified in full. These combinations have been verified against lexical data published in the standard sources.

## 2.2 Suprasegmentals

The scope of Bai non-modal phonation types and the related specifications of co-articulated features have recently been expanded to include both ‘harsh’ and ‘breathy’ voice as well as tense voice (Edmondson and Li 1994: 51–3, 58). All non-modal or laryngeal phonation types do occur with both nasal and oral rhymes, and it goes without saying that both nasal and oral rhymes can co-occur with modal voice or ‘clear’ phonation.

### 2.2.1 Tones and phonation types

The Jianchuan variety of Bai is described in standard sources as having eight tones, two groups correlated, respectively, with modal and non-modal phonation. This situation may be represented as in Figure 40.2, where the symbol {+P} indicating non-modal phonation is associated with relevant pitch and contour parameters to indicate the added features associated with each tonal reflex. Tones 1, 2, and 4 have tense phonation with constriction or foreshortening at the end. Tone 5 has creaky or harsh voice (the term now preferred by Edmondson) and is especially

TABLE 40.1 JIANCHUAN BAI SYLLABARY

	i	e	ɛ	a	o	u	u	v	ao	iɛ	ia	io	iu	ui	ue	ua	iao
p	pi✓ p <sup>h</sup> i✓	pe p <sup>h</sup> e	pe✓ p <sup>h</sup> ɛ	pa✓ p <sup>h</sup> a	po p <sup>h</sup> o✓	pu p <sup>h</sup> u	puu p <sup>h</sup> uu		pao p <sup>h</sup> ao	pie✓ p <sup>h</sup> iɛ	pia✓ p <sup>h</sup> ia	pio✓ p <sup>h</sup> io✓					pio
m	mi	me	mɛ	ma	mo		miu	fu✓	!c	mie	mia	mio	miu				
f		!c	fe✓	!c	!c		!c										
v			vɛ	!c	!c	vu											
t	ti✓	te	tɛ	ta✓	to✓	tu	tuu✓	tu✓	!c	!c			tiu✓	tui✓		tua	
t <sup>h</sup>	t <sup>h</sup> i	t <sup>h</sup> e	t <sup>h</sup> ɛ	t <sup>h</sup> a✓	t <sup>h</sup> o	t <sup>h</sup> u	t <sup>h</sup> uu	t <sup>h</sup> u✓	t <sup>h</sup> ao	!c				t <sup>h</sup> ui	t <sup>h</sup> ue✓		
n	ni	ne	nɛ	na	no		niu	nu		!c	!c						
l	li			la	lo	lu	luu	lu	!c	!c	lia		!c	lui		lua	
ts	tsi* ts <sup>h</sup> i*	tse✓ ts <sup>h</sup> e✓	tse✓ ts <sup>h</sup> ɛ✓	tso✓ ts <sup>h</sup> o✓	tso✓ ts <sup>h</sup> o✓	tsu ts <sup>h</sup> u	tsuu✓ ts <sup>h</sup> uu	tsu✓ ts <sup>h</sup> u✓	!c	!c				tsui✓ ts <sup>h</sup> ui	tsue✓ ts <sup>h</sup> ue	tsua ts <sup>h</sup> ua✓	
s [ʒ]	si* tɕi*	se✓ tɕe✓	se✓ tɕɛ✓	so✓ tɕo✓	so✓ tɕo✓	su tɕu	sui✓ tɕuu	su✓ tɕu	!c	!c				sui✓ tɕui		sua✓	
tɕ	tɕi✓ tɕ <sup>h</sup> i✓	!c	tɕɛ✓ tɕ <sup>h</sup> ɛ✓	tɕa✓ tɕ <sup>h</sup> a✓	tɕo✓ tɕ <sup>h</sup> o✓		tɕuu✓							tɕui			
ɕ	ɕi✓		ɕɛ✓	ɕa✓	ɕo✓	ɕu	!c	ɕu✓	ɕao					ɕui✓	!c		





	Level	Falling	Rising
High	1{+P} 6		8
Non-high	2 {+P} 7	4 {+P}	
Low		5 {+P}	3 {+P}

{+P} Tones (non-modal voice)

- Tone 1 (66) High, level, tense voice
- Tone 2 (44) Non-high, level, tense voice
- Tone 3 (31) Low, falling, breathy voice
- Tone 4 (42) Non-high, falling, tense voice
- Tone 5 (21) Low, falling, harsh voice

Modal voice tones

- Tone 6 (55) High, level
- Tone 7 (33) Non-high, level
- Tone 8 (35) High, rising

Specification of the co-articulated features associated with, especially, tones 3 and 5 is due to the work of Jerold Edmondson as reported in Edmondson and Li 1994.

**FIGURE 40.2** JIANCHUAN BAI TONES

long in its duration, while tone 3 has breathy voice and is the shortest in duration. Tones 6, 7, and 8 have modal or clear voice.

Significant differences obtain between the system that has been used to indicate tones in standard data sources as published by the Chinese Academy of Social Sciences and the system currently employed in the recognized orthographic scheme as revised in 1993. These can be compared in Table 40.2, where each lexical item is transcribed using the two systems and in phonetic representation, with brief tonal descriptors and graphic contour symbols for positive identification.

### 2.2.2 *Literary and colloquial Chinese strata*

Standard sources refer to the specialization of a subset of the Bai tones for use in the reading of Chinese characters (Wiersma 1990: 106–9). According to one authority, the so-called *Han4zi4 bai2du2* (Chinese characters read in Bai) system is the byproduct of a longstanding tradition of bilingual teaching and learning in Bai speaking communities of the Dali region. It is now generally assumed that the earliest Chinese layer in Bai instantiates sounds and phonological distinctions that had become obsolete in Chinese by the third century AD (Starostin 1994: 2–8). On the other hand, Bai may plausibly have undergone a period of later contact with a conservative Chinese dialect resulting from fourteenth-century in-migrations from the Wu and Min areas. An overview of the two Chinese strata is shown in Table 40.3, where the

**TABLE 40.2. REPRESENTATION OF JIANCHUAN TONES AFTER ORTHOGRAPHIC REVISION**

<i>Tone number in sample data</i>	<i>Description and contour cover term</i>	<i>Example</i>	<i>Illustration</i>	<i>1984 orthography</i>	<i>1993 orthography</i>
Tone 1	High, level, tense (55)	attitude	$h^h \varepsilon^1 . u^1$	<i>tairl.dairl</i>	<i>taib.duib</i>
Tone 2	Non-high, level, tense (44)	lungs (pair)	$p^h i a^2 - k^h \varepsilon^2$	<i>piarx-kairx</i>	<i>piak-kai</i>
Tone 3	Low, falling, breathy (31)	tree	$tsu^3$	<i>zet</i>	<i>zet</i>
Tone 4	Non-high, falling, tense (42)	study	$y u^4$	<i>hhert</i>	<i>hhep</i>
Tone 4a	Non-high, falling, modal (32) (Dali)	mountain	$s u^4$ (Jianchuan)	<i>svrrt</i> (Jianchuan)	<i>svz</i> (Dali), <i>svp</i> (Jianchuan)
(Not in Jianchuan data)					
Tone 5	Low, falling, harsh (21)	meat	$k \varepsilon^5$	<i>gai</i>	<i>gaid</i>
Tone 6	High, level, modal (55)	salt	$p i^6$	<i>binl</i>	<i>binl</i>
Tone 7	Non-high, level, modal (33)	sleep	$ts^h \varepsilon^7$	<i>cainx</i>	<i>cainx</i>
Tone 8	High, rising, modal (35)	come	$y u^8$	<i>hhref</i>	<i>hhref</i>

*Notes:* In 1993 Tone 4a was introduced to the scheme, and the symbol 'r' removed, in order to incorporate Dali data into the corpus. Formerly, the symbol 'r' had served solely to mark non-modal (tense voice) phonation on tones 1, 2, and 4, making it possible for tones with similar contour to be symbolically paired (using the same final tone letters) and fewer tone symbols to be used. Removal of 'r' necessitated the introduction of new tone letters 'b', 'p', and 'd'. Note that the 'unmarked' tone was changed from tone 5 to tone 2 on the basis of claim to higher frequency. But note also that the earlier scheme and standard description (right-branching contour symbol ʌ) had implicitly indicated a special type of non-modal phonation for tone 5. Symbols remained the same for tones with modal phonation.

**TABLE 40.3** DISTRIBUTION OF CHINESE TONE GROUPS IN LITERARY AND COLLOQUIAL CHINESE STRATA OF JIANCHUAN BAI

Tone number in sample data	Description and contour cover term	Example	Illustration	Chinese correspondence (literary stratum)	Chinese correspondence (colloquial stratum)
Tone 1	High, level, tense (55)	attitude	$^h\epsilon^1$ , $tu^1$	IIIA, IIIB (modern)	IVA, IIIA (lungs)
Tone 2	Non-high, level, tense (44)	lungs (pair)	$p^hi\epsilon^2-k^he^2$	IIA, IIB	IIIB
Tone 3	Low, falling, breathy (31)	tree	$tsu^3$	IB	IVB
Tone 4	Non-high, falling, tense (42)	study	$yu^4$	IIIA, IIIB	IB
Tone 5	Low, falling, harsh (21)	meat	$ke^5$	IA	IA
Tone 6	High, level, modal (55)	salt	$pt^6$	IVA, IVB	IIA, IIB
Tone 7	Non-high, level, modal (33)	sleep	$ts^he^7$		
Tone 8	High, rising, modal (35)	come	$yu^8$		IIIB? (come)

*Notes:* The symbols IA, IB, IIA, IIB, IIIA, IIIB, IVA, and IVB refer to Chinese tone groupings based on traditional phonological categories; Yin Ping, Yang Ping, Yin Shang, Yang Shang, Yin Qu, Yang Qu, Yin Ru and Yang Ru respectively. The literary layer as reflected by the Jianchuan 'Hanzi baidu' system is basically a five-term system and thus comparable to Mandarin. The seven-term system reflected by the archaic Chinese stratum is possibly comparable to the system reported for e.g. the Yong'an variety of Western Min (Norman 1988: 238), where the old Rusheng group retains its split, and one non-Rusheng group has merged while the other two remain split.

Jianchuan tone reflexes are correlated with historical Chinese tone groups in the ‘literary’ (e.g. modern) and ‘colloquial’ (early) layers.

### 3 WORD CLASSES

The essentials of the Bai word classes may be considered in terms of their usual occurrence in either substantive or predicative expressions. Functors usually associated with nouns or behaving like nouns are classed here as substantives, and those normally occurring in verb phrase contexts or used to facilitate predication are classed as predicatives. There are inevitable compromises in this scheme, however. For example certain functions of the form /no<sup>7</sup>/ (e.g. COMP and NOMZR, as specified in 3.1.3.4 below) may be better illustrated as part of a verbal expression. Adjectives, though behaving syntactically as verbs in many contexts, often simply appear as part of a noun determiner. Sentence-final particles, which add semantic meaning globally to whole utterances, are treated separately.

#### 3.1 Substantives and determiners

One interesting characteristic of the Bai concrete noun is that its association with a related classifier morpheme is relatively close, such that the two morphemes often appear together as a kind of ‘false compound’, which may or may not be qualified by a {NUM+CL} string or other determiner expression. If such a noun form is quantified, this creates a type of ‘CL reduplication’, the CL occurring once attached to the noun and again within the quantifier phrase.

##### 3.1.1 Nouns and proper names

###### 3.1.1.1 Nouns

<i>ko</i> <sup>5</sup>	lake	<i>so</i> <sup>6</sup> . <i>kε</i> <sup>6</sup>	butcher	<i>pi</i> <sup>6</sup>	salt
<i>pi</i> <sup>5</sup> . <i>thv</i> <sup>6</sup>	taro	<i>khu</i> <sup>ε</sup> . <i>pi</i> <sup>6</sup>	thigh	<i>vu</i> <sup>7</sup>	rain

###### 3.1.1.2 Proper names

<i>ye</i> <sup>7</sup> <i>ko</i> <sup>5</sup>	Erhai Lake (in Dali BAP)	<i>yo</i> <sup>4</sup> <i>k<sup>h</sup>ε</i> <sup>3</sup>	Heqing County (east of Jianchuan)
<i>ku</i> <sup>4</sup> . <i>xu</i> <sup>3</sup>	Dali (the old city, not Dali BAP)	<i>tō</i> <sup>3</sup> <i>sv</i> <sup>7</sup> . <i>tci</i> <sup>1</sup>	Party Secretary Dong

###### 3.1.1.3 Locator nouns

<i>tse</i> <sup>4</sup> . <i>fv</i> <sup>7</sup>	right side	<i>ɣu</i> <sup>7</sup> . <i>no</i> <sup>7</sup>	back
<i>pi</i> <sup>6</sup> . <i>fv</i> <sup>7</sup>	left side	<i>tu</i> <sup>5</sup> . <i>ɳv</i> <sup>6</sup>	front
<i>tō</i> <sup>7</sup> . <i>fv</i> <sup>7</sup> . <i>no</i> <sup>7</sup>	top	<i>k<sup>h</sup>u</i> <sup>3</sup> . <i>no</i> <sup>7</sup>	inside
<i>ye</i> <sup>7</sup> . <i>fv</i> <sup>7</sup> . <i>no</i> <sup>7</sup>	bottom	<i>ɳua</i> <sup>2</sup> . <i>no</i> <sup>7</sup>	outside

###### 3.1.1.4 Time expressions

<i>k<sup>h</sup>ε</i> <sup>6</sup> . <i>tu</i> <sup>5</sup>	morning	<i>tci</i> <sup>5</sup> . <i>ji</i> <sup>2</sup>	yesterday
<i>pē</i> <sup>7</sup> . <i>kē</i> <sup>4</sup>	evening	<i>mē</i> <sup>6</sup> . <i>ji</i> <sup>2</sup>	tomorrow
<i>tse</i> <sup>5</sup> . <i>kε</i> <sup>2</sup>	time when	<i>nā</i> <sup>5</sup> . <i>tci</i> <sup>6</sup>	last year

## 3.1.2 Noun classifiers and numerals

The noun in Bai is normally associated with a corresponding classifier which may attach to the noun enclitically to form a non-plural determiner expression, whether definite or indefinite. Bare nouns do also occur, but the occurrence of the structure {noun + CL} in subject or object position without the addition of a determiner 'one' is enough to represent a single member of the category (if it represents new information), or a certain member in particular (if given, e.g. already identified).

## 3.1.2.1 Classifiers

CLASSIFIER	NOUN			
$tu^5$	$ke^6$	chicken	$ke^6-tu^5$	a/the chicken
$j\ddot{i}^5$	$j\ddot{i}^5.k\ddot{e}^6$	person	$j\ddot{i}^5.k\ddot{e}^6-j\ddot{i}^5$	a/the person
$ko^5$	$ko^5$	lake	$ko^5-ko^5$	a/the lake
$p^h\sigma^2$	$khu\epsilon^3.pi^6$	thigh	$khu\epsilon^3.pi^6-p^h\sigma^2$	a/the/my thigh
$ts^h u\epsilon^2$	$s v^6$	book	$s v^6-ts^h u\epsilon^2$	a/the book

## 3.1.2.2 Numerals

$ji^2; j\ddot{i}^8; \alpha^3$	one	$pi\alpha^2$	eight	$\zeta i^3.ne^2$	forty-two
$k\ddot{o}^7; ne^2; si^3$	two	$t\zeta iu^7$	nine	$p\epsilon^2$	hundred
$s\ddot{a}^6$	three	$ts\epsilon^4$	ten	$t\zeta^h i^6$	thousand
$\zeta i^2$	four	$ts\epsilon^4.ji^2$	eleven	$\eta v^4$	10,000
$\eta v^7$	five	$ni^3$	twenty	$\alpha^3.p\epsilon^2$	one hundred
$fv^2$	six	$\zeta i^3$	forty	$ne^2.p\epsilon^2$	two hundred
$t\epsilon^h i^2$	seven	$ni^3.ne^2$	twenty-two	$k\ddot{o}^7.\eta v^4$	20,000

## 3.1.2.3 Measure words

CLASSIFIER	NOUN			
$ya^4$	$so^6.to^5$	sugar	$so^6.to^5 \alpha^3-ya^4$	one pack of sugar (2 cakes)
$t^h\ddot{v}^3$	$vu^7.\zeta ui^7$	rain water	$vu^7.\zeta ui^7-t^h\ddot{v}^3$	a/the pail of rain water
$p\ddot{u}^6$	$\eta\epsilon^5$	go (verb)	$\eta\epsilon^5 \alpha^3-p\ddot{u}^6 k\ddot{o}^7-p\ddot{u}^6$	go (there) once or twice

## 3.1.3 Pronouns and noun particles

## 3.1.3.1 Pronouns

$\eta\sigma^3$	PN.1.SG	$\eta\alpha^6$	PN.1.PL.EX
		$j\ddot{a}^6$	PN.1.PL.1.INC
$no^3$	PN.2	$n\alpha^6$	PN.2.PL
$j\ddot{i}^6$	PN.2.SG.REV		
$mo^3$	PN.3.SG	$m\alpha^6$	PN.3.PL

## 3.1.3.2 Deictic pronouns

$lu^3$	this	$\alpha^6.t\alpha^2$	here	$lu^6 k\ddot{o}^7-ts^h u\epsilon^2$	those two (books)
$mu^3$	that	$mu^6.t\alpha^2$	there	$mu^6 k\ddot{o}^7-ts^h u\epsilon^2$	those two (books)

## 3.1.3.3 Interrogative pronouns

$\alpha^3.to^3$	who	$s\epsilon^3.su^7(ts^h u\epsilon^2)$	what sort of (book)
$\alpha^6.s\epsilon^3$	what, which	$tsi^6.k\epsilon^5(su\alpha^2)$	how to (say)

$a^6na^2$	where, which	$t\tilde{a}^6\tilde{j}^5$	how many (people)
$a^6.na^2.t^h\tilde{a}^6$	when	$s\varepsilon^3.tsu^4 ts^h u\varepsilon^2$	how many (books)

### 3.1.3.4 Noun particles

$no^7$	OBJ	$no^7$	NOMZR
$no^7; \eta v^6$	LOC	$no^7$	COMP
$\eta v^6$	BEN	$no^7$	SUB
$v\tilde{a}^4; \eta v^4$	POSS		

A few examples of the above particles in use may serve to illustrate.

### 3.1.3.5 Examples

$no^3$	$mu^6-no^7$	$t\tilde{c}i^4$	$mi^4$	$t\tilde{a}^4$	$lu^7$
PN.2.SG	PN.3.SG.GEN+SUB	chase	AUX.can	AUX.can	SPRT
‘You’ll be able to catch up with him [for sure].’					
$mo^3$	$sv^3$	$\alpha^3.ne^2-\eta v^6$	$tu^5po^5$		
PN.3.SG	comb (V)	Grandma+LOC	head+CL		
‘He’s combing Grandma’s hair [for her].’					

## 3.2 Predicatives (verbs and verbal adjuncts)

This section illustrates word classes more closely associated with the predicative function, either functioning alone as predicate or occurring in construction with other elements to form a predicate.

### 3.2.1 Verbs and adjectives

Verbs representing political or legal terms and other neologisms are generally borrowed from modern Chinese standard and their usage is generally consistent with Chinese, but many basic monosyllabic verbs can be identified as forming part of the early Chinese layer. Their usage in serial constructions reflects local norms. For some native speakers, this aspect of the language has represented evidence that the Bai people were originally of Chinese origin. For others, it has supported an argument that since the level of Chinese literacy among Bai speakers has historically been high, no vernacular orthography should be developed.

#### 3.2.1.1 Verbs

$ju^2$	eat	$s\tilde{e}^7$	know	$pe^2$	walk
$pi\varepsilon^2$	ask	$\tilde{e}^4$	swallow	$p^hi\tilde{a}^2$	arrive at
$\eta\varepsilon^5$	go	$ku^5$	sell	$yu^4$	study

#### 3.2.1.2 Adjectives

$pie^4$	bland	$\tilde{c}i^2.mi^4, \tilde{c}i^2.kv^2$	square	$yo^4.s\varepsilon^3$	easy	$ph\varepsilon^6$	soft
$k^h v^2$	bent; crooked	$to^4$	large	$tu\tilde{u}^7$	far	$x\tilde{e}^6$	raw
$\eta ui^5$	round	$p\tilde{e}^5$	flat	$t\tilde{c}i^7$	near	$ts^h o^3$	salty

A typical feature of Bai adjectives is shown by reduplication and combination with the subordinating particle /no<sup>7</sup>/ to form various types of relativized noun determiners and verbal adjuncts.

3.2.1.3 Use of -no<sup>7</sup>

<i>phɛ<sup>6</sup>.no<sup>5</sup>.no<sup>5</sup></i>	<i>tɕi<sup>3</sup>.no<sup>7</sup></i>	<i>tse<sup>6</sup>.pe<sup>7</sup>.k<sup>h</sup>v<sup>3</sup></i>
{soft+comfy+comfy}	SITUATION+SUB	bedding+CL
'a set of bedding [or, a bed] that's nice and soft'		
<i>k<sup>h</sup>uã<sup>6</sup>.k<sup>h</sup>uã<sup>6</sup>.no<sup>7</sup></i>	<i>pe<sup>2</sup></i>	
{slow+slow+COMP}	walk	
'walking slowly along'		

## 3.2.2 Auxiliaries, aspect words, coverbs, prepositions and adverbs

The following illustrations suggest the ordering of verbal adjuncts relative to their governing expression.

## 3.2.2.1 Auxiliaries (vV: post-head)

<i>tã<sup>4</sup></i>	can (permission to)	<i>xõ<sup>7</sup></i>	should not
<i>mĩ<sup>4</sup></i>	can (manage to)	<i>tu<sup>2</sup></i>	obtain, succeed
<i>tuã<sup>4</sup></i>	cannot; not OK	<i>juĩ<sup>7</sup></i>	not dare to

## 3.2.2.2 Aspect words (vV: post-verbal)

<i>...ko<sup>4</sup>...lã<sup>4</sup></i>	PRIOR EXPERIENCE	<i>...lã<sup>4</sup></i>	COMPLETED
<i>...tsi<sup>6</sup>.tɕi<sup>3</sup></i>	WHILE IN PROGRESS	<i>...k<sup>h</sup>u<sup>7</sup>.mu<sup>6</sup>.no<sup>7</sup></i>	PROGRESSIVE

## 3.2.2.3 Coverbs (vV: pre-head)

<i>jõ<sup>7</sup></i>	should; want to	<i>ẽ<sup>4</sup></i>	passively undergo
<i>sã<sup>6</sup></i>	mutual action	<i>kã<sup>2</sup></i>	brief action (trial, casual)

## 3.2.2.4 Coverbs (vV: post-head)

<i>ts<sup>h</sup>u<sup>6</sup></i>	finish by doing (DISPOSAL)	<i>phiã<sup>2</sup></i>	do to the extent of
<i>k<sup>h</sup>u<sup>7</sup></i>	start to do	<i>ke<sup>5</sup>.tsi<sup>6</sup></i>	do in a location

## 3.2.2.5 Prepositions (vV: preverbal with noun)

<i>sã<sup>8</sup></i>	from; via	<i>kã<sup>2</sup></i>	DISPOSAL
<i>tã<sup>2</sup></i>	together with	<i>sõ<sup>7</sup>; sẽ<sup>7</sup></i>	CAUSE; MAKE (someone do something)
<i>ko<sup>6</sup></i>	together with	<i>fũ<sup>7</sup></i>	BENEFACTIVE
<i>pu<sup>3</sup></i>	BENEFACTIVE		

## 3.2.2.6 Adverbs (vV: preverbal)

<i>tsã<sup>8</sup>.kẽ<sup>5</sup>.tsi<sup>7</sup></i>	entirely	<i>nu<sup>7</sup></i>	do without stopping; always be doing	<i>li<sup>6</sup></i>	too; also
<i>sẽ<sup>2</sup>.sẽ<sup>2</sup></i>	really	<i>tse<sup>2</sup></i>	still more; not yet	<i>mu<sup>6</sup>.tu<sup>5</sup></i>	just now

## 3.3 Sentence particles

Effects of some frequently encountered sentence-final particles are illustrated in Table 40.4. Not treated here is a corresponding set of topic-advancing words that occur only as utterance-initial functors and clause connectors (Wiersma 1990: 161–4).

