

Mouton Grammar Library

A Grammar of Mongsen Ao



Mouton Grammar Library 39

Editors

Georg Bossong

Bernard Comrie

Matthew Dryer

Mouton de Gruyter
Berlin · New York

A Grammar of Mongsen Ao

by

A. R. Coupe

Mouton de Gruyter
Berlin · New York

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter GmbH & Co. KG, Berlin.

⊗ Printed on acid-free paper which falls within the guidelines of the
ANSI to ensure permanence and durability.

Library of Congress Cataloging-in-Publication Data

<p>Coupe, A. R. (Alexander Robertson) A grammar of Mongsen Ao / by Alec Coupe. p. cm. – (Mouton grammar library ; 39) Includes bibliographical references and index. ISBN-13: 978-3-11-019088-5 (cloth : alk. paper) 1. Ao language – Grammar. I. Title. PL4001.A691C68 2007 495'.4–dc22</p> <p style="text-align: right;">2007024522</p>

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

ISBN 978-3-11-019088-5

ISSN 0933-7636

© Copyright 2007 by Walter de Gruyter GmbH & Co. KG, D-10785 Berlin.
All rights reserved, including those of translation into foreign languages. No part of this book may
be reproduced or transmitted in any form or by any means, electronic or mechanical, including
photocopy, recording or any information storage and retrieval system, without permission in writing
from the publisher.
Printed in Germany.

Acknowledgements

This work was originally submitted at La Trobe University in December 2003 as a doctoral dissertation. Since then it has been through countless revisions and expansions as further insights into the structure of the grammar have been gained. Throughout this process I have enjoyed the support of a number of key people. David Bradley oversaw my doctoral research and provided critical comments on numerous draft chapters. The final draft has been considerably strengthened by incorporating the comments and suggestions of my PhD examiners Robbins Burling, Bill Foley and Martine Mazaudon. Bernard Comrie went to the trouble of reading the entire manuscript and offered pages of detailed criticisms and suggestions for improvements. Randy LaPolla and Carol Genetti shared inspirational suggestions on many aspects of description and analysis. Other people with whom I have had the benefit of fruitful discussions include Sasha Aikhenvald, Andy Butcher, Hilary Chappell, Shobhana Chelliah, Bob Dixon, Antoine Guillaume, Bernd Heine, Jane Hill, Ken Hill, Lars Johanson, Nicole Kruspe, Scott DeLancey, Bill McGregor, Stephen Matthews, Maarten Mous, Willem de Reuse, Fritz Serzisko, Janet Sharp and Tonya Stebbins. It goes without saying that I alone am fully responsible for any remaining errors or misinterpretations.

None of this work would have been possible without the generous support of the Ao community. I owe a profound debt of gratitude to the Mongsen speakers of Nagaland, and in particular, to the very obliging family of my sister-in-law Kashila Gurung. Kashila, her sister Arenla Subong, and many of their extended family went to great trouble to help me understand the intricacies of Mongsen grammar, and their uncle Imdangchiba Imchen was an enthusiastic teacher of his language. Thanks to them, I have been able to write this book. I was also touched by the overwhelming hospitality shown by the people of Mangmetong village, who made my visits so fruitful. I am equally indebted to other members of my extended family in India and Australia for all they have done to ensure that I could complete this research. In particular, I would like to acknowledge everything that my sister-in-law Sumitra Ao did for me in negotiating bureaucratic hurdles and procuring permits for me to enter Nagaland. Thanks also to my wife's family in Shillong and Nagaland for supporting and taking care of me on my numerous visits.

La Trobe University funded the doctoral research with an Australian Postgraduate Award and provided additional grants to cover fieldwork expenses. Subsequent research has been made possible by a three-year Australian Research Council postdoctoral fellowship (DP0450943: A typology

of adverbial subordination and clause linkage in Tibeto-Burman languages). I am deeply grateful to both of these institutions for their support. I thank Pacific Linguistics for granting permission to publish a map that originally appeared in Coupe (2003a). Lastly, Jo Taylor did a brilliant job proof-reading and helping to format the manuscript.

Behind this whole project has been my loyal, loving partner Pavitra, the catalyst for my obsession with the languages of Nagaland. This one is for you.

Table of contents

Acknowledgements	v
List of maps, tables and figures	xvi
Abbreviations and conventions	xviii
1 Introduction	
1.1. The Ao language	1
1.2. Organization of this grammar	1
1.3. Typological overview	2
1.4. Linguistic situation in Nagaland	4
1.5. Genetic classification	7
1.6. Geographical setting	7
1.7. Ethnographic overview	9
1.7.1. Autonym, exonyms and allonyms.....	9
1.7.2. History.....	11
1.7.3. Village life.....	13
1.7.4. Religion.....	14
1.8. Ao dialects.....	16
1.9. Review of literature.....	17
1.10. Ao literature	18
1.11. Fieldwork and sources of data.....	18
1.12. Equipment	20
1.13. Orthography	20
1.14. Theoretical orientation	21
2 Phonology and phonological processes	
2.1. Syllable structure.....	23
2.1.1. Vowel Sequences	25
2.1.2. Absence of phonemic stress	26
2.1.3. Tone	26
2.2. Segmental phonology.....	27
2.2.1. Consonant phonemes	27
2.2.1.1. Voice onset time contrasts	28
2.2.1.2. Stops.....	30
2.2.1.3. Affricates.....	31
2.2.1.4. Fricatives.....	31
2.2.1.5. Nasals	32
2.2.1.6. Approximants.....	38
2.2.1.7. Limited free variation.....	44

2.2.2.	Vowel phonemes	45
2.3.	Segmental differences between Mongsen varieties	47
2.3.1.	Consonants	47
2.3.2.	Vowels	48
2.4.	Word	49
2.4.1.	Phonological criteria	49
2.4.2.	Grammatical criteria	51
2.5.	Clitics	51
2.6.	Phonological processes	52
2.6.1.	Schwa deletion in prefixes	52
2.6.2.	Sporadic vowel harmony	53
2.6.3.	Vowel dissimilation	53
2.6.4.	Glide assimilation	54
2.6.5.	Consonant assimilation	54
2.6.6.	Resyllabification	55
2.6.7.	Other vowel deletions	55
3 Prosody		
3.1.	Tonemes	57
3.1.1.	Lexical tones	57
3.1.2.	Methodology	59
3.1.3.	Grammatical tones	60
3.1.3.1.	Adjectives of degree.....	60
3.1.3.2.	Agentive case marked by a rising tone	64
3.2.	Tonally unspecified prefixes	65
3.3.	Tone sandhi	67
3.3.1.	Tonal combinations in phonological words	68
3.3.2.	Tone sandhi generated by morpheme concatenation	68
3.3.3.	Tone sandhi generated by tones of adjacent words.....	71
3.3.4.	Tone sandhi in nominal compounds	72
3.4.	Postlexical intonation.....	73
3.4.1.	Phrasal intonation.....	74
3.4.2.	Boundary marking intonation on dependent clauses.....	75
3.4.3.	Non-final postlexical intonation on matrix clauses.....	76
3.5.	The glottal stop prosody.....	77
4 Word classes		
4.1.	Open word classes.....	79
4.1.1.	Verbs	79
4.1.2.	Overview of the noun word class.....	81
4.1.3.	Relational bound nouns.....	84
4.1.4.	Non-relational bound nouns.....	86
4.1.5.	Unbound nouns	87

4.2.	Closed word classes	88
4.2.1.	Personal pronouns	89
4.2.2.	Possessive pronouns	94
4.2.3.	Demonstrative pronouns	97
4.2.4.	Interrogative pronouns	98
4.2.5.	Indefinite pronouns	100
4.2.5.1.	Negative indefinite pronouns	101
4.2.5.2.	Specific indefinite pronouns	102
4.2.5.3.	Non-specific indefinite pronouns	103
4.2.5.4.	Reduplicated indefinite pronouns and adverbals	104
4.2.6.	Generic pronoun	105
4.2.7.	Nominal modifiers	106
4.2.7.1.	Nominal deictics	106
4.2.7.2.	Lexical quantifiers	111
4.2.7.3.	Cardinal numerals	112
4.2.7.4.	Decimalization of the overcounting numeral system	117
4.2.7.5.	Case marking clitics	121
4.2.8.	Time words	122
4.2.9.	Adverbs	124
4.2.10.	Discourse connectives	124
4.2.11.	Phrasal conjunction	127
4.2.12.	Interjections	129
4.2.13.	Exclamations and onomatopoeia	129
4.2.14.	Particles and clitics	131
4.2.14.1.	Quotative particle	131
4.2.14.2.	Reportative particle	132
4.2.14.3.	Topic particle	133
4.2.14.4.	Discourse focus particles	137
4.2.14.5.	Illocutionary force particles	138
4.2.14.6.	Optional declarative mood clitic	142

5 Clause structure and grammatical functions

5.1.	Simple clause structure	145
5.1.1.	Declarative clauses	147
5.1.2.	Interrogative clauses	148
5.1.2.1.	Disjunctive questions	148
5.1.2.2.	Polar questions	148
5.1.2.3.	Content questions	150
5.1.3.	Imperative mood categories	150
5.2.	Grammatical functions	150
5.2.1.	Constituent order	151
5.2.2.	Core grammatical marking	154

5.2.3.	Absence of pivot	165
5.2.4.	Absence of passive voice	168
5.2.5.	Summary	169
5.3.	Case marking morphology	169
5.3.1.	Core case marking	173
5.3.1.1.	Agentive	173
5.3.1.2.	Formally unmarked core arguments.....	174
5.3.2.	Oblique case marking.....	175
5.3.2.1.	Allative.....	176
5.3.2.2.	Ablative.....	177
5.3.2.3.	Instrumental	178
5.3.2.4.	Dative	180
5.3.2.5.	Benefactive.....	181
5.3.2.6.	Locative.....	182
5.3.2.7.	Comitative.....	183
5.3.3.	Nascent postpositions.....	184
5.4.	Valency modifying and reciprocal/collective derivations.....	189
5.4.1.	Morphological causative	190
5.4.1.1.	Causative derivation of intransitive verb stems	190
5.4.1.2.	Causative derivation of transitive verb stems	194
5.4.2.	Analytic causative	197
5.4.3.	Reciprocal/collective derivation.....	198

6 The noun phrase, relative clauses and nominalizations

6.1.	Basic noun phrase structure	201
6.1.1.	Particles occurring within the noun phrase	202
6.1.2.	Noun phrase constituents with a variable position.....	202
6.1.3.	Overview of other noun phrase constituents.....	203
6.2.	Functions of noun phrases.....	204
6.3.	Head	205
6.3.1.	Simple head.....	205
6.3.2.	Compound head	206
6.3.3.	Complex head	207
6.4.	Noun phrase modifiers	208
6.4.1.	Attributes.....	208
6.4.2.	Quantifiers.....	211
6.4.3.	Determiners.....	212
6.4.4.	Case marking clitics	213
6.4.5.	Topic particle and discourse focus particles	214
6.5.	Recursively embedded noun phrases	216
6.6.	Relativization and nominalization.....	217
6.6.1.	Headed relative clauses	219

6.6.2.	Headless relative clauses.....	227
6.6.3.	Agentive nominalizations as relative clauses.....	229
6.6.4.	Locative nominalizations as relative clauses	230
6.6.5.	Aspect and mood marking in relative clauses.....	231
6.6.6.	Relative-correlative constructions.....	234
6.6.7.	Nominalization.....	236

7 Nominal morphology

7.1.	Compound nouns	241
7.1.1.	Noun-noun compounds	243
7.1.2.	Noun-verb compounds.....	244
7.1.3.	Nominalized verb-noun compounds	245
7.1.4.	Noun-nominalized verb compounds	245
7.2.	Nominal prefixes.....	246
7.2.1.	Vocative prefix.....	246
7.2.2.	Relational prefix.....	247
7.2.3.	Non-relational prefix	248
7.2.4.	Nominalizing prefix	248
7.3.	Possession	248
7.3.1.	Possession encoded by pronouns	249
7.3.2.	Possession encoded by juxtaposed nouns	251
7.3.3.	Possession encoded by the agentive nominalizer.....	252
7.4.	Nominalization.....	255
7.4.1.	Nominal derivation by prefixation.....	256
7.4.2.	General nominalizer	259
7.4.3.	Nominalizations as adjectives of degree	260
7.4.3.1.	Comparative	261
7.4.3.2.	Superlative	262
7.4.4.	Agentive nominalizer	263
7.4.5.	Locative nominalizer.....	265
7.4.6.	Instrument nominalization	265
7.4.7.	Purposive nominalization.....	265
7.4.8.	Temporal nominalization	267
7.4.9.	Deictic nominalizations.....	268
7.4.10.	The MAN nominalizer	268
7.5.	Other nominal modifiers	269
7.5.1.	Intensifier suffix.....	270
7.5.2.	Semantic gender markers	270
7.5.3.	Diminutive suffix	272
7.5.4.	Terminative suffix.....	273
7.5.5.	Reflexive suffix.....	273
7.5.6.	Locational suffix	274

7.6.	Derivational morphology of numerals	275
7.6.1.	Derivation of distributive numerals	275
7.6.2.	Derivation of ordinal numerals	277
7.6.3.	Derivation of numeral adverbs.....	277

8 Verbs and verbal morphology

8.1.	Verb classes.....	279
8.1.1.	Intransitive verbs	279
8.1.1.1.	Stative intransitive verbs	280
8.1.1.2.	Directional motion verbs.....	280
8.1.2.	Transitive verbs.....	282
8.1.3.	Ambitransitive verbs	285
8.1.4.	Reduplicated verb roots	287
8.2.	Structure of the predicate	287
8.3.	Verbal prefixes	290
8.3.1.	Prohibitive mood.....	290
8.3.2.	Admonitive mood	291
8.3.3.	Negative prefix.....	291
8.3.4.	Nominalizing prefix	292
8.4.	Verb root	294
8.4.1.	Bare roots	294
8.4.2.	Compound verbs	294
8.4.3.	Pseudo-compounds	295
8.4.4.	Compounding versus stem modification.....	297
8.5.	Verbal suffixes	297
8.5.1.	Lexical suffix zone.....	298
8.5.1.1.	<i>-kət</i> (ASCEND).....	301
8.5.1.2.	<i>-lak</i> (DESCEND).....	302
8.5.1.3.	<i>-uk</i> (INTO).....	303
8.5.1.4.	<i>-tən</i> (TO.AND.FRO)	304
8.5.1.5.	<i>-saʔ</i> (SEPARATE)	305
8.5.1.6.	<i>-sa</i> (SPREAD).....	306
8.5.1.7.	<i>-thəŋ</i> (SEVER)	307
8.5.1.8.	<i>-mə</i> (SPLIT)	308
8.5.1.9.	<i>-thəm</i> (END).....	310
8.5.1.10.	<i>-lak</i> (TERM).....	310
8.5.1.11.	<i>-thən</i> (TOGETHER)	312
8.5.1.12.	<i>-tshəŋ</i> (ATTACH)	314
8.5.1.13.	<i>-sət</i> (AFFECT)	316
8.5.1.14.	<i>-tʃəm</i> (ADDIT)	319
8.5.1.15.	<i>-ruk</i> (ASSIST)	320
8.5.1.16.	<i>-phaŋ</i> (COVER).....	320

8.5.1.17.	- <i>tak</i> (REMAIN).....	321
8.5.2.	Reciprocal/collective - <i>təp</i>	321
8.5.3.	Directional suffix - <i>zək</i>	322
8.5.4.	Aspectual suffixes.....	325
8.5.4.1.	Continuative marker - <i>ja</i>	325
8.5.4.2.	Repetitive marker - <i>si</i>	326
8.5.4.3.	Completive marker - <i>ma?</i>	327
8.5.4.4.	Habitual past marker - <i>ja?</i>	328
8.5.5.	Modality suffixes.....	329
8.5.5.1.	Abilitive modality - <i>tʃhət</i>	329
8.5.5.2.	Frustrative modality - <i>phət</i>	330
8.5.5.3.	Conative modality - <i>tsə</i>	332
8.5.5.4.	Desiderative modality - <i>mì</i>	333
8.5.5.5.	Chance modality - <i>tʃha?</i>	334
8.5.6.	Resultant state suffix - <i>tʃak</i>	336
8.5.7.	Perfective aspect suffix - <i>tʃuk</i>	338
8.5.8.	Causative suffix - <i>i?</i>	340
8.5.9.	Negative suffix - <i>la</i>	341
8.5.10.	Positive imperative suffix - <i>aŋ</i>	342
8.5.11.	Converb suffixes.....	343
8.5.12.	Tense and modality.....	344
8.5.12.1.	Categories of past and present.....	346
8.5.12.2.	Categories of irrealis/immediate future.....	349
8.5.12.3.	Anterior.....	353
8.5.12.4.	Summary of TAM categories.....	354
8.5.13.	General nominalizer - <i>pà?</i>	354
8.6.	Concluding comments.....	355

9 Verbless, copula and existential clauses

9.1.	Verbless clauses.....	357
9.1.1.	Identity clauses.....	358
9.1.2.	Negated identity clauses.....	360
9.1.3.	Interrogative identity clauses.....	361
9.1.4.	Ascriptive clauses.....	362
9.1.5.	Negated ascriptive clauses.....	368
9.1.6.	Interrogative ascriptive clauses.....	369
9.1.7.	Cardinal numerals as attributes.....	370
9.1.8.	Expression of tense/aspect/modality in copula clauses.....	371
9.2.	Verbal existential clauses.....	374
9.2.1.	Expression of existence in declarative clauses.....	375
9.2.2.	Expression of existence in interrogative clauses.....	379
9.3.	Extended functions of existentials.....	381

9.3.1.	Mood marking in circumstantial converb clauses.....	381
9.3.2.	Grammaticalized <i>tʃhà</i> ‘do’ as a discourse connective.....	381
9.3.3.	Happening.....	382
9.4.	Possessive verbal clause.....	383
9.5.	Summary.....	385
9.6.	Notes on the origin of copulas.....	386

10 Imperatives

10.1.	Imperative as a distinct clause type.....	387
10.2.	Expression of imperatives.....	389
10.2.1.	Positive imperative.....	390
10.2.2.	Prohibitive.....	391
10.2.3.	Admonitive.....	392
10.2.4.	Semantic verb classes and imperative mood.....	393
10.3.	Dependencies between imperatives and other categories.....	395
10.3.1.	Absence of continuative aspect marking.....	395
10.3.2.	Causative marking in imperatives.....	397
10.3.3.	Other grammatical categories and imperative mood.....	399
10.3.4.	Imperatives and evidentiality.....	402
10.4.	Semantics of imperatives.....	402
10.4.1.	Use of particles in imperative mood.....	403
10.4.2.	Non-command meanings of imperatives.....	404
10.4.3.	Imperatives and discourse.....	405
10.5.	Notes on the origin of imperatives.....	406
10.5.1.	Positive imperative.....	406
10.5.2.	Prohibitive and admonitive.....	408

11 Clause combining

11.1.	Preliminary remarks.....	409
11.2.	Definitions of the converb.....	410
11.3.	The syntactic status of Mongsen converbs.....	413
11.4.	Types of converbs.....	421
11.4.1.	Narrative chaining.....	422
11.4.1.1.	Sequential converb.....	424
11.4.1.2.	Simultaneous converb.....	425
11.4.1.3.	Alternating iterative converb.....	428
11.4.2.	Negation and narrative chaining converbs.....	431
11.4.3.	Specialized converbs.....	435
11.4.3.1.	Conditional converb.....	435
11.4.3.2.	Circumstantial converb.....	437
11.4.3.3.	Concessive converb.....	439
11.4.3.4.	Contemporative converb.....	440
11.4.3.5.	Durative converb.....	441

11.4.3.6.	Causal converb.....	442
11.4.3.7.	Locative converb.....	445
11.4.4.	Notes on the diachronic origin of converbs	446
11.4.5.	Expression of temporal antecedence	447
11.4.6.	Disjunctive expression using converbs	448
11.5.	Space to time grammaticalizations.....	449
11.6.	Causal particle.....	450
11.7.	Complementation	451
11.7.1.	Embedded direct speech complements	452
11.7.2.	Other complement taking predicates.....	453

Texts

1.	The Story of the Longtailed Broadbill	455
2.	The Story of Noksensangba and Imtisangba.....	458
3.	The Founding of Mangmetong Village.....	478
4.	The Birds' Conference	481

English – Mongsen Ao glossary

1.	Quantity, dimension and measurement	487
2.	Natural world	487
3.	Time	488
4.	Animals – land	489
5.	Animals – water	489
6.	Birds	489
7.	Insects, arachnids and invertebrates	490
8.	Plants	490
9.	Human relationships	491
10.	Types of people	491
11.	Body parts and fluids – human and animal	492
12.	Life, death, illness and injury	493
13.	Human activity, experience and belief.....	494
14.	Food	494
15.	Weapons and warfare.....	495
16.	Cultural artefacts, possessions	495
17.	Buildings, structures and locations	496
18.	Cultural terms – community, tradition	497
19.	Language and communication	497
20.	Verbs – stative.....	497
21.	Verbs – sense and cognition.....	499
22.	Verbs – activity	499

References	505
-------------------------	-----

Index of subjects	517
--------------------------------	-----

List of maps, tables and figures

List of Maps

Map 1.	The north-eastern states of India	xxi
Map 2.	Distribution of tribes neighbouring the Ao territory.....	xxii
Map 3.	Location of Ao villages	xxiii

List of Tables

Table 2.1.	Phonotactic distribution of syllable constituents	23
Table 2.2.	Consonant phonemes of Mangmetong Mongsen	28
Table 2.3.	Vowel phonemes of Mangmetong Mongsen.....	45
Table 3.1.	Examples of tonal realizations in selected verbal suffixes	71
Table 3.2.	Examples of words causing tonal alternations.....	72
Table 4.1.	Summary of nominals that function as NP heads.....	82
Table 4.2.	Summary of formal properties of nominals.....	83
Table 4.3.	Citation and vocative forms of some kinship terms	85
Table 4.4.	Personal pronouns	89
Table 4.5.	Possessive pronouns	94
Table 4.6.	Interrogative pronouns.....	99
Table 4.7.	Negative indefinite pronouns.....	101
Table 4.8.	Specific indefinite pronouns.....	102
Table 4.9.	Non-specific indefinite pronouns	103
Table 4.10.	Lexical quantifiers	112
Table 4.11.	Mongsen and Proto-Tibeto-Burman numerals	113
Table 4.12.	Mangmetong Mongsen case marking clitics	121
Table 4.13.	Time words.....	122
Table 4.14.	Terms for divisions of the day.....	123
Table 4.15.	Common exclamations	130
Table 4.16.	Examples of onomatopoeia	130
Table 5.1.	Mangmetong Mongsen case marking clitics	171
Table 5.2.	Nascent postpositions	185
Table 7.1.	Distributive numerals	276
Table 8.1.	Morphological structure of the predicate.....	289
Table 8.2.	Lexical suffixes	299
Table 8.3.	Converb suffixes.....	343
Table 9.1.	Summary of clause types expressing copulative meanings	386
Table 10.1.	Summary of formal criteria distinguishing clause types	389
Table 11.1.	Converb suffixes.....	422
Table 11.2.	Possible diachronic origins of Mongsen converb suffixes	447

List of Figures

Figure 2.1.	Metrical structure of the Mongsen syllable	24
Figure 2.2.	Waveform and spectrogram (313 Hz) of <i>mama</i> ‘breast’	34
Figure 2.3.	Waveform and spectrogram (313 Hz) of <i>hmapaŋ</i> ‘time’	34
Figure 2.4.	Waveform and spectrogram (313 Hz) of <i>nì</i> ‘I’	35
Figure 2.5.	Waveform and spectrogram (313 Hz) of <i>hni</i> ‘followed’	36
Figure 2.6.	Waveform and spectrogram (234 Hz) of <i>ŋaŋsi</i> ‘cursed’	37
Figure 2.7.	Waveform and spectrogram (234 Hz) of <i>hŋa</i> ‘listened’	37
Figure 2.8.	Waveform and spectrogram (234 Hz) of <i>wa</i> ‘went’	39
Figure 2.9.	Waveform and spectrogram (234 Hz) of <i>hwaŋsi</i> ‘scolded’	40
Figure 2.10.	Waveform and spectrogram (469 Hz) of <i>ju</i> ‘carried’	41
Figure 2.11.	Waveform and spectrogram (469 Hz) of <i>hju</i> ‘angled’	41
Figure 2.12.	Waveform and spectrogram (313 Hz) of <i>ɿ</i> ‘stalked’	42
Figure 2.13.	Waveform and spectrogram (313 Hz) of <i>hɿ</i> ‘sewed’	42
Figure 2.14.	Waveform and spectrogram (313 Hz) of <i>li</i> ‘be’	43
Figure 2.15.	Waveform and spectrogram (313 Hz) of <i>hli</i> ‘bought’	44
Figure 3.1.	Waveform and F ₀ trace demonstrating phrasal intonation	74
Figure 3.2.	Waveform and F ₀ trace showing the sequential converb	75
Figure 3.3.	Waveform and F ₀ trace showing intonation on <i>la</i> (TOP)	77
Figure 5.1	Causative derivation of intransitive verb stems	191
Figure 5.2	Causative derivation of transitive verb stems	195
Figure 5.3	Causative derivation of three-place verb stems	196
Figure 6.1	Hierarchical structure of the complex noun phrase	207
Figure 7.1.	Summary of functions of the nominalizing prefix	259
Figure 8.1.	Directional motion verbs	281
Figure 8.2.	Structure of the verb stem <i>kəp-saʔ-zək-tʃuk</i>	300

Abbreviations and conventions

The following abbreviations are used in the text:

A	the agent-like argument of a transitive clause
S	the single argument of an intransitive clause
O	the patient-like argument of a transitive clause
BL	Burmese-Lolo
F ₀	fundamental frequency
Hz	Hertz
Lit.	literally
PTB	Proto-Tibeto-Burman
ms	milliseconds
STC	<i>Sino-Tibetan: a Conspectus</i> (Benedict 1972). Numbers preceded by a hash mark refer to reconstructed Tibeto-Burman roots, as numbered in Benedict, e.g. [<PTB *za ‘child’ STC #59] indicates that the reconstructed form for Proto-Tibeto-Burman ‘child’ is No. 59 in Benedict (1972). The bibliographic details of Benedict (1972) can be found in the References.

Symbols

√	lexical root	NOM	nominalization
+	fused morphemes	NP	noun phrase
σ	syllable	T	any tone
C	consonant	V	vowel
CC	copula complement	VCC	verbless clause complement
CT	copula topic		
G	glide	VCT	verbless clause topic
H	high tone	TAM	tense/aspect/modality
L	low tone	N	noun
ACCOMP	‘ACCOMPANY’ lexical suffix		

Abbreviations used in glosses

1	first person	ADDIT	‘ADDITIVE’ lexical suffix
2	second person	ADM	admonitive mood
3	third person	AFFECT	‘AFFECT’ lexical suffix
ABIL	abilitative modality	AGT	agentive case
ABL	ablative case	ALL	allative case
M	mid tone		

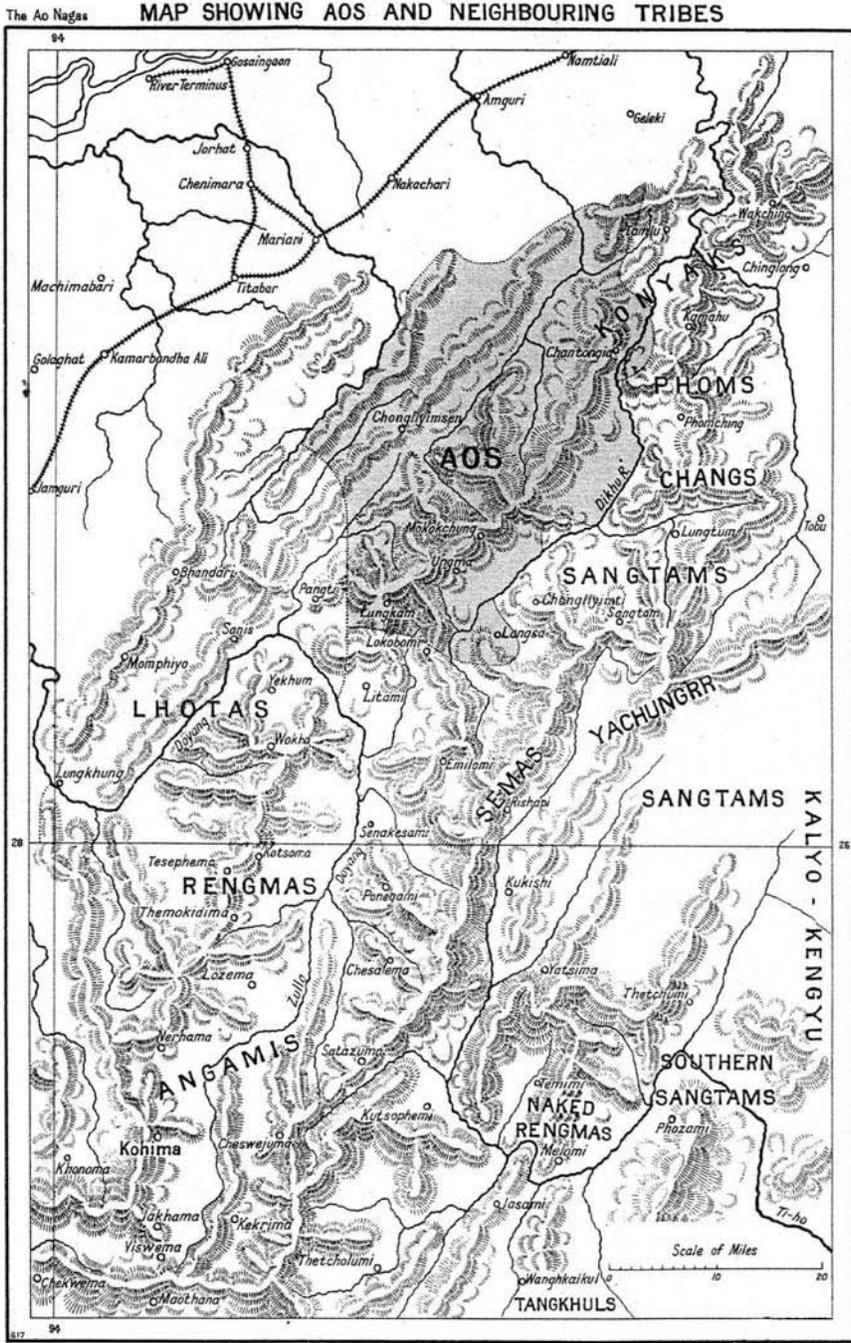
ALT.IT	alternating iterative converb	F	feminine semantic gender
ANAPH	anaphoric demonstrative	FACE	'FACE' nascent postposition
ANOM	agentive nominalization	FRUS	frustrative modality
ANT	anterior tense/aspect	FUT	future tense
ASCEND	'ASCEND' lexical suffix	GPN	indefinite generic pronoun
ATTACH	'ATTACH' lexical suffix	HAB	habitual
BACK	'BACK' nascent postposition	IMM	immediate future tense
BE.COP	copula (< <i>li</i> 'be, stay')	IMP	imperative mood
BEN	benefactive case	IN	inclusive
CAUS	causative	INCEPT	inceptive aspect
CAUSAL	causal converb	INSIDE	'INSIDE' nascent postposition
CHANCE	'CHANCE' modality	INST	instrumental case
CIRCM	circumstantial converb	INTERR	interrogative
CMPL	completive aspect	INTJ	interjection
COLL	collective	INTNS	intensifier
COM	comitative case	INTO	'INTO' lexical suffix
CON	conative modality	IRR	irrealis mood
CONTEMP	contemporative converb	LADV	locative adverbial suffix
CONJ	phrasal conjunction	LNOM	locative nominalization
CONCESS	concessive converb	LOC	locative case
COND	conditional converb	LOC.CV	locative converb
CONT	continuative aspect	LOC.SUFF	locational suffix
COP	copula	M	masculine semantic gender
DAT	dative case	MAN	'MAN' agentive nominalizer
DEC	declarative mood clitic	NEG	negative
DEIC	deictic nominalization	NF	non-final
DESCEND	'DESCEND' lexical suffix	NOM	nominalization
DESID	desiderative modality	NR	general nominalizer
DIM	diminutive	NRL	non-relational prefix
DIR	directional suffix	NZP	nominalizing prefix
DIST	distal demonstrative	NUM	numeral adverb suffix
DU	dual number	ONOM	onomatopoeia
DUR	durative converb	ORD	ordinal
ECHO	echo reduplication	PFV	perfective aspect
EMPHAT	emphatic demonstrative	PL	plural number
END	'END' lexical suffix	PN	personal name
EX	exclusive		
EXCLM	exclamation		

xx *Abbreviations and conventions*

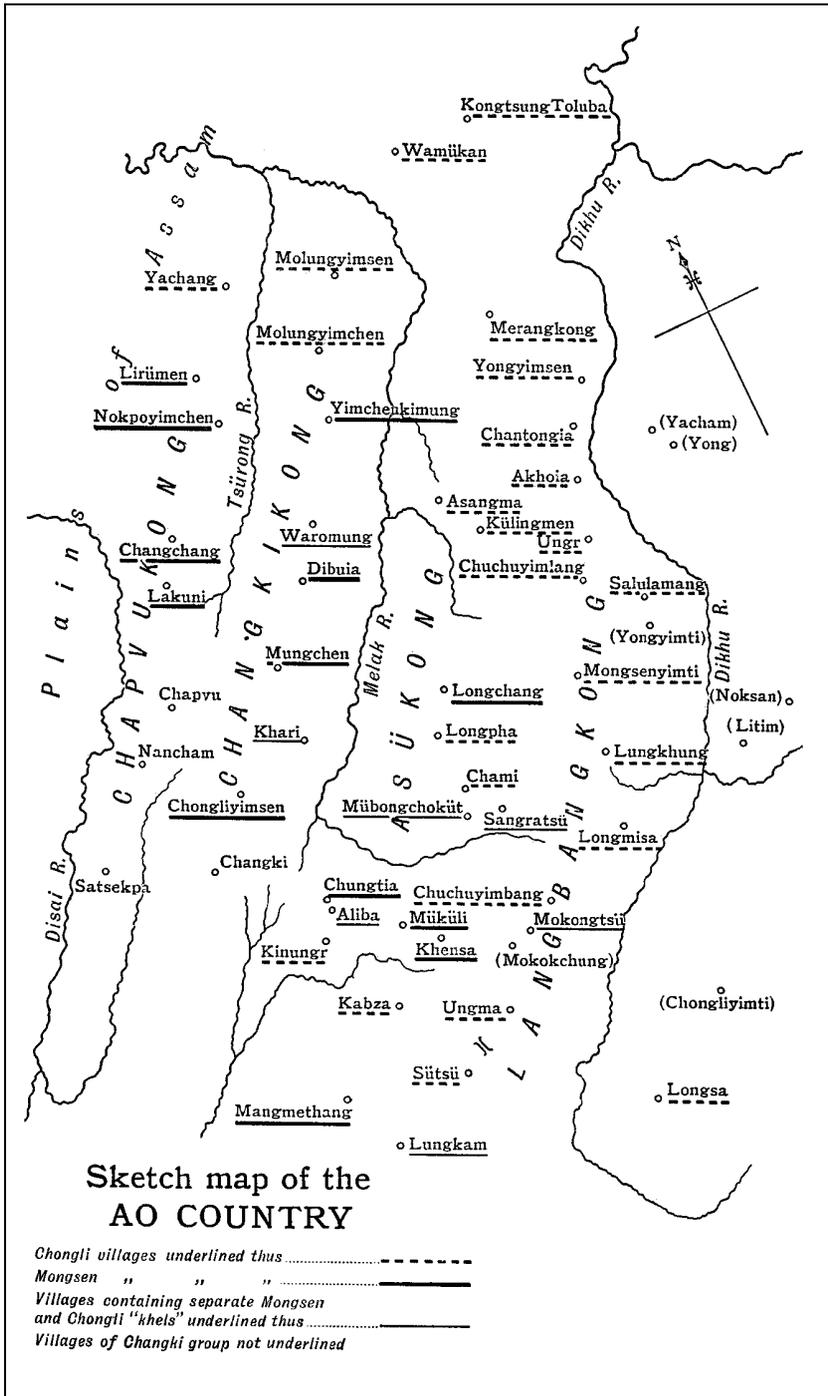
PLACE	'PLACE' nascent postposition	SEQ	sequential converb
POSS	possessive	SG	singular number
PRES	present tense	SIDE	'SIDE' nascent postposition
PROH	prohibitive	SIM	simultaneous converb
PROX	proximate demonstrative	sp.	species
PST	past tense	STOMACH	'stomach' nascent postposition
PTCL	particle	SUP	superlative
PURP	purposive nominalization	TERM	'TERMINATIVE' lexical suffix
QPTCL	interrogative particle	TO.AND.FRO	'TO.AND.FRO' lexical suffix
REACH	'REACH' nascent converb suffix	TOGETHER	'TOGETHER' lexical suffix
REAL	realis mood	TOP	topic particle
RECIP	reciprocal	VOC	vocative
RED	reduplication	VSF	vacuous stem formative
REFL	reflexive marker	SPREAD	'SPREAD' lexical suffix
REP	reportative evidentiality	SPLIT	'SPLIT' lexical suffix
RL	relational prefix	WHEN	'WHEN' nascent converb suffix
RPET	repetitive aspect		
RS	resultant state marker		
SEPARATE	'SEPARATE' lexical suffix		
SEVER	'SEVER' lexical suffix		



Map 1. The north-eastern states of India (from Coupe 2003a: xvii)



Map 2. Distribution of tribes neighbouring the Ao territory (from Mills 1926, map facing p.1)



Map 3. Locations of Ao villages (from Mills 1926, map facing p.4)

Chapter 1

Introduction

1.1. The Ao language

This book describes the grammar of Mongsen, one of two major dialects of the Ao language. According to the 1991 Census of India, Ao is spoken by 170,000 people. It is estimated that perhaps forty percent of that number speak Mongsen as their first language, fifty percent speak the prestige Chungli dialect as their first language and the remaining ten percent speak other Ao dialects.

The Mongsen dialect demonstrates a range of varieties whose lexicons, phonology, phonological processes and morphological forms can differ slightly from village to village. For this reason, each variety of Mongsen is identified by the village where it is spoken when discussed in the text. Language examples are similarly identified, e.g. ‘Longchang Mongsen’ refers to a variety of Mongsen spoken in Longchang village.

The variety of Mongsen described in this work is spoken by approximately 2000 people living in 441 households in the village of Mangmetong, located on the southern edge of Mokokchung district and approximately 10 kilometres south-west of the district town of Mokokchung. The Mangmetong Mongsen data is supplemented by, and on occasion compared to, additional data collected from other Mongsen-speaking villages and individuals; all sources are documented in §1.11.

1.2. Organization of this grammar

The grammatical description is divided into eleven chapters. A treatment of the phonology and phonological processes follows in Chapter Two. The tone system and the use of intonation are then described in Chapter Three. Chapter Four presents criteria for the identification of word classes. Clause structure, grammatical relations, case marking and valency modifying derivations are addressed in Chapter Five. A description of the noun phrase, relativization and nominalization follows in Chapter Six, after which the nominal morphology is discussed in Chapter Seven. Chapter Eight analyzes the prodigious verbal morphology of Mongsen. Chapter Nine presents an analysis of the functions of verbless, copula and existential clauses. Chapter Ten describes imperatives and other types of commands. Chapter Eleven presents a description of clause linkage. A selection of annotated texts demonstrating different speech genres follows the grammatical description. Lastly, a glossary of all lexical items

occurring in the language examples and texts completes the work. These are arranged under sub-headings according to their semantic fields.

Cross-references to section headings in the text are preceded by the symbol §, e.g. §5.4.2. Language examples are numbered separately in each chapter and cross-references are given in parentheses in the text; e.g. (8.15) identifies the fifteenth example of Chapter Eight.

1.3. Typological overview

This section gives a brief overview of the salient features of the language and provides cross-references to the description of each feature in the subsequent chapters. This is designed to serve as a guide to readers who would like to quickly ascertain the salient points of interest in the grammar of Mongsen without having to rummage through the whole of the book in order to find them.

The Mongsen dialect has two principal sub-types based on the presence or absence of a typologically rare voiceless sonorant series, which adds seven phonemes to a basic inventory of twenty consonant segments (§2.2.1). There is a two-way voice onset time contrast in initial stops and affricates. The vowel phoneme inventory presents a simple triangular system of three peripheral vowels and one central vowel (§2.2.2). Some Mongsen varieties have innovated an additional high central rounded vowel in their inventories, and all are characterized by the presence of contrastive creaky voice occurring on just one articulatory position in a limited number of lexical items. The latter necessitates the addition of a fifth phoneme to the basic vowel phoneme inventory.

Like all the Tibeto-Burman languages of Nagaland, Mongsen is a tonal language, with three lexical tones occurring on all syllable types (§3.1.1). Tone has a limited marking function for some grammatical categories, in addition to its lexical contrastive function (§3.1.3). Extensive internal and external tone sandhi processes operate in word formation and across adjacent words (§3.3). Interacting with and often overriding the lexical tone system are patterns of intonation that delineate higher-level constituents of the clause (§3.4). A glottal stop that is restricted to occurring word-finally is analyzed as a word prosody (§3.5), but also demonstrates a contrastive function with allowable final stops; this contrastive function is analyzed as an ancillary corollary of its word-final realization.

There are open word classes of just nouns and verbs (§4.1). The closed word class includes pronouns, nominal modifiers, and peripheral clausal constituents (§4.2). A sizable class of particles and clitics encode case marking, illocutionary force, topic marking, additive or restrictive focus marking, and one category of evidentiality (§4.2.14).

Mongsen demonstrates an unusual system of case-marking on core arguments that is yet to be adequately incorporated into the typology of core case-marking alignments. Case marking of core arguments can be pragmatically and/or semantically motivated in both transitive and intransitive clauses and appears to be sensitive to a number of interrelated parameters that are discussed in detail in §5.2. The language exhibits dependent marking characteristics at the clause level and has an extensive array of grammatical and semantic case-marking clitic particles. A number of body part and locational nouns have undergone recategorization and function as nascent postpositions (§5.3). There is a morphological causative derivation applying to both transitive and intransitive verb roots, and an analytic causative formed with the verb ‘to give’ (§5.4).

The noun phrase allows for recursive embedding (§6.5). Relative clauses function as nominal attributes of their noun phrase heads and are found in both the pre-head and post-head attribute slots of the noun phrase. Pre-head relative clauses potentially encode a referentially restrictive meaning. Both headed and headless relative clauses occur. Headless relative clauses are formally indistinguishable from nominalized clauses and are marked by the same nominal morphology, but can be distinguished by criteria presented in §6.6. While the majority of relative clauses are externally headed, internally-headed relative clauses are also a possible, albeit rarely used strategy (§6.6.1). In common with many Tibeto-Burman languages of the South Asia, a structure resembling a relative-correlative construction offers an alternative means of encoding relativization in Mongsen (§6.6.6).

Nominal compounding is a heavily exploited means of expanding the lexicon and utilizes both nominal and verbal roots (§7.1). Compounding processes appear to have provided the diachronic source for many grammatical morphemes. Mongsen is typically Tibeto-Burman in the way that it makes extensive use of nominalizing derivational morphology (§7.4).

Virtually any transitive verb stem can be used with a single noun phrase argument corresponding to the O argument of a transitive clause. A good number of these these qualify as *bona fide* ambitransitive verbs; i.e. they do not require the antecedent mention of an elided argument to qualify their use. In contrast, the use of ambitransitive verbs in which a single argument corresponds to the A argument of a transitive clause is considerably rarer (§8.1.3). The highly agglutinating nature of the morphology of Mongsen reaches its zenith in the complexity of its clause-final predicate structure (§8.2). Numerous lexical suffixes that have grammaticalized from verb roots are used to express a whole gamut of resultative and directional aspectual meanings (§8.5). These, too, appear to have their diachronic sources in compounding processes.

Equational, ascriptive and existential meanings are expressed by both verbless clauses and verbal copula clauses (§9.1). Two existential verbs

4 Introduction

demonstrate extended functions and are used to encode various grammatical meanings in other areas of the grammar (§9.3).

A characteristic Tibeto-Burman feature of Mongsen discourse is the use of converbs to create chains of dependent clauses. A basic division can be made between narrative chaining converbs (§11.4.1), and specialized converbs that express a range of presuppositional meanings (§11.4.3). Almost all of the converb markers have phonological forms that are either identical to or segmentally very close to morphemes belonging to other mostly nominal grammatical categories (§11.4.4). The diachronic development of the converb system of Mongsen thus proves to be typologically consistent with grammaticalization patterns attested in many verb-final languages of Eurasia.

1.4. Linguistic situation in Nagaland

The north-east of India is one of the most linguistically diverse regions in the world. In Nagaland alone there are at least twenty indigenous minorities living within its 16,579 square kilometres, each speaking a distinct and mutually unintelligible language. The names by which the languages of these minorities are known are Angami, Ao, Chang, Chokri, Kachari, Khezha, Khiamniungan, Konyak, Kuki, Liangmei, Lotha, Phom, Pochuri, Rengma, Rongmei, Sangtam, Sumi (formerly known as Sema), Tikhir, Yimchungrü and Zeme. It is likely that every indigenous community has a number of discrete dialects demonstrating varying degrees of inter-dialect intelligibility, and that every village of every community has its own particular variety of a dialect.

With the exception of Kachari and Kuki, in the past there has been a tendency to refer to all these tribal languages as “Naga” languages, and collectively to the people who speak them as “Nagas” with the indigenous name as a modifier, e.g. ‘Rengma Naga’, (Mills 1937) and more recently ‘Mao Naga’ (Giridhar 1994) and ‘Ao Naga’ (Coupe 1998). At present the only so-called “Naga” languages of Nagaland for which a convincing higher level affiliation can be established are those of the Bodo-Konyak-Jinghpaw branch, viz. Chang, Phom, Kiamniungan and Konyak, spoken in the Mon and Tuensang districts of northern and eastern Nagaland. These jointly form the Konyak sub-group of Bodo-Konyak-Jinghpaw. Following the counsel of Burling (2003: 172), I henceforth make no assumptions about a higher level genetic relationship holding between any of the other Tibeto-Burman languages of Nagaland and make a point of avoiding the use of the term “Naga” as a linguistic label altogether.

The traditional territory inhabited by the Ao is surrounded by communities speaking other languages (see Maps 1 and 2). With the exception of the Indo-Aryan language Assamese and the derivative lingua franca Nagamese, all

neighbouring languages belong to the Tibeto-Burman family. Assamese is spoken on the plain of Assam to the west and north-west in Jorhat and Sibsagar districts respectively. Lotha is spoken in the villages of Wokha district to the south-west. Sumi villages are located in the south and south-east in Zunheboto district, and on bordering lands previously occupied by the Ao. In the east on the far side of the Dikhu River in Tuensang district are Sangtam villages, some also on sites previously inhabited by the Ao. To the east and a little further to the north are Chang villages. The trans-Dikhu Ao village Longsa speaks a dialect of Chungli strongly influenced by Chang, the result of an intensive language contact situation. To the north-east are found Phom villages, and a Phom village previously known as Assiringia in the earlier literature and currently as Nokpo¹ is located deep in the Ao territory. Ao consultants report that this village's language is heavily influenced by Ao, but retains Phom characteristics, making it mostly unintelligible to them. The Ao dialect spoken in the villages of Yong and Yacham at the eastern periphery of the Ao territory is also heavily influenced by a language contact situation with Phom-speaking communities and is said to be unintelligible to the majority of Ao speakers. Lastly, at the north-eastern tip of Mokokchung district in the district of Mon are found communities speaking undocumented Konyak languages.

Throughout Nagaland, in the Tirap district of Arunachal Pradesh to the north and across the border in adjacent areas of Burma, a lingua franca based on an imperfectly learned form of Assamese known as Nagamese is spoken as the language of wider communication. Nagamese has its regional variants, with the phonology of L1 determining the extent of the regional variation (see Sreedhar 1974). Mongsen speakers, for example, pronounce the past tense inflection of the Nagamese verb 'to go' *ja-ise* (go-PST) as [tʃajʃi] when speaking Nagamese, because Mongsen affricates lack a voiced series, and because the nearest vowel in the Mongsen vowel phoneme inventory to the front close-mid vowel of the Nagamese past tense inflection is the high front vowel *i*. This in turn triggers the palatalization of the preceding sibilant, resulting in a characteristic accent.

Nagamese is an important and indispensable means of communication in towns with highly mixed populations, such as Kohima and Dimapur, where it is used as a contact language when native speakers of Assamese, Bengali, Hindi, Nepali and the various Tibeto-Burman languages of the eastern border region speak with anyone outside of their own linguistic community. Some sections of the Kachari community in Dimapur use it exclusively as a first language. Now Nagamese is gradually mounting its threat to the indigenous languages of

1. This village is also known by the exonym Mirinokpo, a label now rejected by the inhabitants. This is because it contains the pejorative *miri*, an unflattering term used by the Ao to describe the trans-Dikhu River tribes (i.e. Phom, Chang, Konyak and Khiamniungan).

Nagaland. Children who are native speakers of Mongsen living in Mongsen-speaking households in Dimapur have been observed to play using Nagamese as the medium of communication, even when the participants in their games are all Mongsen speakers. When people of different tribes marry, it is usual for them to communicate predominantly in Nagamese, which often becomes the first language of their children. Families of government officials that are posted outside of their communities in towns tend to be heavily influenced by Nagamese. Their children use it at school and with their friends, and if the family is affluent, it employs domestic servants from the plains known as “minnies”, who speak with their employers in Nagamese. The children of such families may become semi-speakers of their ancestral indigenous languages, adopting Nagamese as their first language.

It appears that up to now Nagamese has had little impact in the village milieu, however. In Mangmetong village in 1999, for example, all of the children were observed to be monolingual Ao speakers, as were many of the women. The only adults who were bilingual in both Mongsen and Nagamese were those who had spent some time living outside of the village in towns with mixed populations.

It is fairly unusual for marriage partners from different Tibeto-Burman linguistic communities in Nagaland to learn each other’s languages. This is probably attributable to the fact that Nagamese conveniently serves as the medium of communication in mixed marriages. Some residents of the upper ward of Mangmetong village own fields that abut those of the Lotha village of Pangti and are able to speak Lotha. As a result of this contact, some people of this ward have intermarried with Lotha speakers and this has facilitated bilingualism, an otherwise rare occurrence in the Ao community. I have occasionally met an Ao by birth who has grown up in another district of Nagaland and has consequently become fluent in that district’s dominant language but, as noted above, they are usually only semi-speakers of their indigenous language as a consequence of having limited contact with their Ao linguistic community.

There appears to be little awareness of the need for language maintenance amongst the Ao. This is perhaps because they have long been the dominant tribe of the state and do not as yet perceive any threat to the continued existence of their language. Even so, amongst the educated youth there is a growing awareness that the archaic language of poetry and song is gradually becoming incomprehensible to younger generations of Mongsen speakers. This is made patently apparent by the struggle that younger speakers often have when helping to translate the texts of their elders, particularly if the recorded narratives contain traditional songs.

1.5. Genetic classification

A number of genetic classifications of Tibeto-Burman languages have been proposed since the study of this language family began in the nineteenth century. Coupe (2003a: 5–8) provides an overview of the position of Ao in relation to the Tibeto-Burman family according to the interpretations of Shafer (1955, 1974), Benedict (1972), Marrison (1967) and Bradley (1997) and need not be repeated here. With the partial exception of Marrison's (1967) typologically-oriented classification (which also incorporated lexical correspondences), genetic sub-groupings of the languages of north-east India remain based on lexical comparisons and geographic proximity.

As Thurgood (2003: 5) points out, sub-groups based on shared morphological innovations or full historical reconstructions are more convincingly established than are those established on lexical correspondences alone. A perennial obstacle baulking a definitive sub-grouping of the languages of north-east India has been the lack of reliable data. In many cases, comparativists have had to make do with ancient word lists collected over a century ago by colonial administrators, and for many languages whose existence we know of, there is no data available at all.

Burling (2003) proposes a new genetic classification of the Tibeto-Burman languages of north-east India. With respect to the place of Ao in the family, this largely accords with earlier classifications by Grierson ([1903] 1967) and Marrison (1967), and with Shafer (1955), who also suggested a lower-level Ao unit. The basis for Burling's proposed sub-groupings is lexical correspondences, perhaps augmented by the consideration of Marrison's typological data. Some of Marrison's sub-groupings involving Ao may need to be reassessed in light of the findings documented in §2.2. Burling's (2003) classification is justifiably conservative in not attempting to posit any higher level relationships beyond his Ao, Angami-Pochuri, Zeme, Tangkhul and Mizo-Kuki-Chin groups, each named for the most prominent language(s) in the group. He concludes that further progress will not be made on the classification of the Tibeto-Burman languages of north-east India until more descriptive work can be done.

1.6. Geographical setting

Ao is the major language of the Mokokchung district, one of the eleven administrative districts of the hill state of Nagaland, located in north-eastern India. Nagaland lies at the eastern edge of India; it borders Burma (Myanmar) to the east, the hill states of Arunachal Pradesh and Manipur to the north and south respectively, and the state of Assam to the west (see Map 1).

A number of parallel ridges project off the main Himalayan range in an arching south-south-westerly direction and effectively divide the plain of Assam from the Hukawng Valley and the plain of the Chindwin River in Burma to the east. These mountains form an important barrier not only geographically, but also linguistically and culturally, because they divide South Asia from South-East Asia. Varying in height from a few hundred metres in northernmost Nagaland to a maximum of 3826 metres above sea level in southern Nagaland, the range forms a series of ridges on the western side that are known as the Naga Hills. These run through the Mokokchung District (see Map 3). The Naga Hills together with the Patkai range further to the east and north-east create a vast watershed whose rivers drain in a westerly direction into Assam, eventually flowing into the Brahmaputra River. The most westerly ridges of the Naga Hills rise out of the plain of Assam and these delineate the western edge of the traditional Ao territory. To the east, the Dikhu River marks the eastern border of the Mokokchung district, although a few Ao villages are also to be found on the eastern side of this river. All Ao villages lie between 26° 12' N and 26° 45' N, and from 94° 18' E to 94° 46' E.

The ranges of the Mokokchung district have local names (see Map 3). The westernmost inhabited range is called Japukong ('kong' means range), named after Japu village (the range and village names are represented as 'Chapvukong' and 'Chapvu' respectively in Map 3). This low-lying range rises no more than a few hundred metres above sea level. Running parallel with the Japukong is the Changkikong, which is named after Changki village. The Changkikong is a major range and dominates the landscape when viewed from the east. On the eastern edge of the Ao territory is the Langpangkong, running parallel with the western bank of the Dikhu River; this is the location of the largest number of Chungli villages. The Asetkong is a lower range lying between the Changkikong and the Lambangkong and is the location of both Mongsen- and Chungli-speaking villages, some of which are bidialectal. A fifth range called the Ongpangkong is located in the south of the Mokokchung district and runs off the Langpangkong in a roughly westerly direction. This is home to a mix of Mongsen and Chungli villages, with some also found to be bidialectal. The Ongpangkong is not recognized as a separate range in the maps drawn by British cartographers during the colonial era, but this is an important range to the Ao and it figures prominently in their oral tradition, particularly in that of the Mongsen moiety. It is also the highest range in Mokokchung district; the village of Longkum is situated at an altitude of 1846 metres above sea level.

A sixth low-lying range known as the Tsürangkong (named after the Tsürang River) is not shown in Map 3, but can be seen running down the western side of the Japvukong in Map 2. Although this range has no traditional village sites, its river valley has long been used for wet rice cultivation, and it is the location of the administrative town of Mangkolemba.

Because of the constant threat of raids by head-hunting parties in the past, all Ao villages were by necessity strategically located in commanding positions on the tops of hills, so that they could be easily stockaded and defended. As a result of this legacy, many villages are found literally perched on the spine of a range or at the end of a spur. Usually the main street of the village runs down the highest point of a ridge and houses flank its sides. The Ao thus typically live in a mountainous environment, often having to walk considerable distances to reach their lower fields.

1.7. Ethnographic overview

The following subsections provide socio-cultural information on the Ao, including the division of the tribe into linguistically distinguished sub-groups, their folkloric origin, the consequences of their contact with British colonists and then missionaries, their daily existence as subsistence farmers and hunter-gatherers, and their religion.

1.7.1. Autonym, exonyms and allonyms

The established exonym of the tribe is “Ao”, and the autonym is [a³³ʌəɬ³³] or [a³³wəɬ³³] in Mongsen, depending upon the phonology of the variety (see §2.3). The etymology of the word “Ao” is somewhat obscure. Mills (1926: 1) states that, according to his consultants, it is a mispronunciation of *Aor*, meaning ‘those who came’ (presumably across the Dikhu River), as opposed to *Mirir*, ‘those who didn’t come’.

This explanation has certain inconsistencies that are suggestive of a folk etymology. Firstly, the word *Ao* is from the Chungli dialect, in which [aw] is the verb ‘go’, not ‘come’, therefore the deictic orientation of this verb focuses on the wrong side of the river. This makes it seem more like an exonym than an autonym. The retroflex approximant at the end of the word *Aor* is possibly the agentive nominalizer (§7.4.4) that derives a nominal from a verb root, so that this should actually mean ‘the going ones’. As for *Mirir*, it is possible that the first syllable is the negative prefix (§8.3.3), and that the root is the Chungli verb [aɹ] ‘come’, although there are some uncharacteristic vowel changes in the derivation. Overlooking these problems, *Mirir* would mean literally ‘the not coming ones’, which fits with one of the explanations of Mills’ consultants. Currently the term *mizi* is used in Mongsen for a member of any trans-Dikhu River tribe and has somewhat pejorative connotations, hence its rejection by the inhabitants of Nokpo village, who consider themselves to be Ao (see fn. 1). An alternative interpretation of the word *Mirir* is that it is formed with the root *mizi*

suffixed by the agentive nominalizer $-(ə)l$, which would derive a verbal noun meaning something akin to ‘the hostile ones, the enemy’. This assumption is based upon a dictionary entry for *meri* ‘hostile, an enemy, hateful’ in Clark ([1911] 1990: 420).

A further possibility for the source of the Ao autonym $[a^{33}m̩ɛ^{33}]$ or $[a^{33}w̩ɛ^{33}]$, and perhaps the most convincing, is suggested by the morphology and meaning of the Yimchungrü nominalization $a^{33}-ho^{33}-l̩^{11}$ (NZP-pass.by-ANOM), literally ‘the one(s) who passed by’, the term used by the Yimchungrü to refer to the Ao. If this Yimchungrü agentive nominalization is indeed the historical source, then once again it appears to be an exonym that has been adopted by the Ao for self reference. In this case, the inherent deictic orientation of the term used by the Yimchungrü is appropriate. There is an additional piece of evidence that supports this etymology: near the site of the ancestral Yimchungrü village of Langa in Tuensang district is a range that the Yimchungrü call ‘Chungli Hill’, allegedly from the time that the Chungli occupied that site, prior to their migrating further westward to the Mokokchung district. This may well be hotly disputed by the Ao, whose folkloric origin does not extend chronologically beyond their emergence from the earth near Chungliyimti, located further to the west at the edge of the Mokokchung district (see §1.7.2 below).²

The Ao tribe is divided into the Chungli and Mongsen moieties, known as $[tʃuŋ^{11}li^{11}]$ and $[muŋ^{11}sən^{33}]$ respectively. This is the major cleavage in terms of both the social and linguistic organization of the people who identify themselves as Ao. The terms are additionally used by native speakers as linguistic labels for the two major dialects. Three villages situated on the Japukong (Japu, Nancham and Satsekpa) and one on the Changkikong (Changki) speak another Ao dialect known as $[tʃaŋ^{33}ki^{33}]$, which is very close to Mongsen and is mutually intelligible.

An array of alternative names for the numerous languages and peoples of the north-east are to be found in the literature, ranging from pejorative terms used by neighbouring tribes and naïvely adopted by the British, to inconsistent transliterations of the one name (Reid [1942] 1997: 100fn.), to the use of a loconym as an appellative for a whole dialect, e.g. *Khari*, after the village of Khar, previously used for any variety of Mongsen. Some names were coined according to the route that the members of a tribe would take to reach the Assam plain, e.g. *Hattigoria*, so named because Ao traders approached the trading town of Jorhat via Hatigarh, a place where the Disoi River flows out of the hills and formerly the site of an elephant trap (Mackenzie [1884] 1995: 400). These factors have conspired to produce a bewildering number of

2. I am grateful to my Yimchungrü consultant Mr Toshi Wungtung for sharing these insights with me.

exonyms and allonyms for what might actually be the same village or people. Marrison (1967) and Matisoff (1986) are useful for helping to clarify the nomenclatural confusion, and the terms ‘autonym’, ‘exonym’, ‘allonym’ and ‘loconym’ used in this section and elsewhere follow Matisoff’s (1986) usage.

1.7.2. History

Ao folklore holds that the tribe comes from the east, the first ancestors having emerged from the earth at the site of some sacred stones known as *luj tɔɔuk* ‘six stones’ located near the present-day village of Chungliyimti. This is a famous village in the oral history of the Ao and figures prominently in folklore stories. Chungliyimti (lit. ‘big village of the Chungli’) is now a Sangtam village, but its name suggests that it was originally inhabited by the Ao until they were forced to move to their current territory by the pressure of tribes invading from the east.

Mills (1926: 11) observes that the Ao tribe was well past the peak of its power by the time that the British authority began to penetrate the Naga Hills in the mid-nineteenth century. By the 1880s, the Ao were suffering at the hands of marauding Sumis and Lothas who were pushing up from the south and dispossessing them of their lands, and from raids by the Phom and Chang attacking from the east. The village of Mongsenyimti, located on the Lambangkong range just west of the Dikhu River (see Map 2), was reported to have suffered the devastating loss of 148 heads in a single raid by the Chang in 1888, of which 15 were men’s, 30 were women’s and the rest were children’s. The attack on Mongsenyimti was in turn a reprisal for the British raids on the Chang villages Noksan and Litim (see Map 3) in the previous year, in tit-for-tat revenge for earlier Chang attacks on Ao villages. A ‘rabble of some hundreds of friendlies’ had followed the expedition from the Ao country and took the opportunity to indulge in the looting and massacre of their weakened enemy, resulting in a far more severe punishment of those villages than the British had ever intended (Reid 1997: 117–118).

By all accounts (e.g. Mackenzie 1995; Gait [1906] 2003; Shakespear [1929] 1980; Reid 1997) the British were somewhat reluctant to establish a presence in the Naga Hills in the nineteenth century. Annexation was eventually viewed as a cheaper alternative to the costly expeditions known as ‘military promenades’, which were regularly sent into the hills to take revenge for head-hunting raids in the settled districts of Nowgong and Sibsagar, where the commercial interests of the British were located. After an initial period of penetration followed by a change of policy and withdrawal, a new policy of extending political control gradually resulted in the establishment of permanent military outposts, first in the south at Samaguting in 1866–1867 and then later in Kohima in 1878 to

control the Angami, in Wokha in 1882 to subjugate the Lotha, and finally at Mongsenyimti in 1889, accompanied by the incorporation of the whole of the Ao territory into the Naga Hills district in return for protection. A subdivisional headquarters was soon after established nearby at Mokokchung (Reid 1997: 99–124).

The presence of the British eventually put an end to head-hunting and inter-village feuds in the Ao country west of the Dikhu River. Across the frontier in the unadministered territory, however, the practice of head-hunting and feuding went on as it had before, up to and even beyond independence in 1947. Trans-Dikhu tribes were warned not to raid in the annexed territory, and this was enforced by military promenades sent out in the winter to take revenge on villages that resisted the British expansion. Retribution was severe and usually took the form of torching the offending village's houses, grain stores and crops, the confiscation of its livestock, and the jailing of the leaders if they could be captured. Anyone who has spent a winter in the mountains of Nagaland will appreciate the suffering that must have been inflicted by these British incursions.

After independence in 1947, the Naga Hills district was incorporated into the state of Assam, then in 1963 Nagaland was inaugurated as the sixteenth state of the Indian Union. This was ostensibly an attempt by the government of the day to diffuse secessionist rumblings that had been brewing since independence. By this time, the indigenous people of the erstwhile Naga Hills Frontier District and the Tuensang Frontier Division had become highly politicized and were forging a pan-Naga nationalism that sought to incorporate all the disparate tribes of the region to better articulate their demand for sovereignty. Armed conflict, first between the Indian security forces and a united group of Naga nationalists, and then later between splinter groups of Naga nationalists has continued to rip apart the social fabric of Nagaland ever since, but the situation is perhaps more settled now than it has been for decades, demonstrated by the fact that limited tourism is being encouraged in areas of the state previously restricted.

The Ao have fared well in comparison to most other tribes of the area since the coming of the British. An American Baptist missionary and his wife were in the vanguard of the European advance into the Naga Hills (see §1.7.4) and first found acceptance in an Ao village. Christianity and education introduced by the missionaries have subsequently given the Ao a head start over the other minority groups of the state, although the cost has been the suppression and loss of much of their traditional culture. Prior to the patronage of the missionaries, they traditionally benefited financially from their role as middlemen in trade between the plains and the interior (Mackenzie 1995: 400).

1.7.3. Village life

The Ao villager is traditionally a subsistence farmer who supplements his diet with food hunted and gathered in the jungle. Meals are eaten twice a day. A typical meal consists of rice and vegetables, and occasionally meat or fish. The Ao dearly love to eat pork, especially during festivals and other celebrations, but this is observed to be a rare addition to the daily meal in the village and most dishes are vegetarian. Food collected from the streams and forests includes fish, prawns, crabs, jackfruit, mangoes, edible ferns and other greens. Hunting with guns is a favourite pastime of people of the region generally, and this perhaps goes some way towards explaining the relative rarity of birds and animals in the forests, at least in those situated near villages. There is generally little local awareness of conservation, nor a desire by the average villager to preserve what remains. But this may be changing; some village councils are beginning to acknowledge the precarious state of the ecology and have introduced hunting seasons. Deforestation is vastly evident throughout the Mokokchung district and the loss of habitat may also explain the paucity of wildlife near villages. This is mostly a consequence of decades of shifting cultivation, which has resulted in large swathes of land being denuded of verdant jungle. Timber resources in the Mokokchung district have also been heavily exploited, with the result that it is rare to see a tree with a diameter of greater than half a metre.

Streams are fished by a number of methods. One traditional method that can be done by a group of people is known as *juηsəp*. This involves finding a place with a small island that splits the stream into two branches. The smaller of the diverging branches is dammed and the water is quickly bailed out downstream of the dam so that fish, prawns and crabs can be collected in the shallows. The Ao also use a variety of traps and weirs to trap fish and crustaceans.

Every family keeps chickens and some keep one or two pigs. Livestock are mostly kept for breeding and their sale provides a ready source of cash when it is needed. A small market was inaugurated at Mangmetong village during the monsoon of 1999 and provides a place for people to sell the meat of slaughtered pigs, live chickens, and vegetables from their gardens.

Most families have their own plots of land for farming unless they are landless, in which case they will rent a plot of land to grow crops, the rent usually being paid by a proportion of the harvest. The nature of the cultivation is swidden agriculture, known locally by the Indic term *jhum*. The jungle is fired and cleared, and then crops such as rice, maize, yam, sweet potato, chilli and mustard leaf are grown. Because the quality of the soil is relatively poor and fertilizers are not used, a plot of land cleared for swidden agriculture is only productive for two years and then must be left fallow for twelve to fifteen years.

Even landed families will rent other families' fields when they must leave their own land fallow, or if they want to increase the size of their harvest.

Money is little used in the village. Food other than rice is freely shared amongst neighbours, who usually belong to the same clan and live as a sub-community in a separate ward of the village. If a family needs cash, they will sell some of their stored rice. Most families keep a year's supply of rice stored in their granaries. The granaries consist of small huts located together some distance away from the houses, so that in the event of a fire (a common occurrence in the past because of the close proximity of houses in the village's wards) the stored grain is not destroyed in the blaze.

Marriage is traditionally exogamous with respect to one's clan. Potential marriage partners who meet for the first time will often establish their clan membership to ensure that their phratries are compatible (see Mills 1926 for their description). The marriage of couples belonging to incompatible phratries is possible nowadays, but entails serious consequences. The husband is prohibited from ever holding office on the village council, and both husband and wife forfeit the right to speak in public meetings. It usually becomes difficult for such couples to remain living in the village.

Lastly, it is highly noteworthy that the basic divisions of the clan system of the Ao have their closely corresponding equivalents in the Chang, Yimchungrü and Sangtam tribes, and perhaps in all the tribes of Nagaland, although this remains to be confirmed by further research. Marriage outside of one's tribal community is possible, providing that the marriage partner belongs to a correspondingly exogamous clan.

1.7.4. Religion

The Ao are now almost wholly Christian, with the occasional family clinging to the old practices and continuing to make and drink rice beer in some villages. As the vast majority are Baptists, abstention from all forms of alcohol is socially expected and actively enforced in some villages. Mangmetong village, for example, is now completely Baptist and prohibits any drinking of alcohol, even in festivals. Those found breaking the prohibition are liable to be expelled from the village. Of the villages I am familiar with, Longkum still has a few families who make rice beer and practice the ancient religion, and Khensa is also noted for the traditional ways of some of its families, and for the village's observance of traditional festivals and ceremonies. Lycanthropy survives amongst the Ao in some villages, despite the influence of Christianity. A locally famous were-tiger by the name of Tali Meren, an inhabitant of Khensa village, is sometimes consulted by members of the Ao community because of the supernatural powers he is thought to possess.

Christianity was first brought to the Ao country in 1872 by an Assamese evangelist named Godhula, who was encouraged by the American Baptist missionaries Reverend Edward Winter Clark and his wife Mary Mead Clark to try proselytizing amongst the Ao after their mission failed to win many converts amongst the Hindus and Muslims on the Assam plain. A rather rapturous account of the first contact can be found in Clark ([1907] 1978). The American Baptist missionaries eventually succeeded in establishing a bridgehead on the Changkikong Range (see Map 3) in 1876 at Molung Yimchen (also known as Dekha Haimong), literally ‘old village of the Molung’ – the original inhabitants were probably ousted by an earlier Ao invasion. The conversion to Christianity of only some inhabitants created problems in the village of Molung Yimchen, because the converts refused to contribute to the costs of performing traditional ceremonies. This culminated in the converts, together with Clark and his wife, abandoning the village only seven months after the missionaries had arrived. The converts subsequently established a new colony called Molung Yimsen (lit. ‘new village of the Molung’) on a nearby peak. Clark (1978: 19–25) gives a description of the founding of the new colony. Christianity spread fairly rapidly through the Ao villages of the Changkikong range, then more slowly to the villages of the interior ranges. Longkum village, historically one of the most powerful of Ao villages, was the last to convert and probably has the most households that continue to follow the old religion. An account of the traditional belief system of the Ao can be found in Mills (1926).

Much of the traditional ceremony of the original religion remains in the Christianity practiced in Mangmetong village. The pastor, for example, has in many ways an identical role to the ceremonial priest of old and is still called upon to pray for any new venture or inauguration, such as the opening of a new market stall, or the construction of a new house. In pre-Christian times inaugurations were equally important and required sacrifices to be made to appease the spiritual forces. When my wife Pavitra and I first went to Mangmetong village, we asked the village council chairman for permission to teach Pavitra how to drive a jeep on the village football pitch, which was located next to the church. This request was refused on the grounds that it would not be respectful to the church. It was not until I learned that the church is built on the site of a ward’s log drum, the closest thing to a village deity, that I came to make sense of the refusal of this seemingly innocuous request. Before annexation of the Ao territory by the British, human heads taken as war trophies from other villages were often presented to the log drum on the return of the victors, and offerings were made to the log drums of Ao villages to induce the forces of nature to bring rains in time of drought (Mills 1926: 79). The whole area surrounding the modern day church of Mangmetong village is therefore likely to be inhabited by a potentially malevolent spirit, be it Christian or

something else, and hence the village council chairman's reluctance to risk us disturbing it with our driving lessons.

Church services are held every Sunday and are well attended. Sermons and announcements are always given in the Chungli dialect, even in Mongsen villages. Furthermore, all native Mongsen speakers I quizzed uniformly reported that they only ever pray in the Chungli dialect. Christianity has now become an important aspect of the identity of the Ao and other tribes of Nagaland.

1.8. Ao dialects

The Ao recognize two major dialects of their language, viz. Chungli and Mongsen. As mentioned in §1.7.1, a sub-dialect of Mongsen known as Changki is spoken in a village of the same name and in a few other villages located on the Changkikong and Japukong ranges. Changki is sometimes considered a separate dialect, but preliminary investigations suggest a close affinity with Mongsen. The personal pronouns, for example, are identical to those of the Mongsen pronominal system. The distribution of Ao villages according to dialect is set out in Map 3.

Some additional varieties of Ao spoken in Yong-Yacham and Longla villages east of the Dikhu River are recognized as separate dialects in the literature (Mills 1926; Marrison 1967), and are reported to exhibit contact effects from the neighbouring languages of Phom and Chang. Mills (1926: 333) notes that these village varieties bear a strong affinity to Chungli. Their relationship to Chungli is probably somewhat analogous to the relationship of Changki to Mongsen, although a prolonged language contact situation with neighbouring Tibeto-Burman languages may have considerably lessened their similarity. No work has been done on these dialects since their existence was first documented by Mills (1926).

Chungli-Mongsen bidialectalism is an interesting feature of a small number of Ao villages that is yet to be comprehensively investigated and is discussed in Mills (1926: 3) and Coupe (2003a: 49–50). Bidialectal villages are differentiated in Map 3. The usual situation is for one administrative ward (or *khel* in Nagamese) to speak Chungli, and the other Mongsen.

Chungli varieties are believed to demonstrate as much between-village variation as that documented in this work for Mongsen. As noted in Coupe (2003a: 3), the dominance of Chungli as the standard language of the Ao can be attributed to the work of the American Baptist missionaries, who initially lived in Molung Yimchen and its Christian colony Molung Yimsen on the Changkikong range, before moving to Impur to establish a mission on the Asetkong range in 1894. The language of the Molung variety of Chungli was

subsequently spread through the introduction of literacy via the work of the Impur mission, and then by the introduction of the Ao Bible.

All Mongsen speakers appear to have some ability in Chungli, including children, because as the prestige dialect it is taught in schools up to the tenth grade throughout the Mokokchung district, and is also used in church services. Yet the converse is not necessarily found to be the case. A comparison of Text 1, translated into Chungli morpheme by morpheme, demonstrates why a monodialectal Chungli speaker might have trouble understanding Mongsen without sufficient prior exposure to the Mongsen dialect. Even within the one dialect differences can be substantial. One day when working on Mongsen texts with a Longchang Mongsen speaker and a Mebongchukit Mongsen speaker I was astonished to hear them discussing the data in Nagamese. When I asked why they didn't communicate in Mongsen, they replied that their varieties were too different to understand each other easily, and that it was more convenient to use Nagamese! Another consultant reports that the Chungli varieties spoken on the northern end of the Lambangkong range differ significantly from standard Chungli based on the variety spoken in Molung Yimsen, to the extent that they are mutually unintelligible. This report remains to be confirmed.

1.9. Review of literature

A thirty-eight page chapter in the anthropological monograph of Mills (1926) was the first work to be published on the grammar of the Mongsen dialect of Ao. Coupe (1998, 1999, 2003a) describe the phonetics, phonology and tone system of Mongsen, and Coupe (2002) offers a preliminary analysis of the tense and mood system. This is superseded by Coupe (2003c) and the present work. An early Mongsen word list can be found in Grierson (1967: 281–282, 292–327), and a word list by Mills (Typescript n.d.) exists in the SOAS library at the University of London. Mills' data is used in the classification of Marrison (1967).

Published work on the Chungli dialect includes an early grammatical description (Clark [1893] 2002), an early dictionary (Clark [1911] 1990), a phonetic reader (Gowda 1972) and a brief grammatical sketch (Gowda 1975). Other publications that use the Clark materials as sources of data are Grierson (1967), Wolfenden (1929) and Marrison (1967).

1.10. Ao literature

The Ao Bible, written in Chungli, is the most important publication in the language and serves as a standard for the written language. The translation of the Bible began with the efforts of Clark in 1883, culminating in the first edition of the New Testament in 1929, followed by a complete translation of the Old and New Testaments in 1964 (Ao 2001). A thrice-weekly newspaper titled *Ao Milen* (Ao torch) is published in Mokokchung using an orthography based on the Ao Bible.

Attempts were made in the 1990s to revise the Chungli missionary script, with limited success, by linguists from the Central Institute of Indian Languages. Their intentions were well-founded and indeed were justified from an orthographic point of view, but a lack of consultation with the speech community with regard to the advantages of the proposed revisions resulted in a limited acceptance of the reforms.

1.11. Fieldwork and sources of data

The data upon which this description is based was gathered intermittently over a ten-year period. I first chose to work on the Mongsen dialect because it is the less studied of the two major dialects of Ao and I relished the romance of exploring uncharted territory, and secondly, because I knew of a native speaker of Waromung Mongsen who was planning to study in Sydney for two years. He agreed to act as a consultant and I began working with him in 1996 when he arrived in Australia.

At the end of 1996 I undertook a three-month period of fieldwork in the Mokokchung district of Nagaland, concentrating on the Waromung village variety of Mongsen. On my return in 1997 I continued working with the Waromung Mongsen speaker in Sydney. The focus of that research was a phonetic and phonological description of the Waromung village variety of Mongsen, the results of which were published as Coupe (1998) and Coupe (2003a). I returned to Nagaland in 1999 to begin work on the present grammatical description and stayed in north-east India for a year. Initially I lived and worked with consultants in Mangmetong village, and later I worked with Mongsen speakers of various village varieties in the towns of Mokokchung and Dimapur in Nagaland, and in Shillong, Meghalaya. In Mokokchung and Dimapur I had the good fortune to find accommodation in Mongsen-speaking households.

Because of the substantial phonological and morphological differences that can exist between village varieties of the same dialect, a decision was made to base this grammatical description largely on one village variety. This avoided

the initial potential for confusion in the analysis. The variety spoken in Mangmetong village was selected because it has substantially divergent phonological contrasts and morphological forms compared to the Waromung variety I had worked on previously, and because it is the variety spoken by members of my extended family. This facilitated introductions, accommodation and personal security for my stay in Mangmetong village. Limiting my attention initially to just one village variety also minimized the chance of inadvertently identifying variant phonological and morphological forms of different village varieties as allophones or allomorphs of the one system. Together with the previously collected Waromung Mongsen data, this corpus was used to establish a standard against which the data of other Mongsen varieties could then be compared.

I returned to India for a six-month period of fieldwork in 2001. This work was mostly done in Shillong with native speakers of various varieties of Mongsen because of delays in obtaining the necessary restricted area permit required by foreigners to enter Nagaland. During this time an obliging Mangmetong Mongsen speaker came to live with me to teach me his language, and I supplemented this work by employing speakers of Mongsen who were studying in Shillong as additional consultants. Further visits of four to six months' duration were made in 2004, 2005 and 2006 to work on other languages of Nagaland, and I took the opportunity on these field trips to iron out remaining wrinkles in the grammatical analysis whenever I had access to Ao consultants. Lastly, I was fortunate to have the intermittent help of an Ao speaker trilingual in Chungli, Mangmetong Mongsen and Changki (his mother is from Mangmetong village and his father is from Changki village) residing in Brisbane, Australia.

The Mangmetong variety of Mongsen thus forms the bulk of the data presented in this grammar. However, on all fieldtrips I have made a point of visiting other Mongsen villages to collect samples of texts whenever opportunities presented, and as noted above, I have often worked with consultants who speak different Mongsen varieties – for example, those of the villages of Chungtia, Khar, Khensa, Longkum, Longchang, Mekhuli and Mebongchukit – in order to obtain an appreciation of the extent of linguistic diversity that can occur between varieties of the one dialect. Some examples in the text come from Mongsen varieties other than Mangmetong Mongsen. Where this is the case I draw attention to the origin of the example. The reader should assume by default that the examples are representative of Mangmetong Mongsen unless informed to the contrary.

The corpus comprises a 200-item Swadesh list of elicited words in Waromung Mongsen (published as an appendix in Coupe 2003a) and a ca. 800-item word list in Mangmetong Mongsen that is based on the words of the comparative vocabulary of Benedict (1972). The lexical items of the latter were

elicited in a substitution frame so that their tones could be identified; immediately after the item was uttered in the frame by one speaker, I consecutively elicited tokens of each item in isolation from three to six other speakers to reveal any variation.

The grammatical analysis relies on forty folklore narratives, personal histories and procedural texts in the Khar, Khensa, Mangmetong and Waromung varieties of Mongsen, and judicious elicitation based on those texts. The corpus is supplemented by data that I collected from spontaneous conversation, things I overheard, and native speakers' spontaneous corrections of my own attempts to speak Mongsen. I have purposely chosen to represent the language with actual spoken data from the above sources, wherever possible, in preference to using contrived elicited sentences.

A sample of four texts is provided in the final chapter, including a morpheme-by-morpheme comparison of a Mongsen folklore story translated into Chungli. As noted above, this demonstrates why monodialectal speakers of Chungli may have trouble understanding Mongsen, despite the grammars of the two dialects being more or less identical. Substantial differences in morphological form can result in substantial difficulties in comprehension.

1.12. Equipment

The software program 'The Linguist's Shoebox' was used to create a database of interlinearized texts and a lexicon. The Khar, Khensa and Waromung data were recorded in the field using a Sony TCM 5000EV analogue cassette tape recorder and a Sony ECM-FO1 microphone. Some of the Waromung data was previously recorded in the Phonetics Laboratory of the Department of Linguistics at the Australian National University. Most of the Mangmetong data was recorded using a Sony MiniDisc MZ-R50 and a Sony ECM-ZS90 condenser microphone.

The Kay Computerized Speech Laboratory (Kay CSL) instrumentation was used to generate spectrograms and waveforms for the acoustic analysis of the voiceless sonorants in §2.2.1.5–§2.2.1.6, and Speech Analyzer was used to extract pitch contours to demonstrate the impact of intonation on the lexical tone system in §3.4.

1.13. Orthography

The orthography used in this description is based on my phonemic analysis of Mongsen and uses the standard symbols of the 1996 revision of the International Phonetic Association with minor modifications, as described in

Chapter 2. The standard practice of linguistic notation is followed: phonetic transcriptions are given in square brackets, and tone letters (after Chao 1930) indicate pitch values, e.g. [a³³ki³³] ‘house’; phonemic transcriptions are given in italic face and diacritics are used to indicate high and low tones, with the mid tone being unmarked e.g. *aki* ‘house’. The transcription of intonation serving a clause boundary marking function is described in §3.4.

Tone sandhi is extensive in Mongsen, with output tones often differing considerably from underlying tones. In many cases the tone sandhi processes cannot be accounted for exhaustively by rule (§3.3). To demonstrate the extent of tone sandhi, the first line of language examples is presented in italic face with output tones and morphophonological changes demonstrated. The second line, in plain face, gives a morpheme-by-morpheme segmentation showing underlying tones and morphological forms. The third line gives an interlinear gloss identifying lexical items and grammatical categories, and the fourth line gives a free translation in English. Lastly, direct speech in the first line of the phonemic transcription and in the free translation is differentiated from indirect speech, if present, by double quotation marks. The following example demonstrates this presentation of data:

- (1.1) *pùkphula nàŋ mà:ulaw?*
 pùkphu-la nàŋ mə-a:u-la-ù?
 owl-F 2SG NEG-be.good-NEG.PST-DEC
 ‘Owl – you’re no good!’

Capitalization is not used in the initial words of the language examples because the orthography is close to a phonemic transcription and there are no corresponding upper case symbols for some segments, such as the schwa.

1.14. Theoretical orientation

It is hoped that this work might serve as a reference grammar for the study of Tibeto-Burman historical-comparative linguistics, and more generally for the study of linguistic typology; consequently, an overriding consideration has been to ensure that its contents are intelligible to readers of all theoretical persuasions. To achieve this objective, the description has been couched in a broadly functionalist-typological framework, as exemplified by the theoretical and practical approaches adopted in Shopen (1985) and in Newman and Ratliff (2001), with the desire that the results may be equally accessible to both formally- and functionally-oriented linguists. The only exceptions to this are briefly employed in Chapter 3, where an analysis in the framework of autosegmental phonology (Goldsmith 1990) is conducive to accounting for the

effects of grammatical tone spreading (§3.1.3) and the phonological behaviour of the glottal stop (§3.5).

Throughout this description I make continual reference to the many grammaticalization processes at work in Mongsén, and I also make a point of sharing historical-comparative insights where relevant to the analysis of forms. The language is remarkably transparent with respect to the lexical origin of numerous grammatical morphemes. In view of this, it is useful to point out the diachronic sources of these morphemes where they are recognizable, and to relate the emergence of grammar to what is understood about the typology of grammaticalization in the languages of the world.

Chapter 2

Phonology and phonological processes

This chapter describes the segmental phonology and phonological processes of Mongsen, with the focus falling on the variety of this dialect spoken in Mangmetong village of Mokokchung district. The phonological inventories of other Mongsen varieties will then be considered, with a view to establishing the extent of known phonological variation that is to be found within the one dialect. Lastly, criteria for word and clitic are examined, followed by a description of phonological processes applying in word formation.

2.1. Syllable structure

The Mangmetong Mongsen syllable consists of an obligatory vowel and a tone and up to three optional elements having the following linear structure:

$$(2.1) \quad \sigma = (C_1) V (G) (C_2) T$$

Table 2.1. Phonotactic distribution of syllable constituents

(C ₁)	V	(G)	(C ₂)	(T) ¹
p p ^h t t ^h k k ^h				
ts ts ^h tʃ tʃ ^h	i u	w j	p t k	H(igh)
m n ŋ	ə		m n ŋ	M(id)
ᵐᵐ ᵐᵐ ᵐᵐ	a (a)		ɹ	L(ow)
z s h				
w l ɹ j				
ʌ ɿ ɿ				

Allowable constituents of each slot in the syllable are listed in Table 2.1 above. It is relevant to mention at this point that a prosodic glottal stop frequently occurs in words, but is not considered a member of the consonant phoneme inventory.

1. The /High/ toneme is marked in phonemic transcriptions by an acute accent, the /Mid/ toneme is unmarked, and /Low/ toneme is signalled by a grave accent. Pitch is indicated by superscript tone letters 1-5 in phonetic transcriptions, with σ⁵⁵ being the highest and σ¹¹ the lowest pitch level.

The Mangmetong Mongsen syllable can alternatively be represented metrically as having the hierarchical structure demonstrated by Figure 2.1.

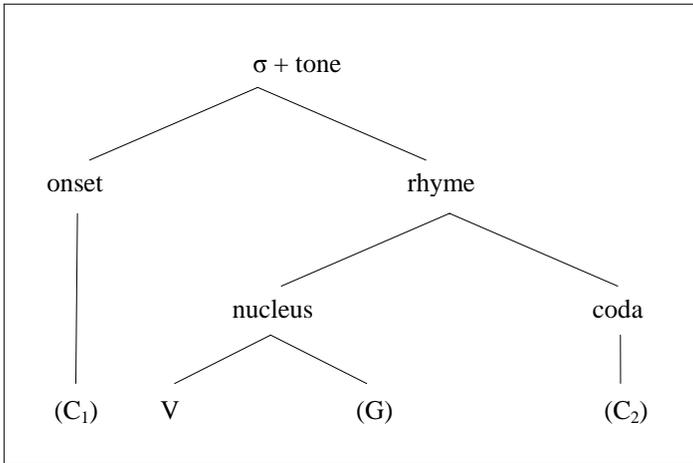


Figure 2.1. Metrical structure of the Mongsen syllable

There are no restrictions on the occurrence of consonants in the onset slot (C_1). Glides that function as syllable onsets behave no differently to consonant segments, thus need not be differentiated when occurring in this position. The nucleus consists of an obligatory syllabic element V and a post-vocalic glide (G) when it is maximally specified. As is typical of all languages of the mainland South-East Asian region, the coda (C_2) is much more constrained than the onset in terms of which phonemes may occur in this position – it only permits the unaspirated stops *p*, *t*, *k*, the voiced nasals *m*, *n*, *ŋ* and the voiced retroflex approximant *ɻ* to occur under its node. A different approach could possibly treat the voiced retroflex approximant as a glide instead of a consonant, seeing that it shares a similar distribution to *w* and *j*. However, the speech of Mangmetong Mongsen speakers is characterized by a tendency to resyllabify a word-final retroflex approximant as an onset of a separate syllable with a schwa nucleus (see §2.6.6 for exemplification), a process that never applies to the semi-vowels *w* and *j* when they occur word-finally. This evidence suggests that the retroflex approximant should rightfully be analyzed as a consonant and not as a fully-fledged member of the glide class, despite its distributional similarities.

It was mentioned above that a glottal stop *ʔ* is analyzed as a word prosody, hence its absence from the phoneme inventory of Table 2.1. An alternative treatment could identify it as a contrastive segment that is limited to occurring under the coda node, but this would need to explain why it is the only consonant

phoneme with a realization restricted to the word-final position, and more importantly, why it is the only one to be deleted before a syllable boundary formed by a concatenated morpheme. Arguments in favour of a word prosody analysis are presented in Coupe (2003a: 24–27) and reviewed in §3.5 with respect to Mangmetong Mongsen.

The analysis of the Mangmetong Mongsen syllable canon differs from that of the Waromung Mongsen syllable presented in Coupe (2003a: 9), principally because the latter variety has an extra vowel $\#$ in its phoneme inventory. This is found mostly to correspond to a high front vowel i in Mangmetong Mongsen (§2.3.2). The high central rounded vowel of Waromung Mongsen can occur in sequence with other non-high central vowels, where it forms the non-syllabic element of a phonetic diphthong, e.g. $a\text{-}\#$ ‘NRL-axe’ and $sə\text{-}\#$ ‘die-IRR’, realized as [a $\#$ ³³] and [sə $\#$ ³³] respectively. Unlike the glides w and j , a non-syllabic counterpart of $\#$ cannot occur in the onset position (C_1); this forces a phonotactic analysis to identify $\#$ as a vowel instead of a glide. The Waromung Mongsen syllable canon must therefore be stated as $(C_1)V_1(V_2)(C_2)T$.

Such a problem for phonotactic analysis is not presented by the structure of the Mangmetong Mongsen syllable. As noted above, w and j freely fill both the onset (C_1) and postvocalic position (G) in the nucleus, allowing them to be recognized as *bona fide* members of a class of glide. This is justified by their behaviour in word formation processes involving resyllabification; when a non-identical vowel is concatenated with a word-final i or u , this always results in the latter’s realization as the offglides j or w respectively. Word formation processes involving glides are described in §2.6.6, and in the following section.

2.1.1. Vowel Sequences

Long vowels do not occur in the data, therefore sequences of identical vowels cannot be assumed to occur under a branching nucleus. If identical vowels are juxtaposed as the result of morpheme concatenation, then phonological processes apply either to delete one of the vowels, or else the sequence is resyllabified, with each vowel forming the nucleus of an independent syllable (see §2.4.2).

Sequences of vowels differing in quality can form tautosyllabic phonetic diphthongs if the second vowel of the V_1V_2 sequence is a high vowel. V_2 is then resyllabified as the non-syllabic element, resulting in a falling diphthong. This is attributable to a language-specific obligatory contour principle applying to sonority conditions in the syllable (Coupe 2003a: 9). Diphthongization processes are exemplified by morpheme concatenations in which a suffix in the form of a high vowel occurs at the morpheme boundary, as in (2.2). The past tense is the unmarked category in Mongsen (§8.5.12.1) and is used here and

elsewhere as the citation form. Translations of bare verbs or those marked only by the declarative mood clitic *-uʔ* thus reflect their inherent past tense specification.

(2.2)	<i>mi</i>	‘swallow’	+	<i>-ùʔ</i>	DEC	>	<i>miwʔ</i>	‘swallowed’
	<i>sisə</i>	‘arise’	+	<i>-ì</i>	IRR	>	<i>sisaj</i>	‘will arise’
	<i>azə</i>	‘challenge’	+	<i>-ùʔ</i>	DEC	>	<i>azəwʔ</i>	‘challenged’
	<i>sə</i>	‘die’	+	<i>-ì</i>	IRR	>	<i>səj</i>	‘will die’
	<i>zəlu</i>	‘write’	+	<i>-ì</i>	IRR	>	<i>zəluj</i>	‘will write’

There is little motivation for treating vowel-glide sequences such as *iw*, *aj*, *əw*, *əj* and *uj* as phonemic diphthongs, due to their lack of syntagmatic cohesion and demonstrated substitutability. Nor is any appreciable difference in rhyme duration found between rhymes consisting of single vowels, versus those consisting of vowel-glide sequences. Since any vowel can occur with a glide to form a phonetic diphthong, treating such VG combinations as separate phonemes would serve only to increase the size of the vowel phoneme inventory. A description of phonological processes involving other vocalic sequences is presented in §2.4.2.

2.1.2. Absence of phonemic stress

A noteworthy characteristic of Mongsen words is their lack of phonemic stress. All syllables of polysyllabic words tend to be equally weighted with respect to the possible perceptual correlates of stress, such as duration, intensity and loudness. Pitch, a fourth possible perceptual correlate of stress commonly utilized by the world’s languages, is monopolized by the lexical tone and intonation systems of Mongsen, thus cannot be exploited for the purpose of giving prominence to a particular syllable in a polysyllabic word.

2.1.3. Tone

Mongsen has a three-level lexical tone system contrasting high, mid and low tonemes (Coupe 1998, Coupe 2003a: 87–110) and this has been observed in all varieties of the Mongsen dialect investigated thus far. All lexical roots are underlyingly specified for tone. They are then subject to variable allotonic realizations, depending upon the tone sandhi environments in which they occur.

It appears that most grammatical morphemes are also underlyingly specified for tone. But because a grammatical morpheme cannot occur in isolation (unlike a verb root in its unmarked past tense form, for example), we can only surmise

what its underlying tone might be by comparing a range of output tones realized on that morpheme when it occurs in different tone sandhi environments created by other concatenated morphemes (e.g. see §3.3.2). In words spoken in isolation, the lexical tones are perceived to be consistently level. This is demonstrated by an acoustic analysis in Coupe (1998; 2003a: 102–110) that plots each toneme's fundamental frequency contour as a function of absolute mean duration.

Interacting with the lexical tone system realized on syllables is a postlexical intonation system that delineates chunks of data in higher domains – such as noun phrases, dependent clauses, and sometimes even independent clauses in extended discourse that we might identify as ‘the discourse paragraph’ for want of a better term. Intonation patterns are superimposed upon the lexical tones of individual syllables in words. This can result in a substantial distortion of the word's underlying level tones, often resulting in output contour tones that function as boundary signals.

The interesting manifestations of tone are addressed in §3.3. Characteristic patterns of postlexical intonation associated with types of dependent clauses and other syntactic units are described in §3.4.

2.2. Segmental phonology

Varieties of Mongsen vary according to the size of their consonant phoneme and vowel phoneme inventories. Mongsen varieties spoken in the south of Mokokchung district have an additional voiceless sonorant series that occurs in contrastive distribution with the ubiquitous voiced sonorant series (described in §2.2.1.5 and §2.2.1.6). Varieties spoken in a few villages of the Changkikong range have an additional vowel phoneme (see §2.3.2), and one of those village varieties (Khar Mongsen) is also found to have the voiceless sonorant series that is common to southern Mongsen varieties.

The following description is limited to the phoneme inventory of Mangmetong Mongsen. Segmental differences observed between other Mongsen varieties will then be discussed in §2.3.

2.2.1. Consonant phonemes

There are twenty-seven consonant phonemes in Mangmetong Mongsen, representing five distinctive places of articulation² and five distinctive manners

2. This ignores a glottal place of articulation, since *h* is rightfully considered a segment that is unspecified for place of articulation (see §2.2.1 for justification).

of articulation. The consonant phonemes are listed in Table 2.2 below according to place and manner of articulation. Phonemic symbols are presented in plain face; orthographic symbols used in the text and in language examples are given in *italic face*, in parentheses.

For reasons mentioned in the description of syllable structure in §2.1, the glottal stop is not recognized as a segmental phoneme, therefore it is not represented in the following table of phonemes (see §3.5 for justification).

Table 2.2. Consonant phonemes of Mangmetong Mongsen

	BILABIAL	DENTAL	POST- ALVEOLAR	PALATAL/ PAL-ALV.	VELAR	GLOTTAL
Plosive						
unaspirated	p (<i>p</i>)	t (<i>t</i>)			k (<i>k</i>)	
aspirated	p ^h (<i>ph</i>)	t ^h (<i>th</i>)			k ^h (<i>kh</i>)	
Affricate						
unaspirated		ts (<i>ts</i>)		tʃ (<i>tʃ</i>)		
aspirated		ts ^h (<i>tsh</i>)		tʃ ^h (<i>tʃh</i>)		
Fricative						
voiced		z (<i>z</i>)				
voiceless		s (<i>s</i>)				h (<i>h</i>)
Nasal						
voiced	m (<i>m</i>)	n (<i>n</i>)			ŋ (<i>ŋ</i>)	
voiceless	m̥ (<i>hm</i>)	n̥ (<i>hn</i>)			ŋ̥ (<i>hŋ</i>)	
Lateral						
voiced		l (<i>l</i>)				
voiceless		l̥ (<i>hl</i>)				
Approximant						
voiced	w (<i>w</i>)		ɹ (<i>ɹ</i>)	j (<i>j</i>)		
voiceless	w̥ (<i>hw</i>)		ɹ̥ (<i>hɹ</i>)	j̥ (<i>hj</i>)		

2.2.1.1. *Voice onset time contrasts*

All varieties of the Mongsen dialect exhibit a two-way contrast in voice onset time for their stop and affricate series. Coupe (2003a: 64–73) presents the results of an instrumental study of voice onset time in Waromung Mongsen and

demonstrates that the phonemic contrast is located between voice onset time lag and coincident voice onset time for syllable onsets. In other words, voiceless aspirated stops and affricates contrast with their corresponding voiceless unaspirated counterparts.

A phonological characteristic that allows for the differentiation of two sub-types of Mongsen is the presence or absence of a voiceless sonorant series. It has not yet been possible to carry out a complete survey of all Mongsen varieties, but it appears that this could contribute a useful criterion for establishing sub-groupings of varieties within the dialect. Spectrographic evidence of voice onset time contrasts in the nasal and approximant series is presented in §2.2.1.5 and §2.2.1.6 respectively.

The village varieties of Mongsen thus far confirmed to have a voiceless sonorant series are Chungtia, Khensa, Mekhuli, Mangmetong and Khar. With the exception of Khar, which is situated on the Changkikong range, all these villages are located on a range known locally as the Ongpangkong. The Ongpangkong range is not marked in Map 3, but is physically located in the south of the Mokochung district and runs off the Langpangkong range in a roughly westerly direction. The locations of the above-mentioned villages help to mentally plot its unmarked position on Map 3. Varieties hitherto confirmed to lack a voiceless sonorant series are Waromung, situated on the Changkikong range, and Longchang and Sangratsü, situated on the Asetkong range. Given this distribution, we might assume from the preliminary findings that the languages of the Ongpangkong range tend to have voiceless sonorants, those of the Asetkong range tend not to have voiceless sonorants, and Mongsen varieties both possessing and lacking a voiceless sonorant series are to be found on the Changkikong range.

The classification of “Naga” languages by Marrison included the phonological criteria of voice onset time contrasts in initial plosives and the presence of voiceless liquids and nasals to establish various sub-groups within the “Naga” languages (1967, I: 25). With respect to Mongsen, the accuracy of this may need to be reassessed in the light of findings of the present work, for the following reasons.

Firstly, the historical sources of data Marrison used to establish the Mongsen phoneme inventory did not contain examples of voiceless sonorants. This was because he only had at his disposal a grammatical sketch (Mills 1926) and word list (Mills n.d) of Longchang Mongsen, a variety that lacks a voiceless sonorant series. As this was the only reliable Mongsen data available in the literature when Marrison made his survey, it was not then known that varieties of Mongsen with a voiceless sonorant series also existed. Secondly, Marrison had done fieldwork on Chungli with a consultant in Assam and concluded that the Chungli dialect did not have a phonemic contrast between a voiceless aspirated and voiceless unaspirated series of stops. This led him to wrongly assume that

the Mongsen transcriptions of Mills, which did suggest two contrastive stop series, were non-phonemic variants that had been inconsistently transcribed (Marrison 1967, I: 9). Marrison subsequently adjusted the Longchang Mongsen transcriptions to fit with his phonological interpretations of Chungli phonology. Thus, Marrison's 'Ao Group' of Type 3C is chiefly characterized on page 26 as not distinguishing phonemic contrasts between $k \sim kh$, $t \sim th$ and $p \sim ph$.

Despite initially claiming that the Chungli initial stops lack a contrast in voice onset time, Marrison then contradicts his earlier assertion by listing both voiced and voiceless initial stop phonemes in a table of Ao phonemes (1967, I: 58). Presumably this was an oversight that escaped his attention, perhaps influenced by the orthographic sources that he relied upon. The table of phonemes is also at odds with an auditory analysis of stops presented by Gowda (1975: 10–11), who similarly finds that Chungli has no phonemic voice onset time contrasts in the stop series.

2.2.1.2. *Stops*

The stop series occurs at three distinctive places of articulation: bilabial, dental and velar. An aspiration contrast – voiceless unaspirated versus voiceless aspirated – occurs in initial and in medial position at each place of articulation.

Examples of voiceless unaspirated and voiceless aspirated bilabials are *pi* 'this', *phiju* (announce.PST) 'announced', *tə-paʔ* (RL-father) 'father' and *tə-pha* (RL-tooth) 'tooth'. Only the voiceless unaspirated bilabial stop may occur in the syllable-final position, e.g. *tə-əp* (NZIP-be.rotten) 'rotten'. A voiceless bilabial fricative [ɸ] occurs as an allophone in free variation with [p^h] in the speech of some speakers, e.g. *ùphəla* [u¹¹ɸə³³la³³] 'red junglefowl' and *pùkphu* [puk¹¹ɸu³³] 'owl'.

The passive articulator for the apical laminal dental stops *t*, *th* (as well as for the sonorants *l*, *hl*, *n* and *hn*) is ideally described as denti-alveolar. In the articulation of these sounds, the apex of the tongue makes light contact at the point where the upper teeth insert into the gums, in addition to the laminal aspect of the tongue making considerable contact with the area of the hard palate immediately anterior to the alveolar ridge. Voiceless unaspirated and voiceless aspirated dental stops occur in initial and in medial position, e.g. *tən* (sing.PST) 'sang', *thən* (sow.PST) 'sowed', *ata* (await.PST) 'awaited' and *athà* (fall.down.PST) 'fell down'. As noted above for bilabials, only the unaspirated dental stop may occur in syllable-final position, e.g. *a-nət* (NRL-two) 'two'.

Voiceless unaspirated and voiceless aspirated dorsal velar stops occur in initial and medial position, e.g. *kəp* (strike.PST) 'struck', *khəp* (depart.PST) 'departed', *a-ku* (VOC-uncle) 'Uncle!' and *a-khu* (NRL-tiger) 'tiger', with the coda position restricted to the just the unaspirated velar stop, e.g. *mà:ùk* 'cup'.

2.2.1.3. Affricates

The affricates are analyzed as forming a separate series from the stops, because: (i) unlike the stop series, affricates do not occur in the coda; (ii) unlike the stops, not all of the places of articulation of the affricates are matched by the nasal series; and (iii) the apical laminal dental affricate and the apical laminal dental stop occur at the same place of articulation. A characteristic shared by the two series is a two-way contrast in voice onset time, but this alone is insufficient evidence for assuming that the stops and affricates should be conflated into a single series.

Voiceless unaspirated and voiceless aspirated laminal dental affricates occur in initial and medial position, e.g. *tsəŋ* (bark.PST) ‘barked’, *tshəŋ* (wear.PST) ‘wore’, *atsə* (look.PST) ‘looked’ and *a-tshə* (NRL-mithun) ‘mithun’ (*Bos frontalis*). An interesting feature of the dental affricate is that it is constrained to occurring in the environment before the schwa. The drastic loss of other vocalic environments in which it may occur could be the harbinger of a phonemic merger with the palato-alveolar affricate, whose environments of realization are unrestricted synchronically.

Voiceless unaspirated and voiceless aspirated laminal palato-alveolar affricates occur in initial and medial position, e.g. *tʃa* (call.PST) ‘called’, *tʃhà* (pick up.PST) ‘picked up’, *a-tʃu* (NRL-DIST) ‘that’ and *a-tʃhu* (NRL-song) ‘song’.

2.2.1.4. Fricatives

The fricative series demonstrates two distinctive places of articulation, viz. dental and glottal. The voiced and voiceless laminal dental fricative phonemes *s* and *z* demonstrate symmetry in the phonetic realizations of their allophones. Both undergo palatalization before a high front vowel, where they are realized as the domed post-alveolars [ʃ] and [ʒ] respectively; elsewhere they are realized as laminal dental fricatives [s] and [z]. The voiced laminal dental fricative and voiced post-alveolar fricative allophones occur in free variation in the environment before a high front vowel for some speakers, while others neutralize the [z]~[ʒ] realizations in favour of the laminal dental fricative in all phonetic environments. The voiced and voiceless laminal dental fricatives occur in initial and medial position, e.g. *za* (enter.PST) ‘entered’, *sa* (say.PST) ‘said’, *məzəm* ‘poison’ and *məsəm* (clean.PST) ‘cleaned’.

The term ‘glottal fricative’ is a somewhat descriptively misleading historical label. It is used in this work to describe a voiceless segment that is unspecified for place of articulation and can fill the onset slot of the syllable. The voiceless glottal fricative occurs in words relatively infrequently compared to the other

consonant phonemes. It is found in both initial and medial position, e.g. *hàw?* ‘yes’, *hija* (level+come.PST) ‘came at the same level’ and *a-hi?* (NRL-rat) ‘rat’.

Some varieties of Mongsen appear to have lost an initial glottal fricative from some words, e.g. Mangmetong *həntsə* ‘egg’, *həmət* (hold.PST) ‘held’ and *hùŋ* (see.PST; obtain.PST) ‘saw, obtained’ are pronounced as [ən³³tsə³³], [ə¹¹mət³³] and [uŋ¹¹] respectively in Waromung Mongsen.

2.2.1.5. *Nasals*

Voiced and voiceless nasals correspond to the stops for place of articulation: bilabial, dental and velar. The voiced nasals occur in all positions of the word, e.g. bilabial: *mən* (sit.PST) ‘sat’, *a-mi?* (NRL-person) ‘person’ and *tə-kəm* (NZP-be alive) ‘(a)live’; laminal dental: *nì* (1SG) ‘I’, *mənaŋ* ‘before’ and *ulən* ‘woodpecker’; dorsal velar: *ŋam* (nod.PST) ‘nodded’, *phaŋa* ‘five’ and *tsəŋsaŋ* (dance.PST) ‘danced’. The velar nasal occurs as a word-initial syllable onset infrequently and is more likely to be found as a coda.

The voiceless nasal series only occurs in initial and medial position as a syllable onset, e.g. bilabial: *hməntfa?* ‘aspidistra sp.’ and *a-hmu* (NRL-eagle) ‘eagle’; dental: *hni* (follow.PST) ‘followed’ and *məhnəm* (smell.PST) ‘smelled’; velar: *hŋàpà?* ‘gills’ and *á-hŋá?* (NRL-fish) ‘fish’. They can be phonetically realized with a dominant breathy voice component throughout the syllable onset, or with a gradual transition from voicelessness to voicing via a breathy stage. These interpretations are supported by spectrographic evidence presented below. Voiceless nasals occur much less frequently in the data than the voiced series. This is partly due to the distributional inequality between the two series. But even if we exclude nasals occurring in the coda position from our calculations, then it is found that in a sample word list of approximately 800 lexical items there are 217 occurrences of voiced nasal onsets, as opposed to just 32 occurrences of voiceless nasal onsets. Voiced nasals thus overwhelmingly outnumber voiceless nasals by almost seven to one in syllable onsets.

The following pairs of spectrograms compare the acoustic characteristics of voiced and voiceless nasals uttered in isolation by a female speaker (AA) in her forties. The data was recorded in the sitting room of the speaker’s house using a portable Sony MZ-R50 MiniDisc and a Sony ECM-ZS90 condenser microphone. A MiniDisc is not considered ideal for acoustic analysis because it uses a digital compression technology called Adaptive TRansform Acoustic Coding (ATRAC) that extracts and records only those frequencies audible to the human ear (Sony MZ-R50 Portable MiniDisc Recorder operating instructions, 1997: 36). However, the quality of the following spectrograms suggests that this recording equipment is adequate for the purpose of demonstrating voice onset

time (hereafter VOT). The data was re-digitized at 16Khz and wide band spectrograms of 313Hz were generated using the Kay Computerized Speech Laboratory (Kay CSL). VOT was calculated using the onset of glottal phonation as the parameter of comparison. It is difficult to be precise when calculating the VOT of Mongsen voiceless sonorants because, unlike the oral stops, there is often a subtle transition from one articulatory gesture to another, and these transitions are not always amenable to precision in acoustic measurement. Nevertheless, there are striking acoustic differences between the voiced and voiceless sonorants of Mangmetong Mongsen, and these differences are amply demonstrated by the following spectrograms.

Figures 2.2–2.3 below compare AA’s voiced and voiceless bilabial nasals in initial position in the words *mama* ‘breast’ and *hmapan* ‘time’ respectively. In the spectrogram of Figure 2.2, a nasal formant consistent with modal voice phonation is observed at approximately 300Hz. Voicing commences at the point marked (a) on the time axis and continues uninterrupted into the vowel at (b). Velic closure and oral release of the bilabial constriction is marked at this point by the transition to the vowel formants and a concomitant increase in amplitude observable in the waveform.

This contrasts with the very different set of acoustic characteristics of the voiceless bilabial nasal in Figure 2.3. The onset begins as acoustic noise, with concentrations around 1700Hz and 4000Hz correlating perceptually to voiceless aspiration. This is consistent with nasal airflow and lasts for approximately 10ms duration, marked by (a) on the time axis. Note that the initial noise observable in the higher frequencies of the spectrogram is not accompanied by any evidence of modal voice phonation at this point in time. Next, (b) marks the start of a transition from a complete absence of modal voice phonation to an incremental increase; this together with the high frequency noise is perceived as breathy voice and strengthens in intensity over the next 10ms. At (c) there is a transition to what is interpreted to represent the start of velic closure, which is completed at (d) and is accompanied by a corresponding increase in amplitude in the waveform. This point also marks the end of the breathiness, represented by the preceding presence of high frequency noise and the transition to the vowel formants. There is a momentary irregularity in the peak-to-peak distances of the waveform corresponding to the following vowel, but this is assumed to have no bearing on the acoustic analysis.

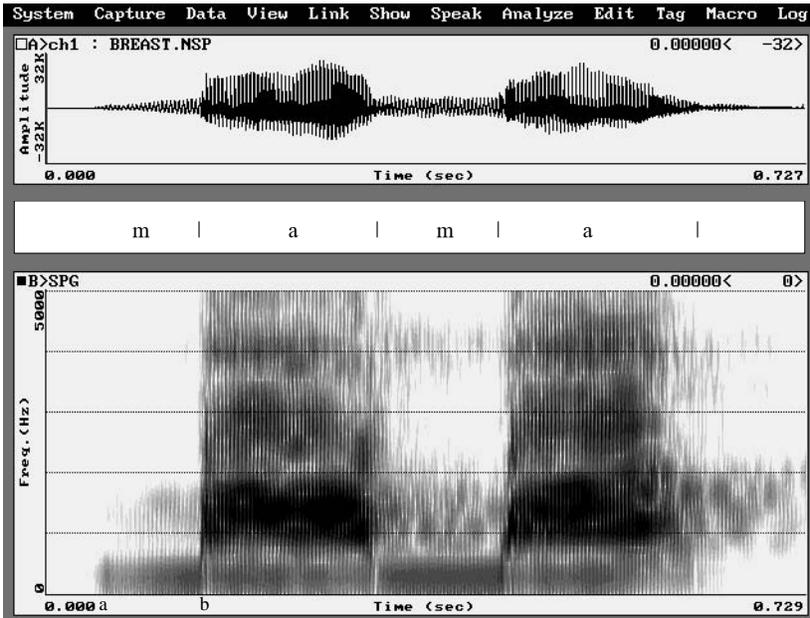


Figure 2.2. Waveform and wide band spectrogram (313 Hz) of *mama* ‘breast’

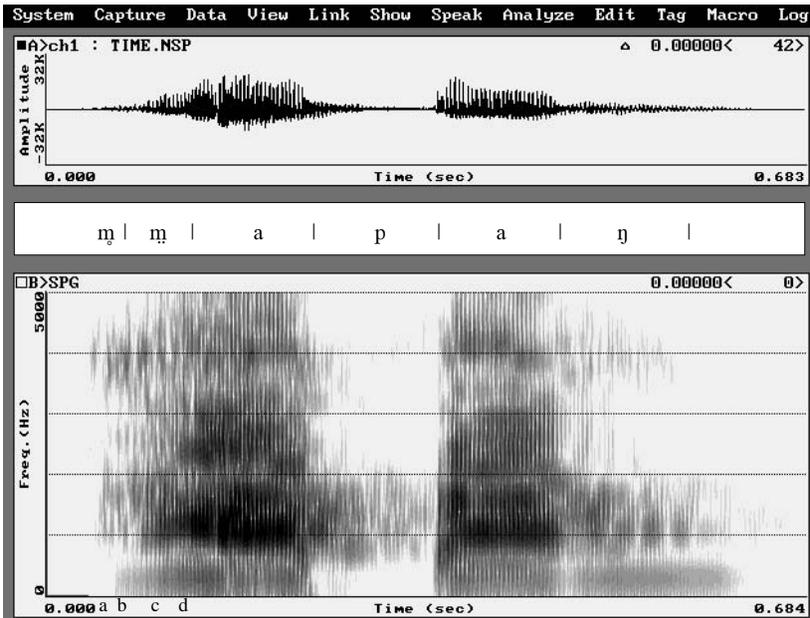


Figure 2.3. Waveform and wide band spectrogram (313 Hz) of *hmapaŋ* ‘time’

The production of the voiceless bilabial nasal is thus characterized by a progression from voicelessness (a), to breathy voice (b), to increasing modal phonation and decreasing breathy voice at the time that velic closure begins (c), and finally to full modal voice phonation at the completion of velic closure and transition to the vowel formants (d).

The next pair of figures compares AA's voiced and voiceless dental nasals in initial position in the words *nì* 'I' and *hni* 'followed' (Figures 2.4 and 2.5 respectively). The voiced nasal of Figure 2.4 is relatively unremarkable, demonstrating a waveform and spectral characteristics consistent with a voiced dental nasal produced with modal voice phonation. There is a clearly delineated transition to the vowel formants as the velum closes and the dental stricture is released. This point in time correlates with an increase in amplitude seen in the waveform.

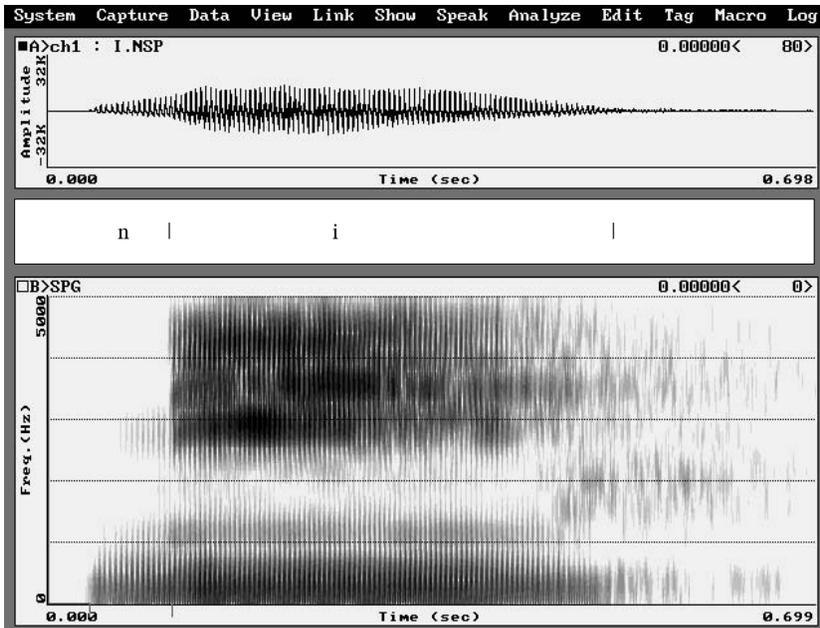


Figure 2.4. Waveform and wide band spectrogram (313Hz) of *nì* 'I'

In Figure 2.5 below, the onset of the voiceless dental nasal is initially voiced for approximately 22ms. Time (a) marks a transition to breathy voice phonation, demonstrated acoustically by the simultaneous occurrence of high frequency noise above 2000Hz, and the presence of regular glottal pulsations indicative of modal voice phonation. This lasts for approximately 65ms. At time (b), velic closure occurs and the oral stricture is released, signalled in the spectrogram by

an increase in amplitude in the waveform and a transition to the vowel formants.

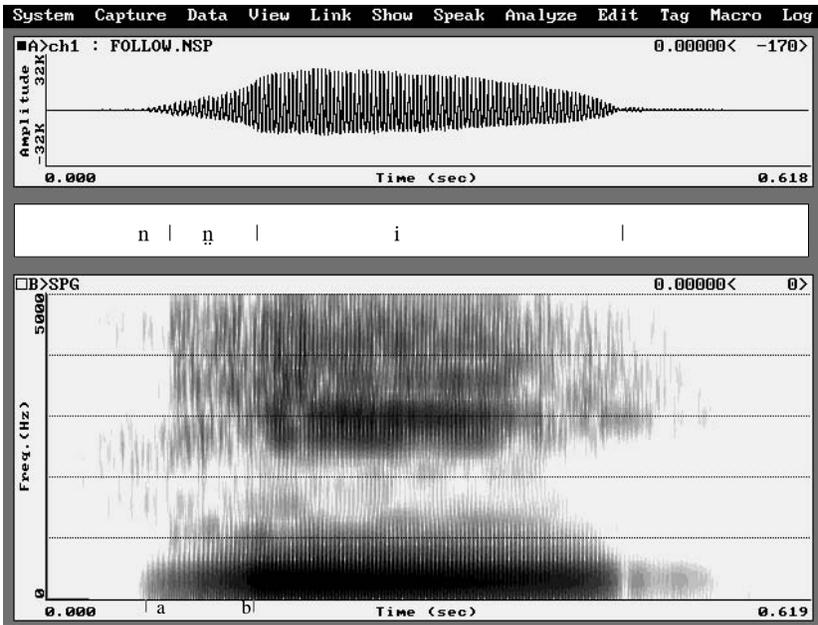


Figure 2.5. Waveform and wide band spectrogram (313Hz) of *hni* ‘followed’

The third pair of spectrograms compares AA’s voiced and voiceless velar nasals in the words *ɲaŋsi* ‘cursed’ and *hɲa* ‘listened’ respectively in Figures 2.6 and 2.7 below. The voiced velar of Figure 2.6 has a very short duration, lasting approximately 15ms. There is an obvious spike marking the release of the velar stricture and the commencement of velic closure, which is completed approximately 25ms later at time (a) and is observed to correspond with an increase in amplitude in the waveform.

The voiceless velar nasal of Figure 2.7 is momentarily voiced at the onset before the characteristic breathiness develops. Once again, breathiness is demonstrated by noise in the frequencies above 2000Hz. It is also observed that the breathiness perseveres for a few milliseconds through the transition to the vowel formants. A breathy vowel has consequently been transcribed and segmented accordingly.

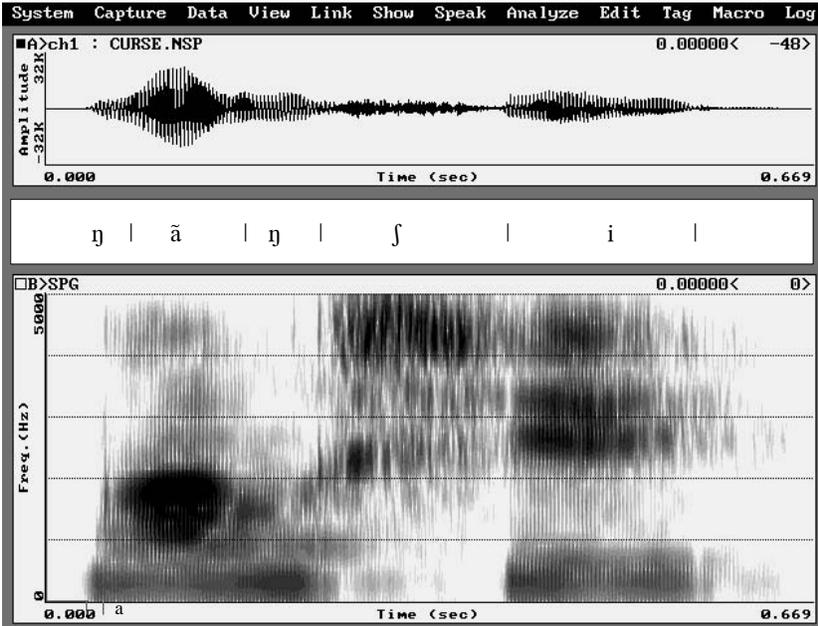


Figure 2.6. Waveform and wide band spectrogram (234Hz) of *ɲaɲsi* ‘cursed’

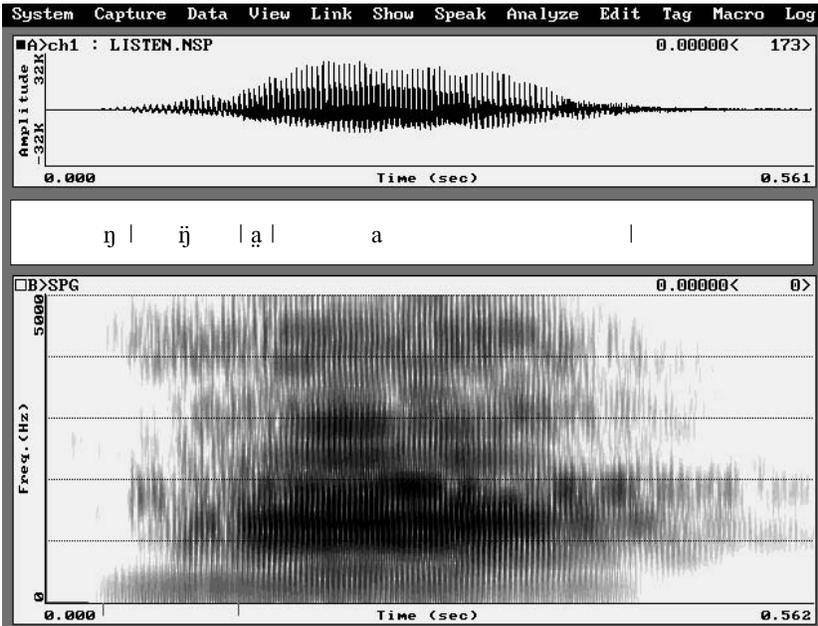


Figure 2.7. Waveform and wide band spectrogram (234Hz) of *hɲa* ‘listened’

In summary, I have demonstrated instrumentally that Mangmetong Mongsen has a series of voiceless nasals, and minimal and sub-minimal pairs prove that these are in contrastive distribution with their voiced nasal counterparts. On phonetic grounds, it may be argued that it is more correct to refer to members of this series as breathy nasals rather than voiceless nasals, since the principal acoustic feature they demonstrate is a breathy onset. However, Ladefoged and Bhaskararao (1991: 82) observed that female speakers articulated the voiceless nasals of Burmese with a slight breathy quality in a study of voiceless nasal segments, yet that did not motivate the authors to refer to them as “breathy nasals”. Secondly, languages reported to have phonemic breathy nasals, e.g. Hindi, Marathi, and Newari (Ladefoged and Maddieson 1996: 107) also have a series of breathy stops, but such a series of breathy consonants is notably absent from the Mongsen phoneme inventory, whereas voicelessness is a heavily exploited phonatory parameter in the phonological system of the language.

2.2.1.6. *Approximants*

There are three central approximants and one lateral approximant, with each articulatory position demonstrating a two-way voice onset time contrast. The results of an acoustic analysis investigating the nature of this voice onset time contrast are presented below.

The voiced labial velar and palatal approximants occur in initial and medial position, both as syllable onsets and as the non-syllabic offglides of phonetic diphthongs, e.g. labial velar: *wa* (go.PST) ‘went’, *sùwì* ‘slow loris’ and *tʃhàwk* (keep.PST) ‘kept’; and palatal: *jim* (fly.PST) ‘flew’, *a-ja* (NRL-night) ‘night’ and *hɲatsə̀laj* ‘tadpole’. Only voiced velar and voiced palatal approximants can occur in final position when forming the offglides of phonetic diphthongs, e.g. *tʃhu.u-ì* [tʃhu³³.uɲ³³] (cook-IRR) ‘will cook’ and *aw* [aw³³] (crawl.PST) ‘crawled’.

The voiced apical post-alveolar approximant is often realized as a sub-laminal retroflex approximant [ɻ], particularly before a non-front vowel. It occurs in all positions, e.g. *ɲuk.ɲa* ‘sixty’, *ləmsìù* ‘comb’, and *thukuɻ* ‘ninety’. The voiced lateral may have an apical alveolar or an apical dental articulation, although this is difficult to establish precisely without the benefit of palatography. It occurs in initial and in medial position, e.g. *lə̀ph.u* ‘cockroach’ and *hɲula* ‘voice’.

A typologically rare set of voiceless approximants never previously recognized in any dialect of Ao is found to contrast with their voiced counterparts. These complement the voiceless nasals described in §2.2.1.5 and form a voiceless sonorant series. Like the voiceless nasals, voiceless approximants carry a lower functional load than do their voiced counterparts. An approximately 800-item word list contained 211 occurrences of voiced

approximants in syllable onset position compared to just 39 occurrences of voiceless approximants. Voiced approximant onsets thus outnumber voiceless approximant onsets by over five to one. The voiceless labial velar approximant is the least common voiceless sonorant and is only encountered in a few words. The voiceless palatal, lateral and post-alveolar approximants occur with roughly the same frequency.

Voiceless approximants occur in syllable onsets in initial and medial position, e.g. voiceless labial velar approximant: *hwaŋ* (roast.PST) ‘roasted’ and *a-hwǎ?* (NRL-bamboo) ‘bamboo sp.’; voiceless palatal approximant: *hjuŋza* ‘ant sp.’ and *ahjak* (be.shy.PST) ‘was shy’; voiceless post-alveolar approximant: *hɹətsə?* ‘borer, insect sp.’ and *ahɹitsʉ* ‘bachelors’ house’; and voiceless lateral approximant: *hlâ* (descend+go.PST) ‘went down’ and *a-hluŋ* (NRL-winged termite) ‘winged termite’.

The following pairs of spectrograms demonstrate acoustic evidence of the phonemic contrast between voiced and voiceless approximants at each of their distinctive places of articulation. The speaker (II) is a male in his late sixties. The data was recorded in a paddy field using the equipment and methodology described above in §2.2.1.5 for the acoustic analysis of the voiceless nasals.

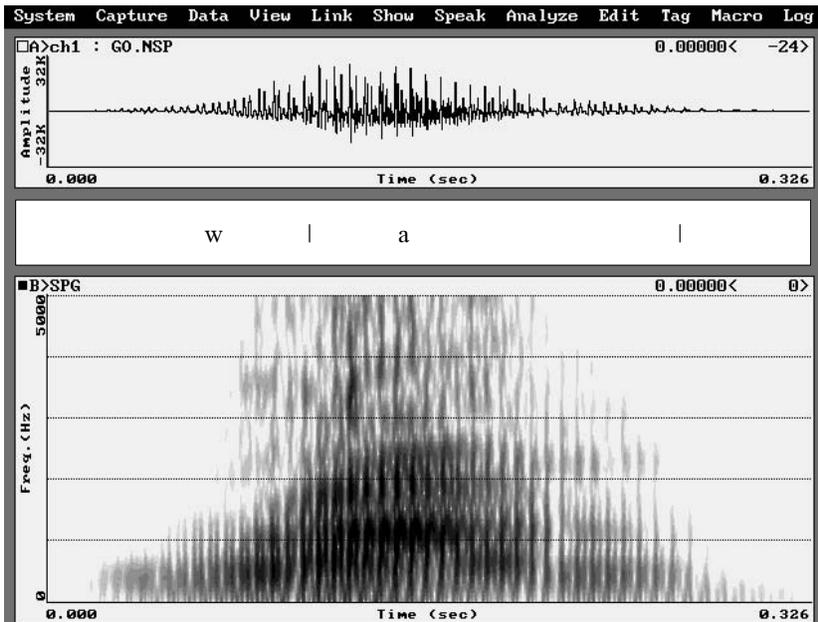


Figure 2.8. Waveform wide band spectrogram (234Hz) of *wa* ‘went’

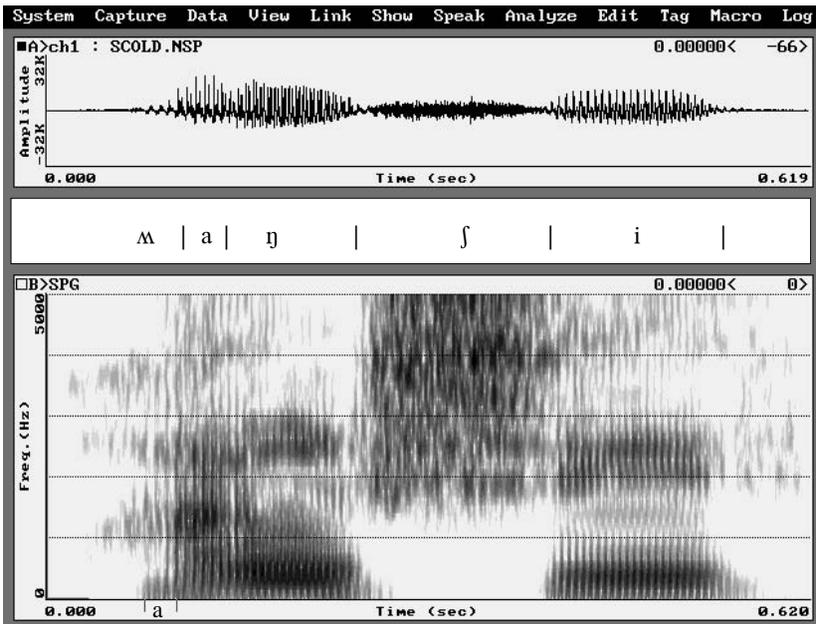


Figure 2.9. Waveform and wide band spectrogram (234Hz) of *hwagsi* ‘scolded’

Figures 2.8 and 2.9 above compare voiced and voiceless labial velar approximants respectively in the sub-minimal pairs *wa* (go.PST) ‘went’ and *hwagsi* (scold.PST) ‘scolded’. The contrast between the two labial velar onsets is signalled by the presence of breathy voice in the spectrogram of Figure 2.9. Evidence of this can be seen as noise in the frequencies above 1000Hz accompanying the modal voice phonation, and in the slightly noise-excited signal of the waveform corresponding with the section marked (a) on the time axis. The breathy voice onset of the voiceless labial velar approximant is quite short in comparison to the modal voice onset of the voiced labial velar and lasts just 26ms.

Figures 2.10 and 2.11 below compare the voiced and voiceless palatal approximants respectively in the words *ju* (carry.PST) ‘carried’ and *hju* (angle.PST) ‘angled’. In Figure 2.10, the change in formant structure indicates a transition to the high back rounded vowel after the prolonged onset of the voiced palatal approximant, marked by (a) on the time axis. In contrast, the onset of Figure 2.11 is substantially breathy, indicated by very flat valleys between peaks in the waveform that are consistent with a prolonged open phase during each glottal pulsation. Voicing commences with the first evidence of excitation in the waveform and the breathy phonation continues up to the transition to the vowel formants.