**TABLE 40.4 SENTENCE PARTICLES IN A JIANCHUAN BAI TEXT**

Particle (sentence-final)	Descriptive meaning and frequency in text	Example	Translation
$la^4$	Inchoative (New situation) (9/90)	$lu^3 \cdot pu^6 \cdot ts^6 \cdot \eta o^3 \cdot ts^h u^7 \cdot j\ddot{o}^2 \cdot \xi a^3 \cdot p\acute{e}^1 \cdot f a^8 \cdot \eta o^3 \cdot ts^h u^7$	This time, I'm going to think of some way to fix him good!
$\xi a^7 \cdot la^4$ ( $\xi a^7$ ) ( $ua^7$ ) ( $\xi a^7$ )	Conjectural (5/90)	$la^8 \cdot mu^6 \cdot no^7 \cdot \eta u^7 \cdot p^h o^2 \cdot ti^6 \cdot tu^6 \cdot ts^6 \cdot mo^3 \cdot tsu^7$ $\eta u^7 \cdot ta^4 \cdot no^7 \cdot \acute{a}^7 \cdot j a^8 \cdot t\acute{c}^h i\ddot{u}^1 \cdot ts^h u^3 \cdot \xi \acute{a}^7 \cdot la^4$ $te^8 \cdot lo^7 \cdot mo^7 \cdot j\ddot{o}^2 \cdot tsu^6 \cdot mu^6 \cdot pi\acute{e}^6 \cdot k\acute{o}^6 \cdot ts^6 \cdot ta^4 \cdot lu^7$	Well, so since he was near-sighted, I suppose he couldn't see things too clearly. Yikes, well you could work for him all right, but ... [you would never get paid].
$lu^7$	Concessive (Gratuitous info) (2/90)	$la^8 \cdot a^6 \cdot ta^2 \cdot ts^6 \cdot mo^3 \cdot ts^6 \cdot k\acute{e}^5 \cdot to^5 \cdot ni^6$	Well now, how do you think he did it?
$ni^6$	Rhetorical Question (1/90)	$mo^3 \cdot t^h u^7 \cdot \acute{a}^7 \cdot tu^2 \cdot ta^4 \cdot t\acute{c}i^7 \cdot ts^h u^7 \cdot ne^5$	He could only see things close up.
$ne^5$	Dismissive (1/90)	$\eta u^7 \cdot f\ddot{u}^7 \cdot no^7 \cdot mo^3 \cdot k^h a\acute{o}^3 \cdot t\acute{c}i\ddot{u}^1 \cdot si^1 \cdot li^6 \cdot ts^3 \cdot iso^7$	Later on, he even got a 'jinshi' degree in the imperial examinations!
$ts^3 \cdot iso^7$	Assertion (1/90)	$ke^6 \cdot ts^6 \cdot mo^3 \cdot ts^h u^7 \cdot k\acute{u}^8 \cdot k\acute{u}^8 \cdot su^7 \cdot x\acute{e}^6 \cdot m\acute{e}^5$	If it were a chicken, why, it would have to cry out 'cluck-cluck-cluck'.
$su^7$ ( $s\acute{e}^7$ ) ( $su^7$ )	Inferential (2/90)	$su^7$	
$pi\acute{e}^4$ ( $pi\acute{e}^4$ )	Imperative (Often reflexive) (2/90)	$la^8 \cdot li^3 \cdot la\acute{o}^3 \cdot ts^h \acute{e}^4 \cdot tsu^3 \cdot \acute{a}^3 \cdot t\acute{c}^h \acute{e}^6 \cdot tu^2 \cdot ts^h u^7 \cdot su\acute{a}^2$ $tsu^7 \cdot t\acute{c}i^4 \cdot tsu\acute{a}^4 \cdot t\acute{c} \acute{e}^2 \cdot t\acute{c}^h \acute{a}^3 \cdot ju^2 \cdot su^2 \cdot mo^3 \cdot pi\acute{e}^4$	So then, as soon as Master Li heard this, he said to himself, 'I'd better grab it and eat it all up quick!'

Notes: The examples shown are drawn from a single narrative (Wiersma 1990: 242-6). The text is segmented into about ninety utterances, of which twenty seven end with particles of the type illustrated. The symbol '9/90' means, then, that one third of the utterances ending with sentence particles used this particle in the same way as illustrated. All data have been checked against Zhao and Xu 1996.



## 4 WORD FORMATION PROCESSES

Bai word formation, along with other aspects of Bai lexicology, has received considerable attention in standard sources published in Chinese over the last twenty years. Users of these standard sources may be puzzled, however, by what may seem to be conflicting claims appearing in publications representing the ongoing work of the same authors: for example, the claim that the ‘inherent’ or native Bai vocabulary consists mainly of monosyllabic words (Xu and Zhao 1984), as against the claim that most native Bai words are polysyllabic forms (Zhao and Xu 1996). It is not safe to assume that all publications appearing after 1980 are updated manifestations of earlier works whose appearance had been simply postponed for political reasons (Wiersma 2001: 24). Despite noteworthy achievements, Bai lexicology and lexicography are still in their infancy, not least because of the significant penetration of Chinese and because loans from the ‘early’ Chinese stratum collocate unexceptionally with autochthonous morphemes according to local word-formation patterns. Chinese loans are also in some sense present in Bai word forms that invite analysis as calques of corresponding Chinese terms.

### 4.1 Inflectional morphology

Inflection for case is present in Bai but its effects are seen only in personal pronouns and only the genitive case is overtly marked. However the subordinating/object-marking particle /no<sup>7</sup>/ which has received attention in recent typological studies of TB languages should be mentioned in connection with genitive marking also, because of the role it plays in forming what might be termed a syntactic genitive and the possible connection of this role to semantic patient marking. This is an interesting area of the grammar where morphology and syntax are apparently intertwined, as will be briefly discussed below.

#### 4.1.1 Pronoun inflection

The Bai personal pronouns at first glance seem anomalous in the language, in that they are subject to systematic changes in form to mark genitive or possessive meaning through rhyme and tone shift. However it is also the case that alternation occurs in other word classes, notably in coverbs related to permission, ability and knowing. Some other cases of morpheme ‘fusion’ can also be found in Bai, for example as occurring in the deictic pronouns illustrated above. In addition to encoding the genitive as opposed to the default case, the Bai personal pronouns also inflect for other contrasts, e.g. singular/plural and exclusive/inclusive. The Bai personal pronoun system is illustrated in Table 40.5, with suppletive forms shown in separate cells.

**TABLE 40.5** CASE INFLECTION OF PERSONAL PRONOUNS IN JIANCHUAN BAI

		<i>First person</i>	<i>Second person</i>	<i>Third person</i>
Singular	default case	<i>ŋo<sup>3</sup></i>	<i>no<sup>3</sup></i>	<i>mo<sup>3</sup></i>
	polite		<i>ǰi<sup>6</sup></i>	
Plural	genitive case	<i>ŋu<sup>6</sup></i>	<i>nu<sup>6</sup></i>	<i>mu<sup>6</sup></i>
	exclusive first	<i>ŋa<sup>6</sup></i>	<i>na<sup>6</sup></i>	<i>ma<sup>6</sup></i>
	inclusive first	<i>ǰã<sup>6</sup></i>		

TABLE 40.6 INCIPIENT SEMANTIC MARKING IN JIANCHUAN BAI

$mo^3 fe^2. iv^4 \eta u^6. no^7$	He's angry at me.
$ku^6 si^3 \eta o^3 mu^6. p^h a^2$	Spoon out some food without sauce [dregs] for me.
$ma^6 fe^2. si^3 \eta a^6. sv^6$	They gave us books.
$no^3 j\ddot{o}^2 \eta u^6. no^7 x\ddot{a}^6. ts\ddot{e}^5 a^6. s\ddot{e}^7$	What do you take me for [see me as, make me into]?
$\eta o^3 ta^4. nu^6. yu^4 k\ddot{e}^6$	I'm worried about you [help you by worrying].
$no^3 mu^6. no^7 t\ddot{c}i^4 mi^4 ta^4 lu^7$	You'll be able to catch up with him [for sure].
$mu^6. no^7 tsu^7 ta^2 p^h \epsilon^6. ts\ddot{i}^6 t\ddot{c}^h i^2 no^7$	S/he's a product of the party [trained up by the party].
$a^3. k^h u\ddot{a}^7. ts\ddot{i}^7 t\ddot{e}^2 ma^6. no^7 tso^4. k^h u\ddot{t}^6$	Little Dog threw stones at them [hit them with stones].
$a^3. mo^7 ts\ddot{e}^2 j\ddot{a}^8 tsu^6 k^h u\ddot{t}^7. mu^6. no^7 ts^h \ddot{a}^6$	Mama hasn't yet started to cook supper.
$li^4 tsu^7 nu^6. va^4 n\epsilon^6$	Is all this stuff yours?

#### 4.1.2 Extended function of genitive marking

Appearing sporadically in published sources, and often in data from the field, are anomalous uses of the genitive third singular pronoun. Frequently the form /mu<sup>6</sup>/ appears alone and gives the impression of playing a role as a hesitation marker. However in certain cases this pronoun form combines syntactically with the locative/subordinating particle /no<sup>7</sup>/ to form a construction that can introduce a patient, undergoer or goal noun, and this usage resembles to some extent the 'mock object' and/or 'possessive object' construction of Mandarin Chinese (Chao 1968: 291–2, 321, 430–1). A brief sampling of the data supporting these observations is shown in Table 40.6. It is interesting that while LaPolla has pointed out extensions of locative case to patient, goal, or anti-ergative marking functions in various TB languages, and extensions of genitive case to agentive marking; no connection between genitive case and patient or goal marking has apparently been observed in any of the languages studied in these initiatives. However the proposal that 'non-systemic marking' of semantic roles through extension of existing case-marking devices, as found in many TB languages, is a manifestation in these languages of an inherited tendency that is grounded in a 'semantically based system of grammatical organization' (as opposed to one based on syntactic relations, e.g. the Chinese model) may offer us a key to understanding what is going on in Bai (LaPolla 1995: 189–90, 210–18).

## 4.2 Derivation

A certain number of polysyllabic word forms in Bai incorporate regular and at least partly productive bound suffixes. Also certain reduplicated forms involve what appear to be a derivational application of the subordinating noun particle /no<sup>7</sup>/ already mentioned at several points above.

#### 4.2.1 Affixation and negative infixing

A few suffixes attach enclitically to eligible nouns and are used productively to derive like sets of terms. These are perhaps not derivational suffixes in the strictest sense, in that they do not generally alter the part of speech of the suffixed noun. Still this process is clearly a type of word formation and may be treated as derivation.

$xo^2$	AGGREGATE/PLURAL (HUMAN)	$x\ddot{a}^4. xo^2$	Chinese (people as a group, or a single member of the group)
		$yo^4. k^h \epsilon^3. xo^2$	people (or a person) from Heqing

<i>po</i> <sup>6</sup>	MALE/RESPECT/LARGE SIZE	<i>tsã</i> <sup>7</sup> <i>sã</i> <sup>7</sup> <i>po</i> <sup>6</sup> <i>ke</i> <sup>6</sup> . <i>po</i> <sup>6</sup>	Zhang San! rooster
<i>mo</i> <sup>7</sup>	FEMALE/RESPECT/LARGE SIZE	<i>ji</i> <sup>6</sup> . <i>tã</i> <sup>6</sup> . <i>mo</i> <sup>7</sup> <i>ke</i> <sup>6</sup> . <i>mo</i> <sup>7</sup> - <i>tu</i> <sup>5</sup>	scythe (big knife for firewood) a/the hen
<i>ja</i> <sup>4</sup>	AGGREGATE/PLURAL/CL (NON-HUMAN/ABSTRACT)	<i>se</i> <sup>3</sup> . <i>vu</i> <sup>7</sup> - <i>ja</i> <sup>4</sup> <i>ɳv</i> <sup>7</sup> . <i>tã</i> <sup>4</sup> <i>lu</i> <sup>3</sup> - <i>ja</i> <sup>4</sup>	business matters, affairs these things, this stuff

The interesting pattern of negative infixing in Bai generally occurs in sentence-final position, and this may be related to the fact that such infixing is a property of partly grammaticized verbs which typically act as coverbs in a sentence. A few verbs that undergo this change, however, can act as the main verb in a predicate, especially the verbs 'know' and 'see.' Nevertheless, the infixed form still generally occurs sentence-finally, and this seems to be one context where the preferred sentence order is for a patient or goal noun to occur preverbally, followed by the appropriate object-marking particle.

<i>tsẽ</i> <sup>5</sup>	become, succeed	<i>tsuẽ</i> <sup>5</sup>	not succeed
<i>tã</i> <sup>4</sup>	able to, OK	<i>tud</i> <sup>4</sup>	not permitted to, unable to
<i>sẽ</i> <sup>7</sup>	know	<i>sũ</i> <sup>7</sup>	not know
<i>kẽ</i> <sup>4</sup>	see	<i>kũ</i> <sup>4</sup>	not see

#### 4.2.2 Reduplication and elaboration

Reduplicative forms include simple forms where complete copying denotes repeated action, intensifies or otherwise extends meaning. More complex forms are also found where reduplication combines with affixation to derive an adverbial or adjectival 'manner' expression. Additionally, Bai has its own complement of four-syllable elaborate expressions. The typical form of such expressions in Bai is built up through insertion of what is probably an epenthetic syllable after each instance of a pair of semantically related terms. Such expressions belong to a literary register most often encountered in artistic oral performances. Elaborate expressions without epenthetic insertion are also frequently found, however, where each syllabic element is subject to morphemic analysis.

Reduplication: repeated action, intensification, quality:

<i>ɳe</i> <sup>5</sup> . <i>ɳe</i> <sup>5</sup> . <i>ja</i> <sup>2</sup> . <i>ja</i> <sup>2</sup>	going back and forth	{go + go + return + return}
<i>l<sup>h</sup>u</i> <sup>6</sup> . <i>l<sup>h</sup>u</i> <sup>6</sup> . <i>tsõ</i> <sup>7</sup> . <i>tsõ</i> <sup>7</sup>	going up and down	{down + down + ascend + ascend}

Combining reduplication and affixation: manner, quality:

<i>mi</i> <sup>4</sup> . <i>mi</i> <sup>4</sup> . <i>su</i> <sup>7</sup>	so dark	{obscure + obscure + SEEM.AFF}
<i>mo</i> <sup>4</sup> . <i>mo</i> <sup>4</sup> - <i>no</i> <sup>7</sup>	(cut it) really fine	{fine + fine + SUB}

Four-syllable elaborate expressions:

<i>ji</i> <sup>4</sup> . <i>li</i> <sup>4</sup> . <i>khv</i> <sup>2</sup> . <i>lv</i> <sup>2</sup>	all crumpled up	{twisted + X + crooked + X}
<i>ko</i> <sup>6</sup> . <i>lo</i> <sup>6</sup> . <i>e</i> <sup>7</sup> . <i>le</i> <sup>7</sup>	dull-witted	{stupid/hollow? + X + mute + X}
<i>ci</i> <sup>2</sup> . <i>mi</i> <sup>4</sup> . <i>ci</i> <sup>2</sup> . <i>ko</i> <sup>2</sup>	square	{four + eye + four + horn}

### 4.2.3 Compounds

Finally, many polysyllabic word forms in Bai neither involve reduplication nor fit the model of elaborate expressions, and these are generally analysed by standard sources in terms of particular syntactic combinations. Binomial compounds are frequent, where two nouns of similar or related meaning combine to derive a third noun (e.g. ‘mouth’). In other compounds, the pattern is comparable to {subject+predicate}, as in ‘greedy’. Another frequent pattern is captured by the {head+modifier} model, as in ‘noodles’. As may be observed in the following examples, the ordering of elements in such compounds is not consistently pre- or post-head.

<i>tɕui<sup>7</sup>.ŋui<sup>7</sup></i>	mouth	{mouth + eye}
<i>ts<sup>h</sup>ã<sup>6</sup>.pẽ<sup>7</sup></i>	meal	{breakfast + supper}
<i>ŋui<sup>7</sup>.xu<sup>2</sup></i>	greedy	{eye + black}
<i>xẽ<sup>6</sup>.miẽ<sup>4</sup></i>	the dark	{sky + obscure}
<i>xẽ<sup>6</sup>.t<sup>h</sup>ẽ<sup>2</sup></i>	thunder	{sky + split}
<i>mo<sup>2</sup>.mi<sup>4</sup>.xẽ<sup>6</sup></i>	noodles	{flour + raw}
<i>fẽ<sup>2</sup>.mo<sup>2</sup>.mi<sup>4</sup></i>	dough	{develop + flour}

## 5 SYNTAX

Flexibility with respect to certain word orders, sporadic marking of grammatical relations, and transparency of certain Bai utterances as calques of their Chinese gloss all speak to the problematic nature of the notion that study of Bai syntax will uncover the ‘original’ Bai grammar. We know that the language has undergone long-term pressure from Chinese, and flexible or mixed word orders and optional/partial marking of relationships are, indeed, features of inscriptional or literary Chinese. Nevertheless, as more and more data from Tibeto-Burman languages are examined comparatively, we see that exactly these features have assumed prominence in an emergent Tibeto-Burman typological framework, within which independent development of similar structures, irrespective of cognacy, is cited as evidence of parallel historical ‘drift’ among the related languages (Egerod 1985: 98–9; LaPolla 1994: 63–8; LaPolla 1995: 214–18). This being the case, the possibility that certain Bai syntactic features may be identified as similar to Chinese should not gainsay their value as linguistic evidence. On the contrary, study of such features in context may in time lead us to understand the mechanisms of contact and penetration whereby ‘relexification’ of a local vernacular by Chinese could possibly have taken place.

### 5.1 Noun phrase structures

From data already cited above we infer that the following orders are typical and may co-occur:

- {N + CL}
- {ADJ + N}

A fully realized noun phrase based on a single substantive then might look something like the following:

((REL/ADJ/ADV) + (SUB)) + (ADJ) + N + ((CL) + ((DET) + (NUM) + (CL)))

A few examples may illustrate:

### 5.1.1

*ts<sup>h</sup>ɛ<sup>2</sup>    thu<sup>3</sup>    tsʊ<sup>2</sup>*  
 red    road    CL  
 ‘a red (dirt) road’

### 5.1.2

*tsu<sup>3</sup>    mu<sup>6</sup>    fʊ<sup>2</sup>    tsu<sup>3</sup>*  
 tree    Dem    NUM.six    CL  
 ‘those six trees’

### 5.1.3

*mu<sup>6</sup>            ku<sup>7</sup>.jo<sup>4</sup>            ji<sup>5</sup>*  
 PN.3.SG.GEN    wife (old.lady)    CL  
 ‘his wife’

### 5.1.4

*ji<sup>5</sup>    xu<sup>3</sup>    no<sup>7</sup>    tɕi<sup>6</sup>.so<sup>2</sup>            ja<sup>4</sup>*  
 boat    inside    SUB    gold.hammers    CL.several  
 ‘the golden hammers that were inside the boat’

### 5.1.5

*li<sup>a</sup>.su<sup>7</sup>    to<sup>4</sup>    no<sup>7</sup>    tɕi<sup>6</sup>.ku<sup>a</sup>    k<sup>h</sup>o<sup>7</sup>*  
 like.that    big    SUB    pumpkin    CL  
 ‘such a big pumpkin’

## 5.2 Verb phrase structures

Space does not allow full discussion on ordering of verbs and verbal adjuncts here, but the remaining sections may serve to illustrate some of the typical patterns.

### 5.2.1 *Word order of predicate and arguments*

According to one authority, the obligatory order for the Dali dialect is simply *SVO*, and this is said to be attributable to greater penetration from Chinese in that dialect, while *SOV* ordering may sometimes be found in the Jianchuan dialect, especially in sentences where the object is a proper name or personal pronoun and it is followed by an object marker of the form /no<sup>7</sup>/ or /ŋʊ<sup>6</sup>/. Also according to this source, such verb-final ordering is the default order in Jianchuan interrogative sentences, regardless of whether an object marker appears. It follows then that in the case of a Jianchuan double object sentence, the human undergoer object will normally be preposed or even exposed to the left of the subject (Wiersma 1990: 194–201).

## 5.2.1.1

$\eta\sigma^3$   $m\epsilon^4$   $l\alpha^4$   $s\nu^6$   $\eta\nu^7$   $ts^h u\epsilon^2$   
 PN.1.SG buy ASP book 5 CL  
 'I've bought five books.'

## 5.2.1.2

$ku^7.p\sigma^6$   $j\tilde{\sigma}^5$   $t\epsilon^h \tilde{\epsilon}^6$   $ku^7.j\sigma^4$   $j\tilde{\sigma}^5$   $mu^6$   $t\tilde{\sigma}^5$   
 old.man CL listen to old.lady CL PN.3.SG.GEN talk  
 'The old man listened to what his wife said.'

## 5.2.1.3 Human undergoer object

$\alpha^3.ne^2$   $su\tilde{a}^6+x\sigma^2$   $no^7$   $li^6$   $ko^5$   $lu^7$   
 KIN.FM SUB SPRT  
 grandma grandkids also love  
 'Grandma loves the grandchildren [and that's a fact].'

## 5.2.1.4 'Disambiguation' with 2 objects

- (i)  $\eta\sigma^3$   $si^3$   $l\alpha^4$   $mu^6$   $no^7$   $\eta\epsilon^5$   $\alpha^3+t\epsilon\tilde{\sigma}^7$   
 PN.1.SG give ASP PN.3.SG.GEN SUB shoes {1+CL}
- (ii)  $\eta\sigma^3$   $mu^6$   $no^7$   $si^3$   $l\alpha^4$   $\eta\epsilon^5$   $\alpha^3+t\epsilon\tilde{\sigma}^7$   
 PN.1.SG PN.3.SG.GEN SUB give ASP shoes {1+CL}
- (iii)  $mu^6$   $no^7$   $\eta\sigma^3$   $si^3$   $l\alpha^4$   $\eta\epsilon^5$   $\alpha^3+t\epsilon\tilde{\sigma}^7$   
 PN.3.SG.GEN SUB PN.1.SG give ASP shoes {1+CL}
- 'I've given her a pair of shoes.'

## 5.2.1.5 Verb-final order with 'resumptive genitive' goal/patient marker

$\epsilon u^6.su^6+x\sigma^2$   $no^7$   $l\alpha\sigma^3.si^6$   $t\alpha^2$   $m\alpha^6$   $no^7$   $u\alpha^4.s\tilde{\epsilon}^4$   
 SUB PREP PN.3.PL/GEN SUB  
 student+PL teacher together play  
 'Teacher horsed around with the students.'

## 5.2.1.6 Verb-final order with NEG and Q

$\eta\sigma^3$   $\alpha^3.\eta\nu^7.si^6$   $no^7$   $tu^7$   $mi^4$   $tu\alpha^4$   
 PN.1.SG KIN.WES SUB AUX {AUX+NEG}  
 Auntie wait able cannot  
 'I cannot wait for Auntie.'

## 5.2.2 Serial verb constructions

The question of whether juxtaposed verbs in Bai represent serialization or concatenation can probably be resolved in favour of serialization based on criteria well established for Loloish (Matisoff 1973: 199–265; Hansson 1985: 287–303; Wheatley 1985: 403–17).

## 5.2.2.1

$\alpha^3.ne^2$   $\eta\epsilon^5$   $t\epsilon^h i^2$   $k\alpha^2$   $pe^2$   
 KIN.FM AUX AUX

grandma go out brief/try walk/leave  
 ‘Grandma’s going out for a little walk’.

## 5.2.2.2

$t\alpha^4 + \eta u^6 + \gamma u^4$	$k\epsilon^3$	$tsi^7$	$\gamma u^8$
{trample + PN.1.SG.GEN + strength}	{VCL.a.bit}	{AUX < lick}	{AUX < come}
help.me	a.bit	do.once; try.to	why.not; come.on

‘How about trying to give me a hand for once?’

## 5.2.2.3

$p^h i \alpha^2$	$lu^3.ku^6$	$li^6$	$tse^2$	$x\tilde{\alpha}^6$	$tu^2$	$mu^6 + ts\upsilon^6$
	{this + ladle;skein}		ADV		AUX	{PN.3.SG.GEN + track}
arrive	now	also	still	see	obtain	traces

‘Until today you can still see traces of it.’

## 5.2.2.4

$j\tilde{\alpha}^6$	$vu^7.k^h\upsilon^6$	$ts^h u^2$	$s\tilde{\alpha}^6$	$mu^7$	$k\epsilon^3$	$tsi^3.t\alpha^4$
PN.1.PL.INC		CL	AUX		VCL.bit	QPRT
	backpack		each.other	exchange	a.bit	ok?

‘How about if we exchange backpacks for a bit?’

5.2.3 *Negation in serial verb constructions*

Whether a given expression can be directly preceded by a negative morpheme has sometimes been used as a criterion for distinguishing main verbs from grammaticized verbs. While this is no doubt a helpful test for Bai, the following examples show that there are at least two critical points in a Bai serial verb construction where negation may be expressed. Pre-head position may be the default, but we see that negation also interacts with certain coversbs in sentence final position.

## 5.2.3.1

$\alpha^3.mo^7$	$tse^2$	$j\alpha^8$	$tsu^6$	$k^h u^7 + mu^6.no^7$	$ts^h \tilde{\alpha}^6$
	ADV	NEG		ASP	
mum	still		do/make	now.in.progress	breakfast

‘Mum hasn’t started to make breakfast yet.’

## 5.2.3.2

$\gamma\alpha^8.mu^7.tsi^6$	$mo^3$	$ju^2$	$j\alpha^8$	$tu^2$	$la^4$
	PN.3.SG		NEG	AUX	ASP
otherwise		eat		obtain	

‘Otherwise, he would not be able to eat (it).’

## REFERENCES

Armijo-Hussein, Jacqueline (1996) ‘Sayyid Ajall Shams al-Din: a Muslim from Central Asia, serving the Mongols in China, and bringing “civilization” to Yunnan’, Unpublished PhD dissertation, Harvard University.

- Benedict, Paul (1981) 'Chinese and Proto-Tai, part II', Presented to the 15th International Conference on Sino-Tibetan Language and Linguistics, Beijing.
- Bradley, David (1982) 'Register in Burmese', in David Bradley (ed.) *Tonation*, 117–32, *Papers in Southeast Asian Linguistics* no. 8, *Pacific Linguistics*, Series A, no. 62.
- Chao, Yuen Ren (1968) *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.
- Dai Qingxia (1980) 'A study of tense and lax vowels in Tibeto-Burman'. In Zhang Yuansheng, Zhang Gongjin and Chen Qiguang (eds) *Academic Papers from the Central Institute of Nationalities*, 71–104, Beijing: [Central Institute of Nationalities] (in Chinese).
- Dell, François (1981) *La langue Bai: phonologie et lexique*. Paris: Ecole des Hautes Etudes en Sciences Sociales.
- Edmondson, Jerold and Li Shaoni (1994) 'Voice quality and voice quality change in the Bai language of Yunnan province'. *Linguistics of the Tibeto-Burman Area* 17.2: 49–68.
- Egerod, Soren (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*, 96–104, *Pacific Linguistics*, Series C, no. 87.
- 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*, 287–309, *Pacific Linguistics*, Series C, no. 87.
- LaPolla, Randy (1994) 'Parallel grammaticalizations in Tibeto-Burman languages: evidence of Sapir's "drift"'. *Linguistics of the Tibeto-Burman Area* 17.1: 61–80.
- (1995) '"Ergative" marking in Tibeto-Burman'. In Yoshio Nishi, James A. Matisoff and Yasuhiko Nagano (eds) *New horizons in Tibeto-Burman morphosyntax*, 189–228. *Senri Ethnological Studies* no. 41, Osaka: National Museum of Ethnology.
- Lee Yeon-Ju and Sagart, Laurent (1998) 'The strata of Bai', Presented to the 31st International Conference on Sino-Tibetan Languages and Linguistics, Lund (handout).
- Li, Charles N. (1983) 'Languages in contact in western China'. *Papers in East Asian Languages* 1: 31–51.
- Liu, Godfrey F.K. and Wang, William S-Y. (1993) 'Baima feima: a case of folk etymology'. *Journal of Chinese Linguistics* 24.1: 128–37.
- Matisoff, James A. (1973) *The grammar of Lahu*. *University of California Publications in Linguistics* vol. 75. Berkeley, CA: University of California Press. (1991) 'Sino-Tibetan linguistics: present state and future prospects'. *Annual Review of Anthropology* 20: 469–504.
- Norman, Jerry (1988) *Chinese*. Cambridge, England: Cambridge University Press.
- Starostin, S.A. (1994) 'The historical position of Bai'. Presented to the 27th International Conference on Sino-Tibetan Languages and Linguistics, Paris.
- Thurgood, Graham (1996) 'Language contact and the direction of internal drift: the development of tones and registers in Chamic'. *Language* 72.1: 1–31.
- Ting Pang-hsin (1983) 'Derivation time of colloquial Min from archaic Chinese'. *Bulletin of the Institute of History and Philology* 54.4: 1–14.
- Wheatley, J.K. (1985) 'The decline of verb-final syntax in the Yi (Lolo) languages of southwestern China'. In Graham Thurgood, James A. Matisoff and David Bradley (eds) *Linguistics of the Sino-Tibetan Area: The State of the Art*, 401–20. *Pacific Linguistics*, Series C, no. 87.
- Wiersma, Grace (1990) 'Investigation of the Bai (Minjia) language along historical lines'. Unpublished PhD dissertation, University of California, Berkeley.
- (2001) 'The Bai (Min Chia) language: a hybrid vernacular of Yunnan'. Unpublished manuscript.
- Xu Lin and Zhao Yansun (1964) 'Outline of the Bai language'. *Zhongguo Yuwen* 4: 321–35 (in Chinese).
- (1984) *Baiyu jianzhi (Survey grammar of the Bai language)*. Beijing: Minzu (in Chinese).
- Yokoyama Hiroko (1994) 'Uxorilocal marriage among the Bai of the Dali Basin, Yunnan'. In Suenari Michio, J.S. Eades and Christian Daniels (eds) *Perspectives on Chinese Society: Anthropological Views from Japan*, 182–90. Tokyo: Institute for the Study of the Languages and Cultures of Asia and Africa.
- Zhao Yansun and Xu Lin (1996) *Bai/Han cidian (A Bai/Chinese word glossary)*. Chengdu: Sichuan Minzu (in Chinese).



# DULONG

*Randy J. LaPolla*

## 1 INTRODUCTION

Dulong [tʃrùŋ] is a Tibeto-Burman language spoken in China, closely related to the Rawang language of Myanmar (Burma). 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 (pop. 5816 according to the 1990 census), or to one part (roughly 6000 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. 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, Third Township, Fourth Township, and Nujiang. In this chapter, we will be using data of the First Township dialect spoken in Gongshan county. See LaPolla 2000 for discussion of the wider affiliations with the Rawang dialects and Anong. Other sources on Dulong, Rawang, and Anong are listed at the end of this chapter. The affiliations of Dulong beyond Rawang and Anong are not yet clear. All three are often put together with Jinghpaw, but this connection does not seem convincing on morphological grounds. More work needs to be done before any conclusion can be reached.

## 2 PHONOLOGY

Dulong has twenty-four initial consonants at six points of articulation (Table 41.1), plus the consonant clusters /pr, br, mr, kr, xr, gr, pl, bl, ml, kl, gl/ in initial position; only the consonants /p, t, ʔ, k, n, m, ŋ, 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 often are pronounced as voiceless stops, e.g. *wən* ‘buy’ > *wət-çü* ‘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, e, ə, a, ɔ, u, u/, and there are three diphthongs, /əi, ai, uɪ/, which only appear in open syllables. The syllable can be CV (*bà* ‘thin’), CVC (*çəm* ‘iron’), 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. *dǎzī* ‘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

TABLE 41.1 THE DULONG CONSONANTS

	<i>Labial</i>	<i>Dental</i>	<i>Retroflex</i>	<i>Palatal</i>	<i>Velar</i>	<i>Labio-velar</i>	<i>Glottal</i>
voiceless stop	p	t		c	k	kw	ʔ
voiced stop	b	d		ɟ	g	gw	
voiceless affricate		ts		tɕ			
voiceless fricative		s		ç	x	xw	
voiced fricative		z		ʒ			
nasal	m	n		ɲ	ŋ	ŋw	
			r				
voiced lateral approximant		l		j			

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 numbers 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 the 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 41.2), demonstrative pronouns (*ɟà* ‘this’, *ʒjà* ‘that’, *kɔ̄* ‘that (remote)’); used as free pronouns and demonstrative adjectives), and interrogative pronouns (*tāŋ(mā)* ‘what’; *ʒni* ‘who’, *kār wā* ‘how’, *kā (dɔ̄)* ‘where’, *tək cā* ‘why’, *kāpē* ‘which’). The latter can also be used as indefinite pronouns, as can *ʒtsəŋ* ‘person’, as in *ʒtsəŋ mǎ-zəŋ-c̄u* [person NEG-enter-R/M] ‘Nobody came in’. There is an emphatic, but not reflexive, pronoun, *ədù*, that can be used to emphasize the agency of the actor (He did it HIMSELF). Many nouns are formed using the prefix *ʒŋ-~ ʒk-*, e.g. *ʒk-saʔ* ‘breath’, ‘steam’, *ʒŋ-ɟà* ‘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 (*rì* inanimate) or group (*maʔ* human; *rà* animate or inanimate) marker. A numeral plus classifier can occur either before or after the head, but a classifier used alone must follow the head, and be used in conjunction with the demonstrative, which precedes the head. Unlike in Rawang, where a noun

TABLE 41.2 THE DULONG PERSONAL PRONOUNS

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
1	ʒgɔ̄	ʒjũŋnī	ʒjũŋ
2	nà	nũŋnī/nǎjũŋ/nǎnī	nǎjũŋ
3	àŋ	əŋnī	əŋjũŋ

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’ (*pɛ*) 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 (*ǎ-/nǎ-/ǎŋ-*) for kinship and relational terms (e.g. *ǎkàŋ* ‘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 *ǎŋ-~ǎk-*. If the adjective follows the head, the prefix is optional. Compare *zǎjè (ǎk)-sār* and *ǎk-sār zǎjè* ‘new book’. In many cases the adjective can be reduplicated as well, e.g. *mrəŋ mrəŋ lūŋ* ‘long stone’. Adjectives also frequently appear as pre-modifiers in the form of the relative clause construction, e.g. *təi tɕi zəŋ gũ kām* [very hard REL bamboo] ‘very hard bamboo’. Following are some more noun phrase examples:

- |   |   |
|---|---|
| (1) a. <i>zǎjè ǎŋi pəŋ</i><br>book two CL<br>‘two books’                        | b. <i>ǎgə zǎjè ri</i><br>1sg book pl<br>‘my books’                |
| c. <i>ǎgə zǎjè cɛʔ pəŋ</i><br>1sg book one CL<br>‘one of my books’              | d. <i>ǎtsəŋ təi cɛ ʃɔʔ</i><br>person big one CL<br>‘a big person’ |
| e. <i>kə təi zǎjè ǎsūm pəŋ</i><br>that big book three CL<br>‘those 3 big books’ | f. <i>ʃə pɛ pəŋ zǎjè</i><br>that NS CL book<br>‘that book’        |

The noun phrase can be followed by a semantic-relation marking postposition to mark the referent as agentive, instrumental, adverbial (*tɛ*); anti-ergative (animate patient, recipient, benefactive), or allative (*lɛ*); locative or temporal (*dɔ*); terminative (*xrɛʔ*); or comitative (*məŋəŋ*). There is also a topic marker (*nū*) and a noun conjunction particle (*ŋi*). 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 *ǎ-*, or by use of the reflexive/middle marking suffix *-ɕi*. The main function of the prefix *ǎ-* is intransitivization (e.g. *təl* ‘roll’, vt.; *ǎtəl* ‘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 (*maʔ*; possibly the same morpheme as the human group-marking noun suffix *maʔ*) 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. *nəŋi ɲa-ə-sət maʔ ɕʉ*  
 2dl NF-RECIP-hit RECIP dl  
 ‘You two are arguing/fighting (with each other).’
- b. *kəlě jàlě sə-zà maʔ*  
 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 *ɲet-ɕʉ* ‘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 *ətəl* ‘roll’, with no specification of agentivity, whereas the suffix implies agentivity, as in *təl-ɕʉ* ‘roll oneself’ (with change of tone). It is also possible to use both affixes together, giving a stative sense, as in *ə-ɲəŋ-ɕʉ* [PREF-look-R/M] ‘be visible’. (See LaPolla 1995b for more detailed discussion.)

- (3) a. *ətə gəʃjūm-ɕʉ*  
 3sg hit.with.fist-R/M  
 ‘He is hitting himself with his fist.’
- b. *ətə mər tɕuʔ-ɕʉ*  
 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 *sə- ~ tə-*, e.g. *dət* ‘broken (of string)’, *sədət* ‘to break (string)’, *ətɕup* ‘pinched’, ‘closed up’, *tətɕup* ‘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) (<ɲi).

- (4) *ətə tě əŋ lě sə-ɲi*  
 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) (= *sə-lūm*).

An analytical causative/missive construction is formed using the verb (*sə-*)-*zūu*, as in *ɲi səzūu* ‘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 transitivizing suffix: *ŋiú* ‘cry’ > *ŋut* ‘mourn (cry for) a dead person’.

- (5) *ətə tě pūŋ lě wà sə-zūu*  
 3sg AGT Pung DAT do CAUS-allow/make  
 ‘He made/allowed Pung do (it).’

There is only one marker of negation, the verbal prefix *mə-*, pronounced *mā-* when the root already has a suffix, e.g. *mā-rənā* ‘not rest’ (with tone change).

A benefactive construction is formed using the verb *sənəŋ* ‘help’ or the benefactive auxiliary *č* after the main verb (again with changed tone on main verb):

- (6) a. *àŋ ɕūŋ ǎgə tɛ̃ rī sənəŋ*  
 3sg wood 1sg AGT carry help+ 1sg  
 ‘I carry wood for him.’ (*lit.* ‘I help carry his wood.’)
- b. *ǎgə tɛ̃ àŋ ɕūŋ rī čŋ*  
 1sg agt 3sg wood carry BEN+ 1sg  
 ‘I carry wood for him.’
- c. *àŋ gwā sā sǎ-tə̃i č*  
 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 *pə-* (pronounced *pā-* when the verb already has a prefix): *pə-kəi* ‘Eat!’ The prohibitive takes the normal negative, but the main verb is followed by *əl* ‘have’: *mɛ-kəi əl* [NEG+NF-eat have] ‘Don’t eat (it)!’ An indirect directive is marked by the prefix *laʔ-*, which is followed by the negative particle for indirect prohibitives: *laʔ-mə-wá* ‘Don’t let him do (it).’ For the hortative the verb *gəl* ‘want’ is used: *kəi gəl* ‘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. *nǎjuŋ ni-kəi jūŋ (gǔ) guī ē ǎ*  
 2pl NF-eat plural NOM taro be Q  
 ‘Is what you(pl) eat taros?’
- b. *jà sərā ǎnī gǔ ē*  
 this thing who NOM be  
 ‘Whose is this thing?’
- c. *ni-kəi mɛ-kəi*  
 NF-eat NEG+NF-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:  $\Sigma-\theta \rightarrow \Sigma-\eta$ ;  $\Sigma-\rho \rightarrow \Sigma-k$ ; other finals with falling tone  $\rightarrow$  level tone; first person dual:  $\Sigma-\zeta i u i$  first person plural: short vowel  $\rightarrow$  long vowel. Second person dual:  $-\zeta i u$ ; second person plural:  $-j i u \eta$ ). 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 NF (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 NF prefix is pronounced  $na-$  and either appears before the other prefix or incorporates the other prefix, e.g. *na-ǎ-sət* and *na-sət* ‘hit yourself’ are both possible. When it appears with consonant-initial prefixes, the NF prefix is simply

marked by a change of the vowel of the other prefix to *-e*, as in (7c), above. The variant *ŋi-* appears as a sandhi form when followed by a syllable with a front vowel.

- (8) a. *ǰgò tě àŋ lě rŋŋ*  
 1sg AGT 3sg DAT carry + 1sg  
 ‘I carried him.’  
 b. *àŋ tě ǰgò lě ɲǰ-rŋŋ*  
 1sg AGT 1sg DAT NF-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-ɣì* [return-go] ‘go back’ and *lɔʔ-rà* [return-come] ‘come back’. These and other direction-related verbs have grammaticalized into tense markers, *ɣì* (< *ɣì* ‘to go’) and *lŋŋ* (< *lŋŋ* ‘to ascend’), both used for recent past actions. The difference between the two forms is an evidential distinction: the use of *lŋŋ* after the verb implies the speaker did not see the action (9a–a’), whereas the use of *ɣì* implies the speaker did see the action (9b–b’). A guess is marked by adding *mɛl* after *lŋŋ*. For a strong assertion, *mũ* is added after *ɣì*. For an action completed some time ago, possibly years ago, *buĩ* is used in place of *ɣì* or *lŋŋ*. Inchoatives take the particle *pəŋ* after the verb or the adverb *tə̀sə̀* before the verb.

- (9) a. *àŋ tě kà:i lŋŋ*                      a’ *àŋ ɣì lŋŋ*  
 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. *àŋ tě kà:i ɣì*                              b’ *àŋ ɣì àm ɣì*  
 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 *ɣì* and *lŋŋ* are not used with first person actors; instead the adverb *zũur* ‘already’ can be used to mark a completed action, as in (10a). In asking someone about their past actions, *lŋŋ* is used (10b).

- (10) a. *ǰgò ɣũŋ zũur*  
 1sg go + 1sg already  
 ‘I went (and came back) already.’  
 b. *nà ɲǰ-ɣì lŋŋ ǎ*  
 2sg NF-go ASP Q  
 ‘Did you go?’

Another evidential distinction is made with the hearsay particle *tɕiwǎ*. This appears most frequently in traditional narrative texts.

- (11) *tɕə̀pǎ.ɪ tɕə̀pǎ.ɪ nũ dǰgĩ kǎ guʔ s̄ tɕiwǎ*  
 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 *mǎtǎl* in (13), below, while others always precede the verb, e.g. *lǎlǎ* in *lǎlǎ mràŋ* ‘very long/tall’. Preverbal adverbial phrases that are not reduplicated adjectives or adverbs usually take the adverb marker *wā* (12a) or sometimes *gū* (12e). A resultative complement also comes before the main verb, marked by the terminative postposition *xreʔ* ‘until’, as in (12e).

- (12) a. *àŋ blakpàì wā àŋzā kài*  
3sg monkey ADV food eat  
‘He eats like a monkey.’
- b. *mǎdǎw ǎbrà-brà ʃì*  
car/truck fast-REDUP go  
‘The car is going quickly.’
- c. *mǎdǎw ʃì ǎbrà*  
car/truck go fast  
‘The car goes (can go) quickly.’
- d. *nǐ-būm mǎ-gǎl gū ǎmrā wà lǎ ʃì tǎiwǎ*  
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. *‘əpè əmè’ guʔ sǎ xreʔ ʔ: ʃì tǎiwǎ*  
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) *ǎgǎ nà mǎdǎm dǎ mràŋ (mràŋ mǎtǎl)*  
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) *ǎjùŋ cē tsūl ʃǎʔ lǎmbrǎʔ ǎl*  
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 ʃì gǎu ē*  
3sg go NOM be+Q  
‘Did he go?’
- b. *pūŋ cǐn tē sət gē (<gǎuē)*  
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) *ǎgǎ sǎlǎp-ǎú sǎ gē*  
1sg teach-R/M NOM NOM+be  
‘I ought to study.’

Subordinators include clause-final *bē* ‘if’, ‘when’, *mǎnǎŋ* ‘because’, ‘when’ (<‘follow’), *sǎnnǎŋ* ‘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. *kǎi 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. *ǎgǎ àŋ lɔʔ jì nit*  
 1sg 3sg return go remember  
 ‘I remember that he went back.’  
 b. *àŋ jì bē ɔgǎ kuʔ jũŋ diu ē*  
 3sg go if 1sg also go+1sg might COP  
 ‘If he comes, I might go.’  
 c. *àŋ ǎgǎ mǎzi ʔmǎliʔ la-jì wā guʔ*  
 3sg 1sg must have.to-go COMP say  
 ‘He says I must go.’

Predication of actions or attributes that occur at the same time are represented in a serial verb construction with *zīn* optionally appearing between the two verbs:

- (18) a. *mǎnzū wà zīn zǎŋ-ǎu*  
 song do LNK enter-R/M  
 ‘He entered singing.’  
 b. *mrǎŋ zīn pǎɕũŋ gũ jɔʔ*  
 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’, *kǎi sā* ‘the thing which can be eaten/food’. Relative clauses, which appear before the head noun, are generally nominalized by the particle *gũ*. In some lexicalized expressions involving relative clauses, no nominalizer is used, e.g. *mǎnzū wà ǎtsǎŋ* [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 tǎ làŋ gũ ɕũŋwǎt*  
 3sg AGT bring NOM flower  
 ‘the flower(s) he brought’  
 b. *ɕũŋwǎt làŋ gũ ǎtsǎŋ*  
 flower bring NOM person  
 ‘the person who brought the flower(s)’  
 c. *àŋ lǎ ɕũŋwǎt làŋ gũ ǎtsǎŋ*  
 3sg DAT flower bring NOM person  
 ‘the person I brought flower(s) for’  
 d. *zǎjǎ brī sā ò gũ (sǎrǎ)*  
 book write NOM do NOM thing  
 ‘the thing I use to write books’  
 e. *ì tɛ ɔ tsǎrī*  
 hemp INST use 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  $V-t\bar{ɔ}n\ m\check{a}n\bar{ɔ}ŋ$  ‘[V finish follow]’ ‘having finished Ving’ or ‘after Ving’ where one or more clauses with reduplicated verbs will lead up to the doing of some act, and then  $V-t\bar{ɔ}n\ m\check{a}n\bar{ɔ}ŋ$  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)  $t\bar{a}n\bar{a}$        $\bar{a}kpl\bar{ɔ}ŋ$   $z\bar{u}$        $t\bar{ɔ}n$        $m\check{a}n\bar{ɔ}ŋ$ ,  
 crossbow body make finish afterwards  
 ‘After finishing the body part of the crossbow’,
- $c\bar{e}l\bar{a}i$        $\bar{ɔ}ŋ$        $t\check{a}l\bar{i}$        $w\bar{a}$        $r\check{a}m\bar{u}-m\bar{u}$   
 one side LOC bow ADV draw-REDUP  
 on one side, (I) draw the shape of a bow,
- $d\check{a}k\check{c}i$        $t\bar{e}$        $puk$ .  
 small.knife INST bore.small.hole + 1sg  
 (and then) use a small knife to bore a small hole.
- $pu?$        $t\bar{ɔ}n$        $m\check{a}n\bar{ɔ}ŋ$ ,  $n\check{a}ts\bar{e}n\check{a}ts\bar{e}$        $t\check{a}l\bar{i}$        $z\bar{a}ŋ$ .  
 bore.small.hole finish after slowly bow fit.into + 1sg  
 ‘After having bored a small hole, (I) slowly fit the bow into the hole.’