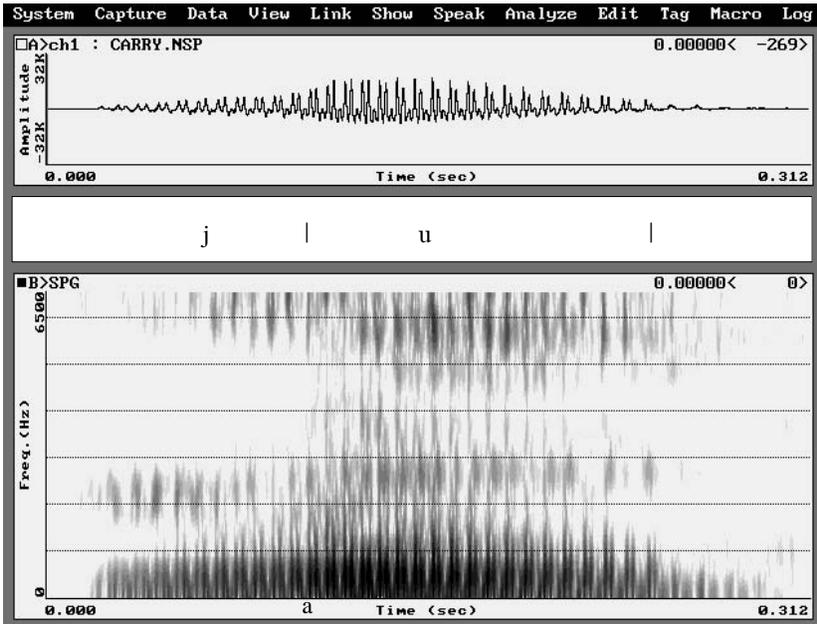


Figure 2.10. Wave form and wide band spectrogram (469Hz) of *ju* ‘carried’

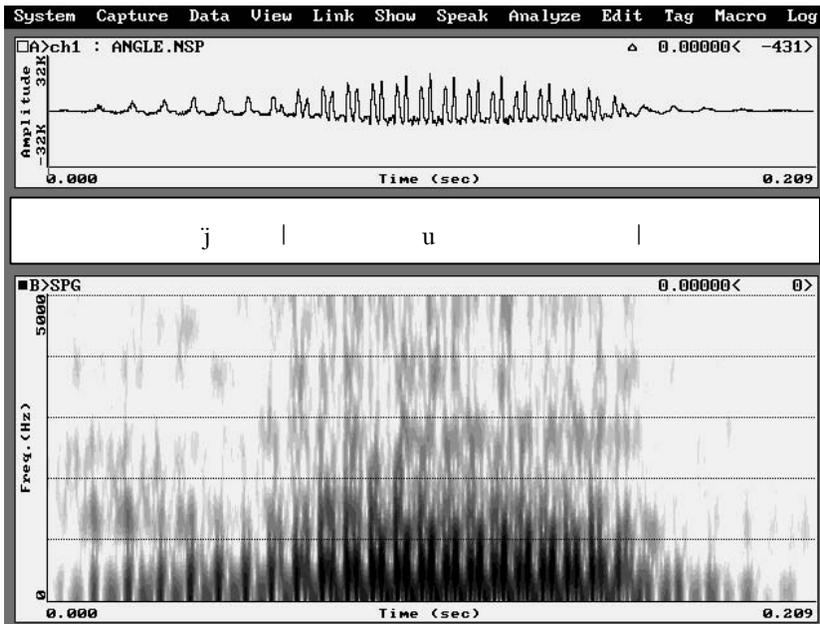


Figure 2.11. Wave form and wide band spectrogram (469Hz) of *hju* ‘angled’

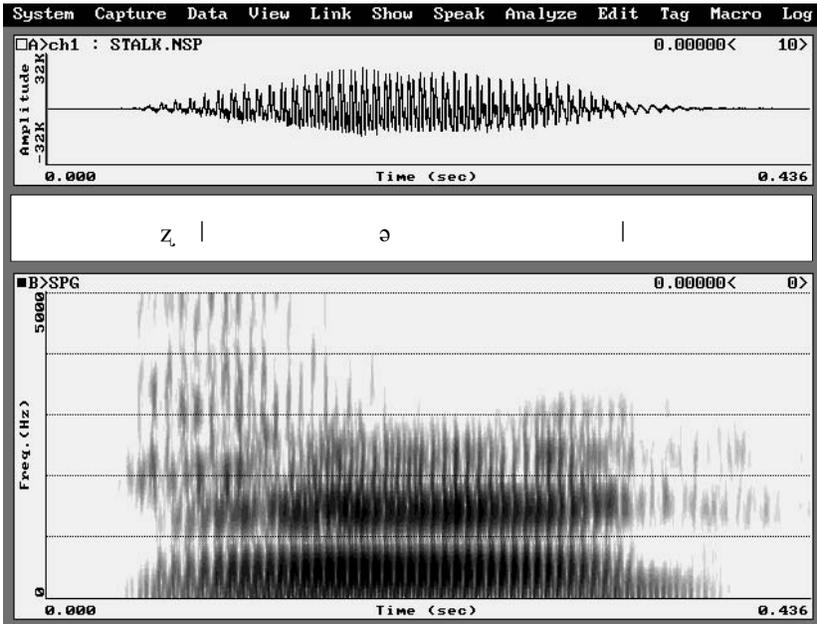


Figure 2.12. Waveform and wide band spectrogram (313Hz) of *ɹə* ‘stalked’

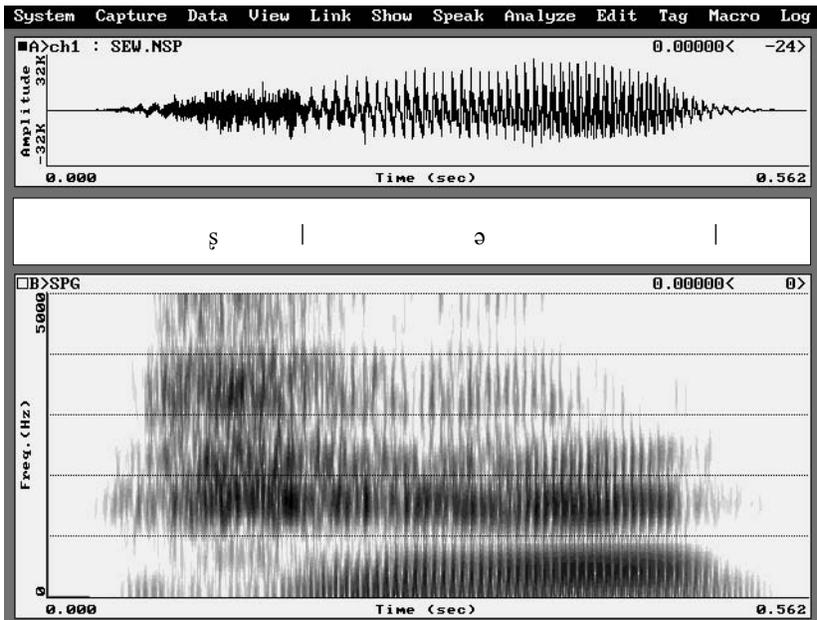


Figure 2.13. Waveform and wide band spectrogram (313Hz) of *hɹə* ‘sewed’

Figures 2.12 and 2.13 above compare the voiced and voiceless post-alveolar approximants in the words *ˌstɑːk* (stalk.PST) ‘stalked’ and *ˌsiːw* (sew.PST) ‘sewed’ respectively. These are noteworthy for demonstrating just how much friction the post-alveolar approximants can generate. The voiced post-alveolar approximant *ɹ* is actually realized as a retroflex fricative [ʒ]. Its onset lasts for 15ms, measured from the first evidence of excitation in the waveform to the first clean peak. The voiceless post-alveolar approximant similarly generates a great deal of friction during its articulation and is realized phonetically as [ʃ]. The abundance of high frequency noise in the absence of modal phonation identifies the onset of this segment as truly voiceless. Voice onset time is calculated to be approximately 200ms.

The final two spectrograms compare the phonetic realizations of the voiced and voiceless lateral approximants in the existential copula *li* ‘be’ and the verb *hli* (buy.PST) ‘bought’. Both spectrograms show an obvious transition from the lateral onset to the vowel, but the voiceless lateral is distinguished by substantial breathiness presenting as noise in the higher frequencies. The manifestation of noise can also be seen in the section of the waveform corresponding to the onset. The entire interval of the breathy onset is calculated to be 95ms.

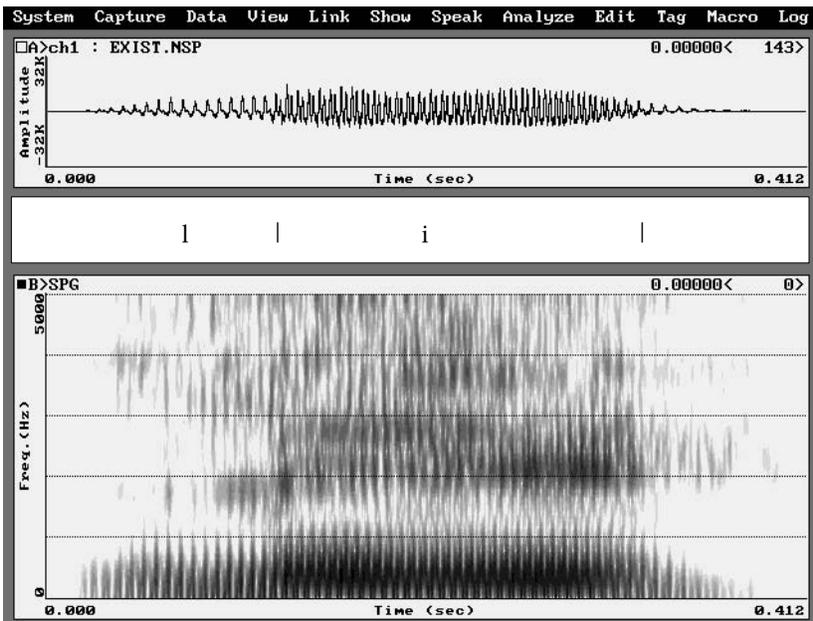


Figure 2.14. Waveform and wide band spectrogram (313Hz) of *li* ‘be’

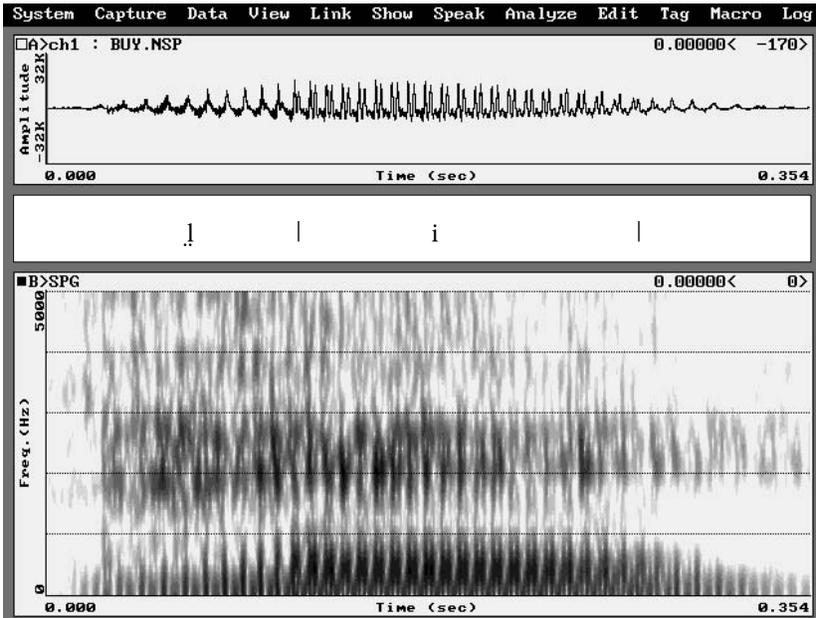


Figure 2.15. Waveform and wide band spectrogram (313Hz) of *hli* ‘bought’

2.2.1.7. *Limited free variation*

Free variation occurs between *tf* and *j* in two verbal suffixes: in allomorphs of the perfective aspect marker *-tfuk*, viz. [tʃuk³³] ~ [juk³³]; and in allomorphs of the resultant state marker *-tfak*, viz. [tʃak³³] ~ [jak³³] (see §8.5.6 and §8.5.7 respectively). Some examples in words are *wa-tfuk* ~ *wa-juk* (go-PFV.PST) ‘went away’, and *mətʊŋ-tʃak-i?* ~ *mətʊŋ-jak-i?* (be.erecst-RS-CAUS.PST) ‘caused to be left standing upright’. This is worth mentioning because the phonemes *tf* and *j* are in contrastive distribution everywhere else, e.g. *jà?* (heard.PST) ‘heard’ and *tʃà?* (consume.PST) ‘ate’, or *a-juŋ* (NRL-river) ‘river’ and *a-tʃuŋ* (NRL-shield) ‘shield’. The anomaly is therefore interpreted as a marginal suspension of an otherwise robust phonemic contrast, limited to just these suffixes.³ Lass (1984: 21) notes a similar phenomenon affecting otherwise contrastive vowel phonemes of English; e.g. /i:/ ~ /ai/ occur in free variation in words such as *either*, *neither*, yet contrast in *beet*, *bite*.

3. What appears to be free variation occurring between the phonemes *z* and *j* in the Waromung variety of Mongsen is attributed to a former intra-village bidialectalism that has resulted in borrowing, either of an alternative pronunciation in certain words, or of whole lexical items. It is discussed at length in Coupe (2003a: 47–50).

2.2.2. Vowel phonemes

There are four modal voice vowel phonemes comprising three peripheral vowels and one mid central vowel. This basic system is augmented by a phonation contrast occurring at the low central articulatory position, adding a fifth contrastive segment to the inventory. The creaky voice phoneme represents an irregularity amongst the vowel phonemes because its contrastive phonation type is restricted to this single articulatory position.

The vowel phonemes are charted in Table 2.3 below. Phonemic and orthographic representations are signalled by the same symbols here and elsewhere in language examples.

Table 2.3. Vowel phonemes of Mangmetong Mongsen

	FRONT	CENTRAL	BACK
High	i		u
Mid		ə	
Low		a ʔ	

The high front vowel *i* occurs in initial, medial and final positions, e.g. *ila* (1PL.IN) ‘we, inclusive of addressee’, *ipu* ‘storm’, *á-lí-má* (NRL-ground-FACE) ‘world’, *a-jim* (NRL-village) ‘village’, *a-hli* (NRL-flea) ‘flea’ and *a-mi* (NRL-spear) ‘spear’.

The high back rounded vowel *u* has two allophonic realizations in free variation, viz. [u] ~ [o]. The close-mid back rounded variant is mostly realized in the proximity of a velar environment, but this is not consistent enough to be stated as a rule, as examples of [u] occurring in the velar environment are also encountered in the data. Furthermore, these allophones can be exchanged in all words without a change or loss of meaning. The high back rounded vowel phoneme occurs in initial, medial and final positions, e.g. *uɣtuk* ‘Indian one-horned rhinoceros’, *uhlu?* ‘parakeet sp.’, *tə-puk* (RL-stomach) ‘stomach’, *khùma* ‘sore on skin’, wound’, *hna.u* ‘flower’ and *upu* ‘ashes’.

The schwa is articulated as a mid central vowel ə with the lips slightly spread. This is the weakest articulatory position in the vowel phoneme inventory and often assimilates in height, backness and rounding to adjacent glides in word formation processes (§2.6.4), or is deleted in the environment of a more stable peripheral vowel (§2.6.1). The schwa occurs infrequently in initial position in Mangmetong Mongsen, some rare examples being *əntʃu* ‘curry’ and *ənti* ‘path’. In the Waromung Mongsen variety, however, schwa-initial words are more common because of the loss of a word-initial initial *h* in a considerable number of words, possibly through the influence of Chungli on

that variety (see Coupe 2003a: 47–50 for discussion of Chungli-Mongsen dialect mixing in Waromung village).

In lexical items, the schwa occurs most commonly in medial position, e.g. *tə-ləm* (RL-head) ‘head’ and *tʃhəmpaŋ mətʃəmpùŋ* ‘yellow-backed sunbird’, and quite rarely in final position in lexical items. All word-final occurrences of ə in lexical items have a fricative, a dental affricate or a retroflex approximant in the syllable’s onset, e.g. *tʃaŋsə* ‘footprint’, *hjuʃtsə* ‘story’ and *hɪə* (perforate.PST) ‘perforated, sewed’. The final position is less restricted in grammatical morphemes, such as in the benefactive case-marker *ataməkə*, in the quotative particle *tə* ‘thus’, and in the conflated agentive/instrumental/allative case marker *nə*. Mangmetong Mongsen speakers often resyllabify a final voiced retroflex approximant, which also results in a schwa occurring in the word-final position (see §2.6.6).

The phonemic contrast between the vowels ə and i is often neutralized in initial position, allowing allophones of either vowel to be realized when the syllable carries a low tone; e.g. *ənti* ‘road, path’ can be realized phonetically as either [ən¹¹ti³³] or [in¹¹ti³³]. Some female speakers of Waromung Mongsen are observed to metathesize, centralize and lower the vowel of the first person singular pronoun *nì* ‘I’ when it occurs in a noun phrase with the agentive case marker, the latter being articulated as *na* with a low central vowel in that variety. The pronoun and the agentive marker are then pronounced together as [ən¹¹na³³]. This suggests that the neutralization could be licensed by a combination of the initial position, the low tone, and perhaps the dental nasal environment.

The low central modal voice vowel *a* occurs in initial, medial and final position in all environments, e.g. *a-tʃa* (NRL-cooked rice) ‘cooked rice’, *a-paŋ* (NRL-water buffalo) ‘water buffalo’, *tə-tʃaŋ* (RL-leg) ‘leg’, *lazati* ‘maiden’, *pa* (3SG) ‘s/he’ and *a-nu-za* (NRL-child-DIM) ‘small child, infant’.

A low central creaky voice vowel phoneme *ǣ* occurs in the absence of a glottal stop environment and in contrastive distribution with *a*, the low central modal voice vowel. Coupe (2003a: 43–45) analyzes it as the idiosyncratic realization of a separate vowel phoneme. Since the creaky voice phonation type does not occur at other articulatory positions, it cannot be treated as a phonatory setting applying to the whole vowel phoneme inventory. The creaky voice vowel carries a very light functional load and occurs in just five lexical roots of the entire corpus, e.g. *wapət* ‘slope’, *hwǣ-tsə* (bamboo-water) ‘bamboo sap’, *wǣ-pàʔ* (slice-NR) ‘to slice’, *tǣnəm wǣpùŋ* ‘Great pied hornbill cock’ and *tǣnəm wǣtsə* ‘Great pied hornbill hen’. The meanings of these words suggest that creaky voice phonation is unlikely to be attributable to onomatopoeia, therefore *ǣ* must be analyzed as a rather unusual but nevertheless independent vowel phoneme. Its acoustic characteristics are investigated in Coupe (2003a: 81–86).

Contrastive creaky voice $\underset{\sim}{a}$ occurs in cognate words in every variety of Mongsen encountered and always in the environment after a labial-velar approximant, where it contrasts with modal voice a , e.g. *wa* (go.PST) ‘went’ and *səŋàwà* ‘leaf’. It therefore seems that the presence of a labial velar approximant is a prerequisite for its realization, although it is difficult to establish the relevance of this with respect to articulatory phonetics.

2.3. Segmental differences between Mongsen varieties

The following sub-sections briefly describe differences encountered in the consonant and vowel phoneme inventories of various village varieties of the Mongsen dialect. As the dialect is not codified, it is highly likely that every single Mongsen-speaking village has its own variety with distinctive phonological, morphological and lexical differences. These shibboleths allow native speakers who are familiar with different varieties of Mongsen to identify the village of origin of other speakers.

2.3.1. Consonants

As noted in §2.2.1.1, a major difference between Mongsen varieties is the presence or absence of a voiceless sonorant series. This potentially adds seven phonemes to a basic consonant phoneme inventory of twenty segments.

Longkum Mongsen has a voiceless labial dental fricative f , e.g. *a-fə* (NRL-snake) ‘snake’ (cf. *a-sə* in Mangmetong Mongsen). Khensa Mongsen and Mekhuli Mongsen have a voiceless labial dental fricative in some words, e.g. *fù* (blow.PST) ‘blew’, *a-fi* (NRL-snake) ‘snake’, *a-fi?* (NRL-rat) ‘rat’ and *ufuŋ* ‘dust’ (corresponding respectively with *phu*, *a-sə*, *a-hi?* and *uphuŋ* in Mangmetong Mongsen). There is additionally a voiced labiodental fricative v in Khensa and Mekhuli, heard in the words *tə-vəti* (NZP-be.big) ‘big’, *məvuŋ* ‘wind’ and *vù?* (bite.PST) ‘bit’ (respectively *təpəti*, *məpuŋ* and *mù?* in Mangmetong Mongsen).

A labialized voiceless unaspirated velar stop $[k^w]$ was heard to contrast with a voiceless unaspirated velar stop in a Khar Mongsen text in the words $[tsə^{11}\eta i^{11}k^w a^{33}]$ ‘day’ and $[tʃho^{33}ka^{33}]$ (emerge.PST) ‘emerged’ (respectively *tsəŋikhu* and *tʃhuwa* in Mangmetong Mongsen), but a full phonological analysis of this variety has not yet been undertaken to identify any other phonological differences that may exist in its consonant phoneme inventory.

2.3.2. Vowels

The one central and three peripheral articulatory positions of Mangmetong Mongsen are, together with the creaky voice vowel, assumed to be representative of the basic vowel phoneme inventory of the Mongsen dialect. The Mongsen varieties spoken in Waromung and neighbouring Khar village, both situated on Changkikong Range, have innovated an additional high central rounded vowel *ʊ* while retaining the three peripheral vowels, schwa and low central creaky voice vowel. Waromung Mongsen *ʊ* is found to be in partial correspondence with Mangmetong Mongsen *i*; it is most frequently encountered in verbal suffixes, and to a much lesser extent in lexical items. The examples of (2.3) below demonstrate these correspondences in words and verbal affixes, while those of (2.4) demonstrate a sample of cognate words in which both reflexes retain *i*.

(2.3)	Mangmetong	Waromung	Gloss
	<i>-iʔ</i>	<i>-ʊʔ</i>	CAUS
	<i>-ì</i>	<i>-ù</i>	IRR
	<i>tə-ni</i>	<i>tə-nʊʔ</i>	(RL-wife) ‘wife’
	<i>a-hiʔ</i>	<i>a-ʊʔ</i>	(NRL-rat) ‘rat’
	<i>tə-jiʔ</i>	<i>tə-ʊʔ</i>	(RL-mother) ‘mother’
	<i>pi</i>	<i>pʊ</i>	(PROX) ‘this’
	<i>á-nítì</i>	<i>á-núʔì</i>	(NRL-woman) ‘woman’

(2.4)	Mangmetong	Waromung	Gloss
	<i>məni</i>	<i>məni</i>	(laugh.PST) ‘laughed’
	<i>hʲùŋli</i>	<i>jùŋli</i>	(stand.PST) ‘stood’
	<i>tə-niŋ</i>	<i>tə-niŋ</i>	(RL-name) ‘name’
	<i>a-mi</i>	<i>a-mi</i>	(NRL-spear) ‘spear’
	<i>a-miʔ</i>	<i>a-miʔ</i>	(NRL-person) ‘person’

These examples suggest that a sound change is in evolution in Waromung Mongsen. At this stage the *i* > *ʊ* transition has involved all of the grammatical morphemes with a high front vowel, but only a small number of Waromung Mongsen lexical items with a high front vowel have thus far been affected. Lastly, it is pertinent to mention that the high central rounded vowel of Waromung Mongsen is additionally found to correspond with two other vowels of Mangmetong Mongsen in just two words. These anomalies are listed below in (2.5) for the sake of completeness.

(2.5)	Mangmetong	Waromung	Gloss
	<i>a-u</i>	<i>a-ʰ</i>	(NRL-axe) ‘axe’
	<i>a-sə</i>	<i>a-hʰ</i>	(NRL-snake) ‘snake’

2.4. Word

A word in Mongsen may minimally consist of just one syllable; this in turn may minimally consist of a monomoraic root forming a single morpheme, which is simultaneously a grammatical word and a phonological word, e.g. *tʃa* (call.PST) ‘called’, *pù?* (carry on back.PST) ‘carried on the back’ or *nì* (1SG) ‘I’. Monosyllabic and disyllabic roots occur with roughly equal frequency in the lexicon. In a sample of approximately 800 lexical items, forty-seven percent of Mangmetong Mongsen words have monosyllabic roots, forty-nine percent have disyllabic roots, and the remainder consist of trisyllabic roots.

The word in Mongsen can be identified on the basis of phonological and grammatical criteria (after Dixon and Aikhenvald 2002), presented below in §2.4.1–§2.4.2. In most cases the two sets of criteria coincide.

2.4.1. Phonological criteria

(i) *Segmental features.* Diphthongs may form across syllable boundaries within phonological words that also constitute grammatical words e.g. *.jà* ‘come’ + *-i?* CAUS → [jaʃi³³] ‘caused to come’, *tʃhà* ‘do’ + *-ukù* ANT → [tʃhaw³³ku¹¹] ‘has done’. Rarely, a single phonological word can consist of two grammatical words. This occurs when two independent grammatical words fuse across a word boundary. For example, the juxtaposed vowels of the separate phonological and grammatical words *á-límá* (NRL-world) ‘world’ and the proximate demonstrative *i* ‘this’ diphthongize to form one phonological word comprising two grammatical words [a⁵⁵li⁵⁵maj³³] ‘this world’, pronounced without a hiatus. Such phenomena are entirely dependent upon vocalic juxtapositions forming diphthongs that are permitted by the phonotactic rules of the language, and are therefore considered marginal.

Sequences of vowels having identical quality and carrying identical tones as a rule do not occur within a phonological or a grammatical word.⁴ If a $V_\alpha V_\alpha$ sequence forms via morpheme concatenation or through the juxtaposition of two phonological words, then usually one of the vowels either undergoes deletion, or the first vowel of the sequence dissimilates to schwa (see §2.6.3 for

4. The only exception to this is presented by the distributive for ‘eight’. See §7.6.1 for discussion.

examples of dissimilation). A sequence of vowels of identical quality and tone in a stream of speech therefore signals a boundary between two words. Sequences of identical consonants, however, may occur both within a stream of speech and within a word. While the latter is quite rare, it is sometimes found to result from suffixation and/or assimilatory processes, e.g. *a-nət* (NRL-two) ‘two’ + *nəm* ‘ply’ → [a³³nən³³nəm³³] ‘two-ply (weaving)’, and thus does not reliably indicate a word boundary. Of the consonant phonemes, only voiceless unaspirated stops, voiced nasals and the voiced retroflex approximant may occur syllable-finally, and therefore word-finally, but this is not a sufficient criterion for the identification of a word boundary because any allowable coda constituent may occur word-internally. Even so, we might use this phonotactic constraint on syllable formation as negative evidence that if a segment is *not* a voiceless unaspirated stop or a voiced nasal or retroflex approximant, then that particular segment cannot be delimiting a word boundary. Other edge phenomena, such as the *i* ~ *ə* neutralization discussed in §2.2.2, are inconsistent and therefore also unreliable indicators of word margins.

(ii) *Prosodic features.* Tone sandhi processes operate across morpheme boundaries in phonological words and across grammatical word boundaries. With respect to suffixes, the tone of the cohered element often spreads leftward and displaces the tone of the root, as exemplified by some of the examples provided in (i). This is seen to raise the tone of the root from low to mid in *ja-jʔ* (come-CAUS.PST) ‘caused to come’ and *tʃha-wkù* (do-ANT) ‘has done’, and to lower it from high to mid in the final syllable of *á-líma-j* (NRL-world-PROX) ‘this world’. With respect to prefixes, tones can similarly spread toward and associate with the root, e.g. *tə-* PROH + *ata* ‘await’ → [ta¹¹ta³³] ‘don’t await’. Here the low tone of the prohibitive suffix spreads rightward after the schwa is deleted and associates with the initial vowel of the verb root, displacing the mid tone of the root *ata* ‘await’. The obligatory association of at least one output tone per syllable is a necessary but not sufficient criterion for wordhood, because most affixes and clitics similarly have tones associated with their syllables. Tone is therefore not a useful criterion for establishing the distinction between word, suffix or clitic.

The glottal stop (discussed in §3.5) is a prosodic element that occurs as the final segment in many verb roots and a considerable number of nominal roots. A few verbal suffixes and some of the nominal suffixes also have glottal stops in their underlying representations, e.g. the causative suffix *-iʔ*, the simultaneous converb suffix *-líʔ*, the reflexive suffix *-jimtàʔ* and the general nominalizer *-pàʔ*. A glottal stop occurring as part of the underlying representation of a root or suffix will only be realized in the output if it is the last element in the linear string of constituents of the word. If it is not, then morpheme concatenation results in the formation of word-internal syllable boundaries, and hence the environmental conditions under which a word-internal glottal stop is deleted.

The word-internal deletion of a glottal stop might accordingly be cited as negative evidence for wordhood, i.e. a morpheme boundary that triggers the deletion of a glottal stop is not also a word boundary. When it *is* realized, the glottal stop prosody of Mongsen is an indicator *par excellence* of a word boundary, and therefore of both a phonological and a grammatical word.

(iii) *Phonological processes*. Certain dissimilatory processes are confined to the word-internal environment (see §2.4.2 for exemplification). For example, a high front vowel occurring as the final linear constituent of a verb root is realized as a schwa in the environment before the high front vowel of a suffixed irrealis marker *-i*, as in *hli* ‘buy’ + *-i* → [l̥əj³³] ‘will buy’ (see §2.6.3). This dissimilation is motivated by the need to preserve the grammatical information that would otherwise be lost through the deletion of one of the two identical vowels. Dissimilatory changes of this nature never occur across phonological or grammatical word boundaries, but apply automatically to any word-internal high front vowel when it is concatenated with a suffix whose initial segment is also [+high, –back].

2.4.2. Grammatical criteria

The grammatical word in Mongsen conforms to criteria frequently cited in the literature (see Dixon and Aikhenvald 2002 and references therein), namely:

- (i) *Isolatability*. Speakers readily identify and assign meanings to grammatical words, but are less likely to be able to segment words into morphemes, or to assign meanings to segmented morphemes.
- (ii) *Immutability*. A root and its affixes that form a grammatical word have a conventionalized order for the concatenation of morphemes. Disruption to this conventionalized order results in nonsense, or minimally, a change in meaning.
- (iii) *Potential pause*. Pauses may occur between words, but not within words.

2.5. Clitics

A clitic in Mongsen generally has the following characteristics:

- (i) it has low selectivity with respect to the lexical category of its host;
- (ii) it cannot occur independently of its host; that is, it cannot function as an independent phonological word;
- (iii) a clitic’s scope of relevance is the word, phrase or clause (this is in contrast to affixes, whose scope of relevance is confined to the lexical stem);

- (iv) it may only occur as the last element of the word, phrase or clause;
- (v) the clitic and its host do not admit an intervening pause.

Mongsen clitics are all enclitic because of their leftward direction of attachment to their hosts. Case-marking clitics are discussed in §5.3, and the optional declarative mood clitic is described in §4.2.14.6.

2.6. Phonological processes

Phonological processes can be divided into those that involve prefixes: schwa deletion (§2.6.1), sporadic vowel harmony (§2.6.2) and glide assimilation (§2.6.4); those that involve roots: vowel dissimilation (§2.6.3); and those that involve suffixes: consonant assimilation (§2.6.5) and resyllabification (§2.6.6). Glottal stop deletion at word-internal morpheme boundaries applies to prefixes, lexical roots and suffixes and is discussed in §3.5.

Some phonological processes occurring in Waromung Mongsen are described in Coupe (2003a: Ch 2) using an autosegmental framework (Goldsmith 1990). These previously investigated patterns are only dealt with briefly in the following sections. Attention is drawn to differences where they are found to occur in Mangmetong Mongsen.

2.6.1. Schwa deletion in prefixes

Four prefixes have inherently unstable schwa nuclei that are deleted when affixed to a stem with a more stable initial low vowel. These are: the nominalizing prefix *tə-* that derives adjectives from verb roots, the homophonous relational prefix *tə-* that occurs on bound kinship terms and body part noun stems, the verbal prohibitive prefix *tə-*, and the verbal negative prefix *mə-*. Some examples are: nominalizing prefix *tə-* + root *asu?* ‘be.small’ → *tasu?* ‘small’; relational prefix *tə-* + root *u* ‘grandfather’ → *tu* ‘grandfather’; prohibitive prefix *tə-* + root *ùŋ* ‘shave’ → *tùŋ* ‘don’t shave’; and negative *mə-* + stem *ilak-àɪ* (pour-PRES) → *milakàɪ* ‘doesn’t pour’.

If the low tone of a prefix differs from the tone associated with a root-initial vowel, then the prefix’s tone spreads to the initial syllable of the root and displaces its underlying tone. Some additional examples demonstrating this are: prohibitive prefix *tə-* + stem *atsə* ‘look’ → *tàtsə* ‘don’t look’, and negative prefix *mə-* + stem *aiəm-ì* (bear-IRR) → *màaiəmì* ‘won’t bear’. The tendency for Mangmetong Mongsen speakers to pronounce what are vowel-initial roots in other varieties with initial *h* limits the extent of this phonological process by assisting the schwa to maintain its syllabic independence.

2.6.2. Sporadic vowel harmony

There is a tendency for the phonologically weak schwa of the above-mentioned prefixes to assimilate to a high back rounded vowel of a neighbouring syllable in the idiolects of some speakers. This can probably be attributed to anticipatory articulatory gestures, such as lip rounding, and the speed of the utterance. Some examples are: relational prefix *tə* + *puk* ‘stomach’ → [tu³³puk³³] and relational prefix *tə* + *kuluk* ‘brain’ → [tu³³ku³³luk³³]. Sporadic vowel harmony is not observed to correlate with any sociolinguistic variables, such as age or sex.

2.6.3. Vowel dissimilation

In Mangmetong Mongsen, a stem-final high front vowel dissimilates to schwa when it is juxtaposed to another high front vowel in word formation processes. This affects stems and affixes in which the segments on either side of a morpheme boundary are both high front vowels, e.g. when a verb stem such as *li* ‘stay’ is concatenated with the high front vowel of the causative suffix *-i?* or the irrealis suffix *-i*. The concatenation of the causative and irrealis suffixes in the same verb stem can also create the environmental conditions under which vowel dissimilation occurs.

The concatenation of such stems and/or affixes in word formation initially results in a sequence of identical high front vowels, which we will distinguish with subscript numbers for ease of exemplification, e.g. V₁V₂. As tautosyllabic sequences of identical vowels are prohibited by the phonotactics (see §2.1.1), this causes V₁ to dissimilate to schwa and V₂ to form an offglide in the output. The following examples are illustrative: *li* ‘stay’ + *-i* (IRR) → [ləj³³] ‘will stay’; *tʃasi* ‘be distressed’ + *-i?* (CAUS) → [tʃa¹¹ʃəj³³] ‘upset (someone)’; *hni-si* (follow-RPET) + *-i* (IRR) → [hni³³ʃəj³³] ‘will keep on following’; and *tʃəm* ‘drink’ + *-i?* (CAUS) + *-i* (IRR) → [tʃəm³³əj³³] ‘will make drink’. The dissimilation of V₁ is motivated by the fact that important grammatical information encoded by the high front vowel of a suffix (represented by V₂) would otherwise be rendered opaque by the fusion of identical phonological features across the morpheme boundary, were the dissimilation not to occur.

A root- or morpheme-final glottal stop is always deleted in the environment of a word-internal syllable boundary (see §3.5), thus its underlying presence does not block the dissimilation process affecting a high front vowel in roots such as *khi?* ‘give’, e.g. *khi?* + *-i* (IRR) → [khəj¹¹] ‘will give’. Additional evidence of its transparency to morphophonological processes is provided by the causativized example with the root *tʃəm* ‘drink’ in the preceding paragraph.

These dissimilation processes do not occur in Waromung Mongsen, because of that variety’s different morphological forms (cf. the examples of [2.3], and

see Coupe [2003a: 53–55], which discusses their very different behaviour in similar word formation processes).

2.6.4. Glide assimilation

If the initial consonant of a stem is a glide, then the schwa of the nominalizing prefix *tə-*, the relational prefix *tə-*, the prohibitive prefix *tə-* and the negative prefix *mə-* undergoes an assimilation in height and lip attitude to the adjacent glide segment under affixation.

Some examples are: nominalizing prefix *tə-* + *wa-pàʔ* ‘go-NR’ → [tu³³wa³³pa^{ʔ33}] ‘(the one) who went/had gone’, *tə-* + *jàk-mì* (beat-DESID) → [ti³³jak³³mi¹¹] ‘thrashable’; relational prefix *tə-* + *ja* ‘mother’ → [ti³³ja³³] ‘mother’; prohibitive prefix *tə-* + *hwàməsa* ‘yawn’ → [tu¹¹ma¹¹mə³³sa³³] ‘don’t yawn’; negative prefix *mə-* + *wa-əɪ* (go-PRES) → [mu¹¹wa³³] ‘does not go’; and *mə-* + *jàʔ-əɪ* (hear-PRES) → [mi¹¹ja¹¹] ‘does not hear’.

2.6.5. Consonant assimilation

Three examples of complete assimilation of a consonant are found in the corpus and all involve the numeral *a-nət* (NRL-two) ‘two’. When the derivational suffix *-pən* is used to derive a numeral adverb, it causes the final dental stop of *anət* to assimilate totally to the bilabial stop of the suffix, e.g. *a-nət* (NRL-two) + *-pən* → [a¹¹nə¹¹pən³³] ‘twice’. At the speed of normal conversation this does not sound geminated, so an alternative analysis could posit that the final dental of the numeral is instead deleted either before or after it assimilates to the bilabial stop. Unexpectedly, the assimilation triggered by the ordinal suffix in this example is not attested with other numerals; for example, *təɲuk* ‘six’ + *-pən* → [tə¹¹ɲuk¹¹pən³³] and *thəɲaɪə àkhət* ‘eleven’ + *-pən* → [thə³³ɲa³³ɲə³³khət¹¹pən³³]. The tone sandhi observable in some of these outputs is triggered by the derivational suffix.

Even more unexpectedly, the dental stop of *anət* disappears when a distributive is formed by reduplicating its final syllable, e.g. *a-nət-nət* (NRL-two-RED) → [a³³nə³³nət³³] ‘two each’. Assimilation of *t* to *n* also occurs in the compound *a-nət* (NRL-two) + *nəm* ‘ply’ → [a³³nən³³nəm³³] ‘two-ply (weave)’. The latter example is pronounced with a dental nasal coda and a dental nasal onset in the penultimate and final syllables respectively. Assimilation does not occur in the formation of other distributives, even for those with final dental stops, e.g. *nuklən* ‘hundred’ + *a-khət-khət* (NRL-one-RED) → [nuk³³lən¹¹a³³khət³³khət³³] ‘one hundred each’. Nor is consonant assimilation found with other compounds formed with *nəm* ‘ply’, e.g. *a-səm* (NRL-three) + *nəm* ‘ply’ → [a³³səm³³nəm³³]

‘three-ply (weave)’. The arbitrary nature of assimilation affecting just the root of the numeral *a-nət* (NRL-two) ‘two’ in a handful of words suggests that this should be analyzed as an instance of fossilization, as opposed to a productive phonological process.

2.6.6. Resyllabification

A characteristic of Mangmetong Mongsen speakers that sets them apart from speakers of other Mongsen varieties is their tendency to resyllabify a word-final retroflex approximant coda. This causes the retroflex approximant to form the onset of a separate syllable with a schwa nucleus; e.g. *sàɹàɹ* ‘animal’ is pronounced either as [sa¹¹.ɹa¹¹], or as [sa¹¹.ɹa¹¹.ɹə³³] after resyllabification. Sometimes there are idiosyncratic pitch realizations on the newly formed syllable, as in this example. Otherwise there is usually a simple mapping of the output tone of the penultimate syllable onto the newly formed final syllable. To illustrate, *wa-əɹ* (go-PRES) ‘goes’ is pronounced with a mid tone as [waɹ³³] after tone sandhi applies, or as [wa³³.ɹə³³] after resyllabification.

Another type of resyllabification process applies to suffixes that form vowel sequences across morpheme boundaries. This is facilitated by the high vowels *i* and *u* forming offglides *j* and *w* respectively in phonetic diphthongs, if they are able to fill the glide slot in the syllable template following morpheme concatenation. Some examples demonstrating the low central vowel of a stem forming a phonetic diphthong with the irrealis suffix *-i* or the anterior suffix *-ukù* are: *ɹà* ‘come’ + *-i* → *ɹàj* ‘will come’; and *sa* ‘say’ + *-ukù* → *sawkù* ‘has said’.

2.6.7. Other vowel deletions

If a sequence of non-high identical vowels forms as a result of morpheme suffixation, then vowel deletion results, e.g. *sana* ‘speak’ + *-aŋ* (IMP) → [sa³³.naŋ³³] ‘speak!’. Usually the tone of a suffix will spread to the root and displace its tone if it is underlyingly different, e.g. *tʃàʔ* ‘consume’ + *-aŋ* → *tʃa-ŋ* (consume-IMP) ‘eat!’, but sometimes aberrations are found where a low tone of a root stays low. This appears to be most likely when a vowel deletion is not involved, e.g. *mùŋ* ‘stay, rest’ + *aŋ* → *mùŋ-aŋ* (rest-IMP) ‘rest!’.

Chapter 3

Prosody

This chapter describes the use of prosody in Mongsen, encompassing lexical tone, grammatical tone, intonation and the glottal stop word prosody. The major function of tone in Mongsen is lexical, although it will be shown that it also has a minor grammatical role. Lexical tone is the focus of Coupe (1998) and Coupe (2003a), therefore this chapter deals only briefly with those areas already described and instead concentrates on outlining the complex characteristics of tone sandhi and the use of postlexical intonation, particularly with respect to its interaction with lexical tone.

3.1. Tonemes

Mongsen is a tonal language with three contrastive pitch levels occurring on all syllable types. Complex internal tone sandhi operates across morpheme boundaries in word formation processes; external tone sandhi is also found to extend across phonological word boundaries. An acoustic and auditory analysis of the lexical tone system of Waromung Mongsen is presented in Coupe (1998) and Coupe (2003a: 87–110). The findings can be considered to be broadly representative of all varieties of the Mongsen dialect, although tone sandhi manifestations in word formation processes can differ, sometimes being emblematic of particular village varieties in the absence of segmental differences in phonology. This chapter will be confined to an investigation of tone sandhi in just one variety of Mongsen – that which is spoken in Mangmetong village.

3.1.1. Lexical tones

Words uttered in isolation have characteristically level, terrace-like pitch registers. A minimal triplet demonstrating all three contrastive tonemes occurring on a monosyllable does not appear to exist in the language. In the absence of a monosyllabic minimal triplet, the following examples serve to demonstrate the use of pitch to differentiate lexical meanings in Mongsen.

The autosegmental analysis presented in (3.3) is superficially able to account for the tone sandhi realizations of the nominalized verb roots in the data of (3.2), but it is an expedient approach that is historically short-sighted. It assumes that originally there was segmental material associated with the high tone that has since been lost, leaving just a floating high tone synchronically. Comparative evidence suggests that a cognate form with a nominalizing function occurs widely in Tibeto-Burman languages – e.g. Meithei *-pə* (Chelliah 1997: 155), Chantyal *-wa* (Noonan 1997: 374), Classical Tibetan *-pa* (Beyer 1992: 299) – and these forms are probably all related to the Proto-Tibeto-Burman (hereafter PTB) masculine gender suffix **-pa* (Benedict 1972: 96), the source of which is **pa* ‘father’ [STC #24]. With so many cognate forms of often identical function, it is highly unlikely that any putative segmental material originally associated with the high tone would have been lost from every single one of the daughter languages without leaving so much as a trace.

In view of this historical evidence, we might abandon the hypothesis that a floating tone is associated with the general nominalizer and instead posit that the high tone is underlyingly associated with the final syllable of the verb roots of (3.2), and secondly, that it is only realized in the environment of the general nominalizer. In view of the fact that the M-H-L tonal pattern of the nominalized verb stems of (3.2) appears to be restricted to a specific subset of verbs, it may be the case that the underlying high tone is associated with these particular verb roots. Other Tibeto-Burman languages are reported to demonstrate morphologically-motivated tonal alternations in their verb stems; sometimes they are conditioned by only a few preceding or following elements. For example, the tones of final verbs in Mpi [Lolo-Burmese] are affected by negation (Bradley 1979: 49), and Ugong [Lolo-Burmese] is reported to have tonal alternations applying to word-final derivational suffixes, but not to word-final inflectional suffixes (Bradley 1992: 56–57).

We will return to the problem of tone sandhi manifestations involving the general nominalizer and other sets of verb roots in §3.3.2.

3.1.2. Methodology

The examples of (3.2) discussed above were uttered in the substitution frame *nàŋ nə _____ tə sa-ŋ* ‘You AGENTIVE _____ thus say-IMP’. The pairs of uninflected and inflected verbs compare the underlying mid tones of unmarked verb roots¹ to their tonal realizations in corresponding nominalized forms. The tones of all the data used to investigate tone sandhi were consistently established with reference to this substitution frame.

1. Recall that verb roots express past tense in their bare form.

A substitution frame is a useful tool for the elicitation of word lists, because it provides a standard of comparison against which the tones of any lexical item may be measured. Of course, there is a chance that the frame itself could create a tone sandhi environment that causes a tonal alternation on the target word. Nevertheless, some invariable yardstick is required for establishing the pitch levels of words in the corpus and this is provided by the mid tone of the agentive case marker, which remains constant regardless of what follows it in the frame. It is relevant to mention that tonal alternations were also sometimes found to occur on words following the target word in the frame. The nature of these perturbations and the tones of the words that triggered the alternations are discussed in §3.3.3.

When a lexical root is not modified by affixes, it is assumed that the tones heard on that lexical root as uttered in the substitution frame are its underlying tones. This is supported by a comparison of the output tones realized on the pairs of inflected and uninflected examples in (3.2). Once the underlying tones of a lexical item have been established, we are then in a position to interpret the tone sandhi processes applying to that lexical item in a methodical manner. Varying tonal realizations motivated by word formation and juxtaposition are examined in §3.3.

3.1.3. Grammatical tones

In addition to its major lexical marking function, tone is employed in a minor way to encode some grammatical categories in Mongsén. It is obligatorily used to help differentiate the comparative and superlative degrees in derived adjectives, and optionally to encode the agentive case when the head of the noun phrase is the first person singular pronoun ‘I’.

3.1.3.1. *Adjectives of degree*

Adjectives of degree are derived from stative verb roots expressing potentially gradable properties (see §7.4.3). Most stative verb roots require the nominalizing prefix *tə-* to derive a deverbal adjective (see §6.4.1, §7.2.4, §8.1.1.1 and §9.1.4); e.g. *tə-pə̀ti* (NZIP-be.big) ‘big’, *tə-asu?* (NZIP-be.small) ‘small’ and *tə-ləm* (NZIP-be.warm) ‘warm’. These express the comparative degree by additionally affixing the general nominalizer *-pà?* respectively to the right margin of the deverbal adjective derivation. The nominalizing prefix is underlyingly unspecified for tone and copies its output tone from an adjacent, tonally specified root syllable (see §3.2 for evidence).

- | | | | | | |
|----------|--|----|---|----|---|
| (3.4) a. | <i>təpətɪpaʔ</i>
tə-pətɪ-pàʔ
NZP-be.big-NR
'bigger' | b. | <i>tasupaʔ</i>
tə-asuʔ-pàʔ
NZP-be.small-NR
'smaller' | c. | <i>tələmpaʔ</i>
tə-ləm-pàʔ
NZP-be.warm-NR
'warmer' |
|----------|--|----|---|----|---|

The examples of (3.4) are illustrative of derivations involving roots with an underlying mid tone on their root-final syllables. The mid tone outputs on the penultimate and final syllables of the comparative derivation are also realized in derivations involving roots with underlying high tones, while the prefix retains the output high tone of the citation form e.g. *tə-nák* (NZP-be.black) 'black' → *tə-nak-paʔ* (NZP-be.black-NR) 'blacker', *tə-hláj* (NZP-be.long) 'long' → *tə-hlaj-paʔ* (NZP-be.long-NR) 'longer'. This consistently suggests that the output tone of a nominalizing prefix is initially assigned by the adjacent syllable of the verb root, which in the next cycle of comparative derivation may itself be subject to tone sandhi perturbations. If the general nominalizer does indeed have an underlying low tone, then it appears to be raised to mid tone by a penultimate syllable with a high or mid tone. One possibility is that the high tone of a root such as *vnák* 'be black' and the low tone of the general nominalizer reciprocally exert sandhi effects on each other when they are maximally different, resulting in a mid tone compromise for both of their syllables.

Stative verb roots that do not require the nominalizing prefix to derive a deverbal adjective, such as *áhná* 'be.near', from an adjective of degree simply by suffixing the general nominalizer *-paʔ* to the verb root and lowering its high tones, if any, to mid. This makes such comparative derivations formally identical to nominalized non-stative intransitive verbs functioning as relative clauses. There appears to be little semantic distinction between *ahna-paʔ* (be.near-NR) functioning as a comparative adjective meaning '(the) nearer (one)', and the identical form functioning as a headless relative clause meaning '(the one) that is near'. Both carry the implication of a comparison against a standard that need not be obligatorily stated, and both are inherently gradable. This semantic correlation is also applicable to the derivation of relativized attributes from stative verbs. These have an identical morphological structure to the examples of (3.4) – see §6.4.1 for further discussion.

Superlative adjectives are formed in either of two ways. One strategy employs the same nominalizing morphology that is used to form comparative adjectives, plus an intensifier suffix *-thi* is inserted next to the root; this is segmentally identical to a nominal intensifier suffix described in §7.5.1 and is

glossed accordingly.² The intensifier of the superlative differs in one important respect, though – it triggers an interesting additional tonal realization. When it is inserted between the verb root and the general nominalizer to derive the superlative, there is an accompanying tone change from a mid to a high tone on the final syllable of the adjacent verb root. This suggests that a floating high tone is part of its underlying representation in this derivation. The high tone does otherwise not surface when the intensifier is used as a noun modifier (see §7.5.1 for examples).

- (3.5) a. *təpətíthipa?*
 tə-pə̀ti-´ thi-pà?
 NZP-be.big-SUP.INTNS-NR
 ‘biggest’
- b. *tasúthipa?*
 tə-asu?´ thi-pà?
 NZP-be.small-SUP.INTNS-NR
 ‘smallest’
- c. *təlémthipa?*
 tə-ləm-´ thi-pà?
 NZP-be.warm-SUP.INTNS-NR
 ‘warmest’

An autosegmental analysis that attempts to explain the output high tone on the final syllable of the verb root might posit a floating tone that is associated with the intensifier suffix, i.e. -´*thi*. This could account for the realization of the high tone, brought about by the displacement of the root-final syllable’s mid tone. Tone spreading, then, could be represented by (3.6). The output mid tone on the general nominalizer is consistent with the assimilation in pitch height seen to occur with comparative adjectives.

- (3.6) $\begin{array}{ccc} \text{M} & \text{M} & \text{H} & \text{M} & \text{L} \\ | & | & & | & | \\ \text{tə-} & \text{ləm} & & \text{-thi} & \text{-pa?} \end{array} \Rightarrow \begin{array}{ccc} \text{M} & \text{M} & \text{H} & \text{M} & \text{M} \\ | & | & & | & | \\ \text{tə-} & \text{ləm} & & \text{-thi} & \text{-pa?} \end{array}$

A superlative adjective can alternatively be marked by the nominalizing prefix and high tones on the final syllable of the root and the general nominalizer alone, as there is no grammatical requirement to include the intensifier suffix for signalling the contrast between the superlative and comparative grammatical categories. The high tones realized on the final syllable of the root and the general nominalizer then carry the entire functional load for differentiating the superlative degree from the comparative degree.

2. Note that in addition to a floating high tone, -´*thi* has an underlying mid tone associated with its syllable when it is used in the superlative derivation, as opposed to a single low tone when it is used as a nominal intensifier.

- (3.7) a. *təpətípaʔ*
 tə-pəti-´-pàʔ
 NZP-be.big-SUP-NR
 ‘biggest’
- b. *tasúpaʔ*
 tə-asuʔ-´-pàʔ
 NZP-be.small-SUP-NR
 ‘smallest’
- c. *tələmpáʔ*
 tə-ləm-´-pàʔ
 NZP-be.warm-SUP-NR
 ‘warmest’

Some Mongsen speakers alternatively produce a low tone on the final syllable of the general nominalizer in superlative adjectives without a resulting change in meaning, e.g. *tə-ləm-pàʔ*. When differing tonal realizations are found to occur on the same reflex in different varieties of Mongsen, or even within the same variety, it is generally the final syllable of the word that demonstrates the aberrant tone. These inter- and intra-dialectal variations further complicate the task of accounting exhaustively for all tonal alternations by rule.

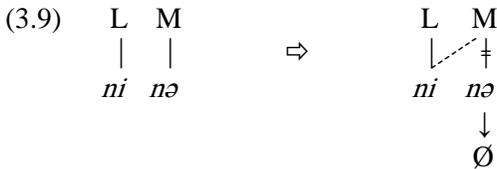
The autosegmental analysis of (3.6) requires some modification in order to account for the outright deletion of the segmental material of the intensifier suffix and its mid tone, and the final realization of a high tone on the general nominalizer. It is assumed that the loss of a syllabic segment also results in the de-linking of an associated tone, as represented by (3.8b). The floating high tone is then able to spread in a perseverative manner to the general nominalizer to displace its low tone, thereby deriving the output of (3.8c). The derivation of adjectives of degree is further described in §7.4.3.

- (3.8) a. $\begin{array}{ccc} M & M & H \\ | & | & \\ t\acute{e}- & l\grave{e}m & \\ | & | & \\ -thi & -paʔ & \end{array} \Rightarrow$ b. $\begin{array}{ccc} M & M & H \\ | & \dagger & \\ t\acute{e}- & l\grave{e}m & \\ | & | & \\ -thi & -paʔ & \end{array}$
- \downarrow
 \emptyset
- \Rightarrow c. $\begin{array}{ccc} M & M & H \\ | & \dagger & \\ t\acute{e}- & l\grave{e}m & -paʔ \end{array}$

Identical tonal realizations and word formation strategies are found in the superlative derivations of roots that do not require the nominalizing prefix in word formation, e.g. *ahna-thí-páʔ* ~ *ahna-thí-pàʔ* and *ahná-páʔ* ~ *ahná-pàʔ*, all of which express ‘nearest’.

3.1.3.2. *Agentive case marked by a rising tone*

A unique characteristic of the agentive case marker *nə* that sets it apart from all the other case-marking morphemes (described in §5.3) is that it can be marked solely by tone. The agentive case marker, like all the case-marking clitics, carries an underlying mid tone. However, when the head of its noun phrase is the first person singular personal pronoun *nì* ‘I’, it is possible to elide the segmental material of the case marker, leaving just its mid tone. The stranded tone then associates in a leftward direction to the adjacent syllable of the pronoun. This results in a low rising output, i.e. *nĩ*, realized phonetically as [ni¹³]. This is the only known instance of a lexical tone coalescence resulting in the creation of a contour tone in Mongsen. The tone spreading process is represented autosegmentally as follows:



Some irrefutable textual evidence of this segmental ellipsis and tone spreading cropped up while I was translating a recorded narrative with the narrator of that story, a Mangmetong Mongsen speaker. When repeating a sentence of his recorded narrative for me to transcribe, he unconsciously inserted the agentive case marker where there wasn’t one originally and pronounced the first person pronoun and the agentive case marker with their usual low and mid tones respectively. The original recording actually had a low rising tone on the first person pronoun *nĩ*, as demonstrated by (3.10).

- (3.10) *nàŋ la nĩ salí? ùŋlíku, tʃépa? tʃhàpà? nəni thən mətəm.*
 nàŋ la nĩ sa-lí? ùŋ-lí?-ku
 2SG TOP 1SG.AGT say-SIM see-SIM-LOC.CV
 tʃépa? tʃhà-pà? nə-ni thən mətəm
 what do-NR 2SG.POSS-wife COM copulate.PST
 ‘I told you so!’³ Why did you have sex with your wife?’

3. *salí? ùŋlíku* is a formulaic expression whose meaning, expressing ‘(I) told you so!’ is not predictable from the sum of its parts. It was used to admonish after a previous warning was given but subsequently ignored, with dire consequences (see Text 2, line 28).

When I asked about the absence of the agentive case marker in the original recording, the speaker explained that a low rising tone produced on the first person singular pronoun carries the same meaning as a first person singular pronoun pronounced with a low tone and overtly case-marked by an agentive marker carrying its independent mid tone. Subsequent testing confirms that the ellipsis of segmental material alone is not possible with any other nominal head, or even with any other personal pronoun. Another example with the first person singular pronoun occurs in the same text (see Text 2, line 76 for the full context).

- (3.11) *nàŋ, itáŋ kəmtaŋsikə táŋ nàŋ nə nǐ sapa? tʃu mələnsi əŋ liəŋ kátʃu.*
 nàŋ i-tə táŋ kəmtaŋ-si-əkə táŋ nàŋ nə
 2SG PROX-thus just be.careful-RPET-SIM just 2SG AGT
 nǐ sa-pà? tʃu mələn-si-əi táŋ li-əŋ kátʃu
 1SG.AGT say-NR DIST change-RPET-SEQ just stay-IMP PTCL
 “You just carefully change [it] again like this, as I told you.”⁴

Ellipsis of the agentive case-marker and tone spreading to a first person singular pronoun is noted to occur in the Khensa Mongsen, Mekhuli Mongsen and Mangmetong Mongsen varieties. Investigations reveal that it is not possible to elide the agentive case-marker and reassociate its tone to a first person singular pronoun in Waromung Mongsen, therefore the process may be confined to just the Mongsen varieties of the Ongpangkong range located in the south of the Mokochung district.

The loss of a grammatical category’s segmental representation has a number of implications for the analysis of the case-marking particles as clitics or grammatical words. A discussion of their status is taken up in §5.3.

3.2. Tonally unspecified prefixes

It is likely that the relational prefix *tə-* of kinship and body part terms, the non-relational prefix *a-* of other bound noun roots, the nominalizing prefix *tə-* and the vocative prefix *a-* are all underlyingly unspecified for tone and are assigned their output tones by an adjacent, tonally specified root syllable. Sometimes the nominal prefixes are pronounced with such short vocalic durations that it is a challenge to identify their output pitches, the reason being that their nuclei may only amount to a few glottal pulsations.

4. VERB-IMP *katʃu* is a formulaic expression encoding impatience and additionally imparts a warning. The particle *katʃu* may originate from a phonological fusion of the additive focus particle *kà?* ‘also’ + *atʃu* ‘that’.

The relational, non-relational and nominalizing prefixes were collectively referred to as nominal prefixes in Coupe (2003a: 21–24), but it is now apparent that their separate morphological categories should be acknowledged, given the different distributions and functions of these prefixes (see §7.2). The relational prefix *tə-* of body parts and kinship terms only occurs with a mid or high tone copied from an adjacent root syllable, principally because no nouns of these semantic classes have an underlying low tone on their root-initial syllables. The tonal realizations of the non-relational prefix *a-* are limited to mid and high tones for the same reason. The following examples demonstrate how the citation tones of the relational and non-relational prefixes coincide with the tones of adjacent root-initial syllables.

(3.12) *Relational prefix tə-*:

<i>tə-khət</i>	(RL-hand/arm)	‘hand, arm’
<i>tə-pázáʔ</i>	(RL-husband)	‘husband’
<i>tə-hɪə</i>	(RL-intestines)	‘intestines’
<i>tə-ti</i>	(RL-older.sibling)	‘older sibling’

(3.13) *Non-relational prefix a-*:

<i>á-sáʔ</i>	(NRL-meat)	‘meat’
<i>á-səŋ</i>	(NRL-wood)	‘wood’
<i>a-zu</i>	(NRL-hail)	‘hail’
<i>a-luŋ</i>	(NRL-stone)	‘stone’

The nominalizing prefix derives deverbal adjectives from stative verb roots and overwhelmingly occurs with high and mid tones, again because most stative verb roots have mid and high tones. The only exception is found on the verb root of *tə-pət̩i* (NZP-be.big) ‘big’. The low tone on the nominalizing prefix is consistent with the low tone on the root-initial syllable of this deverbal adjective, thus supporting the assumption that the prefix copies the tone of an adjacent syllable.

(3.14) *Nominalizing prefix tə-*:

<i>tə-kháʔ</i>	(NZP-be.bitter)	‘bitter’
<i>tə-tʃən</i>	(NZP-be.old)	‘old’
<i>tə-ɟa.ɟaʔ</i>	(NZP-be.rough)	‘rough’
<i>tə-səsə</i>	(NZP-be.smooth)	‘smooth’
<i>tə-pət̩i</i>	(NZP-be.big)	‘big’

The vocative prefix occurs with high, mid and low tones before roots with high, mid and low tones respectively on their initial syllables. Given this distribution, it would appear that the vocative prefix is similarly unspecified for tone and thus copies the tone of an adjacent root syllable in the output. The following examples are illustrative of the relationship between the tones of the vocative prefix and the root-initial syllable:

(3.15) *Vocative prefix a-*:

<i>á-pánuzapa?</i>	(VOC-father's.younger.brother)	'Uncle!'
<i>á-jázəni</i>	(VOC-mother's.older.sister)	'Aunt!'
<i>a-ja</i>	(VOC-mother)	'Mother!'
<i>a-u</i>	(VOC-father's.father)	'Grandfather!'
<i>à-tì</i>	(VOC-wife's.mother, same phratry)	'Mother-in-law!'
<i>à-zə</i>	(VOC-father's.mother)	'Grandmother!'

A tonal alternation affects the roots of the vocatives *àtì* 'Mother-in-law!' and *àzə* 'Grandmother!' in (3.15), because their citation tones with the relational prefix are mid-mid, e.g. *tə-tì* (RL-mother.in.law) 'mother-in-law' and *tə-zə* (RL-father's.mother) 'grandmother'. If the vocative prefix is in fact underlyingly toneless, then it remains to be determined why the tone of the lexical root undergoes an idiosyncratic change of tone, depending upon the prefix type. A possible solution that requires further investigation is that there is a floating low tone associated with the lexical roots of some kinship noun roots, and that it is only realized in the presence of certain affixes.

3.3. **Tone sandhi**

Tonal alternation is known to occur in Tibeto-Burman languages of the north-east India/Burma region, e.g. Tiddim Chin (Henderson 1965), Mpi (Bradley 1979), Maru (Okell 1989), Khezha (Kapfo 1989) and Gong (Bradley 1992). While it is now the established practice to analyze the word as the domain of tone in tonal Bodic languages, the syllable is still assumed to be the prosodic domain of tone for most languages outside of the Bodic branch. Yet Mazaudon (1977: 76–78) draws attention to the poor fit of a monosyllabic domain for tone in a number of Tibeto-Burman languages. She suggests that tone assignment may in fact be based upon a polysyllabic pattern in a number of languages and recognizes the need for a typological study to determine the extent to which languages with a polysyllabic pattern of tone assignment are exceptional. Unfortunately, nearly thirty years after her suggestion was made, we still have little idea of even the basic tonal systems of the overwhelming majority of

Tibeto-Burman languages, let alone an understanding of the function of tone in polysyllabic units.

Complex tonal alternations occurring both within morphologically complex words and between adjacent words independently of constituency suggest that Mongsen belongs to that group of languages in which a polysyllabic pattern of tone assignment may apply. An exhaustive analysis of the tone sandhi patterns of Mongsen is a formidable task and beyond the scope of this grammatical description. In lieu of a comprehensive solution, I offer a preliminary outline of the problem, identify some basic recurring patterns, and suggest some possible approaches that might provide a foundation for subsequent research.

3.3.1. Tonal combinations in phonological words

Low, mid and high tones occur on both smooth and checked syllables.⁵ While all nine combinations of tones on adjacent syllables are possible, maximal differences in pitch tend to occur infrequently within lexical roots.

The following examples demonstrate all the attested combinations of level tones occurring within phonological words, e.g. *tʃànàʔ* ‘dirt’ (L-L), *ikhu* ‘fence’ (M-M), *phátʃhóm* ‘gums’ (H-H), *ùjim* ‘raquet-tailed drongo’ (L-M), *tʃháj-pàʔ* (play-NR) (H-L), *təmáj* ‘all’ (M-H) and *səŋtʃàŋ* ‘fruit’ (M-L). The combinations of L-H and H-M are not found to occur in disyllables, but are attested in trisyllabic words, e.g. *tʃhàmà-páʔ* (disappear-NR) (L-L-H) and *tə-jáŋi* (NZP-be.fitting) (H-H-M).

3.3.2. Tone sandhi generated by morpheme concatenation

The brief reference to tone sandhi in §3.1.1 demonstrated the effect of the general nominalizer *-pàʔ* on an adjacent syllable in a set of verb roots with underlying mid tones. We will now resume the problem of dealing with tone sandhi perturbations triggered by this suffix, after which we will turn to an examination of other verbal morphemes that activate internal tone sandhi in word formation processes.

The following examples demonstrate some melodic patterns that can occur on nominalized verb stems in Mongsen and compare these to the underlying tones associated with their uninflected roots. The first set has an underlying low tone on uninflected monosyllabic roots and consistently realizes L-L patterns on nominalized stems, with one exception: the intransitive verb root *ap* ‘rot’ has a mid tone in the deverbal adjectival form *tə-ap*, phonetically [tap³³]. I assume

5. Checked syllables terminate in a stop; all other syllable types are smooth.

that the tone associated with its root is underlyingly mid, and that the low tone of the general nominalizer causes it to assimilate in height. Since its citation form is morphologically distinct from the other roots of (3.16), it may belong to a different stem class. This would explain its aberrant tonal alternation in the citation form.

(3.16) L > L-L

<i>nəp</i>	(adhere.PST)	<i>nəp-pà?</i>	(adhere-NR)
<i>sù?</i>	(bear.PST)	<i>sù-pà?</i>	(bear-NR)
<i>jàk</i>	(beat.PST)	<i>jàk-pà?</i>	(beat-NR)
<i>mù?</i>	(bite.PST)	<i>mù-pà?</i>	(bite-NR)
<i>pù?</i>	(carry.on.back.PST)	<i>pù-pà?</i>	(carry.on.back-NR)
<i>hlì</i>	(buy.PST)	<i>hlì-pà?</i>	(buy-NR)
<i>tə-ap</i>	(NZN-rot)	<i>àp-pà?</i>	(rot-NR)
<i>tù?</i>	(dig.PST)	<i>tù-pà?</i>	(dig-NR)

The second set of verbs has underlying mid tones on its uninflected monosyllabic roots; these consistently realize mid tones on their nominalized forms. They clearly must belong to a stem class different from the M-M roots of (3.2), because the nominal inflection does not result in the H-L output that was seen, for example, in *tʃháj-pà?* (play.sport-NR).

(3.17) M > M-M

<i>hən</i>	(take.PST)	<i>hən-pa?</i>	(take-NR)
<i>phu</i>	(blow.PST)	<i>phu-pa?</i>	(blow-NR)
<i>ɲuŋ</i>	(burn.PST)	<i>ɲuŋ-pa?</i>	(burn-NR)
<i>pu</i>	(carry.on.shoulder.PST)	<i>pu-pa?</i>	(carry.on.shoulder-NR)
<i>phuŋ</i>	(bloom.PST)	<i>phuŋ-pa?</i>	(bloom-NR)
<i>ɲəm</i>	(bury.PST)	<i>ɲəm-pa?</i>	(bury-NR)
<i>tʃa</i>	(call.PST)	<i>tʃa-pa?</i>	(call-NR)

The consistent mid tone output on the general nominalizer suggests that this stem class of verb causes an assimilation in the height of the suffix's output tone.

Some unexpected sandhi perturbations result when the irrealis suffix is added to some nominalized stems, e.g. *azə-pa?* (dare-NR) has a M-M-M pattern, while *azə-í-pà?* (dare-IRR-NR) has the M-H-L pattern seen in all but one of the nominalized forms of (3.2). Also, the M-M tones of *hən-pa?* (take-NR) change to H-H-H when the perfective suffix is included in the stem, e.g. *hən-tʃúk-pá?* (take-PFV-NR), while the LL tones of *àp-pà?* (rot-NR) change to M-M-M with the addition of the perfective, e.g. *ap-tʃuk-pa?* (rot-PFV-NR). Even the addition

of the putative atonal nominalizing prefix can have striking tone sandhi repercussions in some nominalized stems; for example, a L-M-L pattern is found in the nominalization *nəm-phəŋ-pàʔ* (compress-COVER-NR) when it is used as a relativized attribute, yet the corresponding instrument nominalization *tə-nəm-pháŋ-pàʔ* (NZP-compress-COVER-NR) that can function as a noun phrase head has a M-M-H-L pattern.

Slight variations based on the M-H-L pattern presented in (3.2) are found in verb stems with other suffixing morphology. Unfortunately the analysis is limited by the small corpus of examples that show the M-H-L pattern, but this is sufficiently pervasive to raise the suspicion that specific melodic patterns might apply to the whole of the stem in a systematic manner, e.g.: *laŋli* (reply.PST) ~ *laŋli-təp* (reply-RECIP.PST) ~ *laŋli-təp-i-pàʔ* (reply-RECIP-IRR-NR); *laŋ-saʔ* (cut-SEPARATE.PST) ~ *laŋ-sá-j-pàʔ* (cut-SEPARATE-IRR-NR); *məhnən* (digest.PST) ~ *məhnən-saʔ* (digest-SEPARATE.PST) ~ *məhnən-sá-pàʔ* (digest-SEPARATE-NR); and *ləm-thàŋ* (divide-SEVER.PST) ~ *ləm-tháŋ-pàʔ* (divide-SEVER-NR).

The array of tones realized on the general nominalizer in numerous examples of this chapter serves to demonstrate that it is no trivial matter to establish the underlying tones of affixes, since these often participate in variable tone sandhi perturbations triggered by word formation. The output tones of verbal prefixes are the least problematic for analysis, due to their mostly predictable tonal realizations. For example, the negative prefix *mə-* and the prohibitive prefix *tə-* both have a low tone when they inflect verb roots that have a non-high tone on their initial syllable. When prefixed to a root with an initial high tone, the prefix has a mid tone. This appears to be consistent across all stem classes of verbs (see Coupe [2003a: 18–20] for examples).

Variable tonal alternations can result from word formation. For affixes such as the irrealis marker, the optional declarative mood marker and the present tense marker, a range of realizations might be compared and an affix's underlying tones determined by considering their patterns of realization in various sandhi environments. The paradigms of Table 3.1 below serve to demonstrate the potentially idiosyncratic nature of tonal realizations on a selection of verb stems.

It would appear from this preliminary investigation that a comprehensive analysis of internal tone sandhi in Mongsen must entail the recognition of stem classes. This should be able to account for variations in tonal realizations on the same suffixes when they occur with different verb stems.

Table 3.1. Examples of tonal realizations on selected verbal suffixes

root	gloss	-IRR -ì	-IRR-DEC -ì-ù?	-PRES -(ə)ɪ	-PRES-DEC -(ə)ɪ-ù?	-PST-DEC -Ø-ù?
<i>ajim</i>	‘announce’	<i>ajim-ì</i>	<i>ajim-ì-ù?</i>	<i>ajim-əɪ</i>	<i>ajim-ɪ-ù?</i>	<i>ajim-ù?</i>
<i>apak</i>	‘be flat’	<i>apak-ì</i>	<i>apak-ì-ù?</i>	<i>apak-əɪ</i>	<i>apak-əɪ-ù?</i>	<i>apak-ù?</i>
<i>aɪəm</i>	‘bear’	<i>aɪəm-i</i>	<i>aɪəm-i-u?</i>	<i>aɪəm-əɪ</i>	<i>aɪəm-ɪ-u?</i>	<i>aɪəm-ù?</i>
<i>atsə</i>	‘look’	<i>atsə-i</i>	<i>atsə-i-u?</i>	<i>atsə-ɪ</i>	<i>atsə-ɪ-ù?</i>	<i>atsə-ù?</i>
<i>aɪu</i>	‘be good’	<i>aɪu-i</i>	<i>aɪu-i-u?</i>	<i>aɪu-ɪə</i>	<i>aɪu-ɪ-ù?</i>	<i>aɪ-u?</i>
<i>zɑŋlu</i>	‘make’	<i>zɑŋlu-i</i>	<i>zɑŋlu-i-u?</i>	<i>zɑŋlu-ɪ</i>	<i>zɑŋlu-ɪ-u?</i>	<i>zɑŋl-u?</i>
<i>asà</i>	‘shout’	<i>asà-ì</i>	<i>asà-ì-ù?</i>	<i>asà-ɪ</i>	<i>asà-ɪ-ù?</i>	<i>asà-ù?</i>
<i>hənsət</i>	‘kill’	<i>hənsət-ì</i>	<i>hənsət-ì-ù?</i>	<i>hənsət-əɪ</i>	<i>hənsət-əɪ-ù?</i>	<i>hənsət-ù?</i>
<i>həmət</i>	‘hold’	<i>həmət-ì</i>	<i>həmət-ì-ù?</i>	<i>həmət-əɪ</i>	<i>həmət-əɪ-ù?</i>	<i>həmət-ù?</i>
<i>həmsi</i>	‘use’	<i>həmsə-i</i>	<i>həmsə-i-ù?</i>	<i>həmsi-ɪ</i>	<i>həmsi-ɪ-ù?</i>	<i>həmsi-u?</i>
<i>hən</i>	‘take’	<i>hən-i</i>	<i>hən-i-u?</i>	<i>hən-əɪ</i>	<i>hən-əɪ-u?</i>	<i>hən-ù?</i>
<i>kəm</i>	‘become’	<i>kəm-i</i>	<i>kəm-i-u?</i>	<i>kəm-əɪ</i>	<i>kəm-əɪ-u?</i>	<i>kəm-ù?</i>
<i>jùk</i>	‘sell’	<i>jùk-ì</i>	<i>jùk-ì-ù?</i>	<i>jùk-əɪ</i>	<i>jùk-əɪ-ù?</i>	<i>jùk-ù?</i>
<i>kəp</i>	‘shoot’	<i>kəp-ì</i>	<i>kəp-ì-ù?</i>	<i>kəp-əɪ</i>	<i>kəp-əɪ-ù?</i>	<i>kəp-ù?</i>

3.3.3. Tone sandhi generated by tones of adjacent words

I mentioned in §3.1.2 that the substitution frame (i.e. *nàŋ nə ___ tə sa-ŋ* ‘You AGENTIVE ___ thus say-IMP’) used to determine the lexical tones of words was itself sometimes subject to tonal alternations. The tone on the final syllable of the target word potentially affected the tone of the quotative particle *tə*, which concomitantly affected the tone realized on the final verb *sa-ŋ* (say-IMP). The underlying tones on these two words when uttered in isolation are low and mid respectively.

Examples of the words that triggered different tonal realizations on the quotative *tə* ‘thus’ and imperative verb form *saŋ* ‘say!’ are listed in Table 3.2. A consistent pattern is recognizable in the tonal alternations demonstrated by the first column of words. All have mid tones on their final syllables, triggering aberrant mid and low tone realizations respectively on the quotative particle and the imperative verb form. The second column is less consistent. The first five words have a mid tone on their final syllables, yet this does not cause the low tone of the quotative to assimilate in height; consequently, there is no ancillary effect on the tone of the following verb. Similarly, a low or high tone in the final syllable of the target word consistently results in the usual L-M tonal realization on the quotative and imperative verb respectively. Words with high or low tones on their final syllables thus have no effect on the underlying tones

of the quotative and verb, whereas those with mid tones on their final syllables have a variable tone sandhi effect on the following words of the substitution frame.

Table 3.2. Examples of words causing tonal alternations in the substitution frame

<i>tə səŋ</i> (aberrant M-L pattern)		<i>tə saŋ</i> (unaffected L-M pattern)	
<i>səmtsə</i>	‘ask.PST’	<i>sisə-i-paʔ</i>	‘arise-IRR-NR’
<i>juŋkəm</i>	‘river bank’	<i>məntʃən</i>	‘chair’
<i>məluk</i>	‘basket’	<i>tə-sin</i>	‘NRL-back’
<i>hjaŋkhu</i>	‘bag’	<i>li-ɿ</i>	‘stay-PRES’
<i>hən-paʔ</i>	‘take-NR’	<i>hluli</i>	‘bean’
<i>tʃhàktʃhà-paʔ</i>	‘get.up-NR’	<i>a-hwáʔ</i>	‘NRL-bamboo.sp.’
<i>phu-paʔ</i>	‘blow-NR’	<i>tə-íá-j-pàʔ</i>	‘NZIP-come-IRR-NR’
<i>tə-ɿt</i>	‘RL-bone’	<i>səŋkəp</i>	‘bark of tree’
<i>upu</i>	‘ashes’	<i>tsəŋ-i-pàʔ</i>	‘bark-IRR-NR’
<i>tə-nən</i>	‘NZIP-be.green’		

Taking into account all of the data of Tables 3.1–3.2, it is not yet understood why the final mid tones of most but not all words trigger tone sandhi. Since external tone sandhi cannot always be predicted by rule, the most workable solution to this analytical problem is to recognize stem classes of verbs and divide them according to the tonal alternations they trigger, as advocated above for the analysis of internal tone sandhi. Such an approach lies outside the scope of the present work, however.

3.3.4. Tone sandhi in nominal compounds

Idiosyncratic tonal alternations are often found in nominal compounds. The following words compare tones on nouns in isolation to their realizations in noun compounds formed with *á-sáʔ* (NRL-meat) ‘meat’ as the head of the compound noun. Some combinations of nominal roots allow variable tonal realizations on their compound heads. Where this is found to be the case, both forms are presented.

- (3.18) *á-úk* ‘NRL-pig’ + *á-sáʔ* → *áwk-sàʔ* ‘pork’
a-hən ‘NRL-fowl’ + *á-sáʔ* → *ahən-sàʔ* ‘chicken’
alternative → *ahən-saʔ* ‘chicken’
màsəʔ ‘cow’ + *á-sáʔ* → *màsə-sàʔ* ‘beef’
phàtàk ‘duck’ + *á-sáʔ* → *phàtàk-saʔ* ‘duck meat’

<i>a-khu</i>	‘NRL-tiger’	+	<i>á-sá?</i>	→	<i>akhu-sà?</i>	‘tiger meat’
			alternative	→	<i>akhu-sa?</i>	‘tiger meat’
<i>mətshə</i>	‘deer’	+	<i>á-sá?</i>	→	<i>mətshə-sà?</i>	‘venison’
			alternative	→	<i>mətshə-sa?</i>	‘venison’

The examples of (3.18) demonstrate that it is not entirely possible to predict the output tones on the compound head. Note that the compound heads of ‘chicken’, ‘tiger meat’ and ‘venison’ allow either a low tone or a mid tone to be realized, but the other nominal compounds do not allow this variation, e.g. **awksa?* and **masəsa?*. Output tones on these compound nouns can change again when they are exposed to tone sandhi environments created by adjacent words in clauses, as in (3.19).

- (3.19) *mətshəsá? thən tʃuŋù?*
 mətshə-sá? thən tʃuŋ-ù?
 deer-meat COM eat.meal.PST-DEC
 ‘[I] had venison with my meal.’

The phenomenon of tone sandhi is so extensive and complex in Mongsen that it would make a worthy topic of investigation by itself. Much more work is required to account for all the fascinating sandhi perturbations presented by this language, but the topic will not be explored to a deeper level in this book so that attention can instead be given to the description of the grammar.

3.4. Postlexical intonation

Intonation is an important prosodic feature of Mongsen. Specific patterns are superimposed postlexically over the contrastive tones of individual syllables in lexical items and grammatical morphemes to encode different types of prosodic information. This can result in the underlying tones of those words becoming significantly distorted by its imposition.

Intonation functions as a boundary signal to delineate phrasal and clausal constituents in the prosodic hierarchy. The most commonly occurring and distinctive pattern is the clausal intonation that marks a dependent sequential converb clause (see §11.4.1.1 for a description of its syntax). However, even matrix clauses can be produced with a clause-final rising intonation if the speaker wishes to indicate that more information relevant to the same discourse paragraph is to follow. At a lower level in the prosodic hierarchy, intonation is used to mark off the right edge of noun phrases. Examples of each type of intonation pattern are described in the following sections and demonstrated by waveforms and F_0 traces.

3.4.1. Phrasal intonation

The intonation contours that delineate the right edges of noun phrases in Mongsen discourse create the initial impression that the language has lexical contour tones instead of level tones. However, the elicitation of words in isolation demonstrates that the rising and falling intonation contours that so characterize the language in narrative discourse are actually not part of the lexical tone system. Figure 3.1 presents a waveform and F_0 trace showing the pitch contours realized on the noun phrases of example (3.20). The phrasal intonation is signalled in the following italicized transcriptions by an upward arrow, the final intonation by a downward arrow.

- (3.20) *nì imtʃən tʃâɪ ↗, maŋmətuŋ niŋəɪ ↗, təɪ ↗ nì tənij ↗ imtiluìn ↘.*
 nì imtʃən tʃâ-əɪ maŋmətuŋ niŋ-əɪ
 1SG clan.name offspring-ANOM village.name name-ANOM
 tə-əɪ nì tə-nij imti.lùìn
 thus-SEQ 1SG RL-name personal.name
 ‘I am a descendent of the Imchen clan, a Mangmetong villager, and my name is Imti Luin.’

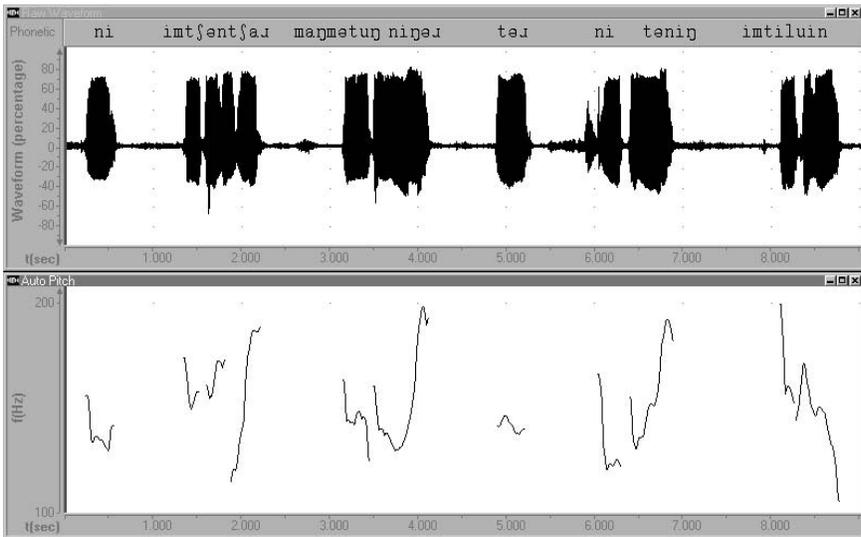


Figure 3.1. Waveform and F_0 trace demonstrating the effects of phrasal intonation

The speaker used this complex sentence to introduce himself prior to narrating a story. With the exception of the first and last noun phrase, the right edges of all the other noun phrases are prosodically marked by a steep phrase-

final rise in intonation that totally dominates and distorts the underlying level tones of the syllables affected. The function of the high rising intonation at the end of each noun phrase is to signal that more information relevant to that particular chunk of discourse is to follow. The end of the intonation phrase is marked by a steep fall, which can be seen on the speaker's name *imti lùn* 'Imti Luin'. This signals prosodically that the end of the discourse paragraph has been reached and warns the audience to expect a new topic to be established.

3.4.2. Boundary marking intonation on dependent clauses

Converbs are a type of non-finite verb form used to mark dependent clauses. They express a range of presuppositional and non-presuppositional meanings, such as condition, concession, iterative activity or state, sequential action, and the manner or attendant circumstances of an activity (see §11.4 for their full description). An important part of the phonological representation of the sequential converb, the most frequently occurring of these non-finite verb forms, is a characteristic prosodic contour.

- (3.21) *maŋmətuŋ ajim kəm-əl* ↗
 maŋmətuŋ a-jim kəm-əl
 village.name NRL-village become-SEQ
 'Having founded Mangmetong village, ...'

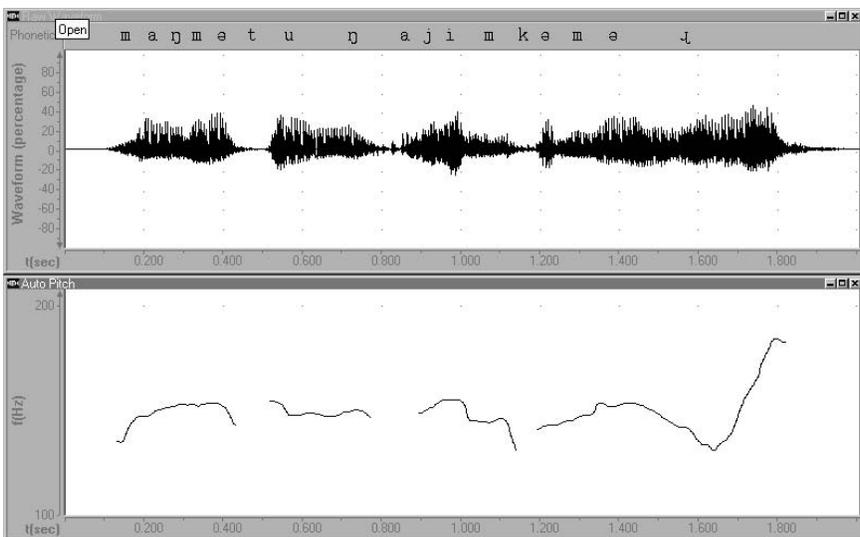


Figure 3.2. Waveform and F_0 trace showing the intonation of the sequential converb

A distinctive intonation pattern marks the right edge of a sequential converb clause and extends prosodically over the whole of the dependent clause. Figure 3.2 illustrates the dramatic rise in pitch that is realized on the converb in the example of (3.21) above. The F_0 contour clearly shows an initial dip on the converb form *kəm-əɪ* (become-SEQ), followed by a significant rise in pitch, which is conveyed almost entirely by the retroflex approximant coda. This is the usual presentation when a pause follows the dependent clause and the speaker has time to prolong the rhyme to convey the rising contour. In rapid conversation, when the rhyme cannot be extended, the right boundary of the sequential converb clause tends to be marked by just a dipping pitch. This can be seen to extend from 1.4 sec to 1.65 sec in Figure 3.2 immediately prior to the sudden rise at the end of the clause.

A verb stem marked by the sequential converb suffix, such as that of (3.21), cannot be directly negated by simply attaching the negative prefix to the converb form. Instead, the discontinuous negating morphology of past tense matrix verbs is used, i.e. *mə-...-la* (NEG-...-NEG.PST), but with one crucial difference: it usually carries a high or rising intonation to indicate that it is a non-final clause (see §11.4.2). This non-final status is marked by the upward arrow in the italicized transcription of (3.22). To achieve the rising intonation, the vowel of the negative suffix may be slightly lengthened to convey the pitch of the contour. The effect of negation on the morphological form of sequential and simultaneous converbs is discussed in greater detail in §11.4.2.

- (3.22) *aki tʃu nə məɪuŋla ↗ litʃhà \.*
 a-ki tʃu la mə-ɪuŋ-la li-tʃhà
 NRL-house DIST TOP NEG-burn-NEG.NF be.COP.PST
 ‘As for the houses, [they] were not burnt down.’
 (Lit: ‘As for the houses, [they] existed without being burned down.’)

3.4.3. Non-final postlexical intonation on matrix clauses

Speakers use postlexical intonation on the clause-final constituent of a matrix clause to signal that more information applicable to the discourse paragraph is to follow, and that the established topic is still relevant. This is realized prosodically as a rising pitch on the final element of the intonation phrase.

- (3.23) “*pi təpaʔ nə sùpàʔ ìla tuku táŋ nuŋla ↗ ...*”
 pi tə-paʔ nə sùpàʔ ìla tuku táŋ
 PROX RL-father INST bear-NR 3PL.IN nine just
 nuŋ la
 NEG.PTCL TOP
 “It is not just nine of us begotten by father ...”

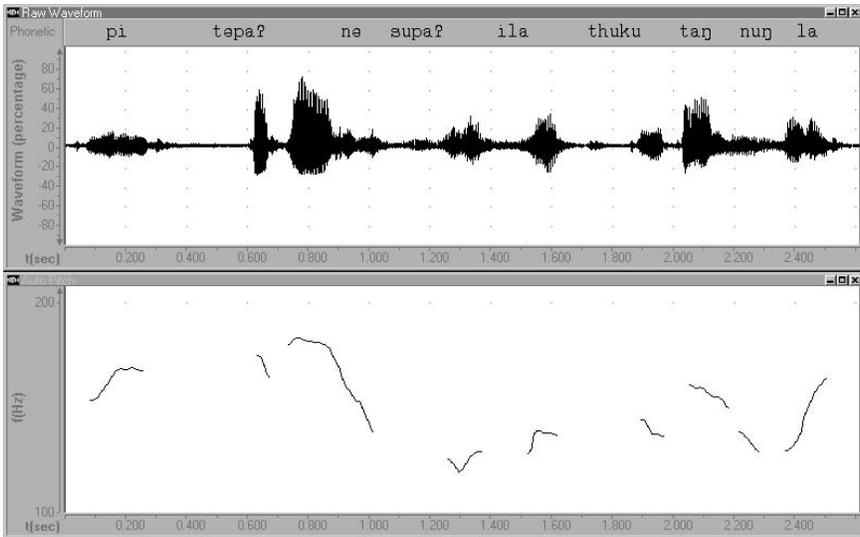


Figure 3.3. Waveform and F_0 trace showing non-final rising intonation on *la* (TOP)

The non-final intonation dominates the lexical tones associated with the syllables over which the prosodic intonation is realized. Figure 3.3 above demonstrates the F_0 contour of the non-final intonation pattern realized on the topic particle *la* in the verbless clause complement of (3.23), signalled by an upward arrow. In the absence of the effects of prosodic intonation, the citation tone of the topic particle is mid, yet it is observed to rise steeply in the F_0 trace under the influence of intonation.

3.5. The glottal stop prosody

A glottal stop is frequently encountered in the underlying representation of verbs, nouns and nominal and verbal morphology. Because of its highly restricted pattern of distribution and unique behaviour under the effect of morpheme concatenation, the glottal stop cannot justifiably be considered a segmental phoneme. For instance, unlike any other consonant phonemes, it is

deleted when occurring in the word-internal environment before a syllable boundary. This constrains it to a word-final realization, as demonstrated by the phonetic outputs of following pairs of words.

(3.24) <i>asi?</i> (deceive.PST)	[a ³³ ʃi ¹¹]
<i>asi?-juk</i> (deceive.PFV.PST)	[a ³³ ʃi ¹¹ juk ³³]
<i>jàksa?</i> (break.PST)	[jak ¹¹ sa ²³³]
<i>jàksa?-i</i> (break.IRR)	[jak ¹¹ saj ³³]
<i>tʃã?</i> (consume.PST)	[tʃã ¹¹]
<i>tʃã?-mà?-juk</i> (consume-CMPL-PFV.PST)	[tʃã ³³ ma ³³ juk ³³]
<i>təp-sa?</i> (strike-SEPARATE.PST)	[təp ¹¹ sa ²³³]
<i>təp-sa?-tʃuk-i?</i> (strike-SEPARATE-CAUS-PST)	[təp ³³ sa ³³ tʃuk ³³ i ²³³]
<i>á-hŋá?</i> (NRL-fish)	[a ⁵⁵ hŋa ²⁵⁵]
<i>hŋakəp</i> ‘fish scale’ (Lit. ‘fish-skin’)	[hŋa ³³ kəp ³³]
<i>a-na?</i> (NRL-honey bee)	[a ³³ na ²³³]
<i>nakùtsə</i> ‘honey’	[na ³³ ku ¹¹ tsə ³³]

The word-final realization of the glottal stop sometimes results in what superficially appears to be contrastive distribution, since it contrasts with \emptyset and the voiceless unaspirated velar stop in this position e.g. *jà* (shine.PST), *jàk* (beat.PST) and *jà?* (hear.PST). Yet a closer examination of its phonotactic distribution clearly shows that any presumed contrastive function, such as that suggested by these examples, is entirely contingent upon the fortuitous word-final alignment of *k* and \emptyset with this prosodic element. Because of this, I see little justification for according it the status of a segmental phoneme. The autosegmental and incidental contrastive segmental functions of the glottal stop are discussed in greater detail in Coupe (2003a: 24–27).

The glottal stop, then, might conveniently be thought of as a prosodic element occupying an independent tier in the prosodic hierarchy of the word and simultaneously having an association with the skeletal tier. This accounts for both its word-final realization and contrastive segmental function. The word-internal deletion before a syllable boundary and word final distribution is handled by a word formula and glottal stop deletion rule originally suggested in Coupe (2003a: 27), repeated below in (3.25).

(3.25) $\omega = \sigma(\sigma^*)(?)$	where:
Deletion rule: $? \rightarrow \emptyset / _ \sigma [$	$\omega =$ word
	$\sigma =$ syllable ⁶

6. Words are rarely found to be longer than six syllables in length in naturally spoken language (see §8.2).

Chapter 4

Word classes

The word classes of Mongsen can be divided into open and closed classes, as follows:

Open classes

1. Verbs
2. Nouns

Closed classes

3. Pronouns: personal, possessive, demonstrative, interrogative, indefinite
4. Nominal modifiers: nominal deictics, quantifiers, case marking clitics
5. Time words
6. Adverbs
7. Discourse connectives
8. Phrasal conjunction
9. Interjections, exclamations and onomatopoeia
10. Particles and clitics

These can be differentiated according to formal criteria, commencing with the open word classes.

4.1. Open word classes

The open word classes of Mongsen are restricted to nouns and verbs. Members of the open classes are formally identified by the types of inflectional and derivational morphology that are able to occur on their stems, and by their distribution in clauses. Only verbs can function as the heads of verbal clauses; only nouns and pronouns can function as the heads of noun phrases.

4.1.1. Verbs

The lexical class of verb prototypically denotes actions, processes or states. Formally, any word belonging to this word class can occur within the discontinuous past tense negative marker *mə̌-...-la*, e.g. *mə̌-wa-la* (NEG-go-NEG.PST) ‘didn’t go’. This alone consistently distinguishes verbs from all other clausal constituents. Other criteria useful for their recognition are their ability to inflect for tense, aspect and modality, and to accept valency-modifying morphology on their stems, such as the causative marker and the reciprocal/

The notion of transitivity is of limited utility for the classification of Mongsen verbs. This is because there are interesting complications in the marking of grammatical functions that tend to operate outside of verb transitivity classes (see §5.2.2). Furthermore, virtually any transitive verb that is able to occur with two core arguments can also be used with just one core argument that corresponds to the O argument of a transitive clause.¹ Some of these qualify as patientive ambitransitive verbs; others that are not genuinely ambitransitive require the antecedent mention of an argument in order to license its pragmatically recoverable elision. A limited number of verbs can function as agentive ambitransitives, in which the sole core argument corresponds to the A argument of a transitive clause. Verbs with omitted core arguments (as well both patientive and agentive ambitransitives) occur without any formal marking of a change in the valency status of the verb (see §5.2.4 and §8.1.3 for a definition of ambitransitivity and further discussion).

While the difference between transitive and intransitive verbs can be formally indistinct, it is possible to distinguish between two functional subclasses of intransitive verbs on the basis of their meanings when their stems are reduplicated. Reduplicated stative verbs express intensified meanings, whereas reduplicated activity verbs express durative meanings (see §8.1.4). The valency of any semantically compatible verb can be increased by causativization or in principle reduced by the reciprocal derivation (§5.4). This freely applies to most activity verbs and, in a more limited manner, to stative verbs.

4.1.2. Overview of the noun word class

The defining characteristic of all members of the noun class is their ability to function as heads of noun phrases. Prototypical nouns are notionally identifiable as lexical items referring to entities that have physical existence in space and time. Mongsen nouns referring to entities with less time-stability, such as abstract nouns, are overwhelmingly derived from verbs using nominalizing morphology (see §7.4 and §8.3.4).

Quantification is a useful criterion for distinguishing between personal names, place names, pronominals and vocatives on the one hand, and all the various sub-classes of common nouns on the other, because quantification only applies to common nouns, e.g. *a-mi? təmáj* (NRL-person all) ‘all (the) people’

1. Core arguments of the clause are represented schematically in this work by A for the agent-like argument of a transitive verb, S for the single argument of an intransitive verb, and O for the patient-like argument of a transitive verb. Justification for using these terms in place of “subject” and “object” is presented in the description of core grammatical marking in §5.3.

and *tə-muwa-j phəli* (NZIP-bless-IRR four) ‘four blessings’. Mass nouns, such as *a-tsə* (NRL-water) ‘water’, cannot be directly quantified by numerals and require a separate nominal that functions as an *ad hoc* classifier, e.g. *a-tsə mə̀rùk a-nət* (NRL-water cup NRL-two) ‘two cups of water’. Nominal quantifiers are described in §4.2.7.2 and §4.2.7.3.

Sub-classes of noun are further distinguished by the grammatical categories they express. The diminutive suffix *-za* [< PTB **za* ‘child’ STC #59] occurs stem-finally on common nouns, e.g. *a-hən-za* (NRL-chicken-DIM) ‘chick of domestic fowl’, and is a fossilized formative in the stems of some kinship terms, e.g. *tə-pázá?* (RL-husband) ‘husband’ (lit. ‘little father’) (see §7.5.2). The masculine and feminine semantic gender markers *-pà?* [<PTB **pa* ‘father’ STC #24] and *-la* are restricted to noun phrases with human referents and people’s names, e.g. *a-mi?-pa?* (NRL-person-M) ‘a man’, *tsəmà?-la-əl* (Assam.plain-F-ANOM) ‘a plainswoman, Indian woman’, *lítjá-pà?* (PN-M) ‘Lichaba’ and *tsə̀ŋrutshə-la* (PN-F) ‘Tsengrutsela’, and animals personified in the context of folklore stories, e.g. *tshə̀luŋ-la* (fox-F) ‘Fox’. By virtue of this morphological distinction, nouns with human or personified referents are recognizable as a discrete subclass of nominal. The occurrence of a lexical root in constituency with a case marking clitic is not infallibly diagnostic of a nominal in Mongsén, due to the fact that some case marking clitics have been reanalyzed as synchronically homophonous converb markers and can occur with verbal roots (see §11.4). Nevertheless, the ability to occur in constituency with a case marking clitic is another property shared by the noun class.

Table 4.1. Summary of nominal sub-classes that function as NP heads

NOMINALS	OPEN CLASSES	Bound nouns	Relational nouns	Kinship terms
				body part terms
				Vocatives
		Non-relational common nouns		
	Unbound nouns	Common nouns		
		Personal names and place names		
	CLOSED CLASSES	Pronominals	Personal pronouns	
			Possessive pronouns	
			Demonstrative pronouns	
			Interrogative pronouns	
Indefinite pronouns				

Table 4.1 above summarizes the classification of words that can function as nominal heads in Mongsen. Formal criteria for the recognition of various types of nominals are set out in Table 4.2 below.

Table 4.2. Formal properties of nominals (+H/P = human/personified referent only)

NOMINAL CATEGORY		Head of a NP	Quantifiable	Occurs with a demonstrative	Occurs with a case marker	Head of possessive NP	Occurs with a gender suffix
OPEN CLASSES	Kinship nouns	✓	✓	✓	✓	✓	some
	Body part nouns	✓	✓	✓	✓	✓	x
	Vocative nouns	✓	x	x	✓	x	some
	Bound common nouns	✓	✓	✓	✓	✓	+H/P
	Personal names	✓	x	✓	✓	✓	most
	Place names	✓	x	✓	✓	✓	x
	Unbound common nouns	✓	✓	✓	✓	✓	+H/P
CLOSED CLASSES	Personal pronouns	✓	some	✓	✓	x	x
	Possessive pronouns	✓	x	x	✓	n/a	x
	Demonstrative pronouns	✓	x	n/a	✓	x	x
	Interrogative pronouns	some	x	x	some	x	x
	Indefinite pronouns	✓	x	x	some	x	x

Table 4.1 makes a primary bipartite division of open and closed classes and secondary tripartite division of bound nouns, unbound nouns and pronominals. Bound nouns can be further subdivided into lower level groups according to their morphological structure and behaviour in word formation. The relational and non-relational sub-classes, described in §4.1.3 and §4.1.4 respectively, are distinguished by the types of prefixes their bound noun roots take in citation form (see §7.2). The forms of the prefixes are:

tə- (RL) relational prefix

a- (NRL) non-relational prefix

Unbound nouns form a semantic class of common nouns. These are independent phonological words whose roots do not require an affix in order to function as noun phrase heads (see §4.1.5). Lastly, pronominals subdivide into

five closed classes. The personal pronouns (§4.2.1) and the possessive pronouns (§4.2.2) share areas of overlap in the marking of possession, but have different clausal distributions according to whether they can form the heads of core noun phrase arguments (for a discussion of restrictions, see §5.3).

It is pertinent to mention that semantic considerations prove to be an unreliable indicator of nominal sub-class membership. While body part and kinship terms group together neatly in a bound relational sub-class on the basis of both semantic and morphological criteria, semantically similar common nouns are found to belong to both the bound non-relational and the unbound common noun sub-classes. The unreliability of semantic criteria alone for establishing nominal sub-class membership is demonstrated by the fact that the referents of some nouns involved in whole-part relationships variously belong to the relational or non-relational bound classes, e.g. *tə-tuŋ* (RL-stem/trunk) ‘stem, trunk’ and *á-sǎŋ* (NRL-wood) ‘wood’, while others are members of the unbound class, e.g. *khuŋsaŋ* ‘branch’, and *hnaɪu* ‘flower’. Similarly, the body fluids *tə-sə* (RL-bile) ‘bile’ and *á-jíʔ* (NRL-blood) ‘blood’ might be considered to belong to the same relational noun class because of their whole-part semantic relationship to the body, yet ‘bile’ takes the relational prefix on its stem and ‘blood’ takes the non-relational prefix.

4.1.3. Relational bound nouns

Relational bound noun roots occur in citation form with the relational prefix *tə-* and include the semantic classes of body parts, e.g. *tə-khǎt* (RL-arm/hand) ‘arm, hand’, *tə-ma* (RL-face) ‘face’, *tə-məsən* (RL-liver) ‘liver’; kinship terms, e.g. *tə-tí* (RL-older sibling) ‘older brother/sister’, *tə-paʔ* (RL-father) ‘father’; and some miscellaneous words that are conceptually involved in a whole-part relationship with another entity, e.g. *tə-zəŋ* (RL-root) ‘root’, *tə-niŋ* (RL-name) ‘name’ and *tə-hmɪla* (RL-shadow/soul/ reflection) ‘shadow, soul, reflection’. The relational prefix *tə-* is homophonous with the nominalizing prefix *tə-* that derives nominals from verbal roots. These two prefixes are treated as being synchronically distinct on the basis of their different functions and different distributions, but possibly have the same diachronic source (see §7.2.2, §7.2.4 and §8.3.4 for description of functions).

The restricted semantic class of relational bound nouns suggests that it could constitute an inalienably possessed class, yet this possibility is not consistently borne out by investigations. Mongsén instead presents a dichotomy of relational and non-relational bound classes of nouns. To a Mongsén speaker’s way of thinking, a relational kinship term like *tə-paʔ* implies ‘someone’s father’, rather than just ‘father’; i.e. its meaning cannot be conceptualized in isolation from its relationship to another entity. But the prefix *tə-* is not a third person form of a

pronoun, nor can it or any other form be reconstructed as a third person pronoun in PTB (Benedict 1972: 93). The specific expression of third person possession therefore demands the presence of the third person pronoun in addition to the relational prefix, e.g. *pa tə-paʔ* (3SG RL-father) ‘his/her father’. The third person pronoun cannot replace the prefix of a kinship term to derive a grammatically acceptable form, e.g. **pa-paʔ*. Thus, the only function of the relational prefix is to fulfil the morphological requirements of word formation for this particular class of bound nominal root.

Kinship and body part nouns can function as the heads of possessive compound nouns formed with any lexical roots belonging to a sub-class of noun identified in Table 4.1 above. When a relational noun occurs as the rightmost element of a possessive compound noun, its relational prefix is optionally deleted at the morpheme boundary if the compound is not fully lexicalized, e.g. *a-miʔ-puk* (NRL-person-stomach) ~ *a-miʔ tə-puk* (NRL-person RL-stomach) ‘a person’s stomach’. According to my consultants, the two types of possessive noun phrase have no difference in meaning.

Table 4.3. Citation and vocative forms of some kinship terms

CITATION FORM	VOCATIVE FORM	SEMANTICS
<i>tə-paʔ</i>	<i>a-paʔ</i>	‘father’
<i>tə-ja</i>	<i>a-ja</i>	‘mother’
<i>tə-u</i>	<i>a-u</i>	‘father’s father’
<i>tə-zə</i>	<i>a-zə</i>	‘father’s mother’
<i>tə-u</i>	<i>a-u</i>	‘mother’s father’
<i>tə-zə</i>	<i>a-zə</i>	‘mother’s mother’
<i>tə-ku</i>	<i>a-ku</i>	‘mother’s brother’
<i>tə-pázəmpaʔ</i>	<i>á-pázəmpaʔ</i>	‘father’s elder brother’
<i>tə-pánuzapaʔ</i>	<i>á-pánuzapaʔ</i>	‘father’s younger brother’
<i>tə-ti</i>	<i>à-ti</i>	‘older sibling’
<i>tə-nu</i>	–	‘younger sibling’
<i>tə-ni</i>	–	‘wife’
<i>tə-pázáʔ</i>	–	‘husband’

Kinship noun roots make up the only semantic class of noun that can take the vocative prefix *a-*. This suggests that they form a lower level, closed sub-class of relational noun. Note that the vocative prefix is homophonous with, but functionally distinct from the non-relational prefix *a-* (§7.2.3). The vocative

prefix obligatorily replaces the relational prefix of the citation form, e.g. *a-pa?* (VOC-father) ‘Father!’

There are some unexpected gaps in the vocative paradigm, as demonstrated by the selection of kinship terms presented above in Table 4.3. A vocative form for the term *tə-nu* (RL-younger.sibling), ‘younger sibling’, for example, is conspicuous for its absence and the homophony its root shares with that of *anu*, the citation form for ‘child’. A fossilized elaborate expression *təza-tənu* ‘children’ is encountered in texts (for an example, see Text 2, line 84), but neither of its constituents can be used in isolation as a lexical noun with an independent meaning of ‘child’. A grammaticalized form functioning synchronically as the diminutive suffix *-za* is the only synchronic trace (see §7.5.3).

Undoubtedly there are historical reasons for *tənu* and *təza* not occurring as citation forms for ‘child’. Matisoff remarks in an annotated footnote that ‘the Old Kuki languages commonly replace TB **za* ‘child’ with forms derived from TB **na·w* ‘younger sibling’ (Benedict 1972: 134fn.). We might speculate that the root of a putative Proto-Ao form **(tə-)za* was similarly replaced by **na·w*, resulting in an intermediate form **tə-nu*. But this would have created homophony with *tə-nu* ‘younger sibling’. Replacing the relational prefix *tə-* with the non-relational prefix *a-* could have resolved this ambiguity, hence *a-nu*.

Lastly, despite the terms for ‘wife’ and ‘husband’ having the morphological structure of relational bound nouns, they too lack corresponding vocative forms. This and other nomenclatural anomalies in the system of kinship terminology can probably also be explained by taking historical factors into account, but must await further historical-comparative research on the Tibeto-Burman languages of north-east India.

4.1.4. Non-relational bound nouns

Non-relational bound nouns occur in citation form with the non-relational prefix *a-* and belong to the semantic classes of cultural artefacts, e.g. *a-mi* (NRL-spear) ‘spear’, *a-nuk* (NRL-machete) ‘machete’ (more commonly known by the Nagamese term *dao* in Nagaland), *a-ki* (NRL-house) ‘house’, and entities belonging to the biosphere, e.g. *á-hlú* (NRL-field) ‘field’, *a-niŋ* (NRL-sky) ‘sky’, *a-khi* (NRL-frog) ‘frog’ and *a-ji* (NRL-dog) ‘dog’. These form possessive noun phrases in an identical manner to the bound nominal roots of the relational nouns described in §4.1.3. They also demonstrate the same structural options for formally encoding a possessive relationship; e.g. *nì á-hlú* (1SG NRL-field), *kə á-hlú* (1SG.POSS NRL-field) and *kə-á-hlú* (1SG.POSS-field) all express ‘my field’.

Older native speakers use both personal and possessive forms of pronominals for marking possession of the same entity by the same referent in

narratives, sometimes only a few clauses apart. This observation supports their claims that there is no difference in meaning between, for example, the three possessive noun phrases of the preceding paragraph. Conversely, speakers aged approximately forty years or younger tend to avoid using personal pronouns to mark possession and often reject constructions such as *nì á-hlú* as ungrammatical when they encounter them in older speakers' texts (see §7.3.1 for additional examples and discussion).

Non-relational bound nouns behave similarly to relational bound nouns in the formation of compound nouns via stem apposition, e.g. *mama* 'breast' + *a-tsə* (NRL-water) 'water' → *mama-tsə* 'milk'. Whether the compound-internal prefix is deleted or retained is an indication of the extent of lexicalization of the compound noun. Fully lexicalized compounds are less likely to allow this option. For example, right-headed compounds formed with the non-relational noun *á-sáʔ* (NRL-meat), such as *áwk-sàʔ* 'pig-meat', do not normally permit the retention of the word-internal nominal prefix. This may be related to a semantic shift in the nature of the relationship holding between the juxtaposed nouns – from one encoding a possessor-possessee relationship to one encoding a modifier-modified relationship.

A final observation concerns the distribution of the relational and non-relational prefixes. These are found to be mutually exclusive with respect to the noun sub-classes with which they can occur, except for one particular bound noun root: *maŋ* 'body'. This normally belongs to the relational class and therefore takes the relational prefix *tə-* in the default situation, e.g. *tə-maŋ*, because of the relationship holding between the entity and its possessor. However, this requires the body to be possessed by a living possessor. To refer to a body after death, the non-relational prefix is used, e.g. *a-maŋ*, thereby signalling a severance of the relationship between the possessor and the possessed entity. This morphologically encoded semantic distinction thus offers strong support for the primary division of bound nouns into relational and non-relational classes.

4.1.5. Unbound nouns

The unbound noun class comprises personal names, place names and common nouns. Personal names are distinguished from other types of unbound nouns by virtue of the fact that they can be used as vocatives.² While it might be notionally possible to use common nouns with human referents as vocatives, this would be unusual because kinship terms are available for this function,

2. Personal names used as vocatives, however, do not take the vocative prefix, e.g. *.ɲʊsənɲɛmpaʔ* 'Rongsenlemba!'.

even if the interlocutors are actually unrelated. Kinship terms are widely used as terms of address to decrease social distance between unrelated people throughout South and South-East Asia.

Unbound common nouns form a miscellaneous group of disparate semantic classes subsumed by flora, e.g. *hnaju* ‘flower’, *khəmphá* ‘cotton’ and *wana* ‘husk’; fauna, e.g. *iɪəm* ‘bear’, *səŋphila* ‘lizard’, *wahɪu* ‘raven’ and *uluʔsa* ‘blue-throated barbet’; insects, e.g. *hjuŋza* ‘ant sp.’, *məɪɪla* ‘mosquito’ and *ʔaŋʔaŋ zɪzɪŋla* ‘cicada’; natural phenomena, e.g. *məpuŋ* ‘wind’, *tsəhŋi* ‘sun’, and *tsəŋi* ‘rain’; and a dustbin category of various entities, e.g. *həntsə* ‘egg’, *ʔfənàʔ* ‘refuse’ and *mətsə* ‘salt’. Possession can be encoded by the simple juxtaposition of unbound roots. The pronominal possession of referents of unbound nouns is encoded by the apposition of a personal or possessive pronoun, e.g. *nì ləmsi.ɪù* (1SG comb) ~ *kə ləmsi.ɪù* (1SG.POSS comb) ‘my comb’, with no apparent difference in meaning, although as noted in §4.1.4, younger speakers reject possessive noun phrases formed with personal pronouns as ungrammatical (see §7.3 for the description of possession).

A comparison of the range of meanings of the examples above with those of the non-relational nouns discussed in §4.1.4 reveals that not all members of the non-relational class and the unbound class of common nouns separate neatly into discrete groups on the basis of semantic or physical similarity. The names of some animals occur as bound roots taking a non-relational prefix, e.g. *a-ji* (NRL-dog) ‘domestic dog’ and *a-paŋ* (NRL-buffalo) ‘water buffalo’, while others are unbound, e.g. *tsəhŋi* ‘wild dog’ (or ‘dhole’, *Cuon alpinus*) and *màsəʔ* ‘domestic ox’. One might assume that a domestic dog and a wild dog ostensibly belong to the same taxonomic class of canine because they are of comparable size, dietary requirements and behaviour. The same might be assumed for a water buffalo and a domestic ox. Yet the nominal forms of these referents evince a very different morphological structure that cannot be accounted for with recourse to any particular semantic parameter, such as a ‘domestic’ versus ‘wild’ contrast. This suggests that there may be an unrecognized system of classification responsible for the division of common nouns that is not transparent to the present analysis.

Alternatively, there may be historical reasons for morphological differences that have become obscured by the passage of time.

4.2. Closed word classes

Under the closed word classes are subsumed: five sub-classes of pronominals; a miscellaneous class of nominal adjuncts functioning as modifiers of noun phrase heads, including deictics, quantifiers and case marking clitics; time words; nascent clausal conjunctions that might be ideally described as

‘discourse connectives’ from the role they play in discourse; a phrasal conjunction; a semi-closed class of interjections and exclamations; and a sizable class of particles.

4.2.1. Personal pronouns

The personal pronouns distinguish three persons and three numbers. There is an inclusive/exclusive contrast in the first person of the dual and plural series. A noun with a human referent can be substituted by a personal pronoun, but the substitution of a noun with a non-human referent is limited to the third person singular. If the non-human referent is non-singular in number, then it is referred to by a collective noun determined by a nominal deictic, e.g. *ípá? luŋ tʃu* (EMPHAT group DIST) ‘that group (of birds, houses, etc.)’.

Table 4.4. Personal pronouns

		1 ST	2 ND	3 RD
SG		<i>nì</i>	<i>nàŋ</i>	<i>pa</i>
DU	INC	<i>inət</i>	<i>nəŋət</i>	<i>panət ~ tuŋnət ~ tuŋət</i>
	EXC	<i>kənət</i>		
PL	INC	<i>isa ~ isəŋla</i>	<i>nəŋla ~ nəŋkhəla</i>	<i>túŋla ~ túŋəla ~ túŋkhəla</i>
	EXC	<i>ila ~ ikhəla</i>		

It is noteworthy that the use of the third person singular for non-human referents occurs in the conversation and narratives of older speakers, yet a speaker in her late twenties considered this to be an incorrect usage when it was investigated. This possibly indicates an evolving age-related shift in the functional reference of the third person singular pronoun. The following sentence illustrates the use of the third person singular *pa* to denote a non-human referent. This was uttered by a consultant aged sixty-nine when I was eliciting the characteristics of a bird from a picture in a field guide.

- (4.2) *pa ténákɪəm.*
 pa tə-nák-ɪəm
 3SG NZP-be.black-be.red
 ‘It’s dark reddish/brown.’

There is considerable free variation in the forms of the third person dual and in the whole of the plural series. The free variation in the plural series is due to the tendency of speakers to reduce trisyllabic pronouns to disyllables in rapid

speech. Syllables containing a plosive are specifically targeted, suggesting a process of phonological assimilation to the continuant environment flanking the syncopated syllable.

Benedict (1972: 93) posits PTB first and second person forms **ŋa* and **naŋ* respectively, but cannot establish a general third person pronoun for the family. This would explain why the structure of a possessive noun phrase with a relational bound noun as its head is so different when it is marked for third person possession, e.g. *kə-ni* (1SG.POSS-wife) ‘my wife’, *nə-ni* (2SG.POSS-wife) ‘your wife’, but *pa tə-ni* (3SG.POSS RL-wife) ‘his wife’, and never **pa-ni*. Many languages lack a distinct third person form, instead employing demonstratives for third person pronominal reference. The diachronic origin of the Mongsen third person singular *pa* is likely to be the masculine semantic gender marker *-pàʔ*, which Benedict (1972: 96) claims to have its source in PTB **pa* ‘father’ [STC #24]. This morpheme is also used synchronically as a suffix for marking nominalization and relativization (see §7.4.2).

The form of the first person singular personal pronoun *nì* with its dental nasal initial may be an Ao innovation, as it is not found in any other Tibeto-Burman language of Nagaland apart from Sumi (known in the earlier literature as Sema), a language spoken on the southern border of the Ao territory (see Map 2). Alternatively, it may be a Sumi innovation that was subsequently borrowed into Ao.

Substantial syncretism exists between the personal pronoun paradigm and the possessive pronoun paradigm (cf. Table 4.5), suggesting the grammaticalization of additional person/number categories to flesh out what was originally a rather simple system, followed by the subsequent sharing of forms to fill paradigmatic gaps. The personal pronouns are formally differentiated from the possessive pronouns by their ability to function as the noun phrase heads of core arguments in A, S or O function (see §5.3). This of course does not apply to the third person singular pronoun or any of the dual series, because the same forms are used as both personal and possessive pronouns and therefore cannot be distinguished.

As noted above, another function of the personal pronouns for some speakers is to encode possession. Many younger speakers find the use of personal pronouns for marking possession to be ungrammatical, yet older speakers freely use both personal and possessive pronouns for marking possession on all types of bound and unbound nouns (see §7.3.1). Examples of personal pronouns of the singular series follow.

- (4.3) *nì aja ku tʃhuwajù.ə.*
 nì aja ku tʃhuwa-ì-ùʔ .ə
 1SG night LOC emerge-IRR-DEC PTCL
 ‘I will surely return at night.’

- (4.4) *nàŋ tʃápá? nə waɪ.*
 nàŋ tʃápá? nə wa-əɪ
 2SG what ALL go-PRES
 ‘Where are you going?’
- (4.5) *təɪ, pa kà? atʃu ku sə.*
 tə-əɪ pa kà? a-tʃu ku sə
 thus-SEQ 3SG also NRL-DIST LOC die.PST
 ‘And, he also died there.’

I initially assumed that the dual series was a Mongsen innovation, because dual pronouns are not recognized in either of the grammatical sketches of Chungli by Clark (2002) or Gowda (1975). However, subsequent investigations with bidialectal Chungli/Mongsen speakers revealed that a dual series does in fact exist in the Chungli dialect as well. Dual pronouns in Mongsen have the morphological structure of left-headed endocentric compounds, each form partly consisting of the bound stem *nət* of the numeral *a-nət* ‘NRL-two’.³

In view of the correspondences between the initial syllables of the plural personal pronouns and the initial syllables of the dual personal pronouns, the other likely diachronic source for the grammaticalization of all but two allomorphs of the dual series is the morphology of the plural series. The two exceptions to this are found in the first person dual exclusive *kənət*, the initial syllable of which is identical to the first person singular possessive pronoun *kə*, and the third person dual allomorph *panət*, the initial syllable of which is identical to the third person singular form *pa* of the personal pronoun paradigm.

The third person dual category has three allomorphs, two of those differing only in the extent of assimilation of the dental nasal of *nət* to the preceding velar nasal. All three allomorphs occur in free variation in casual speech.

- (4.6) *inət khələm táŋ nə ləjù.ə.*
 inət khələm táŋ nə li-ì-ù? ɪə
 1DU.INC together just AGT stay-IRR-DEC PTCL
 ‘We two (inclusive of addressee) will just live together.’
- (4.7) *kənət mukuktʃuŋ nə kəwa.ù?*
 kənət mukuktʃuŋ nə kəwa-əɪ-ù?
 1DU.EXC Mokokchung ALL ascend+go-PRES-DEC
 ‘We two (exclusive of addressee) are going up to Mokokchung.’

3. The Chungli dual series has the forms *anə* 1DU.INC, *kənə* 1DU.EXC, *nənə* 2DU and *tənə* 3DU. Like the dual series of Mongsen, the Chungli dual pronouns partly consist of the root for the numeral TWO, which has the form *na* in this dialect.

- (4.8) *nəŋət tʃu – sitakpà? sɔ́pá nə áhlù ku hmapa tʃhèniŋla,*
 nəŋət tʃu sitak-pà? sɔ́pá? nə a-hlú ku
 2DU DIST be.true-NR who AGT NRL-field LOC
 hmapa tʃhèniŋ la
 work do.work.PST TOP
 ‘You two – as for who really worked in the field, ...’

- (4.9) *təku, tuŋnət awŋ ku áhŋá? tʃu thi.*
 tə-ku tuŋnət a-uŋ ku a-hŋá? tʃu thi
 thus-LOC.CV 3DU NRL-jungle LOC NRL-fish DIST divide.PST
 ‘And then, they (two) divided the fish in the jungle.’

The plural series of personal pronouns similarly demonstrates a left-headed structure and substantial allomorphy, but as noted above, this is attributable to syncope. The phonologically reduced forms exhibit the same distribution as the unreduced forms. A high front vowel whose historical source is unknown occurs recurrently in all allomorphs of the first person plural category.⁴ The initial syllable of the second person plural form is obviously related to the PTB singular form **naŋ* (Benedict 1972: 93). The diachronic source of the initial syllable *túŋ* of the third person form is unknown at present, as is the source of a syllable *khə* that occurs in all person categories of the plural series except the first person inclusive. This syllable may be related to a morpheme *khə* that functions synchronically as a phrasal conjunction in Mongsen and appears to be cognate with the associative case marker *kə* of Meithei (see §4.2.11 for discussion). Examples of plural personal pronouns in sentences follow.

- (4.10) “*isa kà? pi sɔ́pá? nə ajimì*”, *tə səmtsə.*
 isa kà? pi sɔ́pá? nə ajim-ì tə səmtsə
 1PL.INC also PROX who AGT announce-IRR thus ask.PST
 ‘“Who among us (inclusive of addressee) here will make an announce-
 ment?” [she] asked.’

- (4.11) *ila thuku táŋ nuŋla,*
 ìla thuku táŋ nuŋ la
 1PL.EXC nine just NEG.PTCL TOP
 ‘It is not just us (exclusive of addressee) nine, ...’

4. This possibly correlates with the proximate nominal demonstrative *i* in respect to both phonological shape and deictic orientation (see §4.2.7.1).

- (4.12) *nəŋəla tʃu thaku la anu? kúták lima tʃu ku məthəpsi-ə̀-ù?*
 nəŋkhəla tʃu thaku la anu?
 2PL DIST now+LOC TOP still
 kúták lima tʃu ku mə-thəpsi-ə̀-ù?
 heaven country DIST LOC NEG-discuss-PRES-DEC
 ‘You lot are still not being talked about in heaven nowadays.’

- (4.13) *təpa à nə sùpà? túŋəla thuku tə saw.ɪə.*
 tə-pa? à nə sù?-pà? túŋkhəla thuku
 RL-father one INST beget-NR 3PL nine
 tə sa-ù? təɪ
 thus say.PST-DEC REP
 ‘‘They who are begotten of the one father are nine’’, [he] said.’

A stem-final element *la* is common to all the unreduced forms of plural personal pronouns. Mills (1926: 337) suspects that this is an obsolete plural marker found only in the (personal) pronouns and expressions formed with *ami*. He cites *amilānā sar*, translated as ‘men say’, as an instance of its use. The example has been phonemicized and interlinearized in (4.14) to assist in the interpretation of the forms. The morpheme *la* could be the topic particle (§4.2.14.3), although it is fairly unusual for it to occur within the noun phrase rather than at the end of it. I have never come across a form *la* being used to mark nominal plurality like this in the data.

- (4.14) *amila na sa.*
 a-mi? la na sa-ə̀ɪ
 NRL-person ? AGT say-PRES
 ‘Men say.’

A somewhat unusual feature that the personal pronouns share with common nouns is an ability to occur in constituency with a nominal deictic (§4.2.7.1), as in the following example. Note also that the third person singular personal pronoun is used in afterthought by the sixty-nine-year-old speaker of this sentence to denote an inanimate referent.

- (4.15) *tə akhətà hjutsə, akhətà, pa tʃu.*
 tə a-khətà hjutsə a-khətà pa tʃu
 thus NFP-one story NRL-one 3SG DIST
 ‘Thus, that's one story, that.’

One might presume that nominal deictics would never be encountered with first person pronouns, because on conceptual grounds the first person is by nature already strongly deictic. Nevertheless, it is possible for a first person pronoun to occur with a deictic in Mongsén, as in the following narrative example. The meaning encoded by the demonstrative here is most likely to be emphatic.

- (4.16) *inət i məphilamì.ù? na.*
 inət i mə-philamì-à.ù? na
 1DU.IN PROX NEG-separate-DESID-PRES-DEC PTCL
 ‘We two don’t want to be separated, do we?’

4.2.2. Possessive pronouns

The possessive pronouns have two functions. One is to mark possession on a noun phrase head, e.g. *pá.ɿ-ki* (3PL.POSS-house) ‘their (family’s) house’ (see §7.3.1). The other function is to serve as the pronominal head of a noun phrase case marked by an oblique case marker, e.g. *kə li* (1SG.POSS DAT) ‘to me’ (see §5.3.2).

Table 4.5. Possessive pronouns

		1 ST	2 ND	3 RD
SG		<i>kə</i>	<i>nə</i>	(<i>pá</i>)
DU	INC	(<i>inət</i>)	(<i>nəŋət</i>)	(<i>panət ~ tuŋnət ~ tuŋət</i>)
	EXC	(<i>kənət</i>)		
PL		<i>i ~ aj</i>	<i>nìŋ</i>	<i>pá.ɿ</i>

Table 4.5 above lists the possessive pronouns. A comparison of these with the personal pronouns of Table 4.4 reveals significant syncretism between the two pronominal paradigms. Obviously forms have been borrowed from one paradigm to fill gaps in the other. The only forms of possessive pronouns that are in fact distinct from the personal pronouns are the first and second person singular forms *kə* and *nə*, and the plural series *i ~ aj*, *nìŋ* and *pá.ɿ*. The extensive syncretism can be accounted for with recourse to the historical considerations previously discussed in §4.2.1. Forms of possessive pronouns given in parentheses in Table 4.5 are identical to forms of personal pronouns of corresponding person and number presented above in Table 4.4.

Examples of possessive pronouns marking nominal possession are given below. Whether these are bound or free forms is open to interpretation. When they replace a nominal prefix on a bound noun, they give the impression of

being bound morphemes. But they also occur in isolation with oblique case marking clitics, contradicting this impression (see §5.3.2 for examples). Moreover, speakers do not always replace a nominal prefix of a bound root when they mark possession using a possessive pronoun; e.g. ‘my hand’ can be expressed by either *kə tɔ́-khət* (1SG.POSS RL-hand) or *kə-khət* (1SG.POSS-hand), which further suggests that they are morphologically independent. Possessive pronouns will henceforth be glossed as prefixes when replacing a nominal prefix of a bound noun, and as independent words when juxtaposed to an unbound noun or any other independent nominal morpheme.

(4.17) *kəjim maŋmətuŋ.*

kə-jim maŋmətuŋ
1SG.POSS-village village.name
‘My village is Mangmetong.’

(4.18) ... *nì nəkìmə jəɹə ɹəjpà?*

nì nə-kimà jəɹə ɹə-ì-pà?
1SG 2SG.POSS-courtyard folkdance come-IRR-NR
... [when] I was to come to your courtyard for the folkdance, ...

A rare occurrence of the third person singular form juxtaposed directly to the bound root of a body part noun is attested in the textual example of (4.19). This is attributed to a rare elision of the bound stem’s relational prefix *tə-*, because consultants unanimously reject prefixing the third person singular form directly to a bound stem to mark possession. There is an overwhelming preference to retain the nominal prefix on a bound noun when marking possession with the third person singular form. This is typified by the possessive noun phrase of (4.20).

(4.19) *təɹ, hɹətsə? tʃu, pa mija nə athənsipa?, pa nə nəmukzəktʃuk, təɹ tʃətsʃuk.*

tə-əɹ hɹətsə? tʃu pa mija nə athənsi-pà?
thus-SEQ borer.insect DIST 3SG palm ALL assemble-NR

pa nə nəm-uk-zək-tʃuk tə-əɹ tʃəʔ-tʃuk
3SG AGT compress-INTO-SEND-PFV.PST thus-SEQ consume-PFV.PST
‘And then, the borers that assembled in his [i.e. Tiger’s] paw he crushed together and then gobbled up.’

(4.20) ... *pa tələm tʃu əpakapakə kəməɹ, ...*

pa tə-ləm tʃu əpak-apak-əkə kəm-əɹ
3SG RL-head DIST be.flat-RED-SIM become-SEQ
‘... after her head became completely flattened, ...’

Any entity that is collectively possessed is automatically encoded as being possessed in the plural number, regardless of the context. For example, a Mongsen speaker will naturally refer to ‘my family’ as *i-kìtsəŋ* (1PL.EXC.POSS-family) in preference to *kə-kìtsəŋ* (1SG.POSS-family).

- (4.21) *ikìtsəŋ thuku.*
 i-kìtsəŋ thuku
 1PL.EXC.POSS-family nine
 ‘Our family are nine [people].’

This is because he or she is a member of a group at a particular level in the social hierarchy, and recognition of collective ownership at this level always takes precedence over the perception of ownership at the level of the individual, although the latter may be equally true. Consequently entities such as houses, livestock, families, children and fields are by default linguistically encoded as being possessed in the plural, even when the reference might be specifically personal and therefore more accurately encoded as such in the singular.

The textual example of (4.22) below demonstrates this preference. The assertion is addressed to an individual, the father, but the possession of *tfàla* ‘daughter’ is encoded in the plural because the daughter is conceptualized as belonging to the family, not just to the individual addressed. Example sentences further demonstrating the use of plural possessive pronouns to mark possession are presented in §7.3.1.

- (4.22) *mətʃatshəŋ nə mijipəɪ, “nì nìŋ tfàlaj à tshəmì.ùʔ,” tə saw.tə.*
 mətʃatshəŋ nə mijip-əɪ nì nìŋ tʃà-la i
 PN AGT turn-SEQ 1SG 2PL.POSS offspring-F PROX
 à tshə-mì-əɪ-ùʔ tə sa-ùʔ tət
 one take-DESID-PRES-DEC thus say.PST-DEC REP
 ‘Mechatseng turned and said “I want to take one of these daughters of yours.”’

In addition to their function of marking possession on noun phrase heads, possessive pronouns serve as the heads of oblique noun phrase arguments. A possessive form of pronoun is obligatorily selected by the dative case marker. All other oblique case markers optionally select either personal pronouns or possessive pronouns as their noun phrase heads (see §5.3.2 and §5.3.3 for additional examples and discussion of the historical reasons for this).

(4.23) *thúŋkhəla nə nəli kutai khi?*

thúŋkhəla nə nə li kutai khi?
 3PL AGT 2SG.POSS DAT knife give.PST
 ‘They gave you the knife.’

(4.24) *pa nəthən kútə kitsəŋəi.*

pa nə thən kútə kitsəŋ-əi
 3SG 2SG.POSS COM how family-ANOM
 ‘How is she related to you?’

4.2.3. Demonstrative pronouns

A demonstrative pronoun prototypically has a spatial referential function and can stand in place of a noun. Demonstrative pronouns that can function as noun phrase heads make up a closed class of two members: the proximate demonstrative pronoun *ipi* ~ *pi* ~ *i* (PROX) ‘this’, and the distal demonstrative pronoun *a-tfu* (NRL-DIST) ‘that’.

The proximate demonstrative pronoun *ipi* has two phonologically reduced variants and the distal demonstrative pronoun is derivative, formed with the non-relational prefix *a-* on the root of the nominal distal demonstrative *tfu*. Only the derived form of the distal demonstrative can function as a noun phrase head.

(4.25) *pi kútə tʃhà.*

pi kútə tʃhà-əi
 PROX how do-PRES
 ‘How do [I] do this?’

(4.26) *atfu təthəm.*

a-tfu tə-thəm
 NRL-DIST NZP-end
 ‘That’s the end.’

Both the proximate and the distal nominal demonstratives can occur with the emphatic demonstrative *ípá?* (described in §4.2.7.1 below) to form another type of demonstrative pronoun that functions as an independent noun phrase with deictic specification, e.g. *ípá?-i* (EMPHAT-PROX) ‘this (one)’, *ípá?-tfu* (EMPHAT-DIST) ‘that (one)’. These are analyzed as a type of headless noun phrase. The word-internal glottal stop of the emphatic demonstrative is deleted in the output, indicating that these two derivations form single phonological words (see §2.4.1). The proximate and distal demonstrative pronouns can be used to

express the locative adverbial meanings of ‘here’ and ‘there’ respectively when occurring in constituency with an appropriate oblique case marker.

(4.27) *inə hi:ɹə saŋ atə.*

i nə hi:ɹə-əɹ sa-aŋ atə
 PROX ALL level+come-SEQ say-IMP PTCL
 ‘Come here and say [it]!’

(4.28) *tə hjutsə atʃu ku thəm-əɹ.*

tə hjutsə a-tʃu ku thəm-əɹ
 thus story NRL-DIST LOC end-PRES
 Thus, the story ends there.

Sometimes the proximate demonstrative is used without a local case marking clitic, but nevertheless still expresses a locative adverbial meaning.

(4.29) *təmija nə athənsiaŋ i.*

tə-mija nə athənsi-aŋ i
 RL-palm ALL assemble-IMP PROX
 ‘Gather here in the palm [of my paw, said Tiger to the insects].’

4.2.4. Interrogative pronouns

The interrogative pronouns comprise a small class of ten members. These are listed in Table 4.6 below. Interrogative pronouns express indefinite reference and provide the primary means of forming content questions (see §5.1.2). Those expressing the categories of PERSON/ENTITY, PLACE, THING and QUANTITY are formally distinguished from the other ontological categories listed in Table 4.6 by their ability to function as the heads of their own noun phrases.

The category of PLACE has two sub-categories expressing ‘goal’ and ‘location’; these are differentiated by the allative and locative case marking clitics respectively. There are two variants of the LOCATION category. The base of one of these, *tʃəpáʔ* ‘what’, represents the ontological category of THING when it is unmarked by a local case marker.

The bases of the first four categories listed in Table 4.6 below appear to be formed with the ubiquitous general nominalizer *-pàʔ* (see §3.1.3.1 and §7.4.2). If so, it has become a fossilized stem formative in these interrogative pronouns, and as such is unsegmentable. Four of the interrogative pronouns have a specific initial syllable formed with a velar onset and a high back rounded vowel, and all except for one form of the PLACE category are uniformly distinguished by having high tones.

Table 4.6. Interrogative pronouns

CATEGORY	FORM	SEMANTICS
PERSON	<i>səpáʔ</i>	‘who’
PERSON / ENTITY	<i>kúpáʔ</i>	‘which’
PLACE (GOAL)	<i>tʃəpáʔ nə</i>	‘what ALL’
PLACE (LOCATION 1)	<i>tʃəpáʔ ku</i>	‘what LOC’
PLACE (LOCATION 2)	<i>kətsə ku</i>	‘where LOC’
THING	<i>tʃəpáʔ</i>	‘what’
TIME	<i>kújím</i>	‘when’
MANNER	<i>kútə</i>	‘how’
REASON	<i>tʃútə</i>	‘why’
QUANTITY	<i>kújáʔ</i>	‘how many’

Unlike the corresponding categories of many European languages, interrogative pronouns cannot also be used for forming relative clauses, although some marginal constructions that have an Indic-like relative-correlative structure typified by (4.81) of §4.2.14.3 below occasionally crop up in the texts of native speakers (see §6.6.6 for a full account of the relative-correlative strategy). Curiously, younger educated Mongsen speakers have a preference for this type of relative clause when directly translating English relative clauses into Mongsen. I attribute this to their attempts to provide a translation that is both structurally as well as semantically parallel to the English relative clause. Examples of interrogative pronouns follow.