## REFERENCES

- Barnard, J.T.O. (1934) *A Handbook of the Rawang Dialect of the Nung Language*, Rangoon: Superintendent of Government Printing and Stationery.
- Dai, Qingxia, Huang, Bufan, Fu, Ailan, Renzengwangmu and Liu, Juhuang (1991) *Zàng-Miǎnyǔ shíwǔzhǒng (Fifteen Tibeto-Burman languages)*, Beijing: Yanshan Chubanshe (includes chapter on Dulong).
- LaPolla, Randy J. (1995a) ‘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*, Academia Sinica 66.4: 1149–85.
- LaPolla, Randy J. (1995b) ‘Reflexive and middle marking in Dulong/Rawang’. Paper presented to the 28th International Conference on Sino-Tibetan Languages and Linguistics, Charlottesville, VA, October 6–9. To appear in *Himalayan Linguistics*, edited by George van Driem, Berlin: Mouton de Gruyter.
- 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’, 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.
- Liu Juhuang (1988) ‘Dúlóngyǔ dòngcí yánjiū’ (Studies on the Dulong verb), *Yǎyán yánjiū* 1988. 1:176–91.
- Lo Ch’ang-p’ei (1945) ‘A preliminary study of the Trung language of Kung Shan’. *Harvard Journal of Asiatic Studies* 8: 343–48.
- 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: 338–69.
- Sun, Hongkai (1982) *Dúlóngyǔ jiǎnzhi (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.

# KHAM

*David E. Watters*

## 1 INTRODUCTION

Kham, in all its varieties, is spoken in the upper valleys of the Rukum, Rolpa, and Baglung Districts of West Central Nepal by some forty or fifty thousand people. The majority of Kham speakers are Budhas, Puns, Ghartis, and Rokhas – all classified ethnographically as subtribes or clans of the Magar tribe. Geographically, the four Kham speaking subtribes are cut off from the rest of the Magar population by more than a week's walk over rugged, mountainous terrain. Though a comparison of basic vocabulary reveals no more of a special relationship between Kham and Magar than it does between Kham and other Himalayan languages, a careful examination of more innovative vocabulary makes it apparent that Magar is indeed Kham's closest relative.

Kham is known to Nepalis of the region as 'Khamkura', which, roughly translated, means *Kham-talk* or *Kham-speech*. The word *Kham* itself is of unsure origins and means simply 'language' in its broad sense, and '*The Language*' in its strict sense. In West Central Nepal, where the Kham dialects are spoken, the Nepali use of the word *Kham* or *Khamkura* has the more generalized meaning of a local, non-Nepali dialect. Consequently, at least two other languages in the region, Chantyal (see Noonan, Chapter 20 this volume) and Kaike, have received the Nepali appellation *Khamkura*, though neither of them is directly related to the Kham described here.

### 1.1 Kham dialects

Though no extensive survey of Kham has been conducted to date, the general picture emerging from my own fieldwork is that Kham is separated linguistically into three major dialect groupings – Sheshi, Gamale, and Parbate (see Figure 42.1). Typological differences between the three are fairly pronounced, especially in the area of verb morphology, resulting in severe intelligibility restrictions between them. Parbate, the largest group, can be further subdivided into Eastern and Western Parbate, and though intelligibility across the two sub-branches is still fairly low, it is sufficiently high that speakers do not generally resort to secondary languages for communication. Figure 42.1 shows the three major branches of Kham (plus the lower level split in Parbate), along with the names of eleven dialects at the terminal nodes.

The dialect names at the terminal nodes are, broadly speaking, the names of regional dialects composed of several, more specific, village dialects. In a narrower sense, they are the names of specific village dialects. Thus, for example, Takale is both a village dialect, spoken in Taka village, and a regional dialect composed of several village dialects spoken in the same river valley (see Figure 42.2). Regional dialects dominated by the same mother node are all

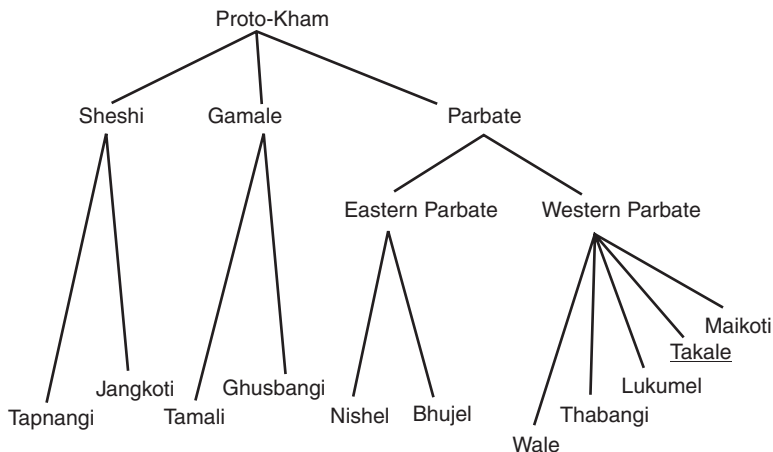


FIGURE 42.1 THE DIALECTS OF KHAM

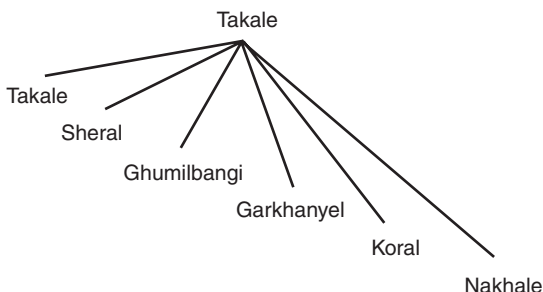


FIGURE 42.2 A PARTIAL LISTING OF VILLAGE DIALECTS BELONGING TO THE TAKALE REGIONAL DIALECT

mutually intelligible, though when situated at opposite ends of a continuum, as are Maikoti and Wale (Figure 42.1), intelligibility may be considerably reduced.

## 1.2 The place of Kham in Tibeto-Burman

Unclassified in any of the major classifications until DeLancey (1987), Kham falls into a western division of the Kiranti group along with Vayu, Chepang, and Magar. Although Kham shares little innovative vocabulary with the Kiranti languages, it does show remarkable correspondences in morphology and the organization of verbal paradigms. Based primarily, then, on the common retention of innovated paradigmatic patterns and cognate forms within those patterns, Kham, along with Magar and Chepang, can be regarded as distant cousins of Kiranti (as suggested also in DeLancey 1987). At the same time, Kham shares a considerable amount of innovative vocabulary with Magar. The Kham languages, then, turn out to be an important link between the East Himalayish languages and the loose grouping of languages Shafer (1966) referred to as West-Central Himalayish.

## 2 PHONOLOGY

### 2.1 Vowels

There are nine vowels in Takale Kham, six of which are basic, and three of which can be traced to secondary developments resulting from the loss of consonants. The six basic vowels are two front, /i/ and /e/, two central, /ə/ and /a/, and two back, /u/ and /o/. Of the three secondary vowels, the front rounded vowels /ü/ and /ö/ are the reflex of a lost /p-/ or /b-/ prefix on syllables with following front vowels /i/ and /e/. In some dialects, notably those from the Maikoti cluster, the prefix still occurs as a relic without semantic function, and in Nishel, prefixed \*p- has become modern o-. In all other dialects, rounding on the following front vowel (i > ü, and e > ö) attests to the presence of the prefix in the proto-language.

The high-back unrounded vowel [u] (written orthographically as /i/) is also a secondary development, but of all the daughter dialects, it occurs only in Takale. The vowel occurs as a reflex of a lost velar coda in the rhyme \*ək or \*əŋ (with \*-əŋ > -ĩ:). The reflex of a lost final velar consonant in any other environment is only compensatory lengthening. Table 42.1 gives the full inventory of vowels in Takale Kham.

All vowels in Takale Kham are contrastive for length and all vowels except /i/ and /ö/ are contrastive for nasalization. All nasalized vowels are phonetically long. All vowels are also contrastive for voice register (or ‘phonation type’). Although voice register is properly a feature of the tonal system, it is manifested on the vowel. There are two contrastive registers – modal register, which is the default, unmarked register, and lax register (a laryngeal quality) which is indicated by an orthographic *h* following the vowel, as in *ah*, *əih*, etc. There are numerous cases in which the lax register can be traced to a lost prefix \*s-.

Diphthongs in Takale Kham occur from any starting point in the oral cavity, but terminate only on front or back vowels. Diphthongs ending in /i/ are, in the majority of cases, the reflex of a lost \*-t or \*-n in the coda (see Michailovsky 1975), the loss of \*-n resulting also in nasalization on the diphthong, as in /əĩ/.

### 2.2 Consonants

The consonantal inventory in Kham is simple, with only three points of articulation having phonemic status – labial, alveolar, and velar (ignoring the consonant /h/). Allophonic alveopalatal articulation occurs on sibilants with following front vowels, as in *fĩ* ‘wood’ or *dʒe*: ‘you plural’. Table 42.2 gives the consonant phonemes in Takale Kham.

Aspiration occurs only with voiceless onsets. What to the uninitiated ear might sound like aspiration on voiced obstruents /b<sup>h</sup>/, /d<sup>h</sup>/, /g<sup>h</sup>/, and /dz<sup>h</sup>/ is, in fact, lax phonation on the syllable. The tonal properties of syllables with lax phonation are entirely different from those with aspirated onsets.

TABLE 42.1 VOWEL PHONEMES IN TAKALE KHAM

	<i>Front</i>		<i>Central</i>	<i>Back</i>	
	<i>-round</i>	<i>+round</i>		<i>-round</i>	<i>+round</i>
High	i	ü		u (i)	u
Mid	e	ö	ə		o
Low			a		

TABLE 42.2 CONSONANT PHONEMES IN TAKALE KHAM

	<i>Labial</i>		<i>Alveolar</i>		<i>Velar</i>		<i>Glottal</i>
	<i>- voice</i>	<i>+ voice</i>	<i>- voice</i>	<i>+ voice</i>	<i>- voice</i>	<i>+ voice</i>	
Stops	p	b	t	d	k	g	
+ asp.	p <sup>h</sup>		t <sup>h</sup>		k <sup>h</sup>		
Fricatives			s	z			
Affricates			ts	dz			
+ asp.			ts <sup>h</sup>				
Nasals		m		n		ŋ	
Approximants:							h
liquids				l r			
glides				y		w	

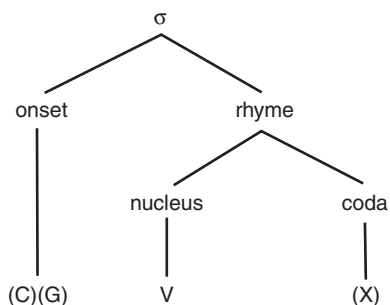


FIGURE 42.3 THE SYLLABLE IN KHAM

### 2.3 The syllable

The basic syllable in Kham is (C)(G)V(X), where ‘G’ is a glide and ‘X’ is a consonant or a vowel. Only the nucleus is obligatory. The presence of a glide adds no weight to the syllable. Weight is determined solely by the rhyme; the presence of X always creates a heavy syllable. Where X is a vowel of the same quality as V, the result is a long vowel (V:), and where X is a vowel of different quality, the result is a diphthong (VV). Where X is a consonant, V is necessarily short. In terms of syllable weight and rhythm groups, then, V:, VV, and VC have the same weight – all are bimoraic.

The minimal word in Kham is bimoraic, while the minimal syllable is monomoraic. Thus, two monomoraic syllables can make a well-formed word, as in *ba-ke* ‘he went’. Monosyllabic open syllable words, on the other hand, can manifest only long vowels – even those with an inherently short vowel. Only in the presence of an additional syllable/affix can the short vowel be manifested.

There are several monomoraic morphemes in Kham consisting of simple vowels, one being *-e*, an ergative, instrumental, genitive, and imperfective morpheme, and the other *-o*, a third person singular, and a nominalizing morpheme. Affixed to long, open syllables (including diphthongs), a following *-e* morpheme is resyllabified as *-ye*, and *-o* as *-wo*. That is, *vvv* is resyllabified to *vvgv*. Affixed to short, open syllables, the suffixes can either

resyllabify as *-ye* or *-wo*, or coalesce with the preceding syllable, as in the following input > output strings: *no-e* > *noy*, *su-e* > *suy/swi:*, *ke-o* > *kyo:* and *si-o* > *syu:*. Where coalescence occurs, the underlying vowels *-e* and *-o* harmonize to the height of the preceding vowel, as in *si-o* > *syu:*.

## 2.4 Tonal patterns

Tone in Kham, as with other languages of the Himalayan region, is best described as an intersection of ‘tonal melody’ and ‘voice register’ (Hale and Pike 1970; Hari 1980). There are two contrastive melodies – a marked, ‘accented’ tone (T1), and an unmarked, ‘neutral’ tone (T2). Voice register has a modifying effect on the fundamental frequencies of tone bearing units, dividing the pitch range into ‘upper’ and ‘lower’ frequencies. Thus, lax register, a marked phonation type characterized by breathiness and laryngeal laxity, induces lower frequencies on T1 and T2 than its unmarked counterpart, modal register. Figure 42.4 illustrates the intersection.

Unlike what has been found for Bodish languages, the marked/low register in Kham cannot be tied to the voicing of initial consonants. Lax register occurs on syllables with voiced and voiceless onsets alike. Furthermore, the evidence suggests that laxity is tied, in at least some cases, to the loss of prefixed *\*s-*. Again, unlike Bodish languages, the tonal melodies cannot be tied to lost finals. Rather, it seems plausible that T1 began as a simple heightened pitch, eventually spreading its features to cover the entire length of the root morpheme, whether one, two, or three syllables.

Some features of tone in Kham are evident only in polysyllabic environments. Across stem-stem boundaries and stem-suffix boundaries, for example, T1 stems are followed by an obligatory ‘downstep’ pitch (see Boxes 1 and 2 of Figure 42.4). This means that T1 stems are always higher than anything that follows. A string of multiple T1 stems in a single intonational phrase, then, results in a series of multiple downsteps. Downstepping is terminated on a suffix, and pitch is reset on the following stem. T2 stems are characterized by the lack of downstep (which is not the same as an atonal syllable).

Downstep never occurs within a single stem, not even when the stem is a morphological compound with multiple T1 morphemes. The entire stem will manifest a single tone, whether T1 or T2, determined by the inherent tone of the first syllable. Thus, where the contributing morphemes in a compound are inherently T2-T1-T1, for example, the entire stem will manifest a T2 pitch, and the tonal transition to a following stem will be a lack of downstep (a T2 feature). The underlying tone of the two non-initial T1s in the compound will be overridden. Only on a following suffix will downstep occur, and that, because T1 occurs in the compound stem. Much like N+N compounds, many derivational affixes on verbs, too, form compounds with the root morphemes to which they are attached.

	Modal		Lax	
	stem	suff	stem	suff
Tone-1	—	— <sup>1</sup>	—	— <sup>2</sup>
Tone-2	—	— <sup>3</sup>	—	— <sup>4</sup>

FIGURE 42.4 THE CONTRASTIVE ‘FOUR-TONE’ SYSTEM IN KHAM

Tone in Kham, then, is characterized by an overwhelming tendency found in most, if not all, of the tonal systems of the Himalayas – the inherent tones of conflicting syllables are suppressed to achieve a ‘one-tone’ constraint over certain stretches of speech. In some languages, the primary domain is the ‘word’, and in others like Kham, the primary domain is the word base (roots and derivational morphemes). Inflectional suffixes manifest a separate tonal character.

### 3 INFLECTIONAL MORPHOLOGY

#### 3.1 Nominal morphology

Inflectional morphology is relatively rich in Kham, more so on verbs than on nouns. Nouns are inflected for a single prefixal position (possession or, with a few nouns, a classifier numeral), and several suffix positions. First and second person possessive prefixes are transparently related to free pronouns, as in the following:

	<i>Free pronoun</i>	<i>Possessive prefix</i>
1st sg	<i>ŋa:</i>	<i>ŋa-</i>
1st dl	<i>gi-n</i>	<i>gi-n-</i>
1st pl	<i>ge:</i>	<i>ge-</i>
2nd sg	<i>nĩ:</i>	<i>nə-</i>
2nd dl	<i>ji-n</i>	<i>ji-n-</i>
2nd pl	<i>je:</i>	<i>je-</i>
3rd sg	<i>no:</i>	<i>o-/u-</i>
3rd dl	<i>no-ni</i>	<i>ni-</i>
3rd pl	<i>no-rə</i>	<i>ya-</i>

Nominal suffixes mark categories like number, grammatical case, and several local/semantic cases, as in (see also 3.1.4):

*Number:*

<i>-∅</i>	singular
<i>-ni</i>	dual
<i>-rə</i>	plural

*Grammatical case:*

<i>-e/-ye</i>	ergative, instrumental, genitive
<i>-lai</i>	primary object (direct object of transitive and indirect object of ditransitive – see Dryer 1986)
<i>-sə</i>	comitative

Number marking is obligatory on nouns, and must (generally) be accompanied by a third person number agreement marker on the verb, as in: *u-lu:-rə ma-ra-dəi-wo* [3S-sheep-PL NEG-3P-find-3S:IMPFV] ‘He didn’t find his sheep.’

The three grammatical cases are mutually exclusive and occupy a position immediately following number, as shown in Table 42.3.

##### 3.1.1 *The ergative/instrumental*

Both the ergative and the instrumental are marked by the suffix *-e/-ye*. Where both occur in a single clause, only the ergative marked NP (which is also the subject) is marked for agreement in the verb, as in *ge: rowa-e ge-pəl-ke* [we axe-INSTR 1P-chop-PFV] ‘We chopped it with an

TABLE 42.3 NUMBER AND CASE MARKING

-1	0	+1	+2
POSS-	root-	NUM-	ERG/INSTR/GEN
			OBJ
			ASC

axe'; *no-ra-e rowa-e pəl-ke-rə* [he/she-PL-ERG axe-INSTR chop-PFV-3P] 'They chopped it with an axe.' Likewise, in intransitive and detransitivized contexts, the subject NP (and not the instrumental argument) is marked for agreement in the verb, as in: *ge-nəi:-rə khī:-ye ba-ke-rə* [1P-friend-PL foot-INSTR go-PFV-3P] 'Our friends went by foot.' A detransitivized example is: *ao-rə kwi:-ye jəi-si-u li-zya-rə* [this-PL hand-INSTR make-DETRANS-NML be-CONT-3P] 'These are made by hand.'

The ergative never occurs on first or second person arguments, nor, so far as I know, on third person intransitive agents – not even for marking emphatic volition.

### 3.1.2 The genitive

The genitive, also marked by *-e/-ye* in Takale (*-lə* in some dialects), marks the modification of one noun by another, as in *sən-e kwa*: [wool-GEN cloth] 'woollen clothing'. The genitive also occurs on the possessor in combination with third person possession, as in *no-e u-zihm* [he-GEN 3S-house] 'his house'. The genitive does not occur with first or second person possession.

The genitive followed by a bare stem (simple modification) and the genitive in combination with third person possession (possessive modification) form contrastive pairs, as in:

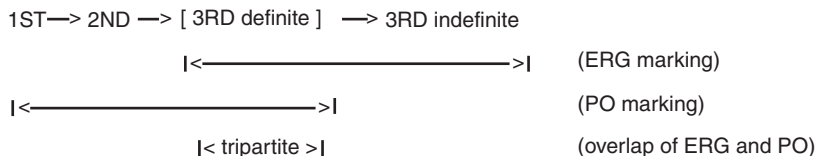
- (1) a. *baza-e kər*  
 bird-GEN wing  
 'a bird wing'  
 b. *baza-e o-kər*  
 bird-GEN 3S-wing  
 'the bird's wing'

### 3.1.3 Alignment of case

Kham employs a split ergative case marking strategy in which only third person subjects in a transitive clause are marked by the ergative. First and second person subjects are always zero marked (nominative). The second participant of a transitive clause, the most patient-like argument, is either zero marked (absolutive) or marked by the 'primary object' marker, depending on the definiteness of the argument. Thus, on the hierarchy in Figure 42.5, only those subject arguments at the lower end of the scale will be marked by the ergative (all third persons, definite or indefinite), and only those object arguments at the higher end will be marked by the PO marker (first or second person and third person definite). Where the scales overlap, a tripartite systems occurs in which S, A, and P are all distinctive.

A high ranking subject (first or second person) and a low ranking object (third indefinite) will result in zero marking on both participants (see 2a), while a low ranking subject (third





**FIGURE 42.5** PERSON HIERARCHY RELEVANT TO KHAM CASE MARKING

person) and a high ranking object (third definite or above) will result in ergative plus PO marking (see 2b), as in the following:

- (2) a. *ŋa:-θ la:-θ ŋa-səih-ke*  
 I-NOM leopard-ABS 1S-kill-PFV  
 'I killed a leopard.'
- b. *gē:h-ye ŋa-lai duhp-na-ke-o*  
 ox-ERG I-PO butt-1S-PFV-3S  
 'The ox butted me.'

It should be noted that in the Himalayan region, 'experiencers' are commonly marked by the equivalent of *-lai* in constructions like *Hunger is to me* (= 'I am hungry'). In Kham the construction is rare, being restricted to a few expressions borrowed complete from Nepali. In native constructions, experiencers assimilate to the unaccusative proto-type, as in: *ŋa:ŋa-kəre:-ke* [I 1S-hunger-PFV] 'I am hungry.'

### 3.1.4 Locative / directional suffixes

Kham employs numerous local cases having to do with static location or directional movement, as in the following:

- (3) *Local cases* (from Takale Kham)
- kə general locative (at)
  - ŋə adessive (around the area)
  - lə inessive (in/inside)
  - tə superessive (on)
  - phə cisative (on the side of)
  - da allative (to/towards)
  - ni ablativ (from)
  - kin elative (away from)/comparative
  - tin delative (down from)
  - pəi lative (up to)
  - səi orientative (in the direction of)

The general locative *-kə* is used primarily to mark location at a place. Used in the context of motion/direction verbs like 'go' or 'come', however, *-kə* implies both 'motion to' and 'arrival at' a place. In such contexts it contrasts significantly with the allative *-da*, as in:

- (4) a. *u-zihm-kə ba-ke*  
 3S-house-LOC go-PFV  
 'He went to (and arrived at) his house.'

- b. *u-zihm-da ba-ke*  
 3S-house-LOC go-PFV  
 ‘He went to (left for) his house.’

The ablative *-ni*, too, has several senses depending, in part, on whether it occurs with motion or stationary verbs. With motion verbs it indicates movement from a place, and with verbs not marked for motion it expresses a notion of ‘on the side of’, as in: *yahm a-ni jəi-ke* [door prox-ABL make-IMP] ‘Make the door on this side!’ The ablative can also mark some instruments, as in: *ŋa: khĩ-ni ŋa-ba-ke* [I foot-ABL 1S-go-PFV] ‘I went by foot’.

Using comparative evidence from other dialects, it is apparent that the Takale elative *-kin* ‘away from’ and delative *-tin* ‘down from’ are contractions of an earlier *-kə-ni* (LOC-ABL) and *-tə-ni* (ON-ABL), respectively. The elative, in addition to ‘motion away from’ also marks the comparative, as in:

- (5) *ŋa:-kin khyo:-wo li-zya*  
 I-ELAT long-NML be-CONT  
 ‘He is taller than I am.’

The locative suffixes listed in (3) occupy four position classes, as shown in Table 42.4. Following slot +4, one of three nominalizing morphemes can also occur.