- (4.30) *pa səpáʔ*
 pa səpáʔ
 3SG who
 ‘Who is he?’

In (4.31) below, the question “What is your name?” obligatorily employs the PERSON category of interrogative pronoun in preference to the THING category in this particular context, but either *səpáʔ* ‘or *tʃəpáʔ* may be used to ask “What are you called?”.

- (4.31) *nə tənɪŋ səpáʔ*
 nə tənɪŋ səpáʔ
 2SG.POSS RL-name who
 ‘What is your name?’

- (4.32) *kúpála tshəj.*
 kúpáʔ la tshə-ì
 which TOP take-IRR
 ‘Which one will [you] take?’

The locative interrogative *tʃəpáʔ ku* often undergoes a phonological contraction to *tʃáku*. The high frequency of occurrence of the reduced form in texts suggests that it is on the way towards full lexicalization.

- (4.33) *ahənzala tʃáku li.*
 ahən-za-la tʃəpáʔ ku li
 chicken-DIM-F what LOC be
 ‘Where is Little Chick?’

- (4.34) *nəŋ tʃəpánə waɪ.*
 nəŋ tʃəpáʔ nə wa-əɪ
 2SG what ALL go-PRES
 ‘Where are you going?’

- (4.35) *kújím .à.*
 kújím .à
 when come.PST
 ‘When did [you] come?’

- (4.36) *tətʃhaɪ “kútə.”*
 tətʃhà-əɪ kútə
 thus-do-SEQ how
 ‘And then, [he asked] “How is it?”’

- (4.37) *lítʃápáʔ nə “nìŋ kìn, nìŋ kìtsəŋ kújáʔ,” tət səm̀tsəwʔ.*
 lítʃá-pàʔ nə nìŋ kìn nìŋ kìtsəŋ
 PN-M AGT 2PL.POSS clan 2PL.POSS family
 kújáʔ tət səm̀tsə-ùʔ
 how.many thus ask.PST-DEC
 ‘“How many in your clan, your family?”, asked Lichaba.’

4.2.5. Indefinite pronouns

Indefinite pronouns are non-referential. They comprise three subclasses: negative indefinite, specific and non-specific. Forms of the negative indefinite

pronouns are listed in Table 4.7, and those of the specific indefinite and non-specific indefinite pronouns are respectively listed in Tables 4.8. and 4.9.

Each sub-class of indefinite pronoun has exponents for the four categories of PERSON, PLACE, THING and TIME. Some categories in the paradigms use the same forms; for example, the PERSON category across three of the four classes of interrogative and indefinite pronouns either has minor variations of the one base form, or is identical. All indefinite pronoun paradigms rely upon case markers for the semantic differentiation of some of their categories.

4.2.5.1. Negative indefinite pronouns

The negative indefinite pronoun series somewhat parallels the interrogative pronominal series in utilizing a basic form for the expression of the PLACE and THING categories. These are similarly differentiated by the presence or absence of a case marker. The use of the allative and locative case markers with the base form further distinguishes a ‘goal’ from a ‘location’, thereby adding a fifth form to the paradigm.

Table 4.7. Negative indefinite pronouns

CATEGORY	FORM	SEMANTICS
PERSON	<i>sɔʔ</i>	‘no-one’
PLACE (GOAL)	<i>tʃáʔ nə</i>	(BASE ALL) ‘nowhere’
PLACE (LOCATION)	<i>tʃáʔ ku</i>	(BASE LOC) ‘nowhere’
THING	<i>tʃáʔ</i>	(BASE) ‘nothing’
TIME	<i>kújímkáʔ</i>	(when+also) ‘never’

The following examples demonstrate that the negative indefinite pronouns always co-occur in a clause with a negated main verb.

- (4.38) *tʃáʔ nə mùwala.*
 tʃáʔ nə m̀-̀wa-la
 nothing ALL NEG-go-NEG.PST
 ‘[I] went nowhere.’

- (4.39) *pa nə tʃáʔ m̀tʃh̀ǹniṅlawʔ.*
 pa nə tʃáʔ m̀-̀tʃh̀ǹniṅ-la-ùʔ
 3SG AGT nothing NEG-do.work-NEG.PST-DEC
 ‘She did nothing.’

- (4.40) *akhì hmapaŋ ku ləp-təp-thùŋ-ku səʔ nə mùkùklà.*
 akhì hmapaŋ ku ləp-təp-thùŋ-ku səʔ nə
 ancient time LOC cut-RECIP-REACH-LOC.CV no-one AGT
 mə-kùk-la
 NEG-succeed-NEG.PST
 ‘In ancient times when [we] cut off each other’s [heads], no-one won.’
- (4.41) *təi, əjim tətsəŋ nə kújímkaʔ əi.ɪə məzà.ɪà.là.*
 tə-əɪ ə-jim tə-tsəŋ nə kújímkaʔ əi.ɪə
 thus-SEQ NRL-village RL-inside ALL never enemy
 mə-za-ɪà-la
 NEG-enter-come-NEG.PST
 ‘And, enemies never entered the village.’

4.2.5.2. *Specific indefinite pronouns*

Specific indefinite pronouns are listed in Table 4.8 below. The categories of PERSON and PLACE both have two forms that occur in free variation. The variant form in the PERSON 1 category can be attributed to phonological reduction, while some speakers prefer to use the noun *amiʔ* ‘person’ with the quantifier *à* ‘one’ for expressing specific indefinite reference in the PERSON 2 category.

Table 4.8. Specific indefinite pronouns

CATEGORY	FORM	SEMANTICS
PERSON 1	<i>səpáʔ ~ səʔ</i>	‘someone’
PERSON 2	<i>a-miʔ à</i>	(NRL-person one) ‘someone’
PLACE 1	<i>təkhì à ku</i>	(BASE one LOC) ‘somewhere’
PLACE 2	<i>kətsə ku</i>	(where LOC) ‘somewhere’
THING	<i>təkhì à</i>	(BASE one) ‘something’
TIME	<i>khən-khən</i>	(once-RED) ‘sometimes’

- (4.42) *amiʔ à nə áwksàʔ tʃu tʃawkù.*
 amiʔ à nə a-úk-sáʔ tʃu tʃàʔ-ukù
 person one AGT NRL-pig-meat DIST consume-ANT
 ‘Someone has eaten the pork.’

One exponent of the category of PLACE is differentiated from the category of THING by a case marker occurring with the base form. This is the same strategy as that which is used to distinguish these two ontological categories in the negative indefinite pronominal paradigm. In the PERSON 1 category, the phonologically reduced form *səʔ* ‘someone’ is identical to the corresponding negative indefinite form. The ambiguity between the two possible meanings of ‘someone’ and ‘no-one’ is resolved by the polarity of the clause (cf. 4.40). The agentive nominalizer seen in *səʔ-əɪ* in (4.43) functions as a genitive marker here to encode the syntactic relation of possessor (see §7.3.3 for description and additional examples).

- (4.43) *jàsala səɪ tʃâɪ la aɪukə tʃəpukù.*
 jàsaʔ la səʔ-əɪ tʃâ-əɪ la aɪu-əkə
 last.night TOP someone-ANOM offspring-ANOM TOP good-SIM
 tʃəp-ukù
 cry-ANT
 ‘Last night someone’s child was really bawling.’

4.2.5.3. Non-specific indefinite pronouns

The non-specific indefinite pronouns are listed in Table 4.9 below. Each of the four ontological categories is partially represented by a recurring base form *masi*, denoting a fundamental meaning of ‘any’ and occurring with the numeral ONE in all but the TIME category. One form of the PERSON category is morphologically indistinguishable from the THING category.

Table 4.9. Non-specific indefinite pronouns

CATEGORY	FORM	SEMANTICS
PERSON	<i>masi à ~ a-miʔ</i>	‘ANY one’ (alt: NRL-person)
PLACE	<i>masi à ku</i>	(ANY one LOC) ‘anywhere’
THING	<i>masi à</i>	(ANY one) ‘anything’
TIME	<i>masi hmapaŋ</i>	(ANY time) ‘anytime’

The category of PERSON can alternatively be represented by the noun *amiʔ* ‘person’ in interrogative clauses. Its non-specific indefinite reference is determinable from the context in the instance of use, as demonstrated by the following example.

- (4.44) *ami? li sè m̀lila.*
 a-mi? li s̀ m̀li-la
 NRL-person be QPCTL NEG-be-NEG.PST
 ‘Is anyone there?’

The remaining categories of non-specific indefinite pronouns combine the base form *masi* ‘ANY’ with a noun or case marker to derive the other non-specific indefinite meanings of PLACE, THING and TIME.

- (4.45) *ma:ruk t̥ʃu masi à ku t̥ʃhìjàkaŋ.*
 ma:ruk t̥ʃu masi à ku t̥ʃhìjàk-aŋ
 cup DIST ANY one LOC put-IMP
 ‘Keep/put the cup anywhere.’
- (4.46) *nàŋ t̥ʃhəmìpà? masi à hli-aŋ.*
 nàŋ nə t̥ʃhə-mì-pà? masi à hli-aŋ
 2SG AGT take-DESID-NR ANY one buy-IMP
 ‘You buy anything you want.’
- (4.47) *masi hmapaŋ ku táŋ t̥ʃùŋì.*
 masi hmapaŋ ku táŋ t̥ʃùŋ-ì
 ANY time LOC just eat.meal-IRR
 ‘[We] will eat anytime.’

4.2.5.4. *Reduplicated indefinite pronouns and adverbals*

Some non-specific indefinite pronouns are derived by a complete reduplication of the base form of an interrogative pronoun. The reduplicated form of the interrogative *t̥ʃəpá?* ‘what’ demonstrates a change in tone from high to mid on the reduplicated morpheme.

- (4.48) *nì nə t̥ʃəpát̥ʃəpa? sapàkukà? túŋla nə ajisà? khi?.*
 nì nə t̥ʃəpá?-t̥ʃəpa? sa-pàkukà?
 1SG AGT what-RED say-CONCESS
 túŋkhəla nə a-ji-sà? khi?
 3PL AGT NRL-dog-meat give.PST
 ‘No matter what I said, they still gave me dog meat.’

The lexical expression of specific indefinite TIME differs significantly from the other categories of specific indefinite pronouns because it is encoded by

reduplication of an adverbial base form, rather than by reduplication of the corresponding interrogative pronoun used for questioning the TIME category. The following reduplicated example expressing ‘sometimes’ is derived from the adverb *khən* ‘once’.

- (4.49) *khənkhənla ləntuŋ tʃhàjù?*
 khən-khən la ləntuŋ tʃhà-ì-ù?
 once-RED TOP accident be-IRR-DEC
 ‘Sometimes, accidents happen.’

4.2.6. Generic pronoun

There is one pronominal with indefinite generic reference whose meaning and function in Mongsen may be equated with *one* of English or *on* ‘one’ of French. The generic pronoun can serve as the head of a noun phrase and does not obligatorily require a coindexed antecedent reference. Its reference is restricted to animate entities.

- (4.50) *tui nə məmələnsipàla màrujù?*
 tui nə mə-mələn-si-pàla mə-əu-ì-ù?
 GPN INST NEG-change-RPET-COND NEG-good-IRR-DEC
 ‘It would not be good if one cannot change by oneself.’

- (4.51) *aji nə tui nə áthútfən nə anəpən wa mətəm.*
 a-ji nə tui nə athùʔ-tʃən nə a-nət-pən
 NRL-dog AGT GPN INST vomit-LNOM ALL NRL-two-ORD
 wa mətəm
 go like
 ‘Like a dog going back to its vomit by itself for a second time.’
 (= to eat one’s words; to reject something and then want it later)

The sentence of (4.51) above is a proverb, which accounts for its lack of temporal specification. It is also noteworthy for having three homophonous case markers in the one sentence. When confronted by such syncretism, a Mongsen speaker must rely upon contextual semantics and their real world knowledge to correctly assign semantic roles to the referents of identically case-marked noun phrases.

4.2.7. Nominal modifiers

The nominal modifiers comprise three closed sub-classes: nominal deictics; numerals and other lexical quantifiers; and case marking clitics. These word classes are formally characterized as standing in a modifying relationship to a noun phrase head.

Mongsen lacks a word class of underived adjectives, but deverbal adjectives derived from stative intransitive verbs are available to fill this role. These are discussed in §6.4.1 with respect to the noun phrase, and in §9.1.4 with respect to their function in verbless clauses.

4.2.7.1. *Nominal deictics*

The nominal deictic class is made up of the following four members:

- (i) emphatic demonstrative
- (ii) proximate nominal demonstrative
- (iii) distal nominal demonstrative
- (iv) anaphoric nominal demonstrative

The emphatic demonstrative *ípaʔ* occurs in a slot preceding the noun phrase head and can determine a noun, but does not inherently express spatial deixis. Three nominal demonstratives occur in the post-head position: the proximate *i*, the distal *tʃu* and the anaphoric *sə*. Only the proximate and the distal nominal demonstratives can be used for specifying spatial deixis. The anaphoric nominal demonstrative is limited to encoding anaphoric reference. Related or identical forms of the proximate and distal nominal demonstratives can function as heads of their own noun phrases. These are discussed in §4.2.3.

(i) *Emphatic demonstrative*. Apart from pre-head relativized attributes and some quantifiers of a handful of temporal nouns (see §6.1.1), the emphatic demonstrative *ípaʔ* is the only other constituent of the noun phrase that occurs in the pre-head position. This may be significant for determining its diachronic source. It functions as a type of determiner, in the narrow sense of uniquely identifying an entity and imparting definiteness to its noun phrase. Native speakers give varying explanations of its function. Some suggest that it is an optional element adding no real semantic content to the noun phrase it occurs in, while others suggest that it serves to particularize the referent of the noun head it modifies. In narrative texts, the emphatic demonstrative tends to occur in noun phrases whose referents have central roles in discourse. For example, the referent of the noun phrase *ípaʔ áwk i* ‘this pig’ in (4.52) is a highly salient entity that has a recurring role in the text from which the example is taken, so its

referential importance to the narrative may be signalled by the presence of the emphatic demonstrative.

- (4.52) *təku ípá? áwk i thijun nə niukəɪ,*
 tə-ku ípá? a-úk i thijun nə
 thus-LOC.CV EMPHAT NRL-pig PROX hulling.room ALL
 nə-uk-əɪ
 lead-INTO-SEQ
 ‘And so, having herded this pig into the paddy-hulling room, ...’

The same may be said of the emphatic referential function of the emphatic demonstrative in *ípá? pəntʃùŋ tʃu* ‘that (very) festival’ in its context of use. The referent of this noun phrase attracts several mentions in preceding discourse by virtue of it being the reason for which a number of important narrative events take place. Its salience also appears to be flagged by the speaker’s use of the emphatic demonstrative.

- (4.53) *tətʃhəɪ, ípá? pəntʃùŋ tʃu ku pətʃu tʃhəɪ.*
 tə-tʃhə-əɪ ípá? pəntʃùŋ tʃu ku pa tʃu tʃhə-əɪ
 thus-do-SEQ EMPHAT festival DIST LOC 3SG that take-PRES
 ‘And then, in that (very) festival, that [stuff] is taken [and used].’

The emphatic demonstrative can occur in constituency with the anaphoric nominal demonstrative, described in (iv) below. Recall that the anaphoric nominal demonstrative is restricted to performing an anaphoric marking function in discourse and cannot be used for spatial deixis. It is used here to signal coreferentiality with an antecedent mention of its noun phrase referent.

- (4.54) *təku ípá? təkʰət sə məzə? ku phəɪukəɪ.*
 tə-ku ípá? tək-ʰət sə məzə? ku phəɪuk-əɪ
 thus-LOC.CV EMPHAT RL-hand ANAPH fire LOC spread-PRES
 ‘And so, [she] spreads the aforementioned hand over the fire.’

These examples show that if the emphatic demonstrative occurs in a noun phrase with a noun head, it is usually accompanied by a nominal demonstrative in the determiner slot following the head. This might account for the varying explanations of its meaning given by native speakers, because a noun occurring with a spatial deictic is interpreted as definite by default. The emphatic demonstrative might therefore appear to some speakers to be a semantically superfluous addition to the noun phrase. This is especially the case when it

occurs, seemingly redundantly, in a noun phrase together with the anaphoric nominal demonstrative.

Lastly, the emphatic demonstrative can occur in constituency with a nominal demonstrative in a headless noun phrase. In this usage it appears to impart referential definiteness, though this again might be presumed to be already encoded by the nominal demonstrative.

- (4.55) *nì nə ípáʔsʉ ku ləpʰtəŋ-ə̀.*
 nì nə ípáʔ tʃu ku ləpʰtəŋ-ə̀
 1SG AGT EMPHAT DIST LOC judge-PRES
 ‘I pass judgement on those [cases and disputes].’
 * I pass judgment there.

The phonological shape of the emphatic demonstrative is suspiciously close to the proximate demonstrative *i* and the general nominalizer *-pàʔ*. The general nominalizer has a variety of extended functions in Mongsen and this could be yet another one (see §7.4 for a description of all its nominalizing functions). The phonological resemblance of the emphatic demonstrative to the general nominalizer suggests that it could be the relict of a type of emphatic nominalization expressing something equivalent to *this one*. This suspicion is further raised by the fact that it occurs in the pre-head position, which it shares with pre-head nominalized modifiers functioning as relative clauses (see §6.6.1). At present it is not possible to determine precisely the meaning encoded by the emphatic demonstrative, or for that matter its diachronic source, but my guess is that it may have originated from a kind of nominalized attribute and is now in the process of grammaticalizing an additional determiner-like function. This assumption is supported by the observation that a noun phrase cannot be modified by both a pre-head relativized attribute and the emphatic demonstrative, presumably because only one nominalized attribute – or a constituent that has grammaticalized from an erstwhile nominalization – can occur in the pre-head position.

(ii) *Proximate nominal demonstrative.* The proximate nominal demonstrative *i* conforms to the typologically widespread pattern of being segmentally represented wholly or in part by a high front vowel. Its basic function is to point to referents that are spatially or contextually proximate. Because the proximate nominal demonstrative consists of just a single high vowel, it obligatorily forms the offglide of a phonetic diphthong when juxtaposed to a word-final vowel of differing quality. This results in the creation of one phonological word consisting of two grammatical words. As noted in §2.4.1, these resyllabification processes are totally dependent upon the juxtaposed vowels forming sequences that conform to possible Mongsen syllable structures.

- (4.56) *pa nə atʃhuj tən.*
 pa nə atʃhu i tən
 3SG AGT song PROX sing.PST
 ‘He sang this song.’

- (4.57) *nì nə atʃak i tɔ̀ɔ̀ ɛntʃhət.rù.*
 nì nə a-tʃak i tɔ̀ɔ̀ ɛntʃhət-rù
 1SG AGT NRL-paddy PROX some take.out-IMM
 ‘I’m going to take out some of this paddy.’

The proximate nominal demonstrative can also be used in a presentational function to introduce new referents, as in (4.58). This example occurred as the first sentence of a folkloric text (see Text 4, line 06).

- (4.58) *akhì hmapaŋ ku wàzàʔ təmáŋ i muŋthaŋ à sɛntəp.*
 akhì hmapaŋ ku wàzàʔ təmáŋ i muŋthaŋ à
 ancient time LOC bird all PROX conference one
 sɛntəp
 gather.PST
 ‘Long ago, all these birds gathered for a conference.’

The proximate nominal demonstrative is used with the quotative complementizer *tə* ‘thus’ (see §4.2.14.1) to derive a meaning ‘like this’. This construction functions as an adverbial modifier encoding the manner of an activity.

- (4.59) *itə mənu-kə jipli-àɪ pà.*
 i-tə mənu-əkə jipli-àɪ pà
 PROX-thus be.late-SIM sleep-PRES QPTCL
 ‘Sleeping in late like this [are you]?’

(iii) *Distal nominal demonstrative.* The basic function of the distal nominal demonstrative *tʃu* is to point to something at a distance from the deictic centre. The entity pointed at can be visible or non-visible.

- (4.60) *nàŋ aki tʃu nə asáwa.*
 nàŋ a-ki tʃu nə asáʔ-wa
 2SG NRL-house DIST ALL ADM-go
 ‘Don’t you go to that house too much!’

The distal nominal demonstrative is overall the most commonly occurring nominal deictic and functions as the default determiner of noun phrases in

narratives. When helping to translate texts or to elicit data, Ao consultants who speak English suggest the English definite article *the* as a functional equivalent of the distal demonstrative. This is how it is used in (4.61), taken from a procedural text on making rice beer.

(4.61) *muli tʃu phàtshəŋì atə.*

muli tʃu phà-tshəŋ-ì atə
 medicine DIST sprinkle-ATTACH-IRR PTCL
 ‘The medicine [i.e. yeast] should be sprinkled and left [on the cooked rice].’

Once a referent has been introduced in a text, it is customary for its noun phrase to be subsequently marked by the distal nominal demonstrative, at least in the early stages of the narrative, until its reference is well established. The distal nominal demonstrative is then gradually replaced by the anaphoric nominal demonstrative as the principal means of determining a noun phrase with an established referent.

Like the proximate nominal demonstrative, the distal nominal demonstrative can be used with the quotative complementizer to form a manner adverbial construction.

(4.62) *tʃutə tʃhaŋ.*

tʃu-tə tʃhà-aŋ
 DIST-thus do-IMP
 ‘Do it like that.’

(iv) *Anaphoric nominal demonstrative.* The anaphoric nominal demonstrative *sə* has a referent tracking function and is used to determine referents representing old or remembered information. It cannot be used for locating entities in space. Consequently it is not grammatical to use the anaphoric nominal demonstrative if the referent is visible, and in discourse it is generally employed only after a referent’s identity has been well established. It tends to be used incrementally in narrative texts with recurring references to the same entity, first overlapping with and then often totally replacing the distal nominal demonstrative as the primary means of determining noun phrases of referents representing old or given information. In the following text example, the referent of the noun phrase *wàzà? təmáŋ sə* ‘all the birds’ is established through prior mentions – the sentence of (4.58) above initially introduces this referent at the beginning of the text. Its old information status is encoded a few clauses later by the anaphoric nominal demonstrative.

- (4.63) *wàzà? təmáj sə jajaɪ, pa thaku jàk.*
 wàzà? təmáj sə jaja-əɪ pa thak ku jàk
 bird all ANAPH be.angry-SEQ 3SG PLACE LOC beat.PST
 ‘All the [aforementioned] birds got angry and beat on her.’

An investigation of narrative texts reveals that there is sometimes an abrupt and early transition to using the anaphoric nominal demonstrative to determine the subsequent noun phrases of a recurring referent. I attribute this to assumptions that a speaker may make with respect to given information and shared socio-cultural knowledge. In the context of a folklore story in which the protagonists are believed to be familiar to an Ao-speaking audience, a speaker may well use the anaphoric nominal demonstrative at a relatively early stage. Overall, however, a pattern of incremental use emerges as being more typical of the anaphoric demonstrative’s distribution in discourse.

While its most common function is to mark the noun phrases of recurring referents in narrative texts, the anaphoric nominal demonstrative can also be used to determine noun phrases in ordinary everyday conversation if the speaker believes that the interlocutor(s) can uniquely identify a referent from a remembered event or shared knowledge.

- (4.64) *hmapa təmáj sə tʃhəniŋaŋ.*
 hmapa təmáj sə tʃhəniŋ-aŋ
 work all ANAPH do.work-IMP
 ‘Do all that work [i.e. that we have previously discussed].’

Lastly, the anaphoric nominal demonstrative cannot be used with the quotative complementizer to form a manner adverbial construction expressing an analogous meaning to those demonstrated by the spatial deictics in the examples of (4.59) and (4.62).

4.2.7.2. Lexical quantifiers

The lexical quantifiers are a smallish sub-class of nominal modifiers comprising less than a dozen lexical items expressing indefinite quantity, e.g. *kúlàŋà* ‘many’ or *təjà* ‘a little’. A mass noun versus count noun distinction is not apparent for these lexical items, but semantic restrictions are found to restrict the range of use of one member of the sub-class. The lexical quantifiers are listed with their glosses in Table 4.10 below. Some are limited to occurring with particular semantic classes of noun; e.g. *muŋjaŋmúŋ* ‘whole’ is only attested as a quantifier of nouns expressing temporal measurement, and *tija?* ‘some’ is only

used to quantify nouns with human referents. The functions of lexical quantifiers in noun phrases are discussed in §6.4.2.

Table 4.10. Lexical quantifiers

QUANTIFIER	SEMANTICS
<i>təmáŋ</i>	‘all’
<i>təluʔ</i>	‘entire’
<i>təɹà</i>	‘some; a little’
<i>təsathàŋ</i>	‘half’
<i>ítʃáʔ</i>	‘few’ (both human and non-human referents)
<i>tijaʔ</i>	‘some, several’ (human referents only)
<i>anət à</i>	‘couple, several,’ (lit. two, one)
<i>kúlàŋà</i>	‘many’
<i>lanpən</i>	‘whole’
<i>muŋ.ɹaŋmúŋ</i>	‘whole’ (nouns of temporal measurement only)
<i>khələm</i>	‘together’

The word *təɹà* ‘some, a little’ is additionally used as a noun synchronically. Its dual word class status is attested by its ability to function both as a noun phrase head in constituency with a case marker, e.g. *təɹà nə* ‘by a bit’, and as a nominal modifier, e.g. *atʃhən təɹà* ‘a little money’. On the basis of this evidence we may infer that this particular quantifier was originally a noun in apposition to another noun, but subsequently became reanalyzed as a modifying nominal adjunct. Note that some of the examples of Table 4.10 have an initial syllable *tə* that could be a relict of the nominalizing prefix (discussed in §6.4.1, §7.2.4 and §8.3.4). However, with the exception of *təsathàŋ* ‘half’, which is feasibly related to the verb root *thàŋ* ‘sever’ and a homophonous lexical verbal suffix (see §8.5.1.7), none of these putative stems suggests a phonological or semantic relationship to any recognizable verb roots.

4.2.7.3. *Cardinal numerals*

The numeral system of Mongsen is now thoroughly decimal. Yet historical sources of data indicate that Mongsen and other Ao dialects originally had an extensive overcounting numeral system, a term first used by Menninger (1969: 76) for typologically unusual cardinal numeral systems found in Mayan languages, Ainu and Old Norse. The Ao overcounting numeral system was eventually replaced by a decimal system through the language engineering

efforts of the American Baptist missionaries beginning in the late nineteenth century (see §4.2.7.4 for further discussion).

Table 4.11. Mongsen and Proto-Tibeto-Burman numerals (where reconstructed)

NUMERAL	MONGSEN FORM	PROTO-TIBETO-BURMAN FORM
ONE	<i>akhət ~ akhəta</i>	<i>*(g-)tyak~*ka~*it~*t(y)ik</i>
TWO	<i>anət</i>	<i>*g-nis</i>
THREE	<i>asəm</i>	<i>*g-sum</i>
FOUR	<i>phəli</i>	<i>*b-liy</i>
FIVE	<i>phaŋa</i>	<i>*b-ŋa~*l-ŋa</i>
SIX	<i>təruk</i>	<i>*d-ruk</i>
SEVEN	<i>thəni</i>	<i>*s-nis</i>
EIGHT	<i>tshət</i>	<i>*b-r-gyat ~ *b-g-ryat</i>
NINE	<i>thuku</i>	<i>*d-kuw</i>
TEN	<i>thəla</i>	<i>*gip</i>
TWENTY	<i>məki</i>	<i>*(m-)kul</i>
THIRTY	<i>səm.la</i>	
FORTY	<i>li.ə</i>	
FIFTY	<i>thənəm</i>	
SIXTY	<i>.luk.la</i>	
SEVENTY	<i>nətəl ~ ni.ə</i>	
EIGHTY	<i>thi.l</i>	
NINETY	<i>thuku.l</i>	
HUNDRED	<i>nuklən</i>	<i>*r-gya</i>
THOUSAND	<i>mi.litfən</i>	<i>*s-toŋ</i>

Cardinal primary numerals ONE to TEN and cardinal numerals for the higher round numbers are listed in Table 4.11, together with the corresponding reconstructed numerals of PTB as proposed by Benedict (1972). Mongsen numbers ONE, TWO and THREE are bound roots occurring with the non-relational prefix *a-* in citation form. Matisoff (1997: 100) refers to recurrent prefixes in consecutive numerals as ‘prefix runs’ and attributes their historical development to analogical levelling and assimilation. The non-relational prefix *a-* of Mongsen numerals ONE to THREE corresponds with the reconstructed **g-* prefix of PTB. Benedict (1972: 94) finds that a single general root cannot be

established for the numeral ‘one’, hence the four reconstructed PTB forms listed in Table 4.11.

A phonologically reduced form *à* ‘one’ is likely to have its source in the Mongsen word for ONE. Mills (1926: 342) transcribes the Mongsen form for ONE as *ākhā* and the Chungli form as *ākhāt*. The synchronic Mongsen form is homophonous with the form for ONE recorded in the previous century for Chungli. When Mongsen speakers count consecutively, they use the form *akhət*; in citation, they give the form *akhəta*. The reduced form *à* is in the process of being reanalyzed as an indefinite article. This is cross-linguistically a widespread historical development, e.g. English *one* > *an*. It is used to introduce new information in narrative discourse. For an example, compare how it is used with a presentational function to determine *munṭhaŋ* ‘conference’ in (4.58). It is also used for determining or enumerating higher round numbers, e.g. *nuklaŋ à* ‘one hundred’ and *miitʃaŋ à* ‘one thousand’. The reduced form can co-occur with a nominal demonstrative, suggesting that synchronically it continues to be treated as a quantifier by the grammar. In (4.65) below it is enumerating ‘one tiger’, if it is assumed that the distal nominal demonstrative is performing the function of determining its noun phrase. But note also that it occurs in the noun phrase *khəʃa tsəŋti à* ‘a woven bamboo wall’, where its function this time seems to be more in keeping with that of an indefinite article than a numeral.

- (4.65) *hmapaŋ à ku akhu à tʃu tʃalipaŋ i ku khəʃa tsəŋti à takja litʃhà.*
 hmapaŋ à ku [a-khu à tʃu]_{NP} tʃalipaŋ i ku
 time one LOC NRL-tiger one DIST June/July PROX LOC
 [khəʃa tsəŋti à]_{NP} tak-ja li-tʃhà.
 flattened.bamboo woven.bamboo.wall one weave-CONT be-COP.PST
 ‘Once upon a time, at this time of June/July, a tiger was weaving a
 bamboo mat.’

The numeral *a-nət* ‘NRL-two’ loses its prefix when it is used to enumerate two entities that are formally treated as conjoined nouns in a complex noun phrase, as in (4.66). Its prefix is otherwise retained when enumerating the referents of a simple noun phrase, e.g. *kùŋhɿə a-nət* ‘two mushrooms’.

- (4.66) *kiphuɿ nə áwkla khə ajila nə áhlù nə zək.*
 kiphuɿ nə [a-úk-la khə a-ji-la nə]_{NP}
 owner AGT NRL-pig-F CONJ NRL-dog-F two
 a-hlú nə zək
 NRL-field ALL send.PST
 ‘An owner sent his dog and his pig to his field.’

The numeral TWO has undergone further grammaticalization in the Mongsen, Chungli and Changki dialects of Ao, where it is used in pronominal compounds to form the dual series (see §4.2.1).⁵

The initial aspirated bilabials of Mongsen *phəli* FOUR and *phaŋa* FIVE are cognate with the PTB *b- prefix and their roots closely correspond to those of the proto numerals. The root of Mongsen *tə-ɯuk* SIX is likewise phonologically close to the PTB form (cf. Table 4.11). Mongsen numerals *thəni* SEVEN, *tshət* EIGHT and *thuku* NINE present a mixed bag of aspirated dental stop initials and an aspirated dental affricate in the case of EIGHT. Assimilation and phonological levelling might account for the aspirated dental initials of the Mongsen numerals *thuku* NINE and *thəja* TEN.

The PTB numerals for *g-nis TWO and *s-nis SEVEN differ only in prefix. While the respectively corresponding Mongsen primary numerals *a-nət* and *thəni* have little in common with each other apart from dental nasals in their respective roots, it is perhaps significant that the root of Mongsen TWO is perfectly preserved in one of the words given for Mongsen SEVENTY, viz. *nətəi*. Furthermore, the now defunct form for Chungli SEVENTY provided by both Clark ([1893] 2002: 45) and Mills (1926: 342)⁶ is *tənəm səi mətsə*, literally ‘fifty and twenty’, a construction we may infer to be the fossilized residue of an ancient quinary system of PTB ($5 + 2 = 7$), as suggested by Benedict (1972: 93).

Mongsen formation is decimal, using *thəja* TEN as the base and a morpheme *-ɬə* as the linking element, e.g. *thəja-ɬə à-khət* ELEVEN, *thəja-ɬə à-nət* TWELVE, *thəja-ɬə à-səm* THIRTEEN ... *thəja-ɬə thùkù* NINETEEN and so forth. These are understood as ‘ten and one’, ‘ten and two’, ‘ten and three’... ‘ten and nine’ etc. The mid tones realized on the primary numerals ONE to NINE become low tones when they are used to form compound numerals. The decimal system described for the teens is consistent for all higher compound numerals, e.g. *səmɬa-ɬə tənùk* THIRTY-SIX, *nətəi-ɬə tshət* SEVENTY-EIGHT. In casual speech the non-relational prefix of ONE to THREE may be dropped in compound numerals, e.g. *nətəi nət* SEVENTY-TWO. Given the correspondence between the linking morpheme *səi* in the obsolete form of Chungli SEVENTY provided above and the morpheme *-ɬə* of the Mongsen compound numerals, it appears that the Mongsen linking morpheme is a phonologically reduced form of the discourse connective *təi* described in §4.2.10. Recall from the discussion of §2.6.6 that Mongsen speakers tend to resyllabify retroflex approximant codas as syllable onsets (also see Text 1 for examples demonstrating a correspondence between Mongsen *təi*

5. Presumably a dual series has also grammaticalised in the Longla and Yong-Yacham dialects of Ao, although this is yet to be confirmed.

6. The orthographies used in the historical sources have been standardized to accord with the present work.

and Chungli *asəi*, both of which express ‘and’ or ‘and then’ and are used as discourse connectives).

The Mongsen numeral *məki* TWENTY is obviously cognate with the proto-numeral **(m-)kul*. That a unitary word for TWENTY can be reconstructed leads Benedict (1972: 93) to surmise that the proto numeral system included a vigesimal unit. The lexical exponent for THIRTY is formed from the root of Mongsen *səm* THREE and part of the root of Mongsen *thəiə* TEN. Mongsen *liə* FORTY similarly uses a combination of parts of the roots of TEN and FOUR.

Mongsen *thənəm* FIFTY is noteworthy for having a form that is not based on the roots of FIVE and TEN. Matisoff (1997: 46) speculates that the word for FIFTY, consistently unusual in many languages of Nagaland, may originate from a different semantic field. He notes that Meithei has a word for FIFTY that might be associated with the word for ‘backbone’, ‘middle of the back’ or ‘ridgepole of a house’. This is suggestive of a half-way point, just as FIFTY is situated between the four two-digit round numbers of TEN, TWENTY, THIRTY and FORTY, and the four two-digit round numbers of SIXTY, SEVENTY, EIGHTY and NINETY. To give some strong support to his speculation, the word for ‘peak, crest of hill’ in Mongsen is *tənəm*. The initial consonant differs from the numeral for FIFTY in being aspirated. However, it corresponds perfectly with the word for ‘peak, crest of hill’ in Chungli, which lacks an aspiration contrast in stops according to Marrison (1967) and Gowda (1975).

Mongsen *jukiə* SIXTY resumes the pattern described above for THIRTY and FORTY. The words *nətəi* and *niə* SEVENTY, as previously noted, contain the root of TWO.⁷ The word for EIGHTY, viz. *thi*, is different again and unrelated to the primary numeral *tshət* EIGHT. The word *thuku* NINETY is almost identical to the primary numeral *thuku* NINE, plus it has a final retroflex approximant related to the synchronic form of the word for TEN.

The synchronic Mongsen forms for EIGHTY and NINETY differ significantly from the forms recorded for Mongsen by Mills (1926: 342–343) eighty years ago. The word he gives for Mongsen EIGHTY is *lira anekhi* ‘forty doubled’, and NINETY is recorded as *telang tüku*. The first constituent of the word for NINETY could have its source in the root *láj* (*hláj* in varieties with a voiceless sonorant series) ‘be long’. If so, NINETY used to be expressed as a ‘long nine’, although the position of *láj* as a pre-head modifier is atypical. The decimalizing influence of the American Baptist missionaries could be responsible for these changes in the forms for EIGHTY and NINETY, in addition to those discussed below in §4.2.7.4.

The non-cognate Mongsen HUNDRED is a compound formed from the nominal root *nuk* ‘machete’ (or *dao* in Nagamese) and the verbal root *láj*. This

7. *nətəi* contains the root of the Mongsen form for TWO, while the initial syllable of *niə* corresponds to **-nis*, the PTB root for TWO.

root is noted above to occur in an obsolete term for NINETY. Mills (1926: 60–61) describes a type of ceremonial sword-like weapon known as *noklang* and provides sketches facing p.61; on p.102 he reports that six- to seven-inch strips of iron shaped like *noklang* and called *chabili* were once used as a form of currency. Furthermore, he notes in the same passage that a bundle of one hundred *chabili* was termed *noklang*, and that this term came to be the means of expressing ‘one hundred’. This proposed etymology might possibly account for the non-cognate form of Mongsen HUNDRED. The older term for NINETY discussed in the preceding paragraph could be attributable to analogical levelling using the word for HUNDRED as the model.

The etymology of Mongsen THOUSAND *mi:itʃãŋ* is unknown. It also appears to be a compound, but I am not sure of the meaning of the root *mi:it*. A segmentally identical word is *mi:it*, the general (albeit pejorative) term for referring to trans-Dikhu River tribes, i.e. the Phom, Chang and Konyak, whose traditional territories are respectively located in the Longleng, Tuensang and Mon districts to the east and north-east. There is also a homophonous form *mi:it*, the name of an ant species. A nominal root *tʃãŋ* has a basic meaning of ‘seed’, but frequently occurs in compound nouns with an inherent meaning of multiplicity. Some examples are *mi-tʃãŋ* (person-seed) ‘population’, *səŋ-tʃãŋ* (wood-seed) ‘fruit’, and *tʃhən-tʃãŋ* (money-seed) ‘price’.

Ordinal numerals, numeral adverbs and distributives are all derived from the cardinal numerals by means of affixing morphology and are described in §7.6.

4.2.7.4. *Decimalization of the overcounting numeral system*

Until eighty or so years ago, all but one Ao dialect had an overcounting numeral system. This was not just confined to the teen numerals or just some dialects of the same language, as assumed by Matisoff (1997: 33), who referred to such systems as ‘subtractive’. The following extract from Clark’s 1893 *Ao Naga Grammar* firstly provides evidence that an overcounting system also existed in Chungli, secondly, that it extended into all of the higher compound numerals, and thirdly, that it was not really subtractive in character.

As will be observed from the above [i.e. a table of numerals], the Aos have distinct names for the digits, and the compounds are regularly formed up to sixteen, as, ten and one are eleven ‘teri ka’, ten and two are twelve ‘teri ana’, &c.; also from twenty to twenty-six, twenty and one ‘metsyri ka, [sic]’ twenty and two ‘metsyri ana’, &c. The same with thirty, forty, &c. But when the six is reached in the compounds, the succeeding ten seems to be anticipated, and we have for sixteen ‘metsy maben trok’ twenty not brought six, equivalent to the sixteen before twenty; ‘metsy; [sic] maben tenet, the seventeen before twenty,

&c. In the same manner from twenty-six to thirty, on reaching the six, thirty is anticipated, thus ‘semyr maben trok’ the six before thirty, twenty-six, ‘lir maben trok’ the six before forty, thirty-six, &c. Clark (2002: 45).

In a table of cardinal numerals on the same page, Clark gives Chungli examples for the numerals THIRTY-SIX, FORTY-SIX and FIFTY-SIX that are consistent with the pattern described in the extract. This indisputably demonstrates that the overcounting system was much more extensive than previously appreciated by Matisoff (1997). Furthermore, a comparison of numerals in the Ao dialects of Mongsen, Chungli, Changki, Longla and Yacham provided by Mills (1926: 342) shows that all but Yacham had an overcounting numeral system, at least in the teens. Based on the evidence from Clark (2002), we now know that this extended into the higher compound numerals in Chungli. By analogy, the overcounting system almost certainly must have extended into the higher compound numerals in all the dialects of Ao but was not documented.

In the Mongsen compound numerals FIFTEEN to NINETEEN that were documented by Mills (1926: 342), e.g. *müyi müpen tükü* NINETEEN, the linking word is the negated verb stem *mə-pən* ‘NEG-complete’. According to my Mongsen consultants, the literal meaning is understood as ‘(the) twenty not completed, (the) nine’, i.e. the nine before the twenty = ‘nineteen’. An elderly consultant, Mr Imti Luin, remembers the American Baptist missionaries replacing the overcounting system of Chungli with a decimal system in the school that was established after they came to Mangmetong village in 1911. The original overcounting system of the Mongsen dialect probably became obsolete due to the spreading influence of the prestige Chungli dialect and its newly crafted decimal system. This process of replacement must have begun shortly after the overcounting compound numerals were documented by Mills (1926). The verb root *pən* is nowadays restricted to the semantic field of temporal quantification, as in the following elicited example.

- (4.67) *pa akəm phəli pənukù.*
 pa a-kəm phəli pən-ukù
 3SG NRL-year four complete-ANT
 ‘He has reached four years of age.’

The reason for Yacham being recorded by Mills (1926: 342) as having a different numeral system can be attributed to its location in the Phom-speaking area (cf. Map 2). Yacham is heavily influenced by the Phom numeral system, a vigesimal pattern that is common to the Phom, Chang and Konyak languages of northern and eastern Nagaland.⁸ Presumably the overcounting system common

8. These three languages belong to a branch of Tibeto-Burman named the ‘sal’ languages, a term coined by Burling (1983) on the basis of a lexical innovation for

to Ao dialects became replaced by the vigesimal system of Phom through language contact. The Yacham term documented for TWENTY is *tāmōng khāt*, literally ‘one body’ which, as Mills (1926: 342) points out, possesses ten fingers and ten toes. The constituents of this compound numeral are cognate with the roots of *tə-maŋ* ‘RL-body’ and *a-khət* ‘NRL-one’ of Mongsen. I suspect that the word for TWENTY in Yacham is a calque on the Phom vigesimal term. In an annotated footnote in Mills (1926: 343fn.), Hutton adds a comparative note describing how he has heard a Phom use the expression ‘a whole man’ when asked how many people were present, meaning ‘at least twenty’.

Since all other Ao dialects apart from Yacham have documented proof of an older overcounting numeral system that uses a negated verb form as the linking word, the weight of evidence supports the assumption that this innovative system could possibly be a genetic feature of Ao. If so, the overcounting numeral system probably existed in the Yacham dialect before it was replaced by the vigesimal system as a consequence of language contact. Mills (1926: 2) reports that the inhabitants of Yacham spoke a dialect resembling Chungli but followed Phom or Konyak customs. This suggests an intensive contact situation. He was told by inhabitants of Yacham that they did not know whether they were Ao, Phom or Konyak, because they were not accepted by any of these tribes as kinsmen. My Chungli-speaking consultants report that they find the language spoken in the present-day Yong and Yacham villages largely unintelligible and attribute this to the influence of Phom.

A question remains: what motivated the replacement of the overcounting numeral system of Ao that was documented in all the dialects except Yacham eighty years ago? The missionary Mary Mead Clark (2002: 45) provides a likely answer.

This method of counting is very objectionable to children learning the use of figures, as in adding up a column if the amount is seventeen – ‘metsy maben tenet’ – the mind catches the twenty, and *two* is very likely carried instead of *one* to the next column. In the schools an effort is being made to discard the above irregularities, and count regularly thus, ‘teri trok’ sixteen, ‘metsyri trok’ twenty-six, &c.

Overcounting numeral systems have been documented in other languages of Nagaland apart from the dialects of Ao, e.g. Sumi (Hutton 1921b),⁹ Lotha (Witter 1888), Khonoma Angami (McCabe 1887), the Pochury dialect of southern Sangtam (on the basis of the Nagaland Bhasha Parishad data in

‘sun’. The other Sal languages of Marrison’s (1967) ‘A’ group – Tangsa, Nocte and Wancho – are spoken in the Tirap and Changlang districts of Arunachal Pradesh. Of these three, only Wancho has a vigesimal numeral system.

9. Hutton (1921b:272fn.) remarks that the overcounting numeral system was rapidly becoming obsolete at this time.

Matisoff 1997) and in the recorded dialects of the eastern Rengma group, the northern group of western Rengma and the southern group of western Rengma (Mills 1937). These have a similar structure to Ao in their use of a negated verb form as the linking word. The dialects of the Eastern Rengma group spoken in the villages of Meluri, Lephori and Sahunya had the most extensive overcounting system I have encountered in the older sources of Tibeto-Burman data. A table of numerals provided by Mills (1937: 293) shows that the overcounting system operates from SIXTEEN to NINETEEN until TWENTY is reached. Mill's description then suggests that from TWENTY-ONE to TWENTY-NINE, THIRTY is anticipated, and likewise for the remaining higher compound numerals up to ONE HUNDRED. The compound numerals given for TWENTY-ONE, THIRTY-ONE, FORTY-ONE and so on up to NINETY-ONE give support to this interpretation. Unfortunately these are the only higher compound numerals documented.

It is puzzling that Marrison (1967) did not document or even comment on the unusual overcounting compound numerals of Chungli and Mongsen, since he cites Clark (1893) and Mills (1926) as sources of data. In any event, his oversights in turn led Matisoff (1997: 33) to believe that dialects of the one language – Mongsen and Chungli being the cited examples – could differ in their numeral systems, and furthermore, that they had no genetic significance. Taking all this into consideration, the evidence suggests that the overcounting numeral systems of a significant number of languages of Nagaland represent a unique innovation that almost certainly has genetic significance. If so, then historical proof of overcounting might be used as a diagnostic tool to augment the linguistic sub-grouping of many languages whose higher level affiliations are yet to be definitively established. This new evidence could make a significant contribution to older sub-grouping methodologies relying principally upon geographical proximity and lexical correspondences based on old word lists of dubious reliability.

Of course, one would need to rule out bilingualism and language contact as the reason for the ubiquity of overcounting numeral systems in these “non-sal” languages of Nagaland, given that terms for numerals are frequently the first lexical items to be learned in the context of trade. But the very fact that overcounting is not found in any of the sal languages – which happen to be spoken in areas contiguous to where Ao, Sema, Sangtam and other Tibeto-Burman languages with erstwhile overcounting numerals systems are spoken – seems to suggest the greater likelihood of a genetic origin. An investigation of this would necessarily entail checking all the historical sources available, including the language engineering policies of the American Baptist missions if they were ever documented, to determine which languages originally had overcounting numeral systems. Such a project would be a highly suitable topic for future research.

4.2.7.5. Case marking clitics

The case marking clitics make up a closed class of eight members. These demonstrate the typical distributional behaviour of clitics in occurring after the last constituent of the noun phrase, and in exhibiting low selectional requirements with respect to their hosts. Forms of the case marking clitics are listed in Table 4.12 below and their functions in clauses are described in §5.3.1–§5.3.2.

Table 4.12. Mangmetong Mongsen case marking clitics

CORE		OBLIQUE			
		Non-local		Primary Local	
<i>nə</i>	Agentive	<i>nə</i>	Instrumental	<i>nə</i>	Allative
		<i>li</i>	Dative	<i>phinə</i>	Ablative
		<i>atəməkə</i>	Benefactive	<i>ku</i>	Locative
		<i>thən</i>	Comitative		

The case marking clitics can be divided into those that perform core marking functions, i.e. the agentive and, in non-canonical syntactic marking function, the dative (see §5.4.1), versus those that perform oblique marking functions, i.e. the instrumental benefactive, comitative, allative, ablative and locative. The case marking of oblique arguments is obligatory, whereas the case marking of core arguments is semantically and/or pragmatically motivated for a substantial number of verbs belonging to both transitive and intransitive classes. The O argument of a transitive clause receives no overt case marking. The A argument of a transitive clause can often be left unmarked for case if its semantic role is clear from the context or pragmatics in the instance of use. Alternatively, in some situations its semantic role must be made explicit by agentive case marking. While the S arguments of most intransitive clauses are unmarked, they too may be assigned agentive case marking under the special conditions discussed in §5.2.

In addition to these eight case marking clitics, there are eleven or so forms of what I will refer to as “nascent” postpositions, nearly all of which have demonstrable etymological sources in relational body part nouns and locational nouns. Most of these co-occur with the locative and allative case markers in an auxiliary-like function. An example is the root of *tə-ma* (RL-face) ‘face’ in *mukhuŋ ma ku* (tobacco.pipe FACE LOC) ‘on top of the tobacco pipe’. The nascent postpositions and their historical origins are discussed in detail in §5.3.3.

4.2.8. Time words

Time words have no fixed position in the clause and thus demonstrate the distributional behaviour of adverbs. They make up a sizable closed class of approximately thirty members. General time words are listed below in Table 4.13. Words specifically used to describe the traditional divisions of the day are given in Table 4.14.

Table 4.13. Time words

FORM	LEXICAL SOURCE (if known)	SEMANTICS
<i>thaku</i>	(probably: now+LOC)	'now'
<i>thani</i>	(now+day)	'today'
<i>thani asaŋ nɪʔ</i>	(now+day NRL-be.light day)	'nowadays'
<i>thajip</i>	(now+morning)	'this morning'
<i>thakaʔ</i>	(now+night)	'tonight'
<i>thalàtà</i>	(now+moon)	'this month'
<i>tóthi</i>		'always'
<i>təsi</i>		'until'
<i>nipa</i>		'cock crow; dawn'
<i>nɪʔ</i>		'one day'
<i>tʃhəmtaŋ</i>		'afternoon'
<i>asaŋ nəp</i>	(?NRL-be.light morning)	'tomorrow morning'
<i>zàkəm</i>	(?+year)	'last year'
<i>thakəm</i>	(now+year)	'this year'
<i>saŋkəm</i>	(?be.light+year)	'next year'
<i>.ràsəʔ</i>		'yesterday'
<i>zàkhəni</i>	(?+day)	'two days ago'
<i>zákhəni</i>	(?+day)	'three or more days ago'
<i>asaŋ</i>	(?NRL-be.light)	'tomorrow'
<i>zani</i>	(?+day)	'two days hence'
<i>zamni</i>	(?three+day)	'three days hence'
<i>məni</i>	(?four+day)	'four days hence'

Seven members of the time word class listed in Table 4.13 occur with the element *tha*. This encodes relative temporal proximity and is semantically

equivalent to ‘now’ when occurring with a nominal root expressing a measurement of time, e.g. *thani* (now+day) ‘today’. It also occurs with the locative case marker in *thaku* (now+LOC) ‘now’. The locative case marker encodes a grammaticalized meaning in this usage and is not so much marking a location in space as marking a location in time. The fact that *tha* occurs with a case marking clitic rules out the possibility of treating it as some kind of semi-productive prefix, yet it is not really a head either because it cannot be modified by any other nominal adjunct. It is probably best analyzed synchronically as a formative in these lexicalized compounds.

The other members of the class present an assortment of morphological structures. Some are monomorphemic. Others are formed with a nominal base *ni(?)* meaning ‘day’ plus a variety of formatives, the majority of which appear to be related to numerals. The unrecognizable formatives are probably now frozen in these lexicalized expressions. Tone is used in one pair of time words to differentiate between meanings, e.g. the L-L-M tones of *zàkhàni* ‘two days ago’ versus the H-M-M tones of *zákhəni* ‘three or more days ago’.

Table 4.14. Terms for divisions of the day

FORM	LEXICAL SOURCE	SEMANTICS
<i>nipa tàŋì</i>	(morning side)	‘mid morning’, 4am–9am
<i>nithuŋ</i>	(day+?reach)	‘day time’ 9am–1pm
<i>jà-tʃhəm taŋi</i>	(night-afternoon side)	‘early afternoon’, 1pm–4pm
<i>a-ja tàŋì</i>	(NRL-night side)	‘early evening’, 4pm–9pm
<i>a-ja t̄-ɪəm</i>	(NRL-night RL-middle)	‘late night’, 9pm– 1am
<i>nipa in</i>	(morning path)	‘early morning’, 1am–4am

Traditional terms for the six divisions of the day are mostly formed by the combination of a time word plus a nascent postposition. The latter are grammaticalized from lexical nouns expressing a location, e.g. *taŋi* ‘side’ and *inti* ‘path’. At an earlier stage of the language, these now-lexicalized expressions formed possessive compound nouns (cf. §5.3.3 for a discussion of the historical origin of nascent postpositions). The time divisions of the day are fairly approximate and allow a good deal of leeway on either side of the time frame to which each division refers.

4.2.9. Adverbs

Adverbs are verbal modifiers that do not have a fixed position of occurrence in the clause. They form a small closed class of just four members: *anukàʔ* ‘again, also, even, yet’; *sáʔ*, which similarly encodes a meaning of ‘again’; and the two verbal quantifiers *íjǎʔ* ‘much’ and *khən* ‘once’. The first of these has two phonologically reduced forms – *anuʔ* is used with the same meaning and distribution as the full form, while the meaning of *kàʔ* appears to have narrowed to ‘also’ in most instances of use. It is additionally used as an focus particle expressing an additive meaning (§4.2.14.4) and is recognizable in the morphological form of both the concessive converb suffix (§11.4.3.3) and the contemporative converb suffix (§11.4.3.4). Its grammaticalization as a marker of concessive clauses is likely to have been facilitated by the ‘even’ sense inherent to this adverb, while its ‘also’ sense provides a semantic pathway to the grammaticalized meaning of the contemporative converb suffix.

The expression of other adverbial meanings is handled by lexical suffixes on verb stems (§8.5), by simultaneous converb constructions used to encode the manner or attendant circumstances in which an activity is performed (§11.4.1.2), and by a small class of specialized converbs (§11.4.3).

4.2.10. Discourse connectives

Being a language that makes great use of dependent clause chaining, Mongsen would appear to have little functional need for the coordination of clauses. Nevertheless, two constructions utilizing subordinating morphology have been reanalyzed as types of discourse connectives and these are employed to create pragmatic structure in discourse. It is important to note that these do not exactly have a paratactic conjoining function that equates to the function of *and* in English. Consequently they are very rarely used for coordinating independent clauses, and never for coordinating nouns within complex noun phrases.¹⁰ As we shall see, the evolved functions of the discourse connectives more closely reflect the clause linking functions of their intermediate morphological sources.

Four prominent examples of discourse connectives are *tə-tʃhà-əɪ* (thus-do-SEQ), *tə-əɪ* (thus-SEQ), *tə-tʃhà-ku* (thus-do-LOC.CV) and *tə-ku* (thus-LOC.CV), all of which are translated as ‘and then’, ‘and’ or ‘after that’ by consultants who speak English, although underlyingly these forms are understood to imply ‘thus having done that...’, or perhaps ‘upon doing that...’ in the case of *tə-tʃhà-ku*

10. Only one of the two dozen or so consultants whose texts I recorded ever used a discourse connective to coordinate nouns within a complex noun phrase. I attribute this to the influence of English, of which he was a fluent speaker.

and *tə-ku*. The latter two connectives are formed with a locative case marker that has been reanalyzed as a marker of a sentential adjunct expressing a temporal presupposition.

If not for the presence of the quotative particle *tə* ‘thus’, the discourse connectives occurring with verbal roots are identical to two types of dependent clause linking devices that are referred to as *converbs* in this grammatical description. They undoubtedly share the same historical origin. The term *converb* was coined by Ramstedt (1903: 55) and subsequently adopted by Altaicists to describe non-finite verb forms in Turkic, but is increasingly being used to describe dependent verb forms with similar functions in other language families (see Haspelmath and König [1995] and references therein). Mongsén *converbs* are defined and described in §11.1–§11.4.

The distribution of discourse connectives formed with a sequential *converb* suffix, i.e. *tə-tʃhà-əI* (thus-do-SEQ) and *tə-əI* (thus-SEQ), was compared to those formed with the locative *converb* suffix, i.e. *tə-tʃhà-ku* (thus-do-LOC.CV) and *tə-ku* (thus-LOC.CV), to determine if there was a semantic basis to their use in texts. It was found that the sequential *converb* forms of discourse connectives are preferred for expressing sequential events, whereas the locative *converb* forms are marginally preferred for expressing a closer temporal relationship between the linked events. But in many cases speakers accept the swapping of one form for the other, suggesting that any original semantic contrast between the two forms has become blurred.

The discourse connective *tə-əI* is a phonologically reduced form of *tə-tʃhà-əI* that has lost its head, leaving just the carcass of the quotative particle and the sequential *converb* suffix behind to carry its meaning. The historical processes by which this has come about are not understood. The headless discourse connective *tə-ku* appears to have undergone the same process of decapitation.

The discourse connectives are restricted to occurring at clause junctures. Thompson and Longacre (1985: 211) report that ‘in some parts of the world verbs of highly generic meaning such as ‘do’ and ‘be’ ... are used as back-reference via adverbial clauses in a highly stylized and reduced manner so that they become in effect conjunctival elements’. This is exactly what is happening to these grammaticalized collocations based on the verb *tʃhà* ‘do’ in Mongsén. Thompson and Longacre (1985: 210) apply the term ‘summary-head linkage’ to the use of such verbs for summarizing the information of a previous paragraph and tying it to the next part of the discourse. The discourse connectives of Mongsén cannot strictly be equated with the coordinative conjunctions of the more familiar European languages, precisely because of this summarizing function.

The summarizing/linking functions of the discourse connectives are demonstrated in (4.68) and (4.69) below. In these extracts from a folklore text, Red Junglefowl’s eggs have just been cracked and scattered, and in her distress

she gives vent to her anger by scratching up the ground in a field that happens to belong to the god Lichaba. A discourse connective occurs at the beginning of each of the (a) and (b) clauses of (4.68), where it summarizes what has preceded and relates that information to what follows in the narrative. This helps to give the text a cohesive structure that consecutively builds on the narrative base by accretion. All discourse connectives involving the sequential converb marker carry the characteristic intonation contour of the sequential converb described in §3.4.2, a feature which further attests to a shared historical origin.

- (4.68) a. *tətʃhəɪ líʃápá? hlù ku təlú? sùsazəktʃuk.*
 tətʃhà-əɪ líʃá-pà? hlú ku təlú? sùsa-zək-tʃuk
 thus-do-SEQ PN-M field LOC entire scratch-SEND-PFV.PST
 ‘And then, in Lichaba’s field, everything was completely scratched up.’
- b. *tətʃhəɪ líʃápá? tʃu áhlú nə walikà?*
 tətʃhà-əɪ líʃá-pà? tʃu a-hlú nə wa-likà?
 thus-do-SEQ PN-M DIST NRL-field ALL go-CONTEMP
 ‘And then, when Lichaba went to the field, ...’

The sentences of (4.69a) and (4.68b) demonstrate the summarizing/linking function of the discourse connective formed with the locative converb. These are taken from the same narrative as the sentences of (4.68a) and (4.67b) and occur a few clauses later in the tale. Rather than merely linking a series of sequential events, here *tətʃhà-ku* suggests a closer temporal relationship between the time that the question of (4.69a) is uttered and the subsequent response of (4.69b). This closely accords with the meaning that the locative case marker-*ku* expresses in its grammaticalized function as a converb suffix (see §11.4.3.7).

- (4.69) a. “*nəŋ tʃútə ùphələla həntsə kəpsàzəkʃuk.*”
 nəŋ tʃútə ùphəla-la həntsə kəp-sa?-zək-tʃuk
 2SG why red.junglefowl-F egg strike-SEPARATE-SEND-PFV.PST
 [Lichaba asks] “Why did you scatter and break Red Junglefowl’s eggs?”
- b. *tətʃhàku “áj, nì təmaku tsəhŋiphu phuja lilikà?”*
 tətʃhà-ku áj nì tə-ma ku tsəhŋiphu
 thus-do-LOC.CV EXCLM 1SG RL-surface LOC sunray
 phu-ja li-likà?
 expose.to.sun-CONT be-CONTEMP
 Upon [Lichaba asking, Earthworm replied] “Hey, when I was on the surface sunning myself, ...”

Discourse connectives can also be formed with lexically specific verb roots. To illustrate, the root of the verb *səmtsə* ‘ask’ could be used in place of the non-specific generic verb *tʃhà* ‘do’ in (4.69b) to express a more explicit meaning, e.g. *tə-səmtsə-ku* ‘Upon asking...’.

Two other unrelated forms of conjunctive elements are occasionally encountered in the speech of older Mangmetong Mongsen speakers. One of these, *alà* ‘and then’, has the same clause juncture distribution as the discourse connectives discussed above. More often than not, however, it appears to correspond with a change in topic. Its etymology is unknown. Clark (1990: 134) mentions in his Chungli dictionary that some (presumably Chungli) villages use *ala* and *alakong* as conjunctions, but offers no clues as to a possible historical origin. Assamese and Nagamese both use the semantically similar form *alo* as a conjunction meaning ‘and’, but the phonological form of this word seems too distant from Mongsen *alà* to suspect that it could have been borrowed from either of those languages.

Lastly, a form *tʃùku* that may have resulted from a fusion of the locative case marking clitic *ku* with an unrecognizable formative *tʃù* is used to link a cause to an effect and is accordingly glossed ‘consequently’. Both *alà* and *tʃùku* occur in the (4.69) below. The use of different types of converb clauses to build structure in narrative discourse is further described in Chapter 11.

- (4.70) *alà tʃùku pa sə məzəm sə nàtʃuk.*
 alà tʃùku pa sə məzəm sə nà-tʃuk
 and.then consequently 3SG ANAPH poison ANAPH cure-PFV.PST
 ‘And, consequently he was cured of the aforementioned poison.’

4.2.11. Phrasal conjunction

The phrasal conjunction *khə* is the sole member of this closed word class. Its function is to encode a paratactic relationship holding between two or more nouns within a complex noun phrase. The coordination of nouns at the level of the noun phrase is its only attested function.

The maximum number of conjoined nouns encountered in texts or natural speech is three. Consultants balk at attempts to elicit more than three conjoined nouns within the same noun phrase. A noun phrase with two conjoined nouns is optionally quantified by the root of the numeral *a-nət* ‘NRL-two’, which has quantifier scope over both the noun constituents of a complex noun phrase. This is demonstrated by the following Waromung Mongsen example, in which the form of the phrasal conjunction is *kha* (see §6.5 for discussion of hierarchy in complex noun phrases).

- (4.71) *í pá? jutfi lùŋkhəm kha nuk.ɿaŋ jim nət ùŋtəpə? jutfi.*
 í pá? jutfi lùŋkhəm kha nuk.ɿaŋ jim nət ùŋtəp-pà?
 EMPHAT story Longkum CONJ Nokrang village two fight-NR
 jutfi
 story
 ‘This story is the story of the war between the two villages of Longkum
 and Nokrang.’

A complex noun phrase composed of up to three nouns conjoined by *khə* is case-marked according to the semantic and/or syntactic role that the complex noun phrase plays within the clause.

- (4.72) *ajà! nì aja khə apa? taŋ nə tʃəpá?*
 ajà nì a-ja khə a-pa? taŋ nə tʃəpá?
 EXCLM 1SG VOC-mother CONJ VOC-father SIDE ALL what
 ‘Aya! What will I [take] to Mother and Father?’

One elderly male speaker optionally marked each of the conjoined nouns individually with the phrasal conjunction, as in the following example. On the basis of the data at hand, this only appears to be possible when a complex noun phrase receive no overt case marking.

- (4.73) *təɿ waɿ, am khə aku? khə niŋəɿ,*
 tə-əɿ wa-əɿ am khə aku? khə niŋ-əɿ
 thus-SEQ go-SEQ aspidistra.sp. CONJ gourd.sp. CONJ be.ripe-SEQ
 ‘And, after going [to the field], after the bottle gourds and aspidistra
 plants had ripened, ...’

The phrasal conjunction is semantically similar to the comitative case marker *thən* (described in §5.3.2.7), in that both morphemes express a meaning of ‘with’ or ‘together with’. They differ crucially in distribution, however, because only the phrasal conjunction can be used to conjoin two nouns within a complex noun phrase. This is, in fact, the only environment in which it may occur.

The morphological form and meaning of the phrasal conjunction *khə* is very close to the associative case marker *kə* of Meithei (Chelliah 1997: 128), suggesting a possible shared diachronic origin. If so, we might speculate that a reanalysis of *khə* has resulted in the development of a paratactic conjoining function in Mongsen, and that the older comitative case marking function is now performed uniquely by the synchronic comitative marker *thən*, which in turn appears to have a verbal origin (see discussion of §5.3.2.7 and §8.5.1.11).

Furthermore, this might explain why the individual marking of conjoined noun phrases with *khə*, such as that seen above in (4.73), is only possible when the complex noun phrase is not overtly case-marked. I infer from this constraint that *khə* once functioned as a relational marker in the same slot as the case marking clitics and encoded the comitative case, prior to its reanalysis as a phrasal conjunction.

4.2.12. Interjections

Interjections are uninflected, single-word responses to interrogative and declarative clauses. The interjection *hāj?* ‘okay’ is used to signal agreement to a proposition, and *hàw?* ‘yes’ gives an affirmative answer to a question. A curt response in the negative to an invitation (to drink tea, smoke a cigarette etc.) can be made with the interjection *mà?* ‘no’ (see §10.4.3). The salutation *saləm* ‘Greetings!’ is a loanword (probably via Persian and Urdu) from Arabic *salām* ‘peace’.

The negative particle *nuŋ* can be used as an interjection to show disagreement to a proposition, in addition to its use as a negator of verbless identity clauses (see §9.1.2). The following example was uttered when I asked a consultant for an explanation about an aspect of the grammar. As he began to explain, another Mongsen speaker animatedly interrupted with the following string of interjections to indicate that my question had been misinterpreted.

- (4.74) *nuŋ nuŋ nuŋ.*
 nuŋ *nuŋ* *nuŋ*
 NEG.PTCL NEG.PTCL NEG.PTCL
 ‘No, no, no.’

4.2.13. Exclamations and onomatopoeia

Exclamations comprise a semi-open class that accepts spontaneously created new members. The recurring forms are listed in Table 4.15 below, together with their associated meanings. Exclamations are not noted to have any characteristic properties associated with their phonological form or phonotactic structure.

Table 4.15. Common exclamations

FORM	SEMANTICS
<i>újàtáj</i>	‘Surely not!’ (incredulity)
<i>ahəj, áj</i>	‘Hey!’ (seeks attention)
<i>ájà</i>	‘Oh no!’ (dismay)
<i>ájít</i>	‘Wow!’ (surprise)
<i>à</i>	‘Ah!’ (satisfaction)
<i>ə</i>	‘Ah, yes’ (agreement)
<i>wá</i>	‘Hey!’ (seeks attention; can also give warning)

Onomatopoeic elements are frequently introduced in the clause by the quotative particle *tə* ‘thus’ (§4.2.14.1). This can make their clauses structurally similar to the embedded complements of direct speech, particularly when the onomatopoeic word functions as the O complement of a matrix clause predicate (see §11.7.1 for discussion and additional examples).

- (4.75) “*pəɪək kəɪək*” *tə phəɪukli təmijuŋ sə à hjùlakli tʃali*,
 pəɪək-kəɪək *tə* phəɪuk-li tə-mijuŋ sə à
 ONOM-ECHO thus spread-ALT.IT RL-finger ANAPH one
 hjù-lak-li tʃàʔ-li
 snap.off-DESCEND-ALT.IT consume-ALT.IT
 “‘Crackle-crackle”, [she] spreads out [the hand over the fire], then
 snaps off one finger, eats it, and then snaps off another, eats it, and so
 on, over and over, ...’

Table 4.16. Examples of onomatopoeia

FORM	SEMANTICS
<i>huhuhu</i>	‘I’m here!’ (informs of presence)
<i>pəɪək kəɪək</i>	crackling sound of something dry being broken up
<i>púŋ</i>	‘Boom!’ (sound of explosion)
<i>phàk</i>	sound of pumpkin, gourd, etc. being chopped open
<i>kələmləm</i>	sound of rolling thunder

4.2.14. Particles and clitics

The particles form a smallish closed class of non-inflecting words that do not belong to any of the other word classes recognized in this chapter. The particle class is in turn subsumed by five sub-classes ranging in size from one member (viz. the quotative, reportative and topic particles) to at least a dozen members (viz. the illocutionary force particles). The syntactic functions and meanings encoded by each sub-class of particle are respectively discussed and illustrated by examples in the following sections.

4.2.14.1. Quotative particle

The quotative particle *tə* ‘thus’ is a complementizer that marks the embedded complements of speech act verbs and other complement-taking predicates. It can be used to introduce a new clause, just as *thus* can in English (also compare the use of the quotative particle in the discourse connectives described in §4.2.10). In its function as a complementizer, the quotative particle occurs immediately after a direct speech complement. Both of these functions are demonstrated in the example of (4.76).

- (4.76) *tə lítʃápá? nə “tshəaŋ” tə məla.*
tə lítʃá-pà? nə tshə-aŋ tə məla
 thus PN-M AGT take-IMP thus permit.PST
 ‘Thus, Lichaba permitted it [saying] “Take her”.’

- (4.77) *anukà? wali “pi mətət sè məmətətla”, təlì “mətətə.ìù?”*,
anukà? wa-li pi mətət sè mè-mətət-la
 again go-ALT.IT PROX know QPTL NEG-know-NEG.PST
tə-li mətət-əɪ-ù?
 thus-ALT.IT know-PRES-DEC
 ‘Again, each time they went [deeper into the jungle he asked] “Do you know this place or not?”, [and each time she answered] “I know it”, over and over ...’

The quotative particle has grammaticalized an additional function as a base by proxy for some verbal morphology, such as the alternating iterative converb suffix *-li* of (4.77) above, and other converb suffixes discussed in §4.2.10. This allows the scope of the proposition encoded by the converb suffix to extend over multiple dependent clauses in a complex sentence. It has a similar function of extending the propositional scope of the subordinating particle *aj* over

embedded clauses encoding reported speech. Further examples and discussion, can be found in §11.6.

The etymological source of the quotative particle cannot be determined at present. A likely suspect is a verb of saying (Saxena 1988), but no obvious candidates have been encountered in the verb class. It may have grammaticalized from a morpheme that no longer functions as a lexical root synchronically.

4.2.14.2. *Reportative particle*

The reportative particle *tàɪ* is the only marker of the category of evidentiality in the grammar of Mongsen. It occurs as the last constituent of the clause and signals that the speaker is not the original source of the information. Its clause-final distribution allows the optional declarative mood clitic *-ùʔ* to select it as a host (see §4.2.14.6).

One day a Mongsen speaker called out the following sentence to her son after I said I would take my leave, the intention being that he could come to say good-bye to me. She encoded this statement as hearsay by using the reportative particle clause-finally.

- (4.78) *pa wa-ɹù tàɹ-ùʔ*
 pa wa-ɹù tàɹ-ùʔ
 3SG go-IMM REP-DEC
 ‘He [said he] is going.’

The reportative particle is suspiciously similar to the quotative particle (§4.2.14.1) in both phonological form and tone, yet the two particles maintain distinct syntactic functions and have different clausal distributions. The quotative particle is used to introduce new clauses, acts as a base by proxy for converb markers and a subordinating particle, and marks complements embedded in the immediate pre-head position of the clause, whereas the reportative particle only occurs clause-finally in the post-head position. Nevertheless, in view of the phonological and suprasegmental resemblance, it is likely that these two categories of particle share a common diachronic source, whatever that may be. It is noteworthy that Jinghpaw [Tibeto-Burman: Bodo-Konyak-Jinghpaw] is also reported to use a homophonous particle (having the form *da* in that language) for marking the categories of quotative and evidentiality, yet a related form is similarly not found to occur synchronically as a verb of saying (Saxena 1988: 377). It seems that it is typologically unusual for a language to maintain distinct functions for quotative and reportative marking, whereas the grammaticalization of evidentials and complementizers from verbs

of saying is a widely attested typological pattern (Saxena 1988, Heine and Kuteva 2002: 261–268).

4.2.14.3. Topic particle

The topic particle *la* marks the topic under discussion, i.e. what the proposition is about. It can have scope over a word, a phrase or a whole clause, depending upon the domain it is marking as a topic.

The sentence of (4.79) presents a textual example of a topicalized noun phrase. Just prior to this assertion, a demand has been made for one of three items. In the addressee’s reply the pragmatically salient noun phrase *ipi* ‘this’ is marked as the topic of the utterance, to which the following predicate and the clause’s illocutionary force particle functions as the comment. The topic and its comment in turn function as the O complement of the matrix clause’s verb of speech.

- (4.79) *“ipi la mək̀hìm̀l̀ù? t̀” t̀ laŋli.*
ipi la m̀-kihì?-m̀-à-ù? t̀ t̀ laŋli
 PROX TOP NEG-give-DESID-PRES-DEC PTCL thus reply.PST
 “‘This’ [he] replied, “[I] don’t want to give.”

Apart from occurring in noun phrases, the topic particle also frequently occurs with time words, e.g. *thani la* ‘today TOP’, *thaka? la* ‘tonight TOP’ and *thajip la* ‘this morning TOP’. Time word + topic particle pairings are so common that they seem to have become conventionalized in the grammar of Mongsen.

The topic particle is used with clausal scope in dependent clauses. Under these circumstances the whole of the dependent clause marked by *la* functions as the topic and its matrix clause functions as the comment.

- (4.80) *t̀tʃha, inisa? – iŋliʃ wət ɹapit t̀ tʃapa? la – s̀à:à:ə t̀ inisala tʃu nə pa la m̀səm̀l̀.*
t̀-tʃhà-ə inisa? iŋliʃ wət ɹapit t̀ tʃa-pà? la
 thus-do-SEQ rabbit English word rabbit thus call-NR TOP
s̀à:à:ə t̀ inisa?-la tʃu nə pa la m̀-ə-m̀-à
 animal thus rabbit-F DIST AGT 3SG TOP NEG-die-DESID-PRES
 ‘And then, *inisa?* – which is called “rabbit” in English – that animal, Rabbit, she doesn’t want to die.’

The complex sentence of (4.80) contains a relativized noun phrase marked by the topic particle. This identifies *inisa?* ‘Rabbit’ as the topic. Immediately

prior to this sentence, the narrator explains how all the animals have been ordered to come one-by-one to Tiger to be eaten each day, so the first use of the topic particle in (4.80) has a contrastive function as well. To further clarify its pragmatic status, the speaker again establishes the topical salience of ‘Rabbit’ after the first mention by rephrasing the topic a number of times, eventually marking the coreferential pronoun *pa* ‘she’ with another topic particle. This is followed by the assertion that she doesn’t want to die, which functions as the comment to both the topic-marked relative clause and the topic-marked pronoun.

The topic particle is used in a similar way to form another type of dependent clause that is functionally equivalent to a relative clause formed with a nominalized verb stem, although it is structurally very different. In (4.81) the direct speech preceded by the topic particle, i.e. *tʃəpá? tʃhàmì.ɪ̀*, functions as the topic, while the imperative of the matrix clause provides the comment.

(4.81) *pa thaku* “*tʃəpá? tʃhàmì.ɪ̀ la tʃhaŋ*”.

pa thak ku tʃəpá? tʃhà-mì-əɪ la tʃhà-aŋ
3SG PLACE LOC what do-DESID-PRES TOP do-IMP

To it [i.e. the lightning, he said] “Whatever you want to do, do it!” [in order to placate it].

The structure of this complex clause bears a number of similarities to the relative-correlative construction found in Indo-Aryan languages (Masica 1991: 410–415) and in other languages discussed in Keenan (1985: 163–170). Firstly, since there is no form class of relative pronoun in Mongsén, an interrogative pronoun must be used instead to introduce the dependent clause. According to Nadkarni (1975), this is also reported to have occurred in Dravidian languages after they borrowed the relative-correlative structure through linguistic convergence with Indo-Aryan languages. Secondly, the dependent clause has the formal structure of an independent clause because it contains a finite verb, yet it is functionally in a dependent relationship to the matrix clause. An intrinsic element of the dependent clause in (4.81) is the topic particle. This separates it from the matrix clause, establishes its pragmatic status as a topic, and encodes its syntactic dependency. This is the characteristic structural presentation of the Mongsén-type of topicalized relative-correlative construction – the dependent clause precedes the matrix clause, an interrogative pronoun functions as the shared argument inside the dependent clause and the topic particle marks the boundary between the subordinate clause and its superordinate clause (see §6.6.6 for further description).

The use of a topicalized dependent clause as the functional equivalent of a relative clause is common amongst urban Mongsén speakers who speak Nagamese and English, but is considerably rarer in the speech of native

speakers living in villages, the latter preferring to relativize using the relativization pattern exemplified by the dependent clause of (4.80) above (see §6.6 for examples and description of all relativization strategies). At present it is not clear what influence Indo-European languages may have had on the choice of relative clause formation in the Mongsen of these bilingual speakers.

A diachronic relationship between topic particles and markers of dependent clauses has been documented in a number of different language families (e.g. see Haiman [1978] and Bickel [1993]) and is attested in Mongsen in the form of the conditional converb suffix *-pàla* and the circumstantial converb suffix *-kùla*. The form *-pàla* is a fusion of the general nominalizer *-pàʔ* (see §7.4.2) and the topic particle *la*, while the circumstantial form is suspected to be a fusion of the locative case marker *ku* with the topic particle *la*. Conditional and circumstantial converbs are discussed under §11.4.2–§11.4.3.

The most frequent use of the topic particle is to mark a contrastive topic. In the folklore narrative from which the following example is taken, an owner sets a task for a dog and a pig in order to determine who deserves to sleep in the house and who should sleep outside. The pig is the topic in the text immediately preceding this clause. To signal a shift in attention to the dog, the speaker now marks the new topic with the topic particle and reinforces its pragmatic status by repeating the topic marking on the following pronoun.

- (4.82) “*təɪ ajila tʃu la, ə nəŋ la, aki táŋ ku tʃukəjùʔ.*”
 tə-əɪ a-ji-la tʃu la
 thus-SEQ NRL-dog-F DIST TOP
 ə nəŋ la a-ki táŋ ku tʃuk-iʔ-i-ùʔ
 EXCLM 2SG TOP NRL-house just LOC keep-CAUS-IRR-DEC
 ‘‘And, as for Dog, Ah! Only you will be kept in the house’’ [he said].’

The topic particle is homophonous with the feminine semantic gender marker *-la*, but is unrelated to this morpheme (see §7.5.2 for description). The feminine semantic gender marker has a single position of occurrence and attaches directly to the nominal head, whereas the topic marker usually occurs, clitic-like, as the last constituent of the topicalized noun phrase. Their different phrasal distributions are demonstrated by the first noun phrase in (4.82) above.¹¹ When the topic particle is used to topicalize a single word, however, it can occur as a noun phrase constituent immediately after the head. This is the structure found in the following text example, in which the narrator is clarifying just who it is that approaches Earthworm after he has some momentary confusion about referents.

11. A convention in Mongsen folklore stories is for all animals to be personified as having feminine gender.

(4.83) a. *tsə̀lə̀la aʔ ... tsə̀lə̀la taŋ nə waɪ̀ə̀, líʔá-páʔ la nə,*

tsə̀lə̀laʔ-la aʔ
earthworm-F INTJ

tsə̀lə̀laʔ-la taŋ nə wa-ə̀ɪ líʔá-pàʔ la nə
earthworm-F SIDE ALL go-PRES PN-M TOP AGT
'Earthworm, ah, ... [Ø_i] goes over to Earthworm, Lichaba._i'

b. “*nàŋ ʔfúʔə̀ uphə̀la hə̀ntsə̀ kə̀psə̀zəkʔʔuk.*”

nàŋ ʔfúʔə̀ uphə̀la hə̀ntsə̀ kə̀p-saʔ-zək-ʔʔuk
2SG why red.junglefowl egg strike-SEPARATE-SEND-PFV.PST
“‘Why did you crack and scatter Red Junglefowl’s eggs?’ [asks
Lichaba].’

The narrator stops abruptly on the first mention of *tsə̀lə̀la* ‘Earthworm’ when he realizes it should be the goal of movement and not the S argument of *waɪ̀ə̀* ‘goes’. In a self-repair on the fly, he reassigns ‘Earthworm’ oblique case marking and continues on, but realizes that the clause now lacks an S argument because he just turned it into an oblique argument. He subsequently adds the noun phrase *líʔá-páʔ la nə* ‘Lichaba’ as an afterthought in the postposed position, following a pause (represented in the italicized line by a comma). The topic particle is used contrastively to make it clear that the referent of the S argument is in fact Lichaba and not another referent that might otherwise be construed as coreferential with the omitted noun phrase. Such a misinterpretation would be highly likely because ‘Lichaba’ does not rate a single mention in the preceding four clauses of the narrative.

The atypical agentive marking on the S argument is probably attributable to the fact that the narrative’s following clause (i.e. [4.83b]) is a direct speech interrogative and syntactically a complement in O function (see §11.7 for the description of complementation). The agentive-marked argument of (4.83a) is therefore Janus-like, functioning as the postposed S argument of that clause and as the understood A argument of the following direct speech clause.¹²

Lastly, it is noted that the topic particle is also homophonous with the suffixing element of the discontinuous past tense negative morpheme and may possibly share a diachronic origin. Evidence in support of this is presented in §8.5.9.

12. It is unusual for S arguments of intransitive verbs of motion to be assigned agentive marking unless licensed by the pragmatically-marked circumstances discussed in §5.3.2.

4.2.14.4. Discourse focus particles

The most textually frequent member of this sub-class is the restrictive focus particle *táj*. This expresses a meaning equivalent to ‘just, only’. A phonologically reduced form *áj* shares the same distribution and meaning. Both reduced and unreduced forms have no distributional constraints with respect to the types of clausal constituents with which they may occur. Within a noun phrase they are usually found in the immediate post-head position, where they restrict their focus to the preceding constituent, e.g. (4.84). Within the clause they may variously occur after a time word, e.g. (4.85), a converb, e.g. (4.86), or even a quotative particle functioning as a discourse connective to introduce a subsequent sentence, e.g. (4.87).

- (4.84) “*nì táj nə tʃhəniŋù?*” *tə sa.*
 nì táj nə tʃhəniŋ-ùʔ tə sa
 1SG just AGT work.PST-DEC thus say.PST
 ‘“Only I worked”, [she] said.’

- (4.85) *asaŋ áj waɪ,*
 asaŋ áj wa-əɪ
 tomorrow only go-SEQ
 ‘Only after going tomorrow ...’

- (4.86) ... *pa ɪəkə táj hjaɪ táj,*
 pa ɪəkə táj hja-əɪ táj
 3SG stalk-SIM just go+level-SEQ just
 ‘... it [i.e. a tiger] just stalking, after just going over ...’

- (4.87) *tə áj ùphəla la tʃu kà? həntsə tʃu məmuktʃhətla.*
 tə áj ùphəla-la tʃu kà? həntsə tʃu
 thus just red.junglefowl-F DIST also egg DIST
 mə-muk-tʃhət-la
 NEG-brood-ABIL-NEG.PST
 ‘So! Red Jungle fowl also wasn’t able to hatch her eggs.’

When the restrictive focus particle *táj* occurs in front of the matrix verb, the clause’s predicate can be interpreted as the target of its focus. In the following example it is additionally intensified by the verbal quantifier *íjǎ?* ‘much’.

- (4.88) *pa íjá? táŋ phìləm.*
 pa íjá? táŋ phìləm
 3SG much just think.PST
 ‘She really thought a lot.’

The discourse particle *kà?* encodes an additive meaning of ‘also, yet, again’. This appears to be a phonologically reduced form of the adverb *anukà?* ‘again, also, even, yet’ discussed in §4.2.9. It is used to focus attention on a clausal constituent, which may be a noun phrase, a time word or a verb. When the focused constituent is a noun phrase, it often serves to distinguish that noun phrase’s referent from another recently mentioned referent, as in (4.89). When the focus of *kà?* is directed at a clausal predicate, its purpose is to highlight the event/state that the predicate represents, usually with implied or actual reference to another event or state, as in (4.90).

- (4.89) *ahənzala kà? pa saləm hən.*
 a-hən-za-la kà? pa saləm hən
 NRL-chicken-DIM-F also 3SG share take.PST
 ‘Little Chick also carried her share.’

- (4.90) “*thajip kà? nən tuɹ au kà? puŋə.ùtə*”, *tə sa tʃu.*
 thajip kà? nən tuɹ au kà?
 this.morning again 2SG beauty good again
 puŋ-ə.ɪ-ù? tə tə sa tʃu
 good-PRES-DEC PTCL thus say.PST DIST
 “‘This morning you’re again looking good and handsome”, she said
 like that.’

4.2.14.5. *Illocutionary force particles*

A sizeable sub-class of illocutionary force particles is used to express the propositional attitude of the speaker. All but one of these occurs in the clause-final position. The exception is the interrogative particle *sə*, which is used in an A-not-A type of disjunctive interrogative construction in addition to occurring clause-finally (see §5.1.2.1 for description). The following example illustrates the use of *sə* in the A-not-A type of interrogative clause.

- (4.91) *hniŋəɪ sə məhniŋla.*
 hniŋ-əɪ sə mə-hniŋ-la
 ripe-PRES QPTCL NEG-ripe-NEG.PST
 ‘Are [they] ripe or not?’

The interrogative particle *sə* may alternatively be used as a clause-final interrogative particle, in common with the other interrogative particles *tʃhaɪú?*, *pà* and *ní*. Examples of interrogative clauses signalled by interrogative particles are presented in the description of clause types in §5.1.2.

The particle *tə* asserts the truth of a proposition and carries a meaning of ‘certainly, indeed’. An example occurs above in (4.90). On the basis of semantic similarity and phonological form, it may be related to the homophonous quotative particle *tə* ‘thus’ (§4.2.14.1). The possibility of it co-occurring with the quotative particle in (4.90) demonstrates its synchronic independence.

The particle *wa* carries a meaning equivalent to ‘as I remember it’. It often occurs at the end of folklore narratives and other types of assertions involving recollections. Taking its meaning into consideration, *wa* might be more correctly analyzed as a type of evidential particle that asserts the truth of a proposition according to the recollections of a speaker.

- (4.92) *tə hjutsə tʃu thəm wa.*
 tə hjutsə tʃu thəm wa
 thus story DIST end.PST PTCL
 ‘Thus the story ended [as I remember it].’

One day when playing a previously recorded narrative I asked a consultant if he could recognize the voice of the narrator, to which he gave the following reply.

- (4.93) *ləmniŋ khàpà? wa.*
 ləmniŋ.khà-pà? wa
 PN-M PTCL
 ‘[I believe it’s] Lemning Kaba.’

The particle *atì* or *atə* imparts an emphatic meaning, often suggestive of mild obligation or recommendation. It figures prominently in procedural texts and other instructive discourse.

- (4.94) *alà íjá? aɹukə tʃhuɹuj atì.*
 alà íjá? a.ɹu-əkə tʃhu.ɹu-ì atì
 and.then much be.good-SIM cook-IRR PTCL
 ‘And then it should be properly cooked.’

The particle *ná* seeks the confirmation of a suggestion or accord to perform an action.

- (4.95) *ɹənəmukù, sà:ùù ná.*
 ɹənəm-ukù sa-ùù ná
 be.ready-ANT say-IMM PTCL
 ‘I’m ready – I’ll tell [the story], if it’s okay.’

The particle *já?* similarly encodes a presupposition and seeks its confirmation. The possibility of the declarative mood clitic occurring in the clause indicates that this particle cannot be interpreted as expressing interrogative mood (see §4.2.14.6 for description of optional declarative mood marking).

- (4.96) *nàŋ pazaɪ nə wajuù? já?*
 nəŋ pazaɪ nə wa-ì-ù? já?
 2SG bazaar ALL go-IRR-DEC PTCL
 [I presume that] ‘You’re going to the bazaar.’

The particle *tʃi* suggests finality, that a conclusion has been reached.

- (4.97) *təɪ tuŋət sətəjpa? sanaw? tʃi.*
 tə-əɪ tuŋət sàti-ì-pà? sana-ù? tʃi
 thus-SEQ 3DU marry-IRR-NR speak.PST-DEC PTCL
 ‘And the two of them talked about the marriage.’

The particle *ɹə* is found in irrealis clauses and attempts to give an assurance that the predicted event or state will actually occur.

- (4.98) *nì aja ku tʃhuwajuù.ɹə.*
 nì a-ja ku tʃhuwa-ì-ù? ɹə
 1SG NRL-night LOC emerge-IRR-DEC PTCL
 ‘I’ll return at night.’

The particle *àw?* expresses surprise, that something contrary to expectation has occurred.

- (4.99) *atʃak sə athi səŋ àw?*
 a-tʃak sə a-thi səŋ àw?
 NRL-paddy ANAPH NRL-field.hut be.full.PST PTCL
 ‘The [aforementioned] field hut was full of rice!’

The particle *ni* adds emphasis to an assertion. It differs from the interrogative particle of the same segmental form in being pronounced with a mid tone. This particle is often encountered in public proclamations and songs.

- (4.100) *auu? ni.*
 au-əɪ-ù? ni
 be.good-PRES-DEC PTCL
 ‘That’s really good!’

The particle *əj* serves to give warning.

- (4.101) *nəŋ la ni nə səjəj.*
 nəŋ la ni nə sa-ɪ əj
 2SG TOP 1SG AGT say-IRR PTCL
 ‘I’ll report you!’

A negative particle *nuŋ* can be used as an interjection to reply to a proposition in the negative, as noted in §4.2.12. It is also used to negate a clause that cannot be negated by the verbal negative prefix because the clause is verbless, or because its predicate is a sequential converb (a verb stem inflected with the sequential converb suffix cannot be directly negated with the negative prefix – see §11.4.2).

- (4.102) *akhu tʃu nùkhu kət-pàkukà? məsən zaɪ nuŋlá, hɪətsə? nə hɪəkə səɪ.*
 a-khu tʃu nùkhu kət-pàkukà?
 NRL-tiger DIST wound possess-CONCESS
 məsən za-əɪ nuŋ la
 maggot enter-SEQ NEG.PTCL TOP
 hɪətsə? nə hɪəkə sə-əɪ
 insect.sp INST perforate-SIM die-PRES
 ‘Though the tiger gets a wound, it is not by maggots entering [the wound that it dies]; it is by insects boring [into its flesh that it] dies.’

The particle *ùku* encodes exasperation. In the naturally spoken sentence of (4.103), a pause occurs between the particle and the proximate nominal demonstrative. The particle is thus still understood to have a clause-final distribution, with the demonstrative postposed as an afterthought.

- (4.103) *nì alaktfukja.ùku, pi.*
 nì alak-tfuk-ja-à.ì-ù? ùku pi
 1SG forget-PFV-CONT-PRES-DEC PTCL PROX
 ‘I’m getting really forgetful of this.’
- (4.104) *nàḡ tənìk mə.ɪəm.à.ìku.*
 nàḡ tə-nìk mə.ɪəm-à.ì ùku
 2SG RL-eye be.red-PRES PROX
 ‘Aah, you’re just jealous!’

A separate sub-class of particles is reserved for use in imperatives and other types of commands, with one exception that is additionally used in other clause types. These are described in §10.4.1.

4.2.14.6. *Optional declarative mood clitic*

The optional declarative mood marker *-ù?* occurs as a clause-final clitic. It demonstrates low selectivity with respect to its host and generally attaches to the last constituent of the clause that is not an illocutionary force particle.¹³ The host may constitute a clause-final interjection (4.105), a finite verb (4.106), a discourse focus particle (4.107), or a clause-final noun phrase constituent, such as a numeral (4.108), a derived adjective inflected for degree (4.109), or a noun (4.110). If the final linear segment of the host is a vowel, then it forms a diphthong with the declarative mood clitic across the morpheme boundary.

- (4.105) *saləmù?*
 saləm-ù?
 salām-DEC
 ‘Greetings!’
- (4.106) *nì m̀.à.ì.à.tʃhət.làw?*
 nì m̀-ì.à-tʃhət-la-ù?
 1SG NEG-come-ABIL-NEG.PST-DEC
 ‘I wasn’t able to come.’
- (4.107) *t̀.à.pà.nə thuku táḡ.ú?*
 t̀-à.pà.nə thuku táḡ-ù?
 thus-CAUSAL nine only-DEC
 ‘That’s why/because of that there are only nine.’

13. The only observed exception to this distribution occurs in Text 2, line 79.

- (4.108) *nì akəm thakəm nə ɹuk.ɹaw?*
 nì a-kəm thakəm nə ɹuk.ɹa-ù?
 1SG NRL-year now+year ALL sixty-DEC
 ‘My age, to this year, is sixty.’
- (4.109) *a, nàŋ ta.ɹúpàw? ní.*
 a nàŋ tə-a.ɹu-’pàʔ-ù? ní
 EXCLM 2SG NZP-be.good-SUP-NR-DEC PTCL
 ‘Ah! You’re the best.’
- (4.110) *atʃu təm.ɹakù?*
 a-tʃu tə-mə.ɹak-ù?
 NRL-DIST NZP-lie-DEC
 ‘That’s a lie.’

The declarative marker optionally occurs as an enclitic to the reportative evidential particle, in addition to occurring as an enclitic to the stems of main verbs in independent clauses. This is demonstrated in the complex sentence of (4.111). Here it serves to affirm not only the facts of both the direct and indirect speech clauses, but also marks the reportative particle that confirms the source of the information as hearsay. The hierarchical structure can be schematically represented as a series of asserted layers in a complex sentence, with each layer in turn marked by the declarative clitic:

[[[direct speech layer -ùʔ] indirect speech layer -ùʔ] evidential layer -ùʔ]

- (4.111) [[[“*nì la, kəni nə ámíkhót tʃà.ùʔ*”] *tə saw?*] *tə.ùʔ atì.*]
 nì la kə-ni nə a-miʔ-khót tʃàʔ-àɹ-ùʔ
 1SG TOP 1SG.POSS-wife AGT NRL-person-hand consume-DEC
 tə sa-ùʔ təɹ-ùʔ atì
 thus say.PST-DEC REP-DEC PTCL
 ‘“Alas! My wife eats human hands”, [he evidently] said, like that’.

The underlying low tone of the optional declarative mood clitic has variable tonal realizations that sometimes correlate to the tone of an adjacent syllable, suggestive of a tonal alternation. However, its clause-final distribution also renders it a target for non-final intonation in discourse (see §3.4.3). This can result in significantly different tonal outputs that are completely independent of the lexical tone sandhi system. Evidence of this is demonstrated by the high tone realization on the optional declarative mood marker of (4.107). The high tone output in this text example can be explained by the fact that the speaker

had not reached the end of his discourse paragraph, so he used the non-final intonation to signal that more information relevant to that topic was to follow. Further proof of the effect of postlexical intonation on the output tone of word-final syllables comes from the comments of consultants, who at first thought that the high tone sounded incorrect when the sentence was played to them in isolation, but subsequently found it acceptable when it was heard in context.

Being a marker of declarative clauses, the optional declarative clitic naturally never occurs in imperative or interrogative clauses (see §5.1), or in converb clauses which, being non-finite, lack the means for the expression of independent mood (see §11.1–§11.4 for description). The diachronic source of this morpheme is not determinable at present.

Chapter 5

Clause structure and grammatical functions

The purpose of this chapter is to demonstrate simple clause structure and to show how grammatical functions are encoded in Mongsen. Basic clause types and criteria for identifying declarative, interrogative and imperative mood clauses are first discussed, after which an analysis of case marking is presented. This is significant for revealing a typologically unusual alignment pattern in which semantic and pragmatic factors play a key role in determining the presence or absence of core case marking. The chapter concludes with a description of valency modifying derivations.

5.1. Simple clause structure

Mongsen clauses may be verbal (5.1–5.3) or verbless (5.4). A verbal clause typically consists of a verb-final predicate¹ and one or two core noun phrases, as determined by the argument structure of the predicate, plus optional clausal adjuncts.

- (5.1) *aji nə tsəŋə̀nù?*
a-ji nə tsəŋ-ə̀-ù?
NRL-dog AGT bark-PRES-DEC
'The dog is barking.'

- (5.2) *tatʃə̀njùpà? nə áwk à hə̀nsət.*
tatʃə̀njù-pà? nə a-úk à hə̀nsət
PN-M AGT NRL-pig one kill.PST
'Tajenyuba killed a pig.'

- (5.3) *tʃuŋukù.*
tʃuŋ-ukù
eat.meal-ANT
'[I've] eaten.'

1. The use of the term 'predicate' reflects current linguistic and not traditional usage, referring to just the head of the clause and its modifiers; the predicate should therefore be understood to exclude all other clausal constituents.

Noun phrases can be freely omitted if their referents are pragmatically recoverable. Thus, the only obligatory constituent of an intransitive or transitive verbal clause is the predicate. This is demonstrated by (5.3), which is the usual way to reply to the question-cum-greeting “Have [you] eaten?” if one does not wish to burden the enquirer with an obligation to provide food.

Noun phrases function as either core or oblique arguments of verbal clauses. We look at how types of clausal arguments are distinguished in §5.3. Noun phrases can also function as independent topics of verbal clauses, but are more commonly found to be coreferential with an understood core argument that has undergone ellipsis. A verbless clause consists of a verbless clause topic and a verbless clause complement. The verbless clause topic can be any nominal, while the verbless clause complement can be any member of the open noun class identified in Table 4.1, a derived nominal, or a personal or interrogative pronoun. Verbless clauses are used to express types of ascriptive and equational meanings. A full account of their functions can be found in §9.1.

- (5.4) *kə təniŋ pəntáŋmuŋla.*
 kə tə-niŋ pəntáŋmuŋ-la
 1SG.POSS RL-name PN-F
 ‘My name is Bendangmongla.’

Clauses can be divided into independent and dependent types. An independent clause is an autonomous syntactic unit whose existence does not rely on other syntactic units, whereas a dependent clause always forms a constituent of another clause. There can be recursive embedding of dependent clauses within dependent clauses, but all embedded clauses ultimately form lower level constituents of their independent superordinate clauses. The types of dependent clauses found in Mongsén are relative clauses (§6.6.1 to §6.6.4, and §6.6.6), nominalized clauses (§6.6.7, §7.4.4, §7.4.7 and §7.4.8), converb clauses (§11.1 to §11.4) and complement clauses (§11.7). Relative clauses are attributes of noun phrases. Nominalized clauses function as embedded arguments of their matrix clause predicates. Converb clauses are a type of clausal adjunct standing in a dependent relationship to a matrix clause, and complements of verbal clauses function as embedded O arguments of their matrix predicates. The morphosyntactic structure of all these types of dependent clauses reflects that of independent clauses, in that all but purposive nominalizations can share the same case marking possibilities for encoding core and oblique argument functions.

We will now consider criteria that can be used for distinguishing the three types of independent clauses: declarative clauses, interrogative clauses and imperative clauses.

5.1.1. Declarative clause

Independent verbal declarative clauses must be specified for absolute tense by one of the absolute tense markers described in §8.5.12. This isolates finite independent declarative clauses from non-finite purposive clauses, relativized or nominalized clauses, and dependent converb clauses that are only specified for relative tense with respect to the absolute tense of their matrix clause predicates. There is a partial exception to this that requires clarification – dependent relative clauses and nominalizations can be formally specified for irrealis mood, e.g. *a-tʃak mələn-ì-pàʔ tʃu* (NRL-paddy transfer-IRR-NR) ‘the paddy that will be transferred’.² In view of this, it is necessary to add the proviso that a declarative clause with irrealis marking on its verb stem cannot be a non-finite relativization or a nominalization.

Secondly, the declarative mood can be formally distinguished from other mood types by the clause-final presence of the optional declarative mood clitic *-ùʔ* (§4.2.14.6). This marker is not obligatory in declarative clauses and speakers vary to the extent with which they use it to indicate their assertions. Nevertheless, the possibility of marking a clause with this morpheme proves to be the definitive criterion for differentiating declarative clauses from other clause types, since it cannot be used to mark the clause-final constituent of an interrogative clause, an imperative clause or a non-finite clause, and it is not affected by polarity.

The following examples respectively demonstrate the optional use of the declarative mood clitic in a verbal clause with negative polarity (5.5), a verbal clause with positive polarity (5.6), a verbless clause (5.7), and lastly, marking an exclamation ([4.104], repeated and renumbered below as [5.8]).

- (5.5) *pa nə tʃáʔ mətʃhəniŋlawʔ.*
 pa nə tʃáʔ mə-tʃhəniŋ-la-ùʔ
 3SG AGT nothing NEG-work-NEG-DEC
 ‘She didn’t do any work.’

- (5.6) *tə hjutsə à azənti nə sawʔ.*
 tə hjutsə à azənti nə sa-ùʔ
 thus story one old.person AGT say.PAST-DEC
 ‘Thus the old people told a story.’

2. Irrealis mood marking occurs in paradigmatic opposition with absolute tense marking. The historical processes by which an absolute tense system grammaticalized and came to incorporate a prior system of mood marking is discussed in §8.5.12.

- (5.7) *atʃu m̀ŋsən.lə hjuʦə tətʂə àw?*
 a-tʃu m̀ŋsən-ə.l hjuʦə tətʂə à-ù?
 NRL-DIST Mongsen-ANOM story NZP-be.short one-DEC
 ‘That’s a Mongsen person’s short story.’
- (5.8) *saləm-ù?*
 salām-DEC
 ‘Greetings.’

5.1.2. Interrogative clauses

An interrogative clause is a type of independent clause that is used for asking questions. Three sub-types are identifiable, according to how the interrogative is morphologically encoded: these variously take the forms of disjunctive, polar and content questions. There is no particular intonation pattern noted for interrogatives that prosodically differentiates the interrogative clause type from other types of independent clauses.

5.1.2.1. *Disjunctive questions*

Disjunctive questions are based on an ‘A-not-A’ template. The first verb of the construction has positive polarity and the second verb has negative polarity. An interrogative particle that is also used in polar questions (see §5.1.2.2 below) intervenes to function as the disjunctive operator. A disjunctive question is answered in the affirmative by repeating the first verb, or in the negative by repeating the second verb. This form of interrogative is widely attested in Sino-Tibetan languages.

- (5.9) *li sè m̀lila.*
 li sè m̀li-la
 be QPTCL NEG-be-NEG.PST
 ‘Was [he] there or not?’

5.1.2.2. *Polar questions*

Polar questions are marked by one of four illocutionary force particles dedicated to marking interrogative clauses (see §4.2.14.5 for a description of other members of this word class). The forms of the interrogative particles are

tʃhaúʔ, *sə*, *pà* and *ní*. Consultants allow these to be interchanged in the same sentence without discerning a change in meaning.

The interrogative particle of polar questions occurs at the end of the interrogative clause. Polar questions are typically answered with a one-word interjection, e.g. *hàwʔ* ‘yes’ or *hàjʔ* ‘okay’, or by repeating the verb of the interrogative clause.

(5.10) *hɲatʃhətətɪ tʃhaúʔ.*

hɲa-tʃhət-ətɪ tʃhaúʔ
listen-ABIL-PRES QPTCL
‘Do [you] understand?’

(5.11) *pi mətətətɪ sə.*

pi mətət-ətɪ sə
PROX know-PRES QPTCL
‘Do [you] know this [place]?’

(5.12) *anu nə kàʔ ípáʔ mətəm i atshəkhəməkə ɲasaŋtəməkə sanaɪ pà.*

a-nu nə kàʔ ípáʔ mətəm i
NRL-child AGT also EMPHAT manner PROX
atshəkhəm-əkə ɲasaŋtəm-əkə sana-ətɪ pà
boast-SIM be.arrogant-SIM speak-PRES QPTCL
‘So the child also speaks boastfully and arrogantly in this manner?’

(5.13) *“nì nə ajimù nì,” tət sa.*

nì nə ajim-ù nì tət sa
1SG AGT cry.out-IMM QPTCL thus say.PST
‘“Will I make a proclamation?”, said [Yellow-backed Sunbird].’

Mongsen lacks an interrogative clause type that corresponds to a tag question, such as those signalled by clause-final *isn't it?* or *don't they?* in English. A tag question encodes a presupposition on the part of the speaker that the content of their question is true. This type of interrogative therefore seeks to confirm the veracity of the speaker's presupposition. In Mongsen, the clause-final illocutionary force particles *jáʔ* and *ná* are available to convey such a presupposition. But since these particles are able to occur in a clause with a constituent marked by the declarative mood clitic, their clause type must be declarative (see §4.2.14.5 for examples).

5.1.2.3. *Content questions*

Content questions are formed with one of the ten interrogative pronouns listed in Table 4.6. Only those expressing the categories of PERSON/ENTITY, PLACE, THING or QUANTITY can function as the heads of their own noun phrases in interrogative clauses.

- (5.14) “*sé pá? nə ajimì,*” *tə səmt sə.*
sé pá? nə ajim-ì t ə səmt sə
 who AGT cry.out-IRR thus ask.PST
 ‘“Who will make a proclamation?”, asked [the leader].’

- (5.15) *túŋkhəla t[é pá? nə wa.ə.*
túŋkhəla t[é pá? nə wa-ə.ɪ
 3PL what ALL go-PRES
 ‘Where are they going?’

5.1.3. *Imperative mood categories*

Imperative mood consists of the positive imperative, the prohibitive and the admonitive categories. Types of imperatives and other commands are the topic of investigation in Chapter 10. Imperative mood forms a discrete clause type on the basis of morphological criteria that are presented in §10.1 and summarized in Table 10.1. Briefly, the salient criteria are that imperative mood clauses cannot be specified for absolute tense and their matrix verbs cannot optionally occur with the declarative mood marker.

5.2. Grammatical functions

This section describes how core grammatical functions are encoded by the grammar of Mongsén. I will begin by noting a useful distinction that Andrews (1985) draws between ‘grammatical functions’ and ‘grammatical relations’. A grammatical function is stated to be any ‘relationship that it might be useful to recognize which is definable over the sentence structures of a language under study, regardless of the extent to which it is important to the grammatical principles of that language’. A grammatical relation, however, is differentiated as one that is ‘generally significant for the workings of the grammatical principles of that language’ (Andrews 1985: 65–66).

If a language consistently targets a specific subset of arguments of the predicate for syntactic operations, such as equi noun phrase deletion, passiv-

ization, ellipsis of infinitival complements or ellipsis under subordination, then the controller of those syntactic operations might be identified as a grammatical relation, such as “subject” or “object”. Whereas grammatical functions are found in all languages, grammatical relations such as “subject” and “object” are language-specific and may not be relevant to the syntax of every language.

To facilitate this discussion, some useful neutral labels are needed to identify the core arguments of the clause, since there is no justification as yet for using the terms “subject” and “object”, in view of the introductory comments. The sentences of (5.1) and (5.2) in §5.1 above demonstrate simple intransitive and transitive clauses respectively, the intransitive clause having one core noun phrase argument, and the transitive clause having two. Following Andrews (1985: 68), the referent of the noun phrase *tatʃənjùpà?* ‘Tajenyuba’ is identified as the participant that controls or causes the activity denoted by the predicate *hənsət* ‘killed’, and therefore expresses the semantic role of agent. The referent of the noun phrase *áwk à* ‘a pig’ is affected by the activity denoted by the predicate and is accordingly identified as expressing the semantic role of patient.

Andrews calls ‘kill’ and other highly transitive two-argument verbs “primary transitive verbs”, because all languages need to have some way of differentiating between semantic roles to determine ‘who does what to whom’, and such verbs usually provide the model for the morphosyntactic treatment of core arguments for all other semantic verb classes. Andrews applies the label A to the noun phrase taking the semantic role of agent of the transitive clause, and O to the noun phrase taking the semantic role of patient. The single argument of the intransitive clause of (5.1) is differentiated from the A and O of a transitive clause by the label S.

Andrews (1985: 71) posits three possible coding strategies that languages use to signal grammatical functions. These are: order and arrangement (i.e. constituent order), case marking and agreement. Being a dependent-marking language at the clausal level (Nichols 1986), Mongsen does not use agreement to encode grammatical functions. The following discussion, then, is limited to a consideration of constituent order and case marking.

5.2.1. Constituent order

The default position for the verb is clause-final. This is a structural characteristic of all members of the Tibeto-Burman family with the exception of the Karen languages, which have undergone a change in constituent order under the areal influence of Mon-Khmer and Tai languages (Solnit 1997: xiv), and Bai, which has been influenced by Chinese (Wiersma 2003). If another constituent of the clause shunts the Mongsen verb out of its clause-final

position, then this will only be the result of self-repair or afterthought, and there is nearly always an accompanying intonation break that isolates the postposed constituent from the rest of the clause. The occurrence of the non-canonical constituent order without an accompanying intonation pause between the predicate and the postposed argument is otherwise extremely rare and can be attributed to the effect of rapid speech.

The following Waromung Mongsen sentence demonstrates a non-canonical constituent order in which the head of an O argument follows the predicate, while its nominal modifiers remain in the usual pre-head position. A comma in the italicized line of the example represents the intonation break that isolates the post-posed head from the rest of the clause. This atypical structure is obviously attributable to afterthought, which belatedly motivates the speaker to clarify the reference of the omitted head. The elided A argument of this transitive clause is recoverable from an antecedent mention in the source narrative.

(5.16) *tà táŋ təmáŋ tsə təpsətma?, tʃùŋlijimti niŋə.*

tà táŋ təmáŋ tsə təpsət-ma? tʃùŋlijimti niŋ-ə.ɪ
 thus just all DIST kill-CMPL.PST village.name name-ANOM
 ‘Thus, just like that [she] killed [them_i] all off, the Chungliyimti
 people_i.’

The A argument of a transitive clause may similarly be postposed to the clause-final position, also often as the result of afterthought, such as when a speaker desires to elucidate the roles of participants in the discourse. As with postposed O arguments or their constituents, there is an accompanying pause between the predicate and the clause-final argument.

(5.17) *tə ahizala taŋku sa, ahən təjila tʃu nə.*

tə a-hi?-za-la taŋ ku sa
 thus NRL-rat-DIM-F SIDE LOC say.PST
 a-hən tə-ji-la tʃu nə
 NRL-chicken RL-mother-F DIST AGT
 ‘Thus to Rat Pup [Ø_i] said [those words], Mother Hen_i.’

The order of noun phrases before a clause-final transitive verb is variable in texts and motivated by pragmatic factors. An AOV constituent order may be considered the default order. This is explained by the fact that the A argument is usually more topical, with its pragmatic importance iconically represented by its position of precedence in the clause. However, if the O argument outranks the A argument in pragmatic relevance, then it is not unusual to find an OAV constituent order in both dependent and independent clauses. This is

demonstrated in (5.18) below. Although it cannot strictly be equated with passive voice, the OAV constituent order has a similar effect in giving somewhat greater pragmatic prominence to the argument that undergoes the activity denoted by the predicate. Because of this, OAV constituent orders are most appropriately rendered by the passive voice in English translations.

- (5.18) *anù? mətshəla tʃu lítʃápá? sə nə tʃa.ə,*
 anù? mətshə-la tʃu lítʃá-pà? sə nə tʃa-əɪ
 again deer.sp.-F DIST PN-M ANAPH AGT call-SEQ
 ‘Again, after Barking Deer was called by Lichaba, ...’

The fronted O argument of a transitive clause requires no special grammatical marking, but it may occur with the topic particle. The topic particle is used to signal a shift in attention to a new referent, to reintroduce an old referent, or to contrast (see §4.2.14.3). Despite the presence of the topic particle marking *nuzaj* in (5.19) below, that noun phrase is still understood to be simultaneously functioning as a core argument of the predicate *tshəjù?* ‘will take’. The A argument of this clause has undergone ellipsis and the two overtly stated noun phrases both refer to the same referent, which is in O function. The topic particle in this sentence thus serves to clarify the referent, who is being differentiated from an older sister previously mentioned in the source narrative.

- (5.19) *nuzaj la tsəŋrutshəla tʃu tshəjù?*
 nuzaj la tsəŋrutshə-la tʃu tshə-ì-ù?
 younger.female TOP PN-F DIST take-IRR-DEC
 ‘The younger one, Tsengrutsela [I’ll] take.’ (said in reply to “Which daughter will you take to marry?”)

A transitive clause with an omitted core argument must be distinguished from structurally identical clauses formed with ambitransitive verbs, the single arguments of which depend upon neither context nor antecedent mention for the correct mapping of their grammatical functions (see §5.2.3 and §8.1.3). Zero anaphora is heavily exploited in Mongsen and it is common for arguments that are recoverable from the discourse to be omitted. Because of this, transitive clauses are more likely to have an elided core argument than to have both A and O arguments overtly represented by noun phrases. Frequent noun phrase ellipsis, the use of ambitransitive verbs and the pragmatically motivated ordering of clausal constituents thus render constituent order an unreliable aid to the determination of grammatical functions in this language.

5.2.2. Core grammatical marking

Case marking via phrase-final clitics is the only formal coding strategy that Mongsen uses for marking grammatical functions. Just one case marking clitic is available for canonical core grammatical marking, viz. the agentive *nə*.³ There is no obligation to use the agentive marker to distinguish the A argument of a transitive clause if the behaviour of its referent falls within social or expected norms, or if the semantic roles of core arguments are clear from the contextual setting or other cues. If this is not the case, however, then agentive marking can be used to disambiguate an atypical A argument from an O argument. Agentive marking can also variably appear on the single core arguments of intransitive clauses under a range of semantically and/or pragmatically motivated circumstances. First we will review the patterns of core argument case marking attested in the languages of the world, since this will provide a useful foundation for the subsequent description of grammatical functions in Mongsen and all the situations which determine case marking or its absence on a core argument.

Languages can be classified according to how they align grammatical functions (e.g. cf. the description of syntactically and semantically motivated marking patterns in Dixon 1979, 1994). Nominative-accusative languages give the same morphosyntactic treatment to the A argument of a transitive clause and the S argument of an intransitive clause, different from the O argument of a transitive clause. Ergative-absolutive languages give the same morphosyntactic treatment to the O argument of a transitive clause and the S argument of an intransitive clause, different from the A argument of a transitive clause.

A third alignment pattern described by Dixon is the split-S pattern, in which the S arguments of one set of intransitive verbs get a morphosyntactic treatment corresponding to the A argument of a transitive clause, while the S arguments of another set of intransitive verbs gets a morphosyntactic treatment corresponding to the O argument of a transitive clause. These arguments are respectively labelled S_a and S_o by Dixon (1994: 72), who states that the class membership of such sub-sets of intransitive verbs is fixed, generally in accordance with their prototypical meanings.

A fourth arrangement is what Dixon (1994: 78) refers to as a fluid-S system. This presents a pattern that resembles a split-S system, except that class membership is not fixed. The core noun phrase arguments of intransitive verbs in fluid-S languages take grammatical marking that is semantically motivated. If the referent controls the activity, its noun phrase takes S_a marking; if control is lacking, then the noun phrase of the referent takes S_o marking.

3. As noted in §4.2.7.5, the dative marker is used for non-canonical core grammatical marking in causativized clauses only – see §5.4.1 for description.

While the fluid-S pattern provides a possible model for the description of case marking, no language has yet been reported to be fully fluid with respect to the choice of marking that can occur on the sole arguments of all its intransitive verbs. Those documented demonstrate a mix of both fluid-S and split-S characteristics (e.g. cf. Holisky 1987, Mithun 1991). Fluid-S marking is most likely to be limited to intransitive verb roots denoting activities that may conceivably be performed by a volitional, controlling participant, or unintentionally effected by a non-controlling participant. The choice of case marking may additionally be influenced by cultural beliefs and practices, notions of marked versus unmarked social behaviour, the expression of personal choice, or the use of derivational and other morphology.

The Austronesian language Acehnese is an example of a language that demonstrates the characteristics of both fluid-S and split-S languages. Durie (1985: 62–67) lists approximately thirty Acehnese verb roots whose intransitive arguments take either S_a or S_o marking, according to the semantics of the particular context. The remaining intransitive verbs demonstrate a grammatical marking pattern characteristic of the split-S system, with S_a or S_o sub-set membership largely determined by a predicate's intrinsic meaning, and whether or not it denotes an activity or a state of being that is inherently under the control of its referent. Interestingly, the normally S_o verb 'die' permits S_a marking in the context of religious martyrdom, demonstrating how a cultural belief can influence the semantic basis of case marking. Durie states that many verb roots furthermore permit the semantic parameter of control to be manipulated by derivational morphology. This allows normally controlled activities to be encoded as uncontrolled, and vice versa for those verbs that denote activities or states that are normally uncontrolled. There is a small number of verb roots whose arguments appear to take counter-intuitive default grammatical case marking, e.g. the roots of 'sneeze' and 'vomit', two inherently uncontrolled activities, belong to the controlled S_a sub-set. Mismatches between the inherent semantics of an intransitive verb and the type of case marking applying to its noun phrase argument appear to be common in many languages with split-S case marking systems and are discussed at length in Mithun (1991).

A Tibeto-Burman language of Manipur known as Meithei or Manipuri (Chelliah 1997, Bhat and Ningomba 1997) is reported to have a typologically unique system of case marking that is distinct from the nominative-accusative, ergative-absolutive, split-S and fluid-S marking patterns described above. Citing Bhat (1988, 1991) and a manuscript based on elicited data that was later published as Bhat and Ningomba (1997), Dixon (1994: 28–29) analyzes Meithei as having a semantically based system of marking that 'directly describes the semantics of the conceptualization of a particular situation without this having to be related to a prototype and filtered through basic syntactic relations'. According to this interpretation, semantic roles are directly marked on noun

phrases: *-nə* for controlling agents, *-bu* for affected animates, and *-də* for indirectly involved beneficiaries, locations, experiencers, goals and patients.

Pace Dixon, in a text-based analysis of case marking in Meithei, Chelliah (1997: 124) points out that '[i]t is also possible to find sentences in which the subject does exercise control over an activity but is not marked by *-nə*'. She claims that other descriptions of Meithei are unduly influenced by elicitation and the traditional grammars of Latin and Sanskrit, which expect a predictable case paradigm. Furthermore, she concludes that pragmatics can delete and/or replace the semantic role markers, with the result that discourse context must be relied upon to determine grammatical relations when more than one interpretation is possible (Chelliah 1997: 129).

Mongsen is a language that demonstrates a system of case marking that is not wholly consistent with any of the patterns of syntactically or semantically motivated grammatical marking outlined above, although it bears a number of similarities that we are now in a position to compare. Consider the following sentences:

(5.20) *nì asəŋ səɬə̀.*
 nì a-səŋ sə-ə̀ɬ
 1SG NRL-wood chop-PRES
 'I'm chopping wood.'

(5.21) *nì nə asəŋ səɬə̀.*
 nì nə a-səŋ sə-ə̀ɬ
 1SG AGT NRL-wood chop-PRES
 'I chop wood.' (i.e. habitually, as an occupation)

Sentence (5.20) was given in reply to the question "What are you doing?", asked when a consultant was chopping firewood. It is a transitive clause with two core arguments. The referent of the A argument *nì* 'I' controls the action denoted by the predicate and the action effects a physical change of state on the O argument *asəŋ* 'wood'. In terms of transitivity, its verb must therefore rank closely to a primary transitive verb, like *hənsət* 'killed' of (5.2), as discussed above in §5.2. Yet the A argument of 'killed' in that example takes agentive case marking, while here it receives no overt case marking.

Later, when discussing the reply with the speaker, I asked what the difference would be if he said the same thing, but with the agentive case marker on the A argument (cf. [5.21] above). This produced a habitual reading, implying that the activity is an occupation. The following unsolicited sentences were then offered as a way of explaining how the core grammatical marking system of Mongsen operates.

(5.22) *ahən atʃak tʃã.ù?*

a-hən a-tʃak tʃãʔ-əɪ-ù?
 NRL-chicken NRL-paddy consume-PRES-DEC
 ‘The chickens are eating paddy.’

(5.23) *ahən nə atʃak tʃã.ù?*

a-hən nə a-tʃak tʃãʔ-əɪ-ù?
 NRL-chicken AGT NRL-paddy consume-PRES-DEC
 ‘The chickens are eating paddy.’ (implying that they are stealing it)

Sentence (5.22) is an assertion that reports on a neutral situation in which the chickens were fed paddy and are eating it. Sentence (5.23), however, has an interesting implication. The agentive case marker on the A argument not only encodes that the chickens are the agents performing the activity of eating the paddy, but additionally insinuates that they are wilfully stealing it. The presence of the agentive case marker in this sentence is accordingly understood to add a nuance of increased agentivity on the part of its referent.

This was explored and confirmed by testing in a more natural setting. Paddy is often left spread out on mats to dry in the sun in villages, and wandering chickens are naturally attracted to the exposed grain if it is left unattended. The fieldworker consequently has ample opportunities to elicit natural responses from Mongsen speakers by drawing their attention to chickens pilfering their drying paddy. Given this particular set of pragmatic circumstances, the agentive case was consistently used to mark the A argument of *ahən* ‘chicken’ in speakers’ responses to the warning *tʃu! tʃépaʔ tʃhàɪ* ‘Look there! What’s happening?’.

Obviously it is the context that helps one to distinguish between the two possible interpretations, i.e. increased wilfulness/intentionality, as opposed to habitual activity (cf. how the addition of agentive marking results in the habitual interpretation of [5.21]). It is conceivable, for example, that (5.23) could be interpreted to mean that chickens habitually eat paddy, given a qualifying context, although this would be peculiar if uttered in the absence of a pragmatically licensed setting, whereas a warning that drying paddy is being eaten by a wandering chicken would not be. The meanings of such sentences would only be potentially ambiguous when the temporal setting is the present tense, which is also used for the expression of habitual activity (see §8.5.12.1 for additional examples). Furthermore, the O argument would need to be non-individuated to allow for the habitual interpretation to emerge.

In both situations presented by the transitive clauses of (5.22) and (5.23), the respective A arguments must be interpreted as controllers of the actions denoted by their predicates. This suggests that the agentive case marker is not specifically marking the controller of the action in (5.23), but that some other

parameter is motivating its inclusion in that sentence. Further evidence that a noun phrase denoting the controller of an activity is not necessarily marked by the agentive case marker – even for a primary transitive verb such as ‘kill’ – comes from a Waromung Mongsen text example in which both noun phrases of a transitive clause containing the verb *ənsət* ‘kill.PST’ are unmarked for case.

(5.24) *təl, pa ʉpʉ ənsət.*

tə-əl pa ʉpʉ ənsət
 thus-SEQ 3SG PROX kill.PST

‘And then, she killed this.’ [i.e. in context, the pig]

The absence of agentive marking on the A argument of the verb *ənsət* ‘killed’ in (5.24) can be explained with respect to the pragmatic redundancy of formally distinguishing an agent from a patient when it is clear from both the context and one’s knowledge of hierarchical relationships in the real world that humans kill domestic pigs, and not vice versa. When semantic roles can be unambiguously assigned to the referents of transitive clauses, it can be pragmatically superfluous to encode the A argument formally via agentive marking. But in a marked situation, such as if it were the case that a domestic pig killed a human, then a statement reporting on this atypical situation would demand the presence of the agentive case marker on the noun phrase that refers to the pig, so as to override any incorrect assumptions about semantic roles based on the presumed docile behaviour of domestic pigs towards humans.

To explore this more deeply, Text 2 was examined to determine the frequency of agentive case marking in dependent and independent transitive clauses. A count reveals that, of the thirty or so transitive clauses with an overtly mentioned noun phrase in A function, almost half receive no agentive case marking. It is significant that the O arguments in all except one of those transitive clauses represent inanimate entities. Arguably it is the semantic nature of those O arguments that renders agentive case marking on the A arguments of their clauses pragmatically redundant. The one exception occurs in a clause describing how a bolt of lightning strikes a bull. But since lightning only ever strikes other things, the absence of agentive marking on this referent’s noun phrase is hardly likely to result in any confusion concerning the mapping of grammatical functions.

Similar motivations for core argument case marking are reported in a number of highlands Papuan languages. This leads Foley (1986: 107–108) to surmise that the disambiguating function of ergative marking provided a model for the development of ergative systems in most highlands Papuan languages. For example, Dani optionally case marks the A argument of a transitive clause with the ergative case marker to disambiguate case relations; ergative case marking is otherwise largely absent if there is no possibility of ambiguous

interpretation. Hua is also reported to use optional ergative case marking for the same purpose (Haiman 1980: 361). And case marking in Folopa (also known as Podopa in the literature) is even more similar to the system of core case marking found in Mongsen because it uses ergative case not just to disambiguate the semantic roles of core arguments in both transitive and intransitive clauses, but also to signal that the referent of an ergative case-marked noun phrase is a volitional, self-motivated performer of an activity. In contrast, its absence can indicate that the referent is acting according to expected social obligations and not according to independently determined motivations (see Anderson and Wade 1988: 5–12). The distribution of the so-called ergative and absolutive case markers in Folopa and the circumstances under which they are used indicate that this pattern is not strictly in accordance with the definition of ergativity suggested by Dixon (1979, 1994), nor for that matter with his model of semantically-based “direct” marking (Dixon 1994: 24). It therefore represents a hitherto unrecognized pattern that is yet to be adequately incorporated into the typology of case marking systems.

Mongsen bears some noteworthy similarities to Folopa that are suggestive of a similar motivation for case marking in transitive clauses. The following anecdotal examples demonstrate how case marking can be used to encode socially marked or personally motivated behaviour, and that this is unrelated to any notion of control over an activity.

One day when working with consultants, I was asked what I had done with the tea that had been made for me, to which I gave the following reply:

- (5.25) *nì nə səŋa tʃəmukù.*
 nì nə səŋà tʃəm-ukù
 1SG AGT tea drink-ANT
 ‘I’ve drunk the tea.’

I was immediately corrected by my Mongsen teachers, who explained that it was inappropriate to use the agentive case marker in this particular situation. But in the context of my intentionally taking someone’s tea and drinking it, especially if it were done in full knowledge that the tea was not actually mine, then the victim of the theft would be justified in asking the question of (5.26) below, encoding the A argument with the agentive case marker in an attempt to force someone to take responsibility for the missing tea. And if I admitted to wilfully taking the tea, then my reply to this question would similarly require the agentive clitic case marking the A argument, as in (5.27).

(5.26) *səpá? nə səŋà tʃəmtʃuk.*

səpá? nə səŋà tʃəmtʃuk
 who AGT tea drink-PFV.PST

‘Who drank up [the] tea?’ (i.e. ‘Who will admit to wilfully drinking my tea?’)

(5.27) *nì nə səŋà tʃu tʃəmtʃukukù.*

nì nə səŋà tʃu tʃəmtʃuk-ukù
 1SG AGT tea DIST drink-PFV.ANT

‘I’ve drunk up the tea.’ (i.e. ‘I admit to wilfully drinking that tea’)

The question of (5.26) could alternatively be framed without agentive case marking on the A argument in a pragmatically neutral situation, in which case the nuance of accusation would not be conveyed. This suggests that the motivation for agentive case marking in transitive clauses is more complex than the referent merely having control over the activity denoted by the predicate. Rather, the presence of agentive case marking in these examples seems to denote wilful or socially marked behaviour on the part of the referent.

Core grammatical marking in Mongsen is not found to be conditioned by tense, aspect or mood distinctions, nor by any other formal grammatical categories or nominal hierarchies discussed in Dixon (1994). The difference in tense/aspect marking between the question of (5.26) and the response of (5.27) is due to the fact that the anterior suffix cannot mark the matrix verb of an interrogative clause. The anterior category is affected by a grammatical dependency that contrasts fewer tense/aspect oppositions in interrogative and negated declarative clauses than in positive declarative clauses (see §8.5.12.3). This has absolutely no impact on core case marking functions, however.

While the notion of control cannot be held responsible for the presence of agentive case marking in the situations presented by the transitive clauses of (5.23), (5.26) and (5.27), there are occasions in which it does appear to contribute a motivating factor, notably when it marks the sole argument of an intransitive clause. Agentive case marking can be used to indicate that a referent is acting volitionally in the performance of an activity that is normally uncontrolled. For example, coughing is a corporeal activity that is inherently spontaneous and therefore uncontrolled in the default situation. Consequently the intransitive argument of ‘cough’ is not overtly marked for case when a simple sentence containing this verb is elicited independently of a context.

(5.28) *nì akhət.*

nì akhət
 1SG cough.PST

‘I coughed.’

But if the simultaneous converb *asáʔ-kə* expressing ‘deliberately’⁴ is used to modify this predicate, then agentive case marking on the noun phrase of the referent becomes an obligatory inclusion.

- (5.29) *nì nə asakə akhət.*
 nì nə asáʔ-əkə akhət
 1SG AGT be.deliberate-SIM cough.PST
 ‘I deliberately coughed.’

Exploratory elicitation suggests that the use of agentive marking is also acceptable in the absence of a modifying converb clause. This carries the implication that the activity was done on purpose, such as in a context in which the referent attempts to get someone’s attention by purposely coughing. Mongsén therefore demonstrates the typical characteristics of a fluid-S case marking pattern.

- (5.30) *nì nə akhət.*
 nì nə akhət
 1SG AGT cough.PST
 ‘I coughed.’ (i.e. on purpose, to get your attention)

Just as the semantic parameter of control can be manipulated by derivational morphology in Acehnese (Durie 1985), so too can derivational morphology play a part in influencing core case marking in Mongsén. The intransitive verb root *sə* ‘die’, for example, denotes a prototypically uncontrolled event. Since it is uncontrolled, the noun phrase argument of ‘die’ accordingly takes no case marking in the default situation. However, the presence of the desiderative suffix on the verb stem in (4.79), repeated for convenience and renumbered below as (5.31), indicates that a personal choice on the part of the referent is also involved. This in turn motivates agentive case marking on the core noun phrase *inisala tʃu nə* (rabbit-F DIST AGT). The example is somewhat complicated by the presence of the topic-marked pronoun *pa la*, which helps to contrast Rabbit with other animals that have meekly gone to their deaths in the narrative (see §4.2.14.3 for the description of topic marking).

4. Given its phonological form, the verb root of this converb may be historically related to the verb *sa* ‘say’. Saxena (1988: 381) reports that in the Tibeto-Burman languages Newar, Meithei and Adi, ‘say’ is used to convey that an activity is performed intentionally or deliberately. Like Mongsén, these languages also appear to use converb clauses to convey the derived meaning. See §10.2.3 for description of a possibly related form used as an admonitive marker in imperative clauses.

(5.31) *tətʃhɑɪ inisaʔ – iŋlif wət ɹapit tət ʃapaʔ la sà:à:ə – tət inisala ʃu nə, pa la məsəmɪ.*

tət-ʃhà-əɪ inisaʔ iŋlif wət ɹapit tət ʃa-pàʔ la
thus-do-SEQ rabbit English word rabbit thus call-NR TOP

sà:à:ə tət inisaʔ-la ʃu nə pa la mə-sə-mì-àɪ
animal thus rabbit-F DIST AGT 3SG TOP NEG-die-DESID-PRES
'And then, *inisaʔ* – the animal which is called “rabbit” in English – that animal, Rabbit, she doesn't want to die.'

While Rabbit does not have control over the event of dying, she does independently assert a desire not to die. Rather than the presence or absence of control *per se*, here it seems that the agentive case marking is encoding the expression of an independently motivated choice.

The following text extract demonstrates how agentive case marking can be used to encode the notion of control and personal choice over another prototypically uncontrolled activity denoted by the intransitive verb root *ʃəp* 'cry'. In elicitation devoid of a context, native speakers never assign case marking to the single argument of this intransitive verb. Yet in the following conversation, every intransitive argument of *ʃəp* takes agentive case marking. This is wholly consistent with the context. In the narrative, a man comes to report a murder to the mother of the victim, but before he divulges the information, he forces the murdered man's mother to give her word that she will not cry upon hearing the news. The data is from the Waromung variety of Mongsen, hence the divergent morphological forms.

(5.32) a. *“nàŋ na ən na aju à sa#əɪ.”*

nàŋ na nì na a-ju à sa-ù əɪ
2SG ALL 1SG AGT NRL-word one say-IRR PTCL
“I'll say something to you.”

b. *“nàŋ na ʃəp-ù məʃəp-ù.”*

nàŋ na ʃəp-ù məʃəp-ù
2SG AGT cry-IRR NEG-cry-IRR
“Will you cry or not?”

c. *tàsaku “ən na məʃəp-ù – saŋ.”*

tà-sa-ku nì na məʃəp-ù sa-aŋ
thus-say-LOC.CV 1SG AGT NEG-cry-IRR say-IMP
'Upon [the man] saying that, [the mother said] “I won't cry – say [it]!”'

- d. “*nàŋ na sitak mətʃəpə pa.*”
 nəŋ na sitak mə-tʃəp-ə pa
 2SG AGT be.true NEG-cry-IRR QPTCL
 “You really won’t cry?”

In the intransitive clause of (5.32b), the man assumes that the mother may be able to exert control over her emotions. He formally encodes this expectation by marking the noun phrase with the agentive case marker in his question. In her reply of (5.32c), the mother confirms that she is able to exert control over her emotions, similarly by encoding the intransitive argument with the agentive case marker. Lastly, in the interrogative of (5.32d) the man again seeks confirmation that the mother will control her emotions and refrain from crying, once more marking the intransitive noun phrase argument with the agentive case marker. Consultants confirm that a hypothetical absence of agentive case marking on the intransitive argument in the mother’s reply of (5.32c) would infer a predicted inability to control her emotions. This exchange thus offers unambiguous textual evidence of the possibility of fluid-S marking on the core arguments of intransitive verbs in Mongsen.