Slots +1 and +3:

*lū:-ra-da*

stone-PL-ALLT

‘towards the stones’

Slots +2 and +3:

*lū:-phə-tə*

stone-CIS-ON

‘on the stone side’ (on the side where the stone is)

Slots +3 and +4:

*ŋah-da-kin*

front-ALLT-ELAT

‘from beforehand’

Slots +3 and +5:

*ŋah-da-ŋao*

front-ALLT-NML3

‘the earlier one’

Slots +3, +4, and +5:

*ŋah-da-kin-kao-rə*

front-ALLT-ELAT-NML2-PL

‘the ones from beforehand’

TABLE 42.4 POSITION CLASSES FOR NOMINAL AFFIXES

-1	0	+1	+2	+3	+4	+5
POSS-	root-	NUM-	CIS-	ALLT- ABLT- LOC- ADS- IN ON	ELAT- DEL LAT ORIENT	(NML1/NML2/NML3)

### 3.1.5 Gender marking

For the most part, there is no gender marking on nouns in Kham, although there is a very old masculine *-pa* and feminine *-ma* suffix lexicalized on certain nouns, as in the following: *khepa* ‘male human’, *dahpa* ‘young man’, *mē:ma* ‘female human’, *bama* ‘hen’. A diminutive *-za* < \*child also occurs occasionally – *riza* ‘younger brother’, *biza* ‘rat’, *baza* ‘chicken’, etc.

The Indic *-a/-i* alternation (m/f) is partially productive on what I call the ‘similative’ class of nouns (see Section 4.2). With similative nouns, some concrete object is conceptualized as representative of an attribute, and the noun with a m/f suffix is used, in part, to identify livestock and nickname people, as in:

(6)	<i>korop</i>	stitch	>	<i>koropchya/kurupci</i>	m/f scar-face
	<i>buchula</i>	adze	>	<i>buchulya/buchuli</i>	m/f adze-tooth
	<i>nakira</i>	wooden pin	>	<i>nakiryana/nakiri</i>	m/f skinny

### 3.1.6 Numeral prefixes

Numeral prefixes occur on a relatively small subset of nouns – those that can in some way be used to quantify other nouns, as in: *to-ro:h harmwī*: [ONE-churn milk] ‘a churnful of milk’, *tu-gur bāhtanje* [ONE-load potato] ‘a load of potatoes’, etc. A major function of such constructions is to impose structure on things that lack structured wholeness. On occasion, nouns that are not normally members of the subset can be used for creating neologisms – *tə-zihm rā:-wo* [ONE-house wide-NML] ‘one house wide’; *tə-kyā: gis-o* [ONE-body heavy-NML] ‘one body heavy’.

Full fledged numerals, those that function as quantifiers in an NP, include an element *-bo/-blo* following the numeral prefix, as in (7). Presumably, *-bo/-blo* itself is some kind of erstwhile classifier unit.

(7)	<i>to-bo</i>	one
	<i>neh-blo</i>	two
	<i>sōh-mlo</i>	three

## 3.2 Deictic roots and relator nouns

The nominal suffixes referred to in Section 3.1.4 also combine with special deictic primitives, all bound roots, with meanings like ‘proximate/distal/remote’, ‘up/down’, ‘front/back’, ‘left/right’, and ‘where’. Together, they form complex locative expressions. The terms for ‘here’ (*a-kə*) and ‘there’ (*nə-kə*), for example, are formed by combining a locative suffix *-kə* with the primitives *a-* ‘proximate’ and *nə-* ‘distal’, respectively. Following are the forms for the roots.

### 3.2.1 Deictic roots

proximate	<i>a-</i>
distal	<i>n-/nə-</i>
remote	<i>h-/ho-</i>
up	<i>ro-</i>
down	<i>me-</i>
front	<i>ŋah-</i>
back	<i>chī:</i>

left	<i>dabərya</i> (Indic)
right	<i>wor-</i>
where	<i>kan-</i>

The demonstratives ‘this’ and ‘that’ in Takale Kham are simply nominalized forms of the deictic primitives ‘proximate/distal/remote’. In other dialects, the demonstratives include also the general locative suffix *-kə*, as in:

<u>Takale:</u>	<u>Gamale:</u>
<i>a-o</i>	<i>a-k-a</i>
prox-NML	prox-LOC-NML ‘this’
<i>n-o</i>	<i>nə-k-a</i>
dist-NML	dist-LOC-NML ‘that’
<i>h-o</i>	<i>ho-k-a</i>
rem-NML	rem-LOC-NML ‘that (yonder)’

Following the deictic primitives and forming compounds with them is a special class of landmark locations – up-country, down-country, side of a mountain (or the side facing outward), and side of a valley (or the side facing inward). The compound stems, like the primitives themselves, require locative suffixes to complete them. The position classes are given in Table 42.5, with examples following.

- (8) a. *ho-ya:-lə*  
 remote-UPCOUNTRY-IN  
 ‘in that upper area’  
 b. *ŋah-ti:-da*  
 front-OUTWARD-ALLT  
 ‘towards the front facing side’

### 3.2.2 Relator nouns

Relator nouns are a special class of nouns that further specify locational relations. Unlike deictic primitives, they get their deictic specification from other nouns, and as such, must either be possessed (in which case they relate to the possessor) or immediately follow the noun they relate to. Relator nouns include notions like ‘on top of’, ‘at the foot of’, ‘at the side of’, etc., as in:

- (9) *u-dū:h-lə*                      *zihm dū:h-lə*  
 3S-under-IN                      house under-IN  
 ‘underneath it’                      ‘under the house’

### 3.3 Verbal morphology

The Kham verb, in addition to marking the categories of tense, aspect, and modality, also marks the person and number of two grammatical roles in transitive and ditransitive

**TABLE 42.5** POSITION CLASSES FOR COMPLEX LOCATIVES

0	+1	+2
DEICTIC	-LNDMRK	-LOC SUFF

**TABLE 42.6** SPLIT ERGATIVE AGREEMENT PATTERNS BASED ON A HIERARCHY OF PERSON

	<i>intrans</i>	<i>1st</i>	<i>2nd</i>	<i>3rd</i>
1st	-ŋ	–	-na	-ŋ
2nd	-na	-ŋ	–	-na
3rd	-u	-ŋ	-na	-u

relationships – subjects and objects. As far as is known, Kham is the only Tibeto-Burman language that consistently marks both. Most Tibeto-Burman languages with person marking strategies follow a pattern in which person takes precedence over grammatical role. In a person hierarchy in which 1/2 ranks over 3, only the highest ranking participant will be marked in the verb, and that regardless of its role. Where both participants are high, i.e. 1 or 2, the object is marked. Thus, in 1-2, 2-1, 3-1 and 3-2 transitive configurations the 1/2 index references the object, while in 1-3 or 2-3 relationships the same 1/2 index references the subject. Because a given person index can reference either a subject or object role, the strategy has sometimes been referred to as a ‘split ergative’ agreement pattern (Bauman 1975; DeLancey 1980). Thus, in Table 42.6, the 1/2 columns are ergative, and the 3 column is nominative – hence, ‘split ergative’.

As a secondary development, some of the languages have devised a means for partially distinguishing the grammatical role of some participants. In some cases, the number of a non-singular participant will also be marked, but this is not the same as indexing two participants. It is frequently the case that the person index and the number index converge on the same participant.

### 3.3.1 Person/number marking in Takale

Takale Kham represents a fairly radical departure from the patterns just described, in that participants are coded in the verb according to their grammatical role. Furthermore, transitive subject indices are the same as the intransitive indices – a purely nominative pattern. Not only do subject and object forms differ from one another in shape, they differ also in affixation type. For first and second persons, subjects are prefixed and objects suffixed, while for third persons the opposite arrangement holds – subjects are suffixed and objects prefixed, as in the following (the arrow shows the direction of action from subject to object):

- (10) a. First and second person participants (prefixed subject, suffixed object)

*subject* → *object*

1s	<i>ŋa-</i>	<i>-na</i>
2s	<i>nə-</i>	<i>-ni</i>
1d	<i>gi-n-</i>	<i>-si-n</i>
2d	<i>ji-n-</i>	<i>-ci-n</i>
1p	<i>ge-</i>	<i>-si</i>
2p	<i>je-</i>	<i>-ci</i>

- b. Third person participants (prefixed object, suffixed subject)

*object* ← *subject*

3s	<i>∅</i>	<i>-o</i>
3d	<i>ni-</i>	<i>-ni</i>
3p	<i>ya-</i>	<i>-rə</i>

The only part of the pattern suggestive of links to the rest of TB is in the differential treatment of third persons in opposition to first and second persons; a variation on a theme of hierarchical agreement patterns. A question that naturally arises is whether or not the agreement patterns in Takale Kham are relatable to other TB agreement paradigms. It turns out that they are. In the more conservative Kham dialects, especially in Gamale and Sheshi, person agreement patterns are clearly reminiscent of Kiranti patterns. In fact, original patterns can be reconstructed for Proto-Kham in which, like Kiranti, person agreement is with the highest ranking person, and number agreement is with the non-singular participant. Furthermore, the actual forms are almost identical to forms found elsewhere in Tibeto-Burman. The situation is quite different in progressive Takale where not only are the patterns different, but the forms, too, are different. Table 42.7 gives the first person free pronominal forms together with their corresponding verbal agreement forms in three Kham dialects – Takale, Mahatale, and Sheshi – as well as in Chepang and Gyarong. Clearly, the forms for Mahatale and Sheshi are original, while the Takale forms are the result of later, secondary developments.

In Takale, not only are the verbal agreement forms for subjects transparently related to free pronouns, they also occur as prefixes. Identical prefixed forms occur elsewhere in the grammar to mark possession on nouns, as in *ŋa-zihm* (1S-house) ‘my house’, *ge-zihm* (1P-house) ‘our house’. Additionally, the prefixes occur in ‘object nominalizations’, as in *ŋa-jəi-wo zihm* (1S-make-NMLZ house) ‘the house that I made/the house of my making’. Though similar object nominalizations occur in all dialects of Kham, only in Takale is there a close correspondence between the regular, declarative paradigm and the nominalized paradigm. The inference to be drawn, of course, is that the modern declarative paradigm in Takale is an erstwhile nominalization and prefixed agents are erstwhile possessors of the action. Much in the anomalous character of the Takale paradigm, in fact, can be accounted for by recognizing a relatively recent verb-noun-verb cycle in which erstwhile verb nominalizations have begun once again to function as verbs of independent, unembedded clauses. Accretions of nominal morphology on one-time nominalized verbs, followed by reanalysis and analogical levelling has produced the modern patterns.

In Sheshi, regular and nominalized paradigms remain entirely distinct. Agreement patterns in the regular paradigm are exclusively suffixing, while in the nominalized paradigm, agent/possessor pronouns are prefixing. The situation is in many ways analogous to what we find in Tiddim Chin – two mutually exclusive paradigms, one of which incorporates old Tibeto-Burman suffixal material, and the other of which incorporates newly prefixed possessive pronouns (Henderson 1957).

Following is a comparison of selected agreement patterns from the regular, declarative and nominalized paradigms in the Takale and Sheshi dialects. Two points should be recognized – one, that the patterns and forms in the Sheshi declarative paradigm bear resemblance to other TB paradigms, and two, that while the declarative and nominalized paradigms are

**TABLE 42.7** AGREEMENT FORMS IN FIVE TB LANGUAGES

Kham	Verbal agreement forms				
	Takale	Mahatale	Sheshi	Chepang	Gyarong
Free forms					
1s ŋa:	ŋa-	-ŋ	-ŋ	-ŋ	-ŋ
1d gin	gin-	-ci	-ci	-cə	-ch
1p ge:	ge-	-e	-i	-i	-i

almost identical for Takale, they remain mutually exclusive for Sheshi (V=verb root, T=tense/aspect):

Sheshi:	<i>nominalized</i>	<i>declarative</i>
1s-2s	<i>ŋa-v-n-u</i>	<i>-na</i>
1s-2p	<i>ŋa-v-ci-u</i>	<i>-cya</i>
2s-1s	<i>nə-v-əŋ-n-u</i>	<i>-ŋ-na</i>
2s-1p	<i>nə-v-si-n-u</i>	<i>-si-na</i>
1s-3s	<i>ŋa-v-ŋ-u</i>	<i>-ŋa</i>
1s-3p	<i>ŋa-v-əŋ-waŋ</i>	<i>-ŋ-ra</i>
Takale:	<i>nominalized</i>	<i>declarative</i>
1s-2s	<i>ŋa-v-ni-u</i>	<i>ŋa-v-ni-T</i>
1s-2p	<i>ŋa-v-ci-u</i>	<i>ŋa-v-ci-T</i>
2s-1s	<i>nə-v-na-o</i>	<i>nə-v-na-T</i>
2s-1p	<i>nə-v-si-u</i>	<i>nə-v-si-T</i>
1s-3s	<i>ŋa-v-wo</i>	<i>ŋa-v-T</i>
1s-3p	<i>ŋa-ra-v-wo</i>	<i>ŋa-ra-v-T</i>

### 3.3.2 Transitivity alternations

Derivational adjustments in verbal valence on inherent transitivity types in Kham belong to two basic types – causativization and detransitivization. Causativization, marked by the prefix *sə-*, is a process by which intransitive verbs are made transitive (and certain types of transitive verbs are made ditransitive). For intransitive verbs, the process is available primarily only to the unaccusative class, for the simple reason that no realignment of semantic roles is required; the semantic status of the single patient/theme argument of the underlying verb remains unchanged under causativization, as in *mī:h-ke/sə-mī:h-ke-o* ‘it cooked’/‘he cooked it’. With unergative verbs, only the periphrastic causative is available, in which case, two verbs are involved and there is no change in the transitivity of the embedded verb. The causing verb is a specialized use of the verb ‘to send’, as in:

- (11) (*no:-ye*) *o-za-lai* *ba-o* *pəri:-ke-o*  
 (she-ERG) 3S-child-PO go-NML send-PFV-3S  
 ‘(She) made her child go’.

Detransitivization is marked by the suffix *-si*, and applies primarily, though not exclusively, to transitive verbs. Semantically, the derivation yields varying results, depending largely on the class of verb to which the derivation applies, but also on secondary factors like grammatical number and tense/aspect. With transitive verbs capable of selecting an animate patient, for example, the resulting derivation yields a reflexive or reciprocal sense depending on the grammatical number of the participants involved – *səih-si-ke* (singular) ‘he killed himself’ vs *khəp-si-ki-ni* (dual) ‘they fought each other’. In an imperfective environment, with its implication of a non-specific, more timeless circumstance, detransitivization yields a kind of ‘impersonal passive’, as in *dəi-si-i* ‘it’s available’/‘one finds it’. The construction applies with equal facility to intransitive verbs, as in *nə-tə ma-ba-si-i* ‘there isn’t go-able’/‘one doesn’t go there’.

Where detransitivization is accompanied by a nominalization the result is functionally equivalent to a prototypical passive. All trace of the agent is deleted, the patient is promoted to subject status, and the overall construction is stativized, as in *ao səih-si-uzə* ‘this has been killed’.

The ‘middle marker’ *-si* occurs obligatorily on certain intransitive verbs like *cuh-si-nya* [sit-MM-INF] ‘to sit’, *cyā:h-si-nya* [stand-MM-INF] ‘to stand’, etc. In such cases, the intransitive is not a detransitivized derivation of a more basic transitive – these are deponent verbs whose middle marking is semantically motivated (Kemmer 1993). For such verbs, a reduction in valence can be obtained only by the introduction of two *-si* morphemes, the first of which is the lexicalized middle marker, and the second of which deletes the subject, as in:

- (12) *hur-si-si-u*                      *po:*  
 bathe-MM-PASS-NML    place  
 ‘the bathing place/the place where (we/someone) bathed’

The marker *-si* does not occur on auxiliaries to make them agree with main verbs in terms of transitivity. Instead, Kham utilizes suppletive pairs, one for the transitive/ditransitive member and the other for the intransitive/transitive member, as in:

- (13) PERFECT AUXILIARY (intransitive vs transitive)  
 a. *hu-də*    *le*  
 come-NF    be:IMPFV  
 ‘He/she has come.’  
 b. *rəi-də*    *nəi-wo*  
 bring-NF    put-3S:IMPFV  
 ‘He/she has brought it.’
- (14) PERMISSIVE COMPLEMENT (transitive vs ditransitive)  
 a. *ba-o*    *dəi-ke-o*  
 go-NML    receive-PFV-3S  
 ‘He/she got to go/was allowed to go.’  
 b. *ba-o*    *e-ke-o*  
 go-NML    give-PFV-3S  
 ‘He/she allowed him to go.’

### 3.3.3 Negative and interrogative

Both the negative and the Yes-no interrogative make use of the same verbal prefix *ma-*, and are distinguished, in large part, by what morphemes co-occur with them. *Ma-* negative, for example, cannot co-occur with the continuitive morpheme *-zya* in non-nominalized verbs. Where *ma-* does occur in such a context it is always interrogative, as in *ma-ba-zya* [INTRG-go-CONT] ‘Is he going?’

I am assuming that the negative *ma-* is basic, and that the interrogative interpretation has derived from it. In one text I have the rare occurrence of a Yes-no question in which the speaker, musing to himself, says: *həi ŋa-do-ke səi: həi ŋa-ma-do-ke* [thus 1S-say-FUT or thus 1S-NEG-say-FUT] ‘Shall I tell him or shall I not tell him?’ The modern interrogative is likely a truncated version of an earlier two clause Yes-no construction.

In addition to the flip-flop between negative and interrogative interpretations of *ma-* in the modern dialects, the tense/aspect markers *-ke* ‘perfective’ and *-ya* ‘future’ in some dialects



appear to switch interpretations with each other according to whether they occur with the negative or the interrogative. Thus, we get the following:

- (15) a. *ba-ke*  
       go-PFV  
       ‘He/she went.’  
    b. *ma-ba-ke*  
       NEG-go-PFV  
       ‘He/she didn’t go.’  
    c. *ma-ba-ke*  
       INTRG-go-FUT  
       ‘Will he/she go?’
- (16) a. *ba-ya*  
       go-FUT  
       ‘He/she will go.’  
    b. *ma-ba-ya*  
       NEG-go-FUT  
       (ungrammatical)  
    c. *ma-ba-ya*  
       INTRG-go-PFV  
       ‘Did he/she go?’

The flip-flop occurs with near perfect symmetry only in Takale Kham. Comparison with other dialects shows that the modern Takale system is something of an historical accident. Future *-ke* (15c), for example, derives from an earlier *\*-te*, and in transitive paradigms can be shown to occur in a different slot than perfective *-ke*, as in:

- (17) a. *ma-jəi-ke-o*  
       NEG-make-PFV-3S  
       ‘He/she didn’t make it’.  
    b. *ma-jəi-wo-ke* < *\*-te*  
       INTRG-make-3S-FUT  
       ‘Will he/she make it?’

In summary, the morpheme *ma-* has only an interrogative interpretation with continuative *-zya* and perfective *-ya*. With perfective/future *-ke* and imperfective *-e*, *ma-* can be either negative or interrogative. The distinction is less ambiguous in transitive paradigms (as in 17a–b) than in intransitive paradigms (as in 15b–c). In nominalized paradigms as well as in tag questions (see below), *ma-* has only a negative interpretation.

Tag questions are marked by the postverbal particle *ro*, and come with negative or positive presupposition. Usually, though not always, tag questions occur in a nominalized form, as in *nəre*: *o-ba-o ro* [2S-husband 3S-go-NML TAG] ‘Your husband left?’ (It appears he did; am I right?); *nə-re*: *o-ma-ba-o ro* [2S-husband 3S-NEG-go-NML TAG] ‘Didn’t your husband leave?’ (It appears he didn’t; am I right?). It should also be noted that neither Yes-no questions nor tag questions have a sentence final rising intonation.

### 3.3.4 Imperatives

Kham has numerous imperative forms, each of which is a gradation on the time and urgency of the command. Here I will discuss only the major types. The ‘immediate imperative’ is the

**TABLE 42.8** POSITION CLASSES IN THE TAKALE KHAM VERB

-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6	+7
12SUB-DL-(NEG/INTRG)-3OBJ-(CAUS)-root-(PASS)-(CAT)-12OBJ-DL-TAM-3SUB-FUT												
ŋa-		ma-	Ø-	sə-	-si	-na	-na		-ke	-Ø	-ya	
nə-			ni-			-hu	-ni		-zya	-o	-ke	
ge-	n-		ya-/ra-			-ta	-si	-n	-e	-ni		
je-			yara-			-duh	-ci			-rə		
						-rih						

Notes: CAT = concatenated verb root, TAM = tense/aspect/modality.

default imperative and the least marked semantically. It addresses a current problem and elicits an immediate response, as in *em ras-ci-ke* [road release-2P-IMP] ‘(You pl) Clear the way!’

By contrast, the ‘non-immediate’ imperative, marked by *-yo*, gives ongoing applicability to the command and is without a terminal boundary – *ca-o-səba-c-yo* [good-NML-ASC go-2P-IMP2] ‘Go well!’

The hortative builds morphologically on the non-immediate imperative by adding a hortative prefix *gəh* to the string – *gəh-ba-c-yo* [HORT-go-2P-IMP2] ‘Please do go (it’ll be okay)’

The negative imperative (the ‘prohibitive’) employs a prefix *ta-* in place of the hortative *gəh-*. It, too, builds morphologically on the non-immediate template, as in: *ta-ba-c-yo* [PROH-go-2P-IMP2] ‘Don’t go!’

In all the imperatives, the agreement morphemes used for marking imperative subjects (you sg/du/pl) are the same as those used for marking objects in the declarative paradigm. In some sense, then, Kham grammaticalizes the implied performative such that the subject of the imperative ‘you’ is morphologically cast as the object of an implied higher verb ‘I order you’.

Optatives are structured more in keeping with nominalized verb forms, and imply less control over the addressee. Furthermore, subject agreement forms in the optative are identical to subject agreement forms in the declarative, implying that the addressee is still, in some sense, a free agent. Following are examples showing the contrast between the two types:

(18) a. IMPERATIVE

*sya: ap-də səih-na-c-yo*  
 meat shoot-NF kill-GO-2P-IMP  
 ‘Go shoot and kill some game!’ (We need the meat)

b. OPTATIVE

*sya: ap-də je-səih-na-kə*  
 meat shoot-NF 2P-kill-GO-OPT  
 ‘May you shoot and kill some game!’ (Good luck)

### 3.3.5 Position classes

In Table 42.8 is a summary of position classes in the regular (non-nominalized, declarative) Takale Kham verb. The table excludes imperative forms.

## 4 ADJECTIVES AND ADJECTIVALS

‘Adjective’, as a natural word class, is almost non-existent in Kham. The entire class is composed of three native words – ‘big’, ‘small’, and ‘short’ (plus ‘low/short in height’ in some dialects) – and a handful of loanwords from Nepali. All other words that serve in a modifying/ adjectival function are derived by nominalization from some other word class.

#### 4.1 The verbal source

All verbs can be nominalized, and once nominalized, can serve as modifiers to nouns, as in: *ba-o mi*: [go-NML person] ‘the person who went’. Only verbs nominalized from the ‘descriptive’ class, however, (those that predicate properties and change of state, not actions) can co-occur with the existential *li-/le-in* predicate–adjective constructions. Thus, the two verb types are easily distinguished on this feature alone:

- (19) a. DESCRIPTIVE VERB  
*ao khyo:-wo li-zya*  
 this long-NML be-CONT  
 ‘This is long.’
- b. ACTION/PROCESS VERB  
 \**ao ba-o li-zya*  
 \*this go-NML be-CONT

The only predicate construction available to the latter class of verb is the predicate nominative, an equative construction, as in: *ao ba-o zə* [this go-NML EMP] ‘This is one who went’. As such, the latter class is exactly parallel to nouns – *ao mi: zə* [this person EMP] ‘This is a person’, but not \**ao mi: li-zya* [this person be-CONT].

#### 4.2 The nominal source

The derivational pathway from nouns to adjectivals is much longer than the one from verbs to adjectivals, and, in fact, passes through verbs as one stage in the process:

- (20) noun → similitive noun → similitive verb → adjectival/nominalization

Only those nouns that can be conceptualized as representative of an attribute/property can serve as a source for this derivation. Thus, for example, *buchula* is a noun meaning ‘adze’. With a similitive noun derivation it becomes *buchul-ya* ‘something with adze-like qualities’. Adding the intransitive formative *-s* we get *buchul-ya-s-* ‘to become adze-like’, and finally the nominalization gives us *buchul-ya-s-o* ‘adze-like/having a tooth protruding from the upper gums’. The first part of the derivation appears to be Nepali influence, and most of the participating nouns are of Nepali origin – *thutya* < Nep. *ThuTe* ‘hornless’, etc. The latter two stages in the derivation, however, are Tibeto-Burman in origin and occur in native Kham words as well – *ju:ca* ‘old woman’ > *ju:ca-s* ‘to grow old (of a woman)’ > *ju:ca-s-o* ‘old (of a woman)’.

### 5 NOUN PHRASE SYNTAX

Constituents of the noun phrase occur in two basic orders – an unmarked, restrictive ordering in keeping with SOV universals, and a more highly marked non-restrictive ordering.

#### 5.1 Restrictive ordering

In the unmarked order the modifier precedes the noun; the quantifier precedes the modifier, and the demonstrative precedes everything else, as in:

- (21) NP → (Dem) (Quant) (Mod) N

The modifier (Mod) is itself a potentially complex unit with an adverbial intensifier preceding the head, as in ‘very tall’:

## (22) Mod → (Adv) Adj/Nml

There are some occasions for more than one restrictive modifier to occur in a noun phrase. Colour, for instance, is sometimes combined with other qualities in the description of an item. Where it does occur, colour typically follows the more general modifier, and hence, is closer to the noun.

Restrictive relative clauses and other nominalizations also occur commonly in noun phrases. In general, ‘nominalized locatives’ (as in ‘the at-here house’, ‘the to-before person’, etc.) are closely associated with the demonstrative class and occur immediately following the demonstrative, but preceding the quantifier/numeral. Other relatives follow the nominalized locatives.

## (23) NP → (Dem) (Loc) (Rel) (Quant) (Mod) N

- (24) *no ɲah-da-ɲao ɲa-zihm-kə hū:-wo gyo:h-wo mi:-ni*  
 that before-ALLT-NML 1S-house-LOC come-NML big-NML person-DL  
 ‘those earlier two people who came to my house’

## 5.2 Non-restrictive ordering

Non-restrictive ordering occurs where some or all of the elements in the noun phrase are in apposition to the head. The noun (with its restrictive elements) occurs first, and everything else follows – a reversal of what we have already seen. The only exception is with the demonstrative; the demonstrative is always restrictive, and hence, always precedes the noun. Thus taking the formula in (21) and reversing all but the demonstrative, we get the following:

## (25) NP → (Dem) N (Quant) (Mod)-num

A critical piece in non-restrictive ordering is the obligatory number agreement (num) on the final element of the appositive, showing, in fact, that we are no longer dealing with one unit, but two juxtaposed units, as in:

- (26) *no mi:-ni, nehblo bənəi gyo:h-wo-ni*  
 that person-DL, two very big-NML-DL(num)  
 ‘those people, the two really big ones’

## 6 NOMINALIZATIONS

There are two basic nominalization types in Kham, one without person/number inflections and the other fully inflected for person and number. In relative clause constructions, the first type functions as an ‘agent nominalization’, and the second as an ‘object nominalization’:

- |                              |                          |
|------------------------------|--------------------------|
| (27) a. AGENT NOMINALIZATION | b. OBJECT NOMINALIZATION |
| <i>jəi-wo</i>                | <i>o-ra-jəi-wo</i>       |
| make-NML                     | 3S-3P-make-NML           |
| ‘the one who made (it/them)’ | ‘those that he made’     |

Both nominalization types occur also as sentential complements, but where they do, the less finite type bears closer syntactic integration with the matrix verb than its inflected counterpart does. Thus, for example, the non-finite nominalization would be used in sentences like ‘He knew how to make it’, while the finite form would be used for ‘He knew that she was making it’.

Nominalized verb forms (of the inflected type) also function as unembedded, main clause predicates. Having begun as verbs, they have come full circle and once again function in the capacity of stand-alone predicates with much the same distributional potential as any finite verb. Their primary role is within the domain of a narrative discourse where their relatively ‘time-stable’ character (Givón 1984) is used to demarcate features like ‘stage-settings’, background information, and other properties of discontinuity.

Nominalized verbs, in fact, play such a major role in the grammatical mechanisms of Kham that they can be seen cutting across the paradigm of every major speech act – not just the declarative, as we have seen, but also of the imperative and the interrogative. Nominalized imperatives have the softened force of an optative, and nominalized interrogatives are less intrusive than their regular counterparts.

## 7 CLAUSE CHAINS AND SWITCH REFERENCE

Clause chains are central to the structure of Kham sentences. In clause chains, all tense/aspect and person/number information is marked on the chain-final verb, and chain-medial verbs are marked with varying degrees of inflection depending on whether they are ‘same-subject’ or ‘different-subject’ clauses. In same-subject chains, the medial verb is non-finite, being composed of a simple root and a ‘non-final’ suffix *-də*, while different-subject chains are marked by the adverbial *-kə* ‘while’ and fully inflected for person and number as well as some tense/aspect specifications.

It turns out, then, that switch reference in Kham is dependent not on a switch reference marker *per se*, but on referential continuity vs discontinuity – and in keeping with universal expectations, continuity is unmarked while discontinuity is marked. In essence, we come away with Haiman’s (1983) general formulation for switch reference in coordinate structures:

(28) DS:SS = agreement marker:  $\emptyset$

In some same-subject chains, medial and final verbs specify different events, while in others, they specify different aspects of the same event. Morphologically there is no distinction. The differences lie, in part, on the level of ‘juncture’ between the two clauses (see Van Valin and LaPolla 1997: 454) and various syntactic tests can be devised for teasing them apart. The perfect, for example, is marked by ‘be’ (intransitive) or ‘put/keep’ (transitive) as the final verb/auxiliary in a chain, but the two verbs are conceptually a part of the same event, as in:

(29) *o-za:*      *zihm-kə*      *rəi-də*      *nəi-wo*  
 3S-child    house-LOC    bring-NF    keep-3S:IMPFV  
 ‘She has brought her child to the house.’

Only where some other sentential constituent comes between the two verbs can the interpretation be forced to a sequence of two events, as in:

(30) *o-za:*      *rəi-də*      *zihm-kə*      *nəi-wo*  
 3S-child    bring-NF    house-LOC    keep-3S:IMPFV  
 ‘She brought her child and kept him in the house.’

In other less grammaticalized chains, separate verbs can only represent sequences of events, regardless of syntactic order, as in *kā: zya-də hu-ke* ‘Having eaten food, he came’.

## 8 EVIDENTIALS

Kham lacks most evidential categories found in many Bodic-type languages. There is, however, a ‘mirative’ and a ‘reportative’ that covers at least some of the semantic space often associated with the evidential categories of hearsay and inference.

The mirative construction is marked by the auxiliary use of the existential copular verb ‘to be’ in third person singular – *oleo*. In fact, both the auxiliary and the main verb occur in nominalized form, and all inflectional categories occur on the main verb, as in:

- (31) *o-ra-bəi-wo*    *o-le-o*  
 3S-3P-take-NML    3S-be-NML  
 ‘(I see) he took them!’

The mirative is about newly discovered information; information not yet integrated into the speaker’s store of knowledge. It makes no claims about the source of information, but only about its newness and the speaker’s apprehension of it. As such, it can occur with first-hand observation, inference, or hearsay (DeLancey 1997). Whatever its source, the speaker takes personal responsibility for the claims – if it is based on inference or hearsay, the conclusions are the speaker’s.

The reportative, by contrast, (which is marked by the postverbal particle *di*) is based on hearsay. It can be newly apprehended knowledge, but the speaker chooses not to convey it as such, but rather according to its source. That is, the speaker makes no claims about the truth of the statement. The apprehension of truths leading up to the conclusion are made by someone else, and the speaker disclaims responsibility for them, as in the following (which contrasts with the mirative marked statement in 31):

- (32) *ya-bəi-ke-o*    *di*  
 3P-take-PFV-3S    REP  
 ‘He took them (or so it’s said).’

## REFERENCES

- Bauman, James (1975) ‘Pronouns and pronominal morphology in Tibeto-Burman’, unpublished PhD dissertation, University of California Berkeley.
- DeLancey, Scott (1980) ‘Deictic categories in the Tibeto-Burman verb’, unpublished PhD dissertation, Indiana University.
- (1987) ‘Sino-Tibetan languages’, in Bernard Comrie (ed.) *The World’s Major Languages*, New York: Oxford University Press, 799–810.
- (1997) ‘Mirativity: the grammatical marking of unexpected information’, *Linguistic Typology* 1: 33–52.
- Dryer, Matthew (1986) ‘Primary objects, secondary objects, and antitativity’, *Language* 62.4: 808–45.
- Givón, T. (1984) *Syntax: A Functional-Typological Introduction*, Vol. 1, Amsterdam/Philadelphia: John Benjamins.
- Haiman, John (1983) ‘On some origins of switch reference marking’, in John Haiman and Pamela Munro (eds) *Switch reference and universal grammar*, Amsterdam/Philadelphia: John Benjamins, 105–28.
- Hale, Austin and Pike, Kenneth L. (1970) ‘Tone systems of Tibeto-Burman languages of Nepal’, *Occasional Papers of the Wolfendon Society on Tibeto-Burman Linguistics*, Volume III, Part I: Studies on tone and phonological segments, Urbana: University of Illinois.
- Hari, Maria (1980) *An Investigation of the Tones of Lhasa Tibetan*, Language Data, Asian-Pacific Series no. 13, Huntington Beach CA: Summer Institute of Linguistics.

- Henderson, Eugénie J. A. (1957) 'Colloquial Chin as a pronominalized language', *Bulletin of the School of Oriental and African Studies* 20: 323–27.
- Kemmer, Suzanne (1993) *The Middle Voice*, Typological Studies in Language, Vol. 23, Amsterdam and Philadelphia: John Benjamins.
- Michailovsky, Boyd (1975) 'On some Tibeto-Burman sound changes', *Berkeley Linguistics Society* 1: 322–32.
- Noonan, Michael (2002) This volume.
- Shafer, Robert (1966) *Introduction to Sino-Tibetan*, Wiesbaden: Otto Harassowitz.
- Van Valin, Robert D., Jr, and LaPolla, Randy J. (1997) *Syntax: Structure, Meaning and Function*, Cambridge/New York: Cambridge University Press.
- Watters, David E. (2002) *A Grammar of Kham*, Cambridge Grammatical Descriptions, ed. by R.M.W. Dixon and Keren Rice. Cambridge: Cambridge University Press.
- Watters, David and Watters, Nancy (1973) *An English-Kham, Kham-English Glossary*, Kathmandu: Tribhuvan University.

# LEPCHA

*Heleen Plaisier*

## 1 INTRODUCTION

The Lepcha people call themselves *Róng* or *Róngkup* ‘children of the Róng’, in full *Mútuncí Róngkup Rumkup* ‘Children of the Róng and of God’, and their language is called *Róng-róng*. 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 probably by upwards of 50,000 speakers. The genetic position of Lepcha within Tibeto-Burman is still unclear, despite the efforts of Hodgson (1857), Waddell (1899), Konow (1909), Shafer (1950), Forrester (1962), Benedict (1972) and Bodman (1988).

Archibald Campbell published a first note on the Lepcha language in 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, Albert Gruenwedel edited and published the Lepcha-English dictionary compiled by George Mainwaring posthumously. 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 has written an unpublished grammar of Lepcha as a Ph.D. dissertation at Deccan College in Pune in 1966. Richard Keith Sprigg has written numerous valuable articles on the Lepcha language, the indigenous Lepcha script and Lepcha history. Native Lepcha scholars such as Khárpú Támsáng and Dóngtshen Luksóm have published important studies on the Lepcha language and culture written in the Lepcha language. More recently, the Lepcha Textbook Department of the Government of Sikkim and the various Lepcha Associations of Sikkim and Darjeeling district have been publishing periodicals, books, plays and collections of poetry in Lepcha. This article is based on fieldwork conducted by myself in the Lepcha-speaking area between 1994 and 1998.

## 2 PHONOLOGY

The specialist literature on Lepcha is replete with accounts of Lepcha phonology. However, most accounts deal only with a single aspect, and there appears to be no single analysis of Lepcha phonology as a whole which is favoured by consensus amongst specialists. The full inventory of Lepcha consonants is given in Table 43.1. The differences in analysis are most evident in the realm of Lepcha vowels, where traditional orthography has contributed to confusing scholars, both native and foreign. Table 43.2 gives my current analysis of Lepcha vowel phonemes, but please note that this analysis may well change in due course.



TABLE 43.1 CONSONANTS

	<i>Voiceless stop</i>	<i>Aspirated voiceless stop</i>	<i>Voiced stop</i>	<i>Voiceless affricate</i>	<i>Aspirated voiceless affricate</i>	<i>Voiced affricate</i>	<i>Voiceless fricative</i>	<i>Voiced fricative</i>	<i>Nasal</i>	<i>Approximant</i>
Labial	p	ph	b				f	v	m	w
Dental	t	th	d				hl		n	l
Alveolar				ts	tsh	z	s			r
Retroflex	tr	thr	dr							
Palatal				c	ch	j	sh		ny	y
Dorso-velar	k	kh	g						ng	
Laryngeal							h			

TABLE 43.2 VOWELS

<i>Vowel phoneme</i>	<i>Phonetic realizations</i>	<i>My transliteration</i>
/a/	[ʌ]	a, â
/á/	[a]	á
/i/	[i]	i, í
/o/	[o ~ o˥ ~ u]	o
/ó/	[ɔ ~ ɔ˥]	ó
/u/	[u]	u
/ú/	[u]	ú
/e/	[e˥ ~ e ~ ε]	e

The transcription that is used here is in fact a transliteration of the native Lepcha orthography. The relationship between the modern pronunciation and the traditional orthography is systematic to the point that a complete and accurate phonological transcription can be readily derived from a faithful transliteration of Lepcha orthography. Native Lepcha orthography makes two distinctions which do not correspond, or no longer correspond, to any real phonological contrast in the modern language. These are the distinctions between the vowel sounds which are transliterated as *a* and *â* and those transliterated as *i* and *í*. My transliteration also represents these two non-phonological or no longer phonological distinctions made in native Lepcha orthography. Wherever the phonological reality of modern Lepcha differs from modern Lepcha orthography, I follow native orthography in my transliteration and give the phonological reality between phoneme brackets along with it. However, whenever a word or expression is so colloquial that it does not yet have a proper Lepcha spelling, especially in transcribing informal conversations, my transliteration is a straightforward phonological transcription. (In general, it may be observed that phonetically a vowel in open syllables has a longer realization than the same vowel in a closed syllable.)

The retroflex consonants /tr/, /thr/, and /dr/ are treated as consonant clusters in the traditional orthography, but in terms of modern phonology these digraphs represent retroflex consonants. The retroflex series are written as 'kr' for /tr/, 'hr' for /thr/, and 'gr' for /dr/ in the native orthography. The phoneme /tr/ is an unaspirated voiceless retroflex plosive [t]. The phoneme /thr/ is an aspirated voiceless retroflex plosive [tʰ], and the phoneme /dr/ is an unaspirated voiced retroflex plosive [d]. To avoid confusion with the consonant clusters 'kr', 'hr', and 'gr', which are not distinguished from the retroflex consonant clusters in the native orthography (or perhaps rather vice versa, as retroflex consonants mainly occur in loans from Tibetan), Mainwaring introduced a small dot, which was to be added below those clusters

**TABLE 43.3** INITIAL CONSONANT CLUSTERS

	-y-	-r-	-ry-	-l-	-ly-
<i>k</i>	ky	kr tr	kry try	kl	kly
<i>kh</i>	khy				
<i>g</i>	gy	gr dr ngr	gry dry	gl	gly
<i>ng</i>					
<i>t</i>	ty				
<i>th</i>	thy				
<i>d</i>	dy				
<i>p</i>	py	pr	pry	pl	ply
<i>ph</i>	phy				
<i>f</i>	fy	fr	fry	fl	fly
<i>b</i>	by	br		bl	bly
<i>m</i>	my	mr	mry	ml	mly
<i>r</i>	ry				
<i>l</i>	ly				
<i>h</i>	hy	hr thr	hry thry		hy
<i>v</i>	vy				

indicating a retroflex sound, to be used both in the native orthography as well as in transliterations. As the initial consonant clusters ‘tr’, ‘thr’, and ‘gr’ are not found in Lepcha, I have chosen to transcribe the retroflex series as /tr/, /thr/, and /dr/ instead of using the dot. In Lepcha, texts written in the native orthography, *Mainwaring’s dot* is occasionally found.

Table 43.3 lists the initial consonant clusters found in Lepcha. It should be noted here that the consonant clusters /kl/, /gl/, /pl/, /fl/, /bl/, and /ml/ are represented by distinct graphemes and are not treated as clusters in the native orthography. The voiceless dental fricative /hl/ is represented by a separate grapheme in native orthography as well, and should be analysed as a separate morpheme in modern Lepcha phonology.

Syllables in Lepcha have the following structure: (Ci) V (Cf), in which Ci represents one of the optional initial consonants or consonantal clusters, and Cf stands for one of the optional final consonants. All syllables contain a vowel as their core. All consonants may appear in initial position, but not all consonants can be combined with -y-, -r-, -ry-, -l- or -ly- to form initial consonant clusters, as can be concluded from Table 43.3. Not all of the vowel phonemes can be combined with the initial consonant clusters listed in Table 43.3, but due to lack of space all of the possible combinations cannot be given here. Only k, t, p, ng, n, m, r, l can occur in final position.

### 3 NOMINAL MORPHOLOGY

Nominals comprise the following parts of speech: nouns, pronouns, adjectives and numerals. Nouns lack grammatical gender distinction and show no agreement with articles, adjectives or verbs. However, Lepcha has rich derivational nominal morphology, which also includes gender-specific derivational suffixes. Affixing, compounding and reduplicating represent the major derivational processes. Adjectives may be used adnominally, predicatively, or independently as nominal heads.

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 existence of a dual suffix <-nyum>, as mentioned by Mainwaring (1876: 27), Sinha (1965: 67) and Tamsang (1978: 10). On the contrary, I have found that the numeral *nyet* 'two' is used to indicate dual number of nouns.

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. 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 and not a genuine indefinite article in the Western European sense 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.

Nominals may take case endings, i.e. suffixes 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. 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), and perhaps the nominalizing suffix <-bú> (NOM) as well, 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. For the different senses of these morphemes, the same glosses are used throughout this article in keeping with the analysis of these morphemes as indexing single grammatical categories with a uniform *Gesamtbedeutung* in all cases. The fact that the different morphemes in its various functions synchronically represent single grammatical categories in modern Lepcha does not necessarily preclude that the

syntactically distinct uses of these morphemes might very well derive from historically distinct etyma through convergent evolution or analogy.

The meaning of the ablative suffix <-nun ~ -nu>, phonetically [nun ~ nu], 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. This is just one facet of a wider socio-linguistic phenomenon in Lepcha, for in Lepcha textbooks, norms of speech are often prescribed which deviate from actual norms observed in natural speech. The fact that these tendencies are not observed consistently and that there is interference from cultivated norms suggest that the situation is in flux. Bilingualism may have exacerbated the process of change because in certain areas fluent monolingual speakers of Lepcha are in the minority.

One sense which the suffix <-nun ~ -nu> expresses when affixed to nouns and adverbs denoting a place or time is the ablative, viz. 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-PRES  
 '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 mediative suffix 'via, by means of which, from', as a noun in the meaning 'road, way' and even as a verb in the meaning 'walk'.

There are a number of Lepcha suffixes which are similar in meaning but distinct from both the ablative and the mediative suffixes. One of these is the suffix <-ren> '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-nyí-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 or ergative function. The classic ergative marks the agent of a transitive verb. Ergativity in Lepcha is manifested by the behaviour and meaning of the suffix <-nun ~ -nu> when attached to the agentive argument of a transitive verb, and in this function <-nun ~ -nu> functions quite differently to a classic ergative. The suffix <-nun ~ -nu> is used to highlight the agentive character of an inanimate entity which acts as the agent of an activity, whether this be transitive or intransitive. In an ergative 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. The ergative 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.

Another Lepcha morpheme which is appropriate to discuss in this context is the morpheme <-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’.

- (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-loose  
 ‘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 instrumental sense of the genitive morpheme <-sá> requires some discussion and must be contrasted with the ablative suffix <-nun ~ -nu>. The genitive 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-PRES  
 ‘I wash my hands with water’.
- (14) *Go nyúgú-sá pi-sho.*  
 1sg pen-GEN write-PRES  
 ‘I write with a pen’.
- (15) *Go ámík-sá ngâk-bám.*  
 1sg eye-GEN look-PROG  
 ‘I look with my eyes’.

The suffix <-nun ~ -nu> marks the unwitting 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 <-nu> 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-PRES  
 ‘What do you do with/by means of a pen?’.

(17) *nyúgú-sá shú zúk-sho?*  
 pen-GEN what make-PRES  
 ‘What do you do with a pen?’.

The comitative suffix <-dep> ‘together with’ (COM) indicates accompaniment, and is often used in conjunction with or as an alternant of the Lepcha genitive suffix <-sá>.

(18) *Hó hu-dep nú o!*  
 2sg 3sg-COM go EMO  
 ‘You, go with him!’

The meaning of the locative suffix <-ká> (LOC), is one of direction, of goal toward which the action or situation is directed, and the locative suffix 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 saryók-ká bám-sho.*  
 tiger jungle-LOC reside-PRES  
 ‘The tiger lives in the jungle’.

(20) *Go Kalenpúng-ká nóng-sho.*  
 1sg Kalimpong-LOC go-PRES  
 ‘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 43.4. 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

**TABLE 43.4 PERSONAL PRONOUNS**

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
1st person	<i>go</i>	<i>kányí</i>	<i>káyú</i>
2nd person	<i>hó</i>	<i>ányí</i>	<i>áyú</i>
3rd person	<i>hu</i>	<i>hunýí</i>	<i>huyú</i>

TABLE 43.5 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'

animals are personified, when the third person pronoun is used to refer to these animals. The form *ka*, which is found in certain expressions such as *ka Mútuncí Róngkup Runkup* '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 (1sg *kásu*, 2sg *ádo*, 3sg *hudo*), e.g. *kasu-sá lí* 'my house', *ádo-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 43.4, e.g. *káyú-sá cho* 'our book', *huyú-sá cho* 'their book', etc.

The major demonstrative and corresponding interrogative pronouns are listed in Table 43.5. The proximate deictic morpheme <á-> refers to something or someone near to the speaker and the distant deictic morpheme <o-> indicates something or someone far from the speaker. 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'.

#### 4 VERBAL MORPHOLOGY

All Lepcha verbs have in principle two stem forms, viz. 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 or stem-final vowel 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 *dít* vs *dí* 'reach'. The inflected stem cannot be predicted on the basis of the form of the verb.

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 knowing how to perform a particular task or activity is expressed by the verb *khut* 'can'. The auxiliary *lyók* (PRB) expresses probability; when used as a main verb, *lyók* means 'resemble, look like'.

When used with verbs, the nominalizing suffix <-bú> (NOM) conveys an imperfective meaning and indicates a state. This suffix has close parallels in other languages of the Himalayas, where there often is a morpheme which both nominalizes verbs and clauses and, when affixed to the main verb of a sentence, marks a type of imperfective aspectual meaning.

There are a number of verbal auxiliaries expressing *Aktionsarten* in Lepcha, such as *nón* 'go', *tho* 'end', and *lel* 'complete'. The perfective auxiliary *nón* 'go', for example, expresses

a past tense activity as completed, or a past tense event or transition, the result of which has been attained. The terminative auxiliary *tho* ‘end’ indicates that an activity or action has come to an end. The completive auxiliary *lel* ‘complete’ indicates that an activity is completed or fulfilled.