Investigations based on elicitation reveal that fluid-S marking applies to any intransitive verb roots expressing states or activities that are not normally controlled by an animate (usually human or personified) referent, but could possibly be controlled in a marked situation, e.g. *aʃhi* ‘sneeze’, *hwàməsa* ‘yawn’, *məni* ‘laugh/smile’ and *tshək* ‘collide’ (if uncontrolled) or ‘bump’ (if controlled). Like Acehnese and other languages reported to have fluid-S marking, Mongsen has some semantically counter-intuitive exceptions. Manner locomotion verbs such as *tʃəli* ‘walk.PST’ and *səm* ‘run.PST’, deictic motion verbs such as *wa* ‘go.PST’, *za* ‘enter.PST’, *tʃhuwa* ‘emerge.PST’ and *lə* ‘come.PST’, and verbs of posture such as *juŋli* ‘stand.PST’ and *mən* ‘sit.PST’ might be assumed to express controlled activities, yet their intransitive arguments do not occur with agentive marking, except in very limited circumstances. For example, it was explained that the only situation in which the intransitive arguments of locomotion verbs like ‘walk’ or ‘run’ could take agentive case marking would be if one were to answer a question such as “Will you walk or run?”. This provides additional evidence that the parameter of wilfulness/intentionality noted to influence agentive case marking in transitive clauses also plays a role in intransitive clauses when the referent is acting assertively and making a conscious personal choice to perform an activity.

It would be misguided to assume that the core case marking system of Mongsen is motivated exclusively by the notion of control. The presence or absence of control is unrelated to the case marking of A arguments in all of the transitive clauses presented above, and it is unlikely to be responsible for the presence of agentive case marking in the intransitive clause of (5.31). With the

exception of the habitual wood-chopper of (5.21), the common denominator of all the referents of agentive marked noun phrases in these sentences is that they are independently motivated entities who express a personal choice to perform an activity, to control an emotion or to entertain a desire. This extends to and encompasses referents who make a personal choice to cough, not to cry, or to run or walk when given the option.

In the default situation of use, the vast majority of intransitive verbs assign no case marking to their single core arguments (i.e. they correspond to the S_o subset, according to Dixon's 1994 approach to classification). In fact, intransitive verb roots whose arguments take default agentive marking are extremely rare – the only examples encountered that are consistently given with agentive case marking in off-the-cuff elicitation are verbs of vocalization, such as *tsəŋ* 'bark', *asà* 'shout', *ajim* 'scream, cry out, announce' and *lanli* 'reply'. Yet even these variably occur with and without agentive marking in texts, suggesting that this semantic class of intransitive verb is also subject to pragmatic and/or semantic influences exerted by a particular contextual setting.

Mongsen thus demonstrates a mix of strategies for encoding grammatical functions through the use of agentive case marking, or its absence. In a pragmatically neutral situation where there is no requirement for encoding wilfulness or personally motivated choice, the behaviour of the referents is prototypical, and the A outranks the O in terms of animacy or some other hierarchically-ordered parameter, both arguments of a transitive clause may be formally unmarked for case. In this situation an interlocutor relies upon their world knowledge of the referents' relationship to one another and the type of activity involved to correctly assign semantic roles to the noun phrases of a transitive clause. But if the referent of the A argument has a semantic role that is atypical, such as one that violates the animacy hierarchy, or if there is increased assertiveness or socially marked behaviour exhibited by the referent of the A argument in the performance of the activity, then its noun phrase is assigned agentive case marking. This typologically unusual system of marking grammatical functions therefore appears to be at times pragmatically motivated, and at other times semantically motivated. Recognition of the influences these two factors exert goes some way towards explaining why agentive case marking can be so variable on the A arguments of transitive clauses in texts.⁵

5. This case-marking pattern may yet turn out to be not all that unusual in Tibeto-Burman languages. Defen Yu (personal communication, 2005) reports that semantic and pragmatic motivations for agentive case-marking in her native language Lisu, a genetically distant Tibeto-Burman language of the Yi branch, correlate closely with those documented here for Mongsen, and my initial investigations of the Chang language of Nagaland also suggest the possibility of semantically/pragmatically agentive marking in both transitive and intransitive clauses.

The use of agentive case marking on A arguments to encode habitual activity is unusual. Presumably the notion of independent volition is also responsible for this, but only weakly so; a habitual performer does not really compare to the resolutely volitional and wilfully acting referents of agentive-marked noun phrases in some of the contrasting examples of transitive clauses discussed above. At present it is difficult to see how the use of agentive marking to signal increased wilfulness or assertiveness correlates with the use of agentive case marking to encode habitual activity in transitive clauses.

The S arguments of some intransitive verbs demonstrate the greatest potential for the presence or absence of agentive marking. The possibility of fluid-S marking applies to those intransitive verb roots denoting activities and states that are normally uncontrolled, but could conceivably be controlled in certain situations by a human referent making an independently motivated choice. Intransitive verbs denoting many activities that are prototypically performed by a volitional referent paradoxically assign no case marking to their referents, with the sole exception of verbs of vocalization. Additionally, intransitive verbs of locomotion whose arguments usually take no case marking in the default situation can assign agentive marking to their arguments if there is a perceived need to express that the referent is acting assertively and expressing a personal choice.

Lastly, it has been demonstrated that derivational morphology and modifying converb clauses encoding the deliberate performance of an activity can also influence core grammatical marking of intransitive arguments if the referent is behaving in an assertive manner, or is exerting a degree of personal control over a usually uncontrolled activity.

5.2.3. Absence of pivot

The grammar of Mongsen does not target a specific grammatical function or semantic role for feeding syntactic operations such as passivization, or for deleting a noun phrase argument that is coreferential with another in multi-clausal sentences. It is therefore consistent with the observations of Foley and Van Valin (1984: 120) and Dixon (1994: 155) that languages demonstrating semantically motivated grammatical marking tend to be pivotless.

The standard test for a syntactic pivot seeks to identify which core noun phrase argument(s) can be deleted under identity with another clausal argument in coordinated clauses. To illustrate, in the sentence *Mary_i met Peter_j and \emptyset _i came home*, only the A argument of the first clause can be coreferential with the deleted S argument of the coordinated clause. This constraint consistently identifies an S/A syntactic pivot and provides evidence of a grammatical relation of 'subject' in English. Being a typical Tibeto-Burman clause chaining

language, Mongsen does not use coordination as a clause linking strategy; consequently the usual diagnostic test for the presence of a syntactic pivot cannot be applied. In lieu of coordinated clauses, complex sentences containing dependent converb clauses with elided noun phrases can be used to demonstrate the absence of a consistent syntactic pivot. The following examples, all selected from narrative texts, reveal that it is contextual pragmatics which determines the cross-clausal coreference of elided arguments, not syntax.

The sentences of (5.33a–b) occur consecutively in the text they are taken from. Both contain a dependent converb clause (henceforth identified by square brackets) and a matrix clause. As is generally characteristic of Mongsen discourse, each clause contains at least one elided noun phrase argument.

(5.33) a. *tə təpəti à hləpəɪ, “nəni məthùŋ təsi áŋ pi təsəlà.”*

tə [Ø_{Ai} tə-pəti à hləp-əɪ]
thus NZP-be.big one wrap-SEQ

nə-ni mə-thùŋ təsi áŋ pi tə-səlà
2SG.POSS-wife NEG-reach until just PROX PROH-untie
‘Thus, [Ø_i] having wrapped up a big [bundle], [Lichaba_i said to Mechatseng] “Don’t untie this before you reach your wife.”’⁶

b. *təɪ həmətəɪ zək.*

tə-əɪ [Ø_{Aj} Ø_{Ok} həmət-əɪ] Ø_{Ai} Ø_{Oj} zək
thus-SEQ hold-SEQ send.PST

‘And then, taking hold [of the bundle], [Lichaba] sent [him].’
(Lit. ‘And then, [Mechatseng_j] took hold [of the bundle_k and Lichaba_i] sent [him_j].)’)’

The complex sentence of (5.33b) in fact has all of its noun phrase arguments elided, leaving just its transitive predicate heads. All these omitted arguments are retrievable from antecedent mentions and the pragmatic context. The elided A and O arguments of both the dependent and matrix clause are represented by Ø_A and Ø_O respectively, according to their respective grammatical functions. These are indexed with lower case letters *i*, *j* and *k*, so as to help the reader track their coreference across clauses. Taking into account both the context provided by (5.33a) and the meaning of (5.33b), the elided A argument of the dependent clause in (5.33b) must be coreferential with the elided O argument of its matrix clause; i.e. there is A=O coreference across the clauses.

The complex sentence of (5.34) similarly consists of a dependent sequential converb clause and its matrix clause. Both clauses are transitive. The matrix

6. Mongsen lacks a lexical exponent for temporal ‘before’. The use of a dependent clause to express temporal antecedence is described in §11.4.5.

clause contains an O argument that is modified by a pre-head relative clause. The A argument of the converb clause is coreferential with the A argument of the matrix clause, which has been deleted under identity. This complex sentence therefore demonstrates A=A coreference. Even if the noun phrase *tujət nə* is alternatively analyzed as a constituent of the matrix clause – e.g. *tujət nə* [*tʃũŋkuli tʃu zaŋlu*] *mətʃatshəŋ nə tsəŋpà? khùmapaŋ tʃu tsətshəŋi?* – A=A coreference would still hold across the clauses.

(5.34) *tujət nə, tʃũŋkuli tʃu zaŋlu, mətʃatshəŋ nə tsəŋpà? khùmapaŋ tʃu tsətshəŋi?*

[*tujət nə_i tʃũŋkuli tʃu zaŋlu-əɪ*]
3DU AGT anthill.soil DIST make-SEQ

Ø_{Ai} mətʃatshəŋ nə tsəŋpà? khùma-paŋ tʃu
PN AGT spear-NR wound-mouth DIST

zə-tshəŋ-i?

spread-ATTACH-CAUS.PST

‘They two_i made a paste of ant-hill soil [and Ø_i] applied it to the wound that was inflicted by Mechatseng’s spearing.’

Lastly, example (5.35) presents a dependent converb clause with a causativized transitive predicate and a matrix clause with an intransitive predicate.

(5.35) *təɪ pa tʃu aki ku mũŋtʃaki, pa áhlù nə wa.*

tə-əɪ [Ø_{Ai} pa_{Oj} tʃu a-ki ku mũŋ-tʃak-i?-əɪ]
thus-SEQ 3SG DIST NRL-house LOC stay-RS-CAUS-SEQ

pa_{Si} a-hlú nə wa
3SG NRL-field ALL go.PST

‘And then, making him stay at the house, she went to the field.’

(Lit. [She_i] caused him_j to stay at the house [and] she_i went to the field)

In (5.35), the elided causer noun phrase functioning as the A argument of the dependent clause must be interpreted as being coreferential with the pronominal S argument of the matrix clause. The third person pronominal argument of the dependent clause is the causee, because the causer noun phrase of a causativized predicate is obligatorily marked by the agentive case marker, and here *pa_{Oj}* is unmarked for case (see §5.4 for the description of causative derivations). This complex sentence therefore presents A=S coreference of cross-clausal core arguments.

The examples of this section demonstrate that a specific core argument of the clause is not targeted as a syntactic pivot, because an A argument of a

dependent clause is variously coreferential with an O, A or S argument of a matrix clause. Contextual pragmatics are therefore held to be wholly responsible for determining the cross-clausal coreference of elided core arguments in complex sentences.

5.2.4. Absence of passive voice

Mongsen lacks a syntactic passive operation for suppressing or omitting an A argument in order to promote an O argument into a subject grammatical relation. Instead, it achieves an equivalent outcome in either of two ways – by deleting a pragmatically recoverable A argument of a transitive clause, or by using a transitive verb in a patientive ambitransitive function.

An ambitransitive verb is a transitive verb that is used with just one core argument. Cross-linguistically, two types can be recognized: a patientive ambitransitive verb has one core argument corresponding to the O of a transitive clause; and an agentive ambitransitive verb has one core argument corresponding to the A of a transitive clause (after Mithun 2000: 87). To illustrate, the verb *break* functions as a patientive ambitransitive verb in English. In the sentence *A window broke*, the core argument *window* is in S function but corresponds to the O argument when this verb is used as the head of a transitive clause, e.g. *Bertil broke a window*. Note that the preceding sentence with the single core argument is immediately comprehensible in the absence of a qualifying context. There is no marking on the verb to suggest a passive derivation, therefore this cannot be identified as a true passive construction. Yet it achieves much the same effect as a passive, so it might be viewed as being functionally equivalent. The syntax of ambitransitivity is demonstrated with examples and discussed in greater detail in §8.1.3.

A clause containing a transitive verb that does not have a corresponding ambitransitive function can still achieve the functional equivalent of a passive voice, simply through eliding its A argument. As noted in §5.2.1, however, the ellipsis of a core argument requires an antecedent mention of its referent and a context that will facilitate the correct mapping of the remaining argument's grammatical function, whereas a sentence formed with a patientive ambitransitive verb has neither requirement.

Two widely reported characteristics of Tibeto-Burman languages are the textually frequent use of zero anaphora, and substantial inventories of ambitransitive verbs. Given these typological features, it is possible on *a priori* grounds to account for the rarity (if not complete absence) of passive voice in the family. If a language permits the O argument of a transitive clause to attain topical prominence merely by suppressing the A argument, or by using a verb in its patientive ambitransitive role, then there is no functional motivation for that

language to grammaticalize a passive voice. Another purpose of passivization in a language with a syntactic pivot is to provide a means of rendering an undergoer argument in S function. In a language with accusative syntax, this subsequently permits the derived S argument to be coreferential with an elided S or A argument in a coordinated clause, e.g. *Jack_i hit Jill_j and Ø_i ran away* → *Jill_i was hit by Jack_j and Ø_i ran away*. But such clausal rearrangements assume no importance whatsoever in a clause chaining language like Mongsen that does not make use of paratactic coordination. Furthermore, we have just seen in §5.2.3 that in the absence of a grammaticalized syntactic pivot, pragmatic context is the sole determinant of coreferentiality across multi-clausal sentences.

5.2.5. Summary

The findings of §5.2.1 to §5.2.3 amply demonstrate that there is little evidence for recognizing a grammatical relation of “subject” or “object” in Mongsen. Grammatical relations cannot be determined on the basis of constituent order, because the order of A and O arguments *vis-à-vis* the predicate can be manipulated by contextual pragmatics. Core case marking is equally unreliable as an indicator of grammatical relations, since agentive case marking can variably occur on the “subject” of an intransitive clause or on the “subject” of a transitive clause, or else the core argument(s) of a clause may not be expressly marked for case at all. Nor is there a syntactic pivot that consistently selects a particular argument of the clause for cross-clausal coreference of deleted noun phrases. Coreference is instead shown to be determined by pragmatic factors alone. And lastly, Mongsen is demonstrated to lack a passivization strategy that targets a specific noun phrase argument of transitive clauses.

Additional tests for the presence of grammatical relations could be attempted, but the already overwhelming evidence against the existence of “subject” and “object” in Mongsen precludes the necessity of this.

5.3. Case marking morphology

The following sections describe the case marking morphemes of Mongsen, commencing with a consideration of arguments for their status as clitics or independent grammatical words.

A case marker in Mongsen is a type of analytic morpheme whose principal function is to signal the semantic and/or syntactic relationship of a noun phrase to its head at the clause level. It is typically juxtaposed to the rightmost element of the noun phrase it marks. This may be a noun, a personal pronoun, a

demonstrative pronoun, or any other allowable constituent of the noun phrase (see §6.1). The low selectivity with respect to the range of word classes they occur in constituency with suggests that the case markers are not word-bound, but phrasally-bound elements. This rules out the possibility of their being nominal suffixes, so it remains to be determined if Mongsen case marking morphemes should be analyzed as clitics, or as independent words constituting a closed word class.

The status of the elements in question is not easily resolved. For instance, a case marker cannot occur independently of a noun phrase that has undergone pragmatically motivated ellipsis. This, together with their phrasally-bound distribution, suggests that they are more clitic-like in their behaviour. On the other hand, with the exception of the agentive case marker discussed in §3.1.3.2, all case markers maintain their phonological independence from juxtaposed constituents. They also carry independent lexical tone, which might be equated with accent for the purposes of diagnostic cross-linguistic comparison. The case markers of Mongsen thus do not present a definitive set of diagnostic “tendencies” or “symptoms” suggested in the literature (e.g. Zwicky 1985, Klavans 1982, Klavans 1985), by which they might be differentiated from independent words. They are therefore best viewed as being located somewhere along a cline extending from independent words at one extreme, to clitics at the other, given that they share the features of both prototypes.

The tone spreading of the agentive marker described in §3.1.3.2, with its resulting de-linking of segments from the skeletal tier, does not occur with other pronominal elements or other case markers. It may represent the early stages of an evolving cliticization process in Mongsen, as the case markers are grammaticalized from a more word-like to a more prototypically clitic-like form. The fact that the process is at present confined to just the agentive case marker does not necessarily refute this assumption. Genetti (1993: 65) notes that research on the historical development of morphological structure indicates that paradigmatic elements may undergo change independently of each other, i.e. not automatically as a unified set. This propensity is already amply demonstrated in Mongsen by the grammaticalized adverbial subordinating functions of the locative and instrumental case markers when they are used as converb suffixes on verb stems (see §11.4).

Forms of the Mangmetong Mongsen case marking morphemes are listed in Table 4.12; for convenience this is repeated and renumbered below as Table 5.1. The case marking clitics can initially be divided into two major sub-sets. The core marking sub-set has just one member, viz. the agentive marker (see §5.3.1). Core arguments of both transitive and intransitive clauses may also be formally unmarked for case under the range of conditions described in §5.2.2. The other sub-set is subsumed by those morphemes that are used for marking oblique arguments of the clause. The oblique case markers, described under

§5.3.2, can again be divided into those that exclusively express the semantic relationship of a noun phrase to the rest of the clause, versus those that have a primary local case marking function. Unlike core arguments, oblique arguments must always be formally specified for a case relation.

Table 5.1. Mangmetong Mongsen case marking clitics

CORE		OBLIQUE			
		Non-local		Primary Local	
<i>nə</i>	Agentive	<i>nə</i>	Instrumental	<i>nə</i>	Allative
		<i>li</i>	Dative	<i>phinə</i>	Ablative
		<i>atəməkə</i>	Benefactive	<i>ku</i>	Locative
		<i>thən</i>	Comitative		

The dative case marking clitic qualifies as a marker of core arguments in limited circumstances. While it has a purely semantic function of marking a recipient in non-causative clauses, it also presents a non-canonical syntactic marking function in causative clauses, where it marks the causee argument.

(5.36) *kawnpu.ɿa nə kəli sanaj?*

kawnpu.ɿa nə kə li sana-iʔ
 village.headman AGT 1SG.POSS DAT speak-CAUS.PST
 ‘The village headman made me speak.’

For the present we will just be concerned with the canonical functions of the case markers. A discussion of the non-canonical case marking functions of the dative in causative constructions is taken up in §5.4.

The agentive, instrumental and allative markers have an identical morphological form in *nə*. Whereas agentive/ergative and instrumental isomorphism is a relatively common pattern within various branches of Tibeto-Burman (LaPolla 1995), as well as cross-linguistically, the conflation of an allative with agentive and instrumental case forms is an extremely rare syncretism in the languages of the world. The ablative marker is also partially represented by the same morpheme in *phinə*. Core agentive marking can be formally distinguished from isomorphic oblique instrumental and allative marking by the types of pronominal heads with which these homophonous case markers may occur in constituency – only a personal pronoun may function as the head of an agentive-marked core argument. Examples such as the proverb of (4.51), repeated below as (5.37), further demonstrate the clear semantic contrast in the functions of the three homophonous case markers. Each encodes a unique

semantic relationship of its referent to the predicate of the clause, viz. agent, instrument and goal.

(5.37) *aji nə tuɿ nə áthúʔfən nə anəpən wa mətəm.*

a-ji nə tuɿ nə athùʔ-fən nə a-nət-pən
 NRL-dog AGT GPN INST vomit-LNOM ALL NRL-two-ORD
 wa mətəm
 go like

‘Like a dog going back to its vomit by itself for a second time.’
 (= to eat one’s words; to reject something and then want it later)

Case marking shows substantial morphological variation across village varieties of Mongsen. The Mongsen variety spoken in the Chungli ward of Waromung village,⁷ for example, has a low central vowel *a* in place of schwa in all the case markers listed above in Table 5.1 with the exception of the ablative, which has a different form in *phəniŋ*, and the benefactive, represented in that variety by *atəma*.

The heavy polysyllabic structure of the benefactive suggests a relatively recent grammaticalization. Also, the stem-final *a* of the Waromung Mongsen form *atəma* and the stem-final *əkə* of the Mangmetong Mongsen form *atəməkə* suggest a verbal source, for the following reason. It is highly likely that *a* and *əkə* are fossilized formatives cognate with the simultaneous converb suffixes *-a* and *-(ə)kə* of respectively identical form in the Waromung and Mangmetong Mongsen varieties (see §11.4.1.2). This seems too much of a coincidence to be merely attributable to a chance correspondence.

The logical corollary of this is that the benefactive case marker has been reanalyzed from a simultaneous converb (*atəm-a* in Waromung Mongsen and *atəm-əkə* in Mangmetong Mongsen). A verb root *atəm* has not been encountered in the data, nor is it recognized synchronically by native speakers. However, Clark ([1911] 1990: 163) lists an entry *atem* in his Chungli dictionary, which he translates as ‘to reiterate a name ..., to eulogize ...’. Furthermore, he gives examples of *atema* with what could be construed to be a nascent benefactive sense, e.g. *Tsungrem atema* ‘in the Lord’s name’, and notes that ‘*atema* is sometimes used as an equivalent of *asoshi*, or *meyong*, meaning “for” in a broad sense’. The original meaning of this verb root seems to provide an entirely credible pathway for the grammaticalization of a benefactive case marker.

7. Waromung village used to be bi-dialectal but is now wholly Mongsen-speaking. The “Chungli ward” refers to that ward in which the Chungli dialect of Ao was once spoken (see Coupe 2003a: 49).

5.3.1. Core case marking

A core argument of a transitive clause in A function either takes agentive marking enclitic to its noun phrase, or receives no formal relational case marking. The latter option creates the possibility of both core arguments of a transitive clause being unmarked for case. In this situation, contextual knowledge contributes to the correct interpretation of ‘who does what to whom’. The single arguments of the vast majority of intransitive clauses also receive no overt case marking by default in pragmatically neutral situations,⁸ but some classes of intransitive verb can assign agentive marking to their arguments under certain circumstances. As we have seen in §5.2.2, core case marking is manipulated by semantic and/or pragmatic factors relating to the noun phrase referent with respect to that specific instance of use. This analysis must therefore assume that agentive case marking on core arguments, or for that matter its absence, is not assigned by the argument structure of the predicate, but is instead determined by non-syntactic factors.

5.3.1.1. Agentive

The agentive *nə* is most typically used to mark the A argument of a transitive clause in which the referent is an independently motivated, wilfully acting animate participant. In some contexts agentive marking serves a disambiguating function, such as when either referent of two core arguments of a transitive clause could conceivably function as the controlling agent of an activity. Semantically motivated agentive case marking occurs on the single arguments of intransitive clauses under the specific circumstances discussed in §5.2.2 (see [5.29–5.32] for examples), and on the A arguments of transitive verbs that denote habitual activities.

- (5.38) *mətʃatshəŋ nə pùŋì tʃu tsəŋ.*
 mətʃatshəŋ nə pùŋì tʃu tsəŋ
 PN AGT wild.pig DIST spear.PST
 ‘Mechatseng speared the wild pig.’

An optionally realized tone spreading process causes the low tone of the first person singular pronoun *nì* ‘I’ to coalesce with the mid tone of the agentive marker after the segmental material of the agentive marker is deleted. This results in a low rising tone that encodes agentive case in the absence of any

8. The only exception to this applies to some intransitive verbs of vocalization, as discussed in §5.2.2.

segmental representation. Tone coalescence only applies to the first person singular pronoun, and is only found in some varieties of Mongsen (see §3.1.3.2 for description).

5.3.1.2. *Formally unmarked core arguments*

The O argument of a transitive clause with an overtly agentive-marked A argument receives no formal case marking. In pragmatically-neutral situations, the single arguments of most intransitive verbs are also formally unmarked for case by default.

- (5.39) *ila nə tʃuŋi taɹu táŋ tʃãɹ.*
 ɪla nə tʃuŋi tə-aɹu táŋ tʃãʔ-əɹ
 1PLEX AGT food NFP-good only consume-HAB.PRES
 ‘We only eat good food.’

- (5.40) *təmhna:k sə tʃha:ɹu-líʔ ɹà.*
 tə-əmhna:k sə tʃha:ɹu-líʔ ɹà
 RL-male.in.law ANAPH be.sick-SIM come.PST
 ‘The son-in-law came in a state of ill health.’

We saw in §5.2.2 that a third possibility in transitive clauses is for both the A and O arguments to be formally unspecified for case relations, as in the sentence of (5.41) below (also cf. [5.20]).

- (5.41) *nì hjutsə à sa:ɹù.*
 nì hjutsə à sa:ɹù
 1SG story one say-IMM
 ‘I’ll tell a story.’

In this situation, one relies upon contextual pragmatics or other cues to correctly determine the mapping of grammatical functions to core arguments of the clause. A survey of narrative texts confirms that in common with (5.41), a significant contrast in animacy between unmarked A arguments and their unmarked O arguments is usually found in transitive clauses, thus there is little likelihood of ambiguity occurring through both core arguments being formally unmarked for case. In situations where ambiguity could arise, agentive marking is employed to clarify the semantic roles of core arguments.

5.3.2. Oblique case marking

The oblique case marking morphemes encode semantic and local case relations. Their forms are listed in Table 5.1 above. In §4.2 it was shown that possession can be encoded using either a possessive pronoun or a personal pronoun; so too can either a possessive pronoun or a personal pronoun fill the head slot of a noun phrase with oblique case marking for many speakers. Any prior division of labour that may have once existed between possessive pronouns and personal pronouns has since become blurred synchronically. This is attributable to historical processes of grammaticalization affecting oblique noun phrase constituents (see §5.3.3 for a full account of these processes). Now the only reliable distributional characteristic that formally differentiates possessive pronouns from personal pronouns is that just the personal pronoun forms can function as heads of noun phrases in A, S or O function. In contrast, either a personal pronoun or a possessive pronoun can be selected as the head of an oblique noun phrase. The only exception to this optionality in oblique marking is found with the dative case, which obligatorily demands a possessive form of pronoun as its head.

Overall it is more common for possessive forms of pronouns to be used as heads of noun phrases with the local case marking clitics, for the following historical reason. The pronominal head of an allative or locative case-marked noun phrase additionally requires a nascent postposition inserted between the head and its case marking clitic, e.g. *kə taŋ nə* (1SG SIDE ALL) ‘to me’. In §5.3.3 I argue that the nascent postposition *taŋ* was the erstwhile head of a possessive compound noun in oblique function and expressed the equivalent of ‘side’, a meaning still preserved in the noun *taŋi*. This would explain why the common choice of pronominal head for an oblique case marking clitic is a possessive form. At an earlier stage of the language, *kə taŋ nə* must have expressed something akin to ‘to my side’.

Because of the substantial syncretism between the two paradigms of pronouns, any distinction drawn between the personal pronouns and possessive pronouns with respect to what can occur as the head of a core or oblique noun phrase argument naturally only applies to the first and second person pronouns *kə* and *nə* of the singular series, and the first, second and third person forms of the plural series, i.e. *i* ~ *aj*, *niŋ* and *páɪ* (cf. the forms of Table 4.4 to those of Table 4.5 to gain an appreciation of the extent of this syncretism).

5.3.2.1. *Allative*

The allative case marker *nə* marks the spatial goal of movement. A spatial goal with a human or personified referent obligatorily occurs with a nascent postposition occurring between the head of the noun phrase and the allative marker (see §5.3.3 for additional examples). An intervening nascent postposition is not required for goals with non-human referents.

(5.42) ... *təpaʔ taŋ nə waɪ*,

tə-paʔ taŋ nə wa-əɪ
 RL-father SIDE ALL go-SEQ
 ‘... after going to the father ...’

(5.43) *nɪʔ áhŋáʔ təpə̀ti à asət tʃu nə za*.

nɪʔ a-hŋáʔ tə-pə̀ti à asət tʃu nə za
 one.day NRL-fish NZP-be.big one fishing.weir DIST ALL enter.PST
 ‘One day, a big fish entered the fishing weir.’

I previously mentioned that if the referent is pronominal, most speakers select a possessive form of pronoun as the head of a noun phrase marked by an oblique case. However, some (usually older) speakers differ in this regard by variously using both personal pronouns and possessive pronouns as heads of oblique noun phrases.

(5.44) *kətaŋ nə ɪaŋ*.

kə taŋ nə ɪà-aŋ
 1SG.POSS SIDE ALL come-IMP
 ‘Come to me.’

(5.45) *tə azənti nə təmənimi tatisi hjutsə à ɪla taŋku itə ɪi hjutsə sajawʔ wa*.

tə azənti nə tə-məni-mi tatisi hjutsə à
 thus old.people AGT NZP-laugh-DESID example story one
 ɪla taŋ ku i-tə ɪi hjutsə sa-jaʔ-ùʔ wa
 1PL.INC SIDE LOC PROX-thus kind story say-HAB.PST-DEC PTCL
 ‘Thus the old folk used to tell this humorous example of a story to us like this, as I remember.’

The referent of the goal can also be non-spatial. In (5.46), the allative clitic case-marks a quantitative goal that results from the activity of cutting up meat.

- (5.46) *tʃùku á-sáʔ tʃu thuku táŋ nə thùŋ.*
 tʃùku a-sáʔ tʃu thuku táŋ nə thùŋ
 consequently NRL-meat DIST nine only ALL reach.PST
 ‘Consequently, the meat only amounted to nine [pieces].’

5.3.2.2. Ablative

The ablative case marker *phinə* marks the source of movement. I noted above in §5.3 that the ablative is partially composed of the morpheme *nə*, which encodes the agentive, instrumental and allative cases. The element *phi* is not similarly found to occur as an independent morpheme synchronically.

- (5.47) *nuksənsaŋpaʔ áhlù phinə tʃhuwaɪ.*
 nuksənsaŋ-pàʔ a-hlù phinə tʃhuwa-əɪ
 PN-M NRL-field ABL emerge-PRES
 ‘Noksensangba returns from the field.’

A highly unusual feature of the ablative case marker is that it allows the nominal distal demonstrative *tʃu* or anaphoric demonstrative *sə* to occur between its two syllables. Both these “infixes” occur synchronically as independent phonological words (see §4.2.7.1), which would seem to suggest a degree of synthesis in Mongsen, a language that is otherwise overwhelmingly agglutinative in its morphological typology. Consultants suggest that an ablative case marker containing a distal nominal demonstrative points to the referent of its noun phrase, or encodes old or given information in the case of the anaphoric demonstrative. Their deictic functions as putative infixes are therefore indistinguishable from their respective deictic functions as independent phonological words (see 4.3.7.1 for description).

- (5.48) ... *lùŋkhəm imsən phisənə ...*
 lùŋkhəm jim-sən phi <sə> nə
 village.name village-new <ANAPH>ABL
 ‘... from the [aforementioned] New Lungkum village ...’

- (5.49) *tuku lítʃápáʔ ki phitʃunə məkhəp tsəŋta ku,*
 tə-ku lítʃá-pàʔ ki phi <tʃu> nə mə-khəp
 RL-mother’s.brother PN-M house <DIST>ABL NEG-depart
 tsəŋta ku
 BETWEEN LOC.CV
 ‘Before departing from the house of Uncle Lichaba, ...’

It seems highly unlikely that these nominal deictics could be operating as true infixes. More plausible is that *phi* originally functioned as a generic head in possessive compounds encoding a location, probably expressing something like ‘source’ at an earlier stage of the language (cf. the analogous position and function of the nascent postposition *taŋ* and its etymological relationship to the noun *taŋi* ‘side’, discussed in §5.3.3).

The infixed positions of the nominal demonstratives in the examples of (5.48) and (5.49) are otherwise difficult to account for, but not if we assume that *phi* originally functioned as the head of these oblique noun phrases, determined by a demonstrative and case marked by a morpheme *nə* that merely marked its argument as being in a semantically unspecified oblique function. As a result of the demotion and reanalysis of *phi* as a phrasal adjunct, the locational ‘source’ meaning previously associated with this compound head subsequently merged with the oblique marking function of *nə*, culminating in a distinct ablative meaning for a newly created case marker *phinə*.

The usual position of a nominal demonstrative in the noun phrase is to the left of the case marking clitic (cf. the noun phrase template of §6.1) and not inside it, as demonstrated by the following example.⁹

- (5.50) ... *tsəŋlijimti lima tʃu phinə* ...
 tsəŋlijimti lima tʃu phinə
 village.name country DIST ABL
 ‘... from the land of Chungliyimti ...’

5.3.2.3. *Instrumental*

The instrumental case marker *nə* marks the means by which an activity is carried out by an agent.

- (5.51) *təɪ məzəʔ nə ɹuŋukù tʃi*.
 tə-əɪ məzəʔ nə ɹuŋ-ukù tʃi
 thus-SEQ fire INST burn-ANT PTCL
 ‘And, [he] cleared [the field] with fire.’

Another function of the instrumental case marker is to mark a body part that is the origin of a physical sensation.

9. The elderly Mangmetong Mongsen speaker whose narrative provided this example had an idiolectal pronunciation for the village name of ‘Chungliyimti’.

(5.52) *nì təpuk nə tʃha.ù?*

nì tə-puk nə tʃha-əI-ù?
1SG RL-belly INST pain-PRES-DEC

‘My stomach is hurting.’ (Lit. ‘By my stomach I am experiencing pain’)

It is not possible to replace the personal pronoun *nì* in this type of construction with the corresponding possessive pronoun *kə* to mark possession, e.g. **kə-puk* (1POSS-belly) or **kə tə-puk* if the relational prefix is retained; consequently this cannot be analyzed as some type of a possessive compound (see §7.3.1), or a ‘double topic’ construction. This might be logically explained by the fact that it is the self or individual as a whole that experiences the unpleasant sensation, not just the body part *per se*. The pronoun *nì* must therefore function semantically as an experiencer in S function, while the noun phrase *təpuk nə* is in oblique function and expresses the means by which the discomfort is experienced.

There is no requirement for the speaker of such sentences to possess epistemic authority, since a third person pronoun or a noun can be used as the head of the noun phrase functioning as the experiencer.

(5.53) *asaŋla tələm nə maŋə.ù?*

asaŋ-la tə-ləm nə maŋ-əI-ù?
PN-F RL-head INST spin-PRES-DEC

‘Asangla has a headache.’

Both the instrumental and the agentive markers can be used to case mark separate noun phrases within the same clause. Instrumental case is not used to mark the noun phrases of animate referents, thus there is little chance of ambiguity arising through two noun phrases both being marked by *nə*. Animates that are semantically equivalent to instruments are instead encoded as causees of morphological causatives (see §5.4.1).

(5.54) *nì ahən təjila nə tsəməlu nə məkhəpù?*

nì a-hən tə-ji-la nə tsə-məlu nə
1SG NRL-chicken RL-mother-F AGT water-boil INST

mə-khəp-ù?

NEG-tip.over-DEC

[Rat Pup said] ‘Mother Hen scalded me by tipping boiling water over me.’

5.3.2.4. *Dative*

The main function of the dative case marker *li* is to mark the recipient argument of a verb of transference, the prototypical exemplar being ‘give’. It has a core case marking role in some causative constructions, a preliminary example of which was provided in the sentence of (5.36). The syntax of constructions involving the morphological causative is described in §5.4.1; the analytical causative is described in §5.4.2.

- (5.55) *tsəŋjutshəla li ípáʔ usət tʃu khiʔ-ukù.*
 tsəŋjutshə-la li ípáʔ usət tʃu khiʔ-ukù
 PN-F DAT EMPHAT stuff DIST give-ANT
 ‘That particular stuff is given to Tsengrutsela.’

It was briefly noted in §5.3.2 above that the dative obligatorily selects the possessive form of a pronoun when it marks a pronominal head functioning semantically as a recipient.

- (5.56) *kə li khiəŋ mə.*
 kə li khiʔ-aŋ mə
 1SG.POSS DAT give-IMP PTCL
 ‘Please give [it] to me.’

One young Mongsen speaker used the personal pronoun *nì* with the dative marker in an elicited causative sentence and claimed that it was interchangeable with the possessive first person form *kə*. This is a case of idiolectal morphological levelling and not *bona fide* free variation, because all other pronominal noun phrases marked by the dative case marker in the corpus have possessive forms. The aberrant example is included below for posterity’s sake. The speaker who uttered this sentence was a tertiary-educated young man aged 22 years who grew up in the town of Mokokchung. It could represent an incipient age-related difference in the use of pronouns with oblique arguments, but all the native speakers I questioned about its grammaticality found it unacceptable.

- (5.57) *asaŋla nə ahən tʃu { *nì ; kə } li məɹəkimiùʔ.*
 asaŋ-la nə a-hən tʃu { *nì ; kə } li
 PN-F AGT NRL-chicken DIST { 1SG ; 1SG.POSS } DAT
 məɹək-iʔ-mì-əɹ-ùʔ
 clean-CAUS-DESID-PRES-DEC
 ‘Asangla wants me to dress the chicken.’

5.3.2.5. Benefactive

The benefactive *atəməkə* is a rarely occurring case marker that encodes the semantic role of beneficiary or purpose. The Waromung variety of Mongsen has a variant form *atəma*, text examples of which are given in (5.58) and (5.60) below. In §5.3 I proposed that the benefactive case marker originates from a verb root marked by a converb suffix that has subsequently been reanalyzed as a case marking clitic. This is suggested by the *kə* and *a* formatives respectively seen in the variant Mangmetong and Waromung Mongsen benefactive forms *atəməkə* and *atəma*, which correspond with the simultaneous converb suffixes *-(ə)kə* and *-a* of those varieties (cf. §11.4.1.2).

Speakers of all ages use both personal pronouns and possessive pronouns as heads of noun phrases case marked by the benefactive marker. The choice of pronominal form does not entail any difference in meaning.

- (5.58) *tʃuphapà? na {nì ~ kə} atəma khiliri zək-ù.*
 tʃupha-pà? na {nì ~ kə} atəma kì-li-əi
 PN-M AGT 1SG ~ 1SG.POSS BEN house-stay-ANOM
 zək-ù
 send-IRR
 ‘Chubaba will send the servant for me.’

In addition to marking the benefactor of an activity denoted by the predicate, the benefactive case expresses a purposive meaning equivalent to ‘for the purpose of, being for’. This usage reflects one of the meanings suggested by the definitions of *atema* in the Chungli dialect, as provided by Clark ([1911] 1990: 163) and discussed above in §5.3. Lastly, the Waromung Mongsen example of (5.60) demonstrates that the benefactive marker is compatible with inanimate as well as animate referents of nominal heads.

- (5.59) *akhu təpətípá? təlu? .uŋku tʃuphà atəməkə,*
 a-khu tə-pəti-´-pà? təlu? .uŋ ku tʃuphà atəməkə
 NRL-tiger NZP-be.big-SUP-NR all AMONG LOC king BEN
 ‘The biggest tiger amongst them, by virtue of being the king, ...’

- (5.60) *ən na nəŋkhəla tətʰi təkəm tʃawŋ atəma ta? à khì?-ù.*
 nì na nəŋkhəla tətʰi təkəm tʃawŋ atəma
 1SG AGT 2PL every.day NZP-be.alive food BEN
 ta? à khì?-ù
 unit.of.measure one give-IRR
 ‘Every day I will give a measure of daily food to you all.’

5.3.2.6. *Locative*

The basic function of the locative case marker *ku* is to encode a spatial location.

- (5.61) *pa lisəʔ təmaŋ tʃu ku mələm mələməkə tiphàŋ.*
 pa lisəʔ tə-maŋ tʃu ku mələm-mələm-əkə tì-phàŋ
 3SG mud RL-body DIST LOC be.thick-RED-SIM daub-COVER.PST
 ‘She smeared mud on [her] body very thickly.’

Verbs of utterance denoting speech acts directed at a specific interlocutor require the noun phrase of that referent to be case marked by a combination of a nascent postposition *taŋ* or *thak* and the locative marker. The nascent postposition *taŋ* is the most common form occurring in the oblique noun phrases of these speech act verbs.

- (5.62) *ahən təjila taŋku təm̀ə̀rak sa.*
 a-hən tə-ji-la taŋ ku tə-m̀ə̀rak sa
 NRL-chicken RL-mother-F SIDE LOC NZP-lie say.PST
 ‘To Mother Hen [Rat Pup] told a lie.’

A nascent postposition *taŋ* co-occurring with the locative is also found in clauses with stative verbs denoting human emotion when the emotion is directed at another referent, rather than just experienced.

- (5.63) a. *nì tʃàsi.ùʔ.*
 nì tʃàsi-ə̀ɪ-ùʔ
 1SG be.distressed-PRES-DEC
 ‘I’m upset.’
- b. *nì nə taŋku tʃàsi.ùʔ.*
 nì nə taŋ ku tʃàsi-ə̀ɪ-ùʔ
 1SG 2SG.POSS SIDE LOC be.distressed-PRES-DEC
 ‘I’m upset at you.’

Activity verbs whose semantics imply physical contact with an animate referent – e.g. *jàk* ‘beat’, *màʔ* ‘slapped’, *mijəm* ‘stroked; patted’, *tsəŋ* ‘punched’ and *pugtshəŋ* ‘pounced upon’ – typically mark their oblique animate arguments with the *thak* form of nascent postposition, in addition to the locative case marker. The nascent postposition *thak* in the oblique noun phrases of these particular verbs seems to additionally specify the orientation or location in which the activity is performed.

The following two examples were spontaneously uttered in response to video clips. The position of the proximate demonstrative between the nascent postposition and the locative in (5.64) is particularly noteworthy for diachronic reasons initially considered in §5.3.2.2, and again discussed in §5.3.3 below. It provides further evidence that the nascent postpositions were once the heads of their noun phrases. In fact, *thak* ‘place’ of (5.65) must be interpreted as the head of its oblique noun phrase case marked by the locative *ku*, seeing that it is the only other constituent of that noun phrase.

(5.64) *apaŋtʃhaŋəɪ ləm̩tshəpà? nə luma təhlaŋpà? thak i ku tsəŋukù.*

apaŋtʃhaŋ-əɪ ləm̩-tshəpà? nə
man-ANOM head-pull-NR AGT

luma tə-hlaŋ-pà? thak i ku tsəŋ-ukù
hair NZP-be.long-NR PLACE PROX LOC punch-ANT

‘The bald man has punched this long-haired one.’

(5.65) *tun̩ət mətəpja.ù? à nə à nə thaku mətəpja.ù?*

tun̩ət mə?-təp-ja-əɪ-ù?
3DU slap-RECIP-CONT-PRES-DEC

à nə à nə thak ku mə?-təp-ja-əɪ-ù?
one AGT one AGT place LOC slap-RECIP-CONT-PRES-DEC

‘They [two] are slapping each other, slapping each other in turns, one-by-one.’

Lastly, the locative marker has grammaticalized as an adverbial subordinator that attaches to non-finite verb stems and expresses relative tense. Its function as a marker of dependent clauses is consistent with the grammatical category of *converb* (see §11.4.3.7 for discussion and additional examples).

(5.66) *aki tʃhàthùŋku,*

a-ki tʃhà-thùŋ-ku
NRL-house make-REACH-LOC.CV

‘When [he] was building his house, ...’

5.3.2.7. Comitative

The comitative case marker *thən* expresses the semantic relation of accompaniment with another referent, the noun phrase of that referent being in core function. It marks the noun phrases of both human and non-human referents. Native speakers of all ages variously use both personal and possessive

pronouns as heads of noun phrases case marked by the comitative marker when the referent is pronominal.

- (5.67) *pa nə thən kútə k̀̀tsəŋəɪ.*
 pa nə thən kútə k̀̀tsəŋ-əɪ
 3SG 2SG.POSS COM how relative-ANOM
 ‘How is s/he related to you?’

The comitative case marker is used to mark the standard of comparison in comparative constructions (see §7.4.3.1).

- (5.68) *ípáɬ la i jiptʃən thən la jiptʃən pi áŋ t́́lɛŋpaw?*
 ípáɬ la i jiptʃən thən la jiptʃən pi áŋ
 EMPHAT TOP PROX bed COM TOP bed prox just
 tə-hlɛŋ-pàʔ-ùʔ
 NZP-be.long-NR-DEC
 ‘This bed is longer than that one.’

The comitative marker is likely to have a verbal diachronic source, on the grounds that its morphological form is identical to the lexical suffix *-thən* (TOGETHER) described in §8.5.1.11. This assumption is further supported by its semantics being very close to the meaning of the lexical suffix.

5.3.3. Nascent postpositions

Eleven or so morphemes, at least half of which originate from body part terms of the relational noun class, function as nascent postpositions. These are used to specify the spatial location of an entity, or to indicate the deictic orientation in which an activity is performed. Forms of the nascent postpositions and the case marking clitics with which they co-occur are listed in Table 5.2.

The vast majority of these co-occur with the allative and/or locative case markers (see §5.3.2.1 and §5.3.2.6 for additional examples). Subtle differences are apparent in the distributions of the *taŋ* and *thak* forms: N *taŋ ku* is preferred with verbs of vocalization when the referent of the noun phrase to which the speech act is directed is human or personified; N *taŋ nə* is mostly used with verbs of deictic motion, also when the oblique noun phrase functioning as the goal has a human or personified referent; and N *thak ku* occurs with a handful of activity verbs expressing physical contact with a human or personified referent of an oblique noun phrase, but in addition is occasionally used with

verbs of vocalization. Thus, in all uses of *taŋ* and *thak* the referent of the noun phrase head belongs to a very specific semantic class of noun. The remaining forms mostly occur with the locative marker and have no restrictions relating to the semantic class of the referent.

Table 5.2. Nascent postpositions

FORM	SOURCE IF KNOWN	SEMANTICS	OCCURS WITH
<i>taŋ</i> (SIDE)	<i>taŋi</i> ('side' + ?PROX)	in direction of N	LOC, ALL
<i>thak</i> (PLACE)	<i>tak</i> (<Chungli 'place')	in direction of N; on N	LOC, ALL
<i>ɹuŋ</i> (AMONG)		among N	LOC
<i>mənaŋ</i> (BEFORE)	<i>mənaŋ</i> 'first, front'	in front of N	LOC
<i>in</i> (PATH)	<i>inti</i> 'path'	in direction of N	LOC, ALL
<i>sin</i> (BACK)	<i>tə-sin</i> (RL-back)	behind N	LOC
<i>ma</i> (FACE)	<i>tə-ma</i> (RL-face)	on top of N; above N	LOC
<i>paŋ</i> (MOUTH)	<i>tə-paŋ</i> (RL-mouth)	upper edge of N	LOC, ALL
<i>pùk</i> (STOMACH) ¹⁰	<i>tə-puk</i> (RL-stomach)	within N; under N	LOC, ALL
<i>təɹəm</i> (MIDDLE)	<i>tə-ɹəm</i> (RL-middle)	middle of N	LOC
<i>tsəŋta</i> (INSIDE)	<i>tə-tsəŋ</i> (RL-inside)	inside/between N	LOC

These combinations represent the most textually frequent pairings found in locative or allative noun phrases with a human or personified referent. While the inclusion of a nascent postposition appears to be common but not obligatory for oblique noun phrases of activity verbs denoting physical contact, it is unusual for locative or allative noun phrases to lack an accompanying nascent postposition if the referent is human and the predicate is a verb of deictic motion or vocalization. The odd exception is occasionally encountered, such as in (5.69) below, in which the allative marker alone is used to mark the human goal of what can only be an underlying verb of speech. Such omissions occur so rarely, however, that they must be considered marginal (cf. [5.32a] for another example).

- (5.69) *təni tʃu nə “aki tʃhapaŋ!”*,
 tə-ni tʃu nə a-ki tʃhapaŋ-aŋ
 RL-wife DIST ALL NRL-house open-IMP
 ‘[When the son-in-law said] to the wife “Open the house!,” ...’

10. This root, referring to the visceral organ ‘stomach’, should not be confused with ‘belly’, the form of which is the compound *tə-puk-làŋ* (RL-stomach-bottom) in Mongsen.

The following examples are representative of the dominant patterns noted for the above-mentioned semantic classes of predicates when the referent of an oblique noun phrase belongs to the human/personified semantic class of nominal.

- (5.70) *ní? la hənsəlila nə sa, ahən təpuŋla taŋ ku,*
 ní? la hənsəli-la nə sa
 one.day TOP leopard.cat-F AGT say.PST
 a-hən təpuŋ-la taŋ ku
 NRL-chicken cock-F SIDE LOC
 ‘One day Leopard Cat said to Rooster, ...’

- (5.71) *ami? təmáŋ taŋ nə mùwala.*
 a-mi? təmáŋ taŋ nə m̀ə-wa-la
 NRL-person all SIDE ALL NEG-go-NEG.PST
 ‘[The baby] didn’t go to any of the people.’

- (5.72) *akhula nə pùŋìla thak ku puŋtshəŋku,*
 a-khu-la nə pùŋì-la thak ku puŋtshəŋ-ku
 NRL-tiger-F AGT wild.pig-F PLACE LOC pounce.on-LOC.CV
 ‘When Tiger pounced on Wild Pig, ...’

Clark ([1911] 1990: 236) lists a lexical noun *dak* ~ *tak* in his Chungli dictionary and mentions that it sometimes means ‘place’ or ‘location’, e.g. *im-tak* ‘place occupied by a village’. Significantly, he also states that verbs of hitting, cutting and spearing etc. frequently require *dak* as a suffix to the object struck if the referent is animate, and sometimes if it is inanimate. This function of *dak* ~ *tak* as a nascent postposition in Chungli is mirrored by the use of *thak* in Mongsen, suggesting that they must share the same historical source. Furthermore, given that a form *taŋi* still occurs as a lexical noun meaning ‘side’, it is highly likely that all [noun/pronoun + locational/body part noun] pairings were originally a type of right-headed possessive compound in which the head expressed a generic locational meaning. As suggested above in §5.3.2, this would explain why possessive pronouns are more frequently used with local case marking clitics. Consider the right-headed [possessor-possessee] syntactic structure of noun phrases involving *taŋi* in the following examples.

- (5.73) *kə taŋi si-aŋ.*
 kə taŋi si-aŋ
 1SG.POSS side turn-IMP
 ‘Turn towards me.’ (Literally: ‘Turn [to] my side’)

- (5.74) *tə lítʃápá? nə álíma tsəmà? líma taŋi jáhəj.*
 tə lítʃá-pà? nə á-líma tsəmà? líma taŋi já?
 thus PN-M AGT NRL-world Assam country side PTCL
 həj
 EXCLM
 ‘Thus, the god Lichaba [made] this Assam Plain side of the world,
 okay?’

The following text example has the allative marker occurring twice in what is referentially the same noun phrase, the first to mark *nuksənsaŋpà?*, and the second to mark *taŋ*. This is extremely unusual. The more common presentation is for the allative to occur just once as the last constituent of the noun phrase. The fact that allative case marking can occur on both the head *and* the nascent postposition indicates that *taŋ* in this example actually has the status of a noun in apposition to the allative case marked noun phrase. It is therefore analyzed as a noun phrase head and glossed in lower case accordingly.

- (5.75) *nuksənsaŋpà? nə taŋ nə walíkà?*
 nuksənsaŋ-pà? nə taŋ nə wa-likà?
 PN-M ALL side ALL go-CONTEMP
 ‘When [the son-in-law] had gone to Noksensangba, ...’

By analogy, examples (5.73–5.75) betray the historical origins of all nominal morphemes used synchronically as nascent postpositions. From its earlier lexical function as an independent apposed noun, and then as the head of a possessive compound noun, *taŋi* has now been reanalyzed and conventionalized as a nascent postposition *taŋ* encoding direction. The final *i* vowel of the noun *taŋi* in (5.73) is probably a fossilized proximate nominal demonstrative that has subsequently fallen victim to the tendency of grammatical morphemes to reduce their polysyllabic structure. It is pertinent to mention that the sentence of (5.74) was accompanied by the speaker motioning towards the Assam plain, as if to reinforce its deictic orientation.

The root of the relational body part noun *tə-puk* ‘RL-stomach’ has grammaticalized as a nascent postposition expressing an inessive or subessive meaning of ‘inside’ or ‘under’ and occurs in combination with either the allative or locative marker. This morpheme must originate as well from a possessive compound in which *puk* functioned as the compound head and expressed a location. Note that the nascent postposition carries a low tone, while the root of the lexical noun carries a mid tone.

- (5.76) *luŋpùk nə zaɪ*,
 luŋ-pùk nə za-əɪ
 stone-STOMACH ALL enter-SEQ
 ‘After going inside a cave, ...’

- (5.77) *miɪi tʃàŋpùk ku liaw?*
 miɪi tʃàŋ-pùk ku li-a-ù?
 ant.sp. foot-STOMACH LOC be-VSF-DEC
 ‘There is a large ant under [your] foot.’

By the same process, the root of the relational body part noun *tə-sin* ‘RL-back’ has developed an auxiliary function meaning ‘behind’ from what was originally a possessive compound noun head. The grammaticalized postposition helps to specify the spatial location of an oblique noun phrase referent.

- (5.78) *tshəluŋla tʃu akhula sin ku mənəɪ*,
 tshəluŋ-la tʃu akhu-la sin ku mən-əɪ
 fox-F DIST tiger-F BACK LOC sit-SEQ
 ‘Fox, having sat down behind Tiger, ...’ (< ‘at Tiger’s back’)

In the sentence of (5.79) below, *ma* and *taŋ* occur juxtaposed to express ‘side’, according to the interpretations of consultants. These morphemes transparently originate from the lexical roots of the relational body part noun *tə-ma* ‘RL-face’ and the locational noun *taŋi* ‘side’. The noun phrase functioning as the goal has been postposed and is separated from the clause by an intonational pause following the matrix verb *wa*.

- (5.79) *təɪ pa thùŋà wa, akhu la mataŋ nə*.
 tə-əɪ pa thùŋ-a wa akhu-la ma taŋ nə
 thus-SEQ 3SG reach-SIM go.PST tiger-F FACE SIDE ALL
 ‘And then, he reached [the front side of] Tiger.’

The root of the body part noun *tə-ma* ‘RL-face’, used as a head in erstwhile possessive compounds, has combined with the dependent element and lexicalized in a few cases, resulting in the formation of new nouns. Some examples are *kimà* ‘courtyard’, from the nominal roots *ki* ‘house’ + *ma* ‘face’, and *lima* ‘country’, formed from the nominal roots *li* ‘soil’ and *ma* ‘(sur)face’.

Sometimes a body part noun functions as the only case marking morpheme of a noun phrase. The noun phrase case marked by *təma* in the following example fills the complement slot of a verbless clause. This was given in reply to the question ‘What’s your age?’.

- (5.80) *nətəɪ təmawʔ*
 nətəɪ tə-ma-ùʔ
 seventy RL-FACE-DEC
 ‘Above seventy.’

DeLancey (1997a) demonstrates that the members of some form classes undergo decategorialization in Tibetan, Tamang, Burmese and English, and consequently grammaticalization processes identical to those described here can result in gradual categorial shifts. In Tibeto-Burman languages in particular, the common pattern is for locational and body part nouns that historically functioned as independent apposed nouns or as heads of possessive compound nouns to be reanalyzed as dependent adpositions. The inherent semantics of these classes of noun is precisely what allows them to change their form class from nouns to adpositions. But since this shift is not abrupt or paradigmatic, it can result in transitional synchronic forms that are categorially in limbo. We see this synchronically with the two functions of *taŋ* – in the sentence of (5.75), it behaves formally as an independently case marked noun head, while its recategorized function as a nascent postposition is apparent in (5.79).

A consequence of this categorial shift is that the pronouns originally used to mark a possessive relationship to the relational noun heads of their oblique noun phrases have now become the heads of those oblique arguments. This provides a logical historical explanation for why possessive pronouns appear to be bound forms elsewhere in the grammar of Mongsén, yet demonstrate what can only be interpreted to be syntactic independence in oblique noun phrases. Historically they *were* dependent bound forms marking the syntactic role of the possessor, but the categorial shift of their possessed relational noun heads to the newer postposition class in oblique arguments has resulted in the reanalysis of possessive pronouns as the heads of these noun phrases.

5.4. Valency modifying and reciprocal/collective derivations

Mongsén has just one type of valency modifying derivation: the morphological causative derivation. A morphological causative always increases the valency of the predicate by adding an argument to the core. Another type of causative construction discussed in §5.4.2 is the analytic causative. This uses a purposive complement and the verb ‘give’ to express a permissive meaning, without affecting the valency of the predicate. The reciprocal/collective derivation also does not formally modify the valency status of the clause, although its semantics may suggest such a change, particularly when a transitive verb marked by the reciprocal/collective suffix expresses a reciprocal meaning and occurs as the head of a clause containing only one core argument.

Reflexives are encoded by a separate reflexive marker *-jimtâ?* attaching to a possessive pronoun that is coindexed to an antecedent mention. The functions of the reflexive marker are described in §7.5.5.

5.4.1. Morphological causative

Morphological causative derivations can be formed from both intransitive and transitive verb stems. The causative suffix *-i?* attaches to the stem in the penultimate slot of the predicate template (see §8.2) and increases the valency of the predicate by one. Semantically, the introduced core argument takes the role of a causer that induces a causee to perform the activity denoted by the predicate. An animate causee may or may not willingly perform the activity. It will be demonstrated that some semantic classes of verbs allow this semantic distinction to be grammatically encoded by the presence or absence of dative case marking on the causee argument. This proves to be somewhat analogous to the semantically motivated marking pattern involving the agentive case marker or its absence on the intransitive arguments of some non-causativized verbs (see §5.2.2 for description).

5.4.1.1. *Causative derivation of intransitive verb stems*

Causativization applies to intransitive activity verb roots, e.g. *.ià* ‘come’ > *.ia-i?* ‘made come’, stative intransitive verb roots, e.g. *pəla* ‘be happy’ > *pəla-i?* ‘made happy’, *sisa* ‘be awake’ > *sisa-i?* ‘woke up’, and intransitive verb roots denoting processes and events, e.g. *sə* ‘die’ > *sə-i?* ‘made die’. The introduced causer argument is obligatorily case marked by the agentive marker and functions as the A argument, while the original S argument of the non-causativized intransitive verb becomes the new O argument of the causativized verb and functions semantically as the causee. The changes in valency and case marking on the core arguments of causativized intransitive verbs are summarized in Figure 5.1 below.

The derived O arguments of causativized intransitive verbs generally receive no case marking in instances of reporting on pragmatically neutral situations. If, on the other hand, a speaker specifically wants to encode that the causee willingly performs the caused activity, and the inherent semantics of the intransitive verb plausibly permits the performance of that activity willingly or by coercion, then the causee noun phrase may instead take non-canonical dative case marking to encode this semantic distinction.

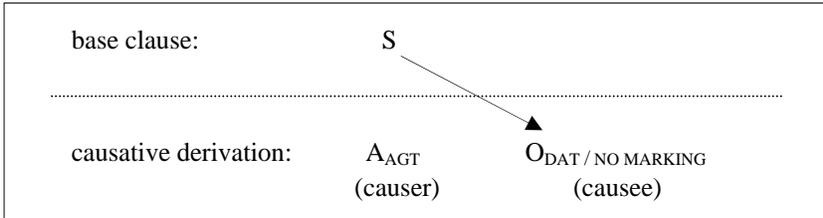


Figure 5.1. Causative derivation of intransitive verb stems

Semantically motivated marking of causee arguments only appears to be possible for some semantic classes of intransitive verb. Causativized verbs of motion, for example, do not permit their causee noun phrases to optionally take dative case marking, even if the causee happens to be a willing participant in the caused event.

- (5.81) a. *nì nə kəni waj?*
 nì nə kə-ni wa-i?
 1SG AGT 1SG.POSS-wife go-CAUS.PST
 ‘I made/let my wife go.’ (she may or may not have wanted to go)
- b. **nì nə kəni li waj?*
 1SG AGT 1SG.POSS-wife DAT go-CAUS.PST
 ‘I let my wife go.’

Judging from consultants’ elicited responses, it seems that the volitional status of a causee argument without dative case marking is in many situations left open to interpretation, irrespective of whether its referent is a willing participant, or forced to perform an activity against its will. Conversely, dative marking on a causee argument of a causativized intransitive verb always unambiguously encodes a participant that retains a significant degree of control over the performance of the caused event.

The elicited sentences of (5.82) below demonstrate how dative case marking or its absence on the causee O argument respectively encodes the presence or absence of control over an activity denoted by the verb root *ajim* ‘scream’. In (5.82a), the actress takes the role of a causee who has no control over her screaming; i.e. it is a spontaneous event that results from a situation in which the film director unexpectedly frightens her. The absence of dative marking on the O argument infers the referent’s lack of control. In contrast, dative marking on the causee argument of (5.82b) explicitly indicates that the actress retains control over a volitional act of screaming. This would be appropriate if the film director instructed the actress to play the part of a character who screams.

- (5.82) a. *filəm di.ɾəktəɾ.pàʔ nə aktɾəs.la ajim-iʔ.*
 filəm.di.ɾəktəɾ.pàʔ nə aktɾəs-la ajim-iʔ
 film.director-M AGT actress-F scream-CAUS.PST
 ‘The film director made the actress scream.’ (by frightening her)
- b. *filəm di.ɾəktəɾ.pàʔ nə aktɾəs.la li ajim-iʔ.*
 filəm.di.ɾəktəɾ.pàʔ nə aktɾəs-la li ajim-iʔ
 film.director-M AGT actress-F DAT scream-CAUS.PST
 ‘The film director made the actress scream.’ (by commanding her)

It appears that the semantic classes of intransitive verb roots that allow the semantically motivated agentive case marking pattern described in §5.2.2 also allow similar case marking possibilities on a causee argument of a causativized verb for the purpose of signalling the presence or absence of control.¹¹ Although this is yet to be systematically investigated according to each semantic class of verb, the data at hand does suggest such a correlation. Compare, for instance, the analogous use of dative marking (or its omission) on the causee noun phrases of (5.83) with the case marking possibilities presented by the same verb root in the elicited examples of (5.28–5.30).

- (5.83) a. *tuktəɾ nə ánítila akhət-iʔ.*
 tuktəɾ nə ánítila akhət-iʔ
 doctor AGT woman-F cough-CAUS.PST
 ‘The doctor made the woman cough.’ (spontaneous, uncontrolled activity)
- b. *tuktəɾ nə ánítila li akhət-iʔ.*
 tuktəɾ nə ánítila li akhət-iʔ
 doctor AGT woman-F DAT cough-CAUS.PST
 ‘The doctor got the woman to cough.’ (controlled, purposeful activity)

The sentence of (5.83a) would be appropriate if the doctor did something to the woman that caused her to have a coughing fit, such as inserting something into her trachea. In contrast, the sentence of (5.83b) would only be appropriate if the doctor were performing a medical examination and requested the woman to cough. The causee of (5.83a) is made to perform an activity over which she has no personal control and its noun phrase is consequently unmarked for case, whereas the dative-marked causee of (5.83b) is a willing participant who performs the activity of coughing by prior arrangement. The use of dative case marking or its absence to encode the amount of agentivity retained by the

11. This does not apply to verbs of motion, however, which only allow semantically motivated case marking in the limited circumstances described in §5.2.2.

causee thus correlates with an identical semantic distinction encoded by core and oblique case marking in Hungarian and Japanese intransitive causative constructions, as discussed by Comrie (1985: 337).

The causative suffix optionally functions as a valency increasing morpheme without necessarily encoding causation when it is suffixed to the verb root of *məni* ‘laugh/smile’. This is probably attributable to the inherent semantics of this particular intransitive verb, and not just in Mongsen. Austin (2005) demonstrates that a number of Australian languages have a split verb class system in which the same transitivity affix results in a causative derivation for one sub-class, and an applicative derivation for the other sub-class. It is noteworthy that the verb ‘laugh’ is found to pattern with the applied sub-class of intransitive verbs in many of these languages.

(5.84) a. *səntipà? nə asaŋla mənəj?*

sənti-pà? nə asaŋ-la məni-i?
PN-M AGT PN-F laugh/smile-CAUS.PST

(i) ‘Sentiba laughed/smiled at Asangla.’

(ii) ‘Sentiba made Asangla laugh/smile.’ (causee may or may not have had control)

b. *səntipà? nə asaŋla li mənəj?*

sənti-pà? nə asaŋ-la li məni-i?
PN-M AGT PN-F DAT laugh/smile-CAUS.PST

‘Sentiba made Asangla laugh/smile.’ (causee had control)

In (5.84a), the free translation of (i) suggests that the causative marker increases the valency of the clause by transitivity affixing the verb stem and adding an argument to the core. An intransitive verb stem causativized by this suffix normally introduces a new A argument in the semantic role of a causer and reassigns the old S argument to a causee semantic role in O function. But in this case, it results in the promotion of what is semantically an oblique argument to the core, in O function. The erstwhile S retains its status as the controller of the activity, but additionally takes agentive case marking to identify it as the A argument of the derived transitive clause. Thus, suffixing *-i?* to the verb root of *məni* ‘laugh/smile’ can optionally produce an applied reading of the same clause.

Alternatively, the free translation of (ii) assumes a causative derivation in which an A argument is introduced in the role of the causer and the original S argument becomes the causee O argument of the causativized verb, as shown in Figure 5.1. Whether or not the causee has control over the activity is left open to contextual interpretation if the O argument receives no formal case marking. On the other hand, dative case marking on the derived O argument in (5.84b)

unequivocally encodes its noun phrase as a causee in the semantic role of a willing participant that retains some degree of control over the caused activity.

The intransitive verb root *məni* ‘laugh/smile’ is the only one encountered in the corpus that can be transitivized by the causative suffix to derive either a causative or an applicative meaning. Dative marking on the derived O argument of (5.84b) blocks the applicative reading and only allows the causative interpretation. Interestingly, Mongsen speakers disagree amongst themselves with respect to the use of dative marking on the causee arguments of some causativized intransitive constructions. One native speaker in his twenties rejected the use of dative marking in (5.84b), claiming that an interpretation of causee control would only be possible with an analytic causative construction using the verb root *khi?* ‘give’ as the matrix verb (cf. the examples of §5.4.2). This may be indicative of an age-related shift in the semantics of the morphological causative derivation.

Lastly, while it is most common to find an animate referent in the role of a causer, it is also possible for an inanimate referent to fill this role if it can function as the effector of an event or process, as in (5.85) below.

- (5.85) *tsəhŋi nə tsəpà? kuŋtʃuki?*
 tsəhŋi nə tsəpà? kuŋ-tʃuk-i?
 sun AGT pond be.dry-PFV-CAUS.PST
 ‘The sun dried out the pond.’

5.4.1.2. *Causative derivation of transitive verb stems*

The morphological causative similarly increases the valency of the predicate by one place when it is suffixed to a transitive verb stem. Causative derivation applies to all transitive activity verb roots, e.g. *tʃā?* ‘consume’ > *tʃā?-i?* ‘made consume’, *təpsət* ‘kill’ > *təpsət-i?* ‘made kill’, *tən* ‘sing’ > *tən-i?* ‘made sing’, transitive verbs of perception and cognition, e.g. *hŋalak* ‘learn’ > *hŋalak-i?* ‘made learn’, and transitive verbs of vocalization, e.g. *sana* ‘speak’ > *sana-i?* ‘made speak’.

When a transitive verb is causativized, the introduced argument functions semantically as the causer and obligatorily receives agentive case marking. The original A argument is demoted and becomes the causee, which either takes dative marking, or receives no overt case marking. Lastly, the original O argument of the base clause remains in O function in the causative derivation and thus is formally unmarked for case. The derivation of causativized transitive verbs is summarized in Figure 5.2 below.

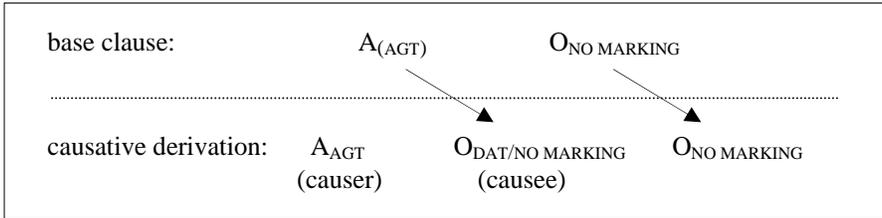


Figure 5.2. Causative derivation of transitive verb stems

It appears that the presence of dative marking on the causee noun phrase of a causativized transitive clause implies a willing referent, while its absence implies that the degree of coercion is left open to interpretation. The dearth of naturally occurring examples in the corpus does not allow confident claims to be made at present, and examples elicited from a number of consultants resulted more in tendencies than consistent patterns. That being said, the sentences of (5.86a–b) generally appeared to encode a fairly clear contrast between a willing and an unwilling causee for the speakers consulted.

(5.86) a. *təja nə anu li muli tʃãʔ-iʔ*?

tə-ja nə anu li muli tʃãʔ-iʔ

RL-mother AGT child DAT drug eat-CAUS.PST

‘The mother got the child take the medicine.’ (the child was willing)

b. *təja nə anu muli tʃãʔ-iʔ*

tə-ja nə anu muli tʃãʔ-iʔ

RL-mother AGT child drug eat-CAUS.PST

‘The mother forced the child take the medicine.’ (the child was unwilling)

The contrast in the degree of agentivity retained by the causee was less apparent for the sentences of (5.87), however. The absence of dative case marking on the causee argument of (5.87b) was reported to imply more of an order, although a nuance of strong coercion was absent. This suggests that the meaning encoded by the presence or absence of dative marking may be sensitive to the inherent semantics of the particular transitive verb. Some causativized transitives, such as *hnàkjak-iʔ* ‘made guard’ in these sentences, logically suggest an inherently amenable causee, whereas the possible contexts in which the causativized verb *tʃãʔ-iʔ* ‘made eat’ could be used freely permit either an acquiescent causee that is cajoled into eating something, or a strongly resistant one that must be forced.

- (5.87) a. *muwapà? nə sipaj luŋ li ai.iə hnàkjak-i?*
 muwa-pà? nə sipaj luŋ li ai.iə hnàk-jak-i?
 PN-M AGT soldier group DAT enemy guard-RS-CAUS.PST
 ‘Moaba got the platoon of soldiers to guard the enemy.’
- b. *muwapà? nə sipaj luŋ ai.iə hnàkjak-i?*
 muwa-pà? nə sipaj luŋ ai.iə hnàk-jak-i?
 PN-M AGT soldier group enemy guard-RS-CAUS.PST
 ‘Moaba ordered the platoon of soldiers to guard the enemy.’

In a non-causative clause formed with potentially three-argument verbs of transference, such as *khi?* ‘give’, the donor receives agentive case marking, the gift receives no case marking, and the recipient receives dative case marking. In basic underived clauses, the order of clausal arguments is determined by pragmatic influences affecting the presentation of information.

- (5.88) *taksi nì-tʃhaŋ nə pulis li atʃhən khi?-ukù.*
 taksi nì-tʃhaŋ nə pulis li atʃhən khi?-ukù
 taxi lead-MAN AGT police DAT money give-ANT
 ‘The taxi driver has given money to the police.’

Now, when the verb root *khi?* ‘give’ undergoes a causative derivation, a fourth argument is added to the core in the semantic role of a causer. This results in an increase in valency by one place and a change in case marking on just the derived causee argument, as demonstrated in Figure 5.3 below.

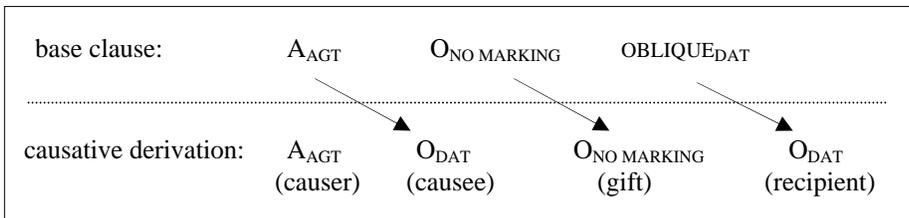


Figure 5.3. Causative derivation of three-place verb stems

The newly introduced noun phrase takes agentive case marking in the role of the causer. The old A argument of the non-causativized verb becomes the dative case-marked causee of the causativized verb, and the gift and the recipient retain their original semantic roles and case marking.

- (5.89) *apaŋtʃhaŋəɪ nə taksi nìtʃhaŋ li pulis li atʃhən khəjukù.*
 apaŋtʃhaŋ-əɪ nə taksi nì-tʃhaŋ li pulis li atʃhən
 man-ANOM AGT taxi lead-MAN DAT police DAT money
 khì?-i?-ukù
 give-CAUS-ANT
 ‘The man has made the taxi driver give money to the police.’

Dative case marking is used twice in the causative derivation – first on the demoted causee, and again on the recipient of the gift. In this situation, word order is used to disambiguate the semantic roles of the two dative-marked noun phrases. Reversing the relative order of *taksi nìtʃhaŋ* ‘taxi driver’ and *pulis* ‘police’ in the clause results in the interpretation that *pulis* is the causee and *taksi nìtʃhaŋ* is the recipient. The position of other arguments in the clause has no bearing on the interpretation of the semantic roles of dative case-marked noun phrases.

5.4.2. Analytic causative

The analytic causative uses the verb root *khì?* ‘give’ as the main verb of a complex sentence that encodes a permissive meaning of ‘let VERB’. This verb is able to take all the usual tense, aspect and modality morphology associated with the verb class. The matrix A argument of the analytic causative construction is assigned agentive case marking, just as it usually would be in the non-causative use of *khì?* ‘give’ as a main verb, especially if the disambiguation of semantic roles is required. The permitted event is expressed by an intransitive or transitive verb root marked with a combination of the irrealis marker *-ì* and the general nominalizer *-pà?*. It occurs with its associated noun phrase argument(s) as an embedded complement of the matrix clause, in O function. This is distinguished in (5.90) below by square brackets.

- (5.90) *nì nə kəni wajpa? khìw?*
 nì nə [kə-ni wa-ì-pà?] khì?-ù?
 1SG AGT 1SG.POSS-wife go-IRR-NR give.PST-DEC
 (My wife wanted to go) ‘I let my wife go.’

The O complement of (5.90) expresses a purposive-like meaning that is functionally equivalent to an infinitive. It is superficially similar to one of the purposive nominalizations described in §7.4.7, except that the noun phrase functioning as the causee argument is syntactically a constituent of the

embedded complement, and therefore functions independently as the notional intransitive argument of the nominalized verb *wajpaʔ*.

An analytic causative formed with a nominalized transitive predicate as the head of the embedded complement has a similar structure. The causee noun phrase of the transitive complement, however, is consistently assigned dative case marking, which distinguishes it from the unmarked O argument of the complement. This may result from a restriction on noun phrases with agentive marking occurring in both the matrix clause and the O complement. The use of dative marking for distinguishing the causee noun phrase mirrors that of case marking on the causee arguments of morphological causative constructions previously discussed in §5.4.1.

- (5.91) *nì nə niŋ tʃàɪ li áhŋáʔ phàjpàʔ khiwʔ.*
 nì nə [niŋ tʃàɪ li á-hŋáʔ phàʔ-ì-pàʔ]
 1SG AGT 2PL.POSS son DAT NRL-fish catch-IRR-NR
 khiʔ-ùʔ
 give.PST-DEC
 (i) (Your son wanted to catch fish) ‘I let your son catch fish.’
 (ii) ‘I gave fish to your son to catch.’

The transitive clause of (5.91) has an alternative meaning if the dative case marked noun phrase is interpreted as a recipient of a non-causative verb ‘give’, rather than as the causee argument of a permissive analytic causative. Unlike the morphological causative, the analytic causative of (5.91) does not change the valency of the clause by adding an argument to the core of the matrix clause. This is demonstrated by the alternative causative and non-causative interpretations of this sentence.

5.4.3. Reciprocal/collective derivation

The verbal suffix *-təp* is used to derive reciprocal and collective meanings. It can mark both transitive and intransitive verb stems. An intransitive verb invariably derives a collective meaning when modified by this suffix.

- (5.92) *íjaʔ pəlatəpja.ùʔ.*
 íjaʔ pəla-təp-ja-àɪ-ùʔ
 much be.happy-COLL-CONT-PRES-DEC
 ‘[They] are very happy together.’

(5.93) *túnəla luŋi àtʃhəmatʃhəməkə məntəpəi liaw?*

túnəla luŋ i àtʃhəm-atʃhəm-əkə mən-təp-əi li-a-ù?
 3PL group PROX be.quiet-REDUP-SEQ sit-COLL-SEQ be-VSF-DEC
 ‘This group is sitting together very quietly.’

Many transitive verbs taking the reciprocal/collective marker allow either a reciprocal or a collective interpretation. The particular meaning that the suffix derives in a given instance of use – whether reciprocal, e.g. *the men are fighting each other*, or collective, e.g. *the men are fighting [the enemy] together* – is largely dependent upon the specific semantics of the predicate involved, the context in which it is used, and most importantly, whether there is ellipsis of the O argument. For example, a reciprocal meaning is not possible with the transitive verb root *atsə* ‘look’ in the text example of (5.94), because this interpretation is blocked by the presence of an overtly stated O argument. But if the core argument *tʃaŋsə* ‘footprint’ were omitted from its O slot in this sentence, then in a different context the sentence could alternatively mean ... *tomorrow [we] will just go and look at each other*.

(5.94) *təpàla asaŋ áŋ waɪ tʃaŋsə atsətəpju?*

tə-pàla asaŋ áŋ wa-əi tʃaŋsə atsə-təp-ì-ù?
 thus-COND tomorrow just go-SEQ footprint look-COLL-IRR-DEC
 ‘If that’s the case, then tomorrow [we] will just go and look at the footprints together.’

On the other hand, either a reciprocal or a collective meaning is potentially possible with the nominalized transitive verb root *ləp* ‘cut’ of (5.95), since the O argument is not overtly mentioned. Although the speaker is actually referring to the time when his village was involved in reciprocal, as opposed to collective head-hunting with other villages, in a different context this could have a collective nuance, i.e. ... *in the time that [we] cut off [heads] together*. A substantial pause following *maŋmətuŋ jim tʃu* identifies this noun phrase as an extra-clausal topic.

(5.95) *təi maŋmətuŋ jim tʃu, ləptəpà? hmapaŋ ku,*

tə-əi maŋmətuŋ jim tʃu ləp-təp-pà? hmapaŋ ku
 thus-SEQ village.name village DIST cut-RECIP-NR time LOC
 ‘And as for Mangmetong village, in the time of head-hunting ...’
 (Lit. ‘... in the time that [we] cut off each others’ heads ...’)

On semantic grounds, the reciprocal interpretation suggests a reduction in valency, whereas the collective meaning does not. In the absence of a formal

mechanism for signalling a reduction in the number of clausal arguments, the valency status of a predicate marked by the reciprocal/collective derivation must be construed with respect to the pragmatic context in which it is used.

Lastly, some transitive verbs can only impart a collective meaning when modified by *-təp*, regardless of the presence or absence of an overt O argument; this restriction is dictated by the semantic entailments of the particular verb involved. The following two examples are illustrative of verb stems that uniquely suggest a collective interpretation of their meanings, independent of context.

(5.96) *tə tɯŋkhəla tʃu aʃhu thəməkə təntəpəi sənsaw?*

tə tɯŋkhəla tʃu a-ʃhu thəm-əkə tən-təp-əi
 thus 3PL DIST NRL-song end-SIM sing-COLL-SEQ

sənsa-ù?

depart.PST-DEC

‘Thus, they all finished singing that song together and departed.’

(5.97) *sətəjpa? tsəŋikhu tʃu zaŋtəpəi, məʃatshəŋ tʃu ajim nə wa.*

sati-ì-pà? tsəŋikhu tʃu zaŋ-təp-əi
 wedding-IRR-NR day DIST arrange-COLL-SEQ

məʃatshəŋ tʃu a-jim nə wa

PN DIST NRL-village ALL go.PST

‘Having arranged [their] day of marriage together, Mechatseng went to the village.’

Additional examples demonstrating the use of the reciprocal/collective suffix can be found in §8.5.2.

Chapter 6

The noun phrase, relative clauses and nominalizations

The focus of this chapter is the syntax of the Mongsen noun phrase. Nominal constituents have been described in the discussion of word classes in Chapter 4. We now look at their syntagmatic arrangement in the noun phrase and consider the types of nominal elements that can function as heads and phrasal adjuncts.

6.1. Basic noun phrase structure

The template of (6.1) sets out the linear order of constituents in a maximally specified noun phrase. Letters enclosed by parentheses in the template indicate optional constituents; those in braces indicate that only one of the enclosed constituents may occupy a particular slot. A subscript number following a letter (e.g. C₂) indicates that a constituent has more than one possible position of a single occurrence in the noun phrase.

(6.1)	({A ; D ₁ })	B	(C ₁)	(D ₂)	(E)	(C ₂)	(F)	(G)	(C ₃)
	DETERMINER ATTRIBUTE	HEAD	TOPIC/FOCUS	ATTRIBUTE	QUANTIFIER	TOPIC/FOCUS	DETERMINER	CASE	TOPIC/FOCUS

Three discourse particles can be used inside the noun phrase to constrain or highlight pragmatic reference to a particular constituent (see following section). Two of these demonstrate an adverb-like freedom of occurrence. All may also occur elsewhere in the clause, where they are similarly used for pragmatically-motivated purposes (see §4.2.14.3 and §4.2.14.4). While discourse particles are not actually nominal elements, their potential appearance within the noun phrase warrants the description of their use here. The same could be said of the case-marking clitics, which are associated with noun phrases, yet are strictly speaking not syntactic constituents of noun phrases, at least synchronically. There is, however, plausible evidence that some members of this class have diachronic origins in body part nouns that once formed the heads of their noun phrases, prior to their grammaticalization as relational morphemes (see §5.3.3).

Apart from discourse particles, some of the nominal adjuncts also demonstrate more than one possible position of occurrence in the noun phrase; these are introduced in §6.1.2 below and cross-referenced to sections in which they are described more fully. In the remaining sections of this chapter I give an overview of noun phrase constituents, then proceed to the description of the functions and types of noun phrases, the possible combinations of various noun phrase constituents, and the functions of relative clauses and nominalizations.

6.1.1. Particles occurring within the noun phrase

Three discourse particles – the topic particle *la* (see §4.2.14.3), the restrictive focus particle (*t*)*áŋ* ‘just, only’, and the additive focus particle *kà?* ‘also, even’ (see §4.2.14.4) – are all represented by (C) in the template. As noted above, these have the same pragmatic functions within noun phrases as they are generally found to have in the clause. With the exception of *kà?*, the discourse particles do not have rigidly fixed positions and can be used noun-phrase internally to contrast, constrain or augment reference to individual constituents, or to just some noun phrase constituents. When used phrase-finally, a particle’s domain of application extends to the whole of the noun phrase, thus its position of occurrence determines its scope of reference. The topic and discourse focus particles are mutually exclusive, so there can only be a single occurrence of either in a noun phrase. These have the greatest freedom of occurrence and appear after the head, after an attribute or quantifier, and in the noun phrase-final position, whereas the additive focus particle is only found at the end of the noun phrase. Its limited distribution may be attributable to a relatively more recent grammaticalization as a discourse focus particle from the adverb *anukà?* (see §4.2.9).

6.1.2. Noun phrase constituents with a variable position of occurrence

Three nominal modifiers show a limited freedom of occurrence. Firstly, the distal and anaphoric nominal demonstratives *tʃu* and *sə* can optionally occur between the two syllables of the ablative case marker *phinə*, rather than to the left of it in the determiner slot, e.g. *phi <tʃu> nə* and *phi <sə> nə*. This is so marginal that I have not attempted to accommodate it in the template (see §5.3.2.2 for discussion of the historical factors responsible for the “infix” position of nominal demonstratives). Secondly, relative clauses function as a type of nominal attribute and occur in the post-head attribute slot (D₂), where they modify the preceding head. But note that it is also possible for a relative clause to occur in front of the head in the (D₁) slot. Relativized modifiers occur

marginally less frequently in the pre-head position than in the post-head position; these positions often correlate with restrictive and non-restrictive reference respectively. Relativized and non-relativized nominal attributes are discussed in §6.4.1, and we look at relative clauses again in more detail in §6.6.

Lastly, one other distributional irregularity deserves special mention. When a quantifier is used with the nouns *ni* ‘day’ or *hmapaŋ* ‘time’ it precedes the head, e.g. *phəli ni* ‘four days’, *laŋpən ni* ‘whole day’ and *khən hmapaŋ* ‘once, one time’. An atypical position is also noted for the word *akhi* ‘ancient’ when it is used to modify *hmapaŋ*, e.g. *akhi hmapaŋ* ‘ancient times’. This is only possible with the semantic class of temporal nouns, and these particular nouns are the only members of this semantic class that permit a quantifying or modifying adjunct to precede the head. The quantifier *tesathəŋ* ‘half’, for instance, can only occur in the usual post-head position when modifying *hmapaŋ*, e.g. *hmapaŋ təsathəŋ ku* ‘in half [the] time’. These quantifier/modifier-noun combinations with their atypical order probably originate from right-headed noun-noun compounds whose dependent constituents have been reanalyzed as quantifiers or modifiers.¹ This would explain the unusual order of their constituents, but it is presently difficult to account for the pre-head position of numerals when they are used to quantify *ni* or *hmapaŋ*. The only logical explanation is that numerals also adopted the pre-head position for quantifying these particular noun heads, perhaps in structural analogy to the erstwhile dependent constituents of right-headed compounds. Like the “infix” position of nominal demonstratives within the ablative case marker, the unusual pre-head position of these few constituents is of such limited scope that we can safely ignore it as an extremely marginal variation of noun phrase structure.

6.1.3. Overview of other noun phrase constituents

A non-spatial determiner *ípáʔ* that I tentatively gloss as the emphatic demonstrative (§4.2.7.1) optionally occurs in the pre-head slot (A), a position it shares with pre-head relativized attributes. The emphatic demonstrative is likely to have its diachronic source in an old nominalization formed with the segmentally similar general nominalizer *-pəʔ*, because a noun phrase determined by *ípáʔ* prohibits the simultaneous presence of a juxtaposed pre-head relative clause. Conversely, a noun phrase modified by a pre-head relative clause cannot be determined by the emphatic demonstrative. This mutually exclusive constraint suggests that these constituents must occupy the same slot in the noun phrase template (see §6.4.3 for further discussion). In the absence of either of these constituents, the head is otherwise the leftmost element. Simple

1. Compare how *akhi* is used alone as a noun phrase head in the sentence of (6.25).

noun phrases overall tend to have a head-first structure. Types of heads are described below in §6.3.

Following the head are three nominal adjunct slots. The first of these, slot (D₂), may be filled by words that function as non-relative descriptive modifiers of the head, such as deverbal adjectives derived from stative verbs by the nominalizing prefix, e.g. *tə-hláj* ‘NZP-be.long’ in *hmapaŋ tə-hláj à* ‘a long time’, or nominalized verb stems that function as post-head relative clauses, e.g. *phù?-pà?* ‘steal-NR’ in *atʃhən phù?-pà? tʃu* ‘the money that was stolen’. Next, in slot (E), come the numerals and lexical quantifiers. Since these share a quantifying function, it is descriptively economical to treat them as members of the same lexical class. Quantifiers can be referred to restrictively by placing a restrictive focus particle in the C₂ slot immediately following the quantifier, e.g. *məɪətsə phaŋa táŋ* ‘only five chillies’. The lexical quantifiers and cardinal numeral word classes are discussed in §4.2.7.2 and §4.2.7.3 respectively. The slot following a (potentially focused) quantifier is filled by post-head nominal deictics functioning as determiners in (F). The deictic word class is described in detail in §4.2.7.1. In the absence of a phrase-final topic or discourse focus particle, a case-marking clitic encoding the case relation of the noun phrase occurs after the final constituent. Some oblique noun phrase heads may additionally occur with a nascent postposition preceding a local case-marking clitic.

In §6.1 I briefly mentioned that case-marking clitics are not normally considered part of the noun phrase. Given that some of the case markers of Mongsen have diachronic origins in relational nouns that have undergone recategorization as grammatical morphemes, it seems indisputable that at least some of these must have been noun phrase constituents at an earlier stage of the language. Furthermore, the fact that the oblique case markers occur with possessive pronouns further strengthens the premise that they originally formed the heads of their noun phrases, prior to their recategorization as case markers. The functions of all case marking clitics are discussed in §5.3; see §5.3.3 specifically for evidence of the origin of oblique case markers in relational nouns of the body part semantic class.

6.2. Functions of noun phrases

A noun phrase is a syntactic constituent whose head is a noun or a pronoun. It can function as:

- (i) a core argument in A, S or O function
- (ii) an oblique argument
- (iii) a verbless clause or copula clause topic

- (iv) a verbless clause or copula clause complement
- (v) an extra-clausal topic

A rising intonation is often superimposed over the individual tones associated with the syllables of words occurring in the noun phrase. This prosody effectively functions as a boundary signal delineating the phrasal domain of the noun phrase. A description of the prosodic functions of postlexical intonation applying to the noun phrase can be found in §3.4.1.

6.3. Head

Noun phrases generally have a form of nominal listed in Table 4.2 as a head.² Noun phrases can also be headless and consist of just noun phrase modifiers, such as when a quantifier is used alone to refer to participants e.g. *anət wa* ‘two went’, *təlu? ɹà* ‘all came’. In such situations the omitted head is implicit, e.g. ‘two (men) went’, ‘all (the villagers) came’. Contextual pragmatics and/or antecedent mention allow the reference of quantifiers and other constituents of a headless noun phrase to be determined. Relative clauses and nominalizations also can be headless; these are discussed in §6.6.

6.3.1. Simple head

A simple noun phrase head is a single free form nominal element, such as a noun, a personal or possessive pronoun, or a demonstrative pronoun. Examples of simple noun heads are provided in the following sentences. Clausal noun phrases relevant to the discussion are henceforth identified by square brackets and subscript labels, and their heads are distinguished by bold face.

- (6.2) *mətʃatshəŋ tʃu əjim nə kəwa.*
 [mətʃatshəŋ tʃu]_{NP} [a-jim]_{NP} nə kəwa.
 PN DIST NRL-village ALL ascend+go.PST
 ‘Mechatseng went up to the village.’

- (6.3) *pə tʃu təsin ku tʃhuwa.*
 [pə tʃu]_{NP} [tə-sin]_{NP} ku tʃhuwa
 3SG DIST RL-back LOC emerge.PST
 ‘She returned behind [the others].’

2. In addition, certain verb stems can be nominalized by affixal morphology and then used as the heads of their own noun phrases – see §7.4.1, §7.4.4, §7.4.5 and §7.4.6 for discussion and examples.

- (6.4) *pi təsəla.*
 [pi]_{NP} tə-səla
 PROX PROH-untie
 ‘Don’t untie this.’

6.3.2. Compound head

Binary compound nouns functioning as noun phrase heads are formed from noun-noun appositions, noun-verb combinations, and nominalized verb-noun or noun-nominalized verb combinations in various kinds of modifier-modified or implicit possessor-possessee relationships, e.g. *ana-tsə* (bee-water) ‘honey’, *hən-tʃha* ‘chicken’s feathers’, and *tɯŋ-tən* (trunk-sing) ‘log drum’. All types of compound nouns are described in §7.1.

Binary compound noun heads function as unitary constituents and are collectively modified by quantifiers or determiners, and case-marked by case-marking clitics. Like noun phrases with simple heads, these have a linear arrangement of constituents.

- (6.5) *aja təɬəm tʃu ku .ɬà.*
 [a-ja tə-ɬəm tʃu]_{NP} ku .ɬà
 NRL-night RL-middle DIST LOC come.PST
 ‘[They] came in the middle of the night.’
- (6.6) *ajila tʃaŋsə táŋ li.*
 [a-ji-la tʃaŋsə]_{NP} táŋ li
 NRL-dog-F footprint only exist.PST
 ‘Only Dog’s footprints were there.’
- (6.7) *luŋkhuŋlaɪ ki ku mùŋì.*
 [luŋkhuŋ-la-əɪ ki]_{NP} ku mùŋ-ì
 PN-F-ANOM house LOC stay-IRR
 ‘[We] will stay at Longkongla’s house.’
- (6.8) *nə tətʃhapùk mənəmàìù?*
 [nə tə-tʃha-pùk]_{NP} mənəm-àì-ù?
 2SG.POSS RL-wing-STOMACH smell-PRES-DEC
 ‘Your armpits reek.’

6.3.3. Complex head

Complex noun phrase heads can be formed with nouns conjoined by the phrasal conjunction *khə* (see §4.2.11). The conjoined nouns and their modifiers (if present) together form a lower level noun phrase, which in turn functions as the complex head of a higher noun phrase. This hierarchical structure allows the constituent(s) of successively higher noun phrases to exert phrasal scope over constituents of lower phrasal nodes. To illustrate, in (6.9) *nət* ‘two’ is a constituent of NP', thus its quantifier scope is limited to the paratactic nouns *am* and *aku?*, conjoined by the phrasal conjunction at a lower level in the hierarchy. The quantifier *təlu?* ‘all’ and the anaphoric nominal demonstrative *sə*, however, are constituents of a higher noun phrase (NP). They therefore quantify and determine any constituents dominated by the node of that higher noun phrase.

- (6.9) *am khə aku? nət təlu? sə ləpmaɪ, athi tʃu nə zək.*
 [[[**am** **khə** **aku?**]_{NP''} nət]_{NP'} təlu? sə]_{NP} ləp-mà-əɪ
 aspидistra.sp. CONJ gourd.sp. two all ANAPH cut-SPLIT-SEQ
 a-thi tʃu nə zək
 NRL-field.hut DIST ALL send.PST
 ‘Having split open all of the [magical] aspидistra fruit and bottle gourds,
 [she] sent [the paddy that poured out of them] to the field hut.’

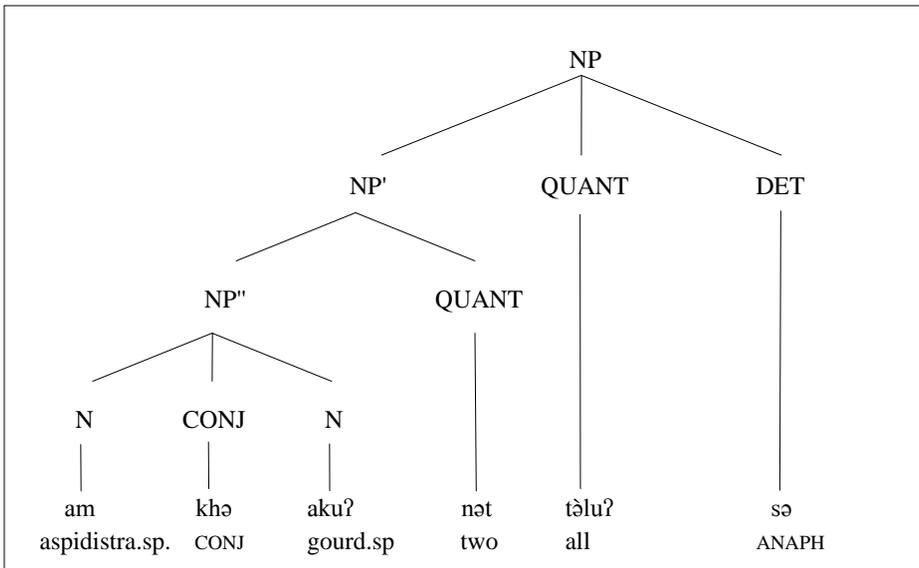


Figure 6.1. Hierarchical structure of the complex noun phrase of (6.9)

The structural relationship holding between the lower and higher noun phrases of (6.9) is represented schematically by the tree diagram of Figure 6.1 above. The tree diagram is useful for showing how two quantifiers can occur in linear order if they belong to different phrasal constituents within a complex noun phrase.

6.4. Noun phrase modifiers

Noun phrase modifiers include nominal attributes, quantifiers, determiners, and by association, case marking clitics and discourse particles, although the latter are not noun phrase constituents *per se*. Quantifiers and determiners have already been described in §4.2.7 in the discussion of word classes. Case marking clitics have been dealt with in the preceding chapter, therefore I will concentrate on describing the functions of nominal attributes in this section and then more briefly discuss other potential modifiers of the noun phrase.

6.4.1. Attributes

Nominal attributes modify their noun phrase heads. There are two types: nominalized attributes that function as relative clauses (see §6.6), and non-relative nominalizations that function as deverbal adjectives. Both types of attributes are derived from verb roots, but via different nominalizing morphology. The type of nominalizer(s) used to derive a nominal attribute will in turn determine how it can be used, as well as its possible position(s) of occurrence in the noun phrase.

A nominal attribute derived by the general nominalizer *-pàʔ* functions as a relative clause. Many relativized attributes can occur either before or after the noun phrase head, with the position of occurrence potentially encoding a subtle difference in meaning. In some cases, such as in the examples of (6.10) below, the meaning of a pre-head relative clause does not differ appreciably from that of a post-head relative clause. In others, the pre-head position can be indicative of restrictive reference, while the post-head position can be indicative of non-restrictive reference. These semantic differences are discussed in §6.6.1 and illustrated with examples.

- (6.10) a. **maŋu** hniŋ-pàʔ tʃu b. hniŋ-pàʔ **maŋu** tʃu
 banana be.ripe-NR DIST be.ripe-NR banana DIST
 ‘that ripe banana’, or ‘that ripe banana’, or
 ‘the banana that is/was ripe’ ‘the banana that is/was ripe’

A further possibility is for a relativized attribute to be derived by both the nominalizing prefix and the general nominalizing suffix, e.g. *maŋu tə-hniŋ-pà?* ~ *tə-hniŋ-pà?* *maŋu* ‘the banana that is/was ripe’. This option most commonly applies to stative intransitive verb stems. According to the intuitions of my consultants, an intransitive verb stem that is nominalized by both of the nominalizing affixes expresses something more akin to a state of existence, whereas those nominalized by the general nominalizing suffix only are more closely aligned to a transitional process or event. Stative intransitive verb stems obligatorily require both derivational affixes when they are used as relativized attributes expressing property concepts. The verb root *hniŋ* of the examples above presumably permits either of these derivational possibilities, because it can also be understood to express a process, e.g. *maŋu tʃu hniŋ* ‘the banana ripened’.

In contrast to the pre- and post-head distributions of relativized attributes, a deverbal adjective derived by the nominalizing prefix *tə-* alone can only occur in the post-head attribute slot. Also, whereas a pre-head relativized attribute derived by *-pà?* can contribute a restrictive meaning to its noun phrase, the meaning of a deverbal adjective is never restrictive, but only descriptive. This constraint correlates to some extent with the non-restrictive meanings of post-head relative clauses.

- (6.11) a. **maŋu** tə-hniŋ tʃu b. *tə-hniŋ **maŋu** tʃu
 banana NZP-be.ripe DIST
 ‘that ripe banana’

A deverbal adjective derived from an intransitive verb by *tə-* can function as a verbless clause complement, but a nominal modifier derived by either the general nominalizer suffix alone or a combination of both the nominalizing prefix and the general nominalizer cannot. Verbless clause topics (VCT) and verbless clause complements (VCC) are identified by square brackets and subscripts in the following pair of examples. The use of deverbal adjectives as verbless clause complements in ascriptive clauses is described in considerable detail in §9.1.4.

- (6.12) a. [**maŋu** i]_{VCT} [tə-hniŋ]_{VCC}
 banana PROX NZP-be.ripe
 ‘This banana is ripe.’
 b. * [**maŋu** i]_{VCT} [(tə-)hniŋ-pà?]_{VCC}
 banana PROX (NZP-)be.ripe-NR

It is very rare to find more than one attribute modifying the same noun phrase head; only a few examples of this occur in the entire corpus of naturally spoken data. Example (6.13) demonstrates the maximum number of nominal modifiers found in one noun phrase. Some of these attributive modifiers are lexicalized elaborate expressions comprising pairs of words denoting similar meanings. The root *mətət*, for example, means ‘know (something or somebody)’ and the root *si* means ‘know (how to do something)’. Their collective meaning is ‘wise’. The paired *təkhàɪ tətəʔ*, derived from the verb roots *khà* ‘possess’ and *təʔ* ‘be strong’, is a lexicalized expression used attributively here to mean ‘wealthy’.

- (6.13) *ami? ta.u təsi təmətət təkhàɪ tətəʔ*
 [a-mi? tə-a.u tə-si-tə-mətət tə-khà-əɪ
 NRL-person NZP-be.good NZP-know-NZP-know NZP-have-ANOM
 tə-tsəʔ]_{NP}
 NZP-be.strong
 ‘[a] good, wise, wealthy person’

Sometimes a nominalized constituent that functions as a modifying attribute is found with its own case marker and determiner after a case-marked head, as in (6.14); this gives the modifying attribute the formal structure of a headless noun phrase in apposition to a headed noun phrase. Substantial pauses (represented by commas in the italicized line) prosodically delineate the phrasal boundaries of each appositive noun phrase. The independent agentive case marking and prosodic independence suggests that the second and third noun phrases of this example are autonomous phrasal constituents of the clause. Semantically, however, the last noun phrase *tə-pə̀ti-‘-pàʔ tʃu nə* describes a qualifying attribute of the referent *wàzàʔ* in the preceding noun phrase. This headless relativization is therefore like a pragmatically-motivated elaboration, as if the speaker is belatedly adding attributes in order to pinpoint reference to the overtly mentioned head in the preceding noun phrase.

- (6.14) *ípáʔ muŋthaŋ tʃu ku, wàzàʔ tʃu nə, təpə̀típàla tʃu nə, ...*
 ípáʔ muŋthaŋ tʃu ku
 EMPHAT meeting DIST LOC
 [wàzàʔ tʃu]_{NP} nə [tə-pə̀ti-‘-pàʔ la tʃu]_{NP} nə
 bird DIST AGT NZP-big-SUP-NR TOP DIST AGT
 ‘In that meeting, that bird, the one that was the leader ...’

This structure could be a strategy to restrict the reference of the initial noun phrase by modifying it using periphrastic means, or to reduce its syntactic

heaviness. Or it may represent a strategy to minimize the cognitive processing demands of interlocutors by limiting the semantic bulk of each noun phrase. Regardless of the purpose, the noun phrases in apposition in (6.14) could alternatively share just one determiner and a case marker. This condensed version would express a similar, if not identical, meaning, e.g. *wàzà? tə-pə̀ti- ˈpà? la tʃu nə* ‘the bird that was the leader [= the biggest]’.

6.4.2. Quantifiers

A noun phrase quantifier can be a lexical word expressing an indefinite quantity (§4.2.7.2), a cardinal numeral expressing a definite quantity (§4.2.7.3), or a nominal derivation of a cardinal numeral (§7.6). The quantifier *tija?* ‘some’ is distinguished by a restriction that limits it to quantifying nouns with human referents, whereas the semantically equivalent *ítʃá?* ‘some’ can be used to quantify all semantic classes of referents. As might be expected, cardinal numerals can be used to express the definite quantity of any semantic class of count noun.

(6.15) *ami? tija?*
 [a-mi? tija?]_{NP}
 NRL-person some
 ‘some people’

(6.16) *phusù? ítʃá? li.ɪə,*
 [phusù? ítʃá?]_{NP} li-ə.ɪ
 century some be-SEQ
 ‘... it being some centuries after, ...’

(6.17) *wàzà? təluk anət*
 [wàzà? təluk a-nət]_{NP}
 bird group NRL-two
 ‘two flocks of birds’

A noun phrase can be headless and yet still occur with a quantifier or other noun phrase modifiers, as demonstrated by the restrictively-focused, headless noun phrase of (6.18) below. This constituent functions as a preposed verbless clause complement of a verbless clause. The understood head of the noun phrase is retrievable from the context, because at the time the speaker was discussing with other Mongsen speakers an appropriate story to narrate.

(6.18) *tətshə à táŋ àj tə, atfu la.*

[tə-tshə à]_{NP} táŋ àj tə [a-tfu]_{NP} la
 NZP-be.short one just PTCL PTCL NRL-DIST TOP

‘[It’s] just a short [story], that.’ (said in response to a suggestion for a story for the speaker to tell)

6.4.3. Determiners

Types of determiners have been discussed at length in §4.2.7.1, and the use of demonstrative pronouns as noun phrase heads was described in §4.2.3. This section therefore concentrates on showing how the pre-head emphatic demonstrative *ípáʔ* is used in noun phrases, either in combination with post-head spatial deictics, or by itself as the sole determiner of the noun phrase.

The common pattern is for the emphatic demonstrative to occur with a post-head nominal demonstrative. Recall that the emphatic demonstrative cannot be used for spatial deixis, which consequently limits it to a determining function when it is used in isolation.

(6.19) *təɪ, ípáʔ mətəm tʃu likə tʃhàwʔ.*

tə-əɪ [ípáʔ **mətəm** tʃu]_{NP} li-əkə tʃhà-ùʔ
 thus-SEQ EMPHAT manner DIST stay-SIM COP.PST-DEC

‘And, [they] lived in that manner.’

Noun phrases determined by the emphatic demonstrative and a post-head nominal demonstrative are sometimes headless in texts. These headless noun phrases are frequently marked by the topic particle and are often followed by a noun phrase containing the omitted constituent of the preceding noun phrase. The following sentence is illustrative of this structural variation. The reference of the headless noun phrase appears to be clarified by the second overtly headed noun phrase, as if the speaker is belatedly making explicit the referent of the omitted head in a recapitulation. In (6.14) above we saw a similar strategy for narrowing reference to the head of a noun phrase by apposing a relativized attribute in what is syntactically a separate noun phrase (as suggested by the independent case marking and the occurrence of a prosodic pause between the two nominal constituents).

(6.20) *ípáʔ la tʃu thəmzəŋ à làtʃhət.*

[ípáʔ la tʃu]_{NP} [thəmzəŋ à]_{NP} làtʃhət
 EMPHAT TOP DIST law one proclaim.PST

‘And that one, a law, was proclaimed.’

In certain contexts, the emphatic demonstrative can be used with just a post-head deictic. When functioning as a verbless clause topic and determined by a nominal demonstrative, it expresses a meaning like a headless nominalization, i.e. ‘this/that one’. In §4.2.7.1 I suggested that it could originate historically from a nominalization, based on its segmental similarity to the general nominalizer *-pàʔ*, its atypical pre-head position of occurrence, and its meaning in sentences such as (6.21).

- (6.21) *ípaj kəinthaŋ.*
 [ípáʔ i]_{NP} [kə inthaŋ]_{NP}
 EMPHAT PROX 1SG.POSS thing
 ‘This one is mine.’

Very occasionally, the emphatic demonstrative is used without a post-head deictic. This only seems possible when there is no requirement for spatial deixis to be encoded in the noun phrase.

- (6.22) *ípáʔ hmapaŋ ku kijupàʔ .ià.*
 [ípáʔ hmapaŋ]_{NP} ku kiju-pàʔ ià
 EMPHAT time LOC owner-M come.PST
 ‘At that time, the owner came.’

6.4.4. Case marking clitics

Case markers demonstrate the distribution of clitics in following the final constituent of the noun phrase, unless the whole of the noun phrase is restrictively focused by a discourse focus particle or topicalized by the topic particle. Recall from the discussion of §5.2.2 that a core argument in O function receives no overt case marking, while the presence or absence of agentive case marking on an A argument is determined by pragmatic and/or semantic factors. In contrast, an oblique argument must be obligatorily specified for a case relation, as determined by the semantic role of its referent in the clause.

- (6.23) *tsəŋh.ɬəm nə álíma ku səŋtuŋlitùŋ sàjaməŋ.ɬià təluʔ zəŋlu.*
 [tsəŋh.ɬəm]_{NP} nə [a-límá]_{NP} ku [səŋtuŋ-litùŋ sàja-məŋ.ɬià
 deity AGT NRL-world LOC tree-ECHO animal-ECHO
 təluʔ]_{NP} zəŋlu
 all make.PST
 ‘God made all the plants and animals and the like in the world.’

There can only be one case marker per noun phrase, although it is potentially possible for multiple embedded noun phrases to each have their own case marking. This is apparent when a case-marked argument of a matrix clause forms the head of an embedded relative clause, which in turn contains case-marked noun phrases. The following example demonstrates the individual case marking of noun phrase constituents at different levels of the phrasal hierarchy. Firstly, *hmapaŋ* forms the head of an oblique (locative) noun phrase in the matrix clause and is simultaneously the shared argument of a relative clause. Within the embedded relative clause, *á-límá* functions as a preposed O argument that receives no overt case marking, while *tə-hmila* is marked for agentive case. A noun phrase is thus case marked according to its semantic role or grammatical function at its respective level of constituency.

- (6.24) *álima təhmila nə nəmphàŋpà? hmapaŋ ku,*
 [[a-límá]_{NP} [tə-hmila]_{NP} nə nəm-phaŋ-pà? hmapaŋ]_{NP} ku
 NRL-world RL-shadow AGT compress-COVER-NR time LOC
 ‘At the time that the world was covered by a shadow, ...’ (speaking of a solar eclipse)

6.4.5. Topic particle and discourse focus particles

The positions within the noun phrase in which the topic particle and the discourse focus particles can occur have been discussed in §6.1.1. Examples demonstrating these positions of occurrence are provided below, starting with the discourse focus particles. A comprehensive description of the functions of the topic particle can be found in §4.2.14.3. Additional examples demonstrating the function of discourse focus particles are provided in §4.2.14.4.

The topic particle most commonly occurs at the end of the noun phrase and, more rarely, immediately after the head, a post-head attribute or a quantifier. Its major function is the marking of contrastive reference.

- (6.25) *akhi ku la, təruk ni mùŋja liŋhà.*
 [akhi]_{NP} ku la təruk ni mùŋ-ja li-tʃhà
 ancient.time LOC TOP six day observe-CONT stay-COP.PST
 ‘In ancient times, [the festival] was observed for six days.’

The most common position of occurrence for the discourse focus particle *táj* is immediately following the head and before a case marker. This restricts the reference of the preceding constituent(s).

- (6.26) *nì ípá? tʃákà táŋ ku làʔ-ì-ù?*
 nì [ípá? tʃákà]_{NP} táŋ ku làʔ-ì-ù?
 1SG EMPHAT place just LOC stay-IRR-DEC
 ‘I’ll stay in just this place.’

Even headless noun phrases can be restrictively focused, because their elided heads are recoverable from context. The position of *táŋ* immediately after the emphatic demonstrative and before a determiner in the following noun phrase suggests that it is restricting the reference of the omitted head. Sometimes the semantic entailment of a noun phrase’s case marker helps to resolve the reference of the elided head, if this is not immediately obvious from the context.

- (6.27) ... *ípá? táŋ tʃu ku ...*
 [ípá? táŋ tʃu]_{NP} ku
 EMPHAT just DIST LOC
 ‘... at just that particular [location] ...’

The next most common position of occurrence for *táŋ* is immediately after an attribute, as in (6.28), or after a quantifier, as in (6.29).

- (6.28) ... *mùŋsənəi təmətʃa táŋ nə ...*
 [mùŋsən-əi təmətʃa]_{NP} táŋ nə
 Mongsen-ANOM NZP-be.pure just INST
 ‘... by just pure Mongsen people ...’

- (6.29) *inət khələm táŋ nə ləʒùə.*
 [inət khələm]_{NP} táŋ nə li-ì-ù? ɹə
 1PL.INC together just AGT stay-IRR-DEC PTCL
 ‘Just we two [clans] will live together.’

It is extremely rare for *táŋ* to occur at the very end of the noun phrase if there is an overt phrase-final case marker. However, example (6.30) demonstrates that it is indeed possible for the discourse focus particle to restrict the reference of an entire noun phrase, even when a case-marker is present. This example also demonstrates the close association that case markers can have with their noun phrases, despite not actually being noun phrase constituents synchronically, at least in most instances. The distal demonstrative *tʃu* occurring inside the ablative marker in *phi <tʃu> nə* forms one phonological word with the other formatives, therefore here we would have to assume that those formatives expressing the ablative case relation are indeed constituents of the noun phrase. An historical explanation for how such an anomaly could have developed is offered in §5.3.2.2.

- (6.30) *ípaʔ phitʃunə áŋ maŋmətuŋ tʃu “maŋmətuŋ” tə tʃatʃhəttʃakiʔ.*
 [ípaʔ phi <tʃu >nə]_{NP} táŋ maŋmətuŋ tʃu maŋmətuŋ
 EMPHAT <DIST>ABL only village.name DIST village.name
 tə tʃa-tʃhət-tʃak-iʔ
 thus call-ABIL-RS-CAUS.PST
 ‘Only from that [time onwards] was Mangmetong village able to be called “Mangmetong”.’

As mentioned earlier, the additive focus particle *kàʔ* is restricted to occurring at the end of the noun phrase.

- (6.31) *... ahən təjala nə kàʔ ahizala tʃu asi,*
 [a-hən tə-ja-la]_{NP} nə kàʔ a-hiʔ-za-la tʃu
 NRL-chicken RL-mother-F AGT also NRL-rat-DIM-F DIST
 asiʔ-əɪ
 deceive-SEQ
 ‘... Mother Hen also tricking Rat Pup, ...’

6.5. Recursively embedded noun phrases

We have already observed in §6.3.3 that noun phrases can be complex, consisting of recursively embedded constituents in dependent relationships to their successively higher phrases. The types of constituents that form complex noun phrases are conjoined nouns, ternary compounds, nominalizations and juxtaposed nouns in multi-layered modifying or possessive relationships. Examples of complex noun phrases follow. Square brackets and subscript labels are again used to distinguish the constituents of embedded and superordinate noun phrases, and heads are distinguished by bold face.

- (6.32) *lazatila təpaʔ khə tija nət tʃu nə məməluŋla tʃhàwʔ*
 [[[lazatì-la]_{NP} **tə-paʔ** **khə** **tə-ja**]_{NP} nət]_{NP} nə
 unmarried.woman-F RL-father CONJ RL-mother two AGT
 mə-məluŋ-la tʃhà-ùʔ
 NEG-be.willing-NEG.CV COP.PST-DEC
 ‘The woman’s father and mother were not willing [to let her marry].’

- (6.33) *... ajim ku lii nə ...*
 [[a-jim]_{NP} ku **li-əɪ**]_{NP} nə
 NRL-village LOC stay-ANOM ALL
 ‘... to the villagers ...’ (lit: to the village-in-dwellers)

- (6.34) *ahizala khə ahənzala nət hjutsə sà-rù.*
 [[[a-hiʔ-za-la khə ahən-za-la]_{NP} nət]_{NP} hjutsə]_{NP} sà-rù
 NRL-rat-DIM-F CONJ chicken-DIM-F two story say-IMM
 ‘[I] will tell the story of Rat Pup and Little Chick.’

6.6. Relativization and nominalization

Nominalization is a derivational process whereby a word or a whole clause, including its syntactic arguments and adjuncts, becomes a noun or a noun phrase. Relativization in Mongsen uses the same affixal morphology as nominalization, but differs in that the resulting derivation functions as an attributive modifier of a noun phrase.

A relative clause can generally be distinguished from a nominalization by the following criteria. Firstly, a relative clause is a dependent modifier of a noun and not an independent matrix clause argument. Secondly, an externally headed relative clause contains a gap corresponding to the head of the overtly mentioned shared argument in the higher clause (this is implicitly true even of headless relative clauses). Thirdly, a nominalized verb stem functioning as a relative clause can often occur in either the pre- or post-head attribute position identified in the template of (6.1), yet still expresses a meaning consistent with relativization.³ In contrast, a nominalization always functions as an embedded argument of the matrix clause and thus does not contain an identifiable gap. Furthermore, rearranging the internal constituent order of a nominalized phrase or clause functioning as an embedded matrix clause argument either results in a change of meaning to one of relativization (e.g., cf. examples [6.73] and [6.74]), or produces an ungrammatical construction.

This last criterion can be useful for distinguishing a relativizing function from a nominalizing one when a relative clause has an internal head, since the head fills the gap corresponding to the shared argument inside the relative clause. In this situation, however, a gap corresponding to the shared argument must then occur in the matrix clause. This offers the means of distinguishing between a nominalization and a headless relative clause. There is always the possibility of restoring an omitted head to either the matrix clause if the head is external, or to the relative clause if it is internal. This is not possible with a

3. That is, a relative clause may have a post-head position $N_{\text{HEAD}}[V_{\text{STEM-NR}}]_{\text{RC}}$, or alternatively, a pre-head position $[V_{\text{STEM-NR}}]_{\text{RC}} N_{\text{HEAD}}$. The only consistent exception to the possibility of a relativized noun phrase having this structural option is found with relative clauses whose locative noun phrase heads express the semantic category of time; see §6.6.1 for examples and discussion.

nominalization functioning as a matrix clause argument, because there is no gapped position corresponding to a shared argument.

Semantics also helps to distinguish the two different functions. The common function of a relative clause in Mongsen is to constrain reference to an entity or entities, e.g. *a-tʃak tʃhuwa-pàʔ tʃu* ‘the paddy that sprouted’, by excluding a subset of related entities (e.g. ‘the paddy that didn’t sprout’), whereas such narrowed reference to a head is never inherently encoded by nominalization (for an appreciation of this difference, compare the nominalization example of [6.78] in §6.6.7). Nominalization and relativization admittedly have certain similarities, but these are directly attributable to the common diachronic origins of nominalizing and relativizing morphology and a shared grammaticalization pathway.

Relativization has been described as just one aspect of nominalization in a number of Tibeto-Burman languages (Matisoff 1972, DeLancey 1986, Noonan 1997). Supporters of this view assume that a relativization is simply an appositive noun phrase, and that the relativization function has developed out of an older nominalization function. LaPolla (2003: 223–227) infers an alternative chronology and supports this with evidence from Qiang (Tibeto-Burman: Qiangic). The forms of the nominalizing clitics in this language suggest that nominalization must have originated from a construction consisting of a (relative) clause or simply a verb that modified a head noun with a very general meaning.

Evidence for this origin is found repeatedly throughout the Tibeto-Burman family, as demonstrated by the ubiquity of nominalizing morphemes with lexical sources in the Proto-Tibeto-Burman roots **mi* ‘man’ or **pa* ‘father’. According to this interpretation, a verb-noun relative construction with a nominal head such as ‘man’ initially expresses a meaning of ‘the man who VERBS’ or ‘the VERB-ing man’. This progresses to a more generic meaning of ‘the one who VERBS’ as the specific semantic entailments of the head are lost and it begins to be used as a nominalizing morpheme. Upon reaching this stage, the now-nominalized verb stem can be used alone as an independent noun meaning ‘the VERB-er’ or ‘the VERB one’. Alternatively, a noun that functions as a notional argument of the nominalized verb stem can occur to its right in apposition, e.g. [VERB-NR] NOUN_{HEAD}. This apposed noun serves as both the head of a complex noun phrase and a notional argument of the nominalized verb stem, which now assumes the modifying functions of a pre-head relative clause. Nominalized verb-noun structures are an attested form of compound noun in Mongsen (see §7.1 for examples), therefore this structure could feasibly provide a foundation for the grammaticalization of a relativizing/nominalizing function out of an earlier compounding process.

Mongsen employs a rather extensive range of morphology for nominalization and relativization:

- (i) general nominalizer *-pàʔ*
- (ii) MAN nominalizer *-tʃhaŋ*
- (iii) agentive nominalizer *-əɪ*
- (iv) locative nominalizer *-tʃən*
- (v) purposive nominalizer *-ɹu*
- (vi) temporal nominalizer *tə...-ì ku*

Suffixes (i–iv) are used for both relativization and nominalization. The general nominalizer *-pàʔ* of (i) is the most textually frequent of all nominalizing morphology and has cognates in many Tibeto-Burman languages. The nominalizers of (ii–vi) have specialized functions that limit their range of application. Although the agentive nominalizer *-əɪ* and the locative nominalizer *-tʃən* can also be used to form relative clauses, they are rarely employed for this purpose. It is more common for them to derive one-word nominalizations that function as nouns. Similarly, the nominalizer *-tʃhaŋ* (-MAN), which bears an obvious etymological relationship to the noun *apaŋtʃhaŋəɪ* ‘man’, is used rather infrequently for any kind of nominal derivation or relativization. The morphology of nominalization is described and further illustrated with examples in §7.4.

In the following examples of this section, embedded relative clauses are identified by square brackets, shared arguments are distinguished by bold face, and matrix clause noun phrases relevant to the description are identified by square brackets and subscripts.

6.6.1. Headed relative clauses

The general nominalizer *-pàʔ* provides the principal means of nominalizing the verb stems of relative clauses in Mongsen. In §6.4.1 we saw that relative clauses functioning as noun phrase modifiers occur either before the head in the D₁ slot, or after the head in the D₂ slot. The elicited examples of (6.35a–b) below demonstrate the pre- and post-head positions of relative clauses in an identical verbless clause.

- (6.35) a. *aji tʃəmja lipaʔ amiʔ tʃu kənùʔ.*
 [[a-ji tʃəm-ja li-pàʔ] **a-miʔ** tʃu]_{NP}
 NRL-rice.beer drink-CONT be-NR NRL-person DIST
 kə-nu-ùʔ
 1SG.POSS-younger.sibling-DEC
 ‘The guy drinking the rice beer is my younger brother.’

b. *ami? aji tʃəmja lipa? tʃu kənù?*

[a-mi? [a-ji tʃəm-ja li-pà?] tʃu]_{NP} kə-nu-ù?

Idem.

The position of the relative clause with respect to the head can sometimes distinguish a restrictive versus a non-restrictive meaning, provided that the relative clause head permits this possibility by virtue of it not having unique reference. The clearest evidence for this comes from the textual example of a post-head relative clause in (6.36a), and the semantic effect that the elicited modification of (6.36b) exerts upon the original meaning.

(6.36) a. *təɪ, tsəŋi .iàpà? táŋ tʃu,*

tə-əɪ [tsəŋi [.ià-pà?] táŋ tʃu]_{NP}

thus-SEQ rain come-NR only DIST

‘And, only the rain that comes,’

b. ?? tə-əɪ [[.ià-pà?] tsəŋi táŋ tʃu]_{NP}

This example comes from a folkloric text explaining why a species of bird only drinks rainwater that has collected on leaves and other *ad hoc* receptacles in the jungle.⁴ When I was exploring the possible structures of relative clauses by attempting to rearrange the constituents of relativized noun phrases in narrative texts, a Mongsen consultant explained that shifting the relative clause to the pre-head position in (6.36b) produced a semantically incongruous result. According to her world view, rain is universal; i.e. there cannot logically be different sub-sets of rain. But the structural modification of (6.36b), with the pre-head relative clause, does in fact suggest that there can be different sub-sets of rain, and that one of these sub-sets is being specifically singled out on this occasion. This explanation seems to be consistent with the notions of restrictive and non-restrictive reference. All other things being equal, the attendant change in meaning – from non-restrictive reference in (6.36a) to restrictive reference in (6.36b) – must be attributable to the change in noun phrase structure.

Pre-head and post-head relative clauses formed with *-pà?* occur in roughly equal numbers in the corpus of naturally spoken data (thirty percent and forty percent respectively; apart from a small number of internally-headed relative clauses, the remainder are headless). Judging from consultants’ responses, in many cases it seems acceptable to use a pre-head relative clause in place of a post-head relative clause and vice versa in texts, although sometimes semantic constraints in a particular context prohibit the alternative structure, as we have just seen in (6.36b) above. The presence of both pre-head and post-head relative

4. For the full context in which this relative clause was uttered, see line 12 of Text 1.

clause structures is not unique to Mongsen and has also been reported in Tibetan and Lahu. DeLancey (1999: 243–244) observes a similar restrictive/non-restrictive difference between pre- and post-head relative clauses in Tibetan texts; Matisoff (1972: 253, 1973: 473–474) notes the possibility of shifting the relative clause to the post-head position in Lahu and states that this results in little or no change in meaning.

The only relative clauses that seem to have a single position of occurrence in Mongsen texts are those with locative-marked noun phrase heads expressing the semantic category of time. Every relativized textual example of *hmapaŋ* ‘time’, for example, has its relative clause in the pre-head position. This is most plausibly explained by the fact that *hmapaŋ* can only have restrictive reference when it is used as a relative clause head. The pattern of restricting relativized attributes to the pre-head slot does not appear to be generally applicable to all relative clauses with locative-marked heads, because post-head relative clauses are also found in narrative texts with locative-marked heads belonging to other semantic classes of noun.

- (6.37) *áhlù jìmpà? hmapaŋ ku,*
 [[a-hlú jìm-pà?] **hmapaŋ**]_{NP} ku
 NRL-field cultivate-NR time LOC
 ‘At the time of cultivating the field, ...’

- (6.38) *tʃùŋlìjìmti lipà? hmapaŋ ku,*
 [[tʃùŋlìjìmti li-pà?] **hmapaŋ**]_{NP} ku
 village.name be-NR time LOC
 ‘In the time that Chungliyimti village existed, ...’

- (6.39) *sàti tʃhàpà? hmapaŋ ku,*
 [[sàti tʃhà-pà?] **hmapaŋ**]_{NP} ku
 marriage do-NR time LOC
 ‘At the time of marrying, ...’

The relative clause of (6.40) below is noteworthy for demonstrating that a relative clause can be used to modify a pronominal head, and that this even applies to first person referents. The function of the relative clause here seems to be one of elaboration. According to consultants’ intuitions, the head could be transferred to the position immediately following the ablative case marker without disturbing the original meaning, but cannot be shifted to the right of the nominalized verb stem to create a pre-head relative clause. I assume that this is because such a structural change would result in a meaning that incorrectly restricts reference to a sub-set of the referents singled out by the relative clause head – i.e. just some of the people of Mangmetong village – when this

restriction is not implied by the original structure of (6.40). No such constraints apply to the relative clauses of (6.41–6.43) though, all of which permit opposite orders with respect to the positions of their heads, and without incurring any change in meaning within the contexts of their source narrative texts. Note that this even applies to the appositive-like structure of (6.41). Syntactically, the constituent *i-tə ka-pà?* (PROX-thus open.mouth-NR) appears to have the status of an independent noun phrase, but semantically it functions as an attributive modifier of *tə-paŋ* ‘mouth’. It is this dependency that permits the relative clause to be shifted to the pre-head position. It would be stretching the imagination to assume that the first line of interlinearization should be construed as expressing ‘the mouth, the like-this opening one’; we should instead view the apposed relative clause of this sentence as providing a means of elaborating the attributes of its understood head. This function is consistent with the post-head position of both non-restrictive relative clauses and non-relativized attributes.

(6.40) *ila tsəŋlijimti lima tʃu phinə m̀ŋsən-əi likə .jà-pà? ...*

[**ila** [tsəŋlijimti lima tʃu phinə m̀ŋsən-əi
1PL.EXC village.name land DIST ABL Mongsen-ANOM

li-əkə .jà-pà?]]_{NP}

be-SIM come-NR

‘We who have come from Chungliyimti country living as Mongsen people ...’

(6.41) *təpaŋ tʃu itə kapa? təpha məkətla likə tʃhà.*

[[**tə-paŋ** tʃu [i-tə ka-pà?]]_{NP}

RL-mouth DIST PROX-thus open.mouth-NR

tə-pha mə-kət-la li-əkə tʃhà

RL-teeth NEG-have-NEG.CV be-SIM COP.PST

‘[The leopard cat saw that the rooster’s] mouth, which was opening like this (speaker motions with thumb and fingers), did not have teeth.’

(6.42) *məʃatshəŋ nə tsəŋpa? áwk sə ...*

[[məʃatshəŋ nə tsəŋ-pà?] **a-úk** sə]_{NP}

PN AGT spear-NR NRL-pig ANAPH

‘The pig that Mechatseng speared ...’

(6.43) *nì aki ípá? mətəm i tʃhà-tʃhət-pà? ...*

[**nì a-ki** [ípá? mətəm i tʃhà-tʃhət-pà?] i]_{NP}

1SG NRL-house EMPHAT manner PROX make-ABIL-NR PROX

‘My house that I am able to build like this here ...’

Consultants disagreed amongst themselves about the acceptability of alternative orders of heads and their relativized attributes in the textual examples of (6.44) and (6.45) below. The pre-head position of the relative clause in (6.44) is consistent with restrictive reference to the head in its particular context of use, therefore a restrictive meaning could possibly be lost by transferring the relative clause to the post-head position. It was not possible to establish a reason for one consultant rejecting a pre-head position for the relative clause of (6.45). This also seems to encode restrictive reference to a single entity in its textual setting, therefore shifting the relative clause to the pre-head position should result in an acceptable alternative structure (see Text 4 for the context in which this example was used). The parameters responsible for determining the position of the relative clause *vis-à-vis* the head remain a topic in need of further investigation, not only in Mongsen, but also in other Tibeto-Burman languages.

- (6.44) *thani asaŋ kà? ípála akhu hmila həmsipà? mi? tʃu li.*
 thani.asaŋ kà? ípála la
 nowadays still EMPHAT TOP
 [[a-khu-hmila həmsi-pà?] a-mi? tʃu]_{NP} li-əɪ
 NRL-tiger-soul use-NR NRL-person DIST be-PRES
 ‘Even nowadays, people who use the spirit of a tiger exist.’
- (6.45) *təɪ, pùkphula tʃu, wàzà? təpuŋ təɪu à tʃhàpà?*
 tə-əɪ pùkphu-la tʃu
 thus-SEQ owl-F DIST
 [wàzà? [tə-puŋ tə-əɪ à tʃhà-pà?]]_{NP}
 bird NZP-be.beautiful NZP-be.good one COP-NR
 ‘And Owl, the bird that was a beautiful one’,

The attentive reader will have observed that the relative clauses presented here and in preceding chapters have all been externally headed. This appears to be a dominant characteristic of Mongsen relative clauses. While the vast majority of the examples in the corpus are externally headed or headless, internally-headed relative clauses do occasionally crop up in texts and in elicitation. We will first consider two textual examples with internally-headed relative clauses, followed by some elicited examples showing both internally- and externally-headed relative clauses. The first of these, demonstrating a relativized O argument, comes from a Waromung Mongsen text.

(6.46) *nì kà? àtì na asə à ənəi .jàpà? pɯ mùk.*

nì kà? [a-ti na a-sə à ən-əi
 1SG also VOC-elder.sibling AGT NRL-shawl one take-SEQ
 .jà-pà? pɯ]_{NP} mùk
 come-NR PROX wrap.PST
 ‘I also wore this shawl that Elder Sister brought.’

(6.47) *təi, tshəluŋla nə asi?juk-zək-pà? sə ní? la phàia?-ì-ù?*

tə-əi [tshəluŋ-la nə asi?-juk-zək-pà? sə]_{NP}
 thus-SEQ fox-F AGT deceive-PFV-SEND-NR ANAPH
 ní? la phàia?-ì-ù?
 one.day TOP catch-IRR-DEC

In context: ‘And, the fox that deceived [him] one day will be caught.’

*The fox that deceived [him] one day will catch [him].

Nominal case marking sometimes gives an indication as to whether the head of a relative clause is a constituent of the matrix clause or the relative clause. In (6.47), for instance, the noun phrase *tshəluŋ-la nə* (fox-F AGT) might at first seem to be a possible constituent of either clause. If it were analyzed as a constituent of the matrix clause, then the agentive case marking would identify its noun phrase as the A argument of the matrix clause headed by the transitive verb *phàia?-ì-ù?* ‘will catch’. In actual fact, the A argument of *phàia?-ì-ù?* has been deleted under identity with an antecedent mention of its referent and is pragmatically recoverable from the context – Fox is contextually understood to be the entity that will be caught, not the entity that will catch an unexpressed referent. The noun phrase *tshəluŋ-la nə* therefore has the wrong case marking for a matrix clause O argument. Its agentive case marking is, however, entirely consistent for an A argument of the subordinate, nominalized transitive predicate *asi?-juk-zək-pà?* (deceive-PFV-SEND-NR), thus it can only be a notional argument of the embedded relative clause. Out of context, the starred reading of (6.47) could be an acceptable interpretation of this sentence’s meaning and would require the relative clause to be analyzed as having an external head. In its context, however, only the internally-headed analysis is possible. Lastly, it would be mistaken to assume that *tshəluŋ-la nə asi?-juk-zək-pà? sə* should be interpreted as two appositive noun phrases expressing something equivalent to ‘the fox, the deceiver’, because then the agentive case marking is completely unmotivated. Obviously the agentive case marker serves to identify the A argument of *asi?-juk-zək-pà?* in the modifying relative clause.

Changing the position of the head with respect to that of the relative clause potentially has implications for the analysis of internally-headed versus externally-headed relative clauses. The elicited sentences of (6.48) show that if

certain conditions are met, transposing the head from behind the relative clause to in front of it can result in an erstwhile externally-headed clause becoming internally headed.

- (6.48) a. *səntipà? nə hənpa? kákót tʃu kəintaŋù?*
 [[sənti-pà? nə hən-pà?] **kákót** tʃu]_{NP} kə intaŋ-ù?
 PN-M AGT take-NR book DIST 1SG.POSS thing-DEC
 ‘The book that Senti carried is mine.’
- b. *səntipà? nə kákót hənpa? tʃu kəintaŋù?*
 [[sənti-pà? nə **kákót** hən-pà?] tʃu]_{NP} kə intaŋ-ù?
 Idem.
- c. *kákót səntipà? nə hənpa? tʃu kəintaŋù?*
 [**kákót** [sənti-pà? nə hən-pà?] tʃu]_{NP} kə intaŋ-ù?
 Idem.

For this structural variation to be possible, the relativized noun must be a core argument of the relative clause and both A and O arguments must be explicitly mentioned, so this additionally requires the relativized verb stem to be transitive. Core arguments represent the most frequently deleted noun phrases under coreference; perhaps this goes some way towards accounting for the dearth of internally-headed relative clauses in the data. In any event, the internally-headed relative clause is a possible, albeit rarely used, relative clause strategy in Mongsen. Note that the head can also be moved to the front of the embedded clause, as shown in (6.48c). This produces an acceptable sentence with no real change in meaning on this occasion, other than a slightly different pragmatic effect.

In certain contexts, both the nominalizing prefix and the general nominalizer suffix are used on a non-stative intransitive verb stem to derive a relative clause. This is more typical of intransitive stems used to express relativized states of existence (e.g. cf. example [6.50] below), or derived adjectives of degree (see §3.1.3.1 and §7.4.3). When both nominalizing affixes are used to create relative clauses from non-stative intransitive verbs, the resulting derivation encodes an additional shade of meaning.