A present tense is expressed by adding the verbal ending <-sho> (PRES) to the verb. The present 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 of which the speaker is certain or convinced that is going to take place soon. In this latter sense, the present future is commonly used to express an activity that is about to take place.

- (23) *Hlo-ká sozóng-sho.*  
 peak-LOC be.cold-PRES  
 ‘It is cold on the hill’.
- (24) *Go vâm-kát theng-sho.*  
 1sg song-one sing-PRES  
 ‘I will sing a song’.

The progressive tense is formed by adding the auxiliary <-bám ~ -wám ~ -ám> (PROG) to the main verb, e.g. *zu-wám* [live-PROG] ‘living, alive’; *Hu lok-bám* [3sg dance-PROG] ‘She is dancing’. The progressive tense expresses an activity or situation presently in progress. Verbs of motion use the auxiliary <-det> ‘move’ to express a progressive tense. When used as a main verb, *bám* means ‘to dwell, reside’.

The preterite tense is expressed by using a simple verb and denotes actions anterior to the speech moment without reference to result or duration. This tense is not overtly marked and the preterite meaning is often indicated by the context or by lexical means in the sentence. When the subject is marked by the ablative suffix <-nun ~ -nu>, sentences can often be interpreted as having a past meaning.

- (25) *Hu nón ma.*  
 3sg go AST  
 ‘He went away’.
- (26) *Go-nun yuk pi ma.*  
 1sg-ABL letter write AST  
 ‘I wrote a letter’.

The verbs in Lepcha which cover the senses of English ‘to be’ are *go* ‘be’ and *nyí* ‘have’. The verb *go* ‘be’ is used as an identity marker, to say that X is Y. The form *gum*, glossed as ‘be + AST’, is a contracted form of <go-ma>. The assertive particle *ma* (AST) is commonly used with both copulas, even with the fully lexicalized form *gum*. The verb *nyí* ‘have’ covers the attributive, locative and perhaps also existential senses of English ‘to be’.

- (27) *Áre kásu-sá lí go ma.*  
 this 1sg-GEN house be AST  
 ‘This is my house’.
- (28) *Lyáng áre rong-kup bâm-lyáng gum.*  
 land this Lepcha-child live-land be + AST  
 ‘This land is the homeland of the Lepcha’s’.



- (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'.

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!*  
 1pl-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. These implications follow from the meaning of the ablative suffix, but are not components of its meaning.

- (35) *Ázóm zóm-lel-nu rok o.*  
 rice eat-complete-ABL read EMO  
 'After you have eaten your food, you should study'.
- (36) *Ámík ok-nu ngâk gang gun shí.*  
 eye open-ABL look if all see  
 'If you open your eyes and look, you will see everything'.

## 5 TEXT

*Renjong lyáng áre-re ká mútuncí-róng-kup rum-kup-sang-sá*  
 Renjong land this-DEF 1pl arch-Lepcha-child god-child-PL.H-GEN

*ayá ayá lúngtyen-nu-do thop-yut-bú-sá*  
 long.ago long.ago legend-ABL-self obtain-send.down-NOM-GEN

*áyít ágyek álát ábám án ámák lyáng gum ma o.*  
 creation birth arrival residence and death land be+AST AST EMO

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 málúk renjong lyáng gum ma o*  
 Ne-Mayel Málúk Renjong land be+AST AST EMO

*án róng-sang-nu-re huyú-do-sang-re-m mútuncí-róng-kup-rum-kup*  
 and Lepcha-PL.H-ABL-DEF they-self-PL.H-DEF-DAT arch-Lepcha-child-god-child

*yang lúngprya-nu lí-bám ma o.*  
 thus repeat-ABL say-PROG ASS EMO

The oldest Lepcha name of this land is Ne Mayel Málúk Renjong, and the Lepchas always call themselves Mútunci Róngkup Rumkup.

*róng lungtyen-sá áyít-sung ká-róng-sáng-sá púmnyothing-re-m-re*  
 Lepcha legend-GEN origin-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-NOM

*cú kíngtsúm zóngbú-cú-sá anyák-sá átsóng-bú sanóng-lóm*  
 mountain Kanchenjunga-mountain-GEN summit-GEN pure-NOM snow-via

*áyítbú-debú-rum-nu zúk-nu áyít-thóm-bú ma o*  
 Áyítbú-Debú-god-ABL make-ABL origin-end-NOM ASS EMO

According to the traditional history of origin, the forefathers of us, Lepchas were created and made by the God Ayítbú Dyebug 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 Kharpú Támsáng and privately published by Lyángsong Támsáng in Kalimpong, India in 1996.)

## ADDITIONAL ABBREVIATIONS

AST	assertive
EMO	emotive
PL.H	plural human
PL.NH	plural non human
PRB	probability

## 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', *Linguistics of the Tibeto-Burman Area* 11 (1): 1–26.
- Bodman, Nicholas (1989) 'Some remarks on Lepcha vowels', in David Bradley *et al.* (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* IX: 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* XVII: 73–83.
- Konow, Sten (1909) 'Tibeto-Burman family', Vol III in Grierson, G.A., *Linguistic Survey of India*, Calcutta: Superintendent of Government Printing, India.
- Luksóm, Dóngxshen (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 Sikkim Hills*, Calcutta.
- Mainwaring, George Byres (1898) *Dictionary of the Lepcha Language*, revised and completed by Albert Grünwedel, Berlin.
- Shafer, Robert (1950) 'Classification of some Languages of the Himalayas', *Journal of the Bihar Research Society* 34: 192–214.
- Sinha, Prabhakar (1966) *A Descriptive Grammar of Lepcha*, unpublished Ph.D. dissertation, Pune: Deccan College.
- Siiiger, Halfdan and Jørgen Rischel (1967) *The Lepchas. Culture and religion of a Himalayan people*, Copenhagen: Gyldendal, Nationalmuseets Skrifter, Etnografisk række 11.
- Tamsang, Khárpú (1978) *A Grammar of the Lepcha Language*, Kalimpong: Lyangsong Tamsang.
- Tamsang, Khárpú (1994) *The Lepcha-English Encyclopedic Dictionary*, Kalimpong: Mayel Clymit Tamsang.
- Waddell, Lawrence Austine (1899) 'The Lepchas or Róng and their songs', *Internationales Archiv für Ethnographie* XII: 41–57.

# INDEX

- Abor 18, 31, 180  
Abor-Miri-Dafla 18, 31, 180, 427  
Achang 8, 47, 195  
Adi 18, 31, 180–1, 456–7, 462  
Ahi 8  
Aimol 13, 188  
Aka 19, 180  
Akha 8–9, 35, 208, 236–51  
Almora 14, 16–17  
Altaic 7, 152, 362  
Amdo 10, 270, 594  
Anal 13, 188  
Andro 11, 175–6, 178, 185  
Angami 12, 44, 173, 183–6, 188  
Angami-Pochuri 12, 183–5, 186, 188  
Anhui 27, 102, 109  
Anlù 105  
Anong 15, 674  
Anqing 105  
Anxiang 88  
Anyáng 86  
Anyì 86  
Anyo 12  
Ao 12, 183–4, 186, 188  
Ao-Chungli 12, 183–4  
Ao-Mongsen 12, 183–4  
Apa Tani 18, 180  
Apatani 18, 31, 33, 179, 181, 456–61,  
463, 465  
Arakan and Arakanese 8, 195  
Arleng 14, 187  
Arunachal Pradesh 31, 169–74, 176–81, 189,  
222, 439, 456  
Assam 169–72, 177, 180, 187, 189,  
387, 427, 456  
Assamese 170  
Athpare 15, 505–15, 539, 546  
A'tong 11, 175–7, 387  
Atsi 195 *see also* Zaiwa  
Aytutthia 196
- Badikhel Pahari 371  
Badong 105  
Baglung District 315, 683  
Bahing 15, 23, 506, 516, 560
- Bai 6, 19–20, 28, 31, 43–5, 50, 54,  
73, 82, 651–72  
Bái Mán 31  
Balti 9, 270  
Bàngbù 109  
Bangladesh 31, 169–70, 177–8, 189,  
387, 399–400, 409  
Bangru 19, 179, 456  
Bangru/Levia 180  
Banjogi 13  
Banpara 11, 177  
Bantawa 15, 16, 505–12, 516, 539  
Bǎoji 86  
Bǎojǐng 88  
Bǎoshan 88, 96, 119  
Bǎoshan Luódiàn 88, 96  
Bǎoshan Shuāngcǎodun 88, 96  
Barish 11, 176  
Bawm 13, 409  
Bayi 456  
Beifang 6  
Beijing 6, 72–3, 76, 78, 81, 126–7, 579  
Belhare 15, 33, 505–6, 512–15, 546–69  
Bengal 178  
Bengali 170, 399–400  
Bengni 20, 31, 179, 181, 456–9, 461–3  
Bete 13  
Bhaktapur 371  
Bhutan 10, 29, 169–70, 179, 270, 439–40,  
456, 705  
bilingualism 12, 18, 237, 356, 371–2,  
399, 469, 518, 546, 651, 658  
Bisoid 8, 236  
Bisu 8  
Black Barbarians 31  
Black Lahu (BL) 208–11 *see also* Lahu  
Black Lisu 8, 222–7, 232, 234 *see also* Lisu  
Black Lo *see* Black Lisu  
Bodic and Bodic languages 9–11, 29, 45, 47,  
52, 54, 255, 518, 703  
Bodish 45, 47, 255, 687  
Bodish-Himalayan 516  
Bodo (Boro) 11–12, 23, 32, 45, 171, 173–7  
Bodo-Garo 11, 32, 45, 176, 427  
Bodo-Koch 11, 176–8, 182, 187

- Bodo-Konyak-Jinghpaw languages 174–8, 188  
 Bokar 19, 179, 181, 456–65  
 Bola 8, 12, 33, 401  
*Book of Poetry* 26  
 Bori 179, 181  
 Boro 11, 175–7  
 Brahmaputra 169–70, 181  
 Brahmi 196, 653  
 Brè 18, 623  
 bronze scripts 161  
 bronzes 160  
 Buddhists 179, 196, 632  
 Bugun 19, 173, 179, 180  
 Bumtang 10, 179  
 Bunan 16  
 Burma 8  
 Burmese-Lolo *see* Lolo-Burmese  
 Burmese-Yiish 195  
 Burmese-Yipho 236  
 Burmish 8–9, 12, 33, 47, 177, 178, 195  
 But-pa 180  
 Bwe 18  
 Bwe Karen 44  
 Byangsi 16, 30  
  
 calligraphy 160, 161  
 Cambodian 196  
 Camling 15–16, 33, 505–16, 533–45  
 Cangluo Menba 179  
 Cangluo Monpa 10, 439  
 Canton 146  
 Cantonese 6, 28, 49–50, 88, 90–1, 95–101,  
 105, 108, 110–13, 116–24, 138, 146–54  
 Caodeng 490–502  
 Caura Khāni 315  
 Central Chin 13, 409  
 Central Institute of Indian Languages  
 (CIIL) 172  
 Central Loloish 8  
 Central Modern Lhasa 9  
 Central Monpa 9–10, 179, 439  
 Central Plains 77, 127  
 Chabao 14, 490  
 Cháling 86  
 Chamba Lahuli 16  
 Chang 11, 175–7, 185, 273  
 Changlí 118  
 Chángsha 76, 80, 107, 116, 119  
 Chángshú 78, 85, 89, 95–6, 117  
 Chángting 86  
 Changzhi 128, 129  
 Chángzhou 78, 85, 87, 88, 94, 96, 113, 119  
 Chantyal 10, 291, 315–34, 336, 338,  
 347, 349, 683  
 Cháoxiàn 106  
 Cháozhou 81, 86–7, 89, 93, 153  
 Chaudangsi 16, 30  
 Chawte 13  
 Chéngbù 85  
 Chengdu 93, 109, 129, 130, 579, 583  
 Chénxi 93  
 Chepang 14, 16–17, 516, 518, 684, 695  
 Chiang Mai 208  
 Chiang Rai 208, 623  
 Chilling 15, 546  
 Chin 12–14, 16, 31, 47, 182, 184–5,  
 188, 409, 516, 695  
 Chin Hills 409, 426  
 Chindwin River 195  
 Chingmengnu 11, 177  
 Chiru 13, 188  
 Chittagong Hills 178, 195  
 Chokri 12, 183–6  
 Chóngmíng 79  
 Chote 13  
 Chothe 13, 188  
 Chulikatta 19, 181  
 Chungli 183, 184  
 Chutia 11, 177  
 Chúxiàn 109  
 Classical Chinese 49, 164  
 Classical Tibetan 9, 10, 27, 255–67  
 Common Chinese 72–3, 77, 79, 81  
 Cóngguà 86  
 Cuona Menba 10, 179  
  
 Daai 13  
 Dafla 18, 180  
 Dali Bai 651  
 Damu 18, 181, 456–7  
 Dangshan 109  
 Danyáng 85–6, 88, 96, 105, 119  
 Daofu 14, 30  
 Dàoxiàn 115  
 Darma 30  
 Darmiya 16  
 Dàtián 86–7  
 Dàxué 73, 81  
 Dazu 456  
 Dehong Dai and Jinghpo Autonomous  
 Prefecture 401  
 Delugong 32  
 Deori 11, 175–7  
 Deuri 11, 45  
 Devanagari alphabet 356  
 Dhammai 19, 173, 179–80, 456  
 Digarish 45, 456  
 Digaro Mishmi 19

- Digaru 19, 179, 181  
 Dimasa 11, 175–7  
 diminutives 27  
 Dingxi 109  
 Dingyuǎn 109  
 Dolakhā Newār 355  
 Dolakhāe 355–70  
 Dongguǎn Qingxi 86  
 Dòngkǒu 119  
 Dravidian 16, 53, 516  
 Dūchāng 79, 87  
 Dulong 14–15, 22, 30, 32, 36, 674–82  
 Dumī 15, 505–6, 512, 514–15, 560  
 Dunganese 85  
 Dun-huang 255  
 Dwags 10, 179  
 Dwāri 315  
 Dzungka 29  
 Dzungkha 10, 282, 439
- Eastern Himalayan 255  
 Eastern Kayah Li 623–6  
 Eastern Lisu 222  
 Eastern Zhou 160  
 Empeo 12, 183, 186  
 Enpíng 86  
 Enshi 105  
 Ergong 14, 30  
 Ersu 17  
 Èzhou 102
- Falam Lai 413  
*fanqie* spellings 164  
 Féidong 109  
 Fěixi 109  
 Fènghuáng 115  
 Fèngtái 109  
 Flowery Lisu 8, 222–7, 232 *see also* Lisu  
 Fúfeng 86  
 Fújiàn 79–80, 86  
 Fùnán 109  
 Fùyáng 109  
 Fúzhōu 81, 87, 93, 103
- Gahri 16  
 Gallong 18, 179, 181, 456–7, 459  
 Gan 6, 72, 76, 79–80, 84–8, 90, 92–4,  
 98–109, 146  
 Gangte 13, 188  
 Gansu 85, 95, 100, 102–3, 107, 109, 118, 126  
 Gari 16  
 Garo and Garos 11, 170–1, 173, 175–7,  
 387–400  
 Garo Hills 176–7, 387, 399–400
- Ghyās Kharkā 315  
 Gokhy 8  
 Gondla 16  
 Gongan 88  
 grammatical relations 27, 29, 561, 564,  
 628, 669, 676  
 Guangdong 27, 80–1, 85–6, 92, 146  
 Guǎngjì 86  
 Guangxi 27, 85, 109, 146  
 Guǎngxī 77, 78  
 Guangzhou 73, 76, 80, 81, 146–7, 149–50  
 Guanhua 127–8  
 Guanyinqiao 14  
 Guanzhong 128  
 Gùchéng 118  
 Guiqiong 17, 30  
 Guiyáng 93, 116  
 Gùiyáng 115  
 Guizhou 77, 78, 93, 126  
 Gupta 36, 255, 459, 462  
 Gurung 9, 10, 29, 291, 315, 336  
 Gǔzhàng 88  
 Gyarong 469–88, 516, 695
- Hǎiyán 86–7, 89, 98  
 Hǎiyáng 92  
 Hakha Lai 13, 409  
 Hakka and Hakkas 6, 49, 50, 72, 76, 79–81,  
 84–6, 89, 94, 103, 105, 113, 146  
 Hallam 13  
 Hándan 86, 107  
 Hángzhou 96, 119  
 Hani 8, 33, 35, 236–7, 239  
 Hānkǒu 105  
 Hānshan 109  
 Hànyáng 105  
 Haoni 8  
 Háoxiàn 109  
 Hayu 15, 505, 509, 512, 514–32, 560  
 Hebei 86, 92, 95, 102, 118, 127  
 Héféi 93, 109  
 Hèfeng 88  
 Hei Lisu *see* Black Lisu  
 Henan 86, 88, 102, 107, 118, 128  
 Héping (Hp) 73  
 Héshan 86  
 Hèshùn 88  
 Héxiàn 109  
 Hill Kachari 11, 177  
 Hill Miri 18, 179, 181, 456–7  
 Himalayas 30, 47, 169, 456, 546, 688, 715  
 Hindus 187  
 Hlota 183  
 Hmar 13, 188

- Hmong-Mien and Hmong-Mienic 7, 44,  
     50, 153  
 Hokkien 6  
 Hong Kong 146, 149–51  
 Hóngtǒng 86  
 honorifics 282, 472  
 Horpa-Shangzhai 14, 490  
 Hpon 177  
 Hrusish 19, 180, 456  
 Hruso 19, 179–80, 456  
 Hua Lisu *see* Flowery Lisu  
 Huái 77  
 Huai River 126  
 Huáinán 109  
 Huáiníng 109  
 Huáiyin 105, 109, 113  
 Huáiyuǎn 109  
 Huáng'an 88, 105  
 Huánggang 105  
 Huángméi 107  
 Huángpí 105  
 Huángyán 85, 88–9, 96, 119  
 Hubei 78, 79, 84, 86, 88, 92–3, 102,  
     105–7, 109  
 Huhehote 95  
 Huizhou 84–6, 90, 93, 95, 107, 109  
 Hunan 76–9, 85–6, 93, 115  
 Huojia 128  
 Huòjia 86, 91, 102  
 Huòlù 118  
 Huòqiu 93–4, 109  
 Huòshan 109  
 Húzhou 85, 96, 119  
 Húzhou Shuanglín 85, 96  
  
 ideographs 157–8  
 Idu 19, 33, 179, 181, 456  
 Idu-Digarú 19, 181  
 Idu Mishmi 19  
 India Missions Association 174  
 Indic 9, 32, 45, 47, 53–4, 170, 692, 693  
 Indo-Aryan 16, 19, 427, 516, 518  
 Inle Lake 195  
 Inner Mongolia 95, 126, 129  
 inscriptions 9, 20, 160, 195–7, 255–6, 653  
*International Encyclopedia of Linguistics* 172  
 Intha 195  
 Irrawaddy River 195, 623  
  
 Jad 9, 47  
 Jahri 16  
 Jaintia 169  
 Jalpaiguri 177  
 Japanese 51, 59, 63, 157, 160, 164, 635–6  
  
 Jiangyong 73, 76  
 Jianchuan 73 *see also* Bai  
 Jianchuan Bai 653–5  
 Jianguá 115  
 Jianguhuai 84, 88, 90, 95–6, 100–2, 105,  
     109, 113, 118  
 Jiang-Huai 126–7, 129–30  
 Jianglè 85  
 Jianglíng 105  
 Jiangu 86, 90, 105, 109  
 Jiangxi 79, 80, 86, 102, 104  
 Jiangyin 87, 95–6  
 Jiangyong 115  
 Jiànníng 80, 86–7  
 Jiàn'ou 81, 85, 87, 88  
 Jiànyáng 81, 85, 88  
 Jiaochéng 93, 107, 109  
*Jīngdiǎn shìwén* 77  
 Jiaodong Peninsula 102  
 Jiaoliao 6, 78  
 Jiashan 109  
 Jiaxìng 86, 119  
 Jiayú 88  
 Jièshou 109  
 Jilu 6  
 Jin 7, 24, 26, 36, 84, 91, 483  
 Jínán 118  
 Jinchéng 109  
 Jinghpaw 11–12, 15, 32–4, 174–8,  
     182, 469, 674  
 Jinghpo 11, 401–8  
 Jinghuai 6  
 Jíngjiāng 85, 96  
 Jínglè 88  
 Jíngmén 92, 105  
 Jingshan 105  
 Jinhua 85, 88, 95–6, 103, 119  
 Jinlu 78  
 Jino 8  
 Jinsha 222  
 Jinuo 8  
 Jínzhài 109  
 Jirel 9, 29  
 Jixi 85–6, 95, 109  
 Jíxiàn 117  
 Jugli 11  
  
 Kabui 12, 183, 186  
 Kachari 11, 175, 177  
 Kachcha 12, 186  
 Kachin 11, 177 *see* Jinghpo  
 Kadu 11, 175, 178  
 Kagate 9, 270  
 Kaipíng 86, 103, 114, 116

- Kaman 19, 181  
 Kamarupan 172, 427  
 Kamphaeng Phet 208  
 Kanashi 16  
 Kanauri 16  
 Kangshan 490  
 Karbi 12, 14, 171, 176, 178, 183–5, 187–8  
 Karbi Anlong 187  
 Karen and Karenic 17–18, 28, 31–2, 43–4,  
     54, 505, 533, 623–4, 632, 638, 641–2, 645  
 Karen State 18, 623, 632  
 Karenni 18, 623, 632  
 Kathmandu 308, 347, 349, 355, 371, 372, 518  
 Kathmandu Newar *see* Nepāl Bhāsā  
 Kathmandu Newari 462  
 Kayah Li 18, 623–31  
 Kazakhstan 53  
 Kengtung 208  
 Khaling 15, 505–6, 508–9, 512, 515–16, 560  
 Kham 14, 17, 30, 36, 516  
 Khams 10, 270  
 Khamti 179, 181–2  
 Khasi 89, 176, 178  
 Khasi Hills 387  
 Khasis 169–70  
 Khatu 8, 236  
 Kheza 12, 183–6  
 Khamngan 11, 175–7, 185  
 Khitan 7  
 Khoa 19, 180  
 Khoirao 12, 183–7  
 Khyang 13  
 Kiao 15, 674  
 Kinauri 14, 16, 30  
 Kinauri-Almora 16  
 Kiranti 4, 5, 14–16, 30–1, 36, 47, 371,  
     505–17, 533, 536, 538–9, 542, 544,  
     546, 549, 560, 562, 684, 695  
 Kiu 15, 674  
 Kiupa 15, 674  
 Kiutze 15, 674  
 Koch 11, 32, 171, 175, 177  
 koinés 76, 82  
 Kokborok 11, 45, 169, 171, 175–7  
 Kom 13, 188  
 Konyak 11, 32, 175–8, 182, 184–5, 188  
 Korean 51, 164–5  
 Kuine Khani 315  
 Kuki 12–14, 31, 44, 47, 170–1, 174,  
     182–5, 187–8  
 Kuki-Chin 12, 31, 47, 188, 409, 427  
 Kuki-Chin-Naga 44, 427  
 Kukish 13, 182, 188  
 Kulung 15, 506, 516  
 Kunming 92, 109  
 Kunshan 89, 95–6, 119  
 Kurukh 16, 516  
 Kwoireng 12, 183, 186, 187  
 Lachi 8, 12, 401  
 Ladakhi 9, 270, 282  
 Lahu 8–9, 27, 33–5, 208–20, 227,  
     237, 249, 558  
 Lahu Shehleh 208  
 Lahul 9, 270  
 Lai 13, 409–26  
 Láí'an 109  
 Laizo 13, 409  
 Lakher 13, 188  
 Lalung 11, 177  
 Lamo 8  
 Lamu 222  
 Lancang 209, 222  
 Lancang Lahu 209  
 Langrong 13  
 Langsu 31, 33  
 Lángxi 109  
 Lánshan 115  
 Lanyin 6, 78  
 Lánzhou 85, 102, 107  
 Lao and Laos 152, 208–9, 222, 236–7  
 Late Zhou Chinese (LZC) 59–70  
 Lavrung 14, 490  
 Lepcha 29, 186, 705–15  
 Leqi 33  
 Levai 19, 456 *see also* Bangru  
 Lhasa 273, 275  
 Lhasa Tibetan 255, 262, 270–85  
 Lhomi 9  
 Lhota 12, 183, 185  
 Liánchéng 103  
 Lianghe 401  
 Liangmai 12, 183–7  
 Liánnán 86  
 Liling 86  
 Limbu 15, 505–16, 518, 546, 560  
 Línfén 77, 86, 88, 107, 109  
 Lingbi 109  
 Língchuan 86, 118  
*Linguistic Encyclopedia* 516  
*Linguistic Survey of India* 176, 182–4, 291  
 Linguistic Survey of Nepal Project 505  
 Língxià 86  
 Línquán 109  
 Línwǔ 86, 109, 115  
 Línxia 100, 118  
 Línxiàn 88  
 Linyi 86