- (6.49) *tuwapa? tʃaŋsə táŋ li, mijipəkə .jàpà? tʃaŋsə məkhàlà.*
 [[tə-wa-pà?] **tʃaŋsə**]_{NP} táŋ li
 NZP-go-NR footprint just be.PST
 [[mijip-əkə .jà-pà?] **tʃaŋsə**]_{NP} mə-khà-la
 return-SIM come-NR footprint NEG-have-NEG.PST
 ‘There were just footprints that had gone; there were no returning footprints.’

A nominalized non-stative intransitive verb stem functioning as a relative clause normally does not take the nominalizing prefix in addition to the general nominalizer – for instance, compare *ɬà-pàʔ* in this example. A consultant suggested that, if the simultaneous converb construction *mijip-əkə* were omitted from the second clause, then it would be possible to use the nominalizing prefix on *ɬà-pàʔ* as well, but this would not be the case if *mijip-əkə* were not first omitted. The same consultant described *tə-wa-pàʔ* in her own words as meaning “gone, action done in the past”, whereas she didn’t recognize this in *wa-pàʔ* when it was exchanged for the prefixed form in this narrative example. What equates to a perfective nuance is independently confirmed by the explanations of other Mongsen speakers, so it would appear that *tə-...-pàʔ* has additionally grammaticalized an aspectual meaning when used to derive relativized attributes from non-stative intransitive verb stems. The reason why *mijip-əkə* and *tə-ɬà-pàʔ* cannot not be used together in the second relativization of (6.49) is because of a mutual incompatibility in their aspectual categories: a verb stem marked by the simultaneous converb suffix expresses the functional equivalent of imperfective aspect, but in this case it would clash with the perfective sense of *tə-ɬà-pàʔ tʃaŋsə*, literally ‘footprints that had already come’.

In §9.1.4 I suggest that non-relativized attributes take either verbal or derived nominal forms in ascriptive clauses according to whether the descriptive state predicates a transitional or permanent property. The following example from a Waromung Mongsen text suggests that the same dichotomy is morphologically formalized in the derivation of relativized attributes from stative intransitive verb stems.

(6.50) *ajà tɛ̀nəm mi taɽupaʔ sə nə luŋzak áŋ kəmʃukukù.*

ajà [tɛ̀nəm mi [tə-aɽu-pàʔ] si]_{NP} na
EXCLM bird.sp tail NZP-be.good-NR ANAPH AGT

luŋzak áŋ kəm-ʃuk-ukù
whetstone just become-PFV-ANT

“Aya! That beautiful Great Pied Hornbill tail feather has just turned into a whetstone!”

According to Keenan and Comrie (1977: 66), the noun phrase positions that can be relativized form an accessibility hierarchy. This is stated as follows:

(6.51) subject > direct object > indirect object > genitive > object of comparison

This represents a hierarchy of constraints on relativization, such that any language permitting a given position on the hierarchy to be relativized also permits the relativization of any positions to its left in the hierarchy. The hierarchy thus predicts the extent of accessibility to relativization.

Mongsen demonstrates a number of constraints on accessibility to relative clause formation. Consultants balk at attempts to elicit relative clauses of dative arguments or standards of comparison, suggesting that these are not relativizable positions, and they are never encountered in texts. The language thus conforms to the observation of Keenan and Comrie (1977: 74) that few languages which distinguish objects of comparison from direct objects or oblique noun phrases permit them to form relativizations on these positions. The following naturally spoken examples demonstrate the only oblique noun phrase positions that permit accession to relativization. These both have the post-head structure used for the elaboration of attributes, as discussed above.

(6.52) *nì tətʃàsi ku jaŋtən-lí? wapa? tsə̀lala tsə̀tʃhaw?*

[**nì** **tə-tʃàsi** ku [jaŋ-tən-lí? wa-pà?]]_{NP}
 1SG NZP-be.distressed LOC hop-TO.AND.FRO-SIM go-NR

tsə̀la?-la tsə̀-tʃha?-ù?
 earthworm-F stamp.on-CHANCE-DEC

‘‘In my distress in which I went hopping to-and-fro, [I] unknowingly stamped on Earthworm’’ [said Barking Deer].’

(6.53) *nì tətʃàsi nə sùsazək-lí? wapa? nəhlú sùsaw?*

[**nì** **tə-tʃàsi** nə [sùsa-zək-lí? wa-pà?]]_{NP}
 1SG NZP-be.distressed INST scratch-SEND-SIM go-NR

nə-hlú sùsa-ù?
 2SG.POSS-field scratch.PST-DEC

‘‘Your field got damaged by my distress with which I went around scratching everything up’’, [said Red Jungle fowl].’

6.6.2. Headless relative clauses

Relative clauses are often headless, accounting for roughly thirty percent of all relative clauses in the corpus of naturally spoken data. This textual frequency is not unexpected in a language that makes great use of zero anaphora.

A headless relative clause is one in which the shared argument is not overtly stated in the matrix clause. A relative clause head can be deleted under coreference with an antecedent mention of its referent, or if the pragmatic context permits the identity of its referent to be established in the absence of a prior mention. They can be distinguished from nominalizations by the possibility of restoring the omitted head to the matrix clause. Also, like externally-headed relative clauses, the embedded clause contains a gap that corresponds to the implicit shared argument of the matrix clause. The following

text examples of two consecutive clauses demonstrate how the head of a relative clause can be omitted under identity with an antecedent mention.⁵ This complex sentence contains a headed and a headless relative clause. The externally-headed relative clause of (6.54a) has previously been discussed in the context of restrictive and non-restrictive reference in §6.6.1 and is repeated and renumbered here.

- (6.54) a. *təɪ, tsəŋi .ià-pà? táŋ tʃu*
 tə-əɪ [tsəŋi [ià-pà?]] táŋ tʃu]_{NP}
 thus-SEQ rain come-NR only DIST
 ‘And, only the rain_i that comes,’
- b. *pa səŋàwà, aɪa ku kət-pà? tʃəm-lí? tʃəm-lí? wa.ù?*
 pa [səŋàwà aɪa ku kət-pà?]_{NP}
 3SG leaf and.the.like LOC have-NR
 tʃəm-lí? tʃəm-lí? wa-əɪ-ù?
 drink-SIM drink-SIM go-PRES-DEC
 ‘she goes around drinking [Ø_i] that collects on the leaves and things.’
 (i.e. ‘she goes around drinking [just the rain that comes_i] that collects on the leaves and things’)

Firstly, an externally-headed relative clause occurs in (6.54a) and functions as a preposed and topicalized O argument of the matrix clause, while the sentence of (6.54b) provides the comment to the preposed and relativized O topic of (6.54a). The relativized noun phrase *tsəŋi .ià-pà? táŋ tʃu* ‘only the rain that comes’ is the implicit relative clause head of (6.54b). But since the gapped argument of this headless relative clause is understood to be coreferential with the preposed topic noun phrase of (6.54a), that argument need not be explicitly mentioned in the matrix clause of (6.54b), and in fact it would be pragmatically redundant to restore it to that clause.

Like their headed counterparts, only some noun phrase positions are amenable to relativization by a headless relative clause. Overall the most commonly relativized position is that of a core argument in O function, followed by one in S function. The fact that the omitted head could be restored to the headless relative clause of (6.55) proves that this is not a nominalization.

- (6.55) *təɪ zək-pà?, ənti.ɪəm ku, sətʃuk.*
 tə-əɪ [zək-pà?]_{NP} ənti.ɪəm ku sətʃuk
 thus-SEQ send-NR path-middle LOC die-PFV.PST
 ‘And then, [the one who was] sent died in the middle of the road.’

5. To appreciate the full context in which this complex sentence is used, see Text 1.

- (6.56) *mətshəla tʃutə, phukə lipa? tʃu ni məmətətla.*
 [mətshə-la tʃu-tə phu-əkə li-pà? tʃu]_{NP}
 deer.sp-F DIST-thus expose.to.sun-SIM be-NR DIST
 ni məmətət-la
 1SG NEG-know-NEG.PST
 ‘‘I didn’t know about [the ornaments] that Barking Deer had left
 exposed to the sun like that,’’ [said Squirrel].’

When a headless relative clause is formed with a verb of speech, it can represent the implicit summation of an entire conversation. The following narrative example was uttered after a prior warning was given to a character that he should make alterations to his house to avoid future problems with lightning strikes, and the advice was not heeded. The omitted head of the relative clause, the prior warning, is retrievable from the antecedent conversation in the narrative. The initial and final distal deictics are used here to express a mild admonition, equivalent to ‘‘See?’’, and the unusual rising tone that occurs on the pronoun is discussed in §3.1.3.2.

- (6.57) *tʃu nǐ sapa? tʃu athənsijaɪ, tatʃu tʃu.*
 tʃu [nǐ sa-pà? tʃu]_{NP}
 DIST 1SG.AGT say-NR DIST
 athənsi-ja-əɪ tə a-tʃu tʃu
 assemble-CONT-PRES thus NRL-DIST DIST
 ‘‘See? [That which] I said is coming together. See that?’’
 (a violent storm can be seen brewing down on the plain of Assam)

6.6.3. Agentive nominalizations as relative clauses

The derivation of agentive nominals applies to words, phrases and whole clauses. Examples of the agentive nominalizer applying to single words and phrases are given in §7.4.4. Here our attention will be limited to examining how it is used to derive relative clauses. The shared argument of an agentive nominalization functioning as a relative clause must be a human referent in A or S function, and all relative clauses formed with the agentive nominalizer are found in the pre-head position in the limited data set at hand. This suggests that the pre-head position is a structural requirement for this type of relative clause.

- (6.58) *mətʃhənəɪ luŋ sə nə...*
 [[mətʃhən-əɪ] **luŋ** sə]_{NP} nə
 NEG-flee-ANOM group ANAPH AGT
 ‘The group that didn’t flee ...’

Semantic ambiguity between a relative and a genitive interpretation is occasionally encountered in dependent clauses formed with the agentive nominalizer, as demonstrated by the alternative relative and genitive meanings of the dependent clause of (6.59). In certain contexts, only a genitive meaning seems possible for the verb stem of a dependent clause marked by *-əɪ* (see §7.3.3 and §7.4.4 for further discussion and examples of this).

- (6.59) *atʃhən pəɪɪja liɪ mi luŋ nə nakalən tʃhukəɪ wajù?*
 [atʃhən pəɪɪ-a li-əɪ] **mi?** **luŋ**]_{NP} nə
 money be.enough-SIM be-ANOM person group AGT
 nakalən tʃuk-əɪ wa-ì-ù?
 Nagaland leave-SEQ go-IRR-DEC
 (i) ‘The people who have sufficient money will leave Nagaland.’
 (ii) ‘The people of sufficient money will leave Nagaland.’ (Literally:
 ‘the enough-money-having’s group of people...’)

Relative clauses formed with the agentive nominalizer are encountered very infrequently compared to those formed with the general nominalizer, accounting for just five percent or so of the total number of relative clauses in the corpus. When a whole clause is relativized to function syntactically as a derived attribute of a noun phrase, as in (6.59), the agentive nominalizer can be replaced by the general nominalizer *-pà?*. This has no impact on its meaning. Replacement of the agentive nominalizer with *-pà?* in derived agentive nominalizations functioning as noun phrase heads, such as *áwŋ ku li-əɪ* (jungle LOC stay-ANOM) ‘member of the underground, insurgent’ is not possible, however.

6.6.4. Locative nominalizations as relative clauses

The locative nominalizer *-tʃən* tends to be used mostly for the lexical nominalization of individual words. When suffixed to a verb stem, *-tʃən* derives a noun meaning ‘the place where VERB’, e.g. *mən-tʃən* (sit-LNOM) ‘chair’. Additional examples of locative nominalizations resulting in the derivation of individual words are provided in §7.4.5.

Locative nominalization may also apply to a clausal predicate, in which case the resulting nominalization has scope over the constituents of the entire clause. This derives a meaning of ‘the place where CLAUSE’. Example (6.60) shows how the locative nominalization can be used as an attribute of a noun phrase head. Because of the semantic entailments of the locative nominalizer, relativizations are usually formed on the heads of oblique arguments. The few examples of locative nominalizations functioning as relative clauses in the corpus occur in the pre-head position. Attempts to transpose the relative clause to a post-head position were rejected as ungrammatical by consultants.

- (6.60) *ami? sətʃən kɪ nə wauù?*
 [[a-mi? sə-tʃən] **ki**]_{NP} nə wa-əɪ-ù?
 NRL-person die-LNOM house ALL go-PRES-DEC
 ‘[We’re] going to the house where the person died.’

The locative nominalization of entire clauses can create headless relative clauses. Despite the absence of an overtly mentioned head, I nevertheless still consider this a case of relativization and not nominalization, because the omitted head could be restored to a gap between the relativized attribute and the locative case marker in the matrix clause.

- (6.61) *kənət litʃən ku mənəŋ.*
 [kənət li-tʃən]_{NP} ku mən-əŋ
 1DU.EXC stay-LNOM LOC sit-IMP
 ‘Sit [at the place] where we two are sitting.’

6.6.5. Aspect and mood marking in relative clauses

Relative clauses allow the encoding of some grammatical categories, such as mood distinctions, aspect and negation. Mood marking allows a realis ~ irrealis distinction to be made in relative clauses, with realis being the default (unmarked) category. The temporal setting of an imperfective event in a relative clause can additionally be encoded independently of the matrix clause. This is achieved by using a simultaneous converb paired with an existential verb and a copula as an auxiliary to specify aspectual and relative tense distinctions.

The following text example uses a grammaticalized past tense copula form of the verb root *tʃhà* ‘do’ in auxiliary function to carry the general nominalizer. This allows another verb stem to take a simultaneous converb suffix in order to express a simultaneous activity relative to the temporal setting inherently encoded by the past tense copula (see §9.2 for description of copula and existential verbs).

- (6.62) *atʃak mələnəkə tʃhàpàʔ tʃu, haj thakala atʃak i təɹà phùʔ-rù.*
 [a-tʃak [mələn-əkə tʃhà-pàʔ] tʃu]_{NP}
 NRL-paddy transfer-SIM COP.PST-NR DIST
 haj thaka la a-tʃak təɹà phùʔ-rù
 INTJ now+night TOP NRL-paddy some steal-IMM
 ‘That paddy they were transferring – Listen! Tonight I’m going to steal some.’

The post-head relative clause in the textual example of (6.62) was used as a template to elicit other temporal and aspectual distinctions in relative clauses:

- (6.63) [a-tʃak [mələn-əkə li-pàʔ] tʃu]_{NP}
 NRL-paddy transfer-SIM be-NR DIST
 ‘The paddy [they] are transferring’

- (6.64) [a-tʃak [mələn-ì-pàʔ] tʃu]_{NP}
 NRL-paddy transfer-IRR-NR DIST
 ‘The paddy [they] will transfer’

The relative clause of (6.64) demonstrates that the irrealis marker can simply be inserted between the verb root and the general nominalizer to encode unrealized events expressed by the relative clause. This does not require any other verb stem in the relative clause construction. Lastly, a verb root that is directly nominalized by *-pàʔ* can have either a timeless habitual interpretation or a realis interpretation.

- (6.65) [a-tʃak [mələn-pàʔ] tʃu]_{NP}
 NRL-paddy transfer-NR DIST
 ‘The paddy [they] transfer (habitually)’; ‘The paddy [they] transferred’

The next example shows the use of the perfective suffix with the irrealis marker to signal that the event of the second relative clause (going to the village of the dead) is perfective with respect to the event of the first relative clause (coming to heaven). This aspectual clarification is necessary, because the verb stems of both relative clauses are marked for irrealis mood, and the chronological order of their events would otherwise be unclear.

- (6.66) *tuj̄nət tʃu, kúták nə .jàjpàʔ, asəjim nə wajukípàʔ, kútákə nə thəpsi.ùʔ, tə sa.*
 [tuj̄nət tʃu [kúták nə .jà-ì-pàʔ]
 3DU DIST heaven ALL come-IRR-NR
 [a-sə-jim nə wa-tʃuk-ì-pàʔ]]_{NP}
 NRL-die-village ALL go-PFV-IRR-NR
 kúták-əɪ nə thəpsi-əɪ-ùʔ tə sa
 heaven-ANOM AGT discuss-PRES-DEC thus say.PST
 ‘‘The heavenly beings are discussing they two who are to come to
 heaven, [they two] who will have already gone to the village of the
 dead,’’ [he] said.’

The sentence of (6.66) represents rather complex layers of embedding. Firstly, everything preceding the quotative complementizer *tə* and the matrix clause verb *sa* ‘said’ is an O complement of that verb of speech. Within the O complement is an OAV transitive clause. The predicate head is *thəpsi-əɪ-ùʔ* ‘discussing’, *kúták-əɪ nə* is its A argument, and *tuj̄nət tʃu* ‘they two’ is its fronted O argument. This O argument forms the head of the first relative clause *kúták nə .jà-ì-pàʔ*, as well as the following relative clause *a-sə-jim nə wa-juk-ì-pàʔ*. The chronological order of events referred to by the two relative clauses is not reflected by their order, but it is clear to consultants that going to the village of the dead must logically precede coming to heaven. This is reinforced by the perfective aspect suffix on the verb stem of the second relative clause. My impression is that the perfective aspect marking would allow the order of relative clauses to be reversed without disturbing the sentence’s meaning. Discrete pauses occur at the clausal margins and are represented by commas in the italicized line of the example. These prosodic boundaries give the impression that the speaker is carefully clarifying the reference of the noun phrase head. He does this by modifying it with the relative clauses, using them as descriptive attributes.

Other verbal suffixes that are commonly found to occur on nominalized verb stems of relative clauses are the directional suffix *-zək* (§8.5.3), the modality suffixes (§8.5.5), the causative suffix (§8.5.8) and various lexical suffixes (§8.5.1). Any verbal suffixes that are semantically compatible may occur on a verb stem, and this includes the vast majority.

Relative clauses are negated with the negative prefix *mə-* that is used for negating verb stems. The negative past suffix *-la* is never found together with the general nominalizer, as they share the same slot of the predicate template (see §8.5.9). Negated relative clauses are used very infrequently in comparison with relative clauses of positive polarity.

- (6.67) *pa tʃu mənən mə.ɪək məkət-pà?, tʃuŋi.*
 [[mənən-mə.ɪək mə-kət-pà?] **tʃuŋi**]_{NP}
 be.dirty-ECHO NEG-have-NR food
 ‘That which is not taboo – food.’
- (6.68) *tənik nə mùŋpà? lima à ku*
 [[tə-nik nə mə-hùŋ-pà?] **lima** à]_{NP} ku
 RL-eye INST NEG-see-NR world one LOC
 ‘In a world that is not seen by the eyes’

6.6.6. Relative-correlative constructions

A structure that shows great similarity to the relative-correlative constructions of Indo-Aryan languages is sometimes encountered in texts, and also occasionally when eliciting relative clauses from some speakers. This was previously discussed in §4.2.14.3, with respect to the function of the topic marker occurring at the end of a dependent clause. The dependent clause of the relative-correlative construction is identified by square brackets in the following examples. The constituents corresponding to the shared argument of the lower and higher clauses are distinguished by bold face. The shared argument can be overtly expressed in the matrix clause, as in (6.69), or may be omitted if its referent can be recovered from the pragmatic context, as in (6.70).

- (6.69) *səpá? nə kùk.ùla, pa təzəmpa? tʃhàjù?*
 [səpá? nə kùk-ù] la **pa** tə-zəm-pà? tʃhà-ì-ù?
 who AGT win-IMM TOP 3SG NZP-be.senior-NR COP-IRR-DEC
 ‘“Whoever wins will be more senior,” [said Wild Pig to Tiger].’
- (6.70) *səpá? nə áhlú ku hməpa tʃhəniŋla, atʃu asaŋ waɪ, asaŋ áŋ mətət-ì-ù?*
 [səpá? nə a-hlú ku hməpa tʃhəniŋ] la
 who AGT NRL-field LOC work labour.PST TOP
 a-tʃu asaŋ wa-əɪ asaŋ áŋ mətət-ì-ù?
 NRL-DIST tomorrow go-SEQ tomorrow only know-IRR-DEC
 ‘“As for who worked hard in the field, will we know only after going there tomorrow,” [said Lichaba].’

Despite the presence of the interrogative pronoun, these topicalized clauses cannot be analyzed or understood as expressing interrogative mood. The interrogative pronoun is instead used to represent the relativized position in the topic-marked clause. The topic-marked clause in this construction cannot stand

in isolation, even though its predicate has the form of a finite verb; therefore its status must be that of a dependent clause. It also narrows the reference of the head, much like a dependent relative clause formed with a nominalized verb stem. In §4.2.14.3 I mentioned that the dependent ‘relative’ clause functions as a topic, to which the correlative higher clause functions as the comment. The analogous structure of the complex sentence of (6.71) below suggests that it, too, qualifies as a relative-correlative construction. In this case, the function of the dependent clause is to narrow the reference to a particular point in time.

(6.71) *kújím áhlú sə thùŋla, ajila sə ahlu sə ku jiplitʃuk.*

[**kújím** a-hlú sə thùŋ] la
when NRL-field ANAPH reach.PST TOP

a-ji-la sə a-hlú sə ku jipli-tʃuk
NRL-dog-F ANAPH NRL-field ANAPH LOC sleep-PFV.PST

‘When [= at the time that they] reached the field, Dog slept in the field.’

A Hindi relative-correlative construction is included below to allow a cross-linguistic comparison to be made to complex structures common to the two languages. A morpheme *jo* can be used as a relative pronoun in the relative clause and corresponds to the shared argument of the correlative pronoun *vah* in the matrix clause. As in Mongsen, the correlative pronoun of the matrix clause may optionally be omitted if its reference is understood.

(6.72) जो इस सून्सान हवेली में रहेगा, वह इनाम पाएगा ।⁶

[**jo** is sūnsān hawelī mē rahegā]
who this deserted house LOC stay.FUT.M.SG

vah inām pāegā
3SG reward receive.FUT.M.SG

‘Whoever will stay in this deserted mansion will get a reward.’

(Barz and Yadava 1993: 210)

The relative-correlative construction of Mongsen is not necessarily more effective than the gapping strategy for relativizing on a core argument of the clause. The structure may have been borrowed into the language to allow for the relativization of oblique arguments, and then became an alternative strategy for deriving relativized attributes of all clausal arguments, including core arguments. I noted earlier that some speakers use it more frequently than others, particularly those who speak English fluently. All of the Mongsen speakers who

6. The Roman transcription of this Hindi example conforms to the orthographic conventions of the South Asian linguistic tradition.

used the relative-correlative construction were also bilingual speakers of Nagamese, the Assamese-based lingua franca of Nagaland. Coupe (2007: 355–358) recognizes a structurally identical relative-correlative construction in this language. The inclusion of a topic particle to mark the dependent clause is unique to Mongsen and Nagamese, and also seems to be particularly diagnostic of linguistic convergence as an explanation for the presence of the relative-correlative structure in the grammar of Mongsen.

6.6.7. Nominalization

Nominalization is a morphological process applying to words, phrases and clauses in Mongsen. The formation of relative clauses by the agentive nominalizer and the locative nominalizer was discussed in §6.6.3 and §6.6.4 respectively. We now look at how the general nominalizer *-pàʔ* can be used to turn whole clauses into noun phrases. Nominalized clauses have a significantly lower textual frequency in comparison to relative clauses, accounting for just five percent or so of verb stems marked by *-pàʔ* (excluding personal names).

A simple form of nominalization derives the head of a noun phrase, as in the following example. This was uttered as the introductory statement of a narrative. In contrast to the relative clause examples presented in §6.6.1 and §6.6.2 above, the derived nominal *zàŋlu-pàʔ* does not form a syntactic constituent of a higher clause, but functions together with its notional argument *aji* as a matrix clause argument.

(6.73) *aji zàŋlupaʔ pà.*

[a-ji zàŋlu-pàʔ]_{NP} pà
 NRL-rice.beer make-NR QPTCL
 ‘[You asked me about] rice beer making?’

Whereas a nominalized verb stem functioning as a relativized attribute can often either precede or follow its head, a nominalization functioning as a matrix clause noun phrase argument does not tolerate a rearrangement of its constituents without incurring a change of meaning consistent with relativization. For example, the nominalized verb stem of (6.73) comes to express a meaning consistent with that of a relativized attribute if its notional argument is moved to the end of the noun phrase, as demonstrated by (6.74).

(6.74) *zàŋlupaʔ aji pà.*

[[zàŋlu-pàʔ] a-ji]_{NP} pà
 make-NR NRL-rice.beer QPTCL
 ‘[Is this the] rice beer that [you] made?’

A nominalization can function as the possessor or the possessee of a complex noun phrase, as in (6.75) and (6.76) respectively (the verbless clause of [6.76] was used to introduce a Waromung Mongsen text). In common with underived nouns, a nominalization is able to co-occur with all manner of nominal adjuncts in the noun phrases of which they form heads.

(6.75) *mùŋsənəɪ təkħələmpa?*

[mùŋsən-əɪ təkħələm-pà?]NP
Mongsen-ANOM NZP-worship-NR
'Mongsen person's [manner of] worship'

(6.76) *ípa? jutʃi lùŋkhəm kha nukɪaŋ jim nət uŋtəpa? jutʃi.*

ípa? jutʃi [lùŋkhəm kha nukɪaŋ jim nət
EMPHAT story village.name CONJ village.name village two
ùŋ-təp-pà? jutʃhi]NP
fight-RECIP-NR story
'This story is the story of the fight between Lungkum village and Nokrang village.' (Lit. ... Lungkum and Nokrang villages' fight's story)

A true nominalization that fills the argument slot of an independent clause is not amenable to an alternative relative clause interpretation, even an internally headed one. To illustrate, it is impossible to interpret the nominalized clause of (6.77) as performing any syntactic function other than that of a noun phrase filling the slot of a copula topic. This sentence was uttered by a Mongsen speaker in the situation of a group of people sitting and warming themselves in the sun one winter morning.

(6.77) *tsəhŋi ku hwaŋkə məŋpai aɪu.ù?*

[tsəhŋi ku hwaŋ-əkə məŋ-pà? i]NP aɪu-əɪ-ù?
sun LOC roast-SIM sit-NR PROX be.good-PRES-DEC
'This sitting [and] bathing in the sun is good.'

Nominalization can involve the multiple layering of embedded clauses. The complex clause of (6.78), for instance, demonstrates three levels of structure. At its heart is the verbless ascriptive clause *wàzà-sá? tə-sən-ù?* '(the) bird-meat is sour'. At the next level of structure, *anu-tʃhaŋ nə* 'the boy' forms an agentive-marked A argument, and the above-mentioned verbless ascriptive clause forms an O complement of the nominalized verb of speech *sa-pà?*. The quotative particle *tə* is in its usual function as a complementizer and links the O complement to its complement-taking utterance verb. At the third level of

structure, the whole of the nominalized clause with its embedded verbless clause *anu-tʃhaŋ nə [wàzà-sá? tə-són-ù?] tə sa-pà? tʃu* occupies the verbless clause topic slot of the superordinate clause, and *sítak* ‘correct’ completes it as the verbless clause complement. The verbless clause topic of the superordinate clause is like an independent clause, with all of the notional arguments of *sa-pà?* present. If not for the nominalizer and the final determiner which identifies this as a clausal noun phrase, *anu-tʃhaŋ nə wàzà-sá? tə-són-ù? tə sa* could stand alone as an independent clause meaning ‘The boy said the bird-meat is sour’.

(6.78) a. *anutʃhaŋ nə wàzàsà? tésánù? tə sapà? tʃu sítak.*

[[a-nu-tʃhaŋ nə [wàzà? -sà? tə-són-ù?]

NRL-child-MAN AGT bird-meat NZP-be.sour-DEC

tə sa-pà?] tʃu]NP sítak

thus say-NR DIST correct

‘The boy’s saying that the bird meat is sour is correct.’

b. [**a-nu-tʃhaŋ** nə wàzà?-sà? tə-són-ù? tə sa-pà? tʃu]NP sítak

‘The boy who said that the bird meat is sour is correct.’

This complex sentence allows us to focus our attention on the overlapping interpretations of nominalization and relativization in Mongsen. Unlike the previous example of (6.77), out of context *anu-tʃhaŋ nə* could possibly be interpreted as the head of an internally headed relative clause. The structural interpretation of (6.78a) assumes that the whole nominalized clause of the bracketed constituent functions as a noun phrase and occupies the verbless clause topic slot of its verbless clause. Alternatively, (6.78b) recognizes *anu-tʃhaŋ* as the internal head of a relative clause, and *wàzà?-sà? tə-són-ù? tə sa-pà?* as its relativized attribute. In this situation, context must be relied upon to determine which of these is the most appropriate interpretation of underlying structure and function. According to my consultants’ intuitions, the relative clause interpretation exemplified by (6.78b) encodes some degree of restrictive reference; i.e. it suggests that the boy who commented on the bird meat being sour is being distinguished from other boys. In the narrative from which this example is taken, reference is in fact made to only one boy, therefore the nominalization interpretation of structure and function is contextually the correct one.

A final example of nominalization comes from a Khensa Mongsen text. This presents a copula sentence in which the first line of interlinearized text functions as a copula topic (CT), and the second line of interlinearized text functions as a copula complement (CC) modified by an oblique noun phrase.

- (6.79) *majaŋtʃhaŋəɪ khə tsəŋsaŋ lùpàŋ nət mijimtəpaʔ tsə itə álíma ku
məkhàpì təmijim litʃhà.*
- | | | | | | |
|---------------|-----------|--------------|--------------|---------------|--------|
| [majaŋtʃhaŋəɪ | khə | tsəŋsaŋlùpàŋ | nət | mijim-təp-pàʔ | tsə]CT |
| PN | CONJ | PN | two | love-RECIP-NR | that |
| [i-tə | a-límá | ku | mə-khà-pì | tə-mijim]CC | |
| PROX-thus | NRL-world | LOC | NEG-have-IRR | NZP-love | |
| li-tʃhà | | | | | |
| be-COP.PST | | | | | |
- ‘Mayangchanger and Tsengsang Lubang’s love for each other was like
no love that exists in this world.’

This is clearly an example of a nominalization that cannot alternatively be analyzed as an internally- or externally-headed relative clause. Firstly, it is not possible to interpret a gap in the nominalized clause that would correspond to the shared argument of that clause. Nor is there a possible candidate for an internal head. Secondly, the nominalized, reciprocal-marked predicate *mijim-təp-pàʔ* ‘love-RECIP-NR’ does not restrict its proposition to any noun phrase head of its clause, as would be expected if it were functioning as a relativized attribute. That is to say, it is impossible to interpret the nominalized clause as an internally-head relative clause expressing something like *Mayangchanger and Tsengsang Lubang who loved each other...*, because this interpretation of structure and meaning sits oddly with the meaning of the rest of the sentence. Thus, the whole construction with its notional syntactic arguments and phrase-final determiner must be functioning as an argument in the copula topic slot of the matrix clause. It is therefore a very elaborate noun phrase consisting of a whole sentential clause.

We return to nominalization again in the following chapter, which looks at the nominal morphology of Mongsen in greater detail.

Chapter 7

Nominal morphology

This chapter deals with the nominal morphology of Mongsen. It begins with a description of compounding processes, then discusses the functions of prefixes and suffixes that can be affixed to noun and verb stems to derive various types of nominal constituents.

7.1. Compound nouns

A compound noun may be defined as the outcome of a lexical process that combines two or more stems to derive a new meaning. The component parts of the newly formed word are also usually found in the language as independent lexical items, but a natural transition is for the compositional meanings of compounds to gradually become lexicalized. Thus, what begins as a juxtaposition of two nominals, one of which is in a descriptive, modifying or possessive relationship to the other, may progressively become more and more compound-like as the original meanings of the juxtaposed constituents are slowly lost in the lexicalization process.

Two examples of this are the words used to name the upper and lower wards of many Ao villages, each of which originally had its own village gate. It is generally the case that the upper and lower wards are referred to as *impaŋ məphu* and *imlaŋ məphu* respectively. The compounds *im-paŋ* and *im-laŋ* are comprised of the nominal roots *(j)im* ‘village’ combining with the relational noun roots *paŋ* ‘mouth’ and *laŋ* ‘bottom, rump, buttocks’ respectively. Despite the familiarity of these lexical items, it is now no longer immediately obvious to the inhabitants that the orifices (i.e. entrances) of Ao villages derive from two anatomical characteristics of a living creature. Nor is it obvious to native speakers that the name for the main thoroughfare of the village, *im-tuŋ*, is also a nominal compound. This often runs along the spine of a ridge and is flanked on either side by houses; its compound is similarly formed from the name for ‘village’ and the relational noun root *tuŋ* ‘stem, trunk’, a meaning that must have originally been limited to tree trunks and the stems of plants. The small tracks that lead off the main road to houses do indeed bear a metaphorical resemblance to branches and leaves attached to a tree trunk or plant stem.

Should one consider these to be examples of nominal juxtapositions expressing possessive relationships, e.g. ‘village’s mouth’ for *im-paŋ* etc., or are they now to be treated as *bona fide* compounds, because for many speakers

they have idiosyncratic meaning? Marginal examples such as these defy attempts at a decisive categorization. It is probably best to view them as being located on a lexicalization continuum that extends from nominal juxtaposition at one extreme, with a meaning predictable from composite parts, to fully lexicalized compounds at the other extreme, often expressing a totally idiosyncratic meaning that may not necessarily be retrievable from its immediate constituents.

Compound nouns are formed from the following types of constituents:

- (i) [N - N]_N
- (ii) [N - V]_N
- (iii) [V_{NOM} - N]_N
- (iv) [N - V_{NOM}]_N

The vast majority of these compounds are endocentric, i.e. they have identifiable heads. But there is also a number of exocentric compounds in which a head is not identifiable. For example, *ləph.u-tśəʔ* ‘mole on skin’, formed from *ləph.u* ‘cockroach’ and *átsəʔ* ‘excreta’ (literally ‘cockroach-shit’) is neither a cockroach nor a type of excrement, *akhu-hmila* ‘were-tiger, lycanthropist’, formed from *akhu* ‘tiger’ and *tə-hmila* ‘shadow, reflection, soul’ is neither a tiger nor a shadow, a reflection or a soul, and *tsəhŋi-təji* ‘watch, clock’, formed from *tsəhŋi* ‘sun’ and *tə-ji* ‘mother’ is neither the sun nor a mother. What these examples serve to demonstrate is that the semantic transparency of compound nouns largely depends upon the extent of their lexicalization.

Compound nouns are formed from both bound and unbound stems. Those formed from bound stems may or may not retain a nominal prefix *tə-* or *a-* in the word-internal environment; e.g. the relational prefix of *tə-ji* ‘RL-mother’ is retained in *tsəhŋi-təji*, but word-internal nominal prefixes are lost in both *ləph.u-tśəʔ* and *akhu-hmila*. Retention or loss of a word-internal nominal prefix is indicative of the extent of the syntactic boundedness of the two constituents in the lexicalized word.

The loss of nominal prefixes and resulting tone sandhi commonly found in compounds is also indicative of the phonological changes that are often associated cross-linguistically with compounding processes. One common manifestation of this in languages with contrastive stress is a rearrangement in stress assignment, e.g. English *key + hole* → ‘keyhole’. Tonal languages are more likely to demonstrate tone sandhi perturbations in compounding processes. The Thai words *sǐi* ‘colour’ and *khǎw* ‘be white’, for example, both have low-rising tones when uttered in isolation, but when they are compounded the tonal output is high on the first constituent and low-rising on the second, e.g. *sǐi-khǎw* ‘white colour’ (Coupe 2001: 735). It is extremely common for idiosyncratic tone sandhi to result from the apposition of nominal stems in Mongsen compounds – for paradigmatic examples, compare the often unpredictable tone sandhi of

noun-noun compounds associated with the root of *á-sáʔ* (NRL-meat) ‘meat’ as the head in §3.3.4.

7.1.1. Noun-noun compounds

Mongsen noun-noun compounds have a right-headed structure. Semantic categories of noun-noun compounds include words specifying a whole-part or possessor-possessee relationship, e.g. *áwk-tʃaŋ* (pig + leg) ‘pig’s leg’, or a modifier-modified relationship, such as village names, e.g. *kapza-im* (Kabza + village) ‘Kabza village’.

Terms referring to inhabitants of localities and ethnic nationalities are encoded by right-headed noun-noun compounds. These consist of the name of the place or the ethnic identity as the modifying element, plus the bound relational noun root *niŋ* ‘name’, which is nominalized by the agentive nominalizer *-əɪ* and functions as the head. Although the stem is already nominal, this nominalizing suffix is required in order to derive a nominal expressing a meaning that characterizes the referent with respect to their origin, habitual activity or ethnicity, e.g. *tsəmàʔ niŋ-əɪ* (Assam Plain + name-ANOM) ‘plainsman, Indian’, literally ‘Assam Plain-named one’ and *hlùtha niŋ-əɪ* (Lotha + name-ANOM) ‘a Lotha’, literally ‘Lotha-named one’. English also derives nouns from noun stems, e.g. *village* > *villager*, *hat* > *hatter* and *mule* > *muleteer*, demonstrating that the derivation of nouns from nouns is not that exotic. As in Mongsen, these derivations usually denote either a type of agent, or a referent whose habitual activity is characterized by the meaning denoted by the nominal base (literally: ‘one who has to do with NOUN’). The derivational functions of the Mongsen agentive nominalizer are discussed further in §7.4.4.

The corporeal whole-part bond that is intrinsic to many members of the relational noun class accounts for their frequency in compound nouns expressing possessor-possessee or modifier-modified relationships. Some of these lexicalizations are formed from a noun stem compounding with the bound root of *tə-puk* ‘stomach’. Recall that this is in the process of grammaticalizing as a nascent postposition that encodes the inessive or subessive case relation of oblique arguments. Examples of compound nouns it occurs in are *tʃaŋ-pùk* (leg/foot-STOMACH) ‘sole of foot’, *tə-tʃha-pùk* (RL-arm length/wing-STOMACH) ‘armpit’ and *luŋ-pùk* (stone-STOMACH) ‘cave’. It is precisely these compounding processes that have created the historical conditions under which *pùk* and the roots of other relational nouns have been reanalyzed as nascent postpositions (see §5.3.3 for discussion).

Sometimes it can be difficult to decide what the head of the compound is if the derived meaning is not unambiguously retrievable from one of the compound’s constituents, or if the categoriality of a constituent is in transition.

For example, *luŋ-pùk* ‘cave’ could be interpreted metaphorically as a type of stone cavity, relating to the meaning of ‘stomach’ (= modifier-modified relationship), or as a location within a stone, relating to the grammaticalized local case-marking function (= modified-modifier relationship). The interpretation depends upon the current status of *pùk*.

7.1.2. Noun-verb compounds

Noun-verb compounds present both left-headed and right-headed structures. Left-headed noun-verb compounds demonstrate a modified-modifier order of constituents in which a noun stem is the head, e.g. *səŋ-kuŋ* (wood + be.dry) ‘dry branch’, *tsə-məlu* (water + boil) ‘boiled water’ and *səŋ-pak* (wood + be.flat) ‘plank’.

Right-headed noun-verb compounds have a verb stem as the head, e.g. *ki-tʃhí?* (house + close) ‘door’, *aja-maŋ* (night + believe) ‘dream’ and *tʃaŋ-tʃhəm* (leg/foot + wear) ‘shoe’. The derived meanings of some of these lexicalizations suggest a modifier-modified relationship, in which the head is somewhat like an agentive or instrument nominalization. This seems reasonable for *ki-tʃhí?* (literally ‘house closer’), but less convincing for *aja-maŋ* or *tʃaŋ-tʃhəm*. Alternatively, these [N-V]_N lexicalizations could be regarded as corresponding to syntactic phrases in which N is considered to be a notional argument of V. Whereas this again seems feasible for *ki-tʃhí?*, such an analysis is a lot less convincing for *aja-maŋ*, because *aja* ‘night’ does not correspond to a notional argument of *maŋ* ‘believe’. The same reservations apply to the assumption of an argument-predicate relationship for *tʃaŋ-tʃhəm*.

There is often internal complexity in compound nouns that involves hierarchical layers of compounding. To illustrate, the first constituent in the compound *maŋmətuŋ jim* ‘Mangmetong village’¹ can be decomposed into a relational noun root, *maŋ* ‘body’, and the stative verb root *mətuŋ* ‘be straight, erect’. This is now lexicalized as one word and is itself analyzable as a left-headed noun-verb compound. The village name therefore represents the right-headed compounding of a left-headed compound and has a [[N-V]_N N]_N hierarchical structure. Similarly, the ancestral Ao village *tʃùŋlijimti* ‘Chungliyimti’, literally ‘big village of the Chungli’, is a compound of a compound. It is composed of the name of the Chungli moiety *tʃùŋli*, the noun *jim* ‘village’ and *ti*, the residue of the stative verb root *pəti* ‘be big’. There are two possible interpretations of its constituent structure: either a left-headed compound [[N-N]_N V]_N in which *ti* is the modifier, or a right-headed compound [N [N-V]_N]_N, in which *tʃùŋli* is the modifier.

1. The origin of the village name is explained in Text 3.

7.1.3. Nominalized verb-noun compounds

Nominalized verb-noun compounds have a right-headed structure and encode a modifier-modified meaning. There are two types. The dependent constituent of the first type is a deverbal adjective derived from a verb root by the nominalizing prefix *tə-*, e.g. *tə-sə-maŋ* (NZP-die + body) ‘corpse’, *tə-sə-nəm* (NZP-die + odour) ‘stench of a rotting carcass’ and *tə-thəm-paŋ* (NZP-end + mouth) ‘edge’. The nominalizing prefix on the dependent constituent in this type of compound noun is obligatory. The second type of nominalized verb-noun compound is formed with a verb stem nominalized by the general nominalizer *-pàʔ*, e.g. *tʃak-hən-paʔ lâtà* (paddy + take-NR moon) ‘October’, literally ‘paddy-carrying month’. This presents a hierarchical [N-V_{NOM}] N]_N structure.

Nominalized verb-noun compounds demonstrate varying degrees of transparency. The composite meanings of the examples provided above are largely interpretable from the meanings of their constituents, whereas that of *tə-məluŋ-tʃaŋ* (NZP-be.willing + seed) ‘heart’ is inextricable from the meanings of its parts.

7.1.4. Noun-nominalized verb compounds

Noun-nominalized verb compounds have a right-headed structure. The head is formed by a verb stem marked by the agentive nominalizer *-(ə)ɪ* or the locative nominalizer *-tʃən*. A noun stem functions as the dependent element and the resulting compound encodes a modifier-modified relationship, e.g. *kì-li-(ə)ɪ* (house + stay-ANOM) ‘servant’, literally ‘(the) house-staying one’. Noun-nominalized verb compounds can resemble syntactic phrases, because the dependent noun stem functions like an argument of the nominalized verb stem. In the case of *kìliɪ*, the argument of the corresponding syntactic phrase would be in oblique function. This may be a reflection of the extent of its lexicalization, because the semantically similar *a-jim ku li-ɪ* (NRL-village LOC stay-ANOM) ‘villager’ retains its compound-internal locative case marker.

The locative nominalizer *-tʃən* is used to derive a nominalization that functions as the head of a compound expressing a location, e.g. *tsəhŋi tʃhuwa-tʃən* (sun + emerge-LNOM) ‘east’, *tsəhŋi-wa-tʃən* (sun + go-LNOM) ‘west’ and *a-nu mən-tʃən* (NRL-child + sit-LNOM) ‘womb’. The various types of nominalizing morphemes are discussed under §7.4.

7.2. Nominal prefixes

There are four nominal prefixes. The vocative prefix *a-* occurs on kinship terms only and the nominalizing prefix *tə-* derives nominals from verb roots. In the discussion of open word classes in §4.1, I proposed that two sub-classes of bound nouns could be established according to the type of nominal prefix attaching to the nominal root – the relational prefix *tə-* occurs on kinship terms and body parts, and the non-relational prefix *a-* occurs on nouns denoting cultural artefacts and objects of the natural environment. Recall that while the forms of the vocative and non-relational prefix are segmentally identical, their functions are distinct. This is also true of the homophonous nominalizing and relational prefixes.

Evidence presented in §3.2 suggests that all nominal prefixes are underlyingly atonal, and that their output tones are copied from the tone of an adjacent root syllable following word formation.

7.2.1. Vocative prefix

The vocative prefix *a-* forms vocative inflections of bound nouns specifically belonging to the sub-class of kinship terms. These are employed for addressing speech act participants, for self reference, and additionally for third person reference between intimates.

- (7.1) *apa? – álímaj ku ami? sɔ? məliláw?, ...*
 a-pa? a-límá i ku a-mi? sɔ?
 VOC-father NRL-world PROX LOC NRL-person no-one
 mɛ-li-la-ù?
 NEG-be-NEG.PST-DEC
 ‘Father! There is nobody in this world, ...’

- (7.2) *aw nə səmpət khèjù.ɛ.*
 a-u nə səmpət khèjù-ì-ù? .ɛ
 VOC-grandfather AGT gift give-IRR-DEC PTCL
 ‘Grandfather [= ‘I’] will give you a present.’

Vocative forms of kinship terms, which can be used for self-reference (as in [7.2] above), may also be used for indirect third person reference in some pragmatically appropriate contexts, such as in conversations between friends or clan members in which reference is made to an absent member of the in-group.

- (7.3) *pa nə “nàŋ apa? taŋ nə waŋ.”*
 pa nə nàŋ a-pa? taŋ nə wa-aŋ
 3SG AGT 2SG VOC-father SIDE ALL go-IMP
 ‘‘You go to Father,’’ she [said].’

The vocative prefix *a-* could be replaced by the first person possessive pronoun *kə-* in (7.3). This would imply greater social distance and formality between the addressee and the absent referent. There are some unexpected gaps in the vocative paradigm – not all kinship terms can be used with the vocative inflection (see Table 4.3 for examples).

The vocative forms of kinship nouns are used for direct address in commands. The vocative noun form functions as a preposed topic in this context, while the imperative clause that follows functions as the comment. The following example demonstrates this pattern.

- (7.4) *phà:ia:ɪ, “aku – aliŋ khiŋ”, tə saw? tə:ɪ.*
 phà:ia?-ə:ɪ a-ku a-liŋ khi?-aŋ
 catch-SEQ VOC-uncle NRL-cane.strip give-IMP
 tə sa-ù? tə:ɪ
 thus say.PST-DEC REP
 ‘Having caught [the pig, he] said ‘‘Uncle! Give me the cane strip’’
 [to bind its legs].’

7.2.2. Relational prefix

The relational prefix *tə-* is limited to occurring on the bound roots of kinship terms, body parts and other entities that conceptually form part of a whole. For examples of kinship terms, see Table 4.3. Some examples of body part terms are *tə-mijuj* (RL-digit) ‘finger, toe’, *tə-nik* (RL-eye) ‘eye’ and *tə-ku* (RL-penis) ‘penis’. Examples of bound roots that do not strictly constitute body parts or kinship terms, but which are nevertheless associated with the body are *tə-sə* (RL-bile), ‘bile’, *tə-hmila* (RL-spirit) ‘reflection, spirit, shadow’ and *tə-saka* (RL-breath) ‘breath’. Other non-body part bound roots involved in a whole-part relationship that consequently qualifies them for the relational prefix are some (but not all) of the terms applying to botanical structures, e.g. *tə-mətshə* (RL-bud), ‘bud’, *tə-tuj* (RL-stem) ‘stem, trunk’ and *tə-zəŋ* (RL-root) ‘root’.

7.2.3. Non-relational prefix

The non-relational prefix *a-* occurs on bound nominal stems that denote meanings not subsumed by the semantic categories of noun described in §7.2.2. As noted above, these mostly belong to the semantic classes of cultural artefacts and objects of the natural environment. Some examples are *a-mi* (NRL-yam) ‘yam’, *a-tʃuŋ* (NRL-shield) ‘shield’, *a-tshəŋ* (NRL-rust) ‘rust’ and *a-ji* (NRL-rice.beer) ‘rice beer’. Additional examples and a discussion of the semantic classes of bound nouns taking the non-relational prefix can be found in §4.1.4.

7.2.4. Nominalizing prefix

The nominalizing prefix *tə-* derives nominals from verb roots. Nominal derivation is much more productive with stative intransitive verbs than with any other semantic class of verb, although this process does not uniformly apply to all semantic classes of stative intransitive verbs (those that undergo this type of derivation are discussed in §9.1.4). Some examples of derivations from non-stative intransitive and transitive verbs are *tə-sə* (NZP-die) ‘dead’, *tə-məlu-tʃəm* (NZP-boil-ADDIT) ‘boiled’ and *tə-məɬək* (NZP-clean) ‘clean’.

Deverbal adjectives derived from intransitive stative verbs are used as nominal attributes in noun phrases, e.g. *a-mi? tə-aiu* (NRL-person NZP-be.good) ‘(a) good man’, *tʃuŋi tə-mənən* (food NZP-be.dirty) ‘dirty/taboo food’ and *a-tʃa tə-tʃha* (NRL-cooked.rice NZP-be.hot) ‘hot cooked rice’. The various functions of the nominalizing prefix as a derivational morpheme are discussed further in §7.4.1 (also see §6.4.1 and §8.3.4). A deverbal adjective derived by *tə-* may be additionally used as a verbless clause complement that expresses a property of a referent. (see §9.1.4 for discussion and further examples).

- (7.5) *ipi təməɬəmù?*
 ipi tə-məɬəm-ù?
 PROX NZP-be.red-DEC
 ‘This is red.’

7.3. Possession

In §4.1.3 I noted that sub-classes of bound nouns could be formally recognized according to the type of nominal prefix their stems take in the citation form. The relational bound class is semantically characterized by nouns denoting body parts, entities expressing a whole-part relationship, and kinship terms. Nouns belonging to these particular semantic domains are sometimes given a distinct

syntactic treatment when marked for possession in many languages of the world, justifying the recognition of a type of possession referred to as inalienable (e.g. see Chappell and McGregor 1996 and references therein).

The very specific semantic domains of the relational bound nouns raises my suspicions that Mongsen may also have the underlying basis of an alienable ~ inalienable contrast, yet possessive marking does not always reflect this distinction in the morphological treatment of some semantic classes of noun (e.g. cf. the variable treatment of nouns denoting kinship terms and artefacts in §7.3.3). Nevertheless, it is interesting that some correlations are to be found. Mongsen certainly presents a system of noun classification that would be conducive to the grammaticalization of an alienable ~ inalienable contrast in its possessive marking.

The language employs three different strategies for encoding possession. Nominal possession can be expressed by:

- (i) the juxtaposition of possessive or personal pronouns (§7.3.1)
- (ii) the juxtaposition of common nouns (§7.3.2)
- (iii) the agentive nominalizer functioning as a genitive marker (§7.3.3)

Each of these strategies is described under its respective sub-heading in the following sections.

7.3.1. Possession encoded by pronouns

Possession can be encoded by forms of either the personal pronouns or the possessive pronouns (see §4.2 for a description of pronominal classes). The pronominal marking of possession applies to any sub-class of noun, under the following conditions.

The relational prefix *tə-* of a bound root can be replaced by a first or second person possessive pronoun to signal possession, e.g. *kə-ləm* (1SG.POSS-head) ‘my head’, *nə-ləm* (2SG.POSS-head) ‘your head’. Possession encoded by the third person singular pronoun *pa* is distinguished by the fact that it cannot replace the nominal prefix of a relational noun, e.g. *pa tə-ləm* (3SG RL-head) ‘his/her head’, never **pa-ləm*. This is attributed to historical factors discussed in §4.1.3. The first and second person possessive pronouns do not obligatorily replace the nominal prefix; e.g. *kə tə-ləm* and *nə tə-ləm* are alternative ways of expressing ‘my head’ and ‘your head’ respectively.

The possession of a relational noun may otherwise be signalled by apposition of a personal pronoun to the noun phrase head. Retention of the relational prefix in these circumstances is obligatory, e.g. *nì tə-ləm* (1SG RL-head) and *nàŋ tə-ləm* (2SG RL-head) ‘my head’, never **nì-ləm* or **nàŋ-ləm*. This suggests that the third person singular pronoun *pa* – the form that expresses

this person/number category in both the personal and possessive pronominal paradigms – historically belongs to the personal pronoun paradigm. It follows the pattern of the first and second person singular personal pronouns in not being able to replace the relational prefix of a bound noun when used to mark a possessive relationship.

I mentioned in §4.1.4 that the use of personal pronouns to encode possession is considered ungrammatical by younger speakers, yet is frequently encountered in the narratives and speech of speakers aged approximately sixty years and above. Older speakers use a variety of morphological means to encode the possession of an identical referent used in identical contexts. This is demonstrated by the three different ways in which three elderly Mongsen speakers introduced themselves at the beginning of their narratives. Remarkably, Mongsen speakers aged approximately forty years or younger uniformly rejected the possessive construction of (7.7) as ungrammatical. The speaker of this particular sentence was a highly respected village elder, a man recognized as an authority on the tribal knowledge of the Ao. The sentence was uttered without a pause between the pronoun and the compound head, therefore *nì* cannot be analyzed as an extra-clausal topic.

- (7.6) *kə təniŋ .ɹuŋsənləmpa? tʃəmì.*
 kə tə-niŋ .ɹuŋsənləm-pà? tʃəmì
 1SG.POSS RL-name PN-M clan.name
 ‘My name is Rongsenlemba Jamir.’

- (7.7) *nì təniŋ imtilùin.*
 nì tə-niŋ imtilùin
 1SG.POSS RL-name PN
 ‘My name is Imtiluin.’

- (7.8) *kəniŋ ləmnij khà-pà?*
 kə-niŋ ləmnij.khà-pà?
 1SG.POSS-name PN-M
 ‘My name is Lemning Kaba.’

Plural possession of an unbound noun is encoded simply by juxtaposition of a possessive pronoun to the head of the noun phrase, e.g. *aj mäsə?* (1PL.POSS cow) ‘our cow’ and *nij kin* (2PL.POSS clan) ‘your (pl.) clan’. Alternatively, plural possession can be encoded by a form of personal pronoun, e.g. *ila maŋmətuŋ lima* (1PL.EX Mangmetong country) ‘our Mangmetong land’, *nəŋkhəla thiŋa?* (2PL wages) ‘your (pl.) wages’, and *túŋkhəla hjaŋa* (3PL generational.age.group) ‘their generational age group’. Both modes of plural possessive marking are found in texts. Note the severe phonological reduction

of the third person plural pronoun in the italicized line of (7.9). This is unrelated to possessive marking and the unreduced form could instead be used without any change in meaning.

- (7.9) *tə ula luŋ tʃu aʃhu tʃu thəməkə tətəpəɪ sənsaw?*
 tə túŋkhəla luŋ tʃu a-ʃhu tʃu thəm-əkə
 thus 3PL group DIST NRL-song DIST end-SIM
 tən-təp-əɪ sənsa-ùʔ
 sing-COLL-SEQ depart-DEC
 ‘Thus, their group finished singing the song together and departed.’

- (7.10) *nuksənsaŋpəʔ tʃu pɑɪ kɪmə tənəkə áhlù nə wa.*
 nuksənsaŋ-pəʔ tʃu pɑɪ kɪmə tən-əkə
 PN-M that 3PL.POSS courtyard cross-SIM
 a-hlú nə wa
 NRL-field ALL go.PST
 ‘Noksensangba was crossing their family’s courtyard as he went to the field.’

The form of the plural possessive pronoun *pɑɪ* deserves special comment. This appears to be based on the third person singular form *pa*. My suspicion is that the retroflex approximant coda is the agentive nominalizer, which is sometimes used as a genitive marker to encode possession, in addition to its nominalizing functions (see §7.3.3 and §7.4.4). If so, the pronominal stem with its genitive suffix *pa-əɪ* (3SG-ANOM) has become lexicalized as *pɑɪ* and now encodes the third person plural category of possessive pronouns. The origin of the high tone cannot be explained at present.

Recall that the possession of a collectively owned entity such as a house or a field is always expressed in the plural number, even when its possession might be more relevantly encoded in the singular number in a particular context of reference. Speakers also hold the very strong impression that the plural possessive pronouns encode possession not just by any unrelated group of people, but specifically by a family unit or clan. The free translation of (7.10) above is therefore an accurate reflection of how Mongsen speakers interpret the reference of the plural possessive pronoun in that sentence.

7.3.2. Possession encoded by juxtaposed nouns

Possessive compounds are formed by the apposition of two noun stems, e.g. *wah.u həntsə* (raven + egg) → ‘raven(’s) eggs’. Context determines whether the

dependent element of the compound is in a possessive or a modifying descriptive relationship to the compound head.

While it is possible to extract recursively embedded examples such as *azənti-paʔ-əi ajiʔ təmi* (old.person-M-ANOM dog RL-tail) ‘[the] old man’s dog’s tail’ from obliging consultants, this type of complex noun phrase is not encountered in the naturally spoken language. In fact, the consultants themselves point out that Mongsen is not spoken in this way. Recursive possession is instead conveyed by periphrastic means using rephrased and reiterated noun phrases in apposition, as if to clarify the reference of possession. The following example is typical of this circumlocution.

- (7.11) *təku atsə tsəŋhɪəm nə, təmhnaʔ tʃu, nuksənsaŋpaʔ təmhnaʔ tʃu, təhmila tʃu tʃatʃuk.*
- | | | | | | |
|-----------------|----------------|----------|----------|----------------|------|
| tə-ku | a-tsə | tsəŋhɪəm | nə | tə-əmhnaʔ | tʃu |
| thus-LOC.CV | NRL-water | deity | AGT | RL-male.in.law | DIST |
| nuksənsaŋ-paʔ | tə-əmhnaʔ | tʃu | tə-hmila | tʃu | |
| PN-M | RL-male.in.law | DIST | RL-soul | DIST | |
| tʃaʔ-tʃuk | | | | | |
| consume-PFV.PST | | | | | |
- ‘And then, a water spirit ate Noksensangba’s son-in-law’s soul.’

As suggested in §6.4.1, this may be a strategy to limit modifying attributes to a maximum of one per noun phrase, so as to reduce the syntactic heaviness of noun phrases and the cognitive processing demands placed on interlocutors. Or, it may represent a discourse convention in which speakers narrow the reference to a particular noun phrase head by postposing attributes and other information in separate noun phrases. This supposition is further supported by the presence of discrete pauses between noun phrases, represented by commas in the italicized line of (7.11).

The other context in which nominal apposition is used to express a possessive relationship is when the possessor is human and the possessee is a body part or other inalienably possessed entity, or an artefact that is individually possessed. This is discussed with reference to genitive marking in the following section.

7.3.3. Possession encoded by the agentive nominalizer

In addition to its primary nominalizing/relativizing function, a secondary extended function as a sequential converb marker (see §7.4.4 and §11.4.1.1 respectively), and possibly even extension to a tense marking function (see

§8.5.12), the agentive nominalizing suffix *-əɪ* is used for marking a genitive relationship between a human possessor and some semantic classes of noun. When used in this genitive marking role, the agentive nominalizer occurs suffixed to the dependent constituent of a possessive noun phrase, as in the following Waromung Mongsen example.

- (7.12) *nì akui ki na tsəŋsaŋ katə̀nku,*
 nì a-ku-əɪ ki na tsəŋsaŋ ka-tə̀n-ku
 1SG VOC-uncle-ANOM house ALL dance ascend-WHEN-LOC.CV
 ‘When I go up to Uncle’s family’s house to dance, ...’

Possession encoded by the agentive nominalizer (often metathesized to *-ɪə* by speakers of the Mangmetong Mongsen variety) is overall relatively rare, probably due to the preference of speakers to encode possession via periphrastic means (see [7.11] above for a demonstration of this). A survey of narrative texts reveals that the agentive nominalizer is never used to encode the personal possession of a body part, nor of any other nouns expressing an inherent whole-part relationship to another entity. Speakers remarked that the elicited example of (7.13b) with the agentive nominalizer sounded peculiar, because it suggests that the possessee is unnaturally disconnected from its possessor.

- (7.13) a. *muwa-pàʔ hmila*
 PN-M shadow/soul/reflection
 ‘Moa’s soul/shadow/reflection’
 b. **muwa-pàʔ-əɪ hmila*

Exploratory elicitation also confirms that the possession of body parts and other entities thought to be bound to their possessors in perpetuity must be encoded by nominal apposition. This correlates with a cross-linguistic tendency for inalienable possession to be expressed by an absence of morphological marking, and for the opposing category of alienable possession to be signalled by morphological means. A possessive noun phrase such as *muwa-pàʔ-əɪ khət* ‘Moaba’s hand’ is still a possible construction in Mongsen, but implies an alienable relationship between the possessor and the possessed entity, such as would be the case if the possessor has taken the severed hand of an enemy as a war trophy.²

There are some peculiarities of genitival marking with respect to the semantic classes of noun to which it applies. While terms denoting body parts

2. This is not quite as contrived as it may first seem. Body parts were indeed taken as war trophies during the pre-Christian era – see Mills (1926: plate facing p.161) for a graphic display of an enemy’s severed limbs.

and other nouns suggestive of a permanent association with the body fall together under what could be analyzed as an inalienable category (and thus are encoded by nominal apposition), possessors of heads denoting kinship terms are sometimes marked by the agentive nominalizer. The following example contrasts the means of encoding a possessive kin relationship between an individual and his daughter (here *-əɪ* marks the possessor in NP₁), against one encoding the relationship between a deceased individual and her corpse (signalled by the absence of morphological marking between the possessor and the possessed head in NP₂).

- (7.14) *aɪəmtʃhaŋ.ə tʃâla jàsà na.ɹu amaŋ tʃu maŋmətuŋ ku mətuŋjakɪ.*
 [aɪəmtʃhaŋ-əɪ tʃâ-la]_{NP1} [jàsà.na.ɹu a-maŋ tʃu]_{NP2}
 PN-ANOM offspring-F PN NRL-body DIST
 maŋmətuŋ ku mətuŋ-jak-iʔ-əɪ
 village.name LOC be.erect-RS-CAUS-PRES
 ‘Aremchang’s daughter Yasa Naro’s corpse is kept propped up there in Mangmetong [village].’

Inconsistency in the expression of possessive kin relationships is occasionally encountered in texts. In this particular example, ‘Aremchang’s daughter’ has genitive marking on the possessor, yet in the textual example of (7.11) above, ‘Noksensangba’s son-in-law’ is encoded just by nominal apposition. In lieu of a confident explanation for this variation, I surmise that the agentive nominalizer may be an optional marker of genitivation for this particular semantic class of possessee, possibly comparable to the morpheme *khǎŋ* that is electively used to signal a possessive relationship between two nouns in Thai. To illustrate, the possessive noun phrase *rót khǎŋ mǎw* (car belong doctor) ‘(the) doctor’s car’ can also be expressed simply by apposition of the nouns denoting ‘car’ and ‘doctor’ e.g. *rót mǎw*. The omission of *khǎŋ* does not affect the meaning. Analogous possibilities for omitting the agentive nominalizer of Mongsen may account for the variation in possessive marking seen in these examples.

An unexpected semantic entailment of genitival marking is that when it is used to signal the possession of an artefact, it often suggests group possession. The difference between the elicited examples of (7.15a–b) is that (7.15a) implies individual possession, whereas (7.15b) implies collective ownership by a family or an associated group of people.

- (7.15) a. *tusi-pàʔ nuk*
 PN-M machete
 ‘Toshiba’s [personal] machete’

- b. *tusi-pàʔ-əɪ nuk*
 PN-M-ANOM machete
 ‘Toshiba’s family’s/gang’s machete’

So it would seem to be the case that if an alienable entity is individually possessed, or if the possessed entity falls under one of the inalienable semantic categories of noun discussed above, then possession is encoded just by nominal apposition. If on the other hand a possessed artefact belongs to a family or other functional group, then the agentive nominalizer must be used to encode the possessive relationship. This may provide an alternative explanation for variation in marking possession with respect to kinship relations; i.e. the agentive nominalizer seen in the possessive noun phrase *aiəmtʃhaŋ-əɪ tʃà-la* of (7.14) might have been used by the speaker to imply group possession, viz. *Aramchang’s family’s daughter*, but this remains to be confirmed.

The agentive nominalizer can be used in a genitive function to mark the possession of a noun phrase that has been relativized by the general nominalizer *-pàʔ*. This morphological accretion is demonstrated in (7.16).

- (7.16) *pa kəiàku, azəntila pa nə “aja” tʃapaɪ ki sə nə walikàʔ,*
pa kəiə-ku
 3SG ascend+come-LOC.CV
 [[azənti-la [pa nə a-ja tʃa-pàʔ] -əɪ]_{NP} ki
 old.person-F 3SG AGT VOC-mother call-NR-ANOM house
 sə nə]_{NP} wa-likàʔ
 ANAPH ALL go-CONTEMP
 ‘At the time he was coming up [to the village], when he went to the
 [aforementioned] house of the old woman whom he called “Mother”, ...’

There are obvious parallels to be drawn between the nominalizing, genitivizing and relativizing functions of the Mongsen agentive nominalizer *-(ə)ɪ* and the similar functions of the Lahu particle *ve* described by Matisoff (1972). This possibly includes the optional deletion of *ve* in possessive noun phrases, although the circumstances under which this might be allowed appear to be more restricted in Mongsen.

7.4. Nominalization

Nominalization is a process by which nouns or noun phrases are derived from individual words or whole clauses. Nominalization applies to noun stems as well as to verb stems. Both prefixation and suffixation are used for verb stem

nominalization, whereas just suffixation is used for noun stem nominalization. Suffixation is overall the more common derivational pattern and has the widest application. The following sections describe the various nominalization strategies in Mongsen, starting with nominal derivation using the nominal prefix.

7.4.1. Nominal derivation by prefixation

The derivation of deverbal adjectives by the nominalizing prefix *tə-* is described in §7.2.4. While many intransitive verb roots can be used as modifiers of noun phrase heads following derivation by *tə-*, only a small number of these can also function as verbal nouns.

Although the nominal morphology of Mongsen could potentially provide the grammatical hardware necessary to derive a verbal noun from virtually any semantically appropriate verb stem, this proves to be a non-productive process in the language. The results of elicitation confirm the premiss that *tə-* cannot be used willy-nilly to nominalize any verb stem that can subsequently function as a verbal noun, and only a select few derivations appear to have undergone the process of lexicalization. Some examples of verbal nouns derived from intransitive verb stems nominalized by *tə-* are *tfäsi* ‘be distressed’ → *tə-tfäsi* ‘distress’, *thəm* ‘end/finish’ → *tə-thəm* ‘end’ (n.), *sə* ‘die’ → *tə-sə* ‘death’, *kəlak* ‘be cunning’ → *tə-kəlak* ‘cunning’ (n.), *pəla* ‘be happy’ → *tə-pəla* ‘happiness’ and *məlak* ‘lie’ → *tə-məlak* ‘lie’ (n.). Some rare examples of nominalized transitive verb stems that function as verbal nouns are *maŋ* ‘believe’ → *ta-maŋ* ‘faith’ (n.),³ *tsəpha* ‘fear’ → *tə-tsəpha* ‘fear’ (n.) and *mijim* ‘love’ → *tə-mijim* ‘love’ (n.). These derivations, based on rather disparate verb classes, are all used as heads of noun phrases in texts.

Most stative intransitive verb roots denoting properties can be nominalized by *tə-* and then used as deverbal adjectives in verbless clause complements. There are some restrictions, however, which are determined according to the semantic class of the verb. These constraints are explained in §9.1.4. The nominalizing prefix is additionally used to derive the non-finite head of one type of purposive complement. Because of their inherent semantics, the derivation of purposive complements mostly applies to the stems of activity verbs (see §7.4.7).

The nominalizing prefix occurs on a verb stem to create a type of nominal derivation that is consistent with what Comrie and Thompson (1985: 355–356)

3. The vowel change in the prefixed syllable of *tamaŋ* probably resulted from a sporadic vowel harmony that was subsequently lexicalized (see §2.6.2 for additional examples).

describe as “objective nominalization”, although perhaps a term like “patientive nominalization” would be more accurate for this kind of construction in Mongsen, given that a grammatical relation of “object” cannot be demonstrated. The following example demonstrates the patientive nominalization.

- (7.17) *ila mùŋsən kìn nə lùŋsà lí phinə təmətʃəj tʃáʔ mətʃä.*
 ìla mùŋsən kìn nə lùŋsà lí phinə
 1PL.EX Mongsen clan AGT Longsa country ABL
 tə-mə-tʃäʔ-ì tʃáʔ mətʃäʔ-ə̀
 NZP-NEG-consume-IRR nothing NEG-consume-PRES
 ‘We Mongsen clans from Longsa country do not eat anything taboo.’
 (Literally: ‘... eat nothing that is not to eat’)

In addition to the nominalizing prefix *tə-*, what I interpret to be the irrealis suffix *-ì* attaches to the stem, and the affixes function together like a discontinuous morpheme embracing the verb root. In contrast to the derivation of verbal nouns, the derivation of patientive nominalizations is a completely productive process applying to any semantically compatible transitive verb root. The derived stem can include the negative prefix, but this is not obligatory to this type of nominalization, as a comparison of the negated forms of (7.18) to the positive polarity examples of (7.19–7.20) demonstrate.

- (7.18) a. *təmésanaj* b. *təməkùkì*
 tə-mə-sana-ì tə-mə-kùk-ì
 NZP-NEG-speak-IRR NZP-NEG-succeed-IRR
 ‘that which is not to be spoken’ ‘that which would not be achieved’

- (7.19) *tətʃhàku pa sə, ámikhót tʃàpàʔ sə nə tʃùku təphiləmì à .jà.*
 tətʃhà-ku pa sə a-miʔ khót tʃàʔ-pàʔ
 thus-do-LOC.CV 3SG ANAPH NRL-person hand consume-NR
 sə nə tʃùku [tə-philəm-ì à]_{NP} .jà-ə̀
 ANAPH INST consequently NZP-think-IRR one come-PRES
 ‘And so, because⁴ she ate a human hand, consequently a burden comes [to him].’

4. A verb stem + general nominalizer (+ optional demonstrative) that is followed by the instrumental case marker has been reanalyzed as a marker of a dependent clause encoding ‘because VERB’ (see §11.4.3.6). For discussion of similar examples of a demonstrative occurring within a suffix, see §5.3.2.2.

- (7.20) *ípaʔ pəntʃũŋ ku tsəŋ.ɾəm pʰəniŋ álú təmuwaj məsà.ɿ*
 ípaʔ pəntʃũŋ ku tsəŋ.ɾəm pʰəniŋ a-lú tə-muwa-ì
 EMPHAT festival LOC deity ABL NRL-field NZP-bless-IRR
 məsàʔ-əɿ
 request-PRES
 ‘In that festival, a field-blessing is requested from the deity.’

An outcome of the patientive nominalization is the emergence of fully-fledged abstract nouns that take determiners or demonstrate other evidence of their nounhood, such as an ability to function as constituents of possessive compounds. In (7.19), the nominalized *tə-philəm-ì* (NZP-think-IRR), based on the root ‘think’, has acquired an abstract nominal meaning of ‘something worrying that requires thought, a mental burden’ and serves as the head of a noun phrase in S function. This is distinguished by square brackets in the interlinearization. In the Waromung Mongsen textual example of (7.20), the nominal derivation *tə-muwa-ì* (NZP-bless-IRR) ‘blessing’ similarly expresses an abstract meaning and functions as a nominal head with respect to its modifier *a-lú* (NRL-field) ‘field’. The structure of this noun phrase is analogous to the possessive noun phrase *a-miʔ-khət* (NRL-person-hand) ‘person’s hand’ seen in the preceding example, in which *amiʔ* ‘person’ functions as a modifier to the head of the bound stem *khət* ‘hand’.

Related to this type of derivation are deverbal adjectives formed with the nominalizing prefix and the desiderative suffix *-mì*, which is more commonly found on non-nominalized verb stems. These word formations apply fairly productively to both transitive and intransitive verb roots to derive an attributive meaning of ‘VERB-able’, some examples being *tə-mijim-mì* (NZP-love-DESID) ‘lovable’ and *tə-məni-mì* (NZP-laugh/smile-DESID) ‘humorous’. The desiderative marker, described in §8.5.5.4, is one of the select few verbal suffixes tolerated on a verb stem nominalized by prefixation.

The nominalizing prefix is an integral component of yet another type of nominalization that is productively derived from verb roots. This has an adverbial meaning. It serves to provide a temporal background expressing ‘at the time of VERB’, upon which matrix clause events are superimposed.

- (7.21) *təʃhuwaj tʃu ku*
 tə-ʃhuwa-ì tʃu ku
 NZP-emerge-IRR DIST LOC
 ‘At the time of returning ...’

The temporal nominalization also uses a combination of the nominalizing prefix *tə-* and the irrealis suffix *-ì* on either side of the stem, but additionally requires the presence of the locative case marker *ku* functioning as a converb

marker. The whole construction derives a meaning consistent with a non-finite temporal clause. The nominal demonstrative that occurs between the nominalized verb stem and the locative marker of (7.21) is an optional constituent, merely providing additional deictic specification with respect to the time of the event. Temporal nominalization is discussed in §7.4.8, and the use of the locative case marking clitic as a marker of non-finite temporal converb clauses is described in §11.4.3.7.

Figure 7.1 below summarizes the diverse range of derivational functions in which the nominalizing prefix participates. The versatility of this prefix is comparable in functional range to the *-wa* suffix of the Bodic language Chantyal described by Noonan (1997), with which it shares many similarities, despite obviously being unrelated diachronically. Naturally this assumes that all of the meanings presented in Figure 7.1 stem historically from the same morpheme. The nominalizing prefix also occurs recurrently in various types of nominalizations also involving the general nominalizer suffix. These derivations are discussed in the following sections.

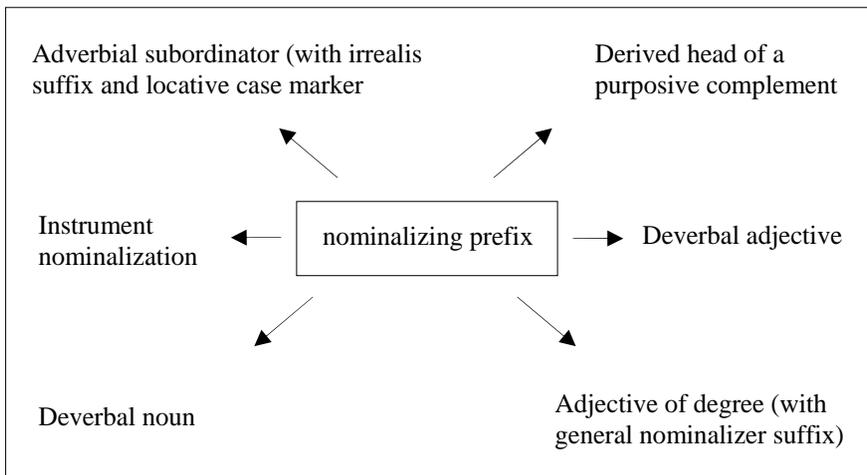


Figure 7.1. Summary of the derivational functions of the nominalizing prefix *tə-*

7.4.2. General nominalizer

The general nominalizer *-pàʔ* carries a functional load exceeding even that of the nominalizing prefix *tə-*. It is used for the nominalization of individual words and whole clauses (§6.6.7), relativization (§6.6.1 and §6.6.2), the derivation of adjectives of degree (§7.4.3), instrument nominalization (§7.4.6) and ordinal numeral derivation (§7.6.2). It also derives the preferred citation form of verbs

and can be used in place of the nominalizing prefix *tə-* to nominalize the stems of verbs in purposive clauses (§7.4.7). Apart from performing these synchronic functions, a diachronically related form is used as the masculine gender marker (§7.5.2), and a fossilized residue is found in the conditional, concessive and causal converb suffixes (§11.4.3).

When used to derive the preferred citation form of verbs, the general nominalizer usually co-occurs with the irrealis suffix.

- (7.22) *tʃəjɾpàʔ tə əŋkɪatʃi nə kútə səj.*
 tʃəjɾ-ì-pàʔ tə əŋkɪatʃi nə kútə sa-ì
 consume-IRR-NR thus English INST how say-IRR
 ‘How would you say “to eat” in English?’

The general nominalizer offers an alternative means of deriving verbal nouns from verb stems. In the following example, it derives a nominalization that functions as the head of its own noun phrase (cf. [7.29], which uses the agentive nominalizer on the same verb stem to derive an identical meaning). Note also how Mongsén avoids syntactically heavy noun phrases and instead uses prosodically independent nominalizations in a periphrastic manner to clarify the identity of a referent. For further evidence of this tendency to limit the number of modifying attributes in the noun phrase, compare the examples of (6.14) and (7.11), and also see the discussion of §6.4.1.

- (7.23) *ɪŋlən nɪŋəɪ, tənɪsɪpàʔ, kəptənɾpàʔ tʃu nə “á pi a-jim tsəŋhɪəm.”*
 ɪŋlən nɪŋ-əɪ [tə-nì-si-pàʔ]_{NP} kəptən-pàʔ tʃu nə
 England name-ANOM NZP-lead-RPET-NR captain-M DIST AGT
 á pi a-jim tsəŋhɪəm
 EXCLM PROX NRL-village deity
 ‘The Englishman, the leader, the captain [said of the old men] “Ah! These are [like] village gods.”’

Descriptions of the other functions of the general nominalizer can be found in the cross-references provided above. Nominalization and relativization are by far the most important functions of the general nominalizer and are discussed in detail in §6.6.

7.4.3. Nominalizations as adjectives of degree

Adjectives of degree are nominal derivations of intransitive verb roots that are capable of expressing gradable properties. Comparative constructions of the

type ‘X is longer than Y’ and superlatives of the type ‘X is the longest amongst all’ have the structure of verbless clauses, consisting of a verbless clause topic and a verbless clause complement. In this respect they are little different from the ascriptive clauses of §9.1.4, except that comparative and superlative constructions allow an oblique argument to occur as the index of comparison (being ‘than Y’ in the schematic representation of ‘X is longer than Y’, and ‘amongst all’ in the case of the superlative).

7.4.3.1 Comparative

A comparative adjective is a type of derived nominal that functions semantically as a graded attribute of a noun. Comparative adjectives are formed by affixing the nominalizing prefix *tə-* and the general nominalizer *-pàʔ* to the margins of an intransitive verb root, e.g. *mijaŋ* ‘be.sweet’ → *tə-mijaŋ-paʔ* ‘sweeter’.⁵ Tone sandhi perturbations usually result from the comparative derivation and are described in §3.1.3.1.

The derived adjective expresses the parameter of comparison. It inherently expresses the index of comparison as well, as there is no morpheme in the comparative construction expressing *more* or *less*. The gradation is conceptualized in greater terms only. If a speaker wishes to express the equivalent of ‘less hard’ as a comparative, for example, then this must be conveyed by its antonym *tə-nəp-paʔ* (NZP-be.soft-NR) ‘softer’. An alternative strategy is for a comparative proposition ‘X is bigger than Y’ to be globally negated by a clause-final negative particle *nuy* (example [4.104] of §4.2.14.5 demonstrates a similar use of the negative particle in a sequential converb clause).

The comitative case marker *thən* is used to mark the noun phrase functioning as the standard of comparison, but this oblique argument is not obligatory and is sometimes omitted if its referent is recoverable from discourse. The noun phrase of the compared entity can also be omitted if its referent is understood from context; for an example of this, see (7.25) below. The compared entity otherwise functions syntactically as a verbless clause topic and the comparative derivation functions as a verbless clause complement. There do not appear to be any structural constraints on the position of the standard of comparison *vis-à-vis* the compared entity. The degree of topicality or other pragmatic considerations determines the relative order of arguments in this type of verbless clause.

When the referent of a noun phrase functioning as the standard of comparison is pronominal, some speakers use either a personal pronoun or a

5. A small number of intransitive verb roots do not additionally require the nominalizing prefix to derive a comparative adjective. Their derivation is discussed in §3.1.3.1.

possessive pronoun as the head. This option is a characteristic of case marking clitics used for marking the pronominal heads of oblique arguments and is discussed in §5.3.2.

(7.24) *nìŋ tʃàɪ i tʃàɪ thən tasupaw?*

nìŋ tʃà-əɪ i tʃà-əɪ thən
 2PL.POSS offspring-ANOM 1PL.POSS offspring-ANOM COM
 tə-asuʔ-pàʔ-ùʔ
 NZP-be.short-NR-DEC
 ‘Your son is smaller than our son.’

(7.25) *nì thən təpətɪpəʔ liɪ pà.*

nì thən tə-pəti-pàʔ li-əɪ pà
 1SG COM NZP-big-NR be-PRES QPTCL
 ‘Is there [one] bigger than me?’ [asked Tiger]

An adjective of degree expressing a gradable property can be used as a headless relative clause (§6.6.2). The relativization functions as an attribute of its omitted head. In the following example, the head of the noun phrase *təpətɪpəʔ à nə* has been deleted, but can be determined from the context. The fact that the elided head could be restored to a position after the nominalized verb stem, e.g. *təpətɪpəʔ Ø à nə*, indicates that this is indeed an instance of relativization. Furthermore, this particular context requires the relative clause to have restrictive reference, therefore it would be inappropriate to restore the head to a position in front of the nominalized stem (see §6.6.1 for discussion of the structural difference between a restrictive and non-restrictive relative clause).

(7.26) *nàŋ thən təpətɪpəʔ à nə juŋthəŋəɪ, nì mət̚àtʃhət̚lâw?*

nàŋ thən tə-pəti-pàʔ à nə juŋ-thəŋ-əɪ
 2SG COM NZP-big-NR one AGT block-SEVER-SEQ
 nì mət̚-à-tʃhət̚-la-ùʔ
 1SG NEG-come-ABIL-NEG.PST-DEC
 ‘A bigger [tiger] than you blocked [the way] and I wasn’t able to come.’

7.4.3.2 *Superlative*

There are two ways of expressing the superlative degree. One uses the same nominalizing morphology as the comparative, plus an intensifier suffix *-thi* with a floating high tone. The intensifier suffix attaches to the stem in the slot preceding the general nominalizer, and the floating high tone displaces the tone

on the syllable preceding the intensifier suffix. An alternative is to not use the segmental intensifier, but to retain the floating high tone and associate it to the preceding syllable (i.e. the penultimate syllable of the derivation). The high tone then becomes a significant phonological cue for differentiating the superlative derivation from the segmentally identical comparative derivation.

An autosegmental analysis of the function of grammatical tone in derived adjectives of degree is presented in §3.1.3.1; the alternative strategies for encoding the superlative degree are demonstrated by the following examples.

- (7.27) *i tʃàɪ təmáj .ɲŋku tasúthipàw?*
 i tʃà-əɪ təmáj .ɲŋ ku
 1PL.POSS offspring-ANOM all AMONG LOC
 tə-asuʔ-ʼ-thi-pàʔ-ùʔ
 NZP-be.small-SUP-INTNS-NR-DEC
 ‘Our boy is the smallest amongst all.’

- (7.28) ... *akhu təpətípàʔ təlɪʔ .ɲŋku ...*
 a-khu tə-pəti-ʼ-pàʔ təlɪʔ .ɲŋ ku
 NRL-tiger NZP-be.big-SUP-NR all AMONG LOC
 ‘... the biggest tiger amongst all ...’

The superlative construction of (7.28) demonstrates that a derived adjective of degree can function just like any other modifying attribute of the noun phrase and similarly occurs in the post-head position, one of the two slots in which nominal attributes normally occur (see §6.1 for a description of noun phrase structure).

7.4.4. Agentive nominalizer

Verbs and nouns can be nominalized by the agentive nominalizer *-əɪ* to derive an agentive noun, ‘the one who VERBS’ in the case of nominalization of a verb, or ‘the NOUN one’ in the case of nominalization of a noun. This type of nominalization is restricted to noun phrases with human referents. The scope of nominalization can be a simple noun (7.29), a compound noun (7.30), a noun phrase (7.31), or a whole clause functioning as a pre-head relativization (7.32).

- (7.29) *tənisɪɪ*
 [tə-nì-si-əɪ]_N
 NZP-lead-RPET-ANOM
 ‘leader’

(7.30) *nùŋsəʔ niŋəɪ*

[nùŋsəʔ niŋ-əɪ]_N
 ancient.time name-ANOM
 ‘ancestor, forebear’

(7.31) *ajim ku liɪ*

[a-jim ku] li-əɪ]_{NP}
 NRL-village LOC stay-ANOM
 ‘villager’ (Literally: ‘in-village-staying one’)

(7.32) *anu skul nə waɪ luŋ li kákət khiukù.*

[[a-nu skul nə wa-əɪ] luŋ li]_{NP} kákət khiʔ-ukù
 NRL-child school ALL go-ANOM group DAT book give-ANT
 ‘The group of children who go to school have been given books.’

The relative clause of (7.32) and its dative-marked head could alternatively be interpreted as a possessive construction linked by *-əɪ* in genitive function and translated as ‘to the children going to school’s group’. This example demonstrates the fuzzy border between genitive and relative clause marking with the agentive nominalizer in Mongsen. Both types of syntactic constructions involve a nominal constituent in a subordinate relationship to another, and as Matisoff (1972), DeLancey (1986) and Genetti (1992) collectively show, the use of the same marker for encoding both types of syntactic relationships is a widely attested grammaticalization pattern in Tibeto-Burman languages.

The nominalization of verbs, nouns and clauses is just one of the varied functions of the agentive nominalizer. Another is a masculine semantic gender marking function in some words, e.g. *tʃä-əɪ* (offspring-ANOM) ‘son’, as opposed to *tʃä-la* (offspring-F) ‘daughter’. A segmentally identical form is also used as a marker of non-finite converb clauses in what is interpreted to be an extended function of the same morpheme. This clause chaining function is examined in §11.4.1. Lastly, it is noteworthy that a segmentally identical morpheme is found on verb stems in polar questions in non-irrealis settings, and is also used as a present tense marker in declarative clauses. It is likely that the agentive nominalizer also provided the diachronic source for this suffix, culminating in the crystallization of an absolute tense system out of an older system of mood oppositions. Cross-linguistic evidence in support of this suspicion is provided in §8.5.12.

7.4.5. Locative nominalizer

The locative nominalizer derives a nominal from a verb stem using the suffix *-tʃən*. The derivation expresses a meaning of ‘the place where VERB’. It is a highly productive process applying to both transitive and intransitive verb roots. Some examples are *tʃuŋ-tʃən* (eat.meal-LNOM) ‘eating place’, an *ad hoc* nominalization used to describe a marquee erected at a wedding reception to provide food for the guests, *jip-tʃən* (sleep-LNOM) ‘bed’ and *wa-tʃən* (go-LNOM) ‘path’. Stems of deverbal nominalizations expressing property concepts retain their prefixes when used to derive a locative nominalization, e.g. *tə-tʃha* > *tə-tʃha-tʃən* (NZIP-be.hot-LNOM) ‘disease’. As these derivations demonstrate, sometimes the locative nominalization results in a completely new meaning. This suggests that they are compound-like. They are not analyzed as compounds, however; firstly, because this nominal derivation seems to be a completely productive process, and secondly, because a morpheme *tʃən* is not found to occur in the language as an independent lexical item.

The use of the locative nominalization to form relative clauses is described in §6.6.4, and a diachronically related converb marker that is used to encode dependent clauses expressing a durative meaning is discussed in §11.4.3.5.

7.4.6. Instrument nominalization

The instrument nominalization affixes the nominalizing prefix *tə-* and the general nominalizer *-pàʔ* to the margins of a verb stem to create a nominal expressing ‘means of VERB-ing’. Sometimes idiosyncratic tone sandhi results from this derivational process. Following are some illustrative examples.

- (7.33) *nəmphaŋ* ‘covered’ → *tə-nəmphaŋ-pàʔ* (NZP-cover-NR) ‘cover’ (n.)
.iuj ‘burned’ → *tə-.iuj-paʔ* (NZP-burn-NR) ‘acid’
mətak ‘stung’ → *tə-mətak-paʔ* (NZP-sting-NR) ‘sting’ (n.)

Derived instrument nominalizations are treated like any other member of the nominal word class and can form compounds with other nominal constituents, e.g. *a-tsə tə-ətʃhət-pàʔ mənùk* (NRL-water + NRL-ladle-NR + cup) ‘water dipper’.

7.4.7. Purposive nominalization

There are four ways that verb roots can be used to form the non-finite complements of purposive clauses: by the root alone, e.g. (7.34); by derivation using the nominalizing prefix *tə-*, e.g. (7.35); by suffixing the purposive

nominalizer *-ɹu*, e.g. (7.36); or by suffixing the irrealis marker *-i* plus the general nominalizer *-pàʔ*, e.g. (7.37). The purposive nominalizer might be dedicated to this complementation strategy, as it is not found to have any other function in the grammar.⁶ Consultants confirm that these purposive complement-forming strategies are equivalent and allow all to be freely exchanged in textual examples without this incurring a change of meaning. Purposive complements are identified in the following examples by square brackets.

(7.34) *wàzàʔ tǝluʔ la tsəpàʔ məsəm wa.*

wàzàʔ tǝluʔ la [tsəpàʔ məsəm] wa
 bird all TOP well clean.PURP go.PST
 ‘All the birds went to clean a well.’

(7.35) *tǝku líʔfápáʔ nə kìmətshə à tǝhni zǝkùʔ tǝɹ.*

tǝ-ku líʔfápáʔ nə kìmətshə à
 thus-LOC.CV PN-M AGT dove.sp. one
 [tǝ-hni] zǝk-ùʔ tǝɹ
 NZP-follow send.PST-DEC REP
 ‘And then, Lichaba sent a Spotted Dove to spy on him.’

(7.36) *tǝ ahizala ʔʃu ləmsiɹù ʔʃhàɹu hlàwʔ.*

tǝ a-hiʔ-za-la ʔʃu [ləmsiɹù ʔʃhà-ɹu] hlà-ùʔ
 thus NRL-rat-DIM-F DIST comb pick.up-PURP descend+go.PST-DEC
 ‘So, Rat Pup went down to pick up the comb.’

(7.37) *nuksənsaŋpaʔ ʔʃu mùŋphu ku áhlù tsəjpaʔ áhlù nə wa.*

nuksənsaŋ-pàʔ ʔʃu mùŋphu ku
 PN-M DIST winter LOC
 [a-hlú tsə-ì-pàʔ] a-hlú nə wa
 NRL-field clear-IRR-NR NRL-field ALL go.PST
 ‘In the winter, Noksensangba went to the field to clear it’ (for swidden agriculture).

6. Although it is rather suspicious that the purposive nominalizer shares an identical segmental form with the immediate future suffix – see the discussion of §8.5.12 concerning the development of an absolute tense system in Mongsen.

7.4.8. Temporal nominalization

The temporal nominalization is expressed by a verb stem that is nominalized by the prefix *tə-* and suffixed by the irrealis marker *-i*, then case marked by the locative marker *ku*. A determiner may optionally intervene between the verb stem and the locative marker. The whole construction forms a type of temporal nominalization that expresses an adverbial meaning corresponding to ‘at the time of VERB’. Native speakers say that the optional inclusion of a nominal demonstrative emphatically points to the temporal location of the subordinated event. Presumably this is a relict left over from the grammaticalization of associated phrasal constituents (compare the optional inclusion of nominal deictics resulting from the reanalysis of an older function in §5.3.2.2 and §11.4.3.6).

- (7.38) a. *təsəj ku*
 tə-sə-ì ku
 NZP-die-IRR LOC
 ‘at the time of dying’
- b. *tətfəmi tʃu ku*
 tə-tʃəm-ì tʃu ku
 NZP-drink-IRR DIST LOC
 ‘at that time of drinking’

The presence of the irrealis marker in the stem of the dependent verb is essential for conveying the meaning of the temporal nominalization. If the irrealis suffix were removed from the stems of (7.38a–b), their meanings would change to ‘upon/while dying ...’ and ‘upon/while drinking ...’ respectively. I view this non-nominal function of the locative case marker as a type of converb, a non-final verb form used for clause linkage in complex sentences. The clause linking function of the locative converb is described in §11.4.3.7, and its closely related function as a discourse connective is described in §4.2.10. Example (7.39) below demonstrates the use of the temporal nominalization in a complex sentence.

- (7.39) *tətfhuwaj tʃu ku, ajila nə áwkla nə tʃhəniŋpa? təlu? tʃu ku tʃəmsilí? táŋ wa.*
 tə-tʃhuwa-ì tʃu ku a-ji-la nə a-úk-la nə
 NZP-emerge-IRR DIST LOC NRL-dog-F AGT NRL-pig-F AGT
 tʃhəniŋ-pà? təlu? tʃu ku tʃəm-si-lí? táŋ wa
 work-NR all DIST LOC step.on-RPET-SIM just go.PST
 ‘At the time of [Pig’s] returning, Dog just went and repeatedly trampled all over the work that Pig had done.’

7.4.9. Deictic nominalizations

A little used, minimally productive suffix *-kə* nominalizes a small number of intransitive verb roots that inherently encode a direction of movement away from the understood reference point. The resulting derivation expresses a spatial location that is distal to the location functioning as the deictic centre. The meanings of these nominalizations thus incorporate the specific deictic information that is fundamental to the semantics of their intransitive verb roots (see §8.1.1.2 for the description of directional motion verbs).

The forms of the deictic nominalizations are: *kəwa-kə* (ascend+go-DEIC) ‘up there’, *hlà-kə* (descend+go-DEIC) ‘down there’, *hja-kə* (level+go-DEIC) ‘over there’, *wa-kə* (go-DEIC) ‘out there’, *tʃuwa-kə* (emerge-DEIC) ‘outside there’ and *za-kə* (enter-DEIC) ‘inside there’. Note that these all encode a distal location. It is not possible to nominalize a directional motion verb whose inherent deixis encodes movement towards a proximate reference point. In other words, a derivation expressing ‘up here’ cannot be formed by nominalizing the root of *kəɹà* (ascend+come) with the deictic nominalizer.

Following are elicited examples in sentences. A rare text example can be found in (7.52). The example sentence of (7.41) demonstrates conclusively that this derivation produces a fully-fledged noun, attested by its ability to function as the head of a noun phrase modified by a demonstrative and a case marker. The diachronic source of the deictic nominalization suffix is unknown.

- (7.40) *pas stɪʃən hjakə li.ù?*
 pas stɪʃən hja-kə li-əɹ-ù?
 bus station level+go-DEIC be-PRES.HAB-DEC
 ‘The bus station is over there.’

- (7.41) *pa hlàkə tʃu nə hlàw?*
 pa hlà-kə tʃu nə hlà-ù?
 3SG descend+go-DEIC DIST ALL DESCEND+go-DEC.PST
 ‘S/he went down to there.’

7.4.10. The MAN nominalizer

A nominalizing suffix *-tʃhaŋ*, which appears to be either a constituent or a phonological reduction of the noun *apaŋtʃhaŋ-əɹ* (man-ANOM) ‘man, male person’, is used fairly infrequently for relativization. Matisoff (1988: 544) reconstructs an obviously cognate Proto-Lolo-Burmese form **tsaŋ* ‘person, man, human being’. The following two examples, demonstrating the use of *-tʃhaŋ* as a nominalizer, are taken from the Waromung village variety of

Mongsen. This morpheme is also found to function as a nominalizing suffix in other Mongsen varieties, generally in the speech of speakers less than forty years of age.

- (7.42) *nì na atʃhən khiʃhaŋ ánúʔtì-la si tʃhəna waukù.*
 nì na atʃhən khi-tʃhaŋ á-núʔtì-la si tʃhən-a
 1SG AGT money give-MAN NRL-woman-F ANAPH flee-SIM
 wa-ukù
 go-ANT
 ‘That woman I gave money to has run away.’

- (7.43) *pa mila kətʃhaŋ amiʔ tsə kàʔ ípáʔ mətəm tsə tʃha.ʉatáj li.*
 pa mila kət-tʃhaŋ a-miʔ tsə kàʔ
 3SG spirit possess-MAN NRL-person DIST also
 ípáʔ mətəm tsə tʃha.ʉ-a táŋ li-əɹ
 EMPHAT manner DIST be.sick-SIM just stay-PRES
 ‘The person who possesses its [= the tiger’s] spirit in this manner will also just remain sick.’

It appears that *-tʃhaŋ* has shed its original masculine lexico-semantic association when it is used to form relative clauses. This permits it to function as a nominalizer in the derivation *khiʔ-tʃhaŋ*, which modifies its noun phrase head *ánúʔtì-la* ‘woman-F’ in (6.42). Unlike the agentive nominalizer *-əɹ* (discussed above in §7.4.4), it is not constrained to deriving relativized attributes of nouns whose referents are agents, but all the examples encountered in texts suggest that it can only be used to modify nouns of the [+human] semantic class at its present stage of grammaticalization.

It is noteworthy that this morpheme is also used like a masculine semantic gender marker on some nouns with human referents, e.g. *anu-tʃhaŋ* (child-male) ‘boy’. Its range of uses suggests a remarkable parallel to some of the functions of the general nominalizer *-pàʔ* (§7.4.2), one of which is to mark masculine semantic gender on nouns with human referents (§7.5.2). Lastly, we can identify yet another correlation with some of the functions of the agentive nominalizer *-əɹ*, one of which is to encode masculine semantic gender in some lexical items.

7.5. Other nominal modifiers

The following sections discuss the forms and functions of six types of suffixes occurring on nominal stems: the intensifier suffix, the semantic gender suffixes,

the diminutive suffix, the terminative suffix, the reflexive suffix and the locational suffix.

7.5.1. Intensifier suffix

The intensifier suffix *-thì* is used as an emphatic marker of physical location, e.g. *mənaŋ-thì* ‘right in front’. This suffix is also used as an emphatic marker of temporal location, because in some situations *mənaŋ-thì* can imply ‘at the very beginning. Context determines whether an emphatic physical or temporal location is implied by the meaning of this suffix, as demonstrated in the following examples.

(7.44) *aki tjuŋthì ku, pi.*

a-ki tə-juŋ-thì ku pi
 NRL-house NZP-be.central-INTNS LOC PROX
 ‘It’s in the very middle of the house, this.’

(7.45) *tətʃhaɪ, təsəthì ku ahizala nə ahənzala tʃu tʃhənsətəɪ,*

tə-tʃhà-əɪ tə-sə-thì ku a-hiʔ-za-la nə
 thus-do-SEQ NZP-be.last-INTNS LOC NRL-rat-DIM-F AGT
 a-hən-za-la tʃu tʃhən-sət-əɪ
 NRL-chicken-DIM-F DIST wring-AFFECT-SEQ
 ‘And then, at the very end, after Rat Pup had strangled Little Chick, ...’

The intensifier suffix has a special marking function in derived adjectives that are inflected for the superlative degree (see §3.1.3.1 and §7.4.3.2 for description).

7.5.2. Semantic gender markers

There are two semantic gender suffixes: the masculine suffix *-pàʔ*, and the feminine suffix *-la*. Most but not all people’s names take a suffix according to their gender, e.g. *asaŋ-la* ‘Asangla’ and *sənti-pàʔ* ‘Sentiba’. Semantic gender markers are frequently omitted from names in casual conversation amongst intimates.

Ao names are usually associated with meaning; *sənti*, for instance, is also an intransitive verb root expressing the meaning of ‘be valued, precious’. In view of this, a semantic gender marker can be considered to have a nominalizing function. This is also consistent with the historical origin of the Mongsen

masculine gender suffix from PTB **pa* ‘father’ [STC #24] (see §4.2.1 for discussion of its grammaticalization as a pronoun, and §7.4.2 for description of the nominalizing functions of a cognate form). The feminine gender suffix *-la* has a less transparent etymological source, since the PTB feminine gender suffix **-ma* [< PTB **ma* ‘mother’ STC #487] reconstructed by Benedict (1972) appears to be unrelated to the synchronic Mongsen form. Nevertheless, its function is also nominalizing, because the above-mentioned name *asaŋ-la* is itself a nominal derivation from an intransitive verb root, viz. *asaŋ* ‘be illuminated’. This name might therefore be elaborately translated as a nominalization expressing ‘(the) illuminated (feminine) one’.

The noun phrase referents of animals are always personified as having feminine semantic gender in folklore stories and their nouns are accordingly marked with *-la*, paradoxically even if the same referent is addressed in direct speech with the vocative form of a masculine kinship term, such as *a-u* (VOC-grandfather) ‘Grandfather’, in the course of the narration. When a semantic gender suffix is used on a common noun with a human referent, e.g. *a-miʔ-pàʔ* ‘a man’, it specifies singular number and individuation of the referent. This is yet further evidence of the nominalizing function of the semantic gender markers.

Certain animals have gender-specific lexical terms. A domestic boar is *áwk təhla* and a domestic sow is *áwk təji*. The modifier expressing feminine gender in the latter example is the relational noun ‘mother’. The masculine gender term for all types of horned wild animals is *jaŋu* and the feminine gender term for horned wild animals is *tətsəʔ* or *tsəla*. Cocks are *təpuŋ* and hens are again *tətsəʔ*. Animals of cultural significance to the Ao, such as the Great Pied Hornbill, have their status acknowledged by specific terms. The cock is called *təhnəm wapaŋ* and the hen is called *təhnəm wətsə*.

A nominalized verb stem can be additionally marked for semantic gender by the masculine gender suffix *-pàʔ*. This is illustrative of the individuating function of the semantic gender suffix mentioned above.

- (7.46) ... *imtisàŋpàʔ*, *təkhàipàʔ*, *təsìipàʔ*, ...
 imtisàŋ-pàʔ tə-khà-əɪ-pàʔ tə-sì-əɪ-pàʔ
 PN-M NZP-have-ANOM-M NZP-know-ANOM-M
 ‘... Imtisangba, the man who was wealthy and learned, ...’

The use of the agentive nominalizer *-əɪ* as a marker of masculine semantic gender in a limited number of lexical items is discussed in §7.4.4 and illustrated with some examples.

7.5.3. Diminutive suffix

A diminutive suffix *-za* has transparently grammaticalized from the relational noun *tə-za* (RL-child) ‘child’ [< PTB **za* STC #59], which is no longer used independently as a lexical noun. It is now only encountered in the fossilized expression *tə-za tə-nu* (RL-child + RL-child) ‘children’. The citation form for ‘child’ is given as *anu*, suggesting an unusual shift from the relational class of bound noun. I offer a speculative diachronic explanation for this in §4.1.3.

The diminutive morpheme is fully grammaticalized as a nominal suffix and occurs on the nouns of both animate and inanimate referents of diminutive size, e.g. *ahən-za* (domestic fowl-DIM) ‘chick’, and *ajuŋ-za* (river-DIM) ‘stream’. In §5.3.3 it was demonstrated that many nascent postpositions have their sources in relational nouns, and that these formed the erstwhile heads of compound nouns. In the process of grammaticalization and the ensuing change in form class, the heads of these oblique noun phrases became reanalyzed as postpositions, while their erstwhile dependent constituents also underwent reanalysis to take their place as the new heads. By an analogous process, it seems logical that compounding has also provided a conduit for the grammaticalization of the diminutive suffix.

Mongsen lacks a morpheme that might unambiguously be identified as an antonymic augmentative, but one does appear to be in the process of grammaticalizing from the lexical noun ‘mother’. This is suggested by the function of *təji* in the following example.

- (7.47) *tə, nuksənsaŋpa? nə “u áhŋá? təji la ləja.ù? .ə.”*
 tə-əɪ nuksənsaŋ-pà? nə u a-hŋá? tə-ji la
 thus-SEQ PN-M AGT EXCLM NRL-fish RL-mother TOP
 ləja-əɪ-ù?
 descend+come-PRES-DEC PTCL
 ‘And, Noksensangba [said] “Ah! A huge fish is coming down [the river].”’

Matisoff (1991) demonstrates that the “mothermorph” is a common lexical source for an augmentative morpheme in languages of East Asia, South-East Asia and beyond, just as ‘child’ is the ubiquitous source of a diminutive morpheme. In (7.47), the meaning that *tə-ji* ‘RL-mother’ conveys is more in keeping with a modifying, metaphorical sense of ‘principal, major’, than with a meaning of ‘mother’. This is in spite of the fact that *təji* is juxtaposed to the right of another noun *a-hŋá?* ‘NRL-fish’, giving it the structural characteristics of a compound noun head in this noun phrase (e.g. compare the compound nouns described in §7.1 above). But it is clear from the context that the construction

áhǵá? tǵji cannot possibly be understood to mean ‘fish’s mother’ in this sentence. Thus, only a modifying augmentative meaning can be assigned to *tǵji*, even though it maintains the formal status of an independent word.

7.5.4. Terminative suffix

The terminative suffix *-lak* specifies the terminal part of the referent when it occurs on a noun stem. Some examples are *tǵ-mi-lak* (RL-tail-TERM) ‘tip of a tail’, *sǵntuǵ-lak* (tree-TERM) ‘apex of a tree’, *tǵ-mijuǵ-lak* (RL-finger-TERM) ‘fingertip’ and *mi-lak* (fire-TERM) ‘flame’. The historical source of this morpheme is almost certainly PTB **lak* ‘hand, arm’ [STC #86]. Like the diminutive suffix discussed above in §7.5.3, can we also attribute the grammaticalization of the terminative suffix to a compounding process, in which it formerly functioned as the relational noun head of a compound expressing ‘terminal part’? This does not seem an unreasonable assumption to make, given that *tǵ-lak* is also attested as a relational noun with a meaning of ‘last, end’. A cognate form is used as a lexical suffix on verb stems and is described in §8.5.1.10.

7.5.5. Reflexive suffix

A reflexive morpheme *-jimtǵ?* can be suffixed to any of the possessive pronouns to derive a reflexive meaning, e.g. *nǵ-jimtǵ?* (2SG.POSS-REFL) ‘yourself’. As is the case in many languages, the reflexive marker of Mongsen can additionally be used to encode an emphatic meaning.

A possessive pronoun marked by the reflexive suffix serves as an anaphoric element coindexed to an antecedent referent in the same clause. In transitive clauses expressing the reflexive meaning, it encodes that the referent that performs the activity denoted by the predicate is also the referent that undergoes the activity. The reflexive marker does not cause a rearrangement in valency because the clause remains transitive, requiring two participants. However, transitivity is in principle reduced by virtue of the fact that the agent acts upon itself. The O argument slot is filled by the reflexive-marked pronominal. It is often possible for the coindexed A argument to be elided if it is recoverable from the context.

The first example sentence below is taken from a Khensa Mongsen text (and hence has a different morphological form for the irrealis marker), and the second is elicited. These examples demonstrate the reflexive meaning of *-jimtǵ?*, in which the entity that is affected by the action is the entity that performs the action.

- (7.48) *nì thanila kəjimtà? sujsajt tʃhàkə sijukpì*
 nì thani la kə-jimtà? suəsajt tʃhà-əkə sə-juk-pì
 1SG now+day TOP 1SG.POSS-REFL suicide do-SIM die-PFV-IRR
 ‘Today I will kill myself by committing suicide.’

- (7.49) *nàŋ nəjimtà? .jàŋlak.*
 nəŋ nə-jimtà? .jàŋ-lak
 2SG 2SG.POSS-REFL cut-DESCEND.PST
 ‘You cut yourself.’

The next two sentences show how the reflexive morpheme can be used to express an emphatic meaning. This is equivalent to the emphatic use of the English reflexive, e.g. *she killed it herself*, versus the purely reflexive *she killed herself*. The possessive pronoun marked by the reflexive marker occurs immediately after the noun phrase it emphasizes.

- (7.50) *pa nə pajimtà? məphəlàtʃhətaj,*
 pa nə pa-jimtà? mə-phəlà-tʃhət aj
 3SG AGT 3SG.POSS-REFL NEG-unravel-ABIL.PST SINCE
 ‘... because he was not able to unravel [it] himself, ...’

- (7.51) *ahizala tʃu pajimtà? pa saləm tʃu múkhúŋ ma ku tshəkəkə tʃà?*
 a-hi?-za-la tʃu pa-jimtà? pa saləm tʃu
 NRL-rat-DIM-F that 3SG.POSS-REFL 3SG share DIST
 mukhuŋ ma ku tshək-əkə tʃà?
 tobacco.pipe FACE LOC roast-SIM consume.PST
 ‘Rat Pup roasted her share [of fish] over a tobacco pipe and ate [it] herself.’

7.5.6. Locational suffix

The locational suffix *-zə* is used with a small number of relational body part nouns to specify the approximate position of a referent relative to a reference point. This suffix may only be relevant to some varieties of Mongsen, because a native speaker of Longchang Mongsen reported that it is not used in the variety spoken in his village.

The only example occurring in the narrative corpus is provided below in (7.52). Some additional examples were elicited, e.g. *tə-ma-zə* (RL-face-LOC.SUFF) ‘towards the top’ and *tə-hlaŋ-zə* (RL-bottom-LOC.SUFF) ‘towards the bottom’, suggesting that it can be used productively with relational body part

nouns. The dearth of examples in the text corpus does not allow any forceful inferences to be made with respect to its function, but it appears that relational body part nouns marked by this suffix are used like locative adverbials to indicate the general location of an entity.

- (7.52) *hlàkə, təpukzə in məphù ku,*
 hlà-kə tə-puk-zə ìn məphù ku
 descend+go-DEIC RL-stomach-LOC.SUFF PATH ward LOC
 ‘Down there, right in the middle of the ward, ...’

Note that the text example of (7.52) is preceded by a deictic nominalization. An example offered by a consultant in explanation of its meaning was also paired with a deictic nominalization, viz. *za-kə tə-tsəŋ-zə* (enter-DEIC RL-inside-LOC.SUFF) ‘inside there’, suggesting that this may be a convention of its use (see §7.4.9 for description of deictic nominalizations).

7.6. Derivational morphology of numerals

Distributives are derived from the cardinal numerals by a partial reduplication of the base. Ordinal numerals and numeral adverbs are derived by attaching suffixes to the stems of cardinal numerals.

7.6.1. Derivation of distributive numerals

The forms of the distributive numerals are listed in Table 7.1 below. Distributives are derived by reduplicating the final syllable of a corresponding cardinal numeral (these are listed in Table 4.11). The distributive for ‘one each’ is irregular and uses a base that is unrelated to the form for the cardinal numeral ONE *akhət(a)*. A consultant aged seventy informs me that in earlier times the distributive ‘one each’ used to be pronounced as *khaŋ-ləŋ-ləŋ*. The obsolete distributive also has little in common with the phonological form of the cardinal numeral *akhət(a)*. The dental stop coda in the root of *a-nət* (NRL-two) ‘two’ undergoes complete anticipatory assimilation to a following dental nasal in the formation of the distributive, e.g. *a-nən-nət* (NRL-two-RED) ‘two each’. Unusual assimilatory processes that are unique to derivations involving the word for two are described in §2.6.5.

Another irregularity in the distributive derivation is demonstrated by the distributive form for ‘eight each’. A consultant pronounced this with three syllables in careful speech, i.e. [tshə³³ə³³ət³³]. This appears to be a reduplication of a reduplication, but only of the rhyme. The underlying form is probably

tshət-ət-ət and the word-internal dental stops are subject to syncope in the output. Seeing that *tshət* EIGHT is the only monosyllabic cardinal numeral, the double reduplication may be motivated by a constraint on the minimal number of syllables in a distributive form, or else it could be attributable to analogical levelling (cf. the impact of analogical levelling in the system of primary cardinal numerals in §4.2.7.3).

Table 7.1. Distributive numerals

FORM	SEMANTICS	FORM	SEMANTICS
<i>taŋ-laŋ-laŋ</i>	‘one each’	<i>thəja-ɬə thuku-khu</i>	‘nineteen each’
<i>a-nən-nət</i>	‘two each’	<i>məki-ki</i>	‘twenty each’
<i>a-səm-səm</i>	‘three each’	<i>səmja-ja</i>	‘thirty each’
<i>phəli-li</i>	‘four each’	<i>lija-ja</i>	‘forty each’
<i>phaŋa-ŋa</i>	‘five each’	<i>thənəm-nəm</i>	‘fifty each’
<i>təjuk-juk</i>	‘six each’	<i>jukja-ja</i>	‘sixty each’
<i>thəni-ni</i>	‘seven each’	<i>nɪja-ja</i>	‘seventy each’
<i>tshə-ət-ət</i>	‘eight each’	<i>tshət-mən zəŋ-lək</i>	‘eighty each’
<i>thuku-ku</i>	‘nine each’	<i>thuku-pən zəŋ-lək</i>	‘ninety each’
<i>thəja-ja-ja</i>	‘ten each’	<i>nuklən akhət-khət</i>	‘one hundred each’
<i>thəja-ɬə a-khət-khət</i>	‘eleven each’	<i>nuklən a-nən-nət</i>	‘two hundred each’
<i>thəja-ɬə a-nən-nət</i>	‘twelve each’	<i>məɪt(ŋ) akhət-khət</i>	‘one thousand each’

Apart from the irregular distributive form of EIGHT, the primary numerals share the same pattern of reduplication of the final syllable. Distributive forms of the TEEN numerals use the linking morpheme *ɬə* ‘and’, which is also found in the TEEN cardinal numerals. This, I suspect, is a metathesized form of the sequential converb suffix *-ət* that is used for dependent clause chaining (see §11.4.1). An extended function of the sequential converb is the discourse connective, which expresses the meaning of ‘and (then)’, and is used to link antecedent independent clauses to the following discourse. The linking morpheme of numeral compounds is discussed in §4.2.7.3 and compared to the discourse connective described in §4.2.10.

Distributive derivation of the higher round numbers from TWENTY through to SEVENTY demonstrates the regular pattern of final syllable reduplication. The pattern then changes radically for *tshət-mən zəŋlək* ‘eighty each’ and *thuku-pən zəŋ-lək* ‘ninety each’. I am not sure of the meanings of the morphemes in these constructions, but am able to offer the following observations. The formative *-mən* ~ *-pən* attaching to the numeral stem could be related to the verb root *pən*

‘complete’. Recall that this is also used as the linking morpheme in the numeral compounds of the obsolete overcounting numeral system. Numeral adverbs are derived with a segmentally identical morpheme *-pən* (see §7.6.3). As for *zəŋ-lək*, this corresponds with a verb root *zəŋ* ‘count’, while *-lak* could be either the ‘terminative’ lexical suffix (§8.5.1.10) or the homophonous DESCEND lexical suffix, although neither of these offers a convincing semantic fit with the distributive meaning. Distributive derivation of higher whole numbers from ‘one hundred each’ onwards resumes the pattern of reduplicating the final syllable. These are like quantified nouns, in which HUNDRED is the head and the partially reduplicated primary numeral is the quantifier.

7.6.2. Derivation of ordinal numerals

Ordinal numerals are derived by attaching the suffix *-pùpàʔ* to the stem of the cardinal numeral, e.g. *a-nəp-pùpàʔ* ‘second’ *a-səm-pùpàʔ* ‘third’, and so forth. As with the distributives, the base of the derivation corresponding to the cardinal numeral ONE, i.e. *mənaŋ-pùpàʔ* ‘first’, has no phonological relation to the cardinal numeral. The root *mənaŋ* can be used with both spatial and temporal reference with the meanings of ‘front, before’ (see §11.5). Given that cross-linguistically the common grammaticalization pathway is from space to time, we can assume that the relational/spatial meaning of *mənaŋ* is prior. The origin and meaning of the penultimate syllable *pù* in these derivations is unknown, but the final syllable *pàʔ* of the ordinal suffix is highly likely to be the ubiquitous general nominalizer discussed in §7.4.2. Ordinal derivations might in fact be recognized as yet another type of nominalization, with a form such as *mənaŋpùpàʔ* understood as expressing ‘the front one’ or ‘the one who is first’.

7.6.3. Derivation of numeral adverbs

Numeral adverbs are derived by attaching the suffix *-pən* to the stem of a cardinal numeral. The numeral adverb *khən* ‘once’ is irregular, in common with both the distributive and ordinal derivations for ONE, and does not use this suffix to express its adverbial meaning. Apart from the anticipatory assimilation that affects the dental stop of the derivation of *a-nət* TWO in *à-nəp-pən* ‘twice’, the derivations of the remaining numeral adverbs are completely regular, e.g. *à-səm-pən* ‘thrice’, *phəli-pən* ‘four times’, *thəɹa nət-pən* ‘eleven times’ and so forth. The derivational morpheme *-pən* consistently lowers the underlying mid tone(s) of the primary numeral stems. It was noted above that this suffix is segmentally identical to the verb root *pən* ‘be complete’, which is only used in

the context of quantification. Its semantics suggest that it has provided the diachronic source for the grammaticalization of the numeral adverb derivation.

Chapter 8

Verbs and verbal morphology

This chapter focuses on the verbal morphology of Mongsen. It commences with a discussion of verb classes, followed by a description of the morphological structure of the predicate. The verbal grammatical categories of each slot and the types of meanings their morphemes express are then described in detail according to their position of occurrence in the predicate.

8.1. Verb classes

Verbs can be divided into intransitive and transitive classes on the basis of syntactic criteria and occur with roughly the same frequency in the lexicon. In a count of 316 elicited verbs based on the comparative vocabulary of Benedict (1972), 155 were intransitive and 161 were transitive, suggesting a very even division between the two major transitivity classes in this restricted corpus.

Recall that the past tense is the unmarked absolute tense category in Mongsen and was chosen as the citation form for verbs. In the following description, glosses and translations of examples that constitute bare roots reflect their inherent past tense, e.g. *mən* (sit.PST) ‘sat’. Henceforth when I refer specifically to the lexical meaning of a verb root, the example is preceded by the symbol √ and the translation is given without temporal specification, e.g. √*mən* ‘sit’.

8.1.1. Intransitive verbs

A strictly intransitive verb can only occur as the head of a verbal clause with a single core argument. Some examples are intransitive verbs of motion, such as *za* ‘entered’, *kəja* ‘came up’ and *tʃəli* ‘walked’; intransitive stative verbs such as *tʃəsi* ‘was distressed’, *átshəkhəm* ‘was boastful’ and *ahjak* ‘was shy’; and intransitive verbs denoting corporeal processes, such as *sə* ‘died’, *tʃəp* ‘wept’ and *sənsi* ‘breathed’.

Some semantic evidence for the existence of two subclasses of intransitive verb is suggested by the meanings of reduplicated roots, as discussed below in §8.1.4. Reduplicated stative intransitive verbs express an intensive meaning in simultaneous converb constructions, e.g. *məjəm-məjəm-əkə* (red-RED-SIM) ‘reddening’, while reduplicated intransitive and transitive activity verbs express

the prolonged duration of an activity, e.g. *tʃhùwa-tʃhuwa-kə* (emerge-RED-SIM) ‘coming, coming out’.

8.1.1.1. *Stative intransitive verbs*

Most semantic classes of stative intransitive verb roots can function as ascriptive complements in verbless clauses or as attributes of their noun phrase heads once they have been nominalized by the nominalizing prefix *tə-* (see §6.4.1 and §7.2.4). Alternatively, stative verbs can be used without undergoing nominalization if the state they predicate is of a temporary nature, or if there has been a recent entry into that state. These two uses are respectively demonstrated by the elicited examples of (8.1a–b) below.

- (8.1) a. *nì tətʃhu təkun.*
 nì tə-tʃhu təkun
 1SG RL-throat NZP-be.dry
 ‘My throat is dry.’
- b. *nì tətʃhu kunʃfukukù.*
 nì tə-tʃhu kunʃfuk-ukù
 1SG RL-throat be.dry-PFV-ANT
 ‘My throat has become dry.’

Deverbal adjectives derived from stative intransitive verbs behave much like the adjective class of English, with the exception that they can be used as complements of verbless clauses. This function is described in §9.1.4, and their use as modifying attributes in noun phrases is described in §6.4.1.

8.1.1.2. *Directional motion verbs*

Mongsen has an elaborate sub-class of intransitive directional motion verbs. These very specifically encode an entity’s path of movement in a three-dimensional space, in addition to encoding its direction of movement relative to a deictic centre. The historical development of this highly specialized set of verbs is a linguistic reflection of the mountainous environment in which the Ao live. Similar systems are reported in other languages, such as the Yimas language of New Guinea (Foley 1991), and are likewise motivated by geographical influences. The following six specialized verbs are used to express directional movement in Mongsen:

- | | | | | |
|-------|----|-------------------------------|--------------------|----------------------------|
| (8.2) | A. | <i>kəwa</i> | (ascend+go.PST) | ‘went up’ |
| | B. | <i>kəja</i> | (ascend+come.PST) | ‘came up’ |
| | C. | <i>hlà</i> | (descend+go.PST) | ‘went down’ |
| | D. | <i>ləja~lala</i> ¹ | (descend+come.PST) | ‘came down’ |
| | E. | <i>hja</i> | (level+go.PST) | ‘went across’ (same level) |
| | F. | <i>hija</i> | (level+come.PST) | ‘came across’ (same level) |

In contexts where the encoding of path in a specific vertical or horizontal plane is redundant, two verbs with meanings that are limited to expressing motion relative to a deictic centre only are used:

- | | | | |
|-------|-----------|------------|--------|
| (8.3) | <i>ja</i> | (come.PST) | ‘came’ |
| | <i>wa</i> | (go.PST) | ‘went’ |

The meanings encoded by the six directional motion verbs are represented schematically in Figure 8.1 below, with reference to A–F listed in (8.2) above.

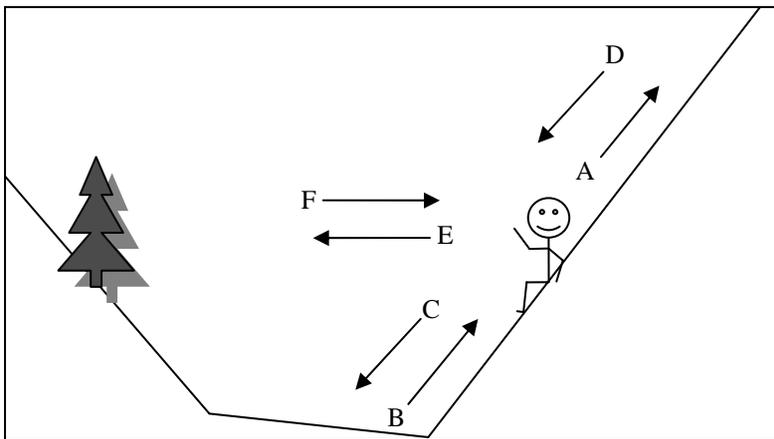


Figure 8.1. Directional motion verbs.

Frozen formatives encoding specific information are transparent in some of these verb roots; for example, *kə* of *kəwa* ‘went up’ and *kəja* ‘came up’ encodes ascent, while the *ja* of *kəja* ‘came up’, *ləja* ‘came down’ and *hija* ‘came across’ is obviously related to the root of the deictic motion verb *ja* ‘came’. The phonological form of the initial syllable of *hija* furthermore suggests a proximate demonstrative source and thus correlates with the deictic orientation of movement expressed by the formative *ja* in the verb root. However, as these

1. These alloforms occur in free variation in the speech of some Mongsen speakers.

are all now fossilized, they are no longer recognized as independent morphemes by native speakers, or for that matter by the grammar. The two formatives recognizable in *kə.ɪa*, for instance, are treated as a single morpheme when negated, e.g. *mə-kə.ɪa-la* (NEG-ascend+come-NEG.PST) ‘didn’t come up’.

The intransitive motion verb *tʃhuwa* ‘emerged’ (from a field, from the jungle, at a house) is superficially similar, both in expressing motion and in having the fossilized formative *wa* as part of its root. It differs in being unspecified for deixis. This allows it to be used with various deictic orientations that are determined by context, e.g. ‘went out’ or ‘came out’. The fundamental sense it encodes is movement from within to without, so I have chosen to gloss it uniformly as ‘emerge’. Speakers do not recognize an independent meaning associated with the syllable *tʃhu* of the root, and nothing that could be semantically related to its phonological form is encountered in the corpus.

Lastly, the intransitive motion verb *za* ‘entered’ functions as the antonym of *tʃhuwa* by expressing movement from without to within. It, too, is unspecified for a particular deictic orientation, consequently allowing it to express movement both towards and away from an understood deictic centre.

8.1.2. Transitive verbs

Transitive verbs can (although do not necessarily only) occur in verbal clauses with two core arguments. Some examples are: transitive activity verbs of effect, in which the semantic patient undergoes a change of state, such as *ɹuŋ* ‘burned’, *hənsət*, ‘killed’ and *tʃâ?* ‘ate’; transitive verbs of manipulation, such as *thəp* ‘threw’, *həmsi* ‘used’ and *hən* ‘took, carried’; transitive verbs of perception, such as *hùŋ* ‘saw’, *jâ?* ‘heard’ and *hŋa* ‘listened’; and transitive verbs of cognition, such as *si* ‘knew (how to do something)’, *alak* ‘forgot’ and *maŋ* ‘believed’.

One factor that somewhat devalues the parameter of transitivity as a precise criterion for verb classification is that a large number of transitive verbs demonstrate ambitransitivity (see §5.2.4, and §8.1.3 below). Other transitive verbs – even those that do not have a demonstrated ambitransitive function – are also frequently used with one or both of their core arguments omitted, especially if the explicit mention of an argument is rendered superfluous by a particular pragmatic context. This leads one to question the relevance of invoking the notion of transitivity in a description of Mongsén. On the other hand, it is at least useful for differentiating strictly intransitive verbs from other transitivity classes, because whereas a large number of transitive verbs may be used with just one core argument, no intransitive verb may be used with two core arguments unless it is marked for a change in valency by the causative suffix (see §5.4.1).

The following example demonstrates how the core argument of a transitive verb can be freely omitted if its referent is pragmatically recoverable from the discourse. In (8.4), the transitive verb *khi?* ‘give.PST’ is used with its O argument omitted. Ellipsis of this argument is licensed by an explicit antecedent mention of its referent in the narrative the example is taken from – two clauses before (8.4) is uttered, Rat Pup implores Little Chick to give her a prawn.

- (8.4) *təɪ, ahənzala tʃu nə khi?*
 tə-əɪ a-hən-za-la tʃu nə khi?
 thus-SEQ NRL-chicken-DIM-F DIST AGT give.PST
 ‘And so Little Chick gave [a prawn (to Rat Pup)].’

Unlike a single-argument clause with an ambitransitive verb that can be used without a qualifying context, the sentence of (8.4) would leave an interlocutor in doubt as to what was given if this clause were not preceded by an antecedent mention of the referent of the omitted core argument. This is therefore not an example of agentive ambitransitivity, but one of zero anaphora. Superficially similar ambitransitivity is described in the following section.

There is probably little descriptive advantage to be gained from recognizing a separate syntactic class of “ditransitive” verb in Mongsen, for the same reasons as those raised above. While native speakers will readily construct elicited sentences with one oblique and two core arguments for verbs such as *khi?* ‘gave’, *hli* ‘bought’ and *zək* ‘sent’, there is no obligatory grammatical requirement for such predicates to occur with all three arguments, and examples with a lesser number are just as common in texts. In the textual example of (8.5) below, for instance, *nəŋkhəla thijà?* ‘your wages’ can only be interpreted as a possessive noun phrase in O function, according to the intuitions of consultants, therefore this clause has a syntactic valency of two. But if *nəŋkhəla* and *thijà?* happened to be separate arguments, then *nəŋkhəla* would require dative case marking to identify it syntactically as the recipient instead of a possessive modifier of the O argument, as in (8.6), resulting in a surface syntactic valency of three. Rampant zero anaphora and the use of verbs in ambitransitive function thus create the frequent potential for a mismatch between semantic and syntactic valence.

- (8.5) *aw nə nəŋkhəla thijà? khəjù.ə.*
 a-u nə nəŋkhəla thijà? khi?-ì-ù? ɪə
 VOC-grandfather AGT 2PL wages give-IRR-DEC PTCL
 ‘Grandfather will give your wages.’

- (8.6) *aw nə nəŋkhàla li thijà? khì?-ì-ù? ɿə*
 a-u nə nəŋkhàla li thijà? khì?-ì-ù? ɿə
 VOC-grandfather AGT 2PL DAT wages give-IRR-DEC PTCL
 ‘Grandfather will give wages to you all.’

When examples of the verb *khi?* ‘gave’ are elicited with three participants, consultants mark the donor with agentive case-marking, the theme argument is distinguished by the absence of case marking, and the recipient takes dative case marking, as we saw above in (8.6). However, in naturally spoken language I sometimes find that it is not only the dative case that can be used to mark the recipient argument of *khi?*. In (8.7), for example, the goal of the theme is distinguished by allative marking. I am not certain if allative marking is used in place of dative marking here because the recipient is unaware of receiving the “gift” of the poison. Nevertheless, it is highly significant that there is no syntactic requirement for a recipient argument of *khi?* to be specifically marked by the dative case if an oblique argument is added to the clause.

- (8.7) *təku pa nə təhjaŋaj jìmlà sə nə hjaŋaj.ɿə sə nə muli khì?-əɿ sə-juk-i?-ì-pà?*
 tə-ku pa nə tə-hjaŋaj jìmlà sə nə
 thus-LOC 3SG ALL RL-lover friend ANAPH ALL
 hjaŋaj-əɿ sə nə muli khì?-əɿ sə-juk-i?-ì-pà?
 lover-ANOM ANAPH AGT drug give-SEQ die-PFV-CAUS-IRR-NR
 ‘And so, the aforementioned lover gave poison to him, to the
 aforementioned lover friend, to cause him to die.’

Furthermore, seeing that sentences such as these are structurally indistinguishable from clauses formed with ordinary monotransitive verbs taking an optional oblique argument, the identification of a ditransitive verb class, then, is reduced to relying upon a native speaker’s intuitions of how many participants are underlyingly represented in the argument structure of a particular verb, irrespective of how many are overtly realized. In contrast to languages such as English, syntactic tests such as dative shift or passivization cannot be applied to the non-agentive arguments of verbs such as ‘give’ in Mongsen to augment a native speaker’s intuitions.

The only situation in which an oblique-marked argument can truly be considered a core syntactic argument of the clause is when a dative-marked noun phrase is used to mark the causee of causativized transitive clauses, but this increase in valency is merely a corollary of the causative derivation, rather than an accurate indication of underlying argument structure (see §5.4.1.2). These observations leave me hesitant to acknowledge any additional categories

of verb beyond intransitive and transitive classes, and even these have a somewhat limited utility.

8.1.3. Ambitransitive verbs

Two types of ambitransitive verb are recognizable in Mongsen. Agentive ambitransitives have a core argument that corresponds to the A argument of a transitive clause, while patientive ambitransitives have a core argument that corresponds to the O argument of a transitive clause. The overwhelming majority of Mongsen transitive activity verbs function as ambitransitives of the patientive type. These are used with just one core argument to express a resultant state. As the language lacks a passive voice, the total suppression of the A argument provides an alternative means of promoting an O argument into pragmatic prominence by focusing on the effect of the action expressed by the predicate. Such intransitive clauses are most appropriately translated into English using the passive voice (see §5.2.4 for earlier discussion).

This is demonstrated by the following narrative examples, which occur a few clauses apart in a text. In (8.8a), the transitive verb *kəp-saʔ-zək-tʃuk* occurs with both A and O core arguments. In the sentence of (8.8b), the same verb functions as a patientive ambitransitive, because the A argument is absent. This allows its predicate to express the functional equivalent of a passive voice. The intransitive S argument of (8.8b) corresponds to the transitive O argument of (8.8a) sentence, and the resulting shift in pragmatic prominence to the semantic undergoer is achieved without any morphological modification of the verb stem, as would be expected in a language with a grammaticalized voice distinction.

(8.8) a. *nəŋ tʃútə̀ ùphəla həntsə̀ kəpsə̀zəkʃuk.*

nəŋ tʃútə̀ ùphəla həntsə̀ kəp-saʔ-zək-tʃuk
2SG why red.junglefowl egg strike-SEPARATE-SEND-PFV.PST
'Why did you crack and scatter Red Junglefowl's eggs?'

b. *ùphəla həntsə̀ tʃu kəpsə̀zəkʃuk.*

ùphəla həntsə̀ tʃu kəp-saʔ-zək-tʃuk
red.junglefowl egg DIST strike-SEPARATE-SEND-PFV.PST
'Red Junglefowl's eggs were cracked and scattered.'

Ambitransitivity must be differentiated from the ellipsis of core arguments, which can result in a superficially identical clausal structure (for examples of this, cf. [8.4] above and [5.19] of §5.2.1). We can be sure that the verb of the intransitive sentence of (8.8b) is indeed a member of the patientive ambitransitive class, because this is in fact the first mention of Red Junglefowl's

eggs in the narrative from which the examples are taken, and because such a statement is comprehensible in the absence of an antecedent mention of the agent of the action.

Note that both core arguments of (8.8a) receive no overt case marking, despite this sentence being a transitive clause. Mongsen is representative of those languages in which core case marking is not syntactically motivated.² Although the grammar possesses the morphological means to differentiate grammatical functions explicitly by using agentive case-marking on the A argument of transitive clauses and no explicit marking on the O argument, it is quite frequently found that both core arguments of transitive clauses are unmarked for case. In these situations, context and/or world knowledge must be relied upon to assign semantic roles correctly to noun phrase referents, particularly when both referents are animate entities. Pragmatically and/or semantically motivated marking of core arguments appears to be a shared characteristic of other Tibeto-Burman languages of the region and is reported to operate in Lahu (Matisoff 1973: 155–156; Matisoff 1976: 423–427) and in Meithei, also known in the literature as Manipuri (Chelliah 1997: 93; Bhat and Ningkomba 1997: 54–56.). Motivations for the explicit marking of core arguments in Mongsen are discussed in detail in §5.2.2.

The agentive ambitransitive verb is much rarer than the patientive type and is limited to just a handful of examples, e.g. *tʃã?* ‘consumed’, *tʃuŋ* ‘ate a meal’, *tʃəm* ‘drank’, *tʃhəniŋ* ‘worked’, *kùk* ‘won’ and *phù?* ‘stole’. The following examples demonstrate that the A argument of the transitive clause of (8.9a) corresponds to the S argument of the intransitive clause of (8.9b) when the same verb is used in ambitransitive function. As with patientive ambitransitive verbs, the agentive type is recognized by its ability to occur with just one core argument in the absence of a qualifying context.

- (8.9) a. *nì nə tùŋàɪ usət məphùtshəŋì.*
 nì nə tùŋàɪ usət mə-phù?-tshəŋ-ì
 1SG AGT other stuff NEG-steal-ATTACH-IRR
 ‘I won’t steal other’s belongings.’
- b. *nì nə məphùtshəŋì.*
 nì nə mə-phù?-tshəŋ-ì
 1SG AGT NEG-steal-ATTACH-IRR
 ‘I won’t steal.’

2. A consultant suggested that the A argument is not marked by the agentive case in this clause, because the character who asks the question of (8.8a) knows that the addressee is the one responsible for Red Junglefowl’s eggs getting cracked. See §5.2.2 for related discussion concerning agentive marking on the noun phrases of referents demonstrating socially-marked behaviour.

The patientive ambitransitive verb serves a crucial pragmatic function in Mongsen, because it equips the language with the ability to focus attention on an undergoer argument in the absence of a formal voice distinction (§5.2.4). In contrast, the agentive ambitransitive has a more limited pragmatic role, as there is much less need to focus on the agent by completely suppressing the patient. By virtue of its greater topicality, the agent of a transitive predicate is inherently more salient, particularly when it occurs in its usual position at the front of the clause. These factors account for the much smaller number of verbs belonging to the agentive ambitransitive type.

8.1.4. Reduplicated verb roots

Complete reduplication is used for intensifying the meanings of intransitive stative verbs, and for encoding the protracted duration of activities expressed by activity verbs. Reduplicated verb roots are typically, although not obligatorily, found in simultaneous converbs (§11.4.1.2); these dependent verb forms are mostly used to impart adverbial-like information encoding the manner or circumstances by which their matrix clause events or states transpire. The whole of the root is reduplicated and then marked stem-finally by the simultaneous converb suffix, e.g. *amu* ‘be.soft’ → *àmu-amu-kə* (be.soft-RED-SIM) ‘very softly’ and *pakla* ‘be.pale’ → *pàkla-pakla-kə* (be.pale-RED-SIM) ‘be extremely pallid’. A tonal alternation in the first syllable of these reduplicated examples is automatic for this type of stem if its underlying tone is mid.

Activity verbs similarly undergo complete reduplication in simultaneous converb constructions, but contrast with stative verb roots in expressing a protracted duration of activity. This word formation process also allows for the reduplication of aspectual suffixes on the verb stem, e.g. *ləp-si* (cut-RPET) → *ləp-si-ləp-si-kə* (cut-RPET-RED-RPET-SIM), lit. ‘repeatedly chopping, chopping, chopping’. As Mongsen has a preference for suffixing, it is assumed that the reduplicated form follows the base.

A few verb root reduplications have become lexicalized and no longer occur in the non-reduplicated form, e.g. *nəmnəm* (whisper.PST) ‘whispered’ *tə-səsə* (NZP-be.smooth) ‘smooth’ and *tə-luŋluŋ* (NZP-be.round) ‘round’.

8.2. Structure of the predicate

The minimal requirement for a predicate is just a verb root, which expresses past tense in its bare form (§8.5.12.1). As is typical of Tibeto-Burman languages generally, verbal morphology is predominantly suffixing. The total number of positions for grammatical categories in the verb template is eleven,

consisting of one prefix position, the root and potentially nine suffixing positions.

Two potential problems are presented by a position class analysis of predicate structure; firstly, not all suffixes are consistently found to occur in a set linear order, and secondly, some suffixes that should constitute a functionally unitary class can occur simultaneously in the same slot of the template, rather than paradigmatically. With respect to the first problem, it is important to note that, even for suffixes found to occur in variable orders, there will always be some semantic selectional restrictions determining the orders in which they occur. That is to say, suffixes cannot be swapped around willy-nilly in the verb stem without regard for the effect this has on meaning. Also, the number of suffixes that do have more than one position of occurrence is extremely limited. A decision has therefore been made to persist with a linear analysis, noting the occasions when a variable order is possible and the implications this has for a change in meaning.

The second problem of more than one member of the same class not occurring in paradigmatic opposition might be addressed by adopting the approach of Kari (1989), who posits “zones” to account for similar analytical problems in Athabaskan verb complexes. The modality suffixes of §8.5.5, for example, demonstrate the possibility of more than one class member co-occurring in a zone. But once again, semantic compatibilities holding between members of the class will determine what verbal categories can or cannot co-occur, and their order of occurrence in the zone always has implications for resulting meanings. Any potential combinations must therefore be licensed by the semantic possibilities of a particular linear order.

Ultimately it might prove more profitable to abandon a linear analysis and assume that verb stem formation proceeds in a cyclical fashion. This would suggest, in effect, that word formation does not proceed so much by item-and-arrangement (Hockett 1958), as by a selectional process, one dictated by the accumulative meaning expressed by the sum of previously selected verbal categories. In other words, the selection of a semantic class of verb root (a) opens up selectional possibilities for a particular set of suffixes (b). The combination of the verb root (a) with a suffix of class (b) subsequently has semantic selectional implications for the affixing of another class of suffix (c) in the next cycle, and so on. Conceivably this kind of approach, freed from the bondage of linear analysis, could exhaustively account for all the selectional possibilities and affix orders that the Mongsen verb stem could present, and could even account for members of the same category occurring together if they happen to be semantically compatible. Appealing as such an approach might be, it is beyond the scope and resources of the present work and therefore will not be attempted.

The following table presents the order of occurrence of suffixes from left to right in the verb template, and provides cross-references to sections discussing the grammatical categories occurring in each of the eleven positions.

Table 8.1. Morphological structure of the predicate (an *asterisk denotes a “zone”, after Kari 1989).

1.	prohibitive mood prefix (§8.3.1), admonitive mood prefix (§8.3.2), negative prefix (§8.3.3), nominalizing prefix (§8.3.4)
2.	Root (§8.4)
3.	*Lexical suffix zone (§8.5.1)
4.	Reciprocal/collective suffix (§8.5.2)
5.	Directional suffix (§8.5.3)
6.	Aspectual suffixes (§8.5.4)
7.	*Modality suffix zone (§8.5.5)
8.	Resultant state suffix (§8.5.6)
9.	Perfective aspect suffix (§8.5.7)
10.	Causative suffix (§8.5.8)
11.	Negative suffix (§8.5.9), positive imperative suffix (§8.5.10), converb suffixes (§8.5.11), tense/mood suffixes (§8.5.12), general nominalizer suffix (§8.5.13)

A number of co-occurrence restrictions ensure that the maximum number of slots that can be filled in any one predicate is limited to considerably less than the eleven discrete positions recognized in this table. The verb stem of (8.10) below presents six morphemes occurring on the one verb stem. While an even more morphologically complex verb stem could be artificially created and might be accepted as being grammatical, a stem with six morphemes represents the usual upper limit encountered in real speech, with the majority of verbal predicates commonly having from two to four morphemes in their stems. The final optional declarative mood clitic (§4.2.14.6) has a clausal distribution and attaches to the final constituent. It is therefore not considered to be part of the verbal morphology.

- (8.10) *məhən-laktsələw?*
 məh-hən-lak-tsə-la-ù?
 NEG-take-DESCEND-CON-NEG.PST-DEC
 SLOT: 1 2 3 7 11
 ‘[She] didn’t try to carry [it] down’

The remaining sections of this chapter describe the various grammatical categories occurring in each slot of the template and discuss their co-occurrence restrictions with respect to other verbal categories.

8.3. Verbal prefixes

Three verbal prefixes – admonitive mood *asáʔ-*, prohibitive mood *tə-*, and negative *mə-* – occur in paradigmatic opposition in the sole prefix slot of (1). This is the position of the nominalizing prefix *tə-* as well, which has been described in §7.4.1.

The nominalizing prefix and the negative prefix are able to occur together in undergoer nominalizations. I treat this as a nominalization of a negated verb stem that is only possible in this specific derivation. The concatenation of synchronically paradigmatic morphemes is probably attributable to historical stratification processes involving the reanalysis of the negative affix as part of the base. A good example of this is the now-frozen deverbal adjective *tamáu* ‘bad’, the underlying structure of which is recognizable as the nominalizing prefix fused with the negative prefix and the root *au* ‘good’. Since the combination of *tə-* and *mə-* cannot otherwise occur productively on verb stems except in patientive nominalizations, the establishment of separate prefix slots for each of these morphemes is not justified.

The verbal prefixes qualify as inflectional morphemes, because (i) they occur in paradigmatic opposition, (ii) they can be prefixed to any member of the lexical class of verb, (iii) their meanings are semantically regular, and (iv) their prefixation produces new word forms of the same lexeme. Problems concerning the application of traditional criteria for distinguishing between derivation and inflection with respect to a number of Mongsen verbal suffixes are discussed in §8.5.

8.3.1. Prohibitive mood

The prohibitive mood prefix *tə-* occurs on the stems of both transitive and intransitive verbs, including those of stative intransitives. It is described in greater detail in §10.2.2.

While the mood of commands generally demonstrates more incompatibilities with other grammatical categories than the declarative or interrogative mood types, from a typological perspective the Mongsen prohibitive mood seems less fastidious about the morphological company it keeps. Its prefix can share the verb stem with the lexical suffixes of slot (3), the reciprocal/collective suffix of slot (4), the directional suffix of slot (5), the aspectual suffixes of slot (6) with

the exception of the continuative and the habitual aspect markers, some modality suffixes of slot (7), the causative suffix of slot (10) and the irrealis suffix of slot (11). Despite these possibilities, it is unusual for the prohibitive to co-occur with affixes of other verbal categories and the vast majority of examples are found with just the verb root. This lack of morphological complexity might be attributable to the cross-linguistically valid observation that commands generally tend to be pithy.

- (8.11) *awŋ ku tətʃəlì.*
 a-uŋ ku tətʃəlì
 NRL-jungle LOC PROH-walk
 ‘Don’t wander around in the jungle.’

8.3.2. Admonitive mood

The admonitive mood prefix *asáʔ-* is used to express dissuasion, rather than outright prohibition, and carries an equivalent force to ‘Don’t VERB too much’. It is given greater attention in §10.2.3. As might be expected, the admonitive category exhibits the same incompatibilities and co-occurrence possibilities as the prohibitive mood prefix.

- (8.12) *nàŋ aki tʃu nə asáʔ-wa.*
 nəŋ a-ki tʃu nə asáʔ-wa
 2SG NRL-house DIST ALL ADM-go
 ‘Don’t you go to that house too much.’

8.3.3. Negative prefix

The negative prefix *mə-* is subject to the fewest collocational restrictions with other verbal categories, the following being identified. Firstly, an incompatibility occurs with the positive imperative marker of slot (11) because the prohibitive prefix is already dedicated to encoding the negated imperative meaning, therefore it must occupy the same initial slot in the stem as the negative prefix. Secondly, an incompatibility is found with the immediate future tense suffix of slot (11). Explanations for the non-occurrence of the negative with immediate future tense are offered in §8.5.12.2. Thirdly, negation is incompatible with the continuative aspect marker (§8.5.4.1) of slot (6) and some of the converb suffixes of slot (11), specifically those involved in dependent clause chaining (see §11.4.1).

The negative suffix *-la* (§8.5.9) must occur in slot (11) in combination with the negative prefix of slot (1) when the temporal setting is in the past. The two negative morphemes could therefore be analyzed together as a discontinuous morpheme straddling the verb stem and functioning as one unit, but only for negation in past tense. The negative suffix never co-occurs with the present tense or irrealis marker of slot (11), nor with the converb suffixes of slot (11). This restriction is demonstrated by the presence of the negative suffix in a verb stem specified (by default) for past tense in (8.13), as opposed to its obligatory absence in negated verb stems marked by the present tense suffix, the irrealis suffix and the conditional converb suffix in (8.14–8.16) respectively.

- (8.13) *tʃuniphitʃünə, hənsəlila sə nə ahən təpuŋla mətsəphala.*
 tʃuniphitʃünə hənsəli-la sə nə a-hən təpuŋ-la
 from.that.day leopard.cat-F ANAPH AGT NRL-fowl cock-F
 mə-tsep̄ha-la
 NEG-fear-NEG.PST
 ‘From that day onwards, [aforementioned] Leopard Cat did not fear Rooster.’

- (8.14) *tətʃhəɪ, əja tʃhəkə mətʃhuwaɪ.*
 tətʃhə-əɪ ə-ja tʃhəkə mə-tʃhuwa-əɪ
 thus-do-SEQ NRL-night reach-SIM NEG-emerge-PRES
 ‘And then, night comes and [she] doesn’t return.’

- (8.15) *təku təpa? nə “u məphùjù?”*
 tə-ku tə-pa? nə u mə-phù?-ì-ù?
 thus-LOC.CV RL-father AGT EXCLM NEG-steal-IRR-DEC
 ‘And so, the father said “Oh, I won’t steal.”’

- (8.16) *mətʃhàpàla aki ku tsəŋla? la-ì-ù?*
 mə-tʃhà-pàla a-ki ku tsəŋla? la-ì-ù?
 NEG-do-COND NRL-house LOC lightning strike-IRR-DEC
 ‘Or if [you] don’t do that, then lightning will strike the house.’

8.3.4. Nominalizing prefix

The nominalizing prefix *tə-* is a versatile derivational morpheme whose function is to nominalize verb roots. This derives a deverbal nominal that can then be used as:

- (i) a verbless clause complement (§9.1.4)
- (ii) an attributive modifier of a noun phrase head (§6.4.1, §7.2.4)
- (iii) an adjective of degree (§7.4.3), which also requires the presence of the general nominalizer (§6.6 and §7.4.2)
- (iv) a purposive nominalization (§7.4.7)
- (v) an instrument nominalization (§7.4.6), which also requires the presence of the general nominalizer

The five uses of the nominalizing prefix listed above are respectively illustrated by examples (8.17–8.21) below.

(8.17) *anu la taua.*

a-nu la tə-a.u-a
 NRL-child TOP NZP-be.good-VSF
 ‘As for the children, they’re good.’

(8.18) *jipʃən təməɹaŋ*

jip-ʃən tə-məɹaŋ
 sleep-LNOM NZP-be.hard
 ‘hard bed’

(8.19) *nì səntipà? thən la təhlaŋpa?*

nì sənti-pà? thən la tə-hlaŋ-pà?
 1SG PN-M COM TOP NZP-be.long-NR
 ‘I am taller than Sentiba.’

(8.20) *tətsəŋtsə walikà?*

tə-tsəŋtsə wa-likà?
 NZP-visit go-CONTEMP
 ‘When [she] had gone to visit, ...’

(8.21) a. *tʃhì*

tʃhì
 shave.PST
 ‘shaved’

b. *təʃhipa?*

tə-ʃhì-pà?
 NZP-shave-NR
 ‘razor’

As the examples of (8.22) below demonstrate, the nominalizing prefix allows for the extremely productive derivation of nominals from verbal roots and is thus a major asset of the language for fleshing out the lexicon.

- (8.22) *luk* > *luk-təp* > *tə-luk-təp*
 join.PST join-COLL.PST NZP-join-COLL
 ‘joined’ ‘joined together’ ‘joint’

The full range of nominalizing morphology is discussed in §7.4. A specific account of the nominalizing functions of *tə-* can be found in §7.4.1.

8.4. Verb root

The verb root forms the head of the predicate and is either mono-, di- or more rarely, trisyllabic.

8.4.1. Bare roots

Bare verb roots frequently occur in the corpus and are inherently specified for past tense when functioning as matrix clause predicates (see §8.5.12 for a complete description of tense marking).

- (8.23) *təsaɾə, nuksənsaŋpa? áhlú nə wa.*
 tə-sa-əɪ nuksənsaŋ-pà? a-hlú nə wa
 thus-say-SEQ PN-M NRL-field ALL go.PST
 ‘And having said that, Noksensangba went to the field.’

- (8.24) *khən hmapaŋ à ku, lífjá-pà? nə álíma zaŋlu-ì-pà? tənlak.*
 khən hmapaŋ à ku lífjá-pà? nə álíma zaŋlu-ì-pà?
 once time one LOC PN-M AGT world make-IRR-NR
 tənlak
 start.PST
 ‘Once upon a time, [the god] Lichaba began to make the world.’

8.4.2. Compound verbs

A verb root can be complex, i.e. made up of two lexical roots that function together as a compound verb. A compound verb forms one grammatical and one phonological word, and its constituents also function as possible simple verb roots of independent grammatical words. Unlike the pseudo-compounds described in the following section, nothing may intrude between the constituents of a true compound verb.

True compound verbs formed from independent verbal roots are rare in comparison with the ubiquitous compound nouns, and only a handful that stand up robustly to criterial tests of verbal compoundhood are attested in the corpus. Some illustrative examples are listed in (8.25).

(8.25)	<i>tshəhlaŋ</i>	<i>tshə?</i> + <i>hláŋ</i>	(pull + be.long)	>	(stretch.PST)
	<i>tshəɹa</i>	<i>tshə?</i> + <i>ɹà</i>	(pull + come)	>	(pull.towards.PST)
	<i>tʃhəɹa</i>	<i>tʃhà</i> + <i>ɹà</i>	(pick.up + come)	>	(collect.PST)
	<i>sakti?</i>	<i>sàk</i> + <i>ati?</i>	(grip.with.nails + release)	>	(slip.PST)
	<i>zaɹà</i>	<i>za</i> + <i>ɹà</i>	(enter +come)	>	(came.in.PST)

All compound verbs originate from two erstwhile independent verbal roots that have come to be treated syntactically as a single unit. For example, it can be demonstrated that *tshəhlaŋ* has lexicalized as one grammatical word, because negation has scope over both constituents of the compound, e.g. *mə-tshəhlaŋ-la* (NEG-stretch-NEG.PST) ‘didn’t stretch’. Unlike the lexical suffixes described in §8.5.1, $\sqrt{hláŋ}$ does not seem to occur productively with any other verb roots, which further suggests that it should be analyzed as a constituent of a *bona fide* compound verb and not as a verbal suffix. Similarly, the compound verb *tshəɹa* ‘pulled towards’ is formed from the independent roots $\sqrt{tshə?}$ ‘pull’ and $\sqrt{ɹà}$ ‘come’ and is negated by the discontinuous negative, e.g. *mə-tshəɹa-la* (NEG-pull.towards-NEG.PST) ‘didn’t pull towards’, with the scope of negation extending over the entire construction. This demonstrates that it functions as a single phonological and grammatical word. Negation is thus a decisive morphological test of verbal compoundhood.

In principle, an alternative analysis could identify these compound verbs as serial verb roots in nuclear juncture, with the negative discontinuous morpheme of the examples above functioning as a nuclear operator that has scope over all the constituents of the nucleus (Foley and Olson 1985: 37). But if verb serialization truly is a feature of the grammar of Mongsen, then surely one would expect to find a lot more examples than the few thus far unearthed, therefore it seems more appropriate to analyze these words as compound verbs.

8.4.3. Pseudo-compounds

Combinations of verb roots occur in what superficially appear to be compounds. One such example is *hənəɹà* ‘brought’. This has resulted from a sequential converb expressing a manner of activity in a dependent clause becoming fused with a matrix verb encoding a direction of movement. Its sources are the two independent grammatical words *hən-əɹ* (take-SEQ) plus *ɹà* (come.PST), which are in the process of undergoing a morphological reanalysis as a single grammatical

word *hənəɹà* with a new meaning of ‘brought’. The negation test proves that this is not a true compound, because it splits the two elements apart, causing the first to be treated as a non-final verb form and the second to take the negative prefix and suffix, e.g. *hən-əɹ mə-ɹà-la* (take-SEQ NEG-come-NEG.PST). Although *hənəɹà* perhaps passes a semantic test for compounds by deriving a new and slightly idiosyncratic meaning, the negation test ultimately demonstrates that it is not yet fully lexicalized. The formation of pseudo-compounds such as *hənəɹà* has undoubtedly been facilitated by the fusion of the final retroflex approximant of the sequential converb suffix *-əɹ* and the retroflex approximant onset of the main verb across a word boundary. Examples of the combination of a converb clause formed with *hən* ‘took’ plus *ɹà* ‘came’, the latter occurring either as a separate converb form or as a main verb, can be found along with examples of the superficially fused form. The following sentences demonstrate this syntactic fluidity. Consultants do not recognize any differences in meaning between the various types of constructions involving the same verb roots.

- (8.26) *təɹ á-sá? həmsəjpa? atʃak həmsəjpa? ahəm hənəɹaŋ.*
 tə-əɹ a-sá? həmsi-ì-pà? a-tʃak həmsi-ì-pà?
 thus-SEQ NRL-meat use-IRR-NR NRL-cooked.rice use-IRR-NR
 am hənəɹà-aŋ
 aspidistra.sp bring-IMP
 ‘And bring aspidistra leaves to use for the meat and rice’ (i.e. as plates).

- (8.27) *pə nə tʃjək-pà? sə takzi sə hənəɹ ɹaɹ khìɹ.*
 pə nə tʃjək-pà? sə takzi sə
 3SG AGT keep.aside-NR ANAPH bamboo.vessel ANAPH
 hən-əɹ ɹà-əɹ khì?-əɹ
 take-SEQ come-SEQ give-PRES
 ‘He brings [the aforementioned things] that he kept aside, those bamboo vessels, and gives them.’

- (8.28) *təɹ, ípá? sin tʃu ku, ɹitu niŋəɹ nə jàsà.naɹu amaŋ hənəɹaɹə, maŋmətuŋ ku mətuŋtʃaki?*
 tə-əɹ ípá? sin tʃu ku
 thus-SEQ EMPHAT BACK DIST LOC
 ɹitu niŋ-əɹ nə jàsà.naɹu a-maŋ hənəɹà-əɹ
 village.name name-ANOM AGT PN NRL-body bring-SEQ
 maŋmətuŋ ku mətuŋ-tʃak-i?
 village.name LOC be.erect-RS-CAUS.PST
 ‘And then, after that, the people of Ritu brought the [smoked and dried] corpse of Yasa Naro and left it propped up in Mangmetong village.’

8.4.4. Compounding versus stem modification

An alternative approach to an analysis of the morphemes found to occur in position (3) (listed in Table 8.2 below) is to identify them all as constituents of compound verbs. As Anderson (1985: 40) observes, it is not always easy to draw a line between compounding and stem modification. Even so, we could perhaps rely upon distribution and semantics to some extent to help tease apart these two different word formations in a synchronic description. As noted above, the morpheme *hlaŋ* in *tshəhlaŋ* ‘stretched’ does not occur juxtaposed to any other lexical root; this limited distribution contrasts strongly with the substantial range of semantic verb classes with which many of the lexical suffixes of Table 8.2 may co-occur. Furthermore, a verb compound’s meaning tends to be non-compositional. This contrasts with the compositional meanings expressed by a verb stem modified by any of the lexical suffixes of §8.5.1.

8.5. Verbal suffixes

I noted above that suffixes are overwhelmingly employed to express the verbal grammatical categories of Mongsen. This is also observed to be the case for the vast majority of Tibeto-Burman languages of north-east India.

As often proves to be the case in many languages, the verbal suffixes of Mongsen do not lend themselves to a tidy classification with respect to traditional definitions of inflectional versus derivational morphology (Bauer 1983, Bybee 1985 and Payne 1990). The problem is that this rigid division is just too inflexible when applied to the verbal morphology of Mongsen and does not allow for the many shades of grey waxing and waning between the extremes of black and white that represent the traditional poles of inflection and derivation.

This is particularly apparent for suffixes that have grammaticalized from lexical roots (see §8.5.1). These optional stem modifiers often retain quite lexical meanings and are restricted to small semantic classes of verbs, thus preventing them from occurring exhaustively in paradigmatic opposition, yet they do not change the word class of the lexical root to which they are suffixed. They therefore demonstrate properties that characterize both inflectional and derivational morphology. Because of this, it is more practical to dispense with the traditional notions of inflection and derivation and instead consider the grammaticalization of verbal suffixes with respect to their degree of abstraction. The more concrete a morpheme’s meaning, the more restricted will be its application as a grammatical suffix to particular semantic classes of verbs; conversely, the more abstract, the less restricted will be its distribution and the greater its range of application as a marker of a grammatical category.

It is not possible to elicit examples of these suffixes in all their potential combinations and be totally confident that the elicited material is reliable, therefore a decision has been made to confine the analysis to data collected from texts and conversations. Where relevant, attention is drawn to aspects of the analysis in need of further research.

8.5.1. Lexical suffix zone

At least seventeen suffixes expressing lexically-oriented adverbial meanings are found to occur in slot (3) adjacent to the root. Many of these are also used to express closely related, if not identical meanings in their functions as main verbs. A few appear to have nominal diachronic sources. The lexical suffixes together with their glosses and meanings are listed below in Table 8.2.

I suspect that most of the morphemes of position (3) have their sources in compound verbs that have subsequently grammaticalized as suffixes, in doing so gradually broadening their respective grammatical meanings and expanding their domain of applicability to lexical roots of divergent semantic classes. They are by no means equal in this respect; some lexical suffixes retain much more semantic content than others, and this in turn limits their productivity. For example, *-tən* (TO.AND.FRO) expresses a meaning that restricts its application to verbs of motion, while other lexical suffixes such as *-sət* (AFFECT) can be used with verb roots of substantially different semantic and transitivity classes. Some have undergone considerable phonological reduction, whereas others are identical in phonological form to their lexical sources. Some retain an independent main verb function. Others are well on the way to developing into fully-fledged verbal suffixes and can no longer be used as main verbs. It is therefore useful to consider the lexical suffixes to be of varying antiquity with respect to their grammaticalization as verbal morphemes.

The lexical suffix slot is arguably the most important conduit for the grammaticalization of new verbal suffixes. Here is a possible *modus operandi* by which consecutive layers of new suffixes emerge from this position in the template.

Juxtaposed constituents (V_1V_2) initially expressing independent meanings become more tightly integrated and are gradually reanalyzed as left-headed compound verbs. The V_2 element of these verb compounds enters into an asymmetrical relationship with V_1 , sharing TAM and other suffixes. This asymmetry hastens the grammaticalization of V_2 as it gradually undergoes a semantic shift from a purely lexical meaning towards a more abstract grammatical meaning. The abstraction of meaning subsequently allows V_2 to be used in combination with a larger number of verbal roots belonging to increasingly diverse semantic classes. This is what has happened to the lexical

suffix *-sət* (AFFECT) described in §8.5.1.13 below, a suffix of venerable Proto-Tibeto-Burman antiquity which is no longer limited to occurring just with semantic classes of verb roots expressing actions that can feasibly result in death.

Table 8.2. Lexical suffixes

SUFFIX	GLOSS	ORIGIN	SEMANTICS
<i>-kət</i>	ASCEND	? $\sqrt{kəwa}$ 'ASCEND+go'	'VERB upwards' (§8.5.1.1)
<i>-lak</i>	DESCEND	cf. PBL * <i>zak</i> 'descend'	'VERB downwards' (§8.5.1.2)
<i>-uk</i>	INTO	unknown	'VERB into' (§8.5.1.3)
<i>-tən</i>	TO.AND.FRO	? $\sqrt{tən}$ 'cross over'	'VERB to-and-fro' (§8.5.1.4)
<i>-saʔ</i>	SEPARATE	? PTB * <i>za</i> 'child'	'separate, reduce to many pieces by VERB-ing' (§8.5.1.5)
<i>-sa</i>	SPREAD	? PTB * <i>za</i> 'child'	'distribute by VERB-ing' (§8.5.1.6)
<i>-thàŋ</i>	SEVER	? $\sqrt{thàŋ}$ 'sever, block'	'reduce to two pieces by VERB-ing' (§8.5.1.7)
<i>-mà</i>	SPLIT	$\sqrt{səma}$ 'split'	'split open by VERB-ing' (§8.5.1.8)
<i>-thəm</i>	END	$\sqrt{thəm}$ 'end'	'end, finish VERB-ing' (§8.5.1.9)
<i>-lak</i>	TERM	PTB * <i>lak</i> 'hand, arm'	'complete; reach the terminal point of VERB-ing' (§8.5.1.10)
<i>-thən</i>	TOGETHER	$\sqrt{athənsi}$ 'gather'	'VERB together' (§8.5.1.11)
<i>-tshəŋ</i>	ATTACH	$\sqrt{tshəŋ}$ 'affix, wear'	'be attached, come into contact with, remain in place by VERB-ing' (§8.5.1.12)
<i>-sət</i>	AFFECT	PTB * <i>g-sat</i> 'kill'	'VERB to death; subject to extreme physical or psychological discomfort' (§8.5.1.13)
<i>-tʃəm</i>	ADDIT	unknown	'VERB in addition, VERB also' (§8.5.1.14)
<i>-ùk</i>	ASSIST	unknown	'assist/accompany in VERB-ing' (§8.5.1.15)
<i>-phàŋ</i>	COVER	unknown	'cover by VERB-ing' (§8.5.1.16)
<i>-tak</i>	REMAIN	unknown	'remain/stay in condition of VERB' (§8.5.1.17)

By this stage, the grammaticalization of erstwhile V_2 as a suffix is almost complete, but older lexical meanings may also be retained, resulting in what Hopper (1991: 22–24) identifies as “layering”. This accounts for the persistence of some lexical meanings alongside related grammatical meanings that are synchronically associated with the same forms.

The position that a morpheme occupies in the predicate template is interpreted as a direct reflection of its relative stage of grammaticalization. Conceivably, the perfective aspect marker of slot (9) and the resultant state marker of slot (8) embarked upon the same grammaticalization pathway that the lexical suffixes of Table 8.2 have set out upon, with their current positions in the verb stem indicative of more advanced stages of grammaticalization. That is, these chronologically older layers have in turn been shunted up the stem hierarchy by the subsequent emergence of more recent layers from slot (3) adjacent to the root.

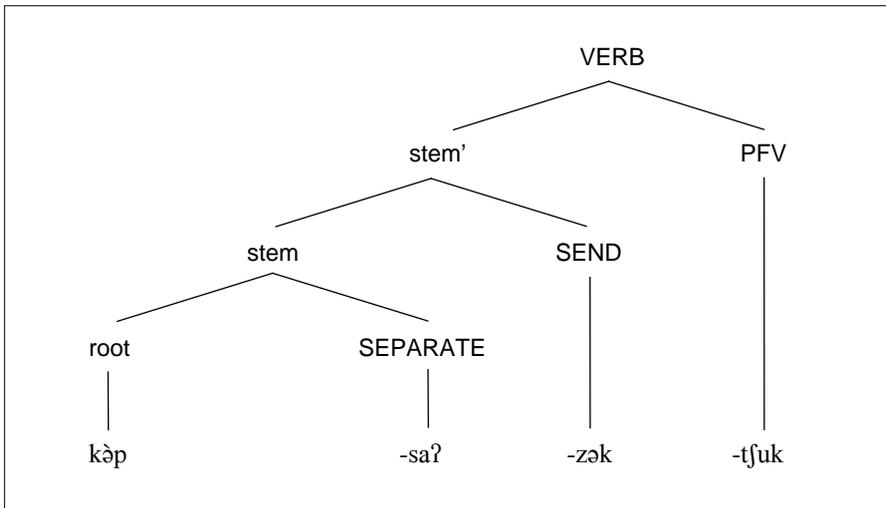


Figure 8.2. Structure of the verb stem *kəp-saʔ-zək-tʃuk* (hit-SEPARATE-SEND-PFV.PST), schematically reflecting the historical development of verbal suffixes.

Figure 8.2 provides a schematic depiction of grammaticalized layers of morphemes in the verb stem *kəp-saʔ-zək-tʃuk* according to this interpretation. The example is taken from the sentences of (8.8) above. This depiction should be understood to be representative of a general scheme that relates to the grammaticalization of all the suffixes of Mongsen that have a verbal origin, not just those demonstrated in the figure.

The grammaticalization of lexical morphemes that originate from a compounding relationship with the verb root creates a hierarchical structure in which the older layers have scope over newly emerging layers. This produces a concomitantly expanding domain of application of verbal categories recognizable in the hierarchy – as new layers emerge from slot (3), forcing older layers to ascend in the hierarchy, the categories encoded by the older layers apply to increasingly varied semantic classes of verbs. For instance, while the lexical suffix *-saʔ* of slot (3) is synchronically restricted to telic, change of state activity verbs expressing resultative meanings, the perfective suffix *-ʔʉk* of slot (8) is now totally unrestricted with respect to the semantic class of verb it can mark and consequently occurs on the stems of both activity and stative verbs. This suggests that the closer a morpheme is to the root, the more lexical its meaning and the greater the number of semantic restrictions limiting its domain of application, while the further away from the root the morpheme is located, the more grammaticalized and abstract is its meaning, and conversely the fewer semantic restrictions limiting its scope of use.

The following sub-sections describe the respective meanings of the lexical suffixes identified in Table 8.2 and give examples of their uses in clauses. A common denominator of all but a few lexical suffixes is the expression of types of resultative meanings; in the most prototypical presentation, the head expresses an activity and the lexical suffix expresses the result of that activity. Those that do not express resultative meanings are used to encode the manner in which an action is performed.

The description of the lexical suffixes commences with three deictics that encode the vector of motion in which an activity takes place. As expected, they occur in paradigmatic opposition with each other, but are compatible with the majority of other lexical suffixes occurring in this zone, particularly those expressing completive meanings of activity predicates.

8.5.1.1. *-kət* (ASCEND)

This lexical suffix expresses ‘in an upward direction’, e.g. *hən-kət-zək-aŋ* (take-ASCEND-SEND-IMP) ‘send [it] up’. It is not found to occur as a main verb, but its phonological shape and semantics imply a possible relationship to the *kə* formative that expresses ‘ascend’ in directional motion verbs (see §8.1.1.2 for discussion). An example of *-kət* is provided in (8.29) below. Its form may be conveniently compared with the directional motion verb *kəwa* ‘went up’, occurring as a verb root in the same sentence.

(8.29) *áhŋáʔ tʃu læɹali, səpɹi táŋ kəwajukli, læɹali səpɹi táŋ pakətəɹi áŋ kəwajukli,*

a-ŋáʔ tʃu læɹa-li səpsi-əɹi táŋ
NRL-fish DIST descend+come-ALT.IT turn-SEQ just

pak-kət-əɹi áŋ kəwa-juk-li
disperse-ASCEND-SEQ just ascend+go-PFV-ALT.IT

‘The fish repeatedly came down [towards the fishing weir], then just turned around and went back up [the river] dispersing as they went up, over and over ...’

Chelliah (1997: 210) describes a “first level derivational suffix” *-kət* in Meithei that similarly expresses ‘in an upward motion’. This suffix surely must be cognate with Mongsen *-kət*. Like the lexical suffixes of Mongsen, the first level derivational suffixes of Meithei have their sources in lexical verb roots, suggesting that they, too, originate from compounding processes.

8.5.1.2. *-lak* (DESCEND)

This lexical suffix is the antonym of *-kət* described in the preceding section. It also does not occur as a main verb. When suffixed to an activity verb, it denotes ‘in a downward direction’, e.g. *hən-lak-zək-əŋ* (take-DESCEND-SEND-IMP) ‘send [it] down’. Its segmental form and semantics suggest a possible correspondence with Proto-Burmese-Lolo **zak* (see Benedict 1972: 30).

(8.30) *tuku nə ah.ɹə læŋpən à hənlakəɹi áŋ thijun sə nə tshəlakzək.*

tə-ku nə a-h.ɹə læŋpən à hən-lak-əɹi áŋ
RL-uncle AGT NRL-cane whole one take-DESCEND-SEQ just

thijun sə nə tshəʔ-lak-zək
hulling.room ANAPH ALL pull-DESCEND-SEND.PST

‘Uncle took down a length of unsplit cane and pulled it down to the [aforementioned] paddy-hulling room.’

Both of the deictic lexical suffixes *-lak* and *-kət* are restricted to occurring with verbs that lack inherent deictic specification for encoding a direction of movement or activity in the vertical plane, be it of a physical entity, or of something less tangible which is nevertheless understood to move through space, such as sound. In the following imperative example, *-lak* is used to indicate the direction in which the commanded announcement is to be made. This is from a high vantage point, the bachelors’ house, down to the villagers below.

- (8.31) *nàŋ la ahɪitʃu nə kəwəɪ, “aki .nɪ.nɪ? ni,” tə phijulakzəkəŋ.*
 nəŋ la ahɪitʃu nə kəwə-əɪ
 2SG TOP bachelors'.house ALL go+ascend-SEQ
 a-ki .nɪ-əɪ-ù? ni tə phiju-lak-zək-aŋ
 NRL-house burn-PRES-DEC PTCL thus announce-DESCEND-SEND-IMP
 ‘You go up to the bachelor’s house and shout down “The houses are
 burning.”’

Most verb stems that are specified for direction of movement by *-lak* or *-kət* are found with the directional suffix *-zək*, although the concomitant presence of *-zək* on the verb stem is not a requirement for the grammatical use of either of these lexical suffixes. The uses of *-zək* as a verbal suffix and *zək* as a main verb are discussed in §8.5.3.

8.5.1.3. *-uk* (INTO)

The lexical suffix *-uk* encodes movement into the spatial domain of another entity. The encoding of movement from outside to inside a container is apparent in the following example, the container being represented by the house.

- (8.32) *tə.ə tənì khə təza anu tʃu tùhŋàɪ ki nə nì.ə wə.ə nìuktʃakì.ə,*
 tə-əɪ tə-ni khə tə-za tə-nu tʃu
 thus-SEQ RL-wife CONJ RL-child RL-child DIST
 tùhŋàɪ ki nə nì-əɪ wə-əɪ nì-uk-tʃak-i?-əɪ
 other house ALL lead-SEQ go-SEQ lead-INTO-RS-CAUS-SEQ
 ‘And, after leading his wife and children into a neighbour’s house and
 leaving them there ...’

In (4.19), repeated and renumbered below as (8.33), *-uk* encodes movement into a container represented by the clenched paw of the protagonist.

- (8.33) *təɪ, hɪətsə? tʃu, pa mija nə athənsipa?, pa nə nəmukzəktʃuk, təɪ tʃatʃuk.*
 tə-əɪ hɪətsə? tʃu pa mija nə athənsi-pà?
 thus-SEQ borer.insect DIST 3SG palm ALL assemble-NR
 pa nə nəm-uk-zək-tʃuk tə-əɪ tʃà?-tʃuk
 3SG AGT compress-INTO-SEND-PFV.PST thus-SEQ consume-PFV.PST
 ‘And then, the borers that assembled in his [i.e. Tiger’s] paw he crushed
 together and then gobbled up.’

sandhi alternations in the output (this particular concatenation of morphemes is realized phonetically as [mi¹¹jaŋ¹¹tən³³law¹]).

8.5.1.5. *-sa?* (SEPARATE)

The lexical suffix *-sa?* indicates the separation of something into smaller components or pieces, usually resulting from an act of destruction, although not necessarily. It often implies the dispersal or scattering of those pieces. This additional meaning can be reinforced by a directional suffix that encodes movement, such as *-zək* (see §8.5.3).

- (8.36) *tsə̀lə̀la tʃu nə kə̀psàzəkʃuk.*
 tsə̀lə̀la-la tʃu nə kə̀p-sa?-zək-tʃuk
 earthworm-F DIST AGT strike-SEPARATE-SEND-PFV.PST
 ‘Earthworm cracked and scattered [Red Junglefowl’s eggs].’

Alternatively, *-sa?* can encode the dissipation of smaller pieces as the result of non-simulfacative events, such as the activity of cutting up of a lump of meat into numerous smaller chunks over a protracted period of time. This may be performed without any associated sense of dispersal. It contrasts with the meaning of the closely related lexical suffix *-sa* (SPREAD), which is formally distinguished by the absence of an underlying glottal stop (see §8.5.1.6 below).

- (8.37) *wà̀samali, thuku táŋ kəm-ə̀l lili,*
 wà̀sa?-sa?-ma?-li thuku táŋ kəm-ə̀l li-li
 slice-SEPARATE-CMPL-ALT.IT nine just become-SEQ be-ALT.IT
 ‘Each time [he] finished slicing up [the magical meat into many pieces, it] kept becoming just nine [pieces], over and over ...’

While there is no indication in the corpus that *-sa?* might be related to a lexical verb root, a similar form occurs in nominal compounds with the meaning of ‘portion’, e.g. *tasu-sa* ‘small portion’, formed from the deverbal adjective *tə-asu?* (NZP-be.small), plus *sa*. This in turn is suspiciously close in phonological form and meaning to the diminutive suffix *-za*, described in §7.5.3.

The affixation of *-sa?* sometimes results in new non-compositional meanings. For example, *kə̀p* in isolation encodes a meaning of ‘struck, shot’, but with the suffix *-sa?* the meaning becomes ‘shattered, broke into pieces’. The activity verbs of (8.38) were elicited in isolation, therefore the presence of *-sa?* with verbs of this semantic class suggests that there is considerable integration of this morpheme with its lexical verb root. In some cases, the root has fully lexicalized with its lexical suffix and no longer occurs independently. The verb

stem *ɿàksaʔ* (break.PST), for example, is unsegmentable. These words were elicited in their nominalized forms, the citation form often automatically chosen by speakers of Tibeto-Burman languages. The tone sandhi manifestations of these nominalized examples are discussed in §3.3.2.

- | | | |
|---------------------------|------------------------|-----------------------|
| (8.38) a. <i>ɿàksapaʔ</i> | b. <i>ɿàksatʃukpaʔ</i> | c. <i>laŋsápàʔ</i> |
| <i>ɿàksaʔ-pàʔ</i> | <i>ɿàksaʔ-tʃuk-pàʔ</i> | <i>laŋ-saʔ-pàʔ</i> |
| break-NR | break-PFV-NR | chop-SEPARATE-NR |
| ‘to break’ | ‘to be broken’ | ‘to chop into pieces’ |

8.5.1.6. *-sa* (SPREAD)

The lexical suffix *-sa* encodes a meaning that is obviously related to *-saʔ* described above in §8.5.1.5, but is nonetheless distinct in both form and semantics. A meaning of ‘spread out, disperse’ in the absence of destruction or an instantaneous change of state is encoded by this suffix in *puŋ-sa* (jump-SPREAD.PST) ‘jumped away’, described as what grasshoppers do when one walks into long grass or, by extension, what flower buds do when they bloom, referring to their petals opening out.

- (8.39) *tʃùku atʃak sə “púŋ!” táŋ tütəɿ in nə puŋsaɿ wawʔ tətə.*
tʃùku *a-tʃak* *sə* *púŋ* *táŋ*
 consequently NRL-paddy ANAPH ONOM just
tütəɿ *in* *nə* *puŋ-sa-əɿ* *wa-ùʔ* *tətə*
 each PATH ALL jump-SPREAD-SEQ go.PST-DEC REP
 ‘Consequently [when the magical spell was broken], the [afore-mentioned] paddy just went “pung” and scattered in each and every direction.’

Because this lexical suffix lacks a glottal stop and has a slightly different meaning to that of *-saʔ*, I analyze it as a separate suffix. This is justified by the sub-minimal pairs *thəp-saʔ* (throw-SEPARATE.PST) ‘smashed to pieces’ and *thəp-sa* (throw-SPREAD.PST) ‘scattered by casting’, and *ləm-saʔ* (divide-SEPARATE.PST) ‘cut into pieces’ and *ləm-sa* (divide-SPREAD.PST) ‘distributed’. Tone sandhi is responsible for the change in tone on the verb roots of these words. Following are additional examples of *-sa* in words, demonstrating the meaning of dispersal encoded by this lexical suffix.

- (8.40) a. *thùksapa?*
 thùk-sa-pà?
 scatter.carelessly-SPREAD-NR
 ‘to scatter rice grains while eating’
- b. *phàsapa?*
 phà-sa-pà?
 scatter-SPREAD-NR
 ‘to give away freely, recklessly’

8.5.1.7. -thàŋ (SEVER)

This lexical suffix denotes ‘into two pieces’. This is usually the consequence of an activity that denotes severing an entity across the grain, resulting in two pieces of what was originally a single unit. Its source may be the lexical root *taŋ* ‘sever in one blow’, although this verb has an initial unaspirated stop consonant, whereas the initial consonant of the lexical suffix is aspirated. A more likely lexical origin is the main verb *thàŋ* ‘block, stop up’, which is related to the meaning conveyed by *-thàŋ* in the sentence of (8.43) and is also of identical phonological form and tone.

Because of the types of activity verbs it occurs with, stems marked by *-thàŋ* often involve the use of an instrument, either the hands, as in *hjà-thàŋ* ‘snap in two with the hands’, or more commonly a cutting tool or other means of physical division.

- (8.41) *imtisàŋpà? tʃu tʃùku anuk sə həməjáj kà? kà? kà? ləpthàŋlítáŋ wa.i.*
 imtisàŋ-pà? tʃu tʃùku a-nuk sə həmə-əi
 PN-M DIST consequently NRL-machete ANAPH hold-SEQ
 áŋ kà?.kà?.kà? ləp-thàŋ-lí? táŋ wa-əi
 just ONOM cut-SEVER-SIM just go-PRES
 ‘Imtisangba consequently takes hold of the machete and Chop! Chop!
 Chop!, he cuts out the knots of wood.’

- (8.42) *mətshəla nə nì taham nət tshəthàŋjakiaj.*
 mətshə-la nə nì taham nət
 barking.deer-F AGT 1SG middle two
 tshə-thàŋ-jak-i? əj
 stamp.on-SEVER-RS-CAUS.PST SINCE
 ‘Since Barking Deer stamped on me and left me severed me in two ...’

The ‘sever’ lexical suffix occurs with semantic classes of verbs expressing division, but without involving actual physical cutting, e.g. *jùŋ-thàŋ-pà?* (block-SEVER-NR) ‘to obstruct’, and *ləm-thàŋ-pà?* (divide-SEVER-NR) ‘to divide’. Some verb roots with this suffix have become lexicalized with new meanings, e.g. $\sqrt{\text{ləpthàŋ}}$ ‘pass judgement’.

- (8.43) *nì ènti.ɪəm ku, nàŋ thən təpətɪpə? à nə juŋthàŋəɪ, nì mət̪àt̪ʃhətlaw?*
 nì ènti ɪəm ku nàŋ thən tə-pət̪i-pà? à nə
 1SG path MIDDLE LOC 2SG COM NZP-be.big-NR one AGT
 juŋ-thàŋ-əɪ nì mət̪-ɪà-t̪ʃhət-la-ù?
 block-SEVER-SEQ 1SG NEG-come-ABIL-NEG-DEC
 ‘On the way one bigger than you obstructed the path and I wasn’t able
 to come.’

Burling (1961: 16) reports that the Bodo-Konyak-Jinghpaw language Garo, spoken in the Garo Hills of Meghalaya, has a verbal suffix *-toŋ* expressing ‘into two pieces’. He describes this as belonging to a sizable class of derivational morphemes that he refers to as “adverbial affixes”. Like the lexical suffixes of Mongsen, these occur immediately to the right of the verb root, express a range of adverbial meanings, and demonstrate varying degrees of productivity, depending upon the particular verb base involved. Burling (1961: 13–17) identifies twenty-nine of the most commonly occurring, non-productive members of this affix class, noting that many can also function independently as lexical verbs.

The Garo suffix *-toŋ* appears to be very close to Mongsen *-thàŋ*, both in form and meaning. One cannot help but be struck by the remarkable similarity in the agglutinative structure of the verb stem and the range of meanings encoded by suffixes that have grammaticalized from main verbs in Garo, Meithei, Mongsen and possibly many other Tibeto-Burman languages of the north-east India region. Whether this is attributable to genetically-shared word formation processes, parallel grammaticalizations resulting from the natural “drift” of languages (Sapir 1921, LaPolla 1994), or areal contact is yet to be determined.

8.5.1.8. *-mà* (SPLIT)

The lexical suffix *-mà* (SPLIT) is similar to both *-sa?* and *-thaŋ* in expressing a particular effect of an action upon an entity that results in its change of state. In the majority of textual examples, *-mà* occurs on verbs of cutting and the entity that is subjected to the action is split into two pieces. This is its meaning in the following dependent clauses, in which the protagonist uses a machete to split open magical gourds containing rice. The end result of this action is that each gourd is split in half.

- (8.44) *ləpmàli, atʃha tʃu nə zəkli,ləpmàli, atʃha tʃu nə zəkli,*
 ləp-mà-li a-tʃha tʃu nə zəkli
 cut-SPLIT-ALT.IT NRL-rice.bin DIST ALL send.ALT.IT
 ləp-mà-li a-tʃha tʃu nə zəkli
 cut-SPLIT-ALT.IT NRL-rice.bin DIST ALL send.ALT.IT
 ‘Each time [she] split open [a magical gourd, she] sent [the rice that
 poured out of it] to the rice bin, over and over, ...’

However, it is also possible for the patient of such an action to become more than two pieces, as in the textual example of (8.45). Here the god Lichaba is testing out his future son-in-law, to whom he passes a length of unsplit cane to bind the legs of a wild boar prior to slaughtering it and distributing its meat to relatives. But first, the prospective son-in-law must hold the boar down while he attempts to tear off a strip of cane that will be pliable enough to use as a thong to bind its legs. While he struggles with the boar, his fiancée is hulling paddy close by. She helps him surreptitiously by letting the end of the cane become pounded by the pestle in her paddy-huller, which is like a solid wooden table containing a mortar. This frays the end, thereby allowing him to use his teeth to tear off a cane strip and complete his task.

- (8.45) *tətʃhàku mitəm nə pi tshəmàzək.*
 tətʃhà-ku mitəm nə pi tshəmà-zək
 thus-do-LOC.CV pestle INST PROX pound-SPLIT-SEND.PST
 ‘And then, this [cane] was split by the pestle.’

In all the uses of *-mà* encountered in the data, the splitting is always along the grain of the entity, never across it. This must be highly significant to its meaning, since it suggests a contrast with the haphazard shattering encoded by *-saʔ* (SEPARATE) and the across-the-grain severing encoded by *-thàŋ* (SEVER). A likely lexical source of *-mà* is the verb root $\sqrt{səmà}$ ‘split’. An example of this verb occurs in the same text as the head of a dependent clause, immediately following the clause of (8.45).

- (8.46) *təɪ, təpaŋ nə muɪ səmàɪ,*
 tə-əɪ tə-paŋ nə mùʔ-əɪ səmà-əɪ
 thus-SEQ RL-mouth INST bite-SEQ split-SEQ
 ‘And, having split [the piece of cane] by biting it with his mouth, ...’

8.5.1.9. *-thəm* (END)

The lexical suffix *-thəm* encodes the completion of an activity. It has probably only recently grammaticalized from its verbal source, because it is not compatible with stative verbs; nor can it freely occur with all activity verbs that are capable of expressing events leading to an endpoint. For instance, *tʃãʔ-thəm-aŋ* (consume-END-IMP) ‘eat (it) up!’ is an acceptable verb stem, but *tʃəm-thəm-aŋ* (drink-END-IMP) is questionable to some speakers, who prefer the root *tʃəm* ‘drink’ to occur with the completive aspectual suffix *-maʔ* of slot (6).

- (8.47) *ajla nə áwkla tʃaŋsə tʃu təmáŋ tʃəmthəmtʃukaj, ajila tʃaŋsə tán li.*
 a-ji-la nə a-úk-la tʃaŋsə tʃu tʃəm-thəm-tʃuk
 NRL-dog-F AGT NRL-pig-F footprint DIST step.on-END-PFV
 aj a-ji-la tʃaŋsə tán li
 SINCE NRL-dog-F footprint just be.PST
 ‘Since Dog had completely trampled all over Pig’s footprints, only
 Dog’s footprints were there.’

The diachronic source of the terminative suffix is a homophonous verbal root *thəm*, expressing a meaning of ‘to end, finish’:

- (8.48) *tə hjutsə atʃu ku thəm-əɪ.*
 tə hjutsə atʃu ku thəm-əɪ
 thus story NRL-DIST LOC end-PST
 ‘There the story ends.’

8.5.1.10. *-lak* (TERM)

The lexical suffix *-lak* encodes the terminal point of an activity. Its meaning is completive, in common with *-thəm* discussed above in §8.5.1.9. The completive meaning of *-lak* is clearly discernible in the textual example of (8.49). In the source narrative, Fox has played a practical joke on Tiger by weaving his tail into some heavy bamboo matting (used for walls of houses) without Tiger being aware of it. When he realizes what has happened, Tiger calls upon some borer insects known as *hɪətsəʔ* to bore through the tightly woven bamboo so that he can extract his tail.

(8.49) *tə tsəŋti tʃu hɪəlakəɪ, pa təmi tʃu tʃhuwa.*

tə tsəŋti tʃu hɪə-lak-əɪ
 thus woven.bamboo.wall DIST perforate-TERM-SEQ

pa tə-mi tʃu tʃhuwa
 3SG RL-tail DIST emerge.PST

‘Thus, after [the borers] had finished boring through the woven bamboo matting, his tail came out.’

This morpheme retains a lexical function, occurring as a relational noun *tə-lak*, meaning ‘last, end’. This is surely related to PTB **lak* ‘hand, arm’, but now expresses a more generic meaning of ‘terminal part’. A terminal sense is also embodied in the term *tə-lak*, which is used to refer to the youngest child of a family, the last of the siblings.

(8.50) *ásəŋ aw tʃu ̀̀ɪaʔ phinə aliŋ təlak phinə ásəŋ tʃu hənəɪ.*

a-səŋ a-u tʃu ̀̀ɪaʔ phinə
 NRL-wood NRL-bamboo.sp DIST afar ABL

aliŋ tə-lak phinə a-səŋ tʃu hən-əɪ
 lower.range RL-end ABL NRL-wood DIST take-SEQ

‘After carrying wood and bamboo from afar, from the end of the lower range, ...’

The terminative suffix *-lak* is homophonous with the suffix *-lak* that encodes ‘descend’ and also occurs in the lexical suffix zone (see §8.5.1.2), therefore speakers must rely upon context to disambiguate the possible two meanings of this morpheme. The two suffixes are semantically compatible and can occur juxtaposed in the same verb stem, e.g. *thəp-lak-lak-əɪ* (throw-DESCEND-TERM-SEQ) ‘after finishing throwing it down’, thus proving their categorial independence. When they co-occur, the initial lexical suffix is interpreted as the directional, which always occurs immediately to the right of the root in the lexical suffix zone.

The terminative lexical suffix *-lak* has fewer restrictions on the semantic classes of verbs it can mark than *-thəm*, suggesting that it may be an older morpheme according to the grammaticalization hypothesis posited in §8.5.1. It occurs with any lexical verb denoting an activity that can be temporally bounded, and thus has much in common with the meaning encoded by the perfective suffix (§8.5.7).

(8.51) *thuku ləplakəɪ tʃhàwkiʔ.*

thuku ləp-lak-əɪ tʃhàwk-iʔ

nine cut-TERM-SEQ keep-CAUS.PST

'After [finishing] cutting the nine [pieces, he] kept them.'

(8.52) *pə nə kákót pi zəŋlakukù.*

pə nə kákót pi zəŋ-lak-ukù

3SG AGT book PROX read-TERM-ANT

'He's read this book.'

(8.53) *tsəŋi ɹaləkliku aki nə wəɪ atì.*

tsəŋi ɹà-lak-li-ku a-ki nə wə-ì atì

rain come-TERM-BE.COP-CIRCM NRL-house ALL go-IRR PTCL

'If it is the case that it stops raining, [we'll] definitely go to the house.'

A related productive suffix *-lak* occurs on noun stems with a meaning of 'terminal part' (see §7.5.4), e.g. *tə-mi-lak* (RL-tail-TERM) 'tip of a tail'. The two suffixes differ not so much in meaning, as in having scope over different dimensions, which is formalized by the word class affiliations of their lexical roots. The verbal terminative suffix denotes a terminal point in time, whereas the nominal terminative suffix denotes a terminal point in space. The semantic link between space and time is well attested cross-linguistically, (e.g. see Haspelmath 1997).

I suggested earlier that these homophonous verbal and nominal suffixes must have a common nominal origin in the PTB root **lak* 'hand, arm'. The crossing of the boundary between spatial and temporal dimensions that allows the one suffix to be used in both domains is comparable to the grammaticalization of the locative case marking clitic as a marker of dependent clauses. Its reanalysis allows the locative to be used for locating events in time, but the same morpheme preserves its original function as a nominal case marker for locating entities in space (see §5.3.2.6 and §11.4.3.7).

8.5.1.11. *-thən* (TOGETHER)

This lexical suffix expresses a resultative meaning consistent with individual entities being grouped into or forming one mass. The lexical root of its stem expresses the manner of activity by which this is accomplished.

(8.54) *təlu? tə hlùthənaɪ, athi ku naŋthənaŋ.*

təlu? tə hlù-thən-əɪ

all thus pick-TOGETHER-SEQ

a-thi ku naŋ-thən-aŋ

NRL-field.hut LOC heap-TOGETHER-IMP

‘After harvesting all [the gourds], heap [them] together in the field hut.’

(8.55) *aji? nə tsəŋthənù?*

a-ji nə tsəŋ-thən-əɪ-ù?

NRL-dog AGT bark-TOGETHER-PRES-DEC

‘The dogs are [all] barking together.’

One possible source for this lexical suffix is the verb root $\sqrt{\text{athənsi}}$ ‘assemble’, which has a very similar meaning to the lexical suffix. If so, a reduction in its phonological bulk would be an expected consequence of its grammaticalization.

(8.56) *hmapaŋ à ku, sàɪaɪ təlu? athənsi, muŋthaŋ à khà.*

hmapaŋ à ku sàɪaɪ təlu? athənsi-əɪ

time one LOC animal all assemble-SEQ

muŋthaŋ à khà

conference one have.PST

‘Once upon a time, all the animals gathered together and had a conference.’

Another possible source is the verb root $\sqrt{\text{ənthən}}$ ‘gather’. I suspect that *-ən* is a fossilized formative that perhaps once expressed ‘by hand’ (it occurs recurrently with this fossilized meaning), but has come to be fused with its lexical root. The root $\sqrt{\text{ənthən}}$ now seems to be fully lexicalized with the simplified meaning of ‘gather’. Its use is demonstrated in the following example, taken from a Waromung Mongsen text.

(8.57) *anjala tsə la, phàɪali ənukli tàtəŋ ənthən.*

an-ja-la tsə la phàɪa?-li ənuk-li

chicken-DIM-F DIST TOP catch-ALT.IT put.in-ALT.IT

tà təŋ ənthən

thus just gather.PST

‘As for Little Chick, each time she caught [a fish] she put it away, over and over, and just gathered them.’

It is relevant to note that the comitative marker *thən* (§5.3.2.7) is suspiciously similar in both meaning and form to the lexical roots $\sqrt{\text{athənsi}}$,

√*ənthən* and the lexical suffix *-thən*. While it does not actually encode the meaning of ‘gather’ that appears to be the essence of meaning inherent in the verbal morphemes discussed above, it does nevertheless express ‘together with’ and is therefore semantically related, perhaps too closely to be attributable to mere chance. An example of the comitative marker occurs in the following sentence.

- (8.58) *nàŋ anu thən mùŋaŋ.*
 nàŋ a-nu thən mùŋ-aŋ
 2SG NRL-child COM rest-IMP
 ‘You stay at home with the baby.’

This prompts the question: could the comitative case marker have grammaticalized from the same verbal source as the lexical suffix? If so, this would be the second unusual example of what appears to be an identical grammaticalized form operating with essentially the same meaning in both the nominal and verbal domains of Mongsen morphology (cf. the verbal and nominal functions of the terminative suffix *-lak* in §8.5.1.10 and §7.5.4 respectively).

8.5.1.12. *-tshəŋ* (ATTACH)

The lexical suffix *-tshəŋ* contributes a meaning of ‘be attached, be in contact with, remain in place’ to its verb stem. It denotes that something attached to or coming into contact with another entity remains in that particular location as the result of the action expressed by the lexical root.

For example, *nəp-tshəŋ-i?* (adhere-ATTACH-CAUS.PST) might be used to describe the act of sticking a piece of chewing gum under a table, with the inclusion of *-tshəŋ* in the verb stem additionally encoding that it remained attached to the surface under the table. The same stem without *-tshəŋ* would simply imply that the chewing gum was stuck somewhere, but without any supplementary understanding that it remained attached in that particular location. One consultant suggested that *-tshəŋ* carries the inference that the thing attached is foreign, that it normally shouldn’t be there. This is suggested by the application of the paste to the wound in the following example, repeated from (5.34) for ease of reference and renumbered as (8.59).

(8.59) *tujət nə, tʃũŋkuli tʃu zàŋlu, mətʃatshəŋ nə tsəŋpa? khùmapaŋ tʃu zətshəŋi?*

tujət nə tʃũŋkuli tʃu zàŋlu-əɪ mətʃatshəŋ nə
3DU AGT anthill.soil DIST make-SEQ PN AGT

tsəŋ-pà? khùm-a-paŋ tʃu zə-tshəŋ-i?
spear-NR wound-mouth DIST spread-ATTACH-CAUS.PST

‘They two made a paste of ant-hill soil [and] applied it the wound that was inflicted by Mechatseng’s spearing.’

Example (8.60) occurs in a narrative immediately before the sentence of (8.45) that was discussed at length in §8.5.1.8. This verb stem also carries the nuance of ‘foreign location’ of the entity, with respect to the piece of cane coming into contact with the top of the paddy-hulling table (and eventually being pounded in it).

(8.60) *atʃak tshəja likə, ípá? ah.ɪə laŋpən tʃu tʃhəmku ma ku .ɪətshəŋ.*

a-tʃak tshə-ja li-əkə ípá? a-h.ɪə laŋpən tʃu
NRL-paddy pound-CONT be-SIM EMPHAT NRL-cane whole DIST

tʃhəmku ma ku .ɪətshəŋ

paddy.hulling.table FACE LOC slide-ATTACH.PST

‘While [she] was pounding the paddy, that length of unsplit cane slid [down] and touched the top of the paddy-hulling table.’

As a main verb, *tshəŋ* expresses a meaning of ‘attach, affix, wear’. The *ani?* ‘sarong’ (or *mekhela* in Assamese) mentioned in the following example is a length of rectangular cloth that is wrapped around the lower half of a woman’s body and secured by partially tucking in the loose end at the waist. An inherent sense of the positive polarity meaning of *tshəŋ* ‘wear’ is that the worn item of clothing remains in a particular location, attached to the body of the wearer.

(8.61) *təsalikà? tənì nə ani mətshəŋlá hi.ɪ.ɪə, aki tʃhapà?*

tə-sa-likà? tənì nə a-ni? mə-tshəŋ-la
thus-say-CONTEMP RL-wife AGT NRL-sarong NEG-wear-NEG.CV

hi.ɪ-əɪ a-ki tʃhapà?
come+level-SEQ NRL-house open.PST

Thus, when [he] said that, his wife, not wearing a sarong (in order to tempt him), came and opened the house.

A related sense of ‘attach’ is suggested by the use of *tshəŋ* as a main verb in the following narrative example for one Mongsen consultant, who thought it conveyed that ‘Prawn wore its head on its butt’.

(8.62) *tətʃhàku, akuŋ la sə təsami sə ku tələm tshəŋʃuk.*

tə-ʃhà-ku a-kuŋ la sə tə-sami sə ku
 thus-do-LOC.CV NRL-prawn TOP ANAPH RL-rump ANAPH LOC

tə-ləm tshəŋ-ʃuk
 RL-head attach-PFV.PST

‘And then, as for [aforementioned] Prawn, its head became stuck to its [aforementioned] rump.’

8.5.1.13. *-sət* (AFFECT)

The source of this morpheme is an attested PTB root (< **g-sat* ‘kill’ STC #58). Although a reflex no longer exists synchronically in Mongsen as an independent lexical item, a grammaticalized vestige *-sət* (AFFECT) survives as a lexical suffix. This encodes a resultative meaning when occurring on verb roots expressing activities affecting animate referents.

When suffixed to verbs of this semantic category, *-sət* commonly implies that death resulted from the manner of the activity expressed by the lexical root of the verb stem. With one exception, viz. (*h*)*ənsət* ‘killed’, all stems formed with *-sət* have identifiable verbal roots with distinct meanings. Whatever prior meaning fossilized (*h*)*ən* once encoded has since become shrouded by time and consequently *hənsət* is now fully lexicalized with a generic meaning of ‘killed’ (it may have formerly encoded ‘by hand’; cf. *ənthən*, discussed above in §8.5.1.11).

The verb (*h*)*ənsət* has a restricted set of referents that can function as O arguments – it only seems to permit referents that are not considered dangerous, or which cannot mount any challenge to the act of killing. It is commonly used to describe the killing of a docile domestic or harmless creature, such as a chicken or a louse, in situations in which one does not wish or need to be specific about the manner in which the act was performed. This lack of relevance concerning the explicit manner of death perhaps explains why (*h*)*ənsət* is commonly used with these types of referents. Of course, it is also possible to be specific about the manner of killing a docile creature, although perhaps there is less pragmatic need to be so contextually precise. Conversely, the killing of a potentially dangerous adversary, such as a tiger, leopard or man, is always encoded by a lexical verb that specifically expresses the manner of activity, such as *təŋ-sət* ‘speared to death’. It would not be unexpected for some derivations involving *-sət* to have a specific semantic class of referents that can function as undergoer arguments of these verbs. We find similar restrictions applying to certain semantically similar verbs in English, such as *assassinate*,

the undergoer argument of which is restricted just to referents belonging to the semantic class of high profile public office holders.

Following are narrative examples of *-sət* occurring as a lexical suffix on verb roots expressing manners of killing that unambiguously result in death. This sense of *-sət* is most closely aligned to the original meaning of its ancient lexical source.

(8.63) *tətʃaŋ ku ɹakəɪ, áwk tʃu thənsət, puŋi sə.*

tə-tʃaŋ ku ɹak-əɪ

RL-leg LOC bind-SEQ

a-úk tʃu thən-sət pùŋì sə

NRL-pig DIST stab-AFFECT.PST wild.pig ANAPH

‘Having bound its legs, the wild pig was slaughtered’ (i.e. by stabbing).³

(8.64) *tə, zaɪ, pa atsə sə ku ahɹəksət.*

tə za-əɪ pa a-tsə sə ku ahɹək-sət

thus enter-SEQ 3SG NRL-water ANAPH LOC drown-AFFECT.PST

‘Thus, after entering the water, he drowned to death.’

The AFFECT lexical suffix can occur productively on any verb stem expressing an activity that could conceivably result in death, e.g. *ɹuŋ-sət* ‘burned to death’, *niiŋ-sət* ‘strangled to death’ and *jək-sət* ‘beat to death’. Some idiosyncratic tonal alternations affect the output tone of the lexical suffix in some of these words. These are mostly of the pattern L → M / M _.

A secondary sense of *-sət* has emerged from the primary sense of ‘VERB to death’ demonstrated in the examples above. This shows a clear semantic shift to a meaning of ‘subject to extreme physical discomfort’ without necessarily terminating in death. The following example, taken from a Waromung Mongsen text, demonstrates this metaphorical meaning. In a traditional Ao folklore narrative, a dog and a pig are sent to the field to work to determine which one deserves to sleep inside the master’s house. The dog lies down and rests, while the pig labours all day under the hot sun but does not die from its exposure, despite the presence of *-sət* on the verb root *phu* ‘expose to sun’. The use of *-sət* in this context therefore simply intensifies the physical discomfort that the hapless pig endures while working in the field, because she lives on in the tale to suffer another fate. Thus, from encoding the manner of an activity that inevitably results in death, a semantically related but less concrete sense of *-sət*

3. Pigs are traditionally slaughtered by first stabbing them in the side of the chest with a bamboo stake.

has evolved to imply ‘extreme physical discomfort’, without actually terminating in a fatal outcome.

(8.65) *áwk tsə la tsəŋi na phusəta álù təmáŋ si pa tənə? na mukuŋa ili.*

a-úk tsə la tsəŋi na phu-sət-a
 NRL-pig DIST TOP sun INST expose.to.sun-AFFECT-SIM
 a-lú təmáŋ si pa tə-na? na mukuŋ-a
 NRL-field all ANAPH 3SG RL-nose INST root-SIM
 ili
 wander.PST

‘As for Pig, burned to the point of death by the sun, she wandered all over the field rooting up the earth with her snout.’

The semantic classes of verbs that this lexical suffix can modify now includes some that do not even denote a physical activity. Example (8.66) shows how *-sət* has extended its semantic domain to verbs of feeling that cannot logically result in physical death. The meaning that *-sət* adds to the members of this semantic class is one of extreme psychological discomfort. This can be seen in the simultaneous converb *tsəpha-sət-əkə*. The source of this example is a folklore story in which Tiger has proclaimed that all the animals must come to him to be eaten, one per day, and Rabbit’s day has arrived.

(8.66) *tsəphasətəkə, pa nə hmapaŋ tətə à mənufuk.*

tsəpha-sət-əkə pa nə hmapaŋ tə-tsə? à
 fear-AFFECT-SIM 3SG AGT time NZP-be.short one
 mənu-tʃuk
 be.late-PFV.PST

‘Being scared to death, she was late for a short time.’

Example (8.67), taken from a Waromung Mongsen text, shows that *-sət* has now fully grammaticalized with a new meaning as a general intensifier that is completely divorced from its original sense of ‘kill’. Here it modifies the stem of a simultaneous converb. This forms a constituent of a headless relative clause that is functioning as the complement of a verbless clause.

(8.67) *pi tʃhimisətáŋ lipa? utà.*

pi tʃhi-mì-sət-a áŋ li-pà? utà
 PROX take-DESID-AFFECT-SIM just be-NR PTCL

‘This is just [the one I’ve] been wanting!’

The various meanings encoded by *-sət* correlate with a suffix *-hət* occurring in Meithei (Chelliah 1997: 207–208). These are undoubtedly reflexes of the same etymon and have shared the same grammaticalization pathway, despite the considerable genetic distance between the two languages. The grammaticalization of *-sət* as a verbal suffix is therefore likely to be of Proto-Tibeto-Burman antiquity.

Lastly, it is relevant to add that the expressions ‘VERB to death’ and ‘dying to VERB’ occur in English, suggesting that the verb DIE may be a common cross-linguistic target for the development of an intensifier meaning. This further suggests that the metaphorical senses that develop out of DIE in Ao, Meithei and English constitute evidence of a specifically semantic drift, although it additionally involves the concomitant grammaticalization of a suffixal morpheme from a cognate lexical verb root in the case of Ao and Meithei.

8.5.1.14. *-tʃəm* (ADDIT)

The lexical suffix *-tʃəm* encodes that an activity or event expressed by a verb is additionally performed or occurred in addition to another event. In the following example, *-tʃəm* implies that the beating was done a second time, in addition to a prior beating that was given. The rarity of its occurrence in the corpus does not allow any statements to be made with respect to restrictions on its use. The examples provided in this section were offered by Mongsen speakers in their illustrative explanations of its meaning.

- (8.68) *pa kàʔ .ɿɿ, awəɿʔ jaktʃəm.*
 pa kàʔ .ɿɿ-əɿ awəɿʔ jàk-tʃəm
 3SG also come-SEQ thief beat-ADDIT.PST
 ‘He also came and beat the thief.’

A suffix of the same form occurs on nominalized verb roots with a completive meaning. The sense it expresses on the nominal form of the verb in (8.69) is metaphorically related, in that it implies that something extra has been done to the water; in this case, that it has also been boiled.

- (8.69) *atsə màrùk ku lipaʔ tʃu təməlutʃəmùʔ.*
 a-tsə màrùk ku li-pàʔ tʃu tə-məlu-tʃəm-ùʔ
 NRL-water cup LOC be-NR DIST NZP-boil-ADDIT-DEC
 ‘The water in the cup is boiled.’

This lexical suffix is productive in nominalized stems of transitive activity verbs, where it appears to encode a completive meaning, e.g. *tə-zarʃlu-tʃəm*

(NZP-make-ADDIT) ‘already made’, *tə-tʃhu.lu-tʃəm* (NZP-cook-ADDIT) ‘already cooked’, and *tə-zəlu-tʃəm* (NZP-write-ADDIT) ‘already written’. Its diachronic source has not been identified.

8.5.1.15. *-ɹùk* (ASSIST)

The “assistive” lexical suffix expresses a meaning of ‘assist, accompany in VERB-ing’. The semantic classes of verbs it occurs with or restrictions limiting its use have not yet been established, but judging from the following text example and its inherent semantics, we might predict that any verb stem that could possibly denote a collectively-performed action or event would tolerate the presence of this suffix. This is one of the only two naturally spoken examples of *-ɹùk* in the corpus.

(8.70) *pa nə “aj aw nì nə kà? tàk.ɹùk.ɹù ní” tə səmtsə.*

pa nə aj a-u nì nə kà?
3SG AGT EXCLM VOC-grandfather 1SG AGT also

tàk-ɹùk-ɹù ní tə səmtsə
weave-ASSIST-IMM QPTCL thus ask.PST

‘Hey, Grandfather! Will I also weave along with you?’, [she] asked.’

Clark ([1911] 1990: 619) describes a Chungli verb *roker* that is very likely to be cognate with the assistive lexical suffix *-ɹùk*. He translates this as ‘to go as a matter of friendship or as a neighborly act and help another in work and expect no compensation whatever’. Its meaning is therefore very close to that expressed by this lexical suffix. In Mongsen, *ɹùk* does not appear to be used as a lexical verb root synchronically.

8.5.1.16. *-phàŋ* (COVER)

A verb stem marked by this highly productive lexical suffix expresses a resultative meaning with a basic sense of ‘cover by VERBing’. As we have seen with other resultative lexical suffixes, the verb root expresses the means by which the result is achieved, e.g. *nəm-phàŋ* (press-COVER) ‘cover (something) by pressing’ and *zə-phàŋ* (smear-COVER) ‘cover (something) by smearing’. The diachronic source of this suffix has not been identified.

- (8.71) *pa lisəʔ təmaŋ tʃu ku mələmmələməkə tìphàn.*
 pa lisəʔ tə-maŋ tʃu ku mələm-mələm-əkə
 3SG mud RL-body DIST LOC thick-RED-SIM
 tì-phàn
 daub-COVER.PST
 ‘He smeared mud on his body very thickly.’

- (8.72) *akhula tʃu ah.ɿə nə khàn-phàn.*
 a-khu-la tʃu a-h.ɿə nə khàn-phàn
 NRL-tiger-F DIST NZP-cane INST wrap-COVER.PST
 ‘Tiger wrapped [his body] thoroughly with cane.’

8.5.1.17. *-tak* (REMAIN)

The lexical suffix *-tak* is an infrequently occurring morpheme expressing a sometimes elusive but generally resultative meaning of ‘remain/stay in condition of VERB’; e.g. *tshək-tak-əɿ* (collide-REMAIN-PRES) describes the act of getting stuck fast while carrying a table through a narrow doorway, and *wak-tak-əɿ* (swell-REMAIN-PRES) describes the condition of a belly that has become swollen from overeating. Somewhat less transparently, *.ɿək-tak* (bind-REMAIN) describes the uncomfortable sensation of wearing an overly tight undergarment.

Following is a rare textual example. Here I assume that the lexical suffix *-tak* implies that the lit torch remains in a state of ignition.

- (8.73) *mìlən təpəti à tsək-tak-əɿ, tsəŋrutshəla, atʃak hənəɿ tʃhuwapaʔ tʃu əntsə waku,*
 mìlən tə-pəti à tsək-tak-əɿ tsəŋrutshə-la
 torch NZP-be.big one ignite-REMAIN-SEQ PN-F
 a-tʃak hən-əɿ tʃhuwa-pàʔ tʃu əntsə wa-ku
 NRL-paddy take-SEQ emerge-NR DIST halfway go-LOC.CV
 ‘After lighting a big torch, upon going halfway [to meet] Tsengrutsela, who was coming out [from the field] carrying the paddy, ...’

8.5.2. Reciprocal/collective *-təp*

The suffix *-təp* encodes either collective activity, i.e. ‘VERB together’, or reciprocal activity, i.e. ‘VERB each other’ (see §5.4.3). The inherent semantics of some verb roots allow for only one of these two interpretations, e.g. *tʃa-təp* (consume-COLL.PST) ‘ate together’ and *wa-təp* (go-COLL.PST) ‘went together’,

while other verb roots only permit a reciprocal meaning, e.g. *mələn-təp* (change-RECIP.PST) ‘exchanged’.

A considerable number of transitive verbs marked with the reciprocal/collective suffix allow for a collective and/or a reciprocal interpretation. Often context must be relied upon to disambiguate potentially different meanings, e.g. *sa-təp* (say-RECIP/COLL.PST) ‘said to each other’ or ‘said together’, and *laŋli-təp* (reply-RECIP/COLL.PST) ‘argued with each other’ or alternatively ‘replied together’.

(8.74) *tə maŋmətuŋ ku imtisəŋpà? khə nuksənsəŋpà? hjutsə nət tʃu tʃutə satəpja?*

tə maŋmətuŋ ku imtisəŋ-pà? khə nuksənsəŋ-pà?
thus village.name LOC PN-M CONJ PN-M

hjutsə nət tʃu tʃu-tə sa-təp-ja?
story two DIST DIST-thus say-COLL-HAB.PST

‘Thus [people] used to tell each other the story of Noksensangba and Imtisangba like that in Mangmetong village.’

Alternatively: ‘Thus [people] collectively used to tell the story of Noksensangba and Imtisangba like that in Mangmetong village.’

(8.75) *tuŋnət tʃu atʃu ku áŋ ləŋlitəpəɪ.*

tuŋnət tʃu a-tʃu ku áŋ a-hju ləŋli-təp-əɪ
3DU DIST NRL-DIST LOC just NRL-word reply-RECIP-PRES

In context: ‘They two argue back and forth with each other there.’

Alternative meaning out of context: ‘They two (collectively) reply there.’

The reciprocal/collective suffix appears to be fossilized in some stems, e.g. *hùntəp*, which has been relexicalized as ‘fight’ and consequently now accepts the reciprocal/collective suffix on its stem, e.g. *hùntəp-təp*. The erstwhile root *hùŋ* no longer occurs as a lexical verb root with an independent meaning of ‘fight’. A homophonous verb root *hùŋ* ‘get, see’ appears to be unrelated.

8.5.3. Directional suffix *-zək*

The directional suffix *-zək* encodes various meanings and occurs in its own slot following the lexical suffixes. One meaning suggests movement away from a deictic centre. This sense is closest to its original lexical meaning of ‘send’. The direction-neutral encoding of motion by *-zək* is often clarified by the presence of a lexical suffix denoting a specific direction. Text counts reveal that seventy

percent of the occurrences of *-zək* in verb stems are accompanied by a lexical suffix signalling a direction of movement.

The following example demonstrates how the combination of a directional lexical suffix and *-zək* on the stem of an activity verb can contribute to a meaning of separation of an entity, movement away from the understood deictic centre, and the direction of that motion. In (8.76), *-zək* encodes the actual movement of the entity, i.e. a piece of wood falling, resulting from Squirrel gnawing through a dry branch.

- (8.76) *thaŋtsəla nə kəwa:áŋ, anak ku səŋkuŋ à məzəplakzəkə:áŋ,*
 thaŋtsə-la nə kəwa-əɪ áŋ
 tree.shrew.sp-F AGT ascend+go-SEQ just
 anak ku səŋ-kuŋ a məzəp-lak-zək-əɪ áŋ
 top LOC wood-dry one gnaw-DESCEND-SEND-SEQ just
 ‘Squirrel, after climbing to the top of a tree and gnawing through a dry branch and sending it down, ...’

In (8.77), the presence of *-zək* reinforces the downward direction in which Fox’s tail is dangling. This example demonstrates that the directional suffix can also be used to encode the direction in which a stationary entity is located.

- (8.77) *tshəluŋla tʃu mənəɪ, təmi saŋlakzəkəɪ,*
 tshəluŋ-la tʃu mən-əɪ tə-mi saŋ-lak-zək-əɪ
 fox-F DIST sit-SEQ RL-tail dangle-DESCEND-SEND-SEQ
 ‘Fox, sitting with her tail dangling down, ...’

I noted earlier in §8.5.1.2 that it is possible for *-zək* to occur on a stem in the absence of a directional lexical suffix, and it is here that a second sense encoded by this suffix comes to the fore. When it occurs with transitive activity verbs whose argument structure potentially allows for a volitional agent – such as *thəp* ‘threw’, *tsə* ‘stamped on’, *kəp* ‘struck, shot’ and *tshə* ‘pounded’ – its presence can enhance the effect of the action upon the semantic patient. One consultant suggested that it ‘completes the task, makes it more severe’. The presence of *-zək* might therefore be interpreted as increasing the transitivity of an activity verb, analogously to the way the verb particles ‘off, away’ in ‘cut off, cut away’ add a certain finality to ‘cut’, in addition to emphasizing the resulting separation of parts.

(8.78) *təkhun tse lanthanzəkpi.*

tə-khun tse lan-thən-zək-pì

RL-neck DIST sever.in.one.blow-SEVER-SEND-IRR

‘The neck [of the chicken] would be severed.’

This is illustrated in the example of (8.78) above, which is taken from a Khensa Mongsen text describing a ritual that used to be performed when a new village was established. In this usage, *-zək* adds to the sense of physical separation of the chicken’s head from the rest of its body as a result of the action of cutting its neck, while simultaneously contributing to the notion of finality encompassing the whole activity.

A third sense of *-zək* is discernible when it occurs as a suffix on the verb root *asi?* ‘deceive’ in (6.47), repeated below and renumbered as (8.79). Although this is the only example of *-zək* occurring on a verb of this semantic class in the corpus, it is indubitably apparent that the meaning it encodes in the relativized verb stem is causation, literally ‘that caused him to be deceived’. The atypical order of suffixes that has the perfective *-tʃuk* preceding *-zək* is discussed in §8.5.7.

(8.79) *təi, tshəluŋla nə asiʔjukzəkpaʔ sə niʔ la phàɹajùʔ.*

tə-əi tshəluŋ-la nə asiʔ-juk-zək-pàʔ sə

thus-SEQ fox-F AGT deceive-PFV-SEND-NR ANAPH

niʔ la phàɹaj-ì-ùʔ

one.day TOP catch-IRR-DEC

In context: ‘And, the fox that deceived [him] one day will be caught.’

LaPolla (1994: 70–71) presents evidence that a verb meaning ‘send on an errand, entrust with a commission’ has independently grammaticalized as a post-verbal marker of causation in scores of languages and dialects of Tibeto-Burman. Few of these forms are cognate, leading him to surmise that their common grammaticalization as markers of causation is the consequence of “drift”, a term coined by Sapir (1921) to describe parallel historical developments in disconnected dialects of a language, long after they have become mutually unintelligible. The divergent and variable meanings of the suffix *-zək*, one of which evinces a causative meaning, suggest that Mongsen can be added to the large number of Tibeto-Burman languages in which a verb meaning ‘send’ has steered the same course of grammaticalization.

A segmentally identical form occurs as a main verb *zək* with the meaning of ‘sent’. The underlying tone of the main verb is low, while the tone of the grammaticalized directional suffix is mid. The following example demonstrates its main verb function and meaning.

- (8.80) *kiphuı nə áwkla khə ajila nət áhlu nə zək.*
 kiphuı nə a-úk-la khə a-ji-la nət a-hlú nə
 owner AGT NRL-pig-F CONJ NRL-dog-F two NRL-field ALL
 zək
 send.PST
 ‘[The] owner sent his pig and his dog to his field.’

8.5.4. Aspectual suffixes

Four suffixes expressing a range of aspectual meanings occur in paradigmatic opposition in the sixth slot of the verb stem. These are: the continuative aspect (§8.5.4.1); the repetitive aspect (§8.5.4.2); the completive aspect (§8.5.4.3); and the habitual past (§8.5.4.4).

8.5.4.1. Continuative aspect *-ja*

The continuative aspect suffix encodes a temporally unbounded state of affairs. It marks both activity and stative verbs.

- (8.81) *áwkla tʃu áhlù tʃu məıaŋəkə táŋ tʃhəniŋja?*
 a-úk-la tʃu a-hlú tʃu məıaŋ-əkə táŋ
 NRL-pig-F DIST NRL-field DIST be.hard-SIM just
 tʃhəniŋ-ja
 work-CONT.PST
 ‘The pig was just working hard in the field.’

- (8.82) *pa nə nəŋ tʃaja.ù?*
 pa nə nəŋ tʃa-ja-əı-ù?
 3SG AGT 2SG call-CONT-PRES-DEC
 ‘He’s calling you.’

The continuative aspect suffix frequently occurs in a dependent clause that indicates a simultaneous state of affairs, such as in *tʃāsija likə* of (8.83) below. The usual pattern is for the verb stem expressing the simultaneous activity or state to be suffixed by the continuative marker *-ja*; this is followed by the existential verb *li* ‘be’, which is marked by the simultaneous converb suffix. The resulting construction of [VERB STEM-CONT + be-SIM] creates a dependent clause and expresses the functional equivalent of imperfective aspect (see §11.4.1.2 for additional discussion and examples). Alternatively, the

continuative aspect marker may be used on a verb stem that forms a constituent of other types of dependent clauses, such as the contemporative converb clause of (8.84).

- (8.83) *tə tʃàsija likə inisala sə nə pa mataŋ sə nə tsəphakə waɪ.*
 tə tʃàsi-ja li-əkə inisaʔ-la sə nə
 thus be.distressed-CONT be-SIM rabbit-F ANAPH AGT
 pa ma-taŋ sə nə tsəpha-əkə wa-əɪ
 3SG FACE-SIDE ANAPH ALL fear-SIM go-PRES
 ‘And, while [Tiger] is upset, [aforementioned] Rabbit fearfully goes over to his side.’

- (8.84) *anu? ùphəla-la həntsə mukja lilikàʔ.*
 anu? ùphəla-la həntsə muk-ja li-lìkàʔ
 again red.jungle.fowl-F egg brood-CONT be-CONTEMP
 ‘Again, when Red Junglefowl was brooding on her eggs, ...’

A verb stem marked for continuative aspect and functioning as the main verb of an independent clause cannot be directly negated. This might be attributed to the fact that if it is asserted that VERB does not occur, then it is conceptually superfluous to also stipulate that the state or event asserted by VERB is continuing at the time that it does not occur. The non-occurrence of continuative aspect with negation suggests that there is a dependency operating between these two grammatical categories. Note how it is absent in the negated verb stem of the disjunctive interrogative construction in (8.85).

- (8.85) *pi kúták ku kàʔ thəpsija.ə sə məthəpsiɪ tə saw? təɪ.*
 pi kúták ku kàʔ thəpsi-ja-əɪ sə
 PROX heaven LOC also discuss-CONT-PRES QPTCL
 mə-thəpsi-əɪ tə sa-ùʔ təɪ
 NEG-discuss-PRES thus say-DEC REP
 ‘“Are [they] also discussing this in heaven?,” [he] asked.’

8.5.4.2. *Repetitive aspect -si*

The repetitive aspect is used to express ‘VERB again’. The question of whether it denotes that an activity is repeated just once or many times over is left open to interpretation. Context must be relied upon to establish the extent of repetition. To illustrate, (8.86) logically implies that the activity is repeated just once, whereas (8.87) implies that the activity must have been performed repeatedly,

an interpretation that is also encouraged by the presence of the simultaneous converb suffix.

- (8.86) *khutʃhaʔ nə ni.ə̀, imlaŋ məpʰù nə hla.ə̀, ajim kəmsi.*
 khutʃhaʔ nə ni-ə̀ im-laŋ məpʰù nə hla-ə̀
 PN AGT lead-SEQ village-bottom ward ALL descend-SEQ
 a-jim kəm-si
 NRL-village become-RPET.PST
 ‘Led by Kucha, they went down to [the site of] the lower ward and again established a village.’

- (8.87) *ajila nə áwk la nə tʃhə̀niŋpaʔ tə̀luʔ tʃu ku tʃə̀msiliʔ táŋ wa.*
 a-ji-la nə a-úk-la nə tʃhə̀niŋ-paʔ tə̀luʔ tʃu ku
 NRL-dog-F AGT NRL-pig-F AGT work-NR all DIST LOC
 tʃəm-si-liʔ táŋ wa
 trample-RPET-SIM just go.PST
 ‘Dog went and repeatedly trampled all [the work] that Pig had done.’

A possible lexical source of this suffix is the verb root \sqrt{si} ‘turn’. No examples of the repetitive suffix are found to occur on a negated verb root in the corpus, which may suggest a neutralization, although this remains to be tested.

8.5.4.3. *Completive aspect -maʔ*

The completive aspect encodes the completion of an activity. It only occurs with transitive verbs whose argument structure permits a patient that can be totally affected by the action expressed by its verb root. The completive aspect does not occur with semantically similar lexical suffixes encoding completive meanings, e.g. *-thəm* (§8.5.1.9) and *-lak* (§8.5.1.10).

- (8.88) *tʃhùwa tʃhuwakə atʃu tʃamaʔ.*
 tʃhùwa-tʃhuwa-əkə a-tʃu tʃàʔ-maʔ
 emerge-RED-SIM NRL-DIST consume-CMPL.PST
 ‘While coming out [from the jungle she], finished eating it up.’
- (8.89) *wə̀samali, thuku táŋ kəmə̀li lili.*
 wə̀saʔ-saʔ-maʔ-li thuku táŋ kəm-ə̀li li-li
 slice-SEPARATE-CMPL-ALT.IT nine just become-SEQ be-ALT.IT
 ‘[Every time he] finished cutting up [the magical meat], it kept becoming just nine [pieces], over and over ...’

Despite being nearly synonymous in meaning to the perfective suffix *-ʃuk* (see §8.5.7), the completive suffix is nevertheless compatible with the perfective aspect. The semantic restrictions applying to the completive and its greater proximity to the verb root suggest a relatively more recent grammaticalization. Conversely, the total lack of semantic restrictions on the use of *-ʃuk* for signalling perfective aspect on all semantic classes of verbs suggests that it belongs to an older layer of grammaticalization.

(8.90) *tə ənti.ɪəm taham ku, ahənzala intaŋ ʃu kàʔ ahizala nə məsàkə ʃfamajuk.*

tə ənti-ɪəm taham⁴ ku a-hən-za-la intaŋ ʃu
 thus path-MIDDLE MIDDLE LOC NRL-chicken-DIM-F thing DIST
 kàʔ a-hiʔ-za-la nə məsà-əkə ʃʃàʔ-maʔ-ʃjuk⁵
 even NRL-rat-dim-F AGT request-SIM consume-CMPL-PFV.PST
 ‘And in the middle of the road Rat Pup asked for and completely ate up
 even Little Chick’s share [of the fish].’

Consultants suggest that the difference between *ʃʃàʔ-maʔ-ʃuk* (consume-CMPL-PFV.PST) and *ʃʃàʔ-ʃuk* (consume-PFV.PST) is a matter of emphasis. The former expresses ‘completely ate up’, whereas the latter is merely ‘ate up’ without additionally implying that the entity is exhaustively affected, while still conveying the notion of perfectivity.

8.5.4.4. *Habitual past aspect -jaʔ*

The habitual past expresses ‘used to VERB’ and occurs with both stative and activity verbs. The habitual past and the continuative aspect have a superficially homophonous form when occurring in the stem-internal environment, because the glottal stop of the habitual past marker is deleted before a word-internal morpheme boundary (see §3.5). The underlying presence of the glottal stop becomes apparent when the habitual past marker is realized as the final syllable of the verb stem, as in (8.91) below.

4. *taham* is probably a variant of the relational body part noun *tə-ɪəm* ‘RL-middle’. It is used here in a nascent function as an auxiliary postposition (see §5.3.3 for description).

5. The perfective suffix has two allomorphs in free variation, viz. *-ʃuk* and *-juk* (see §2.2.1.7 and §8.5.7).

- (8.91) *pa nə, tsəŋkəm jàthùŋku, tsəmà? nə, təluk ni.ì, hjajaŋ waja?*
 pa nə tsəŋkəm jà-thùŋ-ku
 3SG AGT cool.season come-REACH-LOC.CV
 tsəmà? nə tə-luk ni-əɪ hjajaŋ wa-ja?
 Assam.Plain ALL RL-group lead-SEQ contract.wages go-HAB.PST
 ‘When winter was approaching, he used to lead a group to Assam
 looking for work.’
- (8.92) *tə hjutsə tʃu itə sanajaw? wà.*
 tə hjutsə tʃu i-tə sana-ja?-ù? wà
 thus story DIST PROX-thus speak-HAB.PST-DEC PTCL
 ‘Thus the story used to be told like this, as I remember.’
- (8.93) *nì nə pa íjá? tənìk mə.ɪəm.ɟatì.*
 nì nə pa íjá? tə-nìk mə.ɪəm-ja? atì
 1SG AGT 3sg much RL-eye be.red-HAB.PST PTCL
 ‘I used to be very jealous of him.’

8.5.5. Modality suffixes

Five suffixes expressing meanings associated with the grammatical categories of mood and modality occur in the zone of slot (7). These are: the abilitive marker (§8.5.5.1); the frustrative marker (§8.5.5.2); the conative marker (§8.5.5.3); the desiderative marker (§8.5.5.4); and the chance marker (§8.5.5.5). Most of these may occur together within the zone, e.g. *wa-tʃhət-tsə-mì-àɪ* (go-ABIL-CON-DESID-PRES) ‘wants to try to be able to go’. Semantically compatible modality suffixes allow variable orders with different meanings, e.g. *wa-tsə-tʃhət-mì-àɪ* (go-CON-ABIL-DESID-PRES) ‘wants to be able to try to go’. No restrictions on their use have been observed so far.

8.5.5.1. Abilitive modality -tʃhət

The abilitive modality marker expresses ‘can, able to’. This can relate to the performance of a physical activity, or a cognitive ability. The abilitive modality suffix can occur on any semantically-compatible verb stem. As the following examples demonstrate, its matrix clause verb stem can be questioned and negated.

(8.94) *nàŋ aki thùŋtʃhətəɪ sə məthùŋtʃhətla.*

nàŋ a-ki thùŋ-tʃhət-əɪ sə mə-thùŋ-tʃhət-la
 2SG NRL-house reach-ABIL-PRES QPTCL NEG-reach-ABIL-NEG
 ‘Are you able to reach the house, or not?’

(8.95) *təku tsəŋjutshəla sə anak ku li aj, məhlàtʃhət.ɪə.*

tə-ku tsəŋjutshə-la sə anak ku li aj
 thus-LOC.CV PN-F ANAPH top LOC be SINCE

mə-hlà-tʃhət-əɪ

NEG-descend+go-ABIL-PRES

‘And then, since Tsengrustela is at the top [of the tree], she is unable to descend.’

Some verb roots have combined with the abilitive modality suffix to form lexicalized meanings, e.g. *hŋa-tʃhət* (listen-ABIL.PST) ‘understood’ and *philəm-tʃhət* (think-ABIL.PST) ‘remembered, realized’. In the following textual example, the protagonist has grossly underestimated how many traditional gifts of pork he needs to distribute amongst his clansmen for his wedding. At the time that he is taking the meat to them, he realizes that he needs ninety pieces, not just the nine that he is carrying.

(8.96) *təɪ hənəɪ zəkəɪ, inti taham ku pa nə philəm-tʃhət.*

tə-əɪ hən-əɪ zək-əɪ
 thus-SEQ take-SEQ send-SEQ

inti taham ku pa nə philəm-tʃhət
 path MIDDLE LOC 3SG AGT think-ABIL.PST

‘And then, after [he] was sent carrying [the nine pieces of pork], on the way he realized.’

8.5.5.2. *Frustrative modality -phət*

The frustrative modality encodes a basic meaning of ‘VERB incorrectly, VERB ineffectually’. It is also sometimes used in the sense of ‘VERB incompletely’, or ‘VERB repeatedly’. It is perhaps from these other senses that the frustrative meaning has developed. In the following nominalized example from a Waromung Mongsen text, *kəp-phət-pàʔ* (shoot-FRUS.NR) is a derived noun meaning ‘a frustrated shooting’, i.e. a wounding caused by shooting when a kill was anticipated.

- (8.97) *təi kəptʃhəi apawŋpalá, pa ípá? kəpphətpà? tijimthəm tʃhəutəntsəku,*
 tə-əi kəp-tʃhəʔ-əi apawŋ-pála pa
 thus-SEQ shoot-CHANCE-SEQ wound-COND 3SG
ípá? kəp-phət-pà? tə-jimthəm tʃhəu-tən-tsə-ku
 EMPHAT shoot-FRUS-NR NZP-suffer be.sick-WHEN-DIST-LOC.CV
 ‘And then, having perchance been shot, if wounded, then while⁶ he is
 suffering from this wounding ...’

A sense of frustrated activity – in this case, the failure to do something to its completion – is apparent in the following two sentences, which were independently offered by consultants in their explanations of the use of *-phət*.

- (8.98) *pa aki tʃhəphətəi li.*
 pa a-ki tʃhə-phət-əi li
 3SG NRL-house make-FRUS-SEQ stay.PST
 ‘He didn’t finish building the house and left it in that state.’

- (8.99) *nì atʃa tʃəphət.*
 nì a-tʃa tʃəʔ-phət
 1SG NRL-cooked.rice consume-FRUS.PST
 ‘I didn’t eat my fill.’

When the frustrative marker occurs in a V_1 -FRUS *táj* V_1 sequence, as in the text example of (8.100) below, there is an additional sense of ‘just VERB over and over, without success’, with the reduplication of the verb root iconically encoding the repetition of the ineffectual activity.

- (8.100) *“wá aku, wá aku,” tə tʃəphət táj tʃa.*
 wá a-ku wá a-ku
 EXCLM VOC-uncle EXCLM VOC-uncle
 tə tʃə-phət táj tʃa
 thus call-FRUS just call.PST
 ‘“Wa, Uncle! Wa, Uncle!”, [she] ineffectually called out again and again.’

An identical construction was used in an example offered by a Mongsen consultant to demonstrate the various uses and meanings of *-phət*. This encodes

6. The suffix *-tən* of Waromung Mongsen expresses a similar meaning to *-thuŋ* ‘REACH’ in Mangmetong Mongsen when it is used on a dependent verb stem (see 11.4.3.7).

a repeated activity without directly expressing an associated sense of frustration, although perhaps this could be inferred, given a qualifying context.

- (8.101) *pa aki nə tʃhuwaphət táŋ tʃhuwaɪ.*
 pa a-ki nə tʃhuwa-phət táŋ tʃhuwa-əɪ
 3SG NRL-house ALL emerge-FRUS just emerge-PRES
 ‘He comes regularly to the house.’

8.5.5.3. *Conative modality -tsə*

The conative modality fundamentally expresses a meaning of ‘try to VERB’. It is frequently found on verb stems with the imperative suffix, as in the following examples. The first of these comes from a Khensa Mongsen text, hence the different phonological form of the distal nominal demonstrative and other morphemes in this sentence.

- (8.102) *nàŋ nə nəməluŋtsaŋ taɹúkthipa? ku lipa? tsə satsəaŋ átə.*
 nàŋ nə nə-məluŋtsaŋ taɹúk-thi-pà? ku li-pà? tsə
 2SG AGT 2SG.POSS-heart deep.SUP-INTNS-NR LOC be-NR DIST
 sa-tsə-aŋ atə
 say-CON-IMP PTCL
 ‘Please try to tell me what’s in the deepest part of your heart.’

- (8.103) *tsəŋti kupa? phinə itə tʃhaɪ hnihitsə.*
 tsəŋti kupa? phinə i-tə tʃhà-əɪ
 woven.bamboo.wall hole ABL PROX-thus do-SEQ
 hnihi-tsə
 peep-CON.PST
 ‘From a hole in the wall, doing it like this, he tried to peep.’

- (8.104) *tətʃhàku, tənì nə “məkəɹatsəɹu? pi.”*
 tətʃhàku tənì nə məkəɹa-tsə-əɪ-ù?
 thus-do-LOC.CV RL-wife AGT NEG-ascend+come-CON-PRES-DEC
 pi
 PROX
 ‘And so, the wife says [of her husband] “[He] isn’t attempting to come up [from the Assam plain], this one”.’

A closely related probative sense of *-tsə* is ‘test out, see for yourself’, e.g. *tʃəm-tsə-aŋ* (drink-CON-IMP) ‘try a sip’. This sense is implicit in the following textual example.

- (8.105) *təɪ liŋəɪ, “áhlù nə watsəaŋ,” tətʃhàw.ɪə.*
 tə-əɪ liŋ-əɪ a-hlú nə wa-tsə-aŋ
 thus-SEQ plant-SEQ NRL-field ALL go-CON-IMP
 tətʃhà-ù? tətɪ
 thus do.PST-DEC REP
 ‘And then, after [he] had done the planting, [she] said “Go and have a look at the field”.’

The likely diachronic source of this suffix is the root of the verb *atsə* ‘look’, from which a probative meaning of ‘witness, seek proof’, inherent in the meaning of the conative modality, is derived. This is in turn suggestive of a relationship with the grammatical category of evidentiality (see §4.2.14.2). No particular restrictions on its use are noted.

8.5.5.4. Desiderative modality *-mì*

The desiderative modality marker expresses ‘desire to VERB’. Verb stems marked by this suffix are compatible with any tense and can be negated. The desiderative mood suffix occurs without restriction on any semantic class of verb. Its diachronic origin is unknown.

- (8.106) *ipi la məkhimi.ù? tət.*
 ipi la mə-khì?-mì-əɪ-ù? tət
 PROX TOP NEG-give-DESID-PRES-DEC PTCL
 ‘These I don’t want to give!’
- (8.107) *tʃàmìɪ, təpakukà? khatsəpha áŋ məpuŋtshəŋəɪ.*
 tʃà?-mì-əɪ tət-pakukà? kha-tsəpha-a áŋ
 consume-DESID-SEQ thus-CONCESS be.bitter-be.afraid-SIM just
 mə-puŋtshəŋ-əɪ
 NEG-pounce-PRES
 ‘[Leopard cat] wants to eat [Rooster], but being bitterly afraid, [she] doesn’t dare pounce.’

- (8.108) *tʃɔ́páʔ tʃhàmìɪ la tʃhaŋ.*
 tʃɔ́páʔ tʃhà-mì-àɪ la tʃhà-aŋ
 what do-DESID-PRES TOP do-IMP
 ‘Do whatever you want!’

The desiderative suffix is one of the very few verbal markers privileged to occur on a verb stem that has been nominalized by the nominalizing prefix. For examples and description, see §7.4.1.

8.5.5.5. *Chance modality -tʃhaʔ*

The chance marker grammatically encodes an event that transpires by chance. It is compatible with all semantic classes of verb, and the verb stems it occurs on can be questioned and negated. The ‘chance’ meaning is demonstrated in the following elicited example.

- (8.109) *tùhɲàɪ hùŋtəpja likə, nì kàʔ jàktʃhawʔ.*
 tùhɲàɪ hùŋtəp-ja li-əkə nì kàʔ jàk-tʃhaʔ-ùʔ
 other fight-CONT be-SIM 1SG also beat-CHANCE.PST-DEC
 ‘While others were fighting, unexpectedly I was also beaten up.’

The next example ([6.52], repeated and renumbered below as [8.110]) comes from a folklore narrative in which a distressed barking deer is hopping around and accidentally tramples an earthworm that was sunning itself on the ground. In the context of the story, Barking Deer is unaware that she has injured Earthworm, but when told, she protests her ignorance and lack of intention with the help of *-tʃhaʔ*. Surprise, a lack of intention and happenstance are all suggested by the chance modality marker in this sentence.

- (8.110) *nì tətʃàsi ku jaŋtən líʔ wapaʔ tsəlàla tsətʃhawʔ.*
 nì tətʃàsi ku jaŋ-tən-líʔ wa-pàʔ
 1SG NZP-be.distressed LOC hop-TO.AND.FRO-SIM go-NR
 tsəlàʔ-la tsətʃhaʔ-ùʔ
 earthworm-F stamp.on-CHANCE.PST-DEC
 ‘In my distress in which I went hopping to-and-fro, [I] unknowingly stamped on Earthworm’ [said Barking Deer].

The chance marker shows some overlap with the grammatical category of mirativity (DeLancey 1997b, 2001; Lazard 1999) in that it can also encode surprise, although the grammatical encoding of new information is not

consistently central to the meaning of *-tʃhaʔ* in all the contexts of its use. This is particularly apparent in the irrealis examples of (8.111) below, in which any notion of ‘new information’ or ‘surprise’ that is fundamental to the category of mirativity is completely unrelated to the meaning encoded by *-tʃhaʔ*. That notwithstanding, *-tʃhaʔ* does sometimes encode surprise brought about by the assimilation of new information. The following exchange between two characters in a narrative offers evidence that although its range of meanings is often similar, *-tʃhaʔ* cannot specifically be marking mirativity as defined by DeLancey (1997b, 2001). In the following excerpt from a text, Imtisangba has heard that Noksensangba has magical knowledge and decides to test him. After inviting him up to his house, Imtisangba pours rice beer into two cups and asks Noksensangba to predict whether he will get the opportunity to drink the rice beer in his cup.

- (8.111) a. *“thajip i aji tʃəmʃhaj sə mətʃəmʃhaj.”*
 tha-jip i a-ji tʃəm-tʃhaʔ-ì sə
 now-morning PROX NRL-rice.beer drink-CHANCE-IRR QPTCL
 mə-tʃəm-tʃhaʔ-ì
 NEG-drink-CHANCE -IRR
 “‘This morning will you get to drink the rice beer or not?’”
- b. *təlikàʔ “nì tʃəmʃhaj ùku.”*
 tə-likàʔ nì tʃəm-tʃhaʔ-ì ùku
 thus-CONTEMP 1SG drink-CHANCE-IRR PTCL
 ‘When [he asked that, Noksensangba replied] “I’ll get the opportunity to drink alright.”’
- c. *“nàŋ mətʃəmʃhaj atì.”*
 nàŋ mə-tʃəm-tʃhaʔ-ì atì
 2SG NEG-drink-CHANCE-IRR PTCL
 “‘You certainly won’t get the opportunity to drink, though.’”

When Imtisangba goes to pick up his cup to prove Noksensangba wrong, he accidentally knocks it over and spills the rice beer, thus proving the prediction to be correct. There is no surprise or new information encoded in the sentences of this conversation. Rather, what the chance marker is adding to the meaning of sentence (a) is literally ‘What are the chances that VERB will occur?’ and this allusion to chance is encoded by the chance marker on the verb stems of sentences (b) and (c).

Lastly, one day when discussing the location of the Amguri tea garden in Upper Assam with my Mongsen teacher, the conversation reminded me that I had previously read about this particular tea garden in Mills (1926). Upon mentioning this, he rephrased what I had said, but added the chance modality

marker to the verb stem to signal that I had come across this information by chance, since it was not my intention to read specifically about the Amguri tea garden. This information was not new, nor was it surprising; it was merely something that I happened upon by chance in the past.

- (8.112) *nàŋ nə tʃe pi mils nə zəlupa? kakət ku amku.i səŋàni zəŋtʃha?*
 nəŋ nə tʃe.pi.mils nə zəlu-pà? kakət ku
 2SG AGT J.P.Mills AGT write-NR book LOC
 amku.i səŋàni zəŋ-tʃha?
 Amguri tea.garden read-CHANCE.PST
 ‘You read about the Amguri tea garden by chance in the book that J.P. Mills wrote.’

In summary, we have seen that the chance marker and the category of mirative have some aspects of their respective meanings in common, but there are also points of difference, the chief divergence being that the chance marker does not always encode new information, the hallmark of mirativity. The fundamental meaning linking all the attested uses of the chance marker is that VERB happened, happens or will happen because of chance, and that in many contexts such outcomes are unexpected and result in surprise. This is not an obligatory requisite for its use, however.

Some possible lexical sources for the chance marker are the verb roots *taŋtʃhà?* ‘touch’, or *tʃhà?* ‘trap, catch’. Note that the latter verb root can also be used in the sense of ‘be affected by alcohol, be intoxicated’. Examples of their lexical functions are provided in the following sentences (these were uttered spontaneously in response to some video clips that I used to elicit data).

- (8.113) *tujət luma ku sá? taŋtʃhàtəpə.ù?*
 tujət luma ku sá? taŋtʃhà?-təp-əi-ù?
 2DU hair LOC again touch-RECIP-PRES-DEC
 ‘They are touching each other’s hair again.’

- (8.114) *tujət i aji nə tʃhà? sé.*
 tujət i a-ji nə tʃhà? sé
 2DU PROX NRL-rice.beer INST be.intoxicated QPTCL
 ‘Are these two drunk?’

8.5.6. Resultant state suffix *-tʃak*

The suffix *-tʃak* functions synchronically as a resultant state marker. It is used to express the resultant state of activity verbs, stative verbs and verbs of posture,

such as *ɬə-tʃak-kə* (lie-RS-SIM) ‘be lying’ and *mətuŋ-tʃak* (be.erection-RS) ‘was erect’. A verb stem containing the resultant state suffix can be negated and questioned.

- (8.115) *nàŋ tʃəpá? tʃhàpà?, nì áhlú sùsajakí?*
 nàŋ tʃəpá? tʃhà-pà? nì a-hlú sùsa-tʃak-i?
 2SG what do-NR 1SG NRL-field scratch-RS-CAUS.PST
 ‘Why did you do this – leave my field scratched up?’

- (8.116) *atʃu ku tʃutə məzüksajakithuŋtʃükü tənəm tənəm i itəj tʃhuwa.*
 a-tʃu ku tʃu-tə
 NRL-DIST LOC DIST-thus
 məzük-sa-jak-i?-thùŋ-tʃu-ku
 crumple-SPREAD-RS-CAUS-REACH-DIST-LOC.CV
 tənəm tənəm i i-tə i tʃhuwa
 mountain mountain PROX PROX-thus PROX emerge.PST
 ‘There, when [the god Lichaba] caused the earth to be left crumpled up like that, mountains like these ones here emerged.’

The resultant state suffix has grammaticalized from the root of a main verb *tʃak*, meaning ‘leave, abandon, keep’.

- (8.117) *maŋmətuŋ tʃakɪ túnkhəla kinijunəɪ hlàɪ kəm.*
 maŋmətuŋ tʃak-i?-əɪ túnkhəla kinijunəɪ
 village.name leave-CAUS-SEQ 3PL village.name
 hlà-əɪ kəm
 descend+go-SEQ become.PST
 Having left [the corpse of Yasa Naro in] Mangmetong village, they went down and founded Kiniunger village.

The resultant state marker has two allomorphs in free variation: *-jak* ~ *-tʃak*. This makes it suspiciously similar to the perfective suffix (§8.5.7), which also has two allomorphs in free variation: *-juk* ~ *-tʃuk*. Furthermore, both of these aspectual categories appear to have grammaticalized from verbs with similar meanings and use the same initial consonants and codas in their allomorphic variants, differing only in their vowels. The palatal approximant and the palatoalveolar affricate are otherwise in robust contrast, therefore I treat the *tʃ* ~ *j* free variation as a suspension of a phonemic contrast that is limited to just these two aspectual markers (see §2.2.1.7 for further discussion and examples).

The fact that the resultant state marker and the perfective marker can be concatenated indicates that they represent different categories synchronically, even if they do share the same diachronic source. The following elicited example demonstrates their compatibility.

- (8.118) *pa nə usət təu təmáŋ mijimtʃaktʃuk.*
 pa nə usət tə-əu təmáŋ mijim-tʃak-tʃuk
 3SG AGT stuff NZP-good all hide-RS-PFV.PST
 ‘He left all the good stuff hidden.’

The stem position of the perfective marker on the outside of the resultant state marker suggests that it is the older of the two grammaticalized morphemes, according to the proposed inside-out model of grammaticalization involving verb stem morphology (see §8.5.1 for discussion).

8.5.7. Perfective aspect suffix *-tʃuk*

The perfective aspect marker occupies its own slot in the verb template. It is likely to be one of the oldest of layers of morphemes to have grammaticalized from lexical roots, due to its peripheral position of occurrence in the predicate template. It is compatible with all of the lexical suffixes (§8.5.1), the valency-modifying suffixes (§8.5.2 and §8.5.8), all of the aspectual suffixes (§8.5.4), any tense or mood, including irrealis (§8.5.12), and any semantic class of verb.

Like the resultant state marker, the perfective suffix has two allomorphs in free variation – *-tʃuk* ~ *-juk* – with the palatal approximant form being more common, but not obligatory, in the environment after a high front vowel. It encodes two types of related meanings, depending upon the semantic class of verb. For many verbs, the perfective marker denotes closure of the activity or event, indicating that the predication is to be viewed in its bounded entirety.

- (8.119) *təɪ təmhnak tʃu sətʃuk.*
 tə-əɪ tə-əmhnak tʃu sə-tʃuk
 thus-SEQ RL-son.in.law DIST die-PFV.PST
 And, the son-in-law passed away.

- (8.120) *nì kàʔ jah.ɪtʃukukù.*
 nì kàʔ jah.ɪ-tʃuk-ukù.
 1SG also be.one’s.fill-PFV-ANT
 ‘I’m completely fed up [with this].’

When used with stative verbs, the perfective suffix can encode either closure or inception, i.e. entry into the state predicated by its verb stem. This semantic difference might be used as another criterion to justify the recognition of stative intransitive verbs as a separate sub-class of intransitive verb (see §8.1.1). The inception of a state is apparent in the following sentence from a folklore narrative, which explains how the owl's face became flat.

- (8.121) *pa tələm ku tsəpànə, tələm tʃu apaktʃuk.*
 pa tə-ləm ku tsə-pànə
 3SG RL-head LOC peck-CAUSAL
 tə-ləm tʃu apak-tʃuk
 RL-head DIST be.flat-PFV.PST
 'Because [the other birds] pecked on [her] head, it became flat.'

The perfective aspect marker demonstrates a different position of occurrence in the nominalized verb stem of *asiʔ-juk-zək-pàʔ* (deceive-PFV-SEND-NR), seen in the sentence of (8.79), because it normally occurs to the right of the directional suffix *-zək*. Its atypical position here could possibly be attributed to an evolving lexicalization of *tʃuk* as an unsegmentable formative of *asijuk* 'deceive'. Speakers also permit the perfective suffix to follow *-zək* in this verb stem without noticing a change of meaning, and *asiʔ* can still be used independently of the perfective suffix with the meaning of 'deceive'.

The perfective marker is also peculiar for being compatible with the continuative aspect marker *-ja*, as demonstrated by the following example. This was uttered by a narrator in an aside to the audience. He had just reached the part of his narrative where he was supposed to sing a song, but was struggling to remember the lyrics. After a prolonged pause, he uttered this sentence. The combination of perfective and continuative aspectual suffixes on this verb stem suggests an entry into a state of forgetfulness, and that the state of forgetfulness is still current. The resulting difference in semantic scope (as in the sense of Rice 2000) might account for the unusual order of the perfective and continuative suffixes in the verb stem.

- (8.122) *liəŋ ní, nì alaktʃukjaùku pi.*
 li-aŋ ní nì alak-tʃuk-ja-ə̀-ùʔ ùku pi
 stay-IMP PTCL 1SG forget-PFV-CONT-PRES-DEC PTCL PROX
 'Just wait a moment, OK? I'm completely forgetting this.'

The lexical origin of the perfective suffix is the transitive verb root *tʃuk* 'leave, abandon, keep', which is attested synchronically as a main verb, although it is noteworthy that all examples in the text corpus occur with the

causative suffix. While this is not obligatory, it may indicate an incipient narrowing of its lexical function.

(8.123) *ə nàŋ la aki táŋ ku tʃukəjù?*

ə nàŋ la a-ki táŋ ku tʃuk-iʔ-ì-ùʔ
EXCLM 2SG TOP NRL-house just LOC keep-CAUS-IRR-DEC
‘Er, as for you, you’ll be kept just in the house.’

(8.124) *təɪ, pə sə tʃukɪ wawkù.*

tə-əɪ pə sə tʃuk-iʔ-əɪ wa-ukù
thus-SEQ 3SG ANAPH leave-CAUS-SEQ go-ANT
‘And abandoning her, [he] went.’ (Literally: ‘Causing her to be left.’)

8.5.8. Causative suffix *-iʔ*

The causative suffix occurs in the penultimate slot of the verb stem. This is to the right of the lexical, aspectual and modality suffixes, and to the left of the absolute tense markers, the imperative mood suffix and the converb suffixes of slot (11). The syntax of morphological causatives is described in detail in §5.4.1. The causative marker has a number of different phonological shapes in other varieties of Mongsen, e.g. its form is *-ɰʔ* in Waromung Mongsen, but *-piʔ* in Khensa and Mekhuli Mongsen. Its diachronic source is unknown.

(8.125) *səjukəjuʔ, əntiɪəm məthùŋla səjukəjuʔ.*

sə-tʃuk-iʔ-ì-ùʔ ənti-ɪəm mə-thùŋ-la
die-PFV-CAUS-IRR-DEC path-MIDDLE NEG-reach-NEG.CV
sə-tʃuk-iʔ-ì-ùʔ
die-PFV-CAUS-IRR-DEC
‘[I] will make him die; before he arrives [home], in the middle of the road I will make him die.’

The causative suffix shows a different position of occurrence in just one example in the corpus. The following sentence (originally [5.57], renumbered as [8.126] below) was discussed in §5.3.2.4 with respect to its unusual choice of pronoun for a dative-marked noun phrase argument. This example is elicited, but has nevertheless been included here in the event that it proves that the causative suffix has a legitimate alternative position in the verb stem. Given that all speakers consulted considered the choice of the first person singular personal pronoun *nì* to be an ungrammatical head of a dative-marked noun phrase, it may

well be the case that the structure of the verb stem is also of questionable grammaticality, although this is something that remains to be confirmed.

- (8.126) *asaŋla nə ahən tʃu { *nì ; kə } li məɹəkimi.ù?*
 asaŋ-la nə a-hən tʃu { *nì ; kə } li
 PN-F AGT NRL-chicken DIST { 1SG ; 1SG.POSS } DAT
 məɹək-iʔ-mì-àɹ-ùʔ
 clean-CAUS-DESID-PRES-DEC
 ‘Asangla wants me to dress the chicken.’

8.5.9. Negative suffix *-la*

The negative suffix occurs in slot (11), in obligatory combination with the negative prefix (§8.3.3) when the temporal setting is past. Negation is otherwise encoded by the negative prefix alone. It seems that a negative suffix does not occur in the Chungli dialect, judging from the grammatical descriptions of Clark ([1893] 1981) and Gowda (1975), and the Chungli translations of Mongsen negated past tense verbs in Text 1.

- (8.127) *təɹà nə mathàlà.*
 təɹà nə mə-athà-la
 little INST NEG-fall.down-NEG.PST
 ‘[He] missed falling down by a bit.’

It may be significant that the negative suffix *-la* (NEG) has the same phonological form as the Mongsen topic particle *la* (see §4.2.14.3 for description). Hargreaves (1996) observes that Kathmandu Newar uses a clause-final speech act particle *la* to mark a yes/no interrogative clause, and that a homophonous form is also used as a topic particle in that language. He presents evidence that this is a reflex of a Proto-Tibeto-Burman interrogative particle **la*, which Matisoff reconstructs on the basis of comparative evidence from Lahu, Written Burmese and Meithei ([Matisoff 1988: 1347], as cited in Hargreaves [1996: 32]). Like Kathmandu Newar, the initial *l-* of Mongsen demonstrates a fairly consistent phonological correspondence with PTB initial **l-*, e.g. *làtà* ‘moon’ < PTB **s-la~g-la* (STC #144), *a-li* ‘NRL-flea’ < PTB **s-liy* (STC #440), and *ləp* ‘cut’ < PTB **lep* (STC #351).

A widely attested cross-linguistic pattern is for an interrogative particle to correlate with a negative marker, such as the Thai preverbal negative particle *mâj* and the clause-final interrogative particle *máj*, and the English negative tag in the question *You’re staying, no?*. Now, if the negative suffix and topic

particles of Mongsen have followed the same grammaticalization pathway as the negative morphemes of Thai and English, then it is possible that their origin might be traced to the PTB particle **la*. The work of Hargreaves (1996) furthermore suggests that the pathway of reanalysis extends to the development of a topic marking function. Note that there is no conclusive synchronic evidence that *la* functions as an interrogative particle in Mongsen, or that it ever has. Although this morpheme does occur clause-finally in the disjunctive type of interrogative clause, it seems to contribute only to the negated past tense meaning (see §5.1.2.1 for description).

- (8.128) *pi mətət sə məmətətla.*
 pi mətət sə mə-mətət-la
 PROX know QPTCL NEG-know-NEG.PST
 Do you know this or not?"

That being said, the function of *la* in the following narrative example is ambiguous. Here it may be topicalizing the whole of the preceding clause, or it could be contributing to marking the interrogative mood of the clause.⁷ Either way, its sentence-final position of occurrence perhaps offers some insights into how such a morpheme might undergo a reanalysis and extension of function.

- (8.129) *pi mətəti sə məmətəti la.*
 pi mətət-ì sə mə-mətət-ì la
 PROX know-IRR QPTCL NEG-know-IRR TOP?/QPTCL?
 Would you know this or not?"

8.5.10. Positive imperative suffix *-aŋ*

The positive imperative mood suffix occurs in slot (11), in paradigmatic opposition with the negative past suffix, the converb suffixes, the absolute tense markers and the general nominalizer. It is discussed at length in §10.2.1.

The positive imperative mood has some incompatibilities with other verbal morphology, such as those that encode conflicting categories of mood (see Table 10.1 for a summary). Like the prohibitive mood prefix, it may co-occur with an extensive range of lexical and aspectual suffixes (see §10.3.3 for examples).

7. See line 44 of Text 3 for the full context in which this interrogative clause was used.

8.5.11. Converb suffixes

A characteristic pattern of Mongsen clause linkage involves a type of structural dependency formally encoded by what has increasingly come to be known as the “converb” in recent typological literature (e.g. see Haspelmath and König [1995] and references therein).

The Mongsen converb suffixes listed in Table 8.3 below have cross-references to their relevant sections in Chapter 11, where they are described in detail. None of these permit any other categories of slot (11) to co-occur, but the simultaneous converb and the sequential converb freely admit aspectual markers, valency modifying morphemes and lexical suffixes to the left of their position in the verb stem. Only some of the converb markers allow their verb stems to be directly negated by affixal morphology, notably those that are not used for narrative chaining (see §11.4.1).

Table 8.3. Converb suffixes

FORM	GLOSS	SEMANTICS
-əɹ	(SEQ)	sequential activity, anterior event, sometimes temporal qualification or reason (§11.4.1.1)
-əkə; -liʔ; -a	(SIM)	simultaneous activity, attendant circumstance, sometimes sequential activity or reason (§11.4.1.2)
-li	(ALT.IT)	iterative alternation of two activities (§11.4.1.3)
-pàla	(COND)	irrealis conditional (§11.4.3.1)
-kùla	(CIRCM)	circumstantial/hypothetical and realis conditional (§11.4.3.2)
-pàkukàʔ	(CONCESS)	concessive conditional (§11.4.3.3)
-likàʔ	(CONTEMP)	actualized contemporaneous event (§11.4.3.4)
-tʃən	(DUR)	durative activity (§11.4.3.5)
-pànə	(CAUSAL)	consequential circumstance (§11.4.3.6)
-ku	(LOC.CV)	temporal sequence, simultaneous activity/state (§11.4.3.7)

A few of the converb suffixes (such as allomorphs of the simultaneous converb suffix) express aspectual meanings that are similar to the aspectual suffixes described in §8.5.4, yet are able to co-occur with them (for examples and discussion, see §11.4.1.2). What formally differentiates a converb from other verbal morphology denoting similar meanings is the criterion that a converb marker cannot occur on the stem of a matrix verb. This allows a clear

division to be drawn between independent finite and dependent non-finite verb forms.

The converb markers have a variety of diachronic sources. In §7.4.4 I suggested that the sequential converb originates from the agentive nominalizer, on the basis of its phonological form and the cross-linguistic tendency of nominalizers to grammaticalize as clausal subordinators. Nominalizing morphemes also appear to have provided diachronic sources for some of the other converb markers. A formative *pa* that is likely to be related to the general nominalizer *-pàʔ* (§7.5.2), for example, is recognizable in the conditional converb marker *-pàla* and the causal converb marker *-pànə*. The case marking clitics also appear to be represented in the phonological shape of a number of converb suffixes. A form identical to the locative *-ku* functions alone as a marker of dependent temporal clauses, and it also occurs in the circumstantial *-kùla* and the concessive *-pàkukàʔ*. The agentive/instrumental marker *nə* is a likely partial source of the causal converb marker *-pànə*. A comprehensive description of the status and functions of all converb markers is deferred until Chapter 11.

8.5.12. Tense and modality

The fourth verbal category of slot (11) is subsumed by tense and modality. The absolute tense system makes a tripartite temporal distinction between an unmarked past, a formally marked present, and a formally marked immediate future, with the use of the latter category restricted to a speaker possessing the epistemic authority required to assert that an event is imminent. The irrealis category is otherwise used to index non-actualized events (see §8.5.12.2).

I suspect that Mongsen originally had just a simple mood marking system that contrasted a default unmarked realis mood with the overtly marked irrealis mood category. We see evidence of this mood marking function in nominalized verb stems functioning as relative clauses. In §6.6.5 I demonstrated that the irrealis suffix *-i* can be used on the nominalized verb stem of a relative clause to encode a non-actualized event, and that the absence of this marker expresses the default realis mood of an actualized event. It appears that this mood marking system was subsequently re-jigged as part of an absolute tense system, possibly triggered by the reanalysis of one or more nominalizers as absolute tense markers.

A synchronic description of the Mongsen tense system must recognize the function of *-əʔ* as marking present tense on finite verb stems in declarative clauses, and this is also a strong intuition shared by native speakers; but it is nonetheless very suspicious indeed that this suffix has an identical segmental shape to the agentive nominalizer (see §7.4.4). The agentive nominalizer has a

variety of extended functions in addition to the nominalization of verbs and nouns. It is also used as a relativizer (§6.6.3), as a genitivizer (§7.3.3), as a masculine semantic gender marker (§7.4.4), and as a sequential converb marker in dependent clauses (§11.4.1.1). My personal hunch is that the present tense marker *-əɿ* also has its diachronic source in the agentive nominalizer *-əɿ*, and that tense marking represents one more extended function of this morpheme. It is not uncommon cross-linguistically for nominalized verb forms, particularly participles, to contribute to tense and aspect oppositions. The English sentence *She is eating*, for example, uses the present participial form of a verb and a copula to express a present tense event. In a language like Mongsen that does not require a copula for such sentences, it is not unfeasible that an equivalent expression also relied upon a nominalized verb stem, prior to a reanalysis of the function of its nominalizing suffix taking place. Non-embedded nominalizations that function as independent clauses appear to be a common feature of Tibeto-Burman languages and have been reported in Lahu (Matisoff 1972), Athpare (Ebert 1997), Yamphu (Rutgers 1998), and Kham (Watters 2002).

At present I am not entirely certain of how the immediate future marker came to represent its temporal category. It could possibly be diachronically related to the segmentally identical purposive nominalizer *-ɿu*, described in §7.4.7. Conceptually, a purposive marker strikes me as an ideal target for a reanalysis resulting in the emergence of an immediate future, because a purposive similarly makes an allusion to an intention. Furthermore, if an extended function of the agentive nominalizer does turn out to be the marking of present tense, then for parallel reasons it seems logical that another nominalizer might have followed a similar grammaticalization pathway that led, in this case, to marking the immediate future tense.

The irrealis marker continues to represent a mood category, since it can be used to encode non-actualized events or states in the past, in addition to those situated posterior to the moment of speaking. But as Chung and Timberlake (1985: 206) point out, differences in mood and aspect are inherently correlated with temporal distinctions. Thus, correlations can generally be found between irrealis mood and future tense, realis mood and non-future tense, present tense and imperfective aspect, past tense and perfective aspect, and so on. It is these inherent correlations that must have facilitated the incorporation of the older mood system into the grammaticalized absolute tense system of Mongsen. The unmarked category is now associated with past tense, and the irrealis category continues to be used in past tense settings to encode events that were not actualized, as well as those that are predicted to occur after the moment of speech. On the basis of these functions, the mood categories appear to fall into neat paradigmatic opposition with the categories encoded by the present tense and immediate future markers.

Yet another peculiarity of the synchronic tense system of Mongsen is its typologically rare unmarked past tense.⁸ This seems a lot less strange if we assume that the unmarked category originally contrasted with the marked irrealis category in a prior system of mood marking. After the reanalysis of the agentive nominalizer *-əɪ* as a present tense marker and the purposive nominalizer *-ɪu* as an immediate future tense marker, the irrealis marker *-i* continued to be used for encoding non-actualized events, and the erstwhile formally unmarked realis category served to fill the slot of the past tense category in an emergent absolute tense system. Admittedly, much of this analysis is unavoidably speculative. The primary intention here is to shed some light on the character of tense/mood marking in Mongsen, and to sketch an outline of how the synchronic tense/mood system could have historically evolved. It goes without saying that a good deal more historical comparative research in related languages will be required before we have the benefit of the full picture.

Because of the dependencies shared by the markers of past and present tense, these two categories will be discussed together for ease of comparison, after which we will turn to an examination of the immediate future and irrealis markers. A description of the anterior tense/aspect marker concludes the discussion of the tense, aspect and modality categories.

8.5.12.1. *Categories of past and present*

Past tense encodes an event that takes place prior to the time of speech. It is the formally unmarked category in the tense system. Past tense is signalled by a bare verb stem in declarative clauses of positive polarity, and indirectly by the presence of the negative suffix in negative declarative clauses.

Past declarative:

- (8.130) *tətʃhaɪ, wàzàʔ təmáŋ sə jajaɪ, pa thaku jàk.*
 tə-tʃhà-əɪ wàzàʔ təmáŋ sə jaja-əɪ
 thus-do-SEQ bird all ANAPH be.angry-SEQ
 pa thak ku jàk
 3SG PLACE LOC beat.PST
 ‘And then, all the birds got angry and beat on her.’

8. A formally unmarked past tense/realis category is actually not uncommon in Tibeto-Burman languages of Nagaland, being also found in Chang, Khamniungan and Yimchungrü.

Negated past declarative:

- (8.131) *tsəpàʔ məsəməkə ìnti phithùŋʔfuku, pa mə.ɪà.là.*
 tsəpàʔ məsəm-əkə ìnti phi-thùŋ-ʔʃu-ku
 pond clean-SIM path clear-REACH-DIST-LOC.CV
 pa mə-ɪà-la
 3SG NEG-come-NEG.PST
 ‘When [they] were cleaning the pond and clearing the path, she didn’t come.’

Speakers often use the same markers to encode present and past tense in negated clauses, as demonstrated by (8.132) and (8.133) below. The fact that a structurally identical negated stem conveys a past tense meaning in (8.132) and a present tense meaning in (8.133) suggest that a conflation occurs between the two tense categories in negated declarative clauses. This is reported to be a widespread typological phenomenon. In a study of dependencies between grammatical systems, Aikhenvald and Dixon (1998: 64) find that many languages contrast more tense oppositions in the positive than in the negative.

Negated past declarative:

- (8.132) *thaŋtsəla nə “wa! nì məmətətlàw?”*
 thaŋtsə-la nə wa nì mə-mətət-la-ùʔ
 tree.shrew.sp.-F AGT EXCLM 1SG NEG-know-NEG.PST-DEC
 ‘Squirrel said “Wa! I didn’t know!”’

Negated present declarative:

- (8.133) *tətʃhàku, “wa mà.ɪulàw? mà.ɪulàw?, pùkphu-la nàŋ mà.ɪulàw?”*
 tət-ʃhà-ku wa mə-ɪu-la-ùʔ
 thus-do-LOC.CV EXCLM NEG-good-NEG.PST-DEC
 mə-ɪu-la-ùʔ pùkphu-la nàŋ mə-ɪu-la-ùʔ
 NEG-good-NEG-DEC owl-F 2SG NEG-good-NEG-DEC
 ‘Upon [Owl] doing that [they all said] “Wa, Bad! Bad! Owl – you’re bad.”’

In one type of polar interrogative clause, a suffix *-ɪ* is used for both past and present categories, therefore the temporal setting must be interpreted from the context. This may represent a vestige of the original nominalizing function of *-ɪ*, because it is clear that the present tense is not specifically expressed by this marker in this type of clause. I have consequently glossed *-ɪ* as the agentive nominalizer in the following example.

Past/present interrogative:

- (8.134) *atʃa tʃuŋəɪ pà,*
 a-tʃa tʃuŋ-əɪ pà
 NRL-cooked.rice eat.meal-ANOM QPTCL
 (i) ‘Are you eating your meal?’
 (ii) ‘Did you eat / have you eaten your meal?’

In contrast, content questions formed with interrogative pronouns can be formally marked for both past and present.

- (8.135) *nàŋ tʃəpá? nə wa / waɪ.*
 nəŋ tʃəpá? nə wa / wa-əɪ
 2SG what ALL go.PST go-PRES
 ‘Where did you go?’ / ‘Where are you going?’

In declarative clauses of positive polarity, the present tense marker encodes events or states that occur contemporaneously with the speech event, generally with some leeway on either side of the speech act.

Present declarative:

- (8.136) *atʃu tsəŋhɪəm, jəhuwa tsəŋhɪəm nə atsəjaɪù?*
 a-tʃu tsəŋhɪəm jəhuwa tsəŋhɪəm nə atsə-ja-əɪ-ù?
 NRL-DIST deity Jehovah deity AGT look-CONT-PRES-DEC
 ‘That is the god Jehovah watching.’

An alternative means of negating a present tense stem to that demonstrated by (8.133) is to attach just the negative prefix to a stem marked by the present tense morpheme *-əɪ*. Unlike stems negated discontinuously by *mə-...-la*, which permit pragmatically determined past or present temporal interpretations, a negated verb stem that retains the present tense marker unambiguously situates its event in the present.

Negated present declarative:

- (8.137) *pa aki tʃu nə mùwatʃhətəɪ.*
 pa a-ki tʃu nə mə-wa-tʃhət-əɪ
 3SG NRL-house DIST ALL NEG-go-ABIL-PRES
 ‘He is unable to go to his house.’

The present tense suffix is also used as a marker of timeless habitual/generic expressions (cf. the formally distinct habitual past marker of §8.5.4.4). The following example is taken from a Waromung Mongsen text.

Present habitual:

- (8.138) *nì na mənəŋ tʃàɪ.*
 nì na mənəŋ tʃàɪ-əɪ
 1SG AGT first consume-PRES
 ‘‘I eat [the grain] before [anyone else]’’, [said Rat].’

8.5.12.2. *Categories of irrealis/immediate future*

The most frequently used finite marker of non-actualized events in Mongsen is the suffix *-ì*. Since it can also be used for marking non-actualized events in the past, it must represent an irrealis mood category.