- Liphò 8, 222–3, 229–30  
 Lisaw *see* yellow Lo  
 Lǐshan 105  
 Lishpa 19, 180  
 Lisu 8–9, 20, 27, 31, 222–35, 651  
 Liǔlǐn 88  
 liushu 161–2, 165  
 Liyáng 85, 88, 96  
 loanwords 401, 532, 546, 624, 699  
   Akha 237  
   Bai 20, 31  
   Belhare 559  
   Burmese 196–200  
   Cantonese 148–50  
   Chantyal 317–18  
   Garo 399–400  
   Gyarong 469  
   Karen 17–18  
   Lahu 209  
   Qiang 575  
   Tangut 602, 604  
   Tani 18–19, 456  
   Thai 9  
 Lolo and Loloish 5, 8–9, 12, 14, 18, 24,  
   31, 33, 54, 219, 652, 671  
 Lolo-Burmese 5, 8–9, 12, 14, 18, 24, 31, 33,  
   177, 208, 214, 236, 483, 602, 623  
 Lolophò 8, 222  
 Longchang 11  
 Longchuan 33, 401  
 Longchuan Achang 33  
 Lóngdié 109  
 Lóngnán 104  
 Lóngyán 81  
 Lotha 12, 183–4  
 Lu Fayan 163  
 Lùchuan 86  
 Lùfeng 89  
 Luhupa 183  
 Luish 11, 32, 175, 178  
 Lújiang 109  
 Lungcang 11  
 Lungchang 11  
 Luótián 88, 105  
 Luòyáng 102, 107  
 Lushai 13, 171, 188, 483  
 Luxi 88, 401  
 Lyngngam 169  
  
 Mǎ`anshan 109  
 Máchéng 105  
 Mae Hong Son 208, 623  
 Magar 14, 16–17, 516, 518, 683–4  
 Mālāmpāhār 315  
  
 Malkābāng 315  
 Manang 9, 10, 336  
 Manangba 10, 315, 336  
 Manangpa 291  
 Manchatì 16  
 Mǎnchéng 95  
 Manchu 6–7, 89, 127  
 Manchuria 7, 102  
 Manchu-Tungusic 7  
 Mandarin 6, 48–52, 59, 62, 72, 76–81, 84,  
   87, 93–5, 100, 102, 105, 107, 109,  
   115–16, 126–30, 138–43, 146,  
   151–3, 157, 160, 163–4, 654  
   Northern 126, 129–30  
   Southwestern 126, 129–30  
   Standard (SM) 85, 87, 91–2, 96–123  
 Mangale Khani 315  
 Manipur 169–72, 176, 178, 184, 186–8, 427  
 Manipuri *see* Meithei  
 Manipuris 14, 27, 187, 427  
 Mao 12, 183–6  
 Mǎopíng 86, 92  
 Mara 12, 13, 188, 456  
 Maram 12, 13, 186–7  
 Marchcha 16  
 Maring 13, 183–5, 187  
 Marma 195  
 Marpha and Marphali 10, 291  
 Maru 8, 12, 47, 177, 401  
 Mǎyáng 85  
 Mazaudon 10, 291  
 Mech 11, 175, 177  
 Meghalaya 169–71, 176–7, 387  
 Meitei *see* Meithei  
 Meiteilong *see* Meithei  
 Meithei 12, 14, 27, 171, 183–5, 187–8, 427–37  
 Méixiàn 73, 76, 86, 93, 94, 99, 105, 113  
 Mekhong 208, 222  
 Meluri 12, 183, 186  
 Mema 179  
 Menglian County 209  
 Méngshan Xihé 86  
 Miáoli 116  
 Micha 8, 222  
 Middle Chinese 163  
 Mien 208  
 Miji 19, 180  
 Miju 19, 179, 181  
 Mikir 14, 171, 183, 187, 427  
 Mikir Hills 187  
 Milang 18, 179, 181, 456, 457  
 Min 6, 72–3, 79, 81, 84–90, 92–3, 95, 99,  
   100, 102–4, 107, 111, 113, 115–18,  
   146, 153, 160, 164, 653

- Min Chia (Bia) 653  
 Míngxi 85  
 Mǐnjīā 73  
 Miri 18, 31, 171, 180  
 Mirish 18, 31, 180  
 Mishing 31  
 Mising 18, 171, 179–81, 456, 459, 463–4  
 Misingish 180  
 Mizo 12–14, 36, 171, 174, 182–5,  
     187–8, 409  
 Mizo-Kuki-Chin 12–13, 174,  
     182–4, 187–8  
 Mizoram 169–71, 176, 188, 409, 427  
 Mizos 171, 188  
 Modern Lhasa 9  
 Modhupur 387  
 Mon 9, 18, 20, 31, 196–7, 632, 654  
 Mongol 6, 7, 653  
 Mongolia 7, 51, 53, 89, 95  
 Mongsen 183–4  
 Mon-Khmer 6–7, 9, 18, 44–5, 50,  
     52, 54, 169, 623  
 Monpa 47, 179, 439  
     Northern 10, 179, 439  
 Mopwa 18, 623  
 Moshang 11  
 Motuo Menba 10, 179  
 Motuo Monba 10, 439  
 Moulmein 196  
 Móupíng 104, 111  
 Mpi 8, 236  
 Mru 427  
 Munda 16, 169, 516  
     North 16  
 Murmi 291  
 Mustang 291  
 Muya 17, 30  
 Myagdi District 315  
 Myanmar *see* Burma  
 Mzieme 12, 183–6
- Na 18, 46, 61–2, 65, 67, 70, 208,  
     392, 445, 450, 456  
 Na Bengni 18, 456, 457  
 Nachereng 15, 506, 516  
 ‘Naga’ 12, 14, 16, 32, 172, 182–3,  
     186–8, 427, 516  
 Naga-bodo 183, 187  
 Naga-kuki 183, 187  
 Nagaland 169, 170–2, 176–7, 179,  
     182, 184, 186, 188, 427  
 Nagaland Bhasha Parishad (Nagaland  
     Language Agency) 172  
 Nagas 172, 182, 187–8
- Nam 208, 255  
 Nam Tha 208  
 Namsangia 11, 177  
 Namuyi 30  
 Namuzi 17  
 Nánchāng 73, 76, 81, 86, 99  
 Nánchéng 86–7  
 Nánhāi 88  
 Nánhui Zhoupǔ 88, 96  
 Nanjing 127, 129, 130  
 Nánlíng 109  
 Nánníng 109  
 Nánpíng Xiáyáng 85  
 Nántong 109  
 Nar-Phu 10, 291, 315, 336–51  
 Nasu 8, 31, 227, 653  
 Naxi 19–20, 33, 651  
 Nepal 29–30, 52, 291, 294, 315, 336,  
     355–6, 371, 505, 516, 518,  
     533, 546, 566, 591, 683, 705  
 Nepāl Bhāsā 371–83  
 Nepali 10, 16, 302–3, 307–11, 315–23,  
     325–8, 336–40, 356, 358, 360, 371–2,  
     505, 507–8, 514, 518, 524, 532–4,  
     536, 541–3, 545–6, 549–50, 558,  
     563, 568, 683, 690, 699–700, 705  
 New Xiāng 80  
 Newar and Newari 27, 29, 35, 47, 278,  
     355–6, 371–3, 375–6  
 Níngbo 85, 88, 96, 119  
 Níngguó 109  
 Níngguà 86  
 Ninglang Yi 588  
 Ningxia 100  
 Níngxiang 88  
 Níngyuǎn 115  
 Nishi 10, 18, 31, 36, 47, 179–81, 270,  
     456–7, 462  
 Nishing 18, 181, 462  
 Nisu 8  
 Nitu 13  
 Niuwozi Prinmi 17, 588–92  
 Nocte 11, 32, 175–7, 179  
 Northeastern India 169–89, 387,  
     388, 401  
 Northern Lo *see* Black Lisu  
 Nosu 8, 31, 234, 653  
 Nruanghmei 12, 183–7  
 Ntenyi 12, 183, 186  
 Nujiang Lisu 222  
 Nungish 15, 32  
 Nusu 8, 33  
 Nyamkad 9, 47  
 Nyisu 18, 181

- Old Burmese 59, 196–7  
 Old Chinese (OC) 20, 22–8, 59, 96,  
 163, 516, 652  
 Old Kuki 13  
 Old Tibetan 59, 255–7  
 Old Xiāng 80  
 oracle bone inscriptions 160–1
- Padam 18, 36, 179, 181, 457, 460–5  
 Padaung 18, 623  
 Pagan 195–6  
 Paite 13, 188  
 Pakistan 10, 53, 270  
 Pakistan, East 189  
 Palaung 89  
 Palayachi 18, 623  
 Pali 9, 196–9, 211  
 Pani Koch 11, 177  
 Pa-O 17, 623, 632  
 Patan 371  
 Pātle Kharkā 315  
 Pattani 16  
 Pawi 13, 188  
 Phom 11, 173, 175–7, 185  
 Phunoi 8  
 Pijo 8  
 Píngdù 86, 104, 111  
 Píngjiang 86  
 Píngjiang Chángshòu 86–7  
 Pingyao 129  
 Píngyì 118  
 pinyin 223–5, 236  
 Plains Miri 18, 171, 180  
 Pochuri 12, 183–6  
 Prinmi 17, 588–601  
 Pūchéng 85  
 Puiroi 12, 183–6  
 Pumi (Prinmi) 17, 30, 33, 54, 588  
 Pǔqí 86, 93  
 Purik 9, 270  
 Puroit 180  
 Pútián 87  
 Putonghua 6, 72, 126–9, 150, 154  
 Pwo Karen 18, 632–47  
 Pyen 8  
 Pyu 9, 20, 196
- Qiang 17, 23, 27, 30, 33, 35, 54,  
 469, 483, 573–86  
 Qiāngcì 6, 17–18, 30, 33, 498, 573,  
 588, 602  
 Qiánshan 109  
 Qíchun 88  
 Qieyun 72, 77, 163
- Qímen 109  
 Qímén 85–6, 95  
 Qinghai 10, 87, 95, 102–3, 109, 270  
 Qíngjiàn 107  
 Qíngtián 90  
 Qin-Han times 115  
 Qínxiàn 117  
 Qiū 15, 674  
 Qiuzi 15, 674  
 Qíyáng 85  
 Quánjiao 109  
 Quánzhou 81, 90, 99, 103, 113  
 Queyu 17, 30, 33  
 Qúzhou 85, 88, 96, 99, 119
- Rabha and Rabhas 11, 170–1, 175–7  
 radicals of Chinese characters 159  
 Rai 505, 508, 516, 518, 546  
 Ralte 13, 188  
 Rangkas 16  
 Rangloi 16  
 Rawang 14–15, 24, 28, 30–3, 35, 36,  
 469, 483, 674, 675  
 Red Lahu 208 *see also* Lahu  
 Rengma 12, 178, 183–6  
 rGyalrong 14–16, 17, 26, 30, 33, 36, 467, 594  
 Riau Indonesian 60, 64  
 Ribu 490  
 Risiangku Tamang 293–4, 302  
 Rong 186, 613  
 Róngchéng 104  
 Rongmei 12, 186  
 Rǔchéng 86  
 Ruga 11, 175–7, 387  
 Rúgāo 109  
 Ruili 401  
 Rung 5, 14, 17–18, 30–3
- Sak 11, 175, 178, 394  
 Salween 208, 222, 623  
 Salween River 208  
 Sangtam 12, 183–6  
 Sani 8  
 Sanmíng Sanyuán 85  
 Sanshǔi 88  
 Sanskrit 196, 211, 263, 267, 356, 427  
 Sema 12, 183, 186  
 Sengmai 12, 175–6, 178, 185  
 Sgaw 18, 32, 623, 632  
 Shaanxi 86, 91, 102–3, 107, 118, 126, 128–9  
 Shan 179, 181–2, 195, 208, 211, 215, 222,  
 624, 627, 630  
 Shandong 84–6, 88, 92, 102–4, 108, 111, 118  
 Shang dynasty 59, 160

- Shanghai and Shanghaiese 6, 48, 85–8, 94–6,  
106, 112, 118–19, 121, 123, 131–45
- Shanghai Fèngxián 86
- Shangxiàn 86, 91–2
- Shàntóu 86, 87, 90, 93, 104
- Shanxi 77–8, 84–6, 88, 91, 93, 102–3,  
107, 109, 116–18, 128–9
- Shānxiàn 88
- Shanyin 86
- Shàowǔ 81, 86–7
- Shàoxing 96, 103, 119
- Shar chopkha 10, 179
- Shaxiàn 85, 88
- Shèngxiàn 89, 95, 103, 119
- Shèngxiàn Chánglè 89
- Shèngxiàn Chónggrén 88, 96, 119
- Shèngxiàn Tàipíng 88, 96
- Sherdukpen 19, 173, 179, 180
- Sherdukpen-Bugun-Sulung-Lishpa 19
- Sherpa 9, 29, 270
- Shèxiàn 85, 86, 94, 109
- Shèxiàn Túnxi 85, 95
- Shigatse 9, 270
- Shílóu 88
- Shixing 17, 30
- Shō 13
- Shòuguang 88
- Shòuxiàn 109
- Shòuyáng 88
- Shuangfeng 76, 80–1, 99
- Shuanglín 95
- Shuchéng 109
- Shùnchang 85–7
- Shùndé 88
- Shuowen jiezi 161
- Shùyáng 96, 113
- Siamese 89
- Sichuan 10, 30, 31, 77–8, 93, 102, 126,  
222, 270, 469, 490, 573, 588
- Sidaba 14, 490
- Sikkim and Sikkimese 10, 29, 186,  
270, 505, 518, 705
- Simi 12, 183–6
- Simte 13
- Singpho 11, 32, 175–6, 178–9, 182
- Situ 14, 490, 499
- Siyi 86, 88, 90, 93–4, 100–1, 103
- Siyin 13
- Songjiang 88, 96
- Songxi 85, 88
- Sopvoma 12, 183, 186
- Spiti 9, 270
- Srong 255
- Standard Chinese 164
- Standard Mandarin (SM) *see* Mandarin,  
Standard
- Suínfng 85
- Suixi 109
- Súxiàn 105
- Sulung 19, 179–80, 456
- Sunwar 15, 505–6, 516
- Sùqian 86
- Sùsong 109
- Sùxiàn 109
- Suzhou 72–3, 76, 78–9, 81, 86, 88–9,  
95–6, 99, 101, 104, 106, 113–14,  
119, 122
- Syang 10, 291
- Tableng 11, 177
- Tagin 18, 179, 181, 456
- Tai 6–7, 18, 31, 80, 89, 146, 150,  
152–3, 623, 624
- Taigu 77, 118
- Tàihé 86, 109
- Tàihú 109
- Tai-Kadai 6–7, 18, 44–5, 50, 52, 54
- Tai-Lahu-Lawa 209
- Táishan 86
- Taiwan and Taiwanese 104, 115, 157, 160  
*see also* Min
- Taiwan Min 104
- Tàixing 90, 118
- Tàiyuán 76, 88
- Tak 208
- Takpa 9–10, 179
- Tamang and Tamangic 5, 9–11, 29, 33–4, 47,  
291–315, 317, 328, 335–6, 339
- Tamang-Gurung-Thakali (TGT) 255
- Tamang-Gurung-Thakali-Manang  
(TGTm) 291, 293
- Tamlu 11, 177
- Tangbe 10, 291
- Tangkhol 12–14, 183–5, 187–8
- Tangsa 11, 175–9
- Tangut (Xixia) 6, 17, 30, 32, 165, 602–19
- Tani 4–5, 18–19, 31, 35–6, 45, 178, 180–1,  
456–65
- Táoyuán 86
- Taraon 19, 181, 456
- Taron (Dulong) 15, 674
- Tavoy 195, 633
- Tavoyan 8, 195
- Tenasserim 632
- Tengsa 183–4, 186
- Thado 13, 188
- Thai 6–7, 9, 152, 196, 208, 211, 222, 236–7,  
624, 632, 637, 641–2, 646

- Thailand 9, 196, 208–9, 211, 222, 236–7,  
     623, 632, 637, 646  
 Thakali 9–10, 291, 315, 336  
 Thara Khani 315  
 Thini 10, 291  
 Thon-mi Sambhota 255  
 Thukumi 183  
 Thulung 15, 505–16  
 Tianmén 105  
 Tibet 29, 32, 169–70, 179, 262, 281,  
     285, 456  
 Tibetan 9–11, 14–15, 23, 25–6, 35–6, 46–7,  
     54, 179, 255, 259, 261, 270, 276–7, 291,  
     300, 336, 439, 588, 602  
 Tiddim Chin 13, 44  
 Tinan 16  
 Tintinkiya Koch 11, 177  
 Tirap 222  
 Tiwa 11, 175–7  
 Tod 9  
 Tongchéng 105, 109  
 Tónggǔ Sandu 86  
 Tongwèi 109  
 Tongyuán 89  
 Tripura 169–71, 176–7, 387  
 Tripuri 11, 171, 177  
 Tʼrung 14–15, 22, 24, 26, 30  
 Tsangla 33  
 Tshangla 9–11, 179–80, 255, 439–54  
 Tǔ 76  
 Tujia 18, 20, 35, 623  
 Tukumi 12, 184  
 Tungus and Tungusic 6–7, 51, 53  
 Túnxi 86, 109  
 Turkic 51  
 Turkmenistan 53  
  
 Umbule 15, 506, 516  
 U-Tsang Modern Lhasa 9, 270  
 Uzbekistan 53  
  
 Vaipe 188  
 Vietnam 208, 236  
 Vietnamese 5, 32, 89, 164–5, 641  
  
 Wakching 11  
 Waʼnang Koch 11, 177  
 Wancho 11, 175–7, 179  
 Wang Renpo 223  
 Wàngjiang  
     passive construction 109  
 Wànróng 88  
 War 169  
 Wéifang 86, 88, 111, 118  
  
 Wèinán 107  
 Weixi 222–3, 651  
 Wéixiàn 118  
     verbal complements 116  
 Wèixiàn 86, 92  
 Wen Qi Ji 127  
 Wengyuán 86  
 Wenling 79  
 Wenxiàn 102  
 Wenzhou 27, 78–9, 85–7, 90, 93, 95–6,  
     101, 119  
 West Himalayan 14, 16, 30, 36, 47, 255  
 Western Archaic 9, 270  
 Western Zhou 160  
 White Barbarians 31  
 White Lahu 208 *see also* Lahu  
 Woyáng 109  
 writing 156–65  
 writing systems 156, 158, 162, 165, 196,  
     198, 236, 255–7, 401, 632  
 Written Tibetan (WT) 14, 22–6, 278,  
     338, 340, 469, 470, 484, 486, 488  
 Wu 6, 31, 72, 76–80, 84–103, 106–7, 109,  
     113, 116–19, 122, 131, 146  
 Wuʼan 86  
 Wugang 85  
 Wu Mán 31  
 Wuhan 105  
 Wúhú 105  
 Wújiang 96, 119  
 Wújiang Lǐlì 96  
 Wújiang Shèngzé 96  
 Wúwéi 109  
 Wúxi 78, 85, 89, 95–6  
 Wǔyì 103  
 Wǔyuán 85, 95  
  
 Xiàmén 81, 87, 99, 103, 111–13  
 Xiʼan 86  
 Xiandao Achang 33  
 Xiang 6, 72, 76–80, 84–6, 88, 90, 92–5,  
     98–100, 106–7, 109, 114, 116,  
     119, 146  
 Xianggǎng Shènzhèn 86  
 Xiangxiang 88, 93, 106, 111  
 Xiangyuán 117  
 Xiaogan 105  
 Xinfeng 89  
 Xinhua 85  
 Xinhui 86  
 Xíníng 109  
 Xinjiang 126  
 Xintián 115  
 Xinyi 26

- Xinyú 86  
 Xinyuè 105  
 Xinzhou 88  
 Xishūi 88, 93, 109  
 Xiuníng 85–6, 95, 109  
 Xiushui 86  
 Xixia 165, 602  
 Xiyáng 88  
 Xu Shen 161–2  
 Xúnhuà 87  
 Xupu 85  
  
 Yacham 12, 183–4  
 Yachumi 12, 183–4  
 Yakkha 16, 505–6, 515  
 Yamphu 16, 505–6, 509, 512–13, 515  
 Yánchuan 118  
 Yángchéng 88  
 Yángjiang 86, 88, 100  
 Yangon 8  
 Yángqu 88  
 Yangtze 7, 77–8, 126–7, 145, 222  
 Yángxin 86  
 Yángzhou 76, 109  
 Yano 181  
 Yantái 108  
 Yaw 195  
 Yellow Lahu 208 *see also* Lahu  
 Yellow Lo 222 *see also* Lisu  
 Yellow River 3, 127  
 Yi 20, 31, 651–4  
 Yídu 105  
 Yífeng 86–7  
 Yiish 195  
 Yílan 104  
 Yimchungrü 12, 183–4, 185–6  
 Yin Ru tone 146  
 Yíngchéng 105  
 Yíngjiang 401  
 Yingshan 88–9, 105–6, 109  
 Yíntale 18, 623  
 Yípho 236  
 Yixiàn 85, 86, 95  
 Yíxing 85, 96  
 Yìyáng 86, 93  
 Yízhāng 115  
 Yogli 11  
  
 Yǒngān 85, 88, 96  
 Yǒngkāng 85–6, 92, 95  
 Yuan dynasty 7  
 Yuánlíng 88  
 Yuannan 222  
 Yuánpíng 88  
 Yue 6, 29, 72, 76, 79–81, 84–6, 88, 90,  
     92–4, 96–7, 99–101, 103, 107,  
     110–20, 146, 153  
 Yuèxi 109  
 Yúgan 86  
 Yùchéng 86, 102  
 Yúnmèng 105  
 Yunnan 10, 30, 31, 77–8, 92–3, 107, 126,  
     177, 195, 208–9, 211, 222–3, 270, 401,  
     588, 592, 651–3, 674  
 Yúshè 88  
 Yúxiàn 88  
 Yúyáo 85, 96, 119  
  
 Zahao 13, 409  
 Zaiwa 12, 31, 33, 195, 401  
 Zaiwa (Atsi) 8  
 Zaozhuang 102  
 Zauzou 8  
 Zeme 12–13, 183–8  
 Zengchéng 86  
 Zhaba 17, 30  
 Zhang Wei 127  
 Zhang-zhung 255  
 Zhejiang 27, 86, 98  
 Zhènghé 85  
 Zhenqian 73  
 Zhongníng 100  
 Zhongshan 88  
 Zhongxiáng 88  
 Zhongyáng 88  
 Zhongyuan 6  
 Zhongyuan Yayin 127  
 Zhongyuan Yin 127  
 Zhongzhou Yin 127  
 Zhou dynasty 59  
 Zhuchéng 111  
 Zhuji 85, 88, 96, 103, 119  
 Zibó 86, 88, 111, 118  
 Zígui 89  
 Zo 188