Irrealis:

- (8.139) *tsəŋ.tutshəla tʃu tshəj.*
 tsəŋ.tutshə-la tʃu tshə-ì
 PN-F DIST take-IRR
 ‘[I] will take Tsengrutsela.’

- (8.140) *muli tʃu phàtshəŋì atə.*
 muli tʃu phà-tshəŋ-ì atə
 drug DIST sprinkle-ATTACH-IRR PTCL
 (Speaking of making rice beer) ‘The medicine (i.e. yeast) should be sprinkled [on the fermented rice].’

The irrealis marker occurs on a verb stem with an auxiliary verb *liw?* to express the epistemic modality of irrealis events that will possibly occur after the moment of speech, e.g. *wa-ì liw?* (go-IRR might) ‘might go’, but is not used for expressing the possibility of an event having occurred before the moment of speech, e.g. *mə-wa-la liw?* (NEG-go-NEG.PST might) ‘might not have gone’. The formative *li* seen in the auxiliary *liw?* is probably related to the existential copula (see §9.2).

Deontic modality is encoded by the irrealis suffix and the nominalizer suffixed to the verb stem, e.g. *wa-ì-pà?* (go-IRR-NR) ‘should go’. Incidentally, this morpheme combination is identical to one type of nominal inflection used for purposive nominalizations (§7.4.7) and could alternatively mean ‘to go’, given an appropriate context and clausal structure. A clause-final nominalized verb stem that expresses an obligation to perform an act is the only clause type of Mongsen that appears superficially to function as a non-embedded nominalization (example [8.141] comes from a Waromung Mongsen text).

- (8.141) *nì pi la ápá? kha a luɪ taŋ na ənəɪ tshəkaɪpàw?*
 nì pi la a-pá? kha a-~~ɪ~~ luɪ
 1sg PROX TOP VOC-father CONJ VOC-mother and.all
 taŋ na ən-əɪ tshəka-~~ɪ~~-pà-ù?
 SIDE ALL take-SEQ emerge-IRR-NR-DEC
 ‘As for this, I should return, taking [it] to Father and Mother and all.’
- (8.142) *ami? hùŋtəpəɪ ahju pentəŋ thùŋku tʃàjpá?, alà hənsətəɪ, tʃàjukípà?*
 a-mi? hùŋ-təp-əɪ a-hju pentəŋ thùŋ-ku
 NRL-person fight-RECIP-SEQ NRL-word judgement reach-LOC
 tʃà?-ì-pà? alà hənsət-əɪ tʃà?-juk-ì-pà?
 eat-IRR-NR CONJ kill-SEQ eat-PFV-IRR-NR
 ‘When people fight with each other, upon [the village council] reaching a judgement [a pig] must be eaten. And, after it is killed, [it] must be eaten up.’⁹

This structure should in fact be viewed as a type of verbless clause in which the verb stem modified by the irrealis suffix and the general nominalizer functions as a verbless clause complement, because similar examples of deontic modals can be found in clauses using copulas to locate such obligations in non-present tense settings. In (8.143) below, for example, the grammaticalized copula *tʃhà* (from the verb root *tʃhà* ‘do, make’) must be used with the deontic modality inflection to express that an obligation had to be fulfilled in the past (see §9.1.8 for additional examples and description). This sentence could alternatively encode a present tense setting for the obligation expressed by *tshə-ì-pà?* (take-IRR-NR), simply by omitting the clause-final past tense copula. A nominalization encoding deontic modality, then, is a type of verbless clause complement in clauses such as (8.141) and (8.142), or a copula clause complement in clauses with the structure of (8.143).

- (8.143) *mətʃatshəŋ nə tsəŋpà? áwk sə, sati ni ku tsəjpa? tʃhà.*
 mətʃatshəŋ nə tsəŋ-pà? a-úk sə sati ni ku
 PN AGT stab-NR NRL-pig ANAPH wedding day LOC
 tshə-ì-pà? tʃhà
 take-IRR-NR COP.PST
 ‘And then, the pig that Mechtseng speared had to be taken on the day of the wedding.’

9. This sentence makes reference to the settling of litigious disputes. According to customary tribal law, the losing litigant has to pay for the cost of the tribal council eating the pig, which in effect becomes the fine.

I mentioned earlier in §8.3.3 that verbs marked by the irrealis suffix can be negated by the negative prefix. In addition to encoding events posterior to the time of speech, the irrealis suffix is used to encode non-actualized events in the past when its verb stem is negated. In the context of the narrative from which it is taken,¹⁰ the irrealis marker on the verb stem in (8.144) signals an activity (viz. drinking at the pond) that didn't happen posterior to a temporal deictic centre located in the ancient past. Since the temporal setting for this non-actualized activity is situated in the past, it cannot rightly be analyzed here or elsewhere as a marker of future tense.

Negated irrealis:

- (8.144) *pa tsəpà? ku atsə mətʃəməjù?*
 pa tsəpà? ku a-tsə mətʃəm-iʔ-ì-ùʔ
 3SG pond LOC NRL-water NEG-drink-CAUS-IRR-DEC
 'They would not let her drink at the pond.'

The immediate future suffix *-lù* also encodes temporal reference to non-actualized events, but occurs in texts with much less frequency, due to some restrictions limiting its use. Firstly, this tense marker is only ever used when the predicted event immediately follows the speech act. It is often encountered in the introductions of narratives when a speaker states that he or she is going to tell a story, and then immediately begins to relate their narrative. It is also commonly heard when people are about to take their leave. Unlike the irrealis marker, the immediate future suffix cannot be used for encoding non-actualized events in the past.

- (8.145) *ɬənəmukù, sàlù ná.*
 ɬənəm-ukù sa-lù ná
 prepare-ANT say-IMM PTCL
 '[I'm] ready; [I] will say it, OK?'

Secondly, in the very few examples of verbs marked by *-lù* in the corpus, the speech act participant is always a first person referent, with one exception. Compare the following examples, which occurred consecutively in a narrative text.

10. See Text 1.

- (8.146) a. *pa nə “aj aw, nì nə kà? tàk.rùk.rù ní,” t̃ə səmtsə.*
 pa nə aj a-u nì nə kà?
 3SG AGT INTJ VOC-grandfather 1SG AGT also
 tàk-rùk-rù ní t̃ə səmtsə
 weave-ASSIST-IMM QPTCL thus ask.PST
 ‘Hey, Grandfather! Will I also weave along with you?’, [Fox] asked.’
- b. *akhu nə “wa səmtf̃àɪ, nəŋ táŋ nə məluŋəɪla tàk.rùk.rù kəwa.”*
 a-khu nə wa səmtf̃à-əɪ nəŋ táŋ nə
 NRL-tiger AGT EXCLM grandchild-ANOM 2SG just AGT
 məluŋ-əɪ-la tàk-rùk-rù kəwa
 be.willing-ANOM-COND weave-ASSIST-IMM PTCL
 ‘Tiger replied “Ah, grandchild, just weave with me if you are willing, alright?”’

It may be the case that the use of the immediate future suffix requires epistemic authority on the part of the speaker, and that is why it is found almost exclusively with first person speech act participants. The reply of (8.146b) therefore seems to provide a counter-example to this hypothesis. However, in this narrative Tiger has just been asked by Fox if she may also help with the weaving, and so it might be the case that Tiger is merely echoing what was immediately stated, or that he now possesses the epistemic authority that is obligatory for the use of the immediate future suffix.

Out of context, the sentence of (8.146b) was rejected as ungrammatical, as were other hypothetical sentences based on these and other examples in which a first person speech act participant was substituted for a second or third person referent. Interestingly, Miller (1999: 70–71) reports a similar restriction on the use of a future tense marker in the Desano language of Colombia (Tucanoan family). This particular future marker (one of three) encodes the greatest degree of certainty of a predicted event eventuating; therefore its restriction to the first person suggests that epistemic authority is obligatorily required for its use.

A native speaker rejected negated forms with the immediate future suffix that were first put into an appropriate contextual setting. He consistently corrected my modified hypothetical examples that were based on sentences taken from texts, such as that containing the ungrammatical **mə̀sà.rù* in example (8.147) below, by replacing the immediate future marker of negated verb stems with the irrealis marker, i.e. *mə̀-sa-i-ù?* (NEG-say-IRR-DEC) ‘will not say’.¹¹

11. The tone sandhi changes that occur on the verb root here and in other examples have no bearing on the discussion at hand.

- (8.147) *nì nə hjuʦə à { *məsà.ù / mə-sajù? }.*
 nì nə hjuʦə à { *mə-sà-ù / mə-sa-ì-ù? }
 1SG AGT story one NEG-say-IMM NEG-say-IRR-DEC
 ‘I won’t tell a story.’

This is not surprising. Intuitively, it would seem redundant to specify that one is immediately *not* going to perform an activity, when this could just as easily be expressed by the negated irrealis, an example of which is given below.

- (8.148) *nì mətʃāj.*
 nì mətʃāj-ì
 1SG NEG-consume-IRR
 ‘I won’t eat.’ (said by a native speaker when she was offered fruit)

8.5.12.3. Anterior

The anterior category encodes a past action that has relevance to either the current situation, or a point in time that is temporally located in the past. Its precise temporal reference point can only be established within the pragmatic context of use. Like the present perfect and pluperfect categories of English, the anterior of Mongsen combines the characteristics of both temporal and aspectual categories.

The anterior suffix occurs in paradigmatic opposition with the other markers of absolute tense. Because it encodes an event that happened prior to the speech act, but which nevertheless has relevance to another point in time, it might also be analyzed as a marker of relative tense that is able to occur in independent clauses. The fact that it occupies the same slot as the absolute tense markers suggests, on formal grounds, that the anterior should be regarded as part of the absolute tense system. In other ways it is formally distinct. For example, a verb stem marked with the anterior marker cannot be negated.

- (8.149) **mətʃuŋukù.*
 ‘[I] haven’t eaten.’

A question such as *Have you eaten?* must be answered in the negative by the declarative clause of (8.150). It is not grammatically possible for the speaker to specifically express the current relevance of his or her state of not having eaten in his or her reply. If required, this must be done by periphrastic means. A dependency affecting the anterior and past tense categories is therefore apparent under negation.

- (8.150) *mətʃuŋla.*
 mə-tʃuŋ-la
 NEG-eat.meal-NEG.PST
 ‘[I] didn’t eat/haven’t eaten.’

Another restriction applying to the anterior suffix is that it cannot occur on the matrix verb stem of a polar interrogative sentence. A verb stem marked by the anterior suffix does tolerate semantically compatible aspectual suffixes, such as the perfective, e.g. *tʃhuwa-tʃuk-ukù* (emerge-PFV-ANT) ‘has already come out’, but is most commonly found affixed directly to a verb root.

- (8.151) *áhlù ku hmapa sə ajila táŋ nə tʃhəniŋu?, nì nə mətətukù.*
 a-hlù ku hmapa sə a-ji-la táŋ nə
 NRL-field LOC work ANAPH NRL-dog-F only AGT
 tʃhəniŋ-ù? nì nə mətət-ukù
 work.PST-DEC 1SG AGT know-ANT
 ‘Only Dog worked in that field – I know.’

8.5.12.4. *Summary of TAM categories*

Mongsen has an unmarked past tense expressed by a bare verb stem, a formally marked category of present tense encoding current action or state, a formally marked immediate future marker, and a formally marked irrealis marker.

The immediate future marker is restricted to encoding imminent events uttered by a first person speaker, or a speech act participant who possesses the epistemic authority needed to predict the imminent acts of others. Verbs marked by the immediate future marker can be questioned but not negated; conversely, the irrealis marker has no restrictions on temporal proximity or person. Evidence presented in §8.5.12.2 suggests that we should treat *-ì* as a marker of an irrealis category, since it can be used to encode non-actualized events that are not necessarily located in the future. Lastly, the anterior suffix shares some similarities with the immediate future marker in that it, too, does not permit the questioning or negating of its verb stem.

8.5.13. General nominalizer *-pà?*

The general nominalizer is the last of the five categories that occurs in slot (11). It carries the heaviest functional load of all the morphemes of Mongsen. These

many functions have been described in considerable detail in §6.6 and §7.4 and need not be repeated here.

8.6. Concluding comments

This description of the verbal morphology of Mongsen reveals a surprising level of complexity in the verb stem that was previously not fully appreciated for languages of the Ao group, or for that matter any of the Tibeto-Burman languages of Nagaland. It is now apparent from this analysis and other detailed descriptions of Tibeto-Burman languages of north-east India (e.g. Burling 2004, Chelliah 1997) that verb stems in these languages tend to be complex and highly agglutinative, usually incorporating an extensive array of suffixes grammaticalized from lexical verb roots, plus morphological causative affixes, modality markers, aspect and tense markers, types of imperative and other mood markers, and often a range of non-final suffixes that have mostly grammaticalized from case marking and nominalizing morphology.

In particular, the presence of lexical suffixes that have grammaticalized from compounded verb roots stands out as one of the most salient areal characteristics of Tibeto-Burman languages of the north-east. It now appears that this feature, arguably more than any other, serves to distinguish the Tibeto-Burman languages of north-east India from those of the greater Tibeto-Burman family, and irrespectively of their particular branch affiliations.

Chapter 9

Verbless, copula and existential clauses

Three types of clauses are used for the expression of identity, ascription, existence, location and possession. These are:

- (i) verbless clauses
- (ii) copula clauses
- (iii) verbal clauses (encoding existential and possessive meanings)

To facilitate discussion, I will henceforth refer to the core arguments of verbless clauses as verbless clause topics (VCT) and verbless clause complements (VCC); core arguments of copula clauses will be referred to as copula topics (CT) and copula complements (CC). These clausal constituents are identified by square brackets and subscript abbreviations in example sentences. Single arguments of verbal clauses are identified by a subscript S.

9.1. Verbless clauses

Verbless clauses are distinguished by the absence of a predicate;¹ they are used to assert the identity or class membership of a referent, or to ascribe a particular quality or property to the referent(s) of a noun phrase. Such meanings can also be expressed by clauses that have copulas as their heads, but these differ in entailing the obligatory expression of absolute tense, and possibly an aspectual or modal specification as well, as demonstrated by a comparison of examples (9.18) and (9.52) below. The verbless clause is therefore justifiably recognized as a clause type formally distinct from verbal clauses and copula clauses.

Verbless clauses consist of a verbless clause topic and a verbless clause complement. Both the verbless clause topic and the verbless clause complement are morphologically unmarked with respect to case, therefore the linear order of clausal constituents generally determines the status of core arguments. In the default situation (as represented by [9.1] below), the first noun phrase argument of the verbless clause functions as the verbless clause topic, and the second functions as the verbless clause complement. But just as the order of core arguments can be manipulated for pragmatic effect in verbal clauses, so too can the order of constituents be reversed for pragmatic reasons in verbless clauses expressing identity; e.g. compare the reversed ordering of core arguments in

1. Recall from §5.2 that I use the term ‘predicate’ in its modern sense, referring to just the head of the clause and its modifiers.

(9.4). This, however, is a fairly unusual presentation. Normally a deictic functioning as an independent noun phrase fills the verbless clause topic slot and occurs clause-initially (e.g. cf. [9.2]).

As a consequence of lacking a predicate, the verbless clause cannot be specified for tense to locate an equational or ascriptive meaning in time, or in the case of mood, to locate such a meaning in another possible world. If the structure of a Mongsen verbless clause can be likened to an equation represented schematically by the formula NP = NP, for example *he = a farmer*, then it is apparent that tense and mood are simply not relevant grammatical categories to this clause type, because it lacks the verbal means to encode them. Thus, the only way that a verbless clause can express this kind of equational meaning and additionally be specified for temporal or modal reference is through the inclusion of a copula in the clause. The copula then serves as a vehicle for tense or mood marking, thereby allowing temporally specified copulative meanings equivalent to *he was a farmer*, or *he will be a farmer* to find expression. The use of tense/mood marking in clauses encoding equational and ascriptive meanings is discussed in §9.1.8.

The following subsections describe verbless clauses according to the types of meanings they encode.

9.1.1. Identity clauses

This type of verbless clause is used for naming, or for expressing a referent's identity. It has the following default clausal structure in pragmatically neutral contexts:

$$[\text{NP}]_{\text{VCT}} = [\text{NP}]_{\text{VCC}}$$

The verbless clause topic slot can be filled by a pronoun, a noun, a possessed or proper noun, a nominalization, or a nominal demonstrative functioning as the head of its noun phrase. In verbless clauses expressing the identity or the name of a verbless clause topic's referent, the verbless clause complement typically takes the form of a noun of the open word class and its modifiers, but as we shall see, it can also be just a numeral or a derived nominal.

(9.1) *kə tənij pəntáŋmuŋla.*

[kə tə-nij]_{VCT} [pəntáŋmuŋ-la]_{VCC}
1SG.POSS RL-name PN-F

'My name is Bendangmongla.'

- (9.2) *atʃu m̀ŋsənəɪ hjutsə tətʂə àwʔ.*
 [a-tʃu]_{VCT} [m̀ŋsən-əɪ hjutsə tətʂə à-ùʔ]_{VCC}
 NRL-DIST Mongsen-ANOM story NZP-be.short one-DEC
 ‘That’s a Mongsen speaker’s short story.’

- (9.3) *nì imtʃən tʃàɪ.*
 [nì]_{VCT} [imtʃən tʃà-əɪ]_{VCC}
 1SG clan.name offspring-ANOM
 ‘I’m a descendent of the Imchen clan.’

Verbless clauses can be asserted, negated and questioned. In declarative verbless clauses, the declarative mood clitic *-ùʔ* (§4.2.14.6) optionally attaches to the final constituent of the clause. Its host may be a member of any word class apart from that of the illocutionary force particles (§4.2.14.5), whose clause-final position of occurrence is transparent to the declarative mood clitic. The only particles that may freely function as host to the declarative mood clitic are the restrictive focus particle *táj* ‘just, only’ and the negative particle *nun* (e.g. see the example of [9.4] below, and the discussion of §9.1.2).

In the following example, the demonstrative pronoun *atʃu* is not the clause-initial noun phrase argument, but conceptually must be interpreted as the verbless clause topic in its contextual setting. This reversed order of constituents results from the influence of pragmatics over the presentation of information in the form of noun phrases. The reversed constituent order is unlikely to be due to the elision of the verbless clause topic followed by its restoration in the postposed position as an afterthought, since there is no intervening pause between the verbless clause complement and the following verbless clause topic. This is otherwise the hallmark of postposed noun phrase arguments that have been initially elided and then restored to further clarify reference (see §5.2.1 for discussion of constituent order).

- (9.4) *tə hjutsə tʃu, maŋmətuŋ hjutsə tʃu, atʃu táŋúʔ wa.*
 [tə hjutsə tʃu maŋmətuŋ hjutsə tʃu]_{VCC}
 thus story DIST village.name story DIST
 [a-tʃu táŋ-ùʔ]_{VCT} wa
 NRL-DIST just-DEC PTCL
 ‘And, that’s the story, the Mangmetong village story, as I remember.’

(9.5) *pa təsaja.u?*

[pa]_{VCT} [tə-saja-əi-ùʔ]_{VCC}
 3SG N3P-teach-ANOM-DEC
 ‘He’s a teacher.’

The verbless clause complement slot may be filled by a possessive noun phrase (e.g. [9.3] above), a bound nominal root marked by a possessive pronoun, (e.g. [9.6] below), or a complex noun phrase in the form of a nominal head and its modifying relative clause (e.g. [9.7–9.8] below). Note the clause-final attachment of the declarative mood clitic to any constituent that is not an illocutionary force particle, and the pragmatically-motivated constituent order of (9.8).

(9.6) *ípáʔ ka.i tʃu kəintaŋùʔ.*

[ípáʔ ka.i tʃu]_{VCT} [kə-intaŋ-ùʔ]_{VCC}
 EMPHAT car DIST 1SG.POSS-thing-DEC
 ‘That car is mine.’

(9.7) *atʃu kúták thùŋipàʔ miʔ-ùʔ.*

[a-tʃu]_{VCT} [kúták thùŋ-i-pàʔ miʔ-ùʔ]_{VCC}
 NRL-DIST heaven reach-IRR-NR person-DEC
 ‘Those are the people who will reach heaven.’

(9.8) *nì nə sajakə lipaʔ sə piwʔ tə.*

[nì nə sa-ja-əkə li-pàʔ sə]_{VCC} [pi-ùʔ]_{VCT} tə
 1SG AGT say-CONT-SIM be-NR ANAPH PROX-DEC PTCL
 ‘This is what I kept saying would happen.’

9.1.2. Negated identity clauses

Verbless clauses encoding identity are negated by the negative particle *nuy*, which often co-occurs with the topic particle *la*, although this is not an obligatory requirement. When *nuy* and *la* co-occur, they appear to emphasize and reinforce the contrast that the negation adds to the assertion, i.e. ‘It is *not* the case that ...’, which then requires an assertion against which the contrast can be made. Since the scope of the negation is interpreted as applying globally to the whole of the asserted proposition, the clause-final negative particle and its accompanying topic particle cannot be constituents of the verbless clause complement alone.

(9.9) *pa tasaja.ɛ nuŋù?*

[pa]_{VCT} [tə-saja-əɪ]_{VCC} nuŋ-ù?
 3SG NZP-teach-ANOM NEG.PTCL-DEC
 ‘He’s not a teacher.’

(9.10) *át! pi təpanə sùpà? ìla thuku táŋ nuŋla,*

át [pi]_{VCT} [tə-pa? nə sù?-pà? ìla thuku táŋ]_{VCC}
 EXCLM PROX RL-father INST beget-NR 1PL.EXC nine just
 nuŋ la
 NEG TOP
 ‘Alas! It is not just the nine of us begotten by father ...’

(9.11) *sà.ɪ.ɛ tʃu ənsələj nuŋù?*

[sà.ɪ.ɛ tʃu]_{VCT} [hənsələj]_{VCC} nuŋ-ù?
 animal DIST leopard.cat NEG-DEC
 ‘That animal is not a leopard cat.’

The negative particle is used to respond in the negative to a verbless clause phrased as a polar question. An alternative strategy uses a negated form of the copula verb *tʃhà*. The root of this verb has evolved a number of extended functions in addition to its basic lexical function as a transitive verb expressing ‘do, make’. Its role as a vehicle for tense marking is examined in §9.1.8, and the hypothetical meaning it contributes in circumstantial converb clauses is described in §11.4.3.2.

The possible responses in the negative to a polar question are contrasted in (9.12) below.

(9.12) a. *nuŋù?*

nuŋ-ù?
 NEG.PTCL-DEC
 ‘No.’

b. *mətʃhàw?*

mətʃhà-ù?
 NEG-COP-DEC
 ‘No.’

9.1.3. Interrogative identity clauses

Interrogative verbless clauses have the same structure as declarative verbless clauses. In content questions, an interrogative pronoun functions as the verbless clause complement. Mongsén speakers use *səpá?* ‘who’ and not *tʃəpá?* ‘what’ when asking for someone’s name. This is probably a convention that stems from the fact that names are mostly given to human referents.

(9.13) *pa tənij səpá?*

[pa tə-nij]_{VCT} [səpá?]_{VCC}
 3SG RL-name who
 ‘What’s his name?’

(9.14) *atfu səpá? intaŋ.*

[a-tʃu]_{VCT} [səpá?-intaŋ]_{VCC}
 NRL-DIST who-thing
 ‘Whose is that?’

(9.15) *“niŋ kìtsəŋ kújá?” tə səmtsəw.ɪə.*

[niŋ kìtsəŋ]_{VCT} [kújá?]_{VCC} tə səmtsə-ù? təɪ
 2PL.POSS relative how.many thus ask.PST-DEC REP
 ‘“Your relatives number how many?,” [he] asked.’

An oblique argument can optionally occur as a peripheral constituent of a verbless clause. In the following example, a noun phrase case-marked by the comitative marker functions as a non-core argument.

(9.16) *pa nəthən kútə kìtsəŋəɪ.*

[pa]_{VCT} nə thən [kútə kìtsəŋ-əɪ]_{VCC}
 3SG 2SG.POSS COM how relative-ANOM
 ‘How is he related to you?’ (Lit. ‘How is he a house-insider with you?’)

A polar question is formed simply by adding a clause-final interrogative particle to a declarative verbless clause.

(9.17) *nàŋ khristjən sə.*

[nàŋ]_{VCT} [khristjən]_{VCC} sə
 2SG Christian QPTCL
 ‘Are you a Christian?’

9.1.4. Ascriptive clauses

Most attributes are expressed via a derived class of deverbal nominal. As we saw in verbless clauses encoding identity, neither the verbless clause topic nor the verbless clause complement of an ascriptive clause are overtly specified for a case relation. The verbless clause complement expresses a property of a verbless clause topic; it is formed from certain classes of stative intransitive verbs (described below) by means of the nominalizing prefix *tə-* attaching to the verb root. An example is *tə-məsəŋ* ‘white’, which is derived from the root of the

stative verb *məsəŋ* ‘be.white’. The resulting derivation is then used as a clausal complement to express an attribute of the verbless clause topic. Alternatively, stative verbs can be used without undergoing nominalization if the state they predicate is of a transitory nature, if there has been a recent entry into that state, or if the state is in some way exceptional. The stative verb stem is then able to take aspectual suffixes to signal the change of state and its absolute tense setting (also see §8.1.1.1). The clause in this case remains verbal.

It is essential to recognize a number of epistemological classes in order to account for the expression of property concepts in Mongsen, because not all semantic types of stative verbs can be nominalized to derive a nominal that may then be used in verbless clause complement function. Some semantic classes of stative verb used to denote certain attributes of a referent remain unequivocally verbal. The type of clause that is used to express an attribute of a core argument is ultimately determined by the semantic class of the stative verb involved, and in some circumstances this can also be affected by the semantic nature of its referent. That is to say, some property concepts find expression via a verbal clause, while others must be expressed via the medium of a verbless clause.

Dixon (2004: 3–5) recognizes a number of semantic types that are associated cross-linguistically with a class of adjective.² The four core semantic types are identified as (a) DIMENSION, (b) AGE, (c) VALUE and (d) COLOUR. In addition, a number of peripheral semantic types also associated with an adjective class are proposed; these are (e) PHYSICAL PROPERTY, including a subclass referring to CORPOREAL PROPERTY that will be identified separately as (e₁) in this description, (f) HUMAN PROPENSITY, and (g) SPEED.

Most stative intransitive verbs (as well as some ambitransitive activity verbs, such as *məɹək* ‘clean’) can function as verbless clause complements in verbless clauses following nominal derivation by *tə-*. The exceptions – those belonging to the semantic classes of (e₁) CORPOREAL PROPERTY, (f) HUMAN PROPENSITY and (g) SPEED – always express their ascriptive meanings via intransitive verbs in verbal clauses, with their status as verbs verified by their ability to take a range of tense, aspect and modality markers formally associated with the lexical category of verb (see §8.5 for description). These contrast sharply with the semantic types of property concepts expressed by derived nominals, the stems of which are unable to take any form of tense, aspect or mood marking once they have been derived by the nominalizing prefix *tə-*.

The following examples demonstrate the use of nominalized intransitive verbs functioning as verbless clause complements in verbless clauses. These express meanings consistent with Dixon’s semantic classes (a–e).

2. While I disagree with the bold claim by Dixon (2004: 1, 12) that all languages have an underived class of adjective, the recognition of these semantic types is useful to the analysis of ascriptive clauses in Mongsen.

(a) DIMENSION

(9.18) *ípá? aki tʃu tɔ̀pətíaw?*

[ípá? a-ki tʃu]_{VCT} [tɔ̀pətí-a-ù?]_{VCC}
 EMPHAT NRL-house DIST NZP-be.big-VSF-DEC
 ‘That house is big.’

Nominalized verb roots in verbless clause complement function are sometimes accompanied by an optionally realized *-a* added stem-finally, or in the slot immediately before the declarative mood marker, if it is present. The function of this morpheme remains something of a mystery. The meaning of the verbless clause complement does not change if *-a* is omitted, hence its gloss as a vacuous stem formative (VSF) *pro tempore*.

(b) AGE

(9.19) *ípá? aki tʃu íjá? tɔ̀tʃónù?*

[ípá? a-ki tʃu]_{VCT} [íjá? tɔ̀tʃón-ù?]_{VCC}
 EMPHAT NRL-house DIST much NZP-be.old-DEC
 ‘That house is very old.’

(c) VALUE

(9.20) *tə thaŋtsəla tajù?*

tə [thaŋtsə-la]_{VCT} [tə-aj-ù?]_{VCC}
 thus tree.shrew.sp.-F NZP-be.wrong-DEC
 ‘And so Squirrel is wrong.’

One exception to a classificatory system based on semantic type is found with the stative verb root *sitak* ‘be correct, be right’. Although it belongs conceptually to the VALUE semantic class, expressing the antonym of nominalized *tə-aj* ‘NZP-be.wrong’ seen in (9.20) above, *sitak* cannot similarly take the nominal prefix *tə-* to derive a deverbal nominal when it is used as an attribute. This suggests that the direct speech of (9.21) below is a verbal clause.

(9.21) *tətʃhaɪ, “u atʃu sitakù?”*

tə-tʃhà-əɪ u a-tʃu sitak-ù?
 thus-do-SEQ EXCLM NRL-DIST be.right-DEC
 ‘And then, [they said] “Oh, that’s right.”’

However, note that its verb stem lacks the present tense suffix that would locate the temporal reference of this utterance in the present, a temporal setting that would also be consistent with this example’s actual context of use and its status as a verb, if it is one (compare the tense marking on the stative verb stems of (9.30–9.31) below). Furthermore, there are restrictions on its use that are

contingent upon the semantic nature of the referent being ascribed the property. For instance, *sitak* can be used to describe an attribute of an inanimate referent, such as the price of a commodity that is judged to be correct.

- (9.22) *tʃhəntʃəŋ tʃu sitakù?*
 tʃhəntʃəŋ tʃu sitak-ù?
 price DIST be.correct-DEC
 ‘The price is correct.’

Even so, with an inanimate argument such as *tʃhəntʃəŋ* ‘price’, this putative verb cannot be negated as a member of the verb class normally is; i.e. by means of the discontinuous negative morpheme *mə-...-la*. Instead, the negative particle *nuŋ* must be used clause-finally, just as it would be for negating a verbless clause whose verbless clause complement is a nominalized deverbal adjective (see §9.1.5 below). This makes it difficult to determine if the sentence of (9.23) below is verbal, with *sitak* as its stative predicate, or alternatively non-verbal, with *sitak* functioning as the verbless clause complement. An attempt to negate *sitak* with the usual verbal morphology results in an ungrammatical sentence.

- (9.23) *tʃhəntʃəŋ tʃu sitak nuŋù?*
 tʃhəntʃəŋ tʃu sitak nuŋ-ù?
 price DIST be.correct NEG.PTCL-DEC
 ‘The price isn’t correct.’

- (9.24) **tʃhəntʃəŋ tʃu məsitakla*.

Yet if the referent happens to be human and the semantic type of the attribute is interpreted as falling under the epistemological class of HUMAN PROPENSITY, then it is *only* possible to negate *sitak* by means of the usual verbal morphology. The inferred meaning in this case is somewhat different from the VALUE interpretation of the examples given above, and this must license the verbal – as opposed to non-verbal – mode of negation. It also highlights just how sensitive grammatical structure is to both the epistemological class of the attribute and the semantic nature of the argument being ascribed a particular attribute.

- (9.25) *pa məsitakla*.
 [pa]_s mə-sitak-la
 3SG NEG-be.correct-NEG.PST
 ‘He isn’t okay.’ (i.e. ‘He’s dishonest’)

Intransitive verb roots expressing the semantic types of colour and physical property are generally marked as derived nominals when used to express an

fully fledged verbs, they can take tense inflections on their stems in the slot before the declarative clitic.

(e₁) CORPOREAL PROPERTY

(9.30) *amənlə tʃhau.ù?*

[amən-la]_S tʃhau-əɪ-ù?
PN-F be.sick-PRES-DEC
'Amenla is sick.'

(9.31) *nì íjǎ? aɪ.ù?*

[nì]_S íjǎ? aɪ-əɪ-ù?
1SG much be.good-PRES-DEC
'I'm very well.'

An attempt to nominalize the root of an intransitive verb belonging to the semantic types of (e₁) CORPOREAL PROPERTY, (f) HUMAN PROPENSITY or (g) SPEED with the nominalizing prefix *tə-* results in either an ungrammatical construction, or the interpretation of the nominalized verb root as a deverbal adjective in attributive function to a noun head, instead of a verbless clause complement.

(9.32) *ami? taɪ*

[a-mi? tə-aɪ]_{NP}
NRL-person NZP-be.good
'good person'
*[That] person is good.

(f) HUMAN PROPENSITY

(9.33) *təɪ akhula sə tʃâsi.*

tə-əɪ [a-khu-la sə]_S tʃâsi-əɪ
thus-SEQ NRL-tiger-F ANAPH be.distressed-PRES
'And, Tiger is irritated.'

(9.34) *nàŋ tənɪk məɪəm.əɪ.*

[nàŋ]_S tə-nɪk məɪəm-əɪ
2SG RL-eye be.red-PRES
'You're jealous.'

(9.35) *amənlə íjǎ? atsəkhəm.əɪ.ù?*

[amən-la]_S íjǎ? atsəkhəm-əɪ-ù?
PN-F much be.boastful-PRES-DEC
'Amenla's very boastful.'

(g) SPEED

(9.36) *í pá? aji tənəni íjá? pəjə.ù?*

[í pá? a-ji tən-nən i]_S pəjə?-ə.ɪ-ù?
 EMPHAT NRL-dog NZP-brown PROX be.fast-PRES-DEC
 ‘This brown dog is fast.’

(9.37) *í pá? kà.ɪ tʃu ahni.ù?*

[í pá? kà.ɪ tʃu]_S ahni-ə.ɪ-ù?
 EMPHAT car DIST be.slow-PRES-DEC
 ‘That car is slow.’

9.1.5. Negated ascriptive clauses

A verbless clause expressing an attribute of a verbless clause topic in the form of a derived nominal is negated by the negative particle *nuy*. This is placed at the end of the clause. The negative particle is understood to negate the whole of the proposition expressed by the verbless clause, i.e. ‘it is not the case that CLAUSE’.

(9.38) *aki tʃu təpətia la nuyù?*

[a-ki tʃu]_{VCT} [tə-pəti-a la]_{VCC} nuy-ù?
 NRL-house DIST NZP-be.big-VSF TOP NEG.PTCL
 ‘The house is not big.’

This description of the clausal structure of negated attributes applies to all the semantic types identified above except for (e₁) CORPOREAL PROPERTY, (f) HUMAN PROPENSITY and (g) SPEED. Recall that when these three semantic types are used to describe an attribute of a core argument (which is in S function in this context), they are always expressed via stative verbs in verbal clauses. This requirement does not change under negation.

It might be argued that the following example should rightfully be analyzed as belonging to the ‘value’ semantic type; however, in the context of the narrative it was taken from, the personified referent is being described somewhat less with respect to VALUE and more with regard to HUMAN PROPENSITY, thus justifying its inclusion here (compare the use of *sitak* ‘correct’ in [9.25] under a similar set of semantic conditions).

(9.39) *pùkphula, nàŋ mà.rulaw?*

pùkphu-la [nàŋ]_S mà-a.ɪ-ù?
 owl-F 2SG NEG-be.good-NEG.PST-DEC
 ‘Owl! You’re no good!’

The example of (9.34) provided above cannot be negated to derive an equivalent meaning of opposite polarity. Negation affects the predicate only and not the whole of the semi-lexicalized expression, resulting in a meaning unrelated to its positive polarity counterpart. In other words, the scope of negation does not apply globally to the whole of the verbal clause in the same way that negation via the negative particle globally affects verbless clauses.

- (9.40) *nàŋ tənɪk məmɔɪəmɔɪ.*
 [nàŋ]_s tə-nɪk mə-mɔɪəm-əɪ
 2SG RL-eye NEG-be.red-PRES
 ‘Your eye(s) is/are not red.’
 *You’re not jealous.

9.1.6. Interrogative ascriptive clauses

Interrogative ascriptive clauses in which attributes express the core semantic types (a) DIMENSION, (b) AGE, (c) VALUE, (d) COLOUR and the peripheral type of (e) PHYSICAL PROPERTY have the same structure as their respective declarative ascriptive clauses (cf. §9.1.4). Members of these semantic classes remain nominal when they are used to express property concepts in interrogative clauses, thus retain the nominalizing prefix *tə-* on their stative verb roots. The presence of an interrogative particle at the end of the clause identifies the clause type as interrogative.

- (9.41) *aki tʃu tɔpətia pà.*
 [a-ki tʃu]_{VCT} [tə-pətɪ-a]_{VCC} pà
 NRL-house DIST NZP-be.big-VSF QPTCL
 ‘Is the house big?’

Attributes expressing the semantic types of (e₁) CORPOREAL PROPERTY, (f) HUMAN PROPENSITY and (g) SPEED in interrogative clauses remain verbal, allowing their predicates to take tense, aspect and mood markers associated with the lexical class of verbs.

- (9.42) *amənla tʃhɑɪ.ɪ.ə pà.*
 [amən-la]_s tʃhɑɪ.ɪ-əɪ pà
 PN-F be.sick-PRES QPTCL
 ‘Is Amenla sick?’

(9.43) *nàŋ phələməi sə?*

[nàŋ]_s phələm-əi sə
 2SG be.tired-PRES QPTCL
 ‘Are you tired?’

(9.44) *aiənla pəlai pà.*

[aiən-la]_s pəla-əi pà
 PN-F be.happy-PRES QPTCL
 ‘Is Arenla happy?’

9.1.7. Cardinal numerals as attributes

Cardinal numerals can fill the verbless clause complement slot in verbless clauses without undergoing any morphological modification. While these also express an ascriptive property of a verbless clause topic, they are formally very different from verbless clause complements derived from the roots of stative verbs.

In §6.3 I mentioned that numerals can be constituents of headless noun phrases when they are understood to be quantifying the referent of the omitted head, e.g. *pa phaŋa hùŋ* (3SG five see.PST) ‘he saw five [deer]’. The pragmatically driven deletion of the head is demonstrated by the fact that it could be restored to the noun phrase, e.g. *mətshə phaŋa* ‘five deer’. A numeral is performing a similar quantifying function when it occurs as the sole constituent of a headless noun phrase in the verbless clause complement slot of an ascriptive clause. We should interpret the nominalized verb stem of (9.46) to be functioning as a non-relativized noun modifier of *kəm* ‘year’, as the speaker who uttered this sentence was indeed aged seventy-five (and thus could not have been born in ’75).

(9.45) *nì akəm nətəi nət.*

[nì a-kəm]_{VCT} [nətəi nət]_{VCC}
 1SG NRL-year seventy two
 ‘My age is seventy-two.’

(9.46) *təsùpà? kəm tʃu nətəi phàŋà.*

[tə-sù-pà? kəm tʃu]_{VCT} [nətəi phàŋà]_{VCC}
 NZP-bear-NR year DIST seventy five
 ‘[My] birth years are seventy-five.’

Mongsen does not require the equivalent of a dummy subject in its verbless clauses, as demonstrated by (9.47) below. Here a verbless clause complement

occurs in the absence of a verbless clause topic. The preceding converb clause creates the pragmatic conditions under which the verbless clause topic can be elided in this verbless clause.

(9.47) *pa sə.ɔ akəm asəmù?*

pa sə-əɪ [a-kəm a-səm-ù?]_{VCC}
 3SG die-SEQ NRL-year NRL-three-DEC

‘It has been three years since he died.’ (Literally ‘He having died, it is three years.’)

9.1.8. Expression of tense, aspect and modality in copula clauses

Dependencies between different clause types used to express copulative meanings become apparent when native speakers temporally locate the events of verbless clauses outside of their default tenseless settings.

Finite temporal and modal distinctions are made by converting the verbless clause to a verbal clause using a grammaticalized form of a transitive verb originally meaning ‘do, make’ in copular function (described in §9.2 below). This must be specified for tense or mood: the bare stem of *tʃhà* for past tense; *tʃhà-ɹù* for immediate future; and *tʃhà-ì* for irrealis (see §8.5.12 for a description of the tense and mood marking system). In copulative function, *tʃhà* thus serves as a medium for the grammatical expression of tense or mood. This grammaticalized copula does not occur with the present tense marker, presumably because present tense marking is rendered redundant by the verbless clause.

The core marking of arguments in these tense/mood-specified clauses is identical to that of verbless clauses; that is, neither core argument receives any particular morphosyntactic marking. As with verbless clauses, the topic particle *la* can be used on a noun phrase to signal a contrastive topic. In the following text example, the speaker has previously been talking about another referent and now shifts his attention to the pig. This shift in attention is made explicit by the presence of the topic particle marking the copula topic.

(9.48) *áwklai la jimhjutaŋla tʃhàjù?*

[a-úk-la i la]_{CT} [jimhjutaŋla]_{CC} tʃhà-ì-ù?
 NRL-pig-F PROX TOP fine COP-IRR-DEC

‘As for this pig, [it] will be the fine’.⁴

4. The word glossed ‘fine’ is a lexicalized compound *jim-hju-thaŋ-la* ‘village-word-sever-?’. I am not sure of the meaning of the final morpheme. The whole compound refers to a village council’s fine in the form of a pig.

If the pragmatic context or an antecedent mention allows, the noun phrase functioning as the copula topic can be elided.

- (9.49) *jaŋ táŋ tʃhàj sà.*
 [jaŋ táŋ]_{CC} tʃhà-ì sà
 sufficient just COP-IRR QPTCL
 ‘Will [that] be enough?’

The grammaticalized copula *tʃhà* is used with the immediate future tense suffix *-ù* to produce a copulative meaning temporally situated in the immediate future.

- (9.50) *nì tənisiɪ tʃhà.ù.*
 [nì]_{CT} [tə-ni-si-əɪ]_{CC} tʃhà-ù
 1SG NZP-lead-RPET-ANOM COP-IMM
 ‘I’ll be the leader.’

The following example demonstrates the unmarked past tense. This sentence was used at the beginning of a traditional folklore text to introduce the identity of a character that lived long ago.

- (9.51) *hmatsəŋlikla tʃu lazati à tʃhà.*
 [hmatsəŋlik-la tʃu]_{CT} [lazati à]_{CC} tʃhà
 PN-F DIST unmarried.woman one COP.PST
 ‘Matsenglikla was a maiden.’

Ascriptive clauses whose copulative meanings have past tense reference are also encoded by *tʃhà* acting as a vehicle for tense marking. This applies to all the semantic types outlined in §9.1.4, with the exception of the classes of (e)₁ CORPOREAL PROPERTY, (f) HUMAN PROPENSITY and (g) SPEED. Since they are verbal, members of these classes already have the resources for marking tense and mood distinctions directly on their verb stems. Members of all the other semantic classes require a copula to encode their tense or mood distinctions.

- (9.52) *aki sə təpətìà tʃhàw?*
 [a-ki sə]_{CT} [tə-pətì-à]_{CC} tʃhà-ù?
 NRL-house ANAPH NZP-be.big-VSF COP.PST-DEC
 ‘That house was big.’

An existential verb *li* ‘be’ functions as a vehicle for the expression of some aspectual categories in clauses encoding copulative meanings. Its core meaning is ‘stay’ or ‘be at’. It is used in ascriptive clauses with the present tense suffix

-(ə)ɿ to express habitual aspect, and with the past habitual suffix *-jaʔ* to express habitual aspect in the past. The copula *li* in these examples does not assert the existence or location of an entity, but just provides the means for the expression of habitual and past habitual aspect.

(9.53) *atsə tʃu təməkun liɿuʔ.*

[a-tsə tʃu]_{CT} [tə-məkun]_{CC} li-əɿ-ùʔ
 NRL-water DIST NZP-be.cold be-PRES-DEC
 ‘The water is usually cold.’

(9.54) *atsə tʃu təməkun lijaʔ.*

[a-tsə tʃu]_{CT} [tə-məkun]_{CC} li-jaʔ
 NRL-water DIST NZP-be.cold be-HAB.PST
 ‘The water used to be cold.’

The copula *tʃhà* can co-occur with the existential verb *li* in an auxiliary-like function marked by a simultaneous converb suffix *-(ə)kə* (see §11.4.1.2). This gives the functional equivalent of an unbounded, imperfective aspectual reading to a clause whose copulative meaning is located in the past. Sometimes the boundary between clauses expressing identity and existence in the past can be indistinct, depending upon how clausal constituents are structurally interpreted. This ambiguity has possibly provided the diachronic pathway for the reanalysis of an erstwhile existential verb as a tense-carrying copula of identity clauses located in the past.

(9.55) *mətʃatshəŋ tʃu tʃàmɿ kìn à likə tʃhà.*

[mətʃatshəŋ tʃu]_{CT} [tʃàmɿ kìn à]_{CC} li-əkə tʃhà
 PN DIST clan.name clan one be-SIM COP.PST
 ‘Mechatseng was one of the Jamir clan.’

Some speakers pronounce the existential verb *li* and the copula verb *tʃhà* together as one phonological word *litʃhà*, and without the intervening simultaneous converb suffix demonstrated in the example immediately above. The fused form can be used to express a simple past existential meaning, as in the following existential clause. This example is also useful for demonstrating both the existential and grammaticalized copulative function of the verb root *li* in the same sentence.

(9.56) *pa tsəjpaʔ walikàʔ, məlɿla litʃhàwʔ.*

[pa]_S tsə-ì-pàʔ wa-likàʔ mə-li-la li-tʃhà-ùʔ
 3SG meet-IRR-NR go-CONTEMP NEG-be-NEG.CV be-COP.PST-DEC
 ‘When [I] had gone to meet him, [he] wasn’t there.’

The sole noun phrase argument of this complex clause is interpreted as being in core S function with respect to the clause-final matrix verb *litʃhà*. The use of this combined form in the expression of past existence is further investigated in §9.2 below.

In clauses in which an attribute is ascribed to a referent, the semantic classes expressing CORPOREAL PROPERTY, HUMAN PROPENSITY and SPEED remain differentiated as regular intransitive verbs by virtue of the fact that they do not undergo nominalization by *tə-*. Consequently they are able to affix a range of tense, aspect and mood markers directly on their verb stems.

- (9.57) *nì hjimà.ù?*
 [nì]_S hjim-əɪ-ù?
 1SG be.hungry-PRES-DEC
 ‘I’m hungry.’

- (9.58) *nì atsəktʃukukù.*
 [nì]_S atsək-tʃuk-ukù
 1SG be.cold-PFV-ANT
 ‘I’m cold.’

Cross-linguistically, it would appear that the use of copulas to convey non-present tense distinctions in languages with verbless clauses is rather common. Kachru (1968: 44–45), for example, observes that a number of languages of the Indian subcontinent make use of verbless clauses for present tense, but use copulas for other tenses such as past and future.

9.2. Verbal existential clauses

In the preceding section we looked at the tense/mood-carrying functions of two grammaticalized copulas. We will now consider the functions of these verbs in their roles as predicates of intransitive clauses, after which we will examine their grammatical uses beyond those documented in §9.1.8.

Existence and location are expressed in Mongsen by two existential verbs: *li* and *tʃhà*. Since *li* can occur with just one core argument and freely accepts all the tense, aspect and mood markers that are typically associated with verbs, it is analyzed as a sub-class of intransitive verb. The status of *tʃhà* as an existential verb is less clear, because it demonstrates greater limitations on the expression of existence in time. This is likely to be a reflection of its intermediate stage of grammaticalization. In §9.1.8 it was shown that the primary role of *tʃhà* as a copula is to serve as a vehicle for absolute tense/mood marking, and that as a bare stem it encodes past tense. This past tense marking function is also

apparent in existential clauses. There are some complications associated with the analysis of *tʃhà* as an existential verb, particularly when it co-occurs with *li* in the verb form *litʃhà*. In these circumstances its independent lexical status is called into question.

Recall that the intransitive existential *tʃhà* is homophonous with a transitive verb meaning ‘do, make’. Presumably the existential, copulative and other meanings expressed by this polyfunctional form have grammaticalized from an erstwhile transitive verb. For the purposes of this description it is convenient to analyze the transitive and intransitive forms of *tʃhà* as having separate lexical entries on both formal and semantic grounds, given their different valencies, meanings and uses.

9.2.1. Expression of existence in declarative clauses

Declarative verbal existential clauses minimally consist of one noun phrase argument in S function and an intransitive predicate. The S argument of the existential clause is formally unspecified for a case relation. Existence in a specific location can be precisely encoded by an oblique noun phrase marked by locative case, although this is not a necessary requirement for the grammaticality of the existential verbal clause. Existential clauses thus have the following structure:

$$[\text{NP}]_S + ([\text{NP}]_{\text{LOC}}) + \text{EXISTENTIAL VERB}$$

The tense paradigm for *li* is somewhat irregular from speaker to speaker. Some use an unmarked verb root *li* to express past tense, while others consistently use *litʃhà* for past tense, pronounced without a hiatus between the two syllables (i.e. as a single phonological word). Some speakers use a mix of *litʃhà* and *tʃhà* for expressing past existence, occasionally with both forms being used just a few clauses apart in narrative texts.

I am unable to establish a difference in meaning between the existential functions of *li*, *tʃhà* and *litʃhà* and surmise that the three forms of this verb are synchronically in free variation in existential clauses located in the past. This free variation may represent the initial stages of the grammaticalization of *tʃhà* as a marker of past tense in existential clauses, with an earlier meaning retained as an independent existential verb. In *litʃhà*, though, perhaps *tʃhà* has lost its lexical independence and is now marking a past temporal distinction in the guise of a quasi-suffix that is restricted to existential clauses. In these occurrences it is difficult to decide how *tʃhà* should be glossed. As a compromise, it will be analyzed as a past tense marker when occurring together

with *li*, and when occurring independently it will be treated as an existential verb encoding past tense.

The polyfunctionality of *tʃhà* – from a transitive verb meaning ‘do, make’, to an intransitive verb expressing existence, to a marker of past tense in existential clauses, and to a marker of other grammatical functions described below – correlates with Hopper’s (1991) notion of ‘divergence’, in which a lexical form develops new meanings without necessarily discarding its older meanings. Divergence has resulted in the synchronic co-existence of a range of different functions and meanings of *tʃhà*. All of these functional extensions have presumably grammaticalized from a single diachronic source.

The following text examples demonstrate past existence encoded by the three forms of the above-mentioned existentials.

(9.59) *ajila tʃaŋsə táŋ li.*

[aji-la tʃaŋsə táŋ]_s li
dog-F footprint just be.PST
‘Only Dog’s footprints were there.’

(9.60) *ípá? aləmi hmapaŋ ku, təməjəmtəpəɪ tʃhaŋatì ami? ta.ɪu majantʃhaŋəɪ à litʃhà.*

ípá? aləmi hmapaŋ ku [tə-mijim-təp-əɪ tʃhaŋatì
EMPHAT olden time LOC NZP-love-RECIP-ANOM bachelor
a-mi? tə-a.ɪu majantʃhaŋəɪ à]_s li-tʃhà.
NRL-person NZP-good PN one be-PST
‘In olden times there was a lover, a good young man [named] Mayangchanger.’

(9.61) *tʃhaŋatì à tʃhà.*

[tʃhaŋatì à]_s tʃhà
bachelor one be.PST
‘There was a young man.’

In the present tense, *tʃhà* is never used as an existential, either alone or in combination with *li*, whereas *li* can be used alone as an existential verb in both past and present tense settings (e.g. cf. [9.59] with [5.77], repeated and renumbered below as [9.62] below for convenience). In assertions of existence located in the present tense and marked by the declarative mood clitic, *li* is usually accompanied by the vacuous stem formative *-a* (VSF) that often occurs on the stems of derived nominals functioning as verbless clause complements, as discussed in §9.1.4.

- (9.62) *mi:li tʃãŋpùk ku liaw?*
 [mi:li]_s tʃãŋ-pùk ku li-a-ù?
 ant.sp. foot-STOMACH LOC be-VSF-DEC
 ‘There is a large ant under [your] foot.’

The existential *li* also demonstrates different shades of meaning in certain contexts that may represent the emergence of semantic elaboration stemming from its basic meaning. One sense of *li* is ‘to stay’. In the following example it is arguably not so much existence *per se* that is being asserted by *li*, but rather a degree of permanency to the change of state described by the dependent converb clause. This continuity correlates with the habitual aspect meaning expressed by the copula *li* plus the present tense marker *-à*, as demonstrated above in the example of (9.53).

- (9.63) *təi thàku i pa wàzà? təpəti à nə tasu? à táŋ kəməi li.*
 tə-əi tha-ku i [pa]_s wàzà? tə-pəti à nə
 thus-SEQ now-LOC PROX 3SG bird NZP-be.big one AGT
 tə-asu? à táŋ kəm-əi li-à
 NZP-be.small one just become-SEQ be-PRES
 And right now she, a big bird, has become just a small one.⁵

In contrast to ascriptive and identity clauses that must use *tʃhà* as a copula to carry the immediate tense and irrealis mood suffixes (as previously discussed in §9.1.8), only *li* can be used to express existence in the future. No examples of the existential verb *li* marked by the immediate future suffix have been encountered in the corpus.

- (9.64) *inət kələm táŋ nə liju.è.*
 [inət kələm táŋ nə]_s li-ì-ù? .è
 1DU.EX together just AGT be-IRR-DEC PTCL
 ‘We will just live together.’

- (9.65) *nəŋət məliju?*
 [nəŋət]_s mə-li-ì-ù?
 2DU NEG-be-IRR-DEC
 ‘You two will not stay together.’

5. This example from a folklore narrative explains how the Yellow-backed Sunbird, allegedly formerly a large bird, was reduced to a diminutive size as a result of other birds stroking and patting it (see Text 4 for the full context).

(9.66) “*álima tʃu ku, mijimtʃàɪ tʃu, mànli saɲali mànli saɲali tɔ̀ ləju?*,” *tɔ̀ sa.*

a-líma tʃu ku mijimtʃàɪ tʃu
 NRL-world DIST LOC humankind DIST
 màn-li saɲa-li màn-li saɲa-li
 dark-ALT.IT light-ALT.IT dark-ALT.IT light-ALT.IT
 tɔ̀ li-ì-u? tɔ̀ sa
 thus be-IRR-DEC thus say.PST

‘In the world, humankind will have dark, light, dark, light infinitely alternating – it will be thus’ said [Yellow-backed Sunbird].’

The existential *li* behaves just like a typical intransitive verb with respect to the way that it accepts tense, aspect and mood marking on its stem. The following examples demonstrate *li* marked by a range of verbal morphology, in addition to the tense/mood marking possibilities seen in the previous examples. As (9.67) shows, *li* can be marked by the continuative aspectual suffix. This sentence uses a dependent converb clause to encode an anterior event, while the existential encodes a continuing state of existence resulting from that event, literally that the blood has flowed (from a wound) and it is still caked on the body of the pig, according to the interpretations of consultants. There are considerable pauses after both the initial vocative and *áwk* ‘pig’ in the recorded narrative this example is taken from; *áwk* is therefore analyzed as a non-core sentential topic, and *a-jí?* ‘NRL-blood’ is interpreted as being in core S function.

(9.67) “*apa?! áwk – ájí? jiməɪ lija.u?*,” *tɔ̀ sa.*

a-pa? a-úk [a-jí?]_s jim-əɪ li-ja-əɪ-ù?
 VOC-father NRL-pig NRL-blood flow-SEQ be-CONT-PRES-DEC
 tɔ̀ sa
 thus say.PST

‘“Father! The pig – it’s been bleeding”, [she] said.’

The existential *li* can be nominalized to form a relative clause standing in a modifier relationship to its head, which may be explicitly mentioned, or elided.

(9.68) *thijun ku lipa? tʃu khiəŋ.*

thijun ku li-pà? tʃu khi-aŋ
 centre LOC be-NR DIST give-IMP
 ‘Give me [the one] that is in the centre.’

It was previously noted that *li* functions more or less like an auxiliary verb in converb clauses. It is often found in dependent clauses marked by the simultaneous converb suffix *-(ə)kə*, which is in turn preceded by another verb

marked by the same converb suffix. The first converb usually describes a state or an activity, and the second converb formed with *li* expresses the continuity of that state or activity. The two converbs are used together to create a temporal backdrop for matrix verb events (see §11.4.1.2 for further examples).

- (9.69) *təɪ tsəməŋkə likə təmhnak sə ɿà.*
 tə-əɪ tsəməŋ-əkə li-əkə tə-əmhnak sə ɿà
 thus-SEQ wonder-SIM be-SIM RL-son-in-law ANAPH come.PST
 ‘And while [he] was wondering about this, the son-in-law came.’

The existential *li* additionally performs an important main verb function in imperatives involving certain semantic classes of stative verbs: it is required for expressing commands for activities to be performed continuously. These extended functions have evolved out of the continuity inherent in the basic meaning of existence and are discussed in §10.3.1.

As a main verb, *tʃhà* similarly takes a range of tense, aspect and mood marking on its stem; however, in these circumstances it usually encodes a lexical meaning of ‘happen’, rather than ‘existence’. The use of *tʃhà* as a main verb modified by verbal suffixes is described in §9.3.3 below. Arguably the most significant development of *tʃhà* in the grammar of Mongsen is its metamorphosis into a type of connective that links clauses, provides structure to discourse, and helps to advance the plot. This function is touched on in §9.3.2 and described in detail in §4.2.10.

9.2.2. Expression of existence in interrogative clauses

The tense inflections of *li* are irregular in interrogative existential clauses. This correlates with a general dependency between past and present tense noted for interrogative verbal clauses, resulting in the optional conflation of these two grammatical categories (Coupe 2002). This observation can be refined further in the present work with respect to interrogative clauses involving existentials (see §8.5.13 for description of the salient characteristics of past and present tense marking in interrogative clauses).

Existential polar questions formed with *li* are marked with the suffix *-(ə)ɪ* in contextually present tense settings. It is not possible to determine if its underlying tone is mid or low, because the schwa and its associated tone is deleted in the output. This would otherwise help us to determine its function in this type of clause (the present tense marker has a low tone, and the agentive nominalizer has a mid tone). Whether this morpheme is in fact the present tense marker or the agentive nominalization in some older frozen function is open to interpretation (see §8.5.12 for discussion concerning the possible reanalysis of

nominalizing morphemes as markers of absolute tense). I have glossed it as the present tense marker here, based on the temporal settings of the following examples.

- (9.70) *nì thən təpətípa? lí pá.*
 nì thən tə-pətí-pà? lí-əɪ pá
 1SG COM NZP-be.big-NR be-PRES QPTCL
 ‘Is there one bigger than me?’

- (9.71) *nəŋ la làpàki hli təpətípa? həmətəɪ lí pá.*
 nəŋ la làpàki hli tə-pətí-pà?
 2SG TOP bean.sp vine NZP-be.big.SUP-NR
 həmət-əɪ lí-əɪ pá
 hold-SEQ be-PRES QPTCL
 ‘Are you holding onto the biggest bean vine?’

A content question formed with the existential copula *li* that is contextually set in the present is not necessarily formally specified for tense. Recall that the past tense is the formally unmarked absolute tense category in declarative clauses (see §8.4.1).

- (9.72) *ahənzala tʃákú lí.*
 [a-hən-za-la]_s tʃópá? ku lí
 NRL-chicken-DIM-F what LOC be
 ‘Where’s Little Chick?’

Similarly, the existential verbs of disjunctive A-not-A type questions are consistently unmarked for tense, regardless of their actual temporal settings. It is not possible to use *tʃhà* in place of *li* in this type of interrogative to encode either past or present existence.

- (9.73) *pá lí sè mələlila.*
 [pá]_s lí sè mə-li-la
 3SG be QPTCL NEG-be-NEG
 ‘Is he there or not?’ / ‘Was he there or not?’

Content questions that make reference to past existence are formed with *tʃhà*. The following question of (9.74a) and its answer in (9.74b) are taken from a conversation in which a speaker asks another about a church fête she had attended a few days previously.

- (9.74) a. *siwaj niŋəɪ kútó sá tʃhàj.*
 siwaj niŋ-əɪ luŋ sə kútó sá tʃhà i
 C.Y. name-ANOM group ANAPH how PTCL COP.PST PROX
 ‘This C(hristian) Y(outh) group [fête] – how was it?’
- b. *ája íjá? aɹukə táŋ tʃhà atì.*
 ája íjá? aɹu-əkə táŋ tʃhà atì
 EXCLM much be.good-SIM just COP.PST PTCL
 ‘Hey, it was really good.’

9.3. Extended functions of existentials

The following subsections describe additional meanings encoded by existential verbs or their grammaticalized derivatives in Mongsen. This includes mood marking in dependent clauses, the diachronic development of a converb construction into a discourse connective, and the expression of ‘happening’.

9.3.1. Mood marking in circumstantial converb clauses

The existentials *li* and *tʃhà* contribute to a causal versus counterfactual hypothetical irrealis contrast in circumstantial converb clauses (see §11.4.3.2). The description of these grammaticalized functions is deferred until we investigate clause combining in Chapter 11. The following elicited example provides a preview of the hypothetical mood marking function to be discussed.

- (9.75) *aji tʃu məzəm áthúkə tʃhàkùla təkəm lipà? tʃhàw?*
 a-ji tʃu məzəm áthú?-əkə tʃhà-kùla
 NRL-dog DIST poison vomit-SIM HAPPEN.COP-CIRCM
 tə-kəm li-pà? tʃhà-ù?
 NZP-be.alive be-NR be.PST-DEC
 If the dog had vomited the poison, [it] would still be living.

9.3.2. Grammaticalized *tʃhà* as a discourse connective

One of the most commonly encountered construction types in Mongsen narratives consists of the quotative particle *tə*, the verb root *tʃhà* and the sequential converb suffix *-əɪ*. This is used for tail-head linkage in narrative discourse. The resulting collocation *tə-tʃhà-əɪ* (thus-do-SEQ) is always accompanied by an overlaid intonation contour identifying a clausal boundary

marked by the sequential converb, and it is often followed by a considerable pause (the intonation pattern associated with the sequential converb is described in §3.4). The vast majority of speakers tend to pronounce this construction as one phonological word and have difficulty segmenting it into its constituent parts. This suggests that it has become lexicalized for most speakers (see §4.2.10).

9.3.3. Happening

When *tʃhà* ‘do, make’ occurs as an intransitive verb with suffixing morphology, its stem often expresses a meaning of ‘become’ or ‘happen’. This is inferred in (9.76), and explicit in (9.77–9.78).

(9.76) *təɪ tʃhəŋ.ɹutsəla nə mətʃatʃhəŋ təpazaʔ tʃhàjpàʔ la taŋ tʃu ku, ahju tʃu sawkù.*

tə-əɪ tʃhəŋ.ɹutsə-la nə mətʃatʃhəŋ tə-pazaʔ
thus-SEQ PN-F AGT PN RL-husband

tʃhà-ì-pàʔ la taŋ tʃu ku a-hju tʃu sa-ukù
be-IRR-NR TOP SIDE DIST LOC NRL-word DIST say-ANT

‘And Tsengrutsela said those words to Mechtseng, who would become her husband.’

Example (9.77) shows that the only situation in which *tʃhà* can take irrealis marking as a main verb is when it expresses a future sense of ‘happen’. Recall that future existence must be expressed by existential *li* taking the irrealis suffix on its stem.

(9.77) *i sitak məpuŋmaj saʔ tʃhàjù.ɪə.*

i sitak mə-puŋ-maʔ-ì saʔ tʃhà-ì-ùʔ ɪə
PROX correct NEG-be.good-CMPL-IRR PTCL be-IRR-DEC PTCL

‘This is right, something not good at all will happen.’

Lastly, when occurring as a bare stem, *tʃhà* sometimes expresses an incipient meaning of ‘happen’ in its function as the main verb of a content question. In the following example taken from a narrative, a character is beginning to feel the effects of poison that was given to him by stealth.

- (9.78) *thajip i tʃópá? tʃhà pi.*
 thajip i [tʃópá?]_s tʃhà pi
 morning PROX what be PROX
 ‘What’s happening this morning?’

The meaning does not have to be incipient, however. The same speaker utters a sentence identical in structure to (9.78) in another text, this time in the context of a protagonist wondering to himself about the events he is witnessing as he tries to comprehend their significance.⁶ The temporal setting in which both statements are uttered is the present, yet the verb root occurs without the present tense marker that normally is obligatory for marking this tense on verb stems. This is an example of the frequently defective or irregular morphological behaviour of verbs used to express copulative, possessive and existential meanings in Mongsén.

9.4. Possessive verbal clause

Predicated possession is encoded by the verbs *kət* and *khà*, with the existential *li* alternatively able to encode the possession of some semantic classes of referents that occur in clauses with *khà*. Possessive verbs can be negated, specified for finite tense (with some inconsistencies), or nominalized to derive nouns or deverbal modifiers of nouns. For example, *tə-khà-əɪ* (NZIP-have-ANOM) ‘wealthy person’ is derived from the root *khà* ‘have’, which is nominalized by the nominal prefix *tə-* and the agentive nominalizer suffix *-əɪ*. In some clauses, possessive verbs have the formal appearance of transitive verbs with two core arguments, i.e. a possessor and a possessee, both of which are morpho-syntactically unspecified for case relations (e.g. [9.80] below). In other possessive clauses, they resemble intransitive existential clauses with just one core argument and an oblique locative (e.g. cf. [9.79]).

The choice of possessive verb used as a predicate is determined by the semantic nature of the possessee. The possession of body parts or entities that are attached to the body more or less permanently, such as freckles, moles and teeth, or illnesses such as cancerous tumours, tuberculosis or afflictions caused by various types of wounds, are all expressed by *kət*.

- (9.79) *pa thaku ləph.u-tsó? kət.*
 pa thak-ku ləph.u-tsó? kət
 3SG PLACE-LOC cockroach-excreta possess
 ‘He’s got moles.’

6. The example occurs in Text 2, line 23.

(9.80) *nàŋ tətʃhatʃən tamá.u kətù?*

nàŋ tətʃha-tʃən tamá.u kət-ù?
 2SG NZP-hot-LNOM bad have-DEC
 ‘You have a bad disease.’

(9.81) *nàŋ təpha kət sè məkətla la – təpaŋ tʃu kaŋ tʃəj.*

nàŋ tə-pha kət sè məkət-la la
 2SG RL-tooth have QPTCL NEG-have-NEG.PST TOP

tə-paŋ tʃu ka-aŋ tʃəj
 RL-mouth DIST open.mouth-IMP PTCL

‘‘Do you have teeth or not? Just open your mouth,’’ [said Leopard Cat to Rooster].’

However, entities that are in contact with the body rather fleetingly are also expressed with *kət*, as the following example demonstrates. Possessive verbs are defective and sometimes lack the present tense marker when used to encode what are unquestionably present tense contextual settings.

(9.82) *nàŋ asə ku məsən à kətù?*

nàŋ a-sə ku məsən à kət-ù?
 2SG NRL-shawl LOC worm one have-DEC
 ‘You’ve got a worm on your shawl.’

The possession of certain characteristics of individuals, such as personality traits and the possession of supernatural powers, are both encoded by *kət*. Note the inconsistency in tense marking in these sentences.

(9.83) *pa azən kətù? uŋ tə.*

pa azən kət-ù? uŋ tə
 3SG power have-DEC PTCL PTCL
 ‘She certainly has [magical] power.’

(9.84) *pa muli tʃaɪ azən kətijù?*

pa muli tʃaɪ-əɪ azən kət-i-ù?
 3SG drug consume-SEQ power have-IRR-DEC
 ‘After taking the drug she’ll have [magical] power.’

(9.85) *pa təmijimi kətù?*

pa tət-mijim-mi kət-ù?
 3SG NZP-love-DESID have-DEC
 ‘She’s very lovable.’ (i.e. she possesses lovable qualities)

The possession of inanimate objects that are not necessarily in physical contact with one's body and do not fall under the above-mentioned classes – such as money, cars, machetes etc. – tend to be encoded with *khà*.

- (9.86) *nì kà.i à khà-ù?*
 nì kà.i à khà-ù?
 1SG car one have-DEC
 'I have a car.'

- (9.87) *inət atʃhən kúláŋa khà-ù?*
 inət atʃhən kúláŋa khà-ù?
 1DU.IN money many have-DEC
 'We have plenty of money.'

It is often the case that *khà* is also used to express existence. Testing with native speakers reveals that *khà* and *li* can usually be exchanged with no loss or change of meaning. The Waromung Mongsen narrative example of (9.88) demonstrates the use of *khà* as an existential, while the elicited example of (9.89) suggests that *khà* and *li* can be used interchangeably for the expression of existence.

- (9.88) *tə, thani nukɿaŋ jim tsə tʃùku məkhhàlà.*
 tə-əɿ tha-ni nukɿaŋ jim tsə tʃùku
 thus-SEQ now-day name village DIST consequently
 məkhhà-la
 NEG-have-NEG
 'And, consequently today Nokrang village doesn't exist.'

- (9.89) *awŋ ku akhu {məlilaw? / məkhhàlaw?}*.
 a-uŋ ku a-khu {mè-li-la-ù? / məkhhà-la-ù?}
 NRL-jungle LOC NRL-tiger NEG-be-NEG-DEC NEG-have-NEG-DEC
 'There are no tigers in the jungle.'

9.5. Summary

The following table summarizes the functions and meanings expressed by verbless clauses, copular clauses, existential clauses and possessive clauses in Mongsen.

Table 9.1. Summary of clause types expressing copulative meanings

	V-LESS	<i>li</i> / <i>litʃhà</i>	<i>tʃhà</i>	<i>khà</i>	<i>kət</i>
Identity	✓				
Identity + TAM			✓		
Ascription	✓	✓		✓	
Ascription + TAM		✓	✓		
Existence		✓		✓	✓
Happening			✓		
Location		✓		✓	✓
Possession		✓		✓	✓
Becoming			✓		

9.6. Notes on the origin of copulas

Benedict (1972: 62) reconstructs the PTB root of the existential ‘be, exist’ as **s-ri* [STC #264] on the basis of Tibetan *srid-pa* ‘existence’ and Burmese *hrí* ‘to be’. The etymology of this verb root in Lahu and beyond was explored rather reverentially in Matisoff (1985), who posited a semantic association between an existential verb and the term for the Supreme Being.

It may be significant that the Mongsen existential *li* shares a liquid initial consonant and a high front vowel with the reconstructed form, and is also attested in a number of words that relate to existence and creation. The foremost of these occurrences can be found in the name of *lítʃápáʔ*, the deity responsible for creating the world in Ao folklore. The root *lí* also occurs in *a-lí* ‘NRL-ground’, *li-ma* ‘ground-FACE’ and *a-lí-má* ‘NRL-ground-FACE’, with the relational noun root *ma* ‘face’ contributing respectively to the meanings of ‘country’ and ‘earth, world’ in the latter two lexicalized compounds.

Chapter 10

Imperatives

Mongsen employs three imperative mood affixes for encoding different kinds of commands. These are: the positive imperative suffix *-aŋ*, the prohibitive prefix *tə-*; and the typologically rare admonitive prefix *asáʔ-*. Our first task will be to identify the characteristics of imperatives that distinguish this mood from the other types of Mongsen clauses. Following this, our attention turns to the use of the three imperative markers, dependencies between imperative mood and other categories, and the semantics of imperatives.

10.1. Imperative as a distinct clause type

A number of incompatibilities are found between the three markers of imperative mood and some of the grammatical categories encoded by the morphology of interrogative and declarative clauses, thereby justifying the recognition of imperative mood as a separate clause type.

Firstly, the imperative mood differs from the declarative and interrogative moods in not being compatible with the verbal suffixes used to encode absolute tense. Secondly, an imperative clause is incompatible with the declarative mood clitic (§4.2.14.6). Thirdly, verb stems of imperative clauses must encode negative commands (i.e. admonition and prohibition) very differently to the way that negative polarity is encoded in interrogative and declarative clauses. The verb stems of non-imperative clauses are obligatorily negated by the negative prefix *mə-*, whereas verb stems marked with an imperative marker are incompatible with this prefix. Although the admonitive and prohibitive prefixes occur in the same prefix slot of the verb stem as the negative prefix *mə-* of non-imperative clauses, it is still a relevant observation that a verb stem inflected with the positive imperative suffix *-aŋ* cannot also be negated by *mə-* to render a negative command. These morphological co-occurrence restrictions thus isolate the imperative mood as a distinct clause type. Some semantic restrictions affecting the use of certain aspectual and modal categories in imperative clauses are additionally discussed in §10.3 below.

There are no consistent differences between imperatives and other clause types in the expression of core arguments in A or S function, although it is observed that they generally occur less frequently in imperative clauses. This might be attributable to the fact that addressees of commands are by nature already highly referential, which consequently allows for the pragmatically

licensed ellipsis of their noun phrase arguments. When they are overtly expressed, core arguments in A or S function tend to be followed by the topic particle *la*. Sometimes *la* serves a contrastive topic marking function in an imperative clause, as in the textual example of (10.1b), although this is not consistently the case. Native speakers will sometimes mark an A or S argument with the topic particle in an elicited sentence that is uttered in complete isolation from any context that could motivate the need for pragmatic contrast, claiming that it makes a command sound less abrupt. This attenuating function may represent yet another grammaticalized use of the topic particle, in addition to those previously outlined in §4.2.14.3.

(10.1) a. *təku “tùhɲàɪ la atʃak thən-ù? tə.”*

tə-ku tùhɲàɪ la a-tʃak thən-ù? tə
 thus-LOC other TOP NRL-paddy sow.PST-DEC PTCL
 ‘And so, [Mechatseng says] “[But] others sowed rice!”’

b. *tə, “nàŋ la am tʃaŋ khə aku? tʃaŋ nət táŋ áhlú tʃu ku liŋaŋ.”*

tə nàŋ la am tʃaŋ khə aku? tʃaŋ nət
 thus 2SG TOP aspидistra.sp seed CONJ gourd.sp seed two
 táŋ a-hlú tʃu ku liŋ-aŋ
 just NRL-field DIST LOC plant-IMP
 ‘Then [Tsengrutsela replied] “You just plant aspидistra and bottle
 gourd seeds in the field.”’

The following textual examples serve to demonstrate that the topic particle is not required for the grammaticality of an imperative clause with an overtly expressed argument in A or S function.

(10.2) *nàŋ anu thən mùŋaŋ.*

nàŋ a-nu thən mùŋ-aŋ
 2SG NRL-child COM rest-IMP
 ‘You rest at home with the child.’

(10.3) *təɪ, “nàŋ íjáj? kəmtaŋkə táŋ, jipəɪə .ɪaŋ.”*

tə-əɪ nàŋ íjáj? kəmtaŋ-əkə táŋ jip-əɪ .ɪə-aŋ
 thus-SEQ 2SG much be.careful-SIM just sleep-SEQ come-IMP
 ‘And, [he said] “You just sleep very carefully [there] and then come.”’

The influence of pragmatics and semantics over the grammatical marking of clausal arguments in declarative and interrogative clauses also extends to the core arguments of imperative clauses. It is generally the case that A and S arguments of imperative clauses tend to receive no overt case marking,

the scope of imperative mood can extend over associated dependent clauses as well; examples of this are discussed in §10.3.1 below. Intonation is not systematically exploited for the prosodic marking of any type of imperative clause.

10.2.1. Positive imperative

Positive imperative mood is marked by the imperative inflection *-(a)ŋ* on the verb stem. While the default reference of imperatives is the second person, a personal pronoun or other deictic nominal such as personal name or a kinship term can be used in topic function to clarify the reference of a command. The personal name in the following example is used as a form of vocative.

- (10.6) *nuksənsaŋpa? – nàŋ aki nə kə.ɬaŋ.*
 nuksənsaŋ-pà? nàŋ a-ki nə kə.ɬa-aŋ
 PN-M 2SG NRL-house ALL ascend+come-IMP
 ‘Noksensangba – come up to the house.’

Being a wholly dependent-marking language at the clause level (Nichols 1986), Mongsen does not formally distinguish number or person in verbal inflection; consequently context must be relied upon to determine whether a verb marked by *-(a)ŋ* should be interpreted as having hortative first person or imperative second person reference. First person singular reference in commands is highly unlikely, but first person plural reference is occasionally encountered in texts. In the sentence of (10.7) below, for example, there is no morphology by which a hortative mood separate from imperative mood might be identified. Yet in the context of the source narrative, it is clear that the speaker is exhorting the addressee to accompany him on the return journey home, and not merely ordering the addressee to approach him.

- (10.7) “*təm-pà?, ɬaŋ, aki nə tʃhuwa.ɬù,*” *tə saw? tʃu.*
 təm-pà? ɬa-aŋ a-ki nə tʃhuwa-ɬù
 friend-M come-IMP NRL-house ALL emerge-IMM
 tə sa-ù? tʃu
 thus say.PST-DEC DIST
 ‘“Come friend, [we] will go home,” [he] said like that.’

Imperatives encode third person reference relatively rarely. Jussive-type commands tend to occur with overt third person noun phrase arguments,

probably so that any default second person reference of the imperative mood is overridden. The following Khensa Mongsen example illustrates this usage.

- (10.8) *inəti mə̀phila mi.ù? ná. jìmaŋ tsə t́hlàŋ à tʃhàpiaŋ.*
 inət i mə̀-phi-la-mi-ə̀.ù? ná
 2DU.IN PROX NEG-separate-NEG-DESID-PRES-DEC PTCL
 jìmaŋ tsə t́-hlàn à tʃhà-pi?-aŋ
 road DIST NZP-be.long one do-CAUS-IMP
 ‘We two don’t wish to be parted. May the road be a long one for us.’¹

In the absence of an overt first or third person noun phrase, sometimes context alone can be sufficient for interpreting non-second person reference to the argument of an imperative, as in the following Waromung Mongsen example. The existential verb *li* is used here as a matrix verb, marked by the irrealis suffix (which takes the form *-ə̀* in Waromung Mongsen). This signals the continuity of the state commanded by the imperative-marked verb stem beyond the moment of speech. Additional examples of imperatives expressing continuative commands are discussed in §10.3.1.

- (10.9) *pùkphu na “mijimtʃàɪ atəma mànlikà? mànəŋ liə̀, saŋalikà? saŋəŋ liə̀,”
 tà sa.*
 pùkphu na mijimtʃàɪ atəma màn-likà?
 owl AGT humankind BEN be.dark-CONTEMP
 màn-aŋ li-ə̀ saŋa-likà? saŋa-aŋ li-ə̀
 be.dark-IMP stay-IRR be.light-CONTEMP be.light-IMP stay-IRR
 tà sa
 thus say.PST
 ‘‘For humankind, if it becomes dark, may it remain dark; if it becomes
 light, may it remain light.’’ [said Owl].²

10.2.2. Prohibitive

The prefix *tə̀-* encodes the prohibitive mood. It was noted above in §10.1 that this marker occurs in paradigmatic opposition with the negative prefix *mə̀-* of

-
1. The meaning encoded by the causative suffix (*-pi?* in Khensa Mongsen) that co-occurs with the imperative suffix in (10.8) is discussed in detail in §10.3.2.
 2. The contemporative converb *-likà* in this example from a stylized folklore narrative contextually expresses an unexpected conditional meaning. See Text 4, line 11 for an additional example of its use as a conditional, on this occasion uttered by a Mangmetong Mongsen speaker in the narration of the same folktale.

declarative and interrogative clauses, and the admonitive mood prefix *asáʔ-*, described below in §10.2.3. The prohibitive and the admonitive moods are consistent with the positive imperative in not being able to take declarative mood or absolute tense marking; nor can they occur with interrogative pronouns or particles.

As with other types of commands, the prohibitive is only found in direct speech. It occurs on both transitive and intransitive verb stems (including stative verb stems), and on possessive copular verb stems.

- (10.10) *jimli philəməkə, nàŋ na nì tətəpsət.*
 jim-li philəm-əkə nàŋ nə nì tətəp-sət
 law-? think-SIM 2SG AGT 1SG PROH-hit-AFFECT
 ‘Please,³ don’t kill me.’

- (10.11) *“nàŋ əntsə tətà” tətə saw? .iə.*
 nàŋ əntsə tət-à tətə sa-ùʔ tətə
 2SG halfway PROH-come thus say.PST REP
 “‘Don’t come to meet me half-way,” [she] said.’

- (10.12) *nàŋ tʃáʔ sisàj tətəkà.*
 nàŋ tʃáʔ sisàj tət-əkà
 2SG nothing worry PROH-have
 ‘Don’t you worry about anything.’

- (10.13) *nàŋ nə tətʃàsi mà.*
 nàŋ nə tət-ʃàsi mà
 2SG AGT PROH-be.distressed PTCL
 ‘Please don’t be upset.’

10.2.3. Admonitive

The admonitive is encoded by a prefix *asáʔ-* that occurs in paradigmatic opposition with the negative prefix of non-imperative clauses and the prohibitive prefix described in the preceding section. The admonitive mood expresses a meaning akin to ‘Don’t VERB too much’ and is used to dissuade or

3. In Mongsén, ‘please’ is expressed using a formulaic expression *jimli philəməkə*, literally ‘think of the law’, presumably referring to tribal law. I am not sure of the meaning of the suffix *-li*, but its phonological form suggests that it may be related to the existential copula (see §9.2).

limit the performance of an activity. It is compatible with intransitive and transitive activity verbs and intransitive stative verbs.

- (10.14) *asáməni.*
 asáʔ-məni
 ADM-laugh
 ‘Don’t laugh too much.’

- (10.15) *nàŋla ásáʔ asátʃàʔ.*
 nəŋ la á-sáʔ asáʔ-tʃàʔ
 2SG TOP NRL-meat ADM-consume
 ‘Don’t you eat too much meat!’

- (10.16) *asátʃàsi.*
 asáʔ-tʃàsi
 ADM-be.distressed
 ‘Don’t be too upset.’

A possible diachronic source of the admonitive prefix is the intransitive verb root *asáʔ* ‘be deliberate, cunning’, which in turn may be cognate with the verb root *sa* ‘say’ (see fn. 4 under §5.2.2). If so, the historical process by which this prefix has grammaticalized and attained its admonitive connotation is presently not understood.

10.2.4. Semantic verb classes and imperative mood

Verbs can be divided into those that take imperative inflection directly on their stems, and those that must occur in a converb clause in a dependent relationship to a matrix existential verb inflected for the imperative mood.

All activity verbs accept an imperative inflection, regardless of their valency. Stative verbs of cognition also take imperative inflection directly on their stems.

- (10.17) *ahəj imtisàŋpàʔ – təza tənu minəŋ maw.*
 ahəj imtisàŋ-pàʔ tə-za.tə-nu min-aŋ mà ù
 EXCLM PN-M RL-child.RL-child pity-IMP PTCL PTCL
 ‘Hey, Imtisangba! At least have pity on your children.’

- (10.18) *atʃu nə təməluŋ nəmaŋ.*
 a-tʃu nə tə-məluŋ nəm-aŋ
 NRL-DIST INST NZP-be.willing be.satisfied-IMP
 ‘[May your heart⁴] be satisfied with that!’

Stative verbs expressing value or human propensity are not directly marked for imperative mood. These must occur in a dependent converb clause expressing the manner or state, and an existential verb in the matrix clause carries the imperative marking. Under these circumstances, the scope of the imperative extends over both the matrix and the dependent clause. The discourse focus particle in the following examples serves a restrictive function that is not specifically associated with imperative clauses, but frequently follows converb clauses (see §11.4).

- (10.19) *a.uəkə táŋ li-aŋ.*
 a.u-əkə táŋ li-aŋ
 good-SIM just stay-IMP
 ‘Just be good.’

- (10.20) *azən tʃhuwakə táŋ li-aŋ.*
 azən tʃhuwa-əkə táŋ li-aŋ
 power emerge-SIM just stay-IMP
 ‘Just [go ahead and] let your anger out.’ (= Be angry!)

Example (10.9) presented above appears to provide counterevidence to the claim that stative verbs cannot take imperative inflection directly on their stems. There are two extenuating factors that should be taken into account when considering the evidence of this apparent exception, however. Firstly, this example sentence comes from a highly stylized folklore story in which the birds are deciding how to divide up the world into periods of day and night.⁵ Whilst it may be possible to order the world to assume a particular state in this folkloric context, it is unlikely that a rational native speaker would order an individual to assume a state over which they have no possible control, even less a non-human entity. The use of stative verbs with imperative inflection may therefore be restricted to such mythical contexts. Secondly, the reference of the imperative is third person. Jussive commands are likely be subject to different constraints regarding the use of stative verbs, because they can be used to express wishes

4. The citation form for ‘heart’ is a nominalised verb-noun compound *tə-məluŋ tʃaŋ* (NZP-be.willing + seed). The verbal noun *təməluŋ* is perhaps most appropriately equated with ‘mind’.

5. See Text 4.

and desires for outcomes or states that are not rationally possible in the real world, e.g. in blessing formulae and curses.

10.3. Dependencies between imperatives and other categories

The vast majority of grammatical categories have the same meanings and the same forms in imperative clauses as in non-imperatives, but some dependencies are found to operate between certain grammatical categories and the imperative mood. The epistemic modality auxiliary *liw?*, for example, does not occur with the matrix verbs of imperative clauses due to its semantic incompatibility with commands, nor does the continuative aspect suffix *-ja* (see §8.5.14 and §8.5.4.1 respectively for description of their functions in non-imperative clauses). Recall from the discussion of §10.1 that markers of absolute tense cannot be used on stems that also have an imperative affix.

All imperatives are understood to have default present tense reference. Absolute tense distinctions, then, undergo complete neutralization in imperatives, with the consequence that commands for actions to be performed habitually or posterior to the moment of speech must be expressed via lexical means. The sentence of (10.21) illustrates such a strategy.

- (10.21) *təɪ nìŋkhə̀la, ní? à, ní? à tə̀, kətaŋ nə .ɪaŋ.*
 tə-əɪ nìŋkhə̀la ní? à ní? à
 thus-SEQ 2PL one.day one one.day one
 tə̀ kə taŋ nə ɪa-aŋ
 thus 1SG.POSS SIDE ALL come-IMP
 ‘And you all come to me, one per day.’

Dependencies holding between imperatives and other grammatical categories are individually discussed in the following subsections.

10.3.1. Absence of continuative aspect marking

The only aspectual category that never co-occurs with a matrix clause verb inflected for imperative mood is the continuative aspect. In its place, Mongsen uses other means to express continuative commands denoting the semantic equivalent of ‘keep doing VERB’. One strategy makes use of reduplication of the base and any accompanying particles.⁶ The continuative meaning of

6. The reduplicated imperative clause can also be used to hasten the performance of an activity.

reduplicated activity verbs in commands correlates closely with the durative meaning of reduplicated activity verbs marked by the simultaneous converb, described in §8.1.4. In the following example, the interrogative particle is used to soften the command, rather than to signal interrogative mood (see §10.4.1 for discussion of particles used with imperative mood).

- (10.22) *pasiaŋ ní, pasiaŋ ní.*
 pasi-aŋ ní pasi-aŋ ní
 search-IMP QPTCL search-IMP QPTCL
 ‘Keep searching.’

A negated converb clause in a dependent relationship to an imperative matrix clause provides an alternative strategy for expressing a continuative imperative. The negated verb stem of the non-final converb encodes the continuative manner of the activity, while the matrix verb encodes the actual nature of the commanded activity. The dependent status of the non-final clause is often signalled by a high or rising intonation on the negative suffix and its syntactic position before the clause-final matrix verb. The use of intonation as a boundary signal is described in §3.4.2, and a description of negated narrative chaining converbs can be found in §11.4.1.

- (10.23) *məhməntsəla pasiaŋ.*
 mə-hməntsə-la pasi-aŋ
 NEG-stop-NEG.CV search-IMP
 ‘Search without stopping.’

An option in dependent clauses of positive polarity is to use a simultaneous converb to impart an imperfective reading. This occurs with or without the continuative aspectual suffix. The simultaneous converb clause functions as an adjunct of an existential verb taking the imperative inflection in the matrix clause. A core meaning of *li* is ‘stay’; when marked by the imperative inflection in this type of complex sentence, it encodes the expectation that the commanded activity be performed beyond the immediate time frame. In contrast to the preceding example, in (10.24) the dependent converb clause expresses the nature of the activity, while the inherent semantics of the matrix verb *li* contributes to the continuative aspectual nuance, in addition to providing a vehicle for the expression of imperative mood.

- (10.24) *nì mijjə̀kə ɬàthùŋku, tʃhə̀niŋjakə táŋ liaŋ.*
 nì mijjip-əkə ɬà-thùŋ-ku
 1SG turn-SIM come-REACH-LOC.CV
 tʃhə̀niŋ-ja-əkə táŋ li-aŋ
 work-CONT-SIM just stay-IMP
 ‘When I return, just be working.’

The scope of a matrix clause’s imperative mood may or may not extend over its dependent converb clause(s), depending upon pragmatically and/or semantically determined contextual factors and the type of converb involved. In (10.24) above, for instance, the scope of the matrix clause imperative mood encompasses the simultaneous converb *tʃhə̀niŋjakə*, but obviously does not extend to the two preceding converbs in *nì mijjə̀kə ɬàthùŋku*. This contrasts with the scope of the matrix clause imperative mood in (10.25). Here it extends over all the dependent predicates, according to the native speaker intuitions of Mongsen consultants.

- (10.25) *təɪ am tʃu kà? tsəkəɪ, íjá? aɬukə laŋəɪáŋ ʃhàwkəkə liaŋ.*
 tə-əɪ am tʃu kà? tsək-əɪ íjá? aɬu-əkə
 thus-SEQ aspidistra.sp DIST also pluck-SEQ very good-SIM
 laŋ-əɪ áŋ tʃhàwk-əkə li-aŋ
 roll.up-SEQ just keep-SIM stay-IMP
 ‘And, also pluck aspidistra leaves, roll them up properly, and keep them [for use as plates].’

Matrix clause operator scope only extends to a dependent converb clause if that clause is interpreted as encoding a non-presuppositional meaning, as represented by the translation of (10.25). On the other hand, if the converb clause is understood to express a presuppositional meaning that in some way qualifies the matrix clause event with respect to temporal setting, reason, condition, concession etc. – as *nì mijjə̀kə ɬàthùŋku* clearly does in (10.24) – then the scope of the imperative mood is limited to its matrix clause (and any associated non-presuppositional dependent clauses, such as *tʃhə̀niŋjakə táŋ* in that example). The scope of matrix clause operators over dependent clauses is discussed further in §11.3.

10.3.2. Causative marking in imperatives

The underlying phonological form of the causative suffix is *-i?* in the Mangmetong variety of Mongsen. Its valency increasing function is described

in §5.4.1. Recall from §3.5 that the glottal stop is deleted in the environment of a morpheme boundary. This means that the particular word-internal allomorphic realization of the causative marker in Mangmetong Mongsen (either as *-i* or *-j*) depends upon whether the suffix syllabifies before the imperative suffix as a nucleus or as an onset in the process of word formation.

In imperative clauses, the causative suffix potentially expresses a number of different meanings besides the expected valency-changing causative meaning. One of these is a typologically unusual imperative mood that I will refer to as the ‘autobeneficiary imperative’ (see §10.5.1 for description). Another possibility is a permissive meaning. Speakers rely upon context or the type of case-marking on noun phrases (if they are overtly expressed) to arrive at the correct interpretation of a causative-marked imperative clause. Consider the following elicited examples.

- (10.26) a. *atsə tʃəmaŋ.*
 a-tsə tʃəm-aŋ
 NRL-water drink-IMP
 ‘Drink the water.’
- b. *atsə tʃəmjəŋ.*
 a-tsə tʃəm-iʔ-aŋ
 NRL-water drink-CAUS-IMP
 (i) ‘Drink the water [for me/him/us etc].’
 (ii) ‘Make [him] drink the water.’
 (iii) ‘Let [him] drink the water.’

The imperative clause of (10.26a) produces a meaning consistent with that described for positive imperative mood in §10.2.1. Causativizing the verb stem of this example in (10.26b) creates the conditions for three possible meanings to emerge: option (i) assumes the autobeneficiary interpretation; option (ii) is consistent with the valency-increasing causative derivation described in §5.4.1; and option (iii) interprets the causative suffix as encoding a permissive meaning.

These various possibilities are licensed by the absence of case-marked noun phrases. Of course, it is possible to be more precise by including noun phrases with specific case-marking, and this has the effect of forcing a particular meaning.

- (10.27) a. *pa atəməkə atsə tʃəmjəŋ.*
 pa atəməkə a-tsə tʃəm-iʔ-aŋ
 3SG BEN NRL-water drink-CAUS-IMP
 ‘Drink the water for him.’

- b. *pa li atsə tʃəmjaŋ.*
 pa li a-tsə tʃəm-iʔ-aŋ
 3SG DAT NRL-water drink-CAUS-IMP
 ‘Let him drink the water.’

As we can see from these elicited examples, the causative marker does not always perform a valency-increasing function in imperative clauses. Perhaps because it can solicit the performance of an activity for the benefit of someone else, native speakers also feel that the inclusion of the causative suffix encodes somewhat more of a polite request than an overt command when a verb stem is marked by the imperative inflection alone. This is apparent in the following textual examples.

- (10.28) *təku lítʃápáʔ nə “tʃùŋkuli tə tshəɪ, kùmaj ku zəjaŋ,” tə sawʔ təɪ.*
 tə-ku lítʃá-pàʔ nə tʃùŋkuli tə tshə-əɪ
 thus-LOC.CV PN-M AGT anthill.soil thus pound-SEQ
 kùma i ku zə-iʔ-aŋ tə sa-ùʔ təɪ
 wound PROX LOC spread-CAUS-IMP thus say.PST-DEC REP
 ‘And so, Lichaba said “Pound up some ant-hill soil and then spread [the paste] on this wound for me.”’
- (10.29) *təɪ thəplakzəkəɪ, “nàŋ la nì ləmsiɪu tʃu tətʃha hlajaŋ.”*
 tə-əɪ thəp-lak-zək-əɪ nàŋ la nì ləmsiɪu tʃu
 thus-SEQ throw-DIR-SEND-SEQ 2SG TOP 1SG comb DIST
 tətʃha hlà-iʔ-aŋ
 NZP-pick.up descend+go-CAUS-IMP
 ‘And, after throwing down the comb [under the floor of the house, she said] “Go down to pick up my comb for me.”’

Lastly, a causativized imperative can be used to express a jussive-like meaning, such as a wish that a desirable set of circumstances might become a reality. We previously saw this extension of meaning in the Khensa Mongsen sentence of (10.8) above. Such jussive interpretations of causativized imperatives do not appear to be far removed from the permissive sense suggested by interpretation (iii) of (10.26b) above.

10.3.3. Other grammatical categories and imperative mood

Apart from the few exceptions documented above, many other semantically compatible verbal grammatical categories may co-occur with imperative mood.

Any of the lexical suffixes can occur on verb stems inflected with the different types of imperative affixes. The majority of these express lexically-oriented resultative meanings that closely correspond to the meanings they encode when used as lexical verbs, e.g. SEPARATE, END, SEVER, etc. (see §8.5.1 for a full description of the meanings encoded by lexical suffixes).

- (10.30) *“athi ku naṅthənaṅ,” tə saw? tət̪i-ù?*
 a-thi ku naṅ-thən-aṅ tə sa-ù? tət̪i-ù?
 NRL-field.hut LOC heap-GATHER-IMP thus say-DEC REP-DEC
 “‘Heap [the gourds and aspidistra fruit] together in the field hut,”
 [she] said.’

- (10.31) *tʃathəmaṅ.*
 tʃəʔ-thəm-aṅ
 consume-END-IMP
 ‘Finish eating [it] up.’

In addition to the valency increasing causative suffix, a valency modifying collective/reciprocal suffix can occur with imperative inflections (see §5.4 for description of all valency modifying morphology).

- (10.32) *khələm nə tʃhəniṅtəp-aṅ.*
 khələm nə tʃhəniṅ-təp-aṅ
 together INST work-COLL-IMP
 ‘Work together.’

A variety of aspectual categories including perfective, repetitive and completive occur on verb stems with imperative inflection (see §8.5 for a comprehensive description of all verbal suffixes).

- (10.33) *hmapa tʃu tʃhəniṅtʃuk-aṅ.*
 hmapa tʃu tʃhəniṅ-tʃuk-aṅ
 work DIST work-PFV-IMP
 ‘Finish the work.’

- (10.34) *zaṅlusiaṅ.*
 zaṅlu-si-aṅ
 make-RPET-IMP
 ‘Repair [it].’ (Literally: ‘make again’)

(10.35) *tʃamaŋ.*
 tʃãʔ-maʔ-aŋ
 consume-CMPL-IMP
 ‘Eat [it] up completely.’

(10.36) *mətuyʃakaŋ.*
 mətuy-tʃak-aŋ
 erect-RS-IMP
 ‘Leave [it] standing up.’

(10.37) *təzaŋlusi.*
 tə-zaŋlu-si
 PROH-make-RPET
 ‘Don’t repair [it].’

(10.38) *təmáŋ təʃamaʔ.*
 təmáŋ tə-tʃãʔ-maʔ
 all PROH-consume-CMPL
 ‘Don’t eat [it] all.’

The conative marker *-tsə* and the abilitative marker *-tʃhət* are the only two modality suffixes that occur in imperative clauses. These express the same meanings in imperatives as in declarative clauses. The other modality markers, namely the frustrative *-phət*, the desiderative *-mì* and the chance marker *-tʃhaʔ*, do not co-occur on verb stems with the imperative inflection because of their semantic incompatibility with commands (see §8.5.5 for description of their use). As noted above, semantic incompatibility also explains the absence of the epistemic auxiliary *liwʔ* (expressing the degree of certainty of an event taking place) occurring in imperative clauses.

(10.39) *təɪ liŋəɪ, “áhlù nə watsəaŋ,” tə tʃhàwʔ .ìə.*
 tə-əɪ liŋ-əɪ á-hlú nə wa-tsə-aŋ
 thus-SEQ plant-SEQ NRL-field ALL go-CON-IMP
 tə tʃhà-ùʔ təɪ
 thus do.PST-DEC REP
 ‘And then, after [Mechatseng] had done the planting, [Tsengrutsela said] “Go to the field and have a look.”’

- (10.40) *nì mijipəkə .jàthùŋku, zəlutfhətəŋ məw.*
 nì mijip-əkə .jà-thùŋ-ku zəlu-tʃhət-aŋ mà u
 1SG turn-SIM come-REACH-LOC.CV write-ABIL-IMP PTCL PTCL
 ‘By the time I return, [I expect you to] be able to write.’

10.3.4. Imperatives and evidentiality

The reportative evidential particle *təɪ* is used with embedded imperative clauses to signal commands by proxy. This is achieved by using the imperative clause as an O complement of an utterance predicate in an evidential-marked declarative clause.

- (10.41) *akukpà? nə “puəŋ,” tə saw? təjù?*
 akuk-pà? nə pù?-aŋ tə sa-ù? təɪ-ù?
 PN-M AGT carry.on.back-IMP thus say-DEC REP-DEC
 ‘Akokba reportedly said “Carry [him].”’

A verb stem can inflect for imperative mood and then simply be juxtaposed to the reportative particle, as in (10.42).

- (10.42) *.jaŋ təjù?*
 .ja-aŋ təɪ-ù?
 come-IMP REP-DEC
 ‘Come’ [she evidently said].

While this does not require the overt occurrence of a complement-taking utterance predicate and its quotative complementizer in the matrix clause, these constituents are nevertheless understood to be present underlyingly and are restorable to their usual positions between the O complement and the reportative particle (as in [10.41]). Context allows the elided predicate to be recovered from the discourse. The syntactic structure of such complex sentences with embedded imperatives is therefore identical to that of embedded direct speech declarative clauses described in §11.7.1, as both types of embedded clauses form O complements of their explicit or implicit matrix declarative clauses.

10.4. Semantics of imperatives

The imperatives of Mongsen lack register differentiation, having only one basic meaning for each category of positive imperative, negative imperative and

admonitive mood. In each category, a command signalled with an imperative inflection alone is interpreted as being inappropriately curt if used by a social inferior to a superior. In some contexts this does not apply, such as when people of differing social status are working together in a spirit of communal egalitarianism towards a common purpose. Conversely, it is quite common for superiors not to use politeness particles to attenuate commands when speaking with social inferiors.

10.4.1. Use of particles in imperative mood

Several particles can be used with the positive imperative to impart various shades of meaning. With the exception of the interrogative particle *ní*, these form a unique sub-class that is not used to signal illocutionary force in other mood types.

The particle *má?* expresses a mild degree of impatience and the need for urgency. It can be used by a participant lower in the social hierarchy to give an order to someone of higher status in certain contexts; for example, in a situation in which a father and son are working together.

- (10.43) *atsə hənzəkəŋ má?*
 a-tsə hən-zək-aŋ má?
 NRL-water take-SEND-IMP PTCL
 ‘Take the water!’

The particle *atə* softens a command and is often used to implore that an action be performed after an initial request is not granted. It can also impart a tone of slight irritation. Sometimes it is used in a cajoling manner with reluctant children to encourage their participation in an activity.

- (10.44) *nì kà? akuŋ tʃu à khiŋ atə.*
 nì kà? akuŋ tʃu à khiŋ-aŋ atə
 1SG also prawn DIST one give-IMP PTCL
 ‘Give me one of those prawns as well.’

The particle *mà* changes a brusque command to a polite request. This is often used by a social inferior to articulate a command to a social superior, and is also used by non-intimate equals.

- (10.45) *í pá? mətəmi hənəɪaɪ tʃəmaŋ mə.*
 í pá? mətəm i hənəɪà-əɪ tʃəm-aŋ mə
 EMPHAT manner PROX bring-SEQ drink-IMP PTCL
 ‘Drink [holding the cup] in this manner.’

Lastly, an interrogative particle can be used with imperatives to form polite commands, the most commonly occurring being *ní*. Observe that this is articulated with a different tone and meaning from the declarative clause particle *ni* of (10.48) below. As noted in §10.3.1, *ní* is used to attenuate commands in imperatives and does not actually encode an interrogative in this context of use.

- (10.46) *li aŋ ní.*
 li-aŋ ní
 stay-IMP PTCL
 ‘Wait a little.’

10.4.2. Non-command meanings of imperatives

Imperative constructions are used non-prototypically in certain speech formulae, such as in blessings and curses.

- (10.47) *tsəŋhɪəɪm nə muwajaŋ mə.*
 tsəŋhɪəɪm nə muwa-iʔ-aŋ mə
 deity AGT bless-CAUS-IMP PTCL
 ‘May God bless you.’

- (10.48) *phəlak-sət-əkə sə aŋ ní.*
 phəlak-sət-əkə sə-aŋ ní
 be.accidental-AFFECT-SIM die-IMP PTCL
 ‘May you accidentally die.’

This last example has cultural significance. Prior to the changes wrought by the advent of Christianity, the worst fate that could befall an Ao family would be for one of its members to suffer a death resulting from an attack by a wild animal, a snake bite, childbirth, drowning, a fall from a height, or other accident resulting in an untimely demise. Such deaths were considered accursed, requiring the family of the victim to abandon all of their material property – their clothes, money, all household utensils, livestock, their stores of grain and even their standing crops in the fields – in order to rid themselves of the defilement associated with the deceased family member. One strategy for

avoiding this catastrophic loss of wealth was for the family to immediately declare itself Christian, thereby avoiding the need to follow the old customs that would leave them destitute (Mills 1926: 283–286). My impression is that accidental deaths continue to be considered accursed, even by Christians, although they may have forgotten the source of this belief.

10.4.3. Imperatives and discourse

Imperatives are often used in direct speech in narratives, particularly in folklore narratives when there is interaction between people or personified animals. In real life, Mongsen speakers of higher social status tend to use pithy commands without politeness particles when giving orders to addressees occupying a position of lower status in the speech community, a typical example being when parents speak to their children. Conversely, speakers of lower social status tend to use politeness particles to turn commands into requests addressed to their elders. In village life, social standing is principally determined by the traditional Ao social hierarchy based upon generational age groups known as *hjaŋa*.

Commands can be acknowledged with interjections, either affirmatively using *hàj?* ‘okay’ or negatively with *mà?* ‘no’. The following text examples demonstrate the use of both affirmative and negative responses to commands ([10.50] is taken from a Waromung Mongsen narrative).

(10.49) a. “*təɪ tsəŋti tàkəɪ athi tʃu ku ləm-aŋ.*”
 tə-əɪ tsəŋti tàk-əɪ
 thus-SEQ bamboo.matting weave-SEQ
 a-thi tʃu ku ləm-aŋ
 NRL-field.hut DIST LOC lay-IMP
 “‘And then, weave the bamboo matting and lay it in the field hut,’
 [said Tsengrutsela].’

b. *təɪ “hàj?” tət saw? tət.*
 tə-əɪ hàj? tət sa-ù? tət
 thus-SEQ okay thus say.PST-DEC REP
 ‘And then, [Mechatseng] said “Okay.”’

(10.50) a. *áula! nàŋ na àzə tələm ku átsək atsəuaŋ atà.*
 a-á? -la nàŋ na a-zə təl-əm ku
 NRL-rat-F 2SG AGT VOC-grandmother RL-head LOC
 átsək atsə-à-aŋ atà
 louse look-IRR-IMP PTCL
 ‘Rat! Kindly look for lice on Granny’s [i.e. my] head.’

- b. *tʃùku* “*nì anjala tə# na təpsətəpà? əisi. mà?*.”
tʃùku *nì* *an-ja-la* *tə-#* *na*
 consequently 1SG chicken-DIM-F RL-mother AGT
təp-sət-ə-pà? *əisi-əi* *mà?*
 strike-AFFECT-IRR-NR chase-PRES no
 ‘[Rat replied] “Little Chick’s mother is chasing me to kill me.
 No!”’

An affirmative interjection can be used by speakers at all levels in the social hierarchy when agreeing to a command, or to express willingness to partake of an invitation (to eat, to go, etc.). The negative is considered rather abrupt when used in both types of speech act and I was discouraged from using this as a response to invitations when interacting with elders. However, my Mongsen teacher (25 years my senior, whom I addressed using the kinship term for father’s younger brother) freely used the negative interjection when responding to my questions or commands, and so usage does appear to be determined by one’s relative position on the social hierarchy with respect to age.

10.5. Notes on the origin of imperatives

The following discussion compares forms of imperatives in Tibeto-Burman languages of the neighbouring region and suggests etymologies for Mongsen imperatives, where supported by comparative evidence.

10.5.1. Positive imperative

The diachronic origin of the Mongsen positive imperative *-(a)ŋ* is unknown. While it is similar, if not identical, to corresponding forms of a number of closely related dialects and languages of the immediate region, such as Chungli Ao *-(a)ŋ*, Sangtam *-(i)ŋ* and Yimchungrü *-(a)ŋ*, it bears little similarity to the imperative markers of more distantly related languages that have been documented. Therefore it would seem that it developed quite recently, well after the splitting off of the Ao group from a higher branch of Tibeto-Burman. Generally it appears that the languages of southern Nagaland, such as Angami (Giridhar 1980) and Mao (Giridhar 1994) have an extensive, formally marked register system ranging from humble requests to curt commands, whereas the more northerly languages such as Ao and Lotha do not demonstrate this register differentiation, having just one positive imperative form.

The Konyak language Chang [Tibeto-Burman: Bodo-Konyak-Jinghpaw], spoken on the eastern border of the Ao territory, is suspiciously similar to Mongsen in having a special imperative form called the “imperative reflexive” by Hutton ([1929] 1987: 17). This is said to be used ‘only, and always, when the speaker himself is to be the object of the action ordered’.

The curiously named “imperative reflexive” is represented by two morphemes in Hutton’s grammatical sketch – *-ke-* and *-pu* – with the latter also occurring with the negative imperative prefix in a list of verbal inflections (1987: 19). The “imperative reflexive” has the same form as a suffix that Hutton labels the “indicative perfect”, which does indeed appear to have a declarative mood marking function. The Chang imperative reflexive complements a direct imperative *-(a)shi*, a permissive imperative *na...-a* (which is presumably a jussive or third person imperative) and a negative imperative *ta-*, in addition to having special roots for verbs expressing imperatives of ‘come’, ‘go’ and ‘give’. Text examples of the Chang direct imperative and the imperative reflexive are given below. These have been glossed to aid in the recognition of their relevant morphological forms.

Chang direct imperative (Hutton 1987: 50):

- (10.51) *Nyi kuk-kur-gu kou-shi!* *lauke*
 you cock-a-doodle-do crow-DIRECT.IMP said
 ‘[The leopard cat] said “You crow cock-a-doodle-do!”’

Chang imperative reflexive (Hutton 1987: 52):

- (10.52) *Seangtoke “Melbam-shoung, woklong yügan*
 Tortoise Herd.of.wild.pig sago.palm having.turncd.up
apanyo ngo songan-ke”
 having.thrown.away me release-IMP.REFL
 ““Pigs”, said the tortoise, “Root up this sago palm, throw it aside and release me!””

My own preliminary research on the Tuensang dialect of Chang confirms that the language has a prohibitive prefix and three positive imperative suffixes. The form of the prohibitive prefix is $tə^{11}$ - or $tə^{55}$ -, depending upon tone sandhi rules that are yet to be fully investigated. A positive imperative suffix $-a^{11}$ is used by social superiors to give commands to juniors, and a second imperative form $-ji^{11}$ is used in formal situations, or by socially inferior speakers when they give commands to their superiors. Age is observed to be the principal factor dictating a speaker’s choice of positive imperative suffix. Interestingly, a third positive imperative suffix $-pu^{211}$ is reserved for giving a command to perform an action for the benefit of the speaker. This particular imperative therefore

expresses a meaning that is semantically identical to the autobeneficiary imperative meaning potentially expressed by the causativized imperative verb stem of Mongsen, as described above in §10.3.2. I am also able to confirm that Chang does have special forms of imperative verb roots for ‘give’, ‘go’ and ‘come’, and that a suffix *-kəɹ¹¹* occurs as an allomorph of the autobeneficiary imperative, but seemingly only on some verb roots. The specifics of this distribution await further research.

Given the close geographical proximity of Chang to the Ao group of languages (cf. Map 2), it is possible that areal contact is responsible for an autobeneficiary imperative meaning being found in these two rather distantly related languages. But rather than the borrowing of a specific marker, what appears to have occurred is the borrowing of the semantic entailment of the specialized autobeneficiary imperative. Mills (1926: 2) mentions that villages such as Longla and Noksan, located near the eastern border of the Ao territory, had long been under Chang chiefs and had bilingual speakers of a dialect of Ao and Chang. This is precisely the kind of contact situation that could facilitate the borrowing of semantic content associated with a specific grammatical category, even if the actual form that encodes that semantic content is not borrowed. The fact that an existing combination of suffixes has been reinterpreted as marking an additional meaning in Mongsen imperative clauses suggests that the direction of semantic borrowing has been from Chang into Mongsen. Unfortunately very little is known about the grammar of Chang, despite it being over seventy years since Hutton published the first sketch, so additional inferences cannot be made until more data becomes available.

10.5.2. Prohibitive and admonitive

The prohibitive *tə-* has an attested PTB provenance (<**ta*) and is found widely distributed throughout the Burmese-Lolo, Himalayish, and Bodo-Garo branches of the family (Benedict 1972: 97), and on the evidence presented in this description, in the Ao group of languages as well.

As far as I am aware, an admonitive mood marker is unreported in any other Tibeto-Burman language of the north-east India.

Chapter 11

Clause combining

11.1. Preliminary remarks

Arguably the most pervasive and yet striking feature of Mongsen clausal syntax is its use of non-final verb forms to link dependent clauses. In this chapter I will justify using the term *converb* to refer collectively to these clause-linking, dependent verb forms. The converbs of Mongsen can initially be divided into two major types, according to their functions in the language. One function is to chain together potentially multiple sequences of clauses encoding non-presupposed information that advances the plot of a narrative. Another is to express presuppositions analogous to the meanings encoded by the adverbial lexical subordinators of English, such as *if, when, although, since* and *because*.

The salient difference between the two basic functions of converbs resides in the nature of the interpropositional relationship holding between the converb clause and its matrix clause. Because narrative chaining converbs typically do not assert a presupposition, this allows them to express multiple events of equal narrative rank with their matrix clause events (Johanson 1995: 322), potentially resulting in extremely long sentences by European language standards. As Haiman and Thompson (1984: 512) observe, the meanings of such clause combinations seem to parallel those of syntactically coordinate clauses. In contrast, the specialized converbs obligatorily assert adverbial-like presuppositions that qualify their matrix clause events in terms of temporal setting, condition, concession, reason, circumstance, and so forth. The propositional content of such clauses is often found to constrain the specialized converbs to a binary [[dependent clause] matrix clause] type of structure, although it is also possible for multiple concatenations of narrative chaining and specialized converbs (as well as other types of dependent clauses) to be interwoven through the tapestry of a single complex sentence, thereby creating layer upon layer of hierarchical structure.

The analysis of the converb system of Mongsen is complicated by the observation that the narrative chaining converbs can also be used to encode typically “adverbial” propositions that qualify or presuppose the assertions of their matrix clause predicates. The conflation of these two functions in the same converb form is in fact an areal feature of narrative chaining converbs, not only in many languages of Central and South Asia (Bickel 1998: 395), but also in some predicate-final languages of East Asia, such as Japanese. Consequently, one is ultimately dependent upon pragmatics and other contextual information

to determine whether a narrative chaining converb is encoding either a presuppositional adverbial meaning in a given instance of use, or a non-presuppositional meaning associated with the sequential chaining function discussed above. Recall that in the chaining function, the meaning of the converb does not presuppose an adverbial proposition, but merely reports on an event that is semantically coordinate with the event of its matrix clause. Narrative chaining converbs can also be semantically “integrated” with their matrix clause verbs to represent a single unified event, as opposed to two separate events that are chronologically ordered.

The description commences with an overview of definitions of the converb that have been proposed in the recent typological literature. I then consider the syntactic status of Mongsen converbs, taking into account the criteria of the proposed definitions. This is followed by detailed descriptions of each converb category and its diachronic source, where these can be recognized. Complementation is addressed in the final section of the chapter.

Relative clauses and nominalized clauses function syntactically as noun phrase attributes and clausal noun phrase arguments respectively. These have therefore been dealt with under §6.6 in the description of the noun phrase and need not be revisited here.

11.2. Definitions of the converb

I begin by reviewing definitions of the converb before investigating their relevance to the analysis of clause linkage in Mongsen. Perhaps the most important and controversial work to date has been a collection of papers edited by Haspelmath and König (1995). This seeks to establish the validity of a grammatical category of “converb” from a typological perspective.

In his introductory contribution, Haspelmath (1995: 3–8) defines a converb as ‘*a non-finite verb form whose main function is to mark adverbial subordination*’ (the definition is italicized in the original). Haspelmath proposes that converbs:

- (i) form part of the inflectional paradigm of verbs;
- (ii) are non-finite;
- (ii) are adverbial, in the sense that they modify verbs, clauses or sentences but never nouns or noun phrases;
- (iv) are subordinate, i.e. embedded into a superordinate clause.

V.P. Nedjalkov (1995: 97) assumes a similar point of view, but makes no reference to non-finiteness in his set of defining criteria. This has the potential advantage of accounting for languages with finite converbs that would otherwise be excluded by Haspelmath’s italicized assertion, although as Bickel

(1998: 382) notes, it subsequently allows any adjunct clause that is marked by a verbal suffix, rather than a free conjunction, to be recognized as a converb. V.P. Nedjalkov's definition also avoids any reference to subordination (discussed below), a concept whose definition is as problematic for linguistic typology as the notion of the converb.

V.P. Nedjalkov (1995: 97) prefaces his definition with the following statement: 'As a first approximation, we can define a converb as a verb form which depends syntactically on another verb form, but is not its syntactic actant, i.e. does not realize its semantic valencies'. He then proposes that a canonical converb can occupy:

- (i) the position of an adjunct (i.e. an adverbial);

but it cannot occupy the positions:

- (ii) of the only predicate of a simple sentence;
- (iii) of nominal attributes (i.e. participles);
- (iv) of a clausal actant, i.e. it cannot depend on verbs such as *begin*, *order* etc. (i.e. infinitives);
- (v) of a nominal actant (a gerund, i.e. a verbal noun).

These definitions of the converb collectively raise a number of issues that warrant further discussion. Firstly, Haspelmath (1995: 7–8) observes that some languages employ converbs in a non-adverbial (i.e. a non-presuppositional) manner for narrative chaining. In actual fact, *many* languages use converbs for this type of clause linkage. His notion of the converb category is therefore somewhat Eurocentric, considering that narrative chaining appears to be the most common function of converbs in languages of South Asia, Central Asia, East Asia and Ethiopia (Coupe 2005: 147).

Comparable chains of converb clauses terminating in a finite verb are syntactically impossible in European languages. This structural limitation is due to the fact that European converbs are typically constrained to encoding modifying, presuppositional notions in bi-clausal sentences of the [[dependent clause] matrix clause] type discussed in the preceding section, e.g. [[*Moving stealthily*, (= while moving stealthily)] *the soldiers took up positions along the ridge*]. The inclusion of the descriptor "adverbial" as the essence of the prototype is therefore difficult to reconcile with what appears to be the most frequently encountered use of converbs in the verb-final languages of Eurasia spoken beyond the borders of Europe.

Secondly, a number of authors (e.g. Ebert 1999, Koptjevskaja-Tamm 1994 and Lehmann [1988: 193–200]) have argued that finiteness is not strictly absolute, but an inherently gradual concept that varies according to the number of delimiting properties expressed by a verb form. Once again, outside of Europe it is not particularly uncommon for languages to have dependent, clause-linking verb forms that qualify as converbs according to most of the

proposed criteria, and for these to be additionally specified for person and number agreement. Some examples are Greenlandic Eskimo (Fortescue 1984), the Tungusic language Evenki (I. Nedjalkov 1995) and the Paleoasiatic isolate Yukaghir (I. Nedjalkov 1998). Haspelmath (1995: 7) acknowledges that there can be degrees of non-finiteness, but elects to retain the criterion in his definition to ensure that his conceptualization of the converb category accords with the traditional use of the term.

Thirdly, the notion “subordinate” is problematic in Haspelmath’s definition, not least because it is so inconsistently defined in the literature. This is largely attributable to the fact that subordination – much like the notion of the converb – is based upon no single unitary grammatical category (Haiman and Thompson 1984). An unfortunate consequence of this is that different authors entertain different interpretations as to what subordination entails, and by what collection of criteria it may be recognized cross-linguistically.

For Haspelmath (1995: 12), subordination is identified by: (a) disruptive embedding in a superordinate clause; (b) variable position before or after the superordinate clause; (c) the possibility of backwards pronominal anaphora; (d) semantic restrictiveness; and (e) the possibility of extraction. Haspelmath acknowledges in a footnote (1995: 48) that his particular criteria for subordination do not coincide with those proposed by Haiman and Thompson (1984). From this we may infer a justification for Haspelmath dismissing their criticisms of the concept. Foley and Van Valin (1984: 238–263) use yet another set of criteria to distinguish between “coordination”, “subordination”, and what they call “cosubordination” in their typology of nexus relations. According to their definition, a clause intrusively inserted between the constituents of another clause is not evidence of subordination; nor is it what they interpret to be “embedding”.¹ Despite his criteria also not coinciding with Foley and Van Valin’s criteria for subordination, Haspelmath paradoxically uses their terminology to claim that prototypical converb clauses are “subordinate”, and thus categorically distinct from prototypical medial verb clauses, which are “cosubordinate” (Haspelmath 1995: 23).

Genetti (2005: 62–65) is critical of Haspelmath’s criteria for subordination, noting that three of these (i.e. [a], [c] and [e] cited above) can only apply to languages of the Indo-European type that allow clausal coordination.² She finds that the syntactic tests proposed to establish subordination cannot be applied cross-linguistically, as some languages of Asia either lack clausal coordination completely, or else make extremely limited use of conjunctions in clausal syntax. This is particularly apparent in languages with rich converb systems,

1. Foley and Van Valin’s (1984) interpretation of embedding is addressed in §11.3.

2. Genetti also questions the validity of Haspelmath’s positing a semantic criterion, viz. semantic restrictiveness, to define the syntactic notion of subordination.

e.g. Turkic (Johanson 1995: 316), Evenki (I. Nedjalkov 1995: 100), other languages of the Altaic and Paleoasiatic families of eastern Siberia (I. Nedjalkov 1998: 340) and the Omotic language Wolaytta (Lambertini and Sottile 1997: 128), but it can also be true of languages with limited converb inventories, e.g. Dolakhā Newar (Genetti 2005). The notion of subordination therefore remains a nebulous criterion for defining the converb category.

In summary, a considerable number of questions are raised by the proposed definitions of the converb. Most of these stem from the fact that they rely upon a combination of syntactic, morphological and semantic criteria that do not necessarily coincide across all languages employing non-final verbal morphology for signalling clause linkage. Moreover, some of the very criteria that are used to define the converb are themselves poorly defined and variously interpreted in the literature. Despite these problems, we will see that the dependent, clause-linking verb forms of Mongsen do correlate to a high degree with patterns of clause linkage found in languages of the ‘Indo-Altaic area’ (Masica 1976: 137), and that these correlations favour their recognition as converbs. The formal characteristics of Mongsen converbs are explored in the following section.

11.3. The syntactic status of Mongsen converbs

The clausal constituents I identify as converbs are distinguished by their forms and functions. The suffix that marks a verb form as a converb will be referred to as a “converb suffix”. There is some inconsistency in the literature with respect to exactly what the term “converb” denotes. It is used in the present work to refer to a non-finite form of verb that is marked by one of the suffixes listed in Table 11.1 (see §11.4). There is only one exception to this – negation of a narrative chaining converb results in a verb form that has the same morphological structure as a negated past tense verb stem. For this type of dependent constituent, the non-final position of its occurrence serves to identify both its function and syntactic dependency in the absence of distinguishing non-finite morphology. Furthermore, it is semantically differentiated by the privative meaning obligatorily expressed by its negated non-final stem. Apart from its position of occurrence, the privative meaning in particular contrasts it with the various types of meanings otherwise encoded by negated matrix clause verb stems (see §11.4.2 for further discussion and examples of polarity contrasts). Lastly, it is also often accompanied by a distinctive non-final intonation pattern that prosodically marks its non-final status (see §3.4.2).

The formal structure of a converb clause can range from just a single verb form marked by its converb suffix, to a maximally specified clause that contains core and peripheral noun phrase arguments in addition to its converb. As noted

previously in §8.5.11, converb suffixes are able to co-occur with a considerable number of lexical, aspectual, and valency modifying suffixes. The one verbal category with which they can never occur is that of absolute tense (§8.5.12). Nor is a stem marked by a converb suffix able to co-occur with the declarative mood clitic (§4.2.14.6), because only finite forms may be independently specified for illocutionary force. These incompatibilities indicate that the converbs of Mongsen are indeed non-finite, a limitation which prohibits their being employed as predicates of simple verbal clauses.

The following sentence provides a preview of the clause linking functions of the sequential and simultaneous converbs, the two most commonly occurring members of the category. Converbs are henceforth bolded to facilitate their identification in example sentences.

- (11.1) *pi hənlakəɪ, puŋpuŋkə hləp.*
 pi **hən-lak-əɪ** **puŋ-puŋ-əkə** hləp
 PROX take-DESCEND-SEQ be.good-RED-SIM wrap.PST
 ‘Taking this down, [it] was wrapped very well.’

In this particular sentence, the simultaneous converb *puŋ-puŋ-əkə* qualifies the manner in which the noun phrase referent was wrapped, just as the adjunct *very well* does in the English translation. A reduplicated base is not obligatory, but is often used in simultaneous converbs to intensify the meaning encoded by a stative verb root, or to express a protracted duration if the verb root denotes an activity (see §8.1.4 for additional examples). Note that *puŋ-puŋ-əkə* is tightly integrated with its matrix verb *hləp*, to the extent that the simultaneous converb and its matrix verb represent a single event. The way it is used here is reminiscent of the manner in which an adverb derived by the *-ly* suffix is used to modify its verb in English. Because of this, the simultaneous converb and its matrix verb share a coreferential controlling argument, and the simultaneous converb clause also falls under the scope of any matrix clause operators, such as illocutionary force.

The sequential converb of (11.1) is semantically quite different, because it encodes a separate event that logically precedes the one encoded by the matrix clause. Another distinguishing characteristic of the sequential converb in this sentence is that its event shares narrative parity with the event of the matrix clause. Because of this, the sentence could alternatively be translated with a coordinating conjunction linking the two clauses, e.g. ‘[He] took this down and [he] wrapped [it] well’. However, unlike the conjoined clauses in the alternative English translation, the sequential converb clause of (11.1) cannot be independently specified for absolute tense. This prohibits it from standing alone as an independent clause and proves that it is not in a coordinate syntactic relationship with the sentence’s finite clause headed by the predicate *hləp*.

There is no requirement for an argument of the matrix verb to be coreferential with an argument of a converb if the predicates of the higher and lower clauses refer to distinct events. What determines coreferentiality or its absence are pragmatics and context, not syntax. Sequential converb clauses can be sensitive to semantically determined differences in the scope of their sentential operators (note that this also applies to simultaneous converb clauses). That is to say, a sequential converb clause may or may not fall under the scope of matrix clause illocutionary force or other operators, depending upon whether it is interpreted in an instance of use as encoding a presupposition *vis-à-vis* the matrix clause event, or non-presuppositional narrative information that serves to advance the plot sequentially. Thus, the semantic nature of the interpropositional relationship holding between the converb and its matrix verb is ultimately the principal determinant of a matrix clause operator's domain of application (see §11.4.2 below for related discussion and examples).

Neither the sequential nor the simultaneous converb can be used to modify a nominal in the way that a nominalized verb form can occupy the pre- or post-head attribute slots of the noun phrase, as discussed in §6.1; nor can they occupy the head slot of a noun phrase. The description of complementation in §11.7 will furthermore show that the grammar does not have predicates that subcategorize for infinitival object complements, therefore apart from this inapplicable exception, the converbs of (11.1) accord with all the criteria proposed for their recognition by V.P. Nedjalkov (1995).

With respect to Haspelmath's (1995) criteria, the example sentence of (11.1) discussed above does not allow us to determine if its converb clauses are intrusively embedded in the matrix clause, since the demonstrative pronoun *pi* may be an argument of either the sequential converb, or the matrix verb. It is often impossible to tell just which clause a noun phrase argument belongs to if clues cannot be garnered from case marking, verb valency, or pragmatics. Examples (11.2) and (11.3) below, however, clearly show that simultaneous and sequential converb clauses can be intrusively embedded between the constituents of their matrix clauses. The converb clauses of these sentences are enclosed by square brackets to help identify the intrusively embedded elements. According to Haspelmath's criteria, the converb clauses of (11.2) and (11.3) therefore qualify as subordinate constituents, by virtue of their intrusive embedding.

In the context of the narrative that the sentence of (11.2) is taken from, Tiger has decreed that all the animals should come to be eaten one per day, and the designated animal is late for Tiger's dinner appointment. Note that there is an implicit temporal relationship between the event of the matrix clause and the overlapping state of affairs represented by the simultaneous converb clause. Because the simultaneous converb encodes a restrictive temporal qualification in this sentence, it renders it impervious to the illocutionary force of the matrix

clause. This resistance is further enhanced by the S argument of the converb being non-coreferential with the S argument of the matrix verb, although as the example of (11.6) will show, non-coreferentiality is not obligatorily required in order for a dependent clause to be outside of a matrix clause's illocutionary force.

(11.2) *səpá? .jàjpà?, nitə hjiməkə, mətà.ɪ.*

səpá? .jà-ì-pà? [nì i-tə **hjim-əkə** mətə-ɪà-ət
 who come-IRR-NR 1SG PROX-thus be.hungry-SIM NEG-come-PRES
 'The one] who is to come [to be eaten], while I am hungry like this, is
 not coming' [said Tiger].

It is overwhelmingly more common for sequential converb clauses to precede their matrix clauses, due to a preference for the iconic ordering of chronologically prior events. Nevertheless, examples of sequential converbs intrusively embedded between the constituents of their matrix clauses do crop up occasionally in narrative texts.

(11.3) *pa tʃàla nə, hənlakəɪ khiukù.*

pa tʃà-la nə [hən-lak-əɪ] khìʔ-ukù
 3SG offspring-F ALL take-DESCEND-SEQ give-ANT
 'Taking [it] down, [he] gave [it] to his daughter.'

The application of some of Haspelmath's remaining criteria for the recognition of converbs, particularly those of subordination, is not as straightforward. Mongsen is representative of those languages that do not make use of clausal coordination, therefore not all of his tests for subordination are applicable. It is clear thus far, though, that the converbs of Mongsen are non-finite verb forms used to mark dependent clauses, that they can be intrusively embedded into their matrix clauses, and that they may or may not encode adverbial-like presuppositions.

We will now consider the syntactic status of Mongsen converbs with respect to subordination and cosubordination *à la* Foley and Van Valin (1984). In their discussion of clause linkage, Foley and Van Valin (1984: 238–263) use the binary values of [\pm embedded] and [\pm dependent] to set up a three-way contrast between their three nexus types. Coordination is [–embedded, –dependent], cosubordination is [–embedded, +dependent] and subordination is [+embedded, +dependent]. The difference between cosubordination and subordination rests crucially on their notion of embedding which, in contrast to Haspelmath (1995), does not necessarily entail the discontinuity of matrix clauses. According to Foley and Van Valin's model of nexus relations, all of the simultaneous and sequential converb clauses of (11.1) and (11.3) are dependent, but are not

embedded; they are therefore cosubordinate. A [+dependent] clause is only [+embedded] and thus subordinate if it acts as an island, impervious to the scope of matrix clause operators such as illocutionary force, tense and negation. But this leads us to question the syntactic status of the simultaneous converb clauses of (11.2) above and (11.4) below. Should we assume that in these particular sentences the converb clauses are both [+dependent] and [+embedded], i.e. subordinate, because the scope of their matrix clause operators – respectively, negation and illocutionary force – does not extend to their converb clauses? It is clearly the case that Tiger’s state of hunger is not negated by the matrix clause operator in (11.2), while the only event being questioned in (11.4) is that of the matrix clause (i.e. the trampling of Earthworm, not Earthworm’s presence at the time of the trampling). Note that both simultaneous converb clauses in (11.2) and (11.4) express what must be construed as temporal presuppositions in these particular utterances.

- (11.4) *tʃəpá? tʃhàpà? tsəlàla íjá? uŋəkə likə nàŋ nə tsəthàŋjaki?*
 tʃəpá? tʃhà-pà?³ tsəlà?-la íjá? puŋ-əkə li-əkə
 what do-NR earthworm-F much be.good-SIM be-SIM
 nàŋ nə tsə-thàŋ-jak-i?
 2SG AGT stamp.on-SEVER-RS-CAUS.PST
 ‘[Lichaba asked Barking Deer] “Why did you stamp on Earthworm while she was nicely lying [there on the surface]?”’

Further difficulties in the application of the subordinate/cosubordinate distinction are sometimes to be found with sequential converb clauses as well, because like the simultaneous converbs of the examples presented thus far, they are consistently neither “subordinate” nor “cosubordinate”. As we shall see, sometimes the same clause can be either, depending upon the context and its interpretation. While the most common function of the sequential converb is to report a sequential event that does not constitute a speaker’s presupposition (as seen above in [11.1] and [11.3]), occasionally *only* a presupposition can be interpreted contextually. Such a meaning is clearly suggested by the sequential converb of (11.5). The interpretation of temporal overlap encoded by *jim-ja-ai* is additionally constrained by the presence of the continuative aspect marker on the converb stem, and by the relativized locative oblique argument that temporally locates the overlapping events of the higher and lower clauses.

3. The combination of the interrogative pronoun *tʃəpá?* and the verb root *tʃhà* nominalized by the general nominalizer suffix is an alternative means of expressing ‘why?’ (see §4.2.4 for description of interrogative pronouns).

- (11.5) *áhlù jìmjáɿ, atʃak tʃhuwapaʔ hmapaŋ ku, pùŋì nə atʃak tʃu tʃàʔ.*
 a-hlú **jìmj-ja-əɿ**
 NRL-field cultivate-CONT-SEQ
 a-tʃak tʃhuwa-pàʔ hmapaŋ ku
 NRL-paddy emerge-NR time LOC
 pùŋì nə a-tʃak tʃu tʃàʔ
 wild.pig AGT NRL-paddy DIST consume.PST
 ‘When [he] was cultivating the field, at the time the paddy was
 sprouting, a wild pig ate the paddy.’

According to Bickel (1998: 394–395), this is the identifying characteristic of what he terms the “Asian converb”, which systematically conflates adverbial modification with a narrative chaining function. He proposes that the inclusion of modifying and narrative chaining functions in the same verb form(s) differentiates the Asian converb from Papuan *Satzinnenform* (‘sentence-internal forms’, i.e. medial verbs), which can only be used for a sequencing function, and from European converbs, which are consistent with Haspelmath’s (1995) definition. Although the sequential converb of Mongsen is less commonly used for expressing adverbial modification, text examples such as (11.5) nevertheless show that a temporal presupposition can also be encoded by this “Asian converb”.

Furthermore, in some cases it is not clear whether a presuppositional or non-presuppositional meaning is expressed by a sequential converb clause. This is suggested by the alternative translations of (11.6), also taken from a narrative text.

- (11.6) *nàŋ takzi tʃu ləpəɿ, íjáʔ a.uukə tsəthəʔ ku, tʃhàwkaŋ.*
 nàŋ takzi tʃu **ləp-əɿ**
 2SG bamboo.vessel DIST cut-SEQ
 íjáʔ **a.u-əkə** tsəthəʔ ku tʃhàwk-aŋ
 much be.good-SIM marsh LOC keep-IMP
 (i) ‘You make the bamboo cups and store them properly in the marsh.’
 (= non-presuppositional meaning)
 (ii) ‘When/After you make the bamboo cups, store them properly in the
 marsh.’ (= temporal presuppositional meaning)

It is possible to construe from the alternative free translations of (11.6) how matrix clause operator scope, and therefore the proposed subordinate/cosubordinate distinction, is determined by semantic interpretation, not syntax. If a non-presuppositional meaning is interpreted, then the sequential converb clause falls within the scope of the matrix clause imperative mood. Conversely,

if a presuppositional meaning is interpreted, then operator scope does not extend beyond the matrix clause. The point of this observation is that if native speakers allow either interpretation of the same sentence, then surely the linguist must question the value or necessity of drawing a distinction between two types of structural dependency that the speakers themselves do not acknowledge as relevant to the grammar of their language. The distinction may well be a useful one to make for types of dependent clauses in Papuan languages, but it appears to hold little significance for the description of Asian converb clauses. Such problems of application lead me to surmise that the subordinate/cosubordinate refinement does not enhance an analysis of dependent clauses in this language, since it cannot be used to identify the claimed structural differences in a descriptively useful way. Because of this, I adopt the premiss that all converb clauses of Mongsen are simply dependent. I will now briefly review the other structural characteristics of converbs before embarking upon a description of the semantic entailments of each type.

Sequential and simultaneous converb clauses in particular are frequently found to occur with a discourse focus particle (*t*áŋ) (§4.2.14.4); this does not appear to have any relevance to the identification of different types of clausal dependency in Mongsen. Discourse focus particles are overwhelmingly found to occur with narrative chaining converbs that do not express adverbial presuppositions, although it is also possible for them to restrict the meanings of specialized converbs as well.

- (11.7) “*thani muŋ.ɹáŋmən, áhlú mətʃhəniŋlá táŋ, jiplikə táŋ mùŋ*” *tə sa.*
 thani muŋ.ɹáŋmən a-hlu **mə-tʃhəniŋ-la** táŋ
 now+day whole NRL-field NEG do.work-NEG.NF just
jipli-kə táŋ mùŋ tə sa
 sleep-SIM just rest.PST thus say.PST
 “‘The whole day, [Dog] just slept and rested without working the field,’
 said [Pig].’

Context is essential for determining coreferentiality between core arguments of dependent and matrix clauses in discourse, due to the frequent ellipsis of noun phrases and the absence of a syntactic pivot. Often an unstated S/A argument of a converb clause is coreferential with an S/A argument of its matrix clause, as in (11.8) below. Yet this is not always the case, as demonstrated by the lack of coreferentiality between the A arguments of the dependent and matrix clauses of (11.9) and (11.10). These two sentences occur separated by a few clauses in their source narrative. The switch in actor arguments between the converb clause and the matrix clause in these examples is not formally encoded, therefore context must be relied upon to determine if arguments of lower and

higher clauses are coreferential or unrelated. Note that the anaphoric demonstrative in (11.9) at the end of the first line of interlinearized text is serving an anaphoric discourse function (see §4.2.7.1). It is used here to refer back to the earlier event expressed by the converb clause and is not functioning as a determiner normally would in a noun phrase. The fact that a nominal deictic can still occur juxtaposed to a sequential converb betrays the nominal diachronic origin of this converb suffix (see §11.4.1.1 for discussion of historical sources of converb suffixes).

- (11.8) *ípá? áwk i thijun nə nijukəi, mətʃatshəŋ nə áwk tʃu phà.ɿa?*
 ípá? a-úk i thijun nə **nɿ-uk-əɿ**
 EMPHAT NRL-pig PROX hulling.room ALL lead-INTO-SEQ
 mətʃatshəŋ nə a-úk tʃu phà.ɿa?
 PN AGT NRL-pig PROX catch.PST
 ‘And [Ø_i] having herded this particular pig into the paddy-hulling room of the house, Mechatseng_i caught it.’

- (11.9) *təɿ, pa áhlù jìməɿ sə, “aku tʃaŋ khə am tʃaŋ khə tə liŋaŋ” tə saw? tətù?*
tə-əɿ pa a-hlú **jìm-əɿ** sə
 thus-SEQ 3SG NRL-field cultivate-SEQ ANAPH
 aku? tʃaŋ khə am tʃaŋ khə tə liŋ-aŋ
 gour.d.sp seed CONJ aspidistra.sp. seed CONJ thus plant-IMP
 tə sa-ù? tətù-ù?
 thus say.PST-DEC REP-DEC
 ‘And then, after he_i cultivated the field, [Ø_j] said “Plant bottle gourd seeds.”’

- (11.10) *təɿ liŋəɿ, “áhlù nə watsəaŋ” tə tʃhàw? .ɿə.*
tə-əɿ **liŋ-əɿ** a-hlú nə wa-tsə-aŋ tə
 thus-SEQ plant-SEQ NRL-field ALL go-CON-IMP thus
 tʃhà-ù? tətù
 do.PST-DEC REP
 ‘And then, after [he_i] had done the planting, [Ø_j] said “Go to the field [to take a look].”’

Lastly, example (11.11) below shows that the elided S argument of the first converb clause is not coreferential with the A argument of the second converb clause. Additional examples demonstrating the absence of a syntactic pivot in complex sentences can be found in §5.2.3.

- (11.11) *təi kəwalikàʔ, tʃùku, mətʃəmtsə mà:rùk anət ilakəi,*
tə-əi **kəwa-likàʔ** tʃùku
 thus-SEQ ascend+go-CONTEMP consequently
 mətʃəmtsə mà:rùk a-nət **ilak-əi**
 rice.wine cup NRL-two pour-SEQ
 ‘And then when [Ø_i] went up, [Ø_j] consequently having poured two
 cups of rice wine, ...’

The sequential and locative converb suffixes frequently occur in grammaticalized collocations that encode a conjunctive function. These are used for tail-head and summary-head clause linkage, as demonstrated by *tə-əi* (thus-SEQ) in (11.11) above. These discourse connectives are described in detail in §4.2.10. In view of this, I will not include the conjunctive functions of these (often headless) converb constructions in the description that follows, but will continue to mark their collocations with bold face in examples. It is also possible for any of the specialized converbs (described in §11.4.3 below) to occur in headless collocations that encode a clause-linking conjunctive-like function.

11.4. Types of converbs

Mongsen has ten suffixes that mark converb categories. These were listed in Table 8.3 and are repeated below in Table 11.1, for convenience.

V.P. Nedjalkov (1995) makes a useful semantic distinction between narrative converbs, contextual converbs and specialized converbs. As he notes, these represent ideal cases, as many languages show overlapping membership of the three categories. We have already seen that the Mongsen sequential converb correlates with the narrative converb category in respect to its narrative chaining function, but also shows characteristics typical of contextual converbs, which can have a variety of contextually-determined semantic interpretations. Of the narrative chaining converbs, the simultaneous converb has a more specific meaning and is mostly used to express temporal overlap and manner of activity. However, sometimes it occurs in combination with sequential converbs in narrative chains, where it, too, must be interpreted as expressing a sequencing meaning (e.g. cf. [11.12c] of §11.4.1 below). It is also frequently used to refer to the manner in which the event of the matrix verb is performed. Under these circumstances the simultaneous converb and its matrix verb share a referent in A or S function, as was demonstrated by example (11.1) above.

The alternating iterative converb is used in a limited way for narrative chaining, occurring in a minimum of two dependent clauses in addition to the

matrix clause. It differs from the other narrative chaining converbs in consistently expressing a highly specific meaning that is never dependent upon context for its interpretation (see §11.4.1.3). The alternating iterative converb therefore demonstrates characteristics of both the narrative and specialized converb categories.

Table 11.1. Mongsen converb suffixes

FORM	GLOSS	SEMANTICS
-əɪ	(SEQ)	sequential activity, anterior event, sometimes temporal qualification or reason (§11.4.1.1)
-əkə; -liʔ; -a	(SIM)	simultaneous activity, attendant circumstance, sometimes sequential activity or reason (§11.4.1.2)
-li	(ALT.IT)	iterative alternation of two activities (§11.4.1.3)
-pàla	(COND)	irrealis conditional (§11.4.3.1)
-kùla	(CIRCM)	circumstantial/hypothetical and realis conditional (§11.4.3.2)
-pàkukàʔ	(CONCESS)	concessive conditional (§11.4.3.3)
-likàʔ	(CONTEMP)	actualized contemporaneous event (§11.4.3.4)
-tʃən	(DUR)	durative activity (§11.4.3.5)
-pànə	(CAUSAL)	consequential circumstance (§11.4.3.6)
-ku	(LOC.CV)	temporal sequence, simultaneous activity/state (§11.4.3.7)

V.P Nedjalkov's third category of specialized converbs is the most semantically specific; these consistently encode purely presuppositional meanings that equate with adverbial interpretations. The specialized converbs of Mongsen are used to qualify their matrix clause events with respect to condition, circumstance, concession, temporal setting, duration and reason. We will first look at the phenomenon of narrative chaining, before proceeding to an investigation of the semantic functions of the various types of converbs.

11.4.1. Narrative chaining

A conspicuous characteristic of Mongsen narrative discourse is the use of sequential and simultaneous converbs to link multiple sequences of clauses expressing both chronologically ordered and simultaneous activities, with the whole of the complex sentence terminating in a finite verb at the end of the matrix clause. Johanson (1995: 329) eloquently describes the finite verb at the

end of complex clauses in Turkic narrative chains as ‘a finite chain base ... the trunk by which the entire sentence is carried and through which it is rooted in the discourse “ground”’. An exemplary illustration of the multiple use of converbs for clause linkage occurs in Text 2, presented below in (11.12) for ease of discussion. This rather elaborate “sentence”, corresponding to what would be a whole paragraph in English, consists of eleven dependent sequential and simultaneous converb clauses within the one complex structure.

The primary function of the sequential converb is to sequence successive events, but this does not preclude it being used contextually to express other meanings. This is illustrated by (11.12a), in which *nì-əl wa-əl* is more suggestive of a manner of going, rather than the expression of two discrete sequences of events. My impression is that some frequently occurring pairings of converbs (or pairings of converbs with matrix verbs) are becoming conventionalized in the grammar, sometimes resulting in new meanings. This is observed, for example, in *hənəlà* ‘brought’, a meaning that has emerged from the grammaticalized coupling of *hənəlà* (< take-SEQ + come.PST) (see §8.4.3 for further discussion).

In preceding examples it was shown that the simultaneous converb principally describes the manner in which events are carried out, or the simultaneous occurrence of events, rather than their sequence. But sometimes there is contextual ambiguity in which either a manner or sequential interpretation is possible. An example of the simultaneous converb possibly being used to sequence an event occurs in (11.12c).

- (11.12) a. *tələ, tənì khə təza tənù tʃu, tùhɲəl ki nə nì.əl wələ nìuktʃakì.əl,*
tə-əl tə-nì khə tə-za tə-nù tʃu tùhɲəl ki
 thus-SEQ RL-wife CONJ RL-child RL-child DIST other house
 nə **nì-əl** **wə-əl** **nì-uk-tʃak-i?-əl**
 ALL lead-SEQ go-SEQ lead-INTO-RS-CAUS-SEQ
- b. *imtisàŋpà? tʃu, aji tʃu tatsəŋ à atʃət-lakəl,*
 imtisàŋ-pà? tʃu a-ji tʃu
 PN-M DIST NRL-rice.beer DIST
 tatsəŋ à **atʃət-lak-əl**
 bamboo.container one squeeze-DESCEND-SEQ
- c. *tʃəməkə atʃuŋ kà?, kusu ku ami mətʃuŋʃakəl,*
tʃəm-əkə a-tʃuŋ kà? kusu ku
 drink-SIM NRL-shield also fireplace.shelf LOC
 a-mi **mətʃuŋ-tʃak-əl**
 NRL-spear be.erection-RS-SEQ

d. *atʃuŋ kàʔ, anuk tʃu kàʔ h.ruŋsiɹə,*
 a-tʃuŋ kàʔ a-nuk tʃu kàʔ **h.ruŋ-si-əɹ**
 NRL-shield also NRL-machete DIST also hone-RPET-SEQ

e. *həmətəɹ aji tʃu tʃəməkə pa mənəɹ làʔ.*
həmət-əɹ a-ji tʃu **tʃəm-əkə**
 hold-SEQ NRL-rice.beer DIST drink-SIM

pa **mən-əɹ** làʔ
 3SG sit-SEQ wait.PST

‘And after leading his wife and children into a neighbour’s house and leaving them there, Imtisangba squeezed some rice beer into a cup, drank, propped the shield and the spear against the fireplace shelf, the shield also, honed the machete and held [it] drinking the rice beer as he sat and waited.’

Chains of dependent clauses such as these are by no measure unusual in Mongsen, although this is certainly one of the longest so far encountered. Narrative chaining provides the fundamental means of building discourse structure and serves to advance the discourse towards its communicative destination by chronologically ordering sequences of non-finite events.

11.4.1.1. *Sequential converb*

The sequential converb formed with *-əɹ* is the most commonly used converb in the language and occurs recurrently in texts, even recurrently within the one complex sentence, as amply demonstrated by (11.12) above. Sequential converb clauses correlate superficially with paratactic conjoined sentences in languages that have conjunctions; they provide cohesion and structure to discourse and are used to advance the plot. Because their main function is to iconically order sequences of events, sequential converb clauses commonly precede their matrix clauses. The chronology of the events they encode is therefore reflected in this linear arrangement.

The occasional ambiguity in the meanings of narrative chaining converbs noted in the preceding section is apparent in the sequential converbs of (11.13) below. An interpretation of sequential or simultaneous occurrence is equally possible for their events, although on this occasion, the presence of the continuative marker in the converb *atsə-ja-əɹ* perhaps encourages a simultaneous interpretation for the particular event it represents *vis-à-vis* the matrix clause event.

- (11.13) *atsəjaɪ, ahɪɪ, pa hmila tʃu ku pʊntʃhəŋ.*
atsə-ja-əɪ **ahɪə-əɪ** pa hmila tʃu ku pʊntʃhəŋ
 look-CONT-SEQ roar-SEQ 3SG reflection DIST LOC pounce.PST
 ‘Looking and roaring, [Tiger] pounced on her reflection [in the water].’

A verb stem marked by the sequential converb suffix is compatible with the majority of aspectual, modal and lexical verbal suffixes, but cannot be directly negated (see §11.4.2 for discussion). Negation tests demonstrate a dependency that affects both sequential and simultaneous converbs, resulting in a conflated form. Note also that in contrast to simultaneous converbs, a sequential converb juxtaposed to a matrix verb does not necessarily imply reference to a single event, although it can. This was demonstrated above by *nɪ-əɪ wa-əɪ* in (11.12a).

- (11.14) *akuŋ tʃu ɪtʃáʔ səptʃhətəɪ tʃhuwa.*
 a-kuŋ tʃu ɪtʃáʔ **səp-tʃhət-əɪ** tʃhuwa.
 NRL-prawn DIST some bail.out-ABIL-SEQ emerge-PST
 ‘[They] managed to bail out some prawns and returned.’

- (11.15) *áhŋáʔ sə hməntʃakəɪ áŋ tuŋnət atʃu ku jip.*
 a-hŋáʔ sə **hmən-tʃak-əɪ** tán
 NRL-fish ANAPH set-RS-SEQ just
 tuŋnət a-tʃu ku jipli
 3DU NRL-DIST LOC sleep.PST
 ‘He just set the [aforementioned] fish down and they [two] copulated there.’ (In context: *jipli* is euphemistic for ‘copulated’ – see Text 2).

It is very likely that the diachronic origin of the sequential converb is the agentive nominalizer suffix *-əɪ* (see §7.4.4, and also §8.5.12 and §7.3.3 for discussion of related functions). This suspicion rests not just on its identical segmental form, but also upon the analogous nominalizing and converb/dependent sequential clause marking functions of nominalizing morphology in other Tibeto-Burman languages, e.g. Dolakhā Newar (Genetti 1994), Chantyal (Noonan 1997) and Turung (Stephen Morey, personal communication 2006).

11.4.1.2. Simultaneous converb

The basic function of a stem marked by the simultaneous converb suffix is to express the manner in which an activity is performed, or the attendant

circumstances of an activity or state. Three forms of simultaneous converb suffixes are found in Mongsen: $-(ə)kə$, $-liʔ$ and $-a$. The initial schwa of the first form is deleted in the environment after a stem-final vowel, and optionally after nasals. It is somewhat disconcerting that I have not been able to recognize any differences in the meanings of these three suffixes or their phonological environments of realization. Substitution tests based on text examples demonstrate that they can all be exchanged in any context without any change of meaning whatsoever. I consequently gloss them all as SIM and assume that they represent the same formal category.

Additional forms of the one category are not unknown in Tibeto-Burman languages. A form originally restricted to a certain genre or dialect may expand its domain of application into those of other forms expressing the same or a similar meaning in other genres or dialects. This may explain why the $-əkə$ and $-liʔ$ forms are both freely used in the Mangmetong and Khensa village varieties, even within the same complex sentence. An example of this optionality can be seen in (11.55) below. The $-a$ form appears to be the preferred simultaneous converb suffix in the Waromung village variety, but is also sometimes encountered in the narratives of Mangmetong and Khensa Mongsen speakers. All three suffixes occur in the speech of urban speakers of Mongsen, which is probably attributable to homogenization.

It may well be the case that the $-a$ form used in the Waromung variety is originally from the Chungli dialect of Ao (cf. its correspondence to the Mongsen $-əkə$ form in Text 1). If so, its presence in the Waromung variety could result from bidialectalism, as both Chungli and Mongsen were formerly spoken in different wards of Waromung village (see Map 3 for the locations of other bidialectal Ao villages, and Coupe [2003a: 47–50] for evidence of dialect mixing).

The simultaneous converb is found in two very different clause types. In one type, it is tightly integrated with its matrix clause verb and together they refer to a single event. Because of this, they also share a referent in A or S function, and any matrix clause operators. This is apparent for the two examples of simultaneous converbs in (11.16) below. Observe that the sequence $mənukə .jàr$ in the direct speech complement functions as a single unit under the scope of the interrogative mood of that sentence. In this respect the juxtaposed predicates resemble a compound verb or an asymmetrical serial verb construction, in which the converb is the dependent constituent and the matrix verb is the head, specified for absolute tense and other markers of finiteness. The same observations apply to the sequence $tʃàsikə səmtsə$ at the end of that sentence. These simultaneous converb-matrix verb juxtapositions represent the most common use of the simultaneous converb in Mongsen. The juxtaposed constituent does not have to be a matrix verb, however; it may be another converb, as further examples of this section demonstrate.

- (11.16) *“nàŋ tʃápátʃhàli itə mənukə .jà.,” tə tʃàsikə səmtsə.*
 nàŋ tʃápátʃhàli i-tə **mənu-əkə** .jà-ə̀.
 2SG INTERR⁴ PROX-thus be.late-SIM come-PRES
 tə **tʃàsi-əkə** səmtsə
 thus be.distressed-SIM ask.PST
 ‘“What are you doing, coming late like this?”, [Tiger] irritably asked.’

Somewhat less commonly, the simultaneous converb clause expresses a presupposition that qualifies the assertion of its matrix clause event. In this usage, the event represented by the converb clause is distinct from that of the matrix clause, and usually the converb and its matrix clause verb have different actor arguments. To illustrate, the simultaneous converb of (11.17) is semantically indeterminate in suggesting either a reason for its matrix clause event transpiring (i.e. ‘Because of X, ...’), or a temporal qualification of the matrix clause assertion (i.e. ‘When X, ...’), or even one of sequence (X, and Y). In the context of its narrative, the causal interpretation seemed most appropriate (hence the interpretation of *nə* as the instrumental marker, rather than the agentive marker). Structurally and semantically, this usage is substantially different from the tightly integrated juxtaposition of the simultaneous converbs and their respective matrix verbs seen in the preceding example.

- (11.17) *h.ɪtsə? nə h.ɪəkə sə̀.*
 h.ɪtsə? nə **h.ɪ-əkə** sə̀-ə̀.
 insect.sp INST perforate-SIM die-PRES
 ‘[The tiger] dies from boring insects entering [its wound].’

While a sequential converb suffix may sometimes be able to occur in the place of a simultaneous converb suffix to express an equivalent meaning, the converse is not always possible.

- (11.18) a. *tətʃhə. tʃhuwaɪ aki thùŋ.*
tə-tʃhà-ə̀ **tʃhuwa-ə̀** a-ki thùŋ
 thus-do-SEQ emerge-SEQ NRL-house reach.PST
 ‘And then, [he] came out [of the jungle] and reached home.’
 b. **tə-tʃhà-ə̀ tʃhuwa-əkə a-ki thùŋ*
 emerge-SIM

4. This interrogative form is a collocation of *tʃápá?* ‘what’, *tʃhà* ‘do’ and an unrecognizable morpheme *li*.

In the complex sentence of (11.18) above, for example, *tʃhuwaɪ* represents a distinctly punctual and separate event that can only be encoded as such by the anterior meaning of the sequential converb. This suggests that the selection of converb markers may be sensitive to the specific *Aktionsart* of their verb stems.

The simultaneous converb suffix is used to mark the existential copula *li* ‘be’ to help create the functional equivalent of an imperfective aspect. The existential copula behaves like an auxiliary verb to a preceding verb stem in the dependent clause, while the preceding verb stem expresses the semantic content of the simultaneous state or activity and is marked by the continuative aspect suffix. This type of simultaneous converb construction expresses a temporally unbounded overlap with the event of its matrix clause. The existential copula *li* ‘be, stay’ does not otherwise express imperfective aspect, but the meaning of continuity that is inherent to existence makes it a target for this kind of extended function. We find a similar usage of *li* as the matrix verb of one type of imperative clause (see §10.3.1).

- (11.19) *ùphəla həntsə mukja? likə,*
 ùphəla həntsə muk-ja **li-əkə**
 red.junglefowl egg brood-CONT be-SIM
 ‘While Red Junglefowl was hatching her eggs, ...’

Alternatively, both the verb stem expressing the overlapping state/event and the existential verb take simultaneous converb suffixes. In (11.20), the existential copula *li* is marked by the contemporative converb suffix and expresses the relative tense of a realized event, while the preceding verb stem encodes the semantic content, continuative aspect, and contributes to the meaning of overlapping occurrence.

- (11.20) *mətshəla tʃu saɪət khəmpaŋ məzujakə lilikà?*
 mətshə-la tʃu saɪət.khəmpaŋ **məzu-ja-əkə**
 barking.deer-F DIST ivory.armlet clean-CONT-SIM
li-likà?
 be-CONTEMP
 ‘While Barking Deer was cleaning her ivory armllets ...’

11.4.1.3. *Alternating iterative converb*

Converbs formed with the alternating iterative suffix *-li* minimally occur in pairs consisting of different verb roots. They express two (rarely, three) dependent sequential activities or states that alternate, *ad infinitum*. It is not

necessary for the actor arguments of the alternated events to be coreferential, although they often are. Speakers often repeat both clauses of the alternating iterative construction once or twice. This repetition iconically emphasizes the iterativity that is central to this converb's meaning.

- (11.21) *ní? kà? à .jali, pa tʃali, ní? kà? à .jali, pa tʃali, tɔɪ, sà.ɪ.ɪ sə kúlànà tʃakə .jà.*
ní? kà? à .jà-li pa tʃà?-li
 one.day also one come-ALT.IT 3SG consume-ALT.IT
ní? kà? à .jà-li pa tʃà?-li
 one.day also one come-ALT.IT 3SG consume-ALT.IT
tə-əɪ sà.ɪ.ɪ sə kúlànà tʃà?-əkə .jà
 thus-SEQ animal ANAPH many consume-SIM come.PST⁵
 'Each day an animal came and he [i.e. Tiger] ate it, and then the next day another animal came and he ate it, and so on, over and over, and thus many [of the aforementioned] animals came to be eaten.'

- (11.22) *am à nə kà? hləpli, aliŋ nə .jàkli, am à nə hləpli, aliŋ nə .jàkli,*
am à nə kà? hləp-li
 aspidistra.sp one INST also wrap-ALT.IT
a-liŋ nə .jàk-li
 NRL-cane.strip INST bind-ALT.IT
am à nə kà? hləp-li
 aspidistra.sp one INST also wrap-ALT.IT
aliŋ nə .jàk-li
 NRL-cane.strip INST bind-ALT.IT
 '[He] alternately wrapped [the meat] with an aspidistra leaf and tied it with a cane strip, then wrapped it with another aspidistra leaf and tied it with another cane strip and so on, over and over, ...'

The alternating iterative converb clauses of (11.23) below require an explanation of the context in which they were uttered. The protagonist of the narrative has belatedly realized that he miscalculated the number of gifts of meat he needs to distribute to his relatives. He has nine pieces of meat, but in actual fact he needs gifts for his whole clan, not just his immediate family. He attempts to solve his dilemma by cutting the meat into smaller pieces. But the meat is magical, because it came from a wild pig belonging to the god Lichaba. Every time he cuts the nine pieces into a sufficient number of smaller pieces, it

5. The use of *.ja* 'came' in this context suggests a developing perfective meaning for this matrix verb.

linked to subsequent discourse by a pair of plot-advancing sequential converbs that refer to a single event.

- (11.24) *anukà? wali, “pi mətət sə məmətətla” təli, “mətətə.ù?” tə, niɪ waɪ,*
anukà? wa-li pi mətət sə mə-mətət-la
 again go-ALT.IT PROX know QPTCL NEG-know-NEG.PST
tə-li mətət-əɪ-ù? tə ni-əɪ wa-əɪ
 thus-ALT.IT know-PRES-DEC thus lead-SEQ go-SEQ
 ‘Again they went [deeper into the jungle and each time he asked her]
 “Do you know this place?”, [and each time she answered] “I know
 it”, over and over, and so [he] led [her] further, ...’

Haiman (1980: 139) and (1985: 80-81) describe a remarkably similar construction in Hua [Papuan: Eastern Highlands Province, PNG] that encodes a meaning very close to the alternating iterative converb construction of Mongsen and demonstrates an analogous structural symmetry.

The alternating iterative converb suffix has an identical morphological form and tone to the existential verb *li*. As we shall see in the description of specialized converbs in §11.4.3 below, the existential verb appears to be a common lexical target for the grammaticalization of dependent clause-marking morphology. This would be an unusual diachronic source though, because most converbs of Mongsen have demonstrable nominal origins (see §11.4.4). That notwithstanding, the alternating iterative is an extremely unusual converb.

11.4.2. Negation and narrative chaining converbs

It is not possible to directly negate a verb stem marked by the sequential converb or the simultaneous converb. Language consultants also balk at attempts to elicit an alternating iterative converb with a negated stem. On *a priori* grounds, negation seems to be incompatible with the semantics of narrative chaining converbs, because the types of non-presuppositional information they encode potentially render them conceptually incongruous for expressing non-finite sequences of events that are not actualized. Recall that the continuative aspect marker similarly cannot occur on negated stems, probably for related reasons of semantic incongruity (see §8.5.4.1 for discussion).

Direct negation with the prefix *mə-* in the elicited example of (11.25) results in an ungrammatical converb form. Thus, the only way to negate a sequential converb is to replace its marker with both negative affixes used to negate past tense matrix verb stems, viz. *mə...la*. An important prosodic modification often accompanies this change in morphology; the tone on the negative suffix is

trumped by a postlexical high intonation to signal the non-final status of the negated verb stem (see §3.4 for description of postlexical intonation). The non-final position of occurrence, the non-final intonation and the meaning expressed by the negated verb stem in (11.26) all suggest that this cannot be a finite verb stem. This is in spite of the presence of the negative suffix, which otherwise only occurs on negated finite verb stems in past tense (see §8.3.3 and §8.5.9 for description of negating verbal morphology).

(11.25) *atʃa mətʃuŋ-əɪ hjim-əɪ
 NRL-rice NEG-eat.meal-SEQ be.hungry-PRES

(11.26) *atʃa mətʃuŋla hjiməɪ atì.*
 a-tʃa **mətʃuŋ-la** hjim-əɪ atì
 NRL-rice NEG-eat.meal-NEG.NF be.hungry-PRES PTCL
 ‘Without having eaten, [I’m] really hungry.’

Note that the meaning of the negated verb form in (11.26) has now become presuppositional, in that it implies a cause for the effect expressed by its matrix verb. Admittedly, this is perhaps abetted by the semantics of the matrix verb in this example. But that aside, it does not seem possible to negate a sequential converb referring to an event separate to that of its matrix clause while still retaining the non-presuppositional, sequential activity connotation of the positive polarity equivalent. This might be explained by the above-mentioned semantic incongruity of sequencing non-finite events that do not occur.

The elicited sentences of (11.27) below show that the same morphology is used for negating a simultaneous converb when it is tightly integrated with its matrix verb to refer to a single event. Their identical morphological forms suggest that the distinction between the sequential and simultaneous converbs is neutralized under negation. In this case, however, the meaning expressed by the negated non-final verb stem remains non-presuppositional, even though under the influence of negation it expresses a privative connotation. This is also suggested by the meaning of the negated non-final stem in the following textual example.

(11.27) a. *pa səməkə ɪà.*
 pa **səm-əkə** ɪà
 3SG run-SIM come-PST
 ‘He came running.’

b. *pa mət-səm-əkə ɪà
 NEG-run-SIM

- c. *pa məsəm̩la .l̩à.*
 pa **mə-səm-la** .l̩à
 3SG NEG-run-NEG.NF come.PST
 ‘He came without running.’

- (11.28) *inət səŋkhəm à táŋ tʃh̩a.ə̀ inət məh̩iasàl̩əŋ l̩ə̀jù?*
 inət səŋkhəm à táŋ **tʃh̩a-ə̀**
 1DU.IN village.gate one just make-SEQ
 inət **mə-h̩iasà-la** áŋ li-ì-ù?
 1DU.IN not-divide-NEG.CV just stay-IRR-DEC
 ‘We two [clans] will make just one village gate and just live undivided.’ (Literally ‘... live without dividing’)

The spontaneously uttered example of (11.29) below was naturally produced in response to a video clip. Here we see the negated converb in a similar function to the negated non-final stems of (11.27) and (11.28); the negated converb describes a privative manner of activity, while the activity of sitting is encoded by a sequential converb. Both juxtaposed constituents jointly refer to the same event, and thus are tightly integrated and compound-like. This example was not pronounced with a high intonation realized on the negative suffix, possibly because of the presence of the high-toned restrictive focus particle.

- (11.29) *túnkhəla luŋ i məkə̀ptə̀nl̩à táŋ məntə̀pə̀ɪ liaw?*
 túnkhəla luŋ i **mə-kə̀p-tə̀n-la** táŋ
 3PL group PROX NEG-move-TO.AND.FRO-NEG.NF just
 mən-tə̀p-ə̀ɪ li-a-ù?
 sit-COLL-SEQ stay-VSF-DEC
 ‘They are just sitting together without moving.’

Using this sentence as a template, the following sentence of positive polarity was elicited. Predicably, the negating morphology is replaced by the simultaneous converb suffix. This is because the event represented by the converb expresses a manner of activity and refers to the same event as its juxtaposed sequential converb.

(11.30) *túnkhəla luŋ i kəptənəkə táŋ məntəpəi liaw?*

túnkhəla	luŋ	i	kəp-tən-əkə	táŋ
3PL	group	PROX	move-TO.AND.FRO-SIM	JUST

mən-təp-əi	li-a-ù?
sit-COLL-SEQ	stay-VSF-DEC

‘They are sitting together and just wriggling about.’

(Literally: ‘They are just wrigglingly sitting together’.)

The influence of matrix clause negation over narrative chaining converb clauses is potentially only determinable with reference to pragmatics and/or semantics, with the result that the scope of an operator may have more than one possible interpretation. Three native speakers were individually asked to comment on the possible meanings of the elicited sentence of (11.31). A context that might have assisted the consultants to determine the scope of negation was purposely not provided. It is significant that both (i) and (ii) were given as possible meanings by all three speakers. This is attributable to the fact that Mongsen uses narrative chaining converbs for encoding both presuppositional and non-presuppositional meanings, therefore the scope of a negative operator can only be determined by taking contextual semantic and pragmatic factors into consideration.

(11.31) *pa aki nə waɪ, mətʃàlà.*

pa	a-ki	nə	wa-əi	mə-tʃà?-la
3SG	NRL-house	ALL	go-SEQ	NEG-consume-NEG.PST

(i) ‘After going to the house, he didn’t eat.’

(= presuppositional, temporal interpretation)

(ii) ‘He didn’t go to the house and eat.’

(= non-presuppositional, sequential interpretation)

A similar test based on elicited data was done, this time using a sentence with a simultaneous converb clause and a negated matrix verb. This revealed even more possibilities for the scope of negation. Surprisingly, option (iii) of (11.32) allows for the interpretation that negation does not have scope over the matrix verb, even though this is where negation is formally marked.

(11.32) *pa aki nə səməkə məzala.*

pa	a-ki	nə	səm-əkə	mə-za-la
3SG	NRL-house	ALL	run-SIM	NEG-enter-NEG.PST

(i) ‘He ran and did not enter the house.’

(ii) ‘He did not run into the house.’

(iii) ‘He entered the house not running.’

11.4.3. Specialized converbs

The specialized converbs of Mongsen express specific meanings that correlate with adverbial presuppositions, such as condition, circumstance, concession, cause and temporal modification. Because of the highly specific nature of their propositional content, they mostly occur in complex sentences of just two clausal constituents, although it is also possible for them to combine with several plot-advancing, narrative chaining converbs and other types of clausal constituents in complex sentences, such as relative clauses, nominalizations and complements. The specialized converbs of Mongsen closely accord with the definition of the converb as proposed by Haspelmath (1995), precisely because they can only be used to express presuppositional qualifications of their matrix clause assertions.

11.4.3.1. Conditional converb

A verb stem marked by the conditional converb suffix *-pàla* expresses a predictive irrealis proposition.

- (11.33) *nàŋ .jàpàla inət tʃàjù?*
 nàŋ **.jà-pàla** inət tʃàʔ-i-ùʔ
 2SG come-COND 1DU.IN consume-IRR-DEC
 ‘If you come, we’ll eat.’

Unlike the narrative chaining converbs discussed in §11.4.1, a verb stem marked by the conditional converb suffix can be directly negated. I attribute this to the very different nature of the interpropositional relationship holding between a specialized converb and its matrix predicate. This licenses the possibility of a presupposition having negative polarity.

- (11.34) *nàŋ atʃu mələptʃukpàla aki .uŋiʔ?*
 nàŋ a-tʃu **mə-ləp-tʃuk-pàla** a-ki .uŋ-i-ùʔ
 2SG NRL-DIST NEG-cut-PFV-COND NRL-house burn-IRR-DEC
 ‘If you don’t cut out that [knot of wood], the house will burn.’

The conditional converb suffix is similar to the discourse connectives (§4.2.10) and other specialized converb suffixes in being able to occur without a head, attached to just the quotative particle. The grammaticalized collocation of quotative particle + conditional converb suffix has a conjunctive function in discourse and introduces a new clause. It indicates that a summary of what has preceded in the discourse has been taken into account, and that this provides a

frame of reference for subsequent questions or statements. Such collocations may be headed or headless.

- (11.35) *təpàla kúták.ɾə nə kútə tʃhà.ɾə.*
tə-pàla kúták-əɪ nə kútə tʃhà-əɪ
 thus-COND heaven-ANOM AGT how do-PRES
 ‘If that is the case, what are the heavenly beings doing?’⁶

The morphological form of the conditional converb suffix suggests that its diachronic sources are the nominalizer suffix *-pàʔ* and the topic particle *la*, which have subsequently become fused and reanalyzed as a single morpheme. If these are indeed its diachronic sources, then the conditional converb of Mongsen aligns with a well-attested relationship noted between conditionals and topics in the languages of the world (Haiman 1978). I attribute this to the fact that both topics and presuppositions represent “given” factual information. This is also why presuppositional converb clauses cannot be questioned by their matrix clause interrogative operators – it is conceptually illogical to question one’s own presuppositions.

Finally, one other combination of morphemes can be used to encode a conditional meaning, viz. *-əɪla*, which is formed with what is undoubtedly the agentive nominalizer *-əɪ* combining with the topic particle *la*. Its morphological makeup is thus parallel to the more commonly used conditional suffix *-pàla*, as discussed above. Example (11.36) has been repeated from (8.146b) and renumbered, as it is the only one of its kind in the corpus that demonstrates the use of the agentive nominalizer and the topic particle in this conditional marking function. It was uttered by a sixty-year old male speaker in the context of a folklore narrative and has been deemed acceptable by other native speakers, therefore I have no reason to suspect its grammaticality, notwithstanding its textual rarity.

- (11.36) *akhu nə “wa səmtʃà.ɪ, nəŋ táŋ nə muluŋəɪla tàk.ɹùk.ɹù kəwa.”*
 a-khu nə wa səmtʃà-əɪ nəŋ táŋ nə
 NRL-tiger AGT EXCLM grandchild-ANOM 2SG just AGT
 muluŋ-əɪ-la tàk-ɹùk-ɹù kəwa
 be.willing-ANOM-TOP weave-ASSIST-IMM PTCL
 ‘Tiger replied “Ah, grandchild, just weave with me if you are willing, alright?”’

6. Despite the unexpected choice of interrogative, in context *kútə* appears to be expressing “what?”, not “how?” (see Text 2, line 60).

The fact that a conditional meaning can be conveyed by a nominalized verb stem marked by the topic particle offers quite compelling evidence that the conditional converb marker originates from the reanalysis of a topic-marked noun phrase modified by a post-head relative clause. The more common functions of the agentive nominalizer are described in §7.4.4 (also see §8.5.12 for description of other extended functions of this morpheme).

11.4.3.2. Circumstantial converb

The circumstantial converb suffix *-kùla* is used to encode hypothetical and counterfactual conditional meanings. Some overlap is found in the meanings encoded by the conditional and circumstantial converb suffixes. Whereas converbs formed with *-pàla* are restricted to predictive presuppositions, those formed with *-kùla* can express counterfactual and hypothetical imaginary meanings in addition to expressing irrealis conditional presuppositions. The conditional converb suffix *-pàla*, for example, could alternatively be used on the converb stem of the dependent clause in the following elicited sentence without resulting in a change of meaning.

- (11.37) *ikhu tʃu atsə məkhikùla hna:u tʃu məpuŋjù?*
 ikhu tʃu a-tsə **mə-khì?-kùla**
 garden DIST NRL-water NEG-give-CIRCM
 hna:u tʃu mə-puŋ-ì-ù?
 flower DIST NEG-bloom-IRR-DEC
 ‘If [you] don’t water the garden, the flowers won’t bloom.’

For the expression of hypothetical and counterfactual conditional presuppositions, however, only the circumstantial converb suffix may be used. This additionally requires the presence of a copula functioning as an irrealis marker to express the hypothetical nature of the proposition.

- (11.38) *nàŋ tʃha:ukə tʃhàkùla nì nə tuktəi tʃajù?*
 nàŋ **tʃha:u-əkə** **tʃhà-kùla**
 2SG be.sick-SIM HAPPEN.COP-CIRCM
 nì nə tuktəi tʃa-ì-ù?
 1SG AGT doctor call-IRR-DEC
 ‘If you were sick, I’d call the doctor.’

It is not clear if *tʃhàkùla* in the elicited example of (11.38) above forms a grammatical word with the preceding simultaneous converb suffix and its stem,

or if it has the status of an independent phonological word. There are no other examples in the corpus of a simultaneous converb occurring anywhere except stem-finally, so it has been segmented accordingly. The copula is analyzed as a type of auxiliary head that carries the converb marker and expresses the hypothetical counterfactual meaning of this type of converb.

A causal realis versus a counterfactual hypothetical irrealis contrast can be made by using copulas marked by the circumstantial converb suffix. Compare the following two examples: the sentence of (11.39) is taken from a text, and that of (11.40) is elicited, based on the textual example. The output high tone is attributable to the non-final intonation discussed in §3.4. These examples are also noteworthy for the fact that although there is no marker of dependency to be seen on the negated verb stems preceding the auxiliary copulas, both must be considered to be constituents of their respective converb clauses. One possibility is that the negated stems in fact form single phonological words with their converb-marked copulas.

- (11.39) *tətʃhaɪ, mətʃhuwatsəlikùlá,*
tə-tʃhà-əɪ mə-tʃhuwa-tsə li-kùla
 thus-do-SEQ NEG-emerge-CON BE.COP-CIRCM
 ‘And then, since it was the case that she wasn’t trying to return, ...’

- (11.40) *tətʃhaɪ, mətʃhuwatsətʃhàkùlá,*
tə-tʃhà-əɪ mə-tʃhuwa-tsə tʃhà-kùla
 thus-do-SEQ NEG-emerge-CON HAPPEN.COP-CIRCM
 ‘And then, if it were the case that she wasn’t trying to return, ...’

Analogously to the conditional converb (§11.4.3.1) and sequential converb suffixes (§11.4.1.1), a circumstantial converb occurs with the quotative particle and is used in a conjunctive function to introduce subsequent clauses in discourse. And like the collocation formed with the conditional and sequential converb suffixes, this grammaticalized collocation can also be headless. It similarly expresses a summary of what has preceded in the discourse, but implies a subtle difference in meaning – it seems to express greater confidence in the truth value of the summation. I have no examples in the data of just *tə-kùla* occurring without the concomitant presence of the auxiliary ‘BE’ copula *li* as the head, so it is not known at present if this is an obligatory requirement for its use.

- (11.41) *təlikùla mətshəla tʃu sa.ɪət.khəmpan məzujakə lilikà?*
tə-li-kùla mətshə-la tʃu sa.ɪət.khəmpan
 thus-BE.COP-CIRCM deer.sp.-F DIST ivory.armlets
 məzu-**ja-əkə** **li-likà?**
 clean-CONT-SIM be-CONTEMP
 ‘That being the case, while Barking Deer was cleaning her ivory
 armband ornaments, ...’

The irrealis collocation *tə-pàla* (thus-COND) discussed in §11.4.3.1 expresses what may be true, in keeping with its irrealis meaning, while the circumstantial *tə-li-kùla* (thus-BE.COP-CIRCM) implies that what has preceded in the discourse is known to be true. The two types of constructions therefore contrast in terms of epistemic meaning. The difference in meaning between *-pàla* and *-kùla* in discourse connectives seems to correspond to what Thompson and Longacre (1985: 190) refer to as unreality and reality conditionals respectively.

The diachronic sources of the circumstantial converb are suspected to be the locative marker *ku* and the topic particle *la*, which have subsequently become fused and reanalyzed as a morpheme representing a single grammatical category.

11.4.3.3. Concessive converb

The concessive converb suffix *-pàkukà?* expresses a concession towards a proposition asserted by the matrix clause.

- (11.42) *akhu tʃu nùkhu kətpàkukà?*
 a-khu tʃu nùkhu **kət-pàkukà?**
 NRL-tiger DIST wound have-CONCESS
 ‘Even if the tiger has a wound, ...’

It can be used to link successive clauses, just like the other converbs that have developed a conjunctive function as discourse connectives. The morphological form of the concessive can be subject to phonological reduction, particularly in rapid speech, as in (11.43). This example shows its headless use with the quotative particle *tə*, which summarizes the preceding events of the narrative. The protagonist of this story has just chopped out all the knots of wood in the timber of his house in an attempt to stop them attracting lightning strikes (see Text 2 for the full context). The discourse connective expresses this concession in its summarized form and relates it to the assertion of the matrix clause.

Example (11.44) similarly demonstrates the summarizing and clause-linking function of the headless collocation.

- (11.43) *təkukà? tsəŋtəphəla? tʃu məhì.*
tə-pəkukà? tsəŋtəphəla? tʃu mə-hì?-əɪ
 thus-CONCESS lightning DIST NEG-pass-PRES
 ‘But even then [i.e. even after doing all of the aforementioned,] the lightning doesn’t pass [the house].’

- (11.44) *təpəkukà? tsəŋrutsəla nə məsəɪ tāj,*
tə-pəkukà? tsəŋrutsə-la nə məsəɪ-əɪ tə aj
 thus-CONCESS PN-F AGT demand-SEQ thus SINCE
 ‘But even then, since Tsengrutsela had demanded it, ...’

Although there are no examples of negated concessive converbs in the corpus, we may infer on the basis of analogy to the conditional and circumstantial converbs, and by the semantic nature of the proposition expressed, that there would be no constraints on negating the stem of a concessive converb.

Lastly, the morphological form of the concessive marker is suspiciously similar to the nominalizer suffix *-pà?*, the locative marker *ku* and the final syllable of the adverb *anukà?* ‘again, also, even, yet’, which also occurs in the monosyllabic form *kà?* as a phonological word. The ‘even’ sense of the adverb is consistent with the concessive meaning of the grammaticalized converb. These three morphemes appear to have become fused, resulting in their reanalysis as a single morpheme. If so, this must represent a relatively recent grammatical-ization, in view of the ponderous size of this converb suffix. Cross-linguistically, grammatical affixes tend to be no more than a syllable in length.

11.4.3.4. *Contemporative converb*

The contemporative converb suffix *-likà?* is a marker of past tense that is relative to the tense of the matrix verb. It cannot be used for encoding unrealized events. A potentially headless collocation formed with the quotative complementizer *tə* functions as a discourse connective.

- (11.45) *akhula tʃu tshəŋti tʃu takjalikà?*
 a-khu-la tʃu tshəŋti tʃu **tak-ja-likà?**
 NRL-tiger-F DIST bamboo.matting DIST weave-CONT-CONTEMP
 ‘When Tiger was weaving the bamboo wall, ...’
- (11.46) *təɪ hɪəntən walikà?, akunla sə luj mə ku sakti.ə athà.*
tə-əɪ hɪəntən **wa-likà?** a-kun-la sə
 thus-SEQ sacrifice go-CONTEMP NRL-prawn-F ANAPH
 luj ma ku **sakti-əɪ** athà
 stone FACE LOC slip-SEQ fall.down.PST
 ‘And, when [they] went [to do] the sacrifice, [aforementioned] Prawn
 slipped on a stone and fell down.’

The contemporative converb suffix has a morphological form suggestive of a fusion of the existential verb *li* ‘be’ with a phonologically reduced form of the adverb *anukà?* ‘again, also, even, yet’ that I also recognize in the concessive converb suffix of the preceding section. The ‘also’ sense of this adverb is consistent with the relative tense meaning of contemporative converb. No examples of a negated stem marked by the contemporative converb occur in the data, but this is not surprising in view of the semantic incongruity of expressing that an event also didn’t happen at the same time that the matrix clause event happened.

11.4.3.5. Durative converb

The durative converb, formed with *-tʃən*, expresses a perpetual activity or state. In common with the frustrative modality of §8.5.5.2, it occurs in a productive reduplicating V_1 -DUR *táj* V_1 construction that may reinforce the meaning of continuity of an activity or state, e.g. *wa-tʃən táj wa-əɪ* (go-DUR just go-PRES) ‘just keep on going’, *jipli-tʃən táj jipli-əɪ* (sleep-DUR just sleep-PRES) ‘just keep on sleeping’. However, the following example of (11.47) indicates that the expression of continuity is not uniquely dependent upon this construction type, therefore it must be a meaning inherent to this converb form. A noteworthy peculiarity presented by this sentence is that in this specific context, *-likà?* expresses a conditional meaning, rather than the usual contemporative meaning consistently encountered in all other texts (cf. its use in §11.4.3.4). A consultant allowed me to change the contemporative converb suffix to the irrealis conditional converb suffix *-pàla* (§11.4.3.1) without resulting ungrammaticality or a change in meaning. As I am unable to account for this at present, it remains a topic for further investigation.

- (11.47) “*məŋlikà? məŋtʃən məŋəŋ, saŋəlikà? saŋwatʃən saŋəŋ,*” *tə sa pi.*
 məŋ-likà? **məŋ-tʃən** məŋ-əŋ saŋwa-likà?
 be.dark-CONTEMP be.dark-DUR be.dark-IMP be.light-CONTEMP
saŋwa-tʃən saŋwa-əŋ tə sa pi
 be.light-DUR light-IMP thus say.PST PROX
 ‘‘If it is dark, may it stay perpetually dark; if it is light, may it stay
 perpetually light’’, [Owl] said this.’

The durative converb is very rare indeed in narratives or spontaneous language. In over thirty texts, this is the only example encountered. Despite its rarity, a young consultant didn’t hesitate to offer the following example when he was asked to explain the use of *-tʃən* as a converb suffix in the textual example provided above.

- (11.48) *pə tʃəsitʃən tʃəsikə táŋ li.ù?*
 pə **tʃəsi-tʃən** **tʃəsi-kə** táŋ li-ə-I-ù
 3SG be.distressed-DUR be.distressed-SIM just stay-PRES-DEC
 ‘He is continually irritated.’

The durative converb suffix has an identical phonological form to the locative nominalization *-tʃən*, apart from its tone. There is a strong tendency for nominalizing morphology to grammaticalize as converb markers throughout the Tibeto-Burman family, therefore we can quite confidently assume that these two synchronically separate morphemes share a common diachronic source.

11.4.3.6. *Causal converb*

A verb stem marked by the causal converb suffix *-panə* expresses a reason for the proposition expressed by the matrix clause event. The causal converb suffix is formed from a concatenation of the nominalizer suffix *-pà?* and the instrumental case marker *nə*. The nominalizer attaches to the verb stem and is followed by the instrumental marker. The loss of the glottal stop from the nominalizer suffix before the morpheme boundary created by the instrumental marker suggests that the concatenated affixes form one phonological word with the verb root (see §2.4 for criteria to identify the word).

- (11.49) *pa tələm ku tsəpànə, tələm tʃu apaktʃuk.*
 pa tə-ləm ku **tsə-pànə** tə-ləm tʃu apak-tʃuk
 3SG RL-head LOC peck-CAUSAL RL-head DIST flat-PFV.PST
 ‘Because [the other birds] pecked on her head, [her] head became flat.’
- (11.50) *təi pa nə “ù?” tə aɛmpànə pa təpuk təkhuŋlən nə kəɛajukù?*
tə-əɪ pa nə ù? tə **aɛm-pànə**
 thus-SEQ 3SG AGT ONOM thus bear-CAUSAL
 pa tə-puk tə-khuŋ-lən nə kəɛa-juk-ù?
 3SG RL-stomach RL-neck-?⁷ ALL come+ascend-PFV.PST-DEC
 And because “uh!” she bore it (i.e, held her laughter in), her stomach rose up to her throat.’

Interestingly, a nominal deictic may also be inserted between the nominalizer and the instrumental marker. This is synonymous to a nominal demonstrative ostensibly occurring within the ablative marker (see §5.3.2.2), and it offers some valuable evidence for determining the nominal sources of the causal converb. An example of a nominal demonstrative intruding between the nominalized verb stem and the instrumental marker occurs in (7.19), renumbered and repeated below as (11.51) to facilitate comparison.

- (11.51) *tətʃhàku pa sə, ámíkhót tʃàpà? sə nə tʃùku təpìləmì à .jàɪ.*
tə-tʃhà-ku pa sə a-mi? khót
 thus-do-LOC.CV 3SG ANAPH NRL-person hand
tʃà?-pà? sə nə tʃùku tə-phìləm-ì à
 consume-NR ANAPH INST consequently NZP-think-IRR one
 .jà-əɪ
 come-PRES
 ‘And so, because she ate [the aforementioned] human hand, consequently a burden comes [to him].’

How does the concatenation of a nominalizer suffix and the instrumental marker become grammaticalized as a converb marking causality, and furthermore, how does a nominal demonstrative come to be optionally inserted between these two morphemes? A noun phrase modified by a post-head relative clause and potential indeterminacy in the function of a case marker in following

7. Consultants were unable to establish the meaning of this nominal suffix. It could be cognate with PTB **liŋ* ‘neck’ [STC #96], suggesting an elaborate expression for this body part.

elicited example suggest a conducive set of structural and semantic conditions for the reanalysis of erstwhile nominal constituents as a converb clause expressing reason.

- (11.52) *thalata tsəŋi ɰàpàjnə sənaj ɰàksa.ù?*
 thalata tsəŋi ɰà-pà? i nə
 now+month rain come-NR PROX AGT/INST
 sənaj ɰàksa?-əɰ-ù?
 crop break-PRES-DEC
 (i) ‘These rains that come this month damage the crops.’
 (ii) ‘Because these rains come this month, the crops are damaged.’

This sentence has two possible interpretations, according to how the nominalized verb stem and the case marker *nə* are parsed. Recall from §5.3 that the agentive, instrumental and allative case markers have a conflated morphological form in *nə*. Interpretation (i) assumes that *tsəŋi ɰàpà i nə* contains a noun head, a post-head relative clause, a proximal demonstrative, and an agentive marker that encodes its noun phrase as the A argument of that clause’s predicate. But if this case-marker is instead reanalyzed as marking an oblique instrumental argument, then it creates the semantic conditions that allow for the cause and effect adverbial interpretation of (ii) to be read off the same sentence. In other words, from the instrumental interpretation ‘By these rains that come this month ...’ an adverbial causal reading is derived, expressing ‘By cause of these rains that come this month ...’. Thus, from this historical reanalysis emerges the causal converb. Furthermore, if it is accepted that the causal converb construction originates from an erstwhile noun phrase containing a post-head relative clause, then it is entirely reasonable for that noun phrase to also be determined by a nominal deictic, and for that nominal deictic to maintain its function and position in the morphosyntactic reanalysis. We have already seen in §5.3.2.2 that a nominal demonstrative optionally occurring within the ablative case marker can be attributed to historical processes resulting in the reanalysis of function. The nominal demonstrative of causal converb constructions is similarly a relict of reanalysis.

Lastly, the causal converb *-pànə* can be used as a conjunctive element in narrative discourse, attaching to the quotative particle in the absence of a head. This expresses a logical implication holding between preceding and following clauses. Example (11.53) is illustrative of this conjunctive function. The causal discourse connective *təpànə* links an explanation for a state of affairs to the events that have led to that state of affairs.

- (11.53) *təpànə tʃàmìi kìn sə maŋmətuŋ i ku lai ami? sə pulukə la məli.ù?*
tə-pànə tʃàmìi kìn sə maŋmətuŋ i ku
 thus-CAUSAL clan.name clan ANAPH village.name PROX LOC
 la i a-mi? sə **pulu-əkə** la
 TOP PROX NRL-person ANAPH be.abundant-SIM TOP
 mə-li-əi-ù?
 NEG-be-PRES-DEC
 ‘That’s why/Because of that, the Jamir clan is not abundant here in
 Mangmetong village.’ (Literally ‘... does not abundantly exist’.)

11.4.3.7. Locative converb

The locative case marker has grammaticalized as a converb marker that usually encodes a temporal relationship between the events of a dependent clause and its matrix clause.

- (11.54) *tətʃhà.è thaŋtsəla tʃu tʃa.è ahju tʃu səmtsəku,*
tə-tʃhà-əi thaŋtsə-la tʃu **tʃa-əi** a-hju tʃu
 thus-do-SEQ tree.shrew.sp-F DIST call-SEQ NRL-word DIST
səmtsə-ku
 ask-LOC.CV
 ‘And then [Lichaba] calls Squirrel, and upon asking [for her
 explanation], ...’
- (11.55) *tsəŋəi, hniłi? waku, hnikə kəwaku,*
tsəŋ-əi **hni-łi?** **wa-ku** **hni-əkə** **kəwa-ku**
 spear-SEQ track-SIM go-LOC.CV track-SIM go+ascend-LOC.CV
 ‘Having speared [the wild pig], while [he] was tracking and tracking
 [it], when [it] was going up, ...’

While direct suffixation to a verb root is possible, it is more common for the locative converb suffix to co-occur with a grammaticalized form of a main verb meaning ‘reach, arrive’ inserted between the verb root and the locative converb marker. According to consultants, the addition of *thùŋ* (REACH) doesn’t change the meaning of a stem marked by the locative converb alone. This was tested by swapping textual examples of VERB-*thùŋ-ku* with just VERB-*ku*, and vice versa.

- (11.56) *pa pùŋì tʃu atsəlíʔ wathùŋku,*
 pa pùŋì tʃu **atsə-líʔ** **wa-thùŋ-ku**
 3SG wild.pig DIST look-SIM go-REACH-LOC.CV
 ‘When he was going, looking for the wild pig, ...’

Furthermore, a nominal demonstrative can be inserted between the grammaticalized verb *-thùŋ* and the locative *-ku*. This seems to add an emphatic meaning, pointing to the time that the dependent clause event takes place, and corresponds to the way a deictic is used for pointing in the spatial domain. I have argued elsewhere that the ostensibly “infixes” positions of nominal deictic constituents can be attributed to the nominal origins of these structures and historical processes of reanalysis (e.g. see §5.3.2.2 and §11.4.3.6 for further examples and discussion).

- (11.57) *tsəpàʔ məsəməkə ìnti phithùŋtʃùku,*
 tsəpàʔ **məsəm-əkə** ìnti **phi-thùŋ <tʃu > ku**
 well clean-MAN path clear-REACH<DIST>LOC.CV
 ‘At that time when they were cleaning the pond and clearing the path, ...’

11.4.4. Notes on the diachronic origins of converbs

The grammaticalization of adverbial subordinators from postpositions, case marking clitics and nominalizing morphology is a widely attested phenomenon. Genetti (1986) examined the grammars of twenty-six languages of the Bodic branch of Tibeto-Burman and found that syncretism between nominal case marking morphology and adverbial subordinators is not random, but follows regular patterns that are indicative of a grammaticalization trend. Nor is this limited to Tibeto-Burman languages. Cross-linguistically, similar patterns have been reported in a diverse range of language families, including Papuan (Foley 1986), Australian (Austin 1981, Simpson 1988, Dench 1988) and Chibchan (Craig 1991).

Table 11.2 below summarizes the likely diachronic origins of Mongsen converb suffixes. Evidence in support of the proposed historical sources of the converb markers can be found above, in the respective sections describing each type of converb suffix. The development of non-finite clause-linking suffixes from case marking morphemes in Mongsen is generally found to be consistent with the typological patterns of grammaticalization described in a cross-linguistic study of the phenomenon by Genetti (1991). To this I would add that nominalizers also appear to provide just as important a historical source for their diachronic development, based upon the frequency with which cognate

forms are found to contribute to a converb-marking function. This is not surprising, given the nominal origins of much of the converb morphology.

Table 11.2. Diachronic origins of converbs

FORM	GLOSS	DIACHRONIC SOURCE
-əɪ	(SEQ)	<agentive nominalizer/relativizer/genitivizer -əɪ
-(ə)kə; líʔ; -a	(SIM)	< ? benefactive case marker <i>atəməkə/atəma</i>
-li	(ALT.IT)	< source unknown, perhaps existential <i>li</i> ‘be at, stay’
-ku	(LOC.CV)	< locative case marker <i>ku</i>
-pàla	(COND)	<general nominalizer -pàʔ + topic particle <i>la</i>
-kùla	(CIRCM)	< locative case marker <i>ku</i> + topic particle <i>la</i>
-pàkukàʔ	(CONCESS)	< general nominalizer -pàʔ + locative case marker <i>ku</i> + final syllable of <i>anukàʔ</i> ‘even, yet, still, also’
-likàʔ	(CONTEMP)	< ? existential <i>li</i> ‘be at, stay’ + final syllable of <i>anukàʔ</i> ‘even, yet, still, also’
-tʃən	(DUR)	< locative nominalizer -tʃən
-pànə	(CAUSAL)	< general nominalizer -pàʔ + agentive/instrumental case marker <i>nə</i>

11.4.5. Expression of temporal antecedence

Mongsen lacks a lexical exponent for temporal ‘before’. In compensation for this lexical gap, temporal antecedence is expressed by a dependent clause formed with the negative prefix and the locative converb suffix attached to a verb stem. A nascent postposition in auxiliary function may occur between the verb root and the locative converb suffix.

(11.58) *pa nə á-sáʔ sə hənəɪ, tuku lítʃápáʔ ki phitʃunə məkhəptsəŋtaku,*

pa nə á-sáʔ sə **hən-əɪ**
 3SG AGT NRL-meat ANAPH take-SEQ

tə-ku lítʃá-pàʔ ki phi <tʃu> nə
 RL-uncle PN-M house <DIST>ABL

mə-khəp-tsəŋta-ku

NEG-depart-BETWEEN-LOC.CV

‘Before he departs from the house of Uncle Lichaba, taking the [aforementioned]meat, ...’

The locative converb suffix can be omitted or replaced by a time word that is restrictively focused by the discourse focus particle. It is not feasible to analyze *təsi* in the following example as a postposition, because it demonstrates an adverb-like distribution and meaning and occurs with other constituents besides noun phrases. It is most commonly found with other time words, e.g. *thaku təsi* ‘until today’. It is accordingly analyzed as a member of the time word lexical class (see §4.2.8).

- (11.59) *nəni məthùŋ təsi áŋ pi tèsəlà.*
 nə-ni **mə-thùŋ** **təsi** **áŋ** pi tès-əlà
 2SG-wife NEG-reach until just PROX PROH-untie
 ‘Don’t untie this before reaching your wife.’

11.4.6. Disjunctive expression using converbs

A form of the verb ‘do’ that is negated and marked by the conditional converb has grammaticalized a disjunctive meaning as a type of discourse connective and is used to express ‘or’. The derived disjunctive meaning requires the presence of the negative prefix on the converb stem.

- (11.60) *alà tsəŋikhu asəm ni mətʃhàpàla atsəkjìm ku phəli ni,*
 alà tsəŋikhu a-səm ni
 and day NRL-three day
mə-tʃhà-pàla atsəkjìm ku phəli ni
 NEG-do-COND winter LOC four day
 ‘And then, [in] three days, or in winter, four days, ...’
 (speaking of how long rice beer should be left to ferment)

Example (11.61), taken from a Waromung Mongsen text, uses a similar structure to express a disjunctive concessive meaning.

- (11.61) *mətʃhàpàkukà? ánətì tʃhàpàla, .iatʃhənlaɪ tà tʃaɪ.*
 mə-tʃhà-pàkukà? ánətì tʃhà-pàla .iatʃhən-la-əɪ
 NEG-do-CONCESS woman do-COND soothsayer-F-ANOM
 tà tʃa-əɪ
 thus call-PRES
 But even if [that were-tiger happens not to be a man], if [that were-tiger is a] woman, then [she] is called a “rachenlar.”

Other clause-linking discourse connectives that are used to express conjunctive meanings in discourse are described in §4.2.10.

11.5. Space to time grammaticalizations

The root of the relational noun *tə-sin* (RL-back) is developing a sequential meaning equivalent to ‘after VERB’. The following examples demonstrate this extended function. Note, however, that in (11.62) *sin* occurs juxtaposed to a sequential converb, which is also used to express temporal sequence, while in (11.63) it co-occurs with the locative case marker in order to express a location in time. This indicates that the function of *sin* as a fully-fledged marker of temporal antecedence has not yet completely grammaticalized to the extent that it can encode a temporal presupposition in the absence of supporting morphology.

- (11.62) *tuŋət liɪ sin áhlù jìm.*
 tuŋət li-əɪ sin á-hlú jìm
 3DU stay-SEQ BACK NRL-field cultivate.PST
 ‘After living [together], they cultivated fields.’

- (11.63) *tətʃhàɪ sin ku tʃhəmpanj mətʃəmpùŋ à tʃu, ahən təpuŋ mətəmi, təpəti à tʃhà.*
 tə-tʃhà-əɪ sin ku tʃhəmpanj.mətʃəmpùŋ à tʃu
 thus-do-SEQ BACK LOC bird.sp one DIST
 a-hən təpuŋ mətəmi tə-pəti à tʃhà
 NRL-chicken cock manner NZP-be.big one COP.PST
 ‘And then, after that, there was Yellow-backed Sunbird, who was a big one, like a rooster.’

This extension of its function is a corollary of the recategorization of *sin* from a relational noun functioning as an oblique noun phrase head, to a postposition marking a location in space, and finally to its evolving reanalysis as a temporal subordinator marking a location in time (see §5.3.3 for discussion of the recategorization of relational nouns).

A grammaticalization involving *mənaŋ* ‘front’ appears to be evolving a similar function. The inherent meaning of this locational noun also lends itself to expressing location both in space and time. The developing ‘location in time’ meaning allows it to be used to indicate temporal anteriority, as suggested by the following example.

- (11.64) *maŋmətuŋ tə tʃapaʔ tʃu, lùŋkhəm phinə, mənəŋthì, ləmtu, aláŋ, ləmtu, puŋən, tsəti, tə nə ni.ɪə, ajim za.ɪ.ɪə kəm.*

maŋmətuŋ	tə	tʃa-pàʔ	tʃu	lùŋkhəm	phinə
village.name	thus	call-NR	DIST	village.name	ABL
mənəŋ-thì	ləmtu	aláŋ	ləmtu	puŋən	
FRONT-INTNS	clan.name	clan.name	clan.name	clan.name	
tsəti	tə	nə	nì-ə.ɪ		
clan.name	thus ⁸	AGT	lead-SEQ		
a-jim	za-ɪ.ɪ-ə.ɪ	kəm			
NRL-village	enter-come-SEQ	become.PST			

‘The village that is called Mangmetong first became a village after the Lemtu, Alang, Pongen and Tseti clans, leading from Lungkum village, entered the village [site] and founded it.’

11.6. Causal particle

Dependent clauses expressing presuppositional reasons can be encoded by *aj*, a particle expressing ‘since’ and accordingly glossed. This morpheme often has a distinctive boundary marking intonation associated with its syllable.

While the meaning it encodes is more or less equivalent to that encoded by the causal converb suffix (§11.4.3.6), the broad distribution of the causal particle suggests that it is not a converb marker, nor some other kind of bound verbal suffix. For example, it is able to occur after a deverbal adjective that has been derived by the nominal prefix *tə-*, as demonstrated by (11.65). This is the first characteristic that distinguishes it from all the converb forms.

- (11.65) *amiʔ ta.ɪu aj, mələpjuʔ.*

a-miʔ	tə-ɪ.ɪu	aj	mə-ləp-ì-ùʔ
NRL-person	NZP-be.good	SINCE	NEG-cut-IRR-DEC
‘Since [they’re] good men, [we] won’t kill them.’			

Secondly, the causal particle can be juxtaposed to a verb stem inflected by what I assume is the present tense marker (it could possibly be the agentive nominalizer, which has an identical form, but a nominalization seems semantically unlikely here). Similarly, it can also occur in a clause with a verb stem marked by the irrealis suffix. This theoretically disqualifies it from

8. This is the only instance in the corpus of the quotative particle occurring within a noun phrase.

membership of the converb category, since no converb suffix is able to mark a verb stem that is independently specified for absolute tense or mood.

- (11.66) *“nì kà akuŋ tʃu a khiŋ atə, nì la tʃamajukəɾaj,” tə sa.*
 nì kàʔ a-kuŋ tʃu à khiŋ-aŋ atə
 1SG also NRL-prawn DIST one give-IMP PTCL
 nì la tʃäʔ-maʔ-tʃuk-əɾ aj tə sa
 1SG TOP consume-CMPL-PFV-SEQ SINCE thus say.PST
 ‘‘Give me one of those prawns as well,’’ [she] said, ‘‘seeing that I have eaten up all mine.’’

- (11.67) *mə̀həni aj, mə̀khiʔ-ì-ùʔ.*
 mə̀-hən-ì aj mə̀-khiʔ-ì-ùʔ
 NEG-take-IRR SINCE NEG-give-IRR-DEC
 ‘Since [I] won’t have any [prawns] to take, [I] won’t give any.’

Thirdly, while the causal particle requires neither an underived verb stem nor the quotative complementizer as a base, that does not preclude it from co-occurring with these constituents, as we see in the following examples. Observe how its juxtaposition to the quotative complementizer in (11.68b) allows the scope of the proposition encoded by the causal particle to extend over multiple dependent clauses in a complex sentence. The unusual distribution and function of this particle thus makes it unique in the language. Its diachronic origin is unknown at present.

- (11.68) a. *nàŋ azən kətaj.*
 nəŋ azən kət aj
 2SG power possess SINCE
 ‘Since you possess power, ...’
- b. *nàŋ azən kət-ùʔ tə sataj.*
 nəŋ azən kət-ùʔ tə sa tə aj
 2SG power possess-DEC thus say.PST thus SINCE
 ‘Since it is said that you possess power, ...’

11.7. Complementation

A complement functions as an argument of the predicate. The structure of the embedded complement ranges from one word, as in the use of onomatopoeia and interjections, to full verbal clauses consisting of a predicate and its core

arguments. To aid in their recognition, the embedded complements of the examples of this section are marked by square brackets.

11.7.1. Embedded direct speech complements

Embedded direct speech complements function as O arguments of their matrix clauses. The embedded complement is marked by the quotative complementizing particle *tə̀*, and the matrix clause contains a verb of utterance. The complement occurs in the preverbal position of the clause, the unmarked position for O arguments of independent clauses.

- (11.69) *tsə̀ŋrutshə̀la nə́ “áhlù tə̀pə̀ti à tsə̀aŋ,” tə̀ saw? tə̀jù?*
 tsə̀ŋrutshə̀-la nə́ [a-hlú tə̀pə̀ti à tsə̀-aŋ]
 PN-F AGT NRL-field NZP-be.big one clear-IMP
 tə̀ sa-ù? tə̀I-ù?
 thus say-PST.DEC REP-DEC
 ‘Tsengrutsela said “Clear a big field.”’

A direct speech complement may be embedded into a dependent clause that is itself in a dependent relationship to a matrix clause.

- (11.70) *lítfápá? nə́ “nàŋ tfə̀pá? nə́ waɪ,” tə̀ sə̀mtsə̀ku,*
 lítfá-pà? nə́ [nàŋ tfə̀pá? nə́ wa-ə̀ɪ] tə̀ sə̀mtsə̀-ku
 PN-M AGT 2SG what ALL go-PRES thus ask-LOC.CV
 ‘When Lichaba asked “Where are you going?”, ...’

Embedded clauses expressing onomatopoeia are superficially similar to direct speech and other types of complements, because they can also be marked by the quotative particle. They differ in not necessarily functioning as core arguments of the clause, although they can fulfil this role. The onomatopoeia of (11.71) is not a sentential complement, whereas that of (11.72) definitely does function as the O argument of its matrix clause.

- (11.71) *sə̀ŋnik sə̀ŋji tʃu ku “tʃə̀ɪət! tʃə̀ɪət!” tə̀ phə̀latshə̀ŋə̀ɪ.*
 sə̀ŋ-nik sə̀ŋji tʃu ku [tʃə̀ɪət tʃə̀ɪət]
 wood-eye wood.grain.around.knot DIST LOC crack! crack!
 tə̀ phə̀la?-tshə̀ŋ-ə̀ɪ
 thus strike-ATTACH-PRES
 ‘“Crack! Crack!” [the lightning] strikes the grain around the knot in the wood.’

- (11.72) *təɪ thani təsi “káku káku” tɛ̀ asà.ùʔ.*
 tə-əɪ thani təsi [káku káku] tɛ̀ asà-əɪ-ùʔ
 thus-SEQ now+day until ONOM ONOM thus shout-PRES-DEC
 ‘And until today, [that cuckoo bird] still cries “Kaku, Kaku” (that is, “Uncle, Uncle”).’

11.7.2. Other complement taking predicates

Other complement taking verbs include verbs of thinking, e.g. *maŋ* ‘believed’, *philəm* ‘thought’ and verbs of permission, e.g. *mə̀lə̀* ‘permitted’. These all share the same syntactic structure as embedded direct speech complements of utterance predicates, as described in the preceding section. They also use the quotative complementizer to separate the O complement from the predicate, despite their predicates not belonging to the semantic class of utterance predicates.

- (11.73) *ahwəɪ nə álíma tʃu líʃápàʔ tsə̀ŋh.ɪəm nə zaŋlu tɛ̀ maŋ.*
 ahw-əɪ nə [álímá tʃu líʃápàʔ tsə̀ŋh.ɪəm nə zaŋlu]
 Ao-ANOM AGT world DIST PN-M deity AGT make.PST
 tɛ̀ maŋ
 thus believe
 ‘The Ao believe that the god Lichaba made the world.’

- (11.74) *təpaʔ à nə sùpàʔ táŋ sə̀mtsə̀.ùʔ tɛ̀ philəmùʔ.*
 [tə-paʔ à nə sùʔ-pàʔ táŋ sə̀mtsə̀-əɪ-ùʔ]
 RL-father one INST beget-NR just ask-PRES-DEC
 tɛ̀ philəm-ùʔ
 thus think.PST-DEC
 ‘[He] thought [Lichaba] was just asking [him about the number of sons] begotten of one father.’

Texts

1. The story of the Long-tailed Broadbill

Speaker: †Lemning Kaba, aged 69 years at time of recording

Recorded: May 1999 at Mangmetong village, Nagaland

This unedited, traditional Mongsen folk tale explains why the Long-tailed Broadbill only drinks rainwater that has collected on leaves and other *ad hoc* receptacles in the jungle. The narrative has additionally been translated into Chungli with the help of a bidialectal speaker to demonstrate the morphological differences and correspondences between the two major dialects of Ao. The first text line in italic face represents the Mongsen dialect with output tones and morphophonological processes shown, the second tier of interlinearization (marked by an initial M) shows segmentation of Mongsen morphology, and the third tier of interlinearization (marked by an initial C) represents the equivalent forms of the Chungli dialect. As this is the first contact I have had with Chungli, segmentation into morphemes has not been attempted. Output tones have been marked, however. My initial impression is that Chungli has three lexically contrastive pitches, in common with Mongsen.

The narrator of the story was a man who lived in *imlàn məphù*, the lower ward of Mangmetong village. The consultant who assisted with the translation into Chungli was a bidialectal speaker of the Longchang variety of Mongsen and the standard variety of the Chungli dialect.

01 *wàzà? təlʉ? la tsəpà? məsəm wa.*

M wàzà? təlʉ? la tsəpà? məsəm wa

C ùzə? atʃak ?¹ tsəpu? məsami àw
bird all TOP well clean.PURP go.PST

All the birds went to clean a well.

02 *təɪ, khəlɪŋlɪ, təɪ wàzà? tənɪŋ à li.*

M tət-əɪ khəlɪŋlɪ tət-əɪ wàzà? tənɪŋ à li

C asəɪ kəlɪŋlɪ asəɪ ùzə? tənəŋ ka liasə
thus-SEQ bird.sp. thus-SEQ bird RL-name one be.PST

And, there was a bird named the Longtailed Broadbill.

1. My bidialectal consultant was not sure which Chungli morpheme corresponded to the topic particle *la* of Mongsen. Question marks elsewhere in the Chungli data indicate the lack of an equivalent morpheme.

03 *təi khəliŋlila tʃu,*

M tə-əi khəliŋlì-la tʃu

C asəi kəliŋlila tʃi
thus-SEQ bird.sp.-F DIST

And, that Long-tailed Broadbill,

04 *tsəpà? məsəməkə inti phithùŋtʃùku,*

M tsəpà? məsəm-əkə inti phi-thùŋ-tʃu-ku

C tsəpu? məsama ləmaŋ apitəŋtʃi
well clean-SIM path clear-REACH-DIST-LOC.CVat the time [the other birds] were in the process of cleaning the pond
and clearing the path,05 *pa mə.ɪàlà.*

M pa mə-ɪà-la

C pa maɪù
3SG NEG-come-NEG.PST
she didn't come.06 *təpànə wàzà? təlulʃ səntəpəi,*

M tə-pànə wàzà? təlulʃ səntəp-əi

C ípá? tʃaki ùzə? atʃak atənsiɪ
thus-CAUSAL bird all gather-SEQ

That's why all the birds, after gathering together,

07 *pa tələm ku tsəpànə tələm tʃu apaktʃuk.*

M pa tə-ləm ku tsə-pànə tə-ləm tʃu apak-tʃuk

C pa tukulak nuŋ atʃəmpaaki tukulak tʃi mətəsəkatak
3SG RL-head LOC peck-CAUSAL RL-head DIST flat-PFV.PST
because [they] pecked on her head, [her] head became flat.²08 *təi pa tsəpà? ku atsə mətʃəməjù?*

M tə-əi pa tsəpà? ku a-tsə mə-tʃəm-i?-i-ù?

C asəi pa tsəpu? nuŋ tsə mətʃəmtaktsətsə
thus-SEQ 3SG pond LOC NRL-water NEG-drink-CAUS-IRR-DEC
And, [they] would not let her drink at the pond.

2. Here the speaker appears to be confusing the character of the Long-tailed Broadbill with that of the Owl in another story (cf. Text 4).

09 *təɪ wàzàʔ nə tsəpàʔ ku atsə mətʃəmilà.*

M tə-əɪ wàzàʔ nə tsəpàʔ ku a-tsə

C asəɪ ùzàʔ i tsəpuʔ nuŋ tsə
thus-SEQ bird AGT pond LOC NRL-water

M mə-tʃəm-iʔ-la

C mətʃəmtaktsəʔ

NEG-drink-CAUS-NEG.PST

And, the birds didn't let [her] drink at the pond.

10 *təɪ pa tsəhɪàthùŋku tsəŋí tʃakə “tsəŋí tsəŋí” tə asəɪà.*

M tə-əɪ pa tsəhɪà-thùŋ-ku tsəŋi tʃa-əkə

C asəɪ pa tsəɪàtaŋ tsəŋlu tʃaa
thus-SEQ 3SG be.thirsty-REACH-LOC.CV rain call-SIM

M tsəŋi tsəŋi tə asà-əɪ

C tsəŋlu tsəŋlu tà asataŋ
rain rain thus shout-SEQ

And so, when she becomes thirsty, [she] calls the rain, chirping “Rain, rain”,

11 *atʃu ku pa nə ásáthùŋku, tsəŋi ɪà.ɪà.*

M a-tʃu ku pa nə asà-thùŋ-ku

C itaŋ tʃi pa i asataŋtʃi
NRL-DIST LOC 3SG AGT shout-REACH-LOC.CV

M tsəŋi ɪà-əɪ

C tsəŋlú aɪ.ɪà
rain come-PRES

and when she chirps, the rain comes there.

12 *təɪ tsəŋi ɪàpàʔ táŋ tʃu,*

M tə-əɪ tsəŋi ɪà-pàʔ táŋ tʃu

C asəɪ tsəŋlu aɪpàʔ táŋ tʃi
thus-SEQ rain come-NR only DIST

And, only the rain that comes,

13 *pa səŋàwà aia ku kətpà? táŋ tʃəm-lí-tʃəm-lí? wau?*

M pa səŋàwà aia ku kətpà? tʃəm-lí?
 C pa səŋù.u ? nuŋ akətpà? tʃəmijùŋa
 3SG leaf and.the.like LOC possess-NR drink-SIM

M tʃəm-lí? wa-əi-ù?

C tʃəmijùŋa aw
 drink-SIM go-PRES-DEC

she goes around drinking just [the rain that comes] that collects on the leaves and things.

14 *tə atʃu táŋú?*

M tə a-tʃu táŋ-ù?
 C tà atʃi táŋà?
 thus NRL-DIST only-DEC
 And that's all [she drinks].

15 *hju-tsə tʃa tʃu.*

M hju-tsə tʃa a-tʃu
 C utsə tʃa atʃi
 story PTCL NRL-DIST
 That's the story.

2. The story of Noksensangba and Imtisangba

Speaker: Imdangchiba Imchen, aged 69 years

Recorded: June 2001 at Shillong, Meghalaya

This is an unedited traditional folktale about a man possessing magical powers and his rich neighbour, as told by a Mangmetong Mongsen speaker. The speaker was a resident *imlāŋ məphù*, the lower ward of Mangmetong Village for most of his life but had been living in an Ao colony in Dimapur, Nagaland for approximately ten years at the time of recording.

01 *ɹənəmukù sà.ù ná.*

ɹənəm-ukù sa-ù ná
 prepare-ANT say-IMM PTCL
 [I'm] ready. [I] will tell [the story], OK?

- 02 *hàj?*
 hàj?
 okay
 [interlocutor]: Okay.
- 03 *a, maŋmətuŋ əjim kəməi, phusù íťǎ? li.ə,*
 a maŋmətuŋ a-jim kəm-əi
 INTJ village.name NRL-village become-SEQ
 phusù íťǎ? li-əi
 century some be-SEQ
 After the founding of Mangmetong village, it being some centuries
 after,
- 04 *azən kətəi, à, nuksənsaŋpa?, tə ťǎ.ə sù?*
 azən kət-əi à nuksənsaŋ-pà? tə ťǎ-əi sù?
 power possess-ANOM one PN-M thus call-SEQ beget.PST
 one possessing magical powers called Noksensangba was born.
- 05 *təi, pa thən ťǎ maŋmətuŋ ku, təsì à imtisàŋpà?, təkħà, təsì, pa íťǎ?
 tán tətʃhatsə ami? à ťhà.*
 tə-əi pa thən ťǎ maŋmətuŋ ku tə-sì-əi
 thus-SEQ 3SG COM DIST village.name LOC NZP-know-ANOM
 à imtisàŋ-pà? tə-kħà-əi tə-sì-əi
 one PN-M NZP-have-ANOM NZP-know-ANOM
 pa íťǎ? tán tətʃhatsə a-mi? à ťhà
 3SG much just NZP-be.talented NRL-person one COP.PST
 And, with him there in Mangmetong village lived a wealthy, learned
 and very talented person, a man by the name of Imtisangba.
- 06 *pa kà? li.*
 pa kà? li
 3SG also stay.PST
 He also lived there.
- 07 *tətʃhà, nuksənsaŋpa? nə təmħnak ni.ə,*
 tə-ťhà-əi nuksənsaŋ-pà? nə tə-əmħnak ni-əi
 thus-do-SEQ PN-M AGT RL-son.in.law lead-SEQ
 And then, [one day] Noksensangba took his son-in-law,

- 08 *khapza in tájǎ nə wa.ɪ̀.*
 khapza in taŋi nə wa-əɪ
 village.name PATH side ALL go-SEQ
 [and they] went towards the Kabza village side,
- 09 *ajuŋ ku, tsəj̄ ajuŋ ku, asət à zaŋlu.*
 a-juŋ ku tsəj̄ a-juŋ ku asət à
 NRL-river LOC river.name NRL-river LOC fishing.weir one
 zaŋlu
 make.PST
 [and] built a fishing weir in a river, the Tsei River.
- 10 *təɪ pa nə təmhnak ni.ɪ̀ wa.ɪ̀ at̚su ku zaŋlukə áhǎǎ? phà?*
 tə-əɪ pa nə tə-əmhnak ni-əɪ wa-əɪ
 thus-SEQ 3SG AGT RL-son.in.law lead-SEQ go-SEQ
 a-t̚su ku zaŋlu-əkə a-hǎǎ? phà?
 NRL-DIST LOC make-SIM NRL-fish catch.PST
 And, after leading his son-in-law there and building [a fishing weir,
 they] caught fish.
- 11 *nǐ?, áhǎǎ? təpəti à asət t̚su nə za.*
 nǐ? a-hǎǎ? tə-pəti à asət t̚su nə
 one.day NRL-fish NZP-be.big one fishing.weir DIST ALL
 za
 enter.PST
 One day, a big fish entered the fishing weir.
- 12 *təɪ, nuksənsaŋpa? nə “u! áhǎǎ? təji la ləɪa.ù? ɪ̀.”*
 tə-əɪ nuksənsaŋ-pà? nə u a-hǎǎ? tə-ji
 thus-SEQ PN-M AGT EXCLM NRL-fish RL-mother³
 la ləɪa-əɪ-ù? ɪ̀
 TOP descend+come-PRES-DEC PTCL
 ‘And, Noksensangba [said] “Ah! A huge fish is coming down [the
 river].”’

3. The use of *təji* ‘RL-mother’ here suggests an emerging augmentative function for this morpheme. See discussion of §7.5.3.

- 13 *təɪ atʃu phətʃhətəɪ, təmhnak li “aki nə hənəɪ tʃhuwaŋ.”*
 tə-əɪ a-tʃu phà.ɪaʔ-tʃhət-əɪ
 thus-SEQ NRL-DIST catch-ABIL-SEQ
 tə-əmhnak li a-ki nə hən-əɪ tʃhuwa-aŋ
 RL-son.in.law DAT NRL-house ALL take-SEQ emerge-IMP
 And, having managed to catch it, [he] gave it to his son-in-law [saying]
 “Take it back to the house”.
- 14 *təɪ, “nəŋ ɪjáʔ kəmtaŋkə táŋ, jipəɪə ɪaŋ.”*
 tə-əɪ nəŋ ɪjáʔ kəmtaŋ-əkə táŋ jipli-əɪ ɪà-aŋ
 thus-SEQ 2SG much be.careful-SIM just sleep-SEQ come-IMP
 And, [he said] “You just sleep very carefully [there] and then come”
 (Implying “don’t have sex with your wife tonight”).
- 15 *“hàjʔ.” təɪ təmhnak nə áhŋáʔ tʃu ajim nə hənəɪ tʃhuwa.*
 hàjʔ tə-əɪ tə-əmhnak nə a-hŋáʔ tʃu
 okay thus-SEQ RL-son.in.law AGT NRL-fish DIST
 a-jim nə hən-əɪ tʃhuwa
 NRL-village ALL take-SEQ emerge.PST
 “O.K.” [he said]. And then, the son-in-law took the fish and returned to
 the village.
- 16 *təɪ ... likàʔ tənì tʃu nə, “aki tʃhapəŋ!”*
 tə-əɪ ...-likàʔ tən-ì tʃu nə a-ki tʃhapəʔ-aŋ
 thus-SEQ -CONTEMP⁴ RL-wife DIST ALL NRL-house open-IMP
 And then, when [the son-in-law] said to his wife⁵ “Open the house!,”
- 17 *təsəlikàʔ tənì nə anìʔ mətshəŋla hɪ.ɪ.ɪ.ə, aki tʃhapàʔ.*
 tə-sə-likàʔ tən-ì nə a-nìʔ mət-tshəŋ-lá
 thus-say-CONTEMP RL-wife AGT NRL-sarong NEG-wear-NEG.CV
 hɪ.ɪ.ə-əɪ a-ki tʃhapàʔ
 level+come-SEQ NRL-house open.PST
 Thus, when [he] said that, his wife, not wearing a sarong (in order to
 tempt him), came and opened the house.

4 There is a substantial pause preceding the cotemporal realis converb suffix (represented by the ellipsis) in which time the speaker regathers his thoughts. This accounts for the unusual absence of both the quotative particle and a head.

5. This is a rare instance of a speech act directed at a human interlocutor without the noun phrase of that referent being marked by an auxiliary postposition in addition to the allative (see §5.3.3 for discussion).

- 18 *təlikula áhǵá? sə hməntʃakəi áŋ, tuŋnət atʃu ku jip.*
 tə-li-kula a-hǵá? sə hmən-tʃak-əi táŋ
 thus-BE-CIRCM NRL-fish ANAPH set-RS-SEQ just
 tuŋnət a-tʃu ku jip
 3DU NRL-DIST LOC lie.down.PST
 That being the case, he just set the [aforementioned] fish down and they [two] copulated there.
- 19 *təku atsə tsəŋhɪəm nə, təmhnaŋ tʃu, nuksənsaŋpa? təmhnaŋ tʃu, təhmila tʃu tʃatʃuk.*
 tə-ku a-tsə tsəŋhɪəm nə tə-əmhnaŋ tʃu
 thus-LOC.CV NRL-water deity AGT RL-son.in.law DIST
 nuksənsaŋ-pà? tə-əmhnaŋ tʃu tə-hmila tʃu
 PN-M RL-son.in.law DIST RL-soul DIST
 tʃà?-tʃuk
 consume-PFV.PST
 And then, a water spirit ate Noksensangba's son-in-law's soul.
- 20 *alà nipa ku, nuksənsaŋpa? asət ku lipa? tʃu, nuksənsaŋpa? nə taŋ nə walikà?*
 alà nipa ku nuksənsaŋ-pà? asət ku li-pà? tʃu
 and dawn LOC PN-M fishing.weir LOC be-NR DIST
 nuksənsaŋ-pà? nə taŋ nə wa-likà?
 PN-M ALL SIDE ALL go-CONTEMP
 And, at dawn, when [the son-in-law] had gone to Noksensangba, who was at his fishing weir,
- 21 *áhǵá? tʃu, nuksənsaŋpa? tʃu áhǵá? tʃu, ləɪali, səpsiɪ táŋ kəwajukli,*
 a-hǵá? tʃu nuksənsaŋ-pà? tʃu a-hǵá? tʃu
 NRL-fish DIST PN-M DIST NRL-fish DIST
 ləɪa-li səpsi-əi táŋ kəwa-tʃuk-li
 descend+come-ALT.IT turn-SEQ just ascend+go-PFV-ALT.IT
 Noksensangba [noticed that] the fish repeatedly came down [towards the weir], then just turned around and went back up [the river] again,

- 22 *ləialì səpsɪ táj pakkətəi áŋ kəwajukli,*
 ləia-li səpsi-əɪ táŋ pak-kət-əɪ táŋ
 descend+come-ALT.IT turn-SEQ just disperse-ASCEND-SEQ just
 kəwa-tʃuk-li
 ascend+go-PFV-ALT.IT
 then came back down before turning again, just dispersing as they went
 back up, over and over,
- 23 *“thajip i tʃəpá? tʃhà pi.”*
 thajip i tʃəpá? tʃhà pi
 morning PROX what HAPPEN.COP PROX
 [causing Noksensangba to wonder] “What’s going on this morning?”
- 24 *“áhá? təji lala.ə hɣaza la təlu? səpsijuk.ə pà pi.”*
 a-hhá? tə-ji lala-əɪ hɣá?-za la təlu?
 NRL-fish RL-mother descend+come-SEQ fish-DIM TOP all
 səpsi-tʃuk-əɪ pà pi
 turn-PFV-PRES QPTCL PROX
 “After the big fish come down, all the small fish turn back, is it?” (the
 water spirit has been offended by the sexual activities of the son-in-law
 and his wife on the previous night, so it prevents the fish from entering
 the trap)
- 25 *təɪ tsəŋmaŋkə likə təmhnaɪ sə .ià.*
 tə-əɪ tsəŋmaŋ-əkə li-əkə tə-əmhnaɪ sə .ià
 thus-SEQ wonder-SIM be-SIM RL-son.in.law ANAPH come.PST
 And, while [he] was wondering about this, his [aforementioned] son-in-
 law came.
- 26 *təku təmhnaɪ sə tʃhaɪuli? .ià.*
 tə-ku tə-əmhnaɪ sə tʃhaɪ-uli? .ià
 thus-LOC.CV RL-son.in.law ANAPH be.sick-SIM come.PST
 And, his son-in-law came in a state of ill health (he is bewitched by the
 water spirit).
- 27 *“a məkhàw?”*
 a mə-khà-ù?
 EXCLM NEG-have-DEC
 “It's hopeless!” [says Noksensangba].

- 28 “*nàŋ la, nǐ salí? ùŋlí? ku*”
 nàŋ la nǐ sa-lí? ùŋ-lí?-ku
 2SG TOP 1SG.AGT say-SIM see-SIM-LOC.CV⁶
 “See? I warned you!”
- 29 “*tʃópá? tʃhàpà?, nənì thən mətəm?*”
 tʃópá? tʃhà-pà? nə-ni thən mətəm
 what do-NR 2SG.POSS-wife COM copulate.PST
 “Why did [you] have sex with your wife?”
- 30 “*ajukə .iaŋ!*”
 aju-əkə .ià-aŋ
 be.quick-SIM come-IMP
 “Come here quickly!”
- 31 “*təməphà ku tshətaŋ.*”
 tə-məphà ku tshət-aŋ
 RL-foot LOC stand.on-IMP
 “Stand on my feet!”
- 32 *təi təməphà ku tshəti, mətʃakkətəkpa? ajim thùŋ.*
 tə-əi tə-məphà ku tshət-i?-əi
 thus-SEQ RL-foot LOC stand.on-CAUS-SEQ
 mətʃak-kət-zək-pà? a-jim thùŋ
 step-ASCEND-DIR-NR NRL-village reach.PST
 And then, making [the son-in-law] stand on [his] feet, the stepping
 [ones] reached the village (in one magical step).
- 33 *təi, təmhnak tʃu sətʃuk.*
 tə-əi tə-əmhnak tʃu sə-tʃuk
 thus-SEQ RL-son.in.law DIST die-PFV.PST
 And, the son-in-law passed away.
- 34 *təi, aməŋ tʃu həntʃukəkə tə tʃhà.ə,*
 tə-əi a-məŋ tʃu həntʃuk-əkə tə tʃhà-əi
 thus-SEQ NRL-body DIST discard-SIM thus do-SEQ
 And, after disposing of the body,

6. *salí? ùŋlí? ku*s a frozen expression used to chide.

- 35 *nuksənsaŋpa? tʃu, m̀n̄p̄hu ku áhlù tsəj áhlù nə wa.*
 nuksənsaŋ-pà? tʃu m̀n̄p̄hu ku a-hlú tsəj
 PN-M DIST winter LOC NRL-field clear.PURP
 a-hlú nə wa
 NRL-field ALL go.PST
 in the winter, Noksensangba went to the field to clear it [in preparation
 for cultivation].
- 36 *tətʃhàlikà? imtisàŋpà? tək̀hà.ɪpà?, təs̄.ɪpà? sə, aki tʃhàth̀n̄ku,*
 tətʃhà-likà? imtisàŋ-pà? tək̀hà-əɪ-pà?
 thus-do-CONTEMP PN-M NZP-have-ANOM-M
 tət̄-s̄-əɪ-pà? sə a-ki tʃhà-th̀n̄ku
 NZP-know-ANOM-M ANAPH NRL-house make-REACH-LOC.CV
 And while [he] was doing this, Imtisangba – the aforementioned one
 who was wealthy and learned – when it came time to build his house,
- 37 *ásəŋ aw tʃu, ̀̀ɪa? phinə, aliŋ təlak phinə, ásəŋ tʃu hənəɪ tʃhuwa.ə, pa*
aki tʃhà.
 a-səŋ a-u tʃu ̀̀ɪa? phinə aliŋ
 NRL-wood bamboo DIST afar ABL lower.range
 tək̀-ɪak phinə a-səŋ tʃu hən-əɪ tʃhuwa-əɪ
 RL-end ABL NRL-wood DIST take-SEQ emerge-SEQ
 pa a-ki tʃhà
 3SG NRL-house make.PST
 [he] carried wood and bamboo from afar, from the end of the lower
 range, and returned and built his house.
- 38 *tə aki tʃu tʃhà ... tʃhalakəɪ,*
 tət̄ a-ki tʃu tʃhà tʃhà-lak-əɪ
 thus NRL-house DIST make make-TERM-SEQ
 Thus, having finished (narrator stutters) building [his] house,
- 39 *n̄ɪ? anənni à likə*
 n̄ɪ? a-nət-ni à li-əkə
 one.day NRL-two-day one be-SIM
 it being a few days later,

- 40 *nuksənsaŋpa? tʃu páɪ kìmà tənəkə áhlú nə wa.*
 nuksənsaŋ-pà? tʃu páɪ kìmà tən-əkə
 PN-M DIST 3PL.POSS courtyard cross-SIM
 a-hlú nə wa
 NRL-field ALL go.PST
 Noksensangba was crossing their [i.e. Imtisangba's] family's courtyard
 as he went to the field.
- 41 *“nuksənsaŋpa?, nàŋ aki nə kəɪaŋ.”*
 nuksənsaŋ-pà? nàŋ a-ki nə kəɪa-ŋ
 PN-M 2SG NRL-house ALL ascend+come-IMP
 “Noksensangba, come up to the house” [said Imtisangba].
- 42 *“ɪaŋ” tɛ̀.ù?*
 ɪà-ŋ tɛ̀.ù?
 come-IMP REP-DEC
 “Come”, [he] said.
- 43 *“hàj?.” tɛ̀ɪ kəwalikà?, tʃùku mətʃəmtsə mà.rùk anət ilakəɪ,*
 hàj? tɛ̀-əɪ kəwa-likà?
 okay thus-SEQ ascend+go-CONTEMP
 tʃùku mətʃəmtsə mà.rùk a-nət ilak-əɪ
 consequently rice.wine cup NRL-two pour-SEQ
 “OK.” And then, when [Noksensangba] went up, [Imtisangba]
 consequently having poured two cups of rice wine,
- 44 *“nàŋ azən kətù? tɛ̀ sa taj pi mətətɪ sɛ̀ məmətətɪ la.”*
 nàŋ azən kət-ù? tɛ̀ sa tɛ̀ aj
 2SG power possess-DEC thus say thus SINCE
 pi mətət-ì sɛ̀ mɛ̀-mətət-ì la
 PROX know-IRR QPTCL NEG-know-IRR TOP
 [Imtisangba said] “Since it is said that you possess magical power,
 would you know this or not?”
- 45 *“thajip i, aji tʃəmtʃhaj sɛ̀ mətʃəmtʃhaj?”*
 thajip i a-ji tʃəm-tʃha?-ì sɛ̀
 morning PROX NRL-rice.beer drink-CHANCE-IRR QPTCL
 mɛ̀-tʃəm-tʃha?-ì
 NEG-drink-CHANCE-IRR
 “Will you get the opportunity to drink this rice beer this morning, or
 not?”

- 46 *təlikàʔ, “nì tʃəmtʃhaj ùku.”*
 tə-likàʔ nì tʃəm-tʃhaʔ-ì ùku
 thus-CONTEMP 1SG drink-CHANCE-IRR PTCL
 When [asked that, Noksensangba] replied “I’ll get the opportunity to drink alright.”
- 47 *“nàŋ mətʃəmtʃhaj atì.”*
 nàŋ mətʃəm-tʃhaʔ-ì atì
 2SG NEG-drink-CHANCE-IRR PTCL
 “You certainly won’t get the opportunity to drink, though.”
- 48 *“ajà! màrùk ku aji sáʔ mətʃəmtʃhaj pà.”*
 ajà màrùk ku a-ji sáʔ
 EXCLM cup LOC NRL-rice.beer PTCL⁷
 mətʃəm-tʃhaʔ-ì pà
 NEG-drink-CHANCE-IRR QPTCL
 “Ayaa! Even though rice beer is in the cup I won’t get to drink [it], eh?”
- 49 *luk.ɿa.əɿ təmijun nə màrùk ku tùrə tùuktʃuk.*
 luk.ɿa-əɿ tə-mijun nə màrùk ku tùʔ-əɿ
 snatch-SEQ RL-finger INST cup LOC poke-SEQ
 tùʔ-uk-tʃuk
 poke-INTO-PFV.PST
 And snatching at the cup, he poked it with his fingers and knocked it over.
- 50 *“wá. tʃəj. nàŋ mətʃəmtʃhaj təpaʔ sə [nì ...] màrùk tùuktʃuk tà ti tət.”*
 wá tʃəj nàŋ mətʃəm-tʃhaʔ-ì tət-pàʔ sə
 INTJ PTCL 2SG NEG-drink-CHANCE-IRR thus-NR ANAPH
 [nì] màrùk tùʔ-uk-tʃuk tà ti tət
 1SG cup poke-INTO-PFV.PST like.that PTCL PTCL
 “See? You won’t get the opportunity to drink. Just as [I said], (narrator makes false start), you’ve gone and knocked the cup over, just like that.”

7. My consultant seemed to think that this particle implied a concessive meaning. It is the only example of its kind encountered in the corpus.

- 51 *təi nuksənsaŋpa? nə, “ípa? mətəm i hənəiaɪ tʃəmaŋ mə.”*
 tə-əɪ nuksənsaŋ-pà? nə ípá? mətəm i
 thus-SEQ PN-M AGT EMPHAT manner PROX
 hənəià-əɪ tʃəm-aŋ mə
 bring-SEQ drink-IMP PTCL
 And Noksensangba said “Drink holding the cup in this manner”,
- 52 *pa nə anəzə nə màiùk tʃu lukia.ə pa la tʃəm.*
 pa nə anəzə nə màiùk tʃu lukia-əɪ
 3SG AGT pair INST cup DIST snatch-SEQ
 pa la tʃəm
 3SG TOP drink.PST
 and he snatched up the cup with [his] pair [of hands] and drank.
- 53 *“alà, nì aki ípá? mətəm i tʃhàtʃhətpaj,”*
 alà nì a-ki ípá? mətəm i tʃhà-tʃhət-pà? i
 and 1SG NRL-house EMPHAT manner PROX make-ABIL-NR PROX
 “And so, my house that I am able to build like this,” [said Imtisangba
 boastfully]
- 54 *“pi, kúták ku kà? thəpsija.ə sə məthəpsi,” tə saw? təi.*
 pi kúták ku kà? thəpsi-ja-əɪ sə
 PROX heaven LOC also discuss-CONT-PRES QPTCL
 mə-thəpsi-əɪ tə sa-ù? təi
 NEG-discuss-PRES thus say.PST-DEC REP
 “Are [they] also talking about this in heaven, or not?” [he] asked.
- 55 *“újàtəj. nəŋəla tʃu thaku la anu? kúták lima tʃu ku məthəpsi.ù?”*
 újàtəj nəŋkhəla tʃu thaku la
 EXCLM 2PL DIST now+LOC TOP
 anukà? kúták lima tʃu ku mə-thəpsi-əɪ-ù?
 still heaven country DIST LOC NEG-discuss-PRES-DEC
 “Come off it! You lot (i.e. people like you) are still not being discussed
 in heaven nowadays.”

- 56 *“hlàkə, təpukzə in məphù ku, intən məphù ku”*
 hlà-kə tə-puk-zə in məphù ku
 descend+go-DEIC RL-stomach-LOC.SUFF PATH ward LOC
 intən məphù ku
 plain ward LOC
 “Way down there inside of the ward, in the plain ward,”
- 57 *“məraŋtʃəmpa? khə imtʃəntʃhàŋ nət thəpsi-ja? təli.”*
 məraŋtʃəm-pà? khə imtʃəntʃhàŋ nət thəpsi-ja tə li
 PN-M CONJ PN two discuss-CONT thus be
 “Merangjemba and Imchenchang, those two are being discussed”.
- 58 *“atʃu kúták thùŋpà? miw?”*
 a-tʃu kúták thùŋ-ì-pà? mi?-ù?
 NRL-DIST heaven reach-IRR-NR person-DEC
 “They are the people who will reach heaven” [said Noksensangba].
- 59 *“a, azənti asaŋ tʃùku təsəjuki hmapaŋ thùŋəi.”*
 a azənti.asa-əŋ tʃùku tə-sə-tʃuk-ì hmapaŋ
 INTJ old.person-ANOM consequently NZP-die-PFV-IRR time
 thùŋ-əi
 reach-PRES
 “Ah, [they’re] old people, so [they] are reaching the time to die.”
- 60 *“təpàla, kúták.ɪ nə kútá tʃhà-əi.”*
 tə-pàla kúták-əi nə kútá tʃhà-əi
 thus-COND heaven-ANOM AGT how do-PRES
 “If that is the case (i.e. that they are soon to die), what are the heavenly beings doing?” [asked Imtisangba]
- 61 *“a, tuŋnət anakəkə atʃhu tən-ja? təli.”*
 a tuŋnət anak-əkə a-tʃhu tən-ja tə li
 EXCLM 3DU praise-SIM NRL-song sing-CONT thus be
 “Ah, [the heavenly beings] are singing their praises” [says Noksensangba].

- 62 “*ú sàmsì tði li ni mǎɲtʃǝmpá? imtʃǝntʃhàŋ hǝntʃhà kùh.ɲàŋ nǝ sàmsì tði ni.*”

ú sàmsì tði li ni mǎɲtʃǝm-pà? imtʃǝntʃhàŋ
EXCLM sway thus be PTCL PN-M PN

hǝn-tʃha kùh.ɲàŋ nǝ sàmsì tði li ni
fowl-feather head-dress INST sway thus be PTCL

(Narrator sings a song with figurative meaning:) The feathers in a head-dress worn by a dancer are compared with heavenly beings. The feathers sway and move like heavenly beings dancing and singing the praises of Merangjemba and Imchenchang.

- 63 *tð, “tuŋnǝt tʃu, kúták nǝ .ɲàjpà?, asǝjim nǝ wajukípà?, kúták.ɲ nǝ thǝpsi.ɲu?”*, *tð sa.*

tð tuŋnǝt tʃu kúták nǝ .ɲà-ì-pà? a-sǝ-jim nǝ
thus 3DU DIST heaven ALL come-IRR-NR NRL-die-village ALL
wa-tʃuk-ì-pà? kúták-ɲ nǝ thǝpsi-ǝ-ù? tð
go-PFV-IRR-NR heaven-ANOM AGT discuss-PRES-DEC thus
sa

say.PST

“The heavenly beings are discussing they two who are to come to heaven, [they two] who will have already gone to the village of the dead”, said [Noksensangba].’

- 64 “*alà, aki itð tʃhàtʃhǝtpaj*”

alà a-ki i-tð tʃhà-tʃhǝt-pà? i
and NRL-house PROX-thus make-ABIL-NR PROX

“And so, this house that [I] am able to build like this”,

- 65 “*nàŋ nǝ kútá atsǝ.ɲ.*”

nàŋ nǝ kútá atsǝ-ǝ.ɲ
2SG AGT how look-PRES

“What do you think [of it]?”

- 66 *tðpà? “u, aki tʃu atsǝli? wai, pǝ.ɲ?”*

tð-pà? u a-ki tʃu atsǝ-lí? wa-ɲ
thus-NR EXCLM NRL-house DIST look-SIM go-SEQ

pǝ.ɲ?

horizontal.beam

Because he asked him that, [Noksensangba replied] “Well! While looking around the house, the horizontal beam ...”

- 67 *“aki ku pəɹəʔ tɨjuŋ ku, á-səŋ ípá? laj ku, səŋɿ kət-ù?”*
 a-ki ku pəɹəʔ tə-juŋ ku a-səŋ
 NRL-house LOC horizontal.beam NZP-be.central LOC NRL-wood
 ípá? la i ku səŋɿ kət-ù?
 EMPHAT TOP PROX LOC gnarly.wood possess-DEC
 [I noticed that] “In the house, in the centre of the horizontal beam supporting the roof, right here is a gnarly knot in the wood.”⁸
- 68 *“a, læptʃhətəɹ thùktʃukpə? táŋkusi nàŋ atʃu mələptʃukpàla aki .ɹuŋiw?”*
 a læp-tʃhət-əɹ thùktʃuk-pə? táŋkusi nàŋ a-tʃu
 INTJ cut-ABIL-SEQ throw.away-NR⁹ unless 2SG NRL-DIST
 mə-ləp-tʃuk-pàla a-ki .ɹuŋ-ì-ù?
 NEG-cut-PFV-COND NRL-house burn-IRR-DEC
 “Hey, unless you cut and throw [that part] away, if you don’t cut that out, the house will burn.”
- 69 *“mətʃhàpàla aki ku tsəŋla? là-ì-ù?”*
 mə-tʃhà-pàla a-ki ku tsəŋla? là-ì-ù?
 NEG-do-COND NRL-house LOC lightning strike-IRR-DEC
 “Or if you don’t do that, then lightning will strike the house.”
- 70 *“ajà, nàŋ tənɨk məɹəməɹ ùku!”*
 ajà nàŋ tə-nɨk məɹəm-əɹ ùku
 EXCLM 2SG RL-eye be.red-PRES PTCL
 “Ayaa! You’re just jealous!”, [says Imtisangba].
- 71 *“aki təkítəna táŋ ku nuŋla.”*
 a-ki təkɨ-təna táŋ ku nuŋ la
 NRL-house outermost.side just LOC not TOP
 “[The knot surrounded by wood grain] is not on the outermost side of the house”, [repeated Noksensangba].
- 72 *“aki tɨjuŋthì ku pí.”*
 a-ki tə-juŋ-thì ku ípá? i
 NRL-house NZP-be.central-INTNS LOC EMPHAT PROX
 “It’s right here in the centre of the house.”

8. The Ao believe that knots of wood in the timber of houses attract lightning strikes.

9. The nominalizer suffix *-pə?* is not marking a relative clause here, so I assume that this is a phonologically reduced form of the conditional converb suffix *-pàla*. Compare the full form in the second line of interlinearization.

- 73 *tətʃhɑɪ “hàj? tət tjuŋthì ku ípá? la i nàŋ nə tshətfəm əj”,*
 tət-ʃhà-əɪ hàj? tət tət-juŋ-thì ku
 thus-do-SEQ okay PTCL NZP-be.central-INTNS LOC
 ípá? la i nàŋ nə tshə-ʃəm əj
 EMPHAT TOP PROX 2SG AGT take-ADD SINCE
 And then, [he said] “Yeah, since you’ve additionally gone and put this
 [the knot with wood grain] right in the centre,” [lit. had it additionally
 take the place of the centre]
- 74 *“tìjàku, mà.nu.u? tət sa.ə atì.”*
 tìjàku mà-ə.nu-ə.ɪ-ù? tət sa-ə.ɪ atì
 consequently NEG-good-PRES-DEC thus say-PRES PTCL
 “That’s precisely why [I’m] saying that it’s no good at all.”
- 75 *tətʃhàku, tuŋnət tʃu atʃu ku áŋ əhju laŋlitəpə.ɪ.ə,*
 tət-ʃhà-ku tuŋnət tʃu a-tʃu ku táŋ a-hju
 thus-do-LOC.CV 3DU DIST NRL-DIST LOC just NRL-word
 laŋli-təp-əɪ
 reply-RECIP-SEQ
 And thus, the two of them just argue back and forth there,
- 76 *“nàŋ, itáj kəmtaŋsikə táŋ nàŋ nə nǐ sapa? tʃu mələnsi.ɪ áŋ liəŋ ká.tʃu.”*
 nàŋ i-tət táŋ kəmtaŋ-si-əkə táŋ
 2SG PROX-thus just be.careful-RPET-SIM just
 nàŋ nə nǐ sa-pà? tʃu
 2SG AGT 1SG.AGT say-NR DIST
 mələn-si-əɪ táŋ li-əŋ kátʃu
 change-RPET-SEQ just stay-IMP PTCL
 [with Noksensangba warning] “You just carefully change [it] again like
 this, as I told you also,”
- 77 *tət sa.ɪ nʊksənsaŋpa? əhlú wa.*
 tət-sa-əɪ nʊksənsaŋ-pà? a-hlú nə wa
 thus-say-SEQ PN-M NRL-field ALL go.PST
 And having said that, Noksensangba went to the field.

- 78 *təɪ, ajatàŋi tʃu ku nuksənsaŋpa? áhlù phinə tʃhuwaɪə.*
 tə-əɪ ajatàŋi tʃu ku nuksənsaŋ-pà? a-hlú
 thus-SEQ early.evening DIST LOC PN-M NRL-field
 phinə tʃhuwa-əɪ
 ABL emerge-PRES
 And then, in the early evening, Noksensangba returns from the field.
- 79 *“ahəj imtisàŋpà?, təzatənu minaŋ màw?”*
 ahəj imtisàŋ-pà? tə-za.tə-nu min-aŋ mà-ù?
 EXCLM PN-M RL-child.RL-child pity-IMP PTCL-DEC
 “Hey, Imtisangba! At least have pity on your children, won’t you!” [he says]
- 80 *“tʃu! ní sapa? sə athənsijaɪ tatʃu tʃu!”*
 tʃu ní sa-pà? sə athənsi-ja-əɪ tə
 DIST 1SG.AGT say-NR ANAPH assemble-CONT-PRES thus
 a-tʃu tʃu
 NRL-DIST DIST
 “See? [That which] I said is coming together. See that?” (a violent storm can be seen brewing down on the plain of Assam)
- 81 *“thaku thùŋiliw?”*
 thaku thùŋ-i liw?
 now+LOC reach-IRR might
 “[The storm] might reach us now.”
- 82 *tə, nuksənsaŋpa? nə saɪə hlà.*
 tə nuksənsaŋ-pà? nə sa-əɪ hlà
 thus PN-M AGT say-SEQ descend+go.PST
 Thus, Noksensangba said that and went down.
- 83 *“i, sitak, mìpuŋmaj sa? tʃhàjù? .ə.”*
 i sitak mə-puŋ-ma?-i sa? tʃhà-i-ù? .ə
 PROX be.correct NEG-be.good-CMPL-IRR PTCL be-IRR-DEC PTCL
 [Imtisangba says] “This is right. Something not good at all will happen.”

84 *tə.ɪə, tə.nì khə tə.za tə.nu tʃu, tùhɣàɪ ki nə ni.ɪə wa.ɪə ni.juktʃakɪ.ɪə,*
 tə-əɪ tə-nì khə tə-za.tə-nu tʃu
 thus-SEQ RL-wife CONJ RL-child.RL-child DIST
 tùhɣàɪ ki nə nì-əɪ wa-əɪ nì-uk-tʃak-i?-əɪ
 other house ALL lead-SEQ go-SEQ lead-INTO-RS-CAUS-SEQ
 And, after leading his wife and children into a neighbour's house and
 leaving them there,

85 *imtisàŋpà? tʃu, aji tʃu, tatsəŋ à atʃətlakəɪ,*
 imtisàŋ-pà? tʃu a-ji tʃu tatsəŋ à
 PN-M DIST NRL-rice.beer DIST bamboo.container one
 atʃətlak-əɪ
 squeeze-DESCEND-SEQ
 Imtisangba squeezed some rice beer into a cup,

86 *tʃəməkə, atʃuŋ kà?, kusu ku ami mətuŋtʃakəɪ,*
 tʃəm-əkə
 drink-SIM
 a-tʃuŋ kà? kusu ku a-mi mətuŋ-tʃak-əɪ
 NRL-shield also fireplace.shelf LOC NRL-spear be.erection-SEQ
 drank, propped the shield and the spear against the fireplace shelf,

87 *atʃuŋ kà?, anuk tʃu kà? h.ɪuŋsi.ɪə*
 a-tʃuŋ kà? a-nuk tʃu kà? h.ɪuŋ-si-əɪ
 NRL-shield also NRL-machete DIST also hone-RPET-SEQ
 the shield also, honed the machete

88 *həmətəɪ aji tʃu tʃəməkə pa mənəɪ là?*
 həmət-əɪ a-ji tʃu tʃəm-əkə pa mən-əɪ là?
 hold-SEQ NRL-rice.beer DIST drink-SIM 3SG sit-SEQ wait.PST
 and held [it] drinking the rice beer as he sat and waited.'

- 89 *təku, tsəmà? phinə, məpuŋ tʃu ip.ɿ tʃu kələmləm kələmləm kələmləm kələmləm məkə athənsi kəɿ.*
 tə-ku tsəmà? phinə məpuŋ tʃu ip.ɿ tʃu
 thus-LOC.CV Assam.plain ABL wind DIST storm DIST
 kələm kələm kələm-ləm kələm-ləm-əkə athən-si-əɿ
 ONOM ONOM ONOM-RED ONOM-RED-SIM shake-RPET-SEQ
 kəɿ
 ascend+come.PST
 And then, a violent storm ascended from the plain of Assam, rumbling and shaking everything as it came up.
- 90 *a tʃùku tsəŋtəphəla? tʃu nə nuksənsəŋpə? nə sapa? sə,*
 a tʃùku tsəŋtəphəla? tʃu nə
 er... consequently lightning DIST AGT
 nuksənsəŋ-pà? nə sa-pà? sə
 PN-M AGT say-NR ANAPH
 And consequently the lightning, just as Noksensangba had said,
- 91 *səŋnik səŋi tʃu ku “tʃəɿət tʃəɿət” tə phəlatshəŋəɿ.*
 səŋ-nik səŋi tʃu ku tʃəɿət-tʃəɿət
 wood-eye gnarly.wood DIST LOC ONOM-RED
 tə phəla?-tshəŋ-əɿ
 thus strike-ATTACH-PRES
 “Crack! Crack!”, [the lightning] strikes the the knot in the wood
- 92 *imtisəŋpà? tʃu tʃùku anuk sə həməɿ áŋ, kà? kà? kà? ləp-thəŋli? táŋ waɿ.*
 imtisəŋ-pà? tʃu tʃùku a-nuk sə həmət-əɿ
 PN-M DIST consequently NRL-machete ANAPH hold-SEQ
 táŋ kà?.kà?.kà? ləp-thəŋ-lí? táŋ wa-əɿ
 just ONOM cut-SEVER-SIM just go-PRES
 So Imtisangba, consequently just takes hold of the [aforementioned] machete and Chop! Chop! Chop!, [he] just cuts [the knot] out of the wood.
- 93 *təkukà?, tsəŋtəphəla? tʃu mihi.*
 tə-kukà? tsəŋtəphəla? tʃu mə-hi?-əɿ
 thus-CONCESS lightning DIST NEG-pass-PRES
 But even then, [i.e. after doing all that] the lightning strikes do not diminish.

- 94 *ləpsi ləpsi ləpsikə aki sə təlu? ləpsatʃuk.*
 ləp-si ləp-si ləp-si-əkə
 cut-RPET cut-RPET cut-RPET-SIM
 a-ki sə təlu? ləp-saʔ-tʃuk
 NRL-house ANAPH all cut-SEPARATE-PFV.PST
 Chopping, chopping, chopping, [he] chopped that whole house up.
- 95 *pa kà? nìtʃuk.*
 pa kà? nì-tʃuk
 3SG also be.tired-PFV.PST
 He was also exhausted.
- 96 *təi likə pái mäsəʔ puŋtʃa təpəti à, tsəŋi nə zakə tʃu tʃənəkə tʃhuwaɪə*
 tə-əɪ li-əkə pái mäsəʔ puŋtʃa tə-pəti à
 thus-SEQ be-SIM 3PL.POSS ox bull NZP-be.big one
 tsəŋi nə za-əkə tʃu tʃən-əkə tʃhuwa-əɪ
 rain ABL saturate-SIM DIST flee-SIM emerge-SEQ
 And while he was there, one of their family’s big bulls, saturated by the rain and fleeing the downpour, emerged [from the rain],
- 97 *kiləm tʃu nə hiɪaɪə*
 kiləm tʃu nə hiɪa-əɪ
 eaves DIST ALL level+come-SEQ
 and came over to the eaves of the house,
- 98 *mäsəla tʃu hjùŋli.*
 mäsəʔ-la tʃu hjùŋli
 ox-F DIST stand.PST
 and that bull¹⁰ stood [there].
- 99 *“ajà. tʃükü nì kà? jah.ɪtʃukukù.”*
 ajà tʃükü nì kà? jah.ɪ-tʃuk-ukù
 EXCLM consequently 1SG also be.one’s.fill-PFV-ANT
 “Aya! I’m completely fed up with this” [says Imtisangba].

10. All animals in Mongsen folklore stories are assigned feminine semantic gender, paradoxically even when their referents are clearly masculine.

- 100 *pa thaku “tʃɔ́pá? tʃhàmì.ɪə la tʃhəŋ!”*
 pa thak ku tʃɔ́pá? tʃhà-mì-əɪ la tʃhà-aŋ
 3SG PLACE LOC what do-DESID-PRES TOP do-IMP
 To it [i.e. the lightning, in order to placate it, he said] “Do whatever you want!”
- 101 *təpa?, tsəŋla? tətʃhatsə à lai tán, mäsə? la tʃu altijùŋ nə səmà.ɪə thəptʃaki?*
 tə-pà? tsəŋla? tə-tʃhatsə à là-əɪ tán
 thus-NR lightning NZP-be.dangerous one strike-SEQ just
 mäsə?-la tʃu alitijùŋ nə səmà-əɪ thəp-tʃak-i?
 ox-F DIST half INST split-SEQ drop-RS-CAUS.PST
 Upon saying which, a dangerous bolt of lightning just struck the bull, splitting and felling [it] in halves.
- 102 *“atʃu nə təməluy nəmaŋ!”*
 a-tʃu nə tə-məluy nəm-aŋ
 NRL-DIST INST NZP-be.willing be.satisfied-IMP
 “[May your heart¹¹] be satisfied with that!” [he said].
- 103 *təpa?, tsəŋi tʃùku ip.ɪu tʃu hməntsəw?*
 tə-pà? tsəŋi tʃùku ip.ɪu tʃu hməntsə-ù?
 thus-NR rain consequently storm DIST stop.PST-DEC
 Upon saying which, the rain and the storm consequently abated.
- 104 *tə, maŋmətuŋ ku imtisəŋpà? khə nuksənsəŋpà? hjutsə nət tʃu tʃutə satəpja?*
 tə maŋmətuŋ ku
 thus village.name LOC
 imtisəŋ-pà? khə nuksənsəŋ-pà? nət tʃu hjutsə
 PN-M CONJ PN-M two DIST story
 tʃu-tə sa-təp-ja?
 like.that say-COLL-HAB.PST
 Thus, [we] used to tell the story of Noksensangba and Imtisangba to each other like that in Mangmetong village.

11. The citation form for ‘heart’ is a nominalised verb-noun compound *tə-məluy tʃaŋ* (NZP-be.willing + seed). The verbal noun *təməluy* is perhaps most appropriately equated with ‘mind’.

3. The founding of Mangmetong Village

Speaker: Imdangchiba Imchen, aged 69 years

Recorded: July 2001 at Shillong, Meghalaya

The following unedited story is an oral history describing how Mangmetong village was established, and how it came to get its distinctive name, which literally means ‘upright corpse village’. Details concerning the speaker can be found in the introduction to Text 2. It is pertinent to mention that this is the history of the founding of Mangmetong Village according to the residents of the lower ward. Residents of the upper ward of Mangmetong assert a different account.

- 01 *maŋmətuŋ tə tʃapa? tʃu, lùŋkhə̀m phinə, mənaŋthì, lə̀mtu, aláŋ, lə̀mtu, puŋən, tsəti, tə nə ni.ə̀, əjim za.ɪ.ə̀ kəm.*
 maŋmətuŋ tə tʃa-pà? tʃu lùŋkhə̀m phinə mənaŋ-thì
 village.name thus call-NR DIST village.name ABL FRONT-INTNS
 lə̀mtu aláŋ lə̀mtu puŋən tsəti tə nə
 clan.name clan.name clan.name clan.name clan.name thus AGT
 nì-ə̀ɪ a-jim za-ɪ.ə̀ɪ kəm
 lead-SEQ NRL-village enter-come-SEQ become.PST

The village that is called Mangmetong first became a village after the Lemtu, Alang, Pongen and Tseti clans, leading from Lungkum village, entered the village [site] and founded it.

- 02 *tə, ípá? táŋ tʃu ku lùŋkhə̀m imsən tə tʃakə tʃhà.*
 tə ípá? táŋ tʃu ku lùŋkhə̀m jim-sən tə
 thus EMPHAT only DIST LOC village.name village-new thus
 tʃa-əkə tʃhà
 call-SIM COP.PST

And that particular place was called New Lungkum village.

- 03 *anə̀pən-pùpa?, khutʃha? nə ni.ə̀, jimlaŋ məphù nə hlá.ə̀, əjim kəmsi.*
 a-nət-pən-pùpà? khutʃha? nə nì-ə̀ɪ jim-laŋ
 NRL-two-NUM-ORD PN AGT lead-SEQ village-bottom
 məphù nə hlà-ə̀ɪ a-jim kəm-si
 ward ALL descend+go-SEQ NRL-village become-RPET.PST
 Secondly, led by Kucha, they went down to the [site of the] lower ward and again established a village.

- 04 *ípá? tʃu ku, aláŋ nə lə.ɿ.ə.ə*
 ípá? tʃu ku aláŋ nə lə.ɿ.ə.ə
 EMPHAT that LOC clan.name AGT descend+come-SEQ
 There, after the Alang clan came down,
- 05 *“inət səŋkhəm à táŋ tʃh.ə.ə, inət məh.ɿ.sà.láŋ lə̀jù?”*
 inət səŋkhəm à táŋ tʃh.ə.ə
 1DU.IN village.gate one just make-SEQ
 inət mə-h.ɿ.sà-la áŋ li-ì-ù?
 1PL.IN NEG-divide-NEG.CV only stay-IRR-DEC
 [they said] “We will make just one village gate and live undivided.”
- 06 *“təj, nəŋ tʃá? sisàj təkha”*
 tək aj nəŋ tʃá? sisà-ì tək-hà
 thus SINCE 2SG nothing worry-IRR PROH-have
 “Because of that, you’ve got nothing to worry about.”
- 07 *“inət khələm táŋ nə lə̀jù.ə.”*
 inət khələm táŋ nə li-ì-ù? ɿə
 2DU.IN together just INST stay-IRR-DEC PTCL
 “Just we two [clans] will live together.”
- 08 *tə, ípá? táŋ tʃu ku, ‘məh.ɿ.sà’.*
 tək-ə.ɿ ípá? táŋ tʃu ku mə-h.ɿ.sà
 thus-SEQ EMPHAT just that LOC NEG-divide
 And, that [village] was [named] ‘Undivided’.
- 09 *tə, lùŋkhəm imsən phisənə ‘məh.ɿ.sà ajim’ təkəm.*
 tək lùŋkhəm jim-sən phi <sə> nə
 thus village.name village-new <DIST>ABL
 mə-h.ɿ.sà a-jim tək kəm
 NEG-divide NRL-village thus become.PST
 And from New Lungkum village, [it] became [known as] ‘Undivided village’.
- 10 *tə, ípá? sin tʃu ku,*
 tək-ə.ɿ ípá? sin tʃu ku
 thus-SEQ EMPHAT BACK DIST LOC
 And then, after that,

- 11 *jitu niŋəi nə jàsà.na.u amaŋ hənə.ɬə.ə, maŋmətuŋ ku mətuŋtʃaki?*
 jitu niŋ-əi nə jàsà.na.u a-maŋ hənə.ɬə-əi
 village.name name-ANOM AGT PN NRL-body bring-SEQ
 maŋmətuŋ ku mətuŋ-tʃak-i?
 village.name LOC be.erection-CAUS.PST
 the people of Ritu Village brought the [smoked and dried] corpse of
 Yasa Naro and left it propped up in Mangmetong village.
- 11 *maŋmətuŋ tʃaki, túŋkhəla, kinijunəi hlə.ə kəm.*
 maŋmətuŋ tʃak-i?-əi túŋkhəla kinijunəi
 village.name leave-CAUS-SEQ 3PL village.name
 hlə-əi kəm
 descend+go-SEQ become.PST
 Having left [the corpse] in Mangmetong village, they went down and
 founded Kiniunger village.
- 12 *a.əmtʃhaŋ təluk a.əmtʃhaŋ.ə tʃà.ɬə jàsà.na.u amaŋ tʃu maŋmətuŋ ku
 mətuŋtʃaki.*
 a.əmtʃhaŋ tə-luk a.əmtʃhaŋ-əi tʃà-la jàsà.na.u
 PN RL-group PN-ANOM offspring-F PN
 a-maŋ tʃu maŋmətuŋ ku mətuŋ-jak-i?-əi
 NRL-body DIST village.name LOC be.erection-CAUS-PRES
 [As for] Aremchang's party, Aremchang's daughter Yasa Naro's corpse
 is kept upright there in Mangmetong Village.
- 13 *təpənə, ípá? phi <tʃu > nə táŋ maŋmətuŋ tʃu, “maŋmətuŋ”, tə tʃatʃhəttʃaki?*
 tə-pənə ípá? phi <tʃu > nə táŋ maŋmətuŋ tʃu
 thus-CAUSAL EMPHAT <DIST>ABL just village.name DIST
 maŋmətuŋ tə tʃa-tʃhət-tʃak-i?
 village.name thus call-ABIL-CAUS.PST
 Because of that, only from [that time] onwards was it able to be called
 “Mangmetong Village” (lit. “upright corpse village”).
- 14 *təi thàku təsi maŋmətuŋ tʃa.u?, atʃu.*
 tə-əi thaku təsi maŋmətuŋ tʃa-əi-ù? a-tʃu
 thus-SEQ now+LOC until village.name call-PRES-DEC NRL-DIST
 And, up to now it's [still] called Mangmetong.

- 15 *tə hjutsə tʃu, maŋmətuŋ hjutsə tʃu atʃu táŋú? wa.*
 tə hjutsə tʃu maŋmətuŋ hjutsə tʃu
 thus story DIST village.name story DIST
 a-tʃu táŋ-ù? wa
 NRL-DIST just-DEC PTCL
 And that's the story of Mangmetong village, as I recall.

4. The birds' conference

Speaker: Bendangmongla, aged 72 years

Recorded at: Mangmetong Village 1999

This unedited story was narrated by a woman who is a resident of the lower ward of Mangmetong village. The narrative gives a folkloric explanation for how the world came to be divided into periods of light and dark, how the owl came to have a flat face, and why the Yellow-backed Sunbird is such a small bird. A similar story is found in other, distantly related Tibeto-Burman languages, such as Rawang (Randy LaPolla, personal communication 2005) and Khamniungan. While the details and the characters may vary slightly from language to language, the story's plot remains sufficiently constant in these different languages to recognize a number of commonalities suggestive of a shared origin.

- 01 *kə tənín pəntán.muŋla.*
 kə tənín pəntán.muŋ-la.
 1SG.POSS RL-name PN-F
 My name is Bendangmongla.
- 02 *inə .jà.ə saŋ tə.*
 i nə .jà-ə sa-aŋ tə
 PROX ALLAT come-SEQ say-IMP PTCL
 [Interjector:] Come here and say [the story]!
- 03 *nì akəm nətəɪ nət.*
 nì a-kəm nətəɪ nət
 1SG NRL-year seventy two
 My age is seventy-two.

- 04 *saj tʃharú? la.*
 sa-ì tʃharú? la
 say-IRR QPTCL TOP
 Will I tell it?
- 05 *hjuṣə tʃu saj tʃharú?*
 hjuṣə tʃu sa-ì tʃharú?
 story DIST say-IRR QPTCL
 Will I tell the story?'
- 06 *akhi hmàpàṅ ku, wàzà? təmáṅi, muṅthaṅ à səntəp.*
 akhi hmàpàṅ ku wàzà? təmáṅ i muṅthaṅ à
 ancient time LOC bird all PROX conference one
 səntəp
 meet.PST
 Long ago, all these birds gathered for a conference.
- 07 *ípá? muṅthaṅ tʃu ku,*
 ípá? muṅthaṅ tʃu ku
 EMPHAT conference DIST LOC
 In the conference,
- 08 *wàzà? tʃu nə təpətípàla tʃu nə, “isa kà? pi səpá? nə ajimì,” tə səmtə.*
 wàzà? tʃu nə tə-pəti-’-pà? la tʃu nə
 bird DIST AGT NZP-big-SUP-NR TOP DIST AGT
 isa kà? pi səpá? nə ajim-ì tə səmtə
 1PL.IN also PROX who AGT cry.out-IRR thus ask.PST
 the bird that was the leader asked “Who among us here will make a
 proclamation?”
- 09 *təku, pùkphuḷa sə nə “nì nə ajim-ì w, ” tə sa.*
 tə-ku pùkphu-la sə nə
 thus-LOC owl-F ANAPH AGT
 nì nə ajim-ì w tə sa
 1SG AGT cry.out-IMM PTCL thus say.PST
 Upon asking that, Owl said “I will make a proclamation.”
- 10 *tətʃhai pa nə ajimpa? ... a,*
 tətʃhà-əi pa nə ajim-pà? a
 thus-do -SEQ 3SG AGT cry.out-NR INTJ
 And then, [that which] she proclaimed was ... ah,

- 11 *màṅlikà? màṅtʃón táṅ màṅaṅ, saṅalikà? saṅwátʃón saṅwaṅ*
 màṅ-likà? màṅ-tʃòn táṅ màṅ-aṅ
 be.dark-CONTEMP be.dark-DUR just dark-IMP
 saṅwa-likà? saṅwa-tʃòn saṅwa-aṅ
 be.light-DUR be.light-DUR be.light-IMP
 “If it is dark, just let it remain dark; if it is light, let it remain light.”
- 12 *tə sa pi.*
 tə sa pi.
 thus say.PST PROX
 She said this.
- 13 *tətʃhàku “wa! màrulaw? màrulaw?”*
 tətʃhà-ku wa mət-aru-la-ù? mət-aru-la-ù?
 thus-do-LOC EXCLM NEG-good-NEG-DEC NEG-good-NEG-DEC
 ‘And upon saying that [they all said] “Oh, Bad! Bad!”
- 14 *“pùkphula nàṅ màrulaw?”*
 pùkphu-la nàṅ mət-aru-la-ù?
 owl-F 2SG NEG-good-NEG-DEC
 “Owl – you’re bad.”
- 15 *tətʃha.ət wàzà? təmáṅ sə jaja.ət pa thàku jàk.*
 tətʃhà-ət wàzà? təmáṅ sə jaja-ət
 thus-do-SEQ bird all ANAPH be.angry-SEQ
 pa thak ku jàk
 3SG PLACE LOC beat.PST
 ‘And then, all the aforementioned birds got angry and beat on her.’
- 16 *tətʃhàku, pa tələm sə, íjá? puṅkə tʃhàpà?*
 tətʃhà-ku pa təl-əm sə íjá? puṅ-kə tʃhà-pà?
 thus-do-LOC 3SG RL-head ANAPH much be.good-SIM be-NR
 ‘And upon doing that, her aforementioned head, which was very beautiful,’
- 17 *àpakapakə kəmtʃuk, pùkphula tʃu.*
 apak-apak-kə kəm-tʃuk pùkphu-la tʃu
 be.flat-RED-SIM become-PFV.PST owl-F DIST
 ‘[the head of] that Owl became flattened.’

- 18 *tətʃhaɪ sin ku,*
 tətʃhà-əɪ sin ku
 thus-do-SEQ BACK LOC
 ‘Then after that,’
- 19 *tʃhəmpaŋ mətʃəmpùŋla tʃu ahən təpuŋ mətəm i təpət i à tʃhà.*
 tʃhəmpaŋ.mətʃəmpùŋ-la tʃu a-hən təpuŋ mətəm i
 yellow.backed.sunbird-F DIST NRL-fowl cock like PROX
 tət-pət i à tʃhà.
 NZP-be.big one COP.PST
 ‘there was Yellow-backed Sunbird, who was a big one, like a rooster.’
- 20 *tətʃhàku, “aj! nì nə ajim.ɹù ní,” tə sa.*
 tətʃhà-ku aj nì nə ajim.ɹù ní
 thus-do-LOC EXCLM 1SG AGT cry.out-IMM PTCL
 tət sa
 thus say.PST
 ‘And so, she said “Hey, will I make a proclamation?”’
- 21 *“hàj? hàj? hàj?” təpa?*
 hàj? hàj? hàj? tət-pà?
 okay okay okay thus-NR
 And because [the birds replied] “Yes, yes, yes”,
- 22 *“álíma tʃu ku mijimtʃàɪ tʃu mànli saŋali mànli saŋali tət ləju?” tət sa.*
 a-líma tʃu ku mijimtʃàɪ tʃu
 NRL-world DIST LOC humankind DIST
 màn-li saŋa-li màn-li saŋa-li
 dark-ALT.IT light-ALT.IT dark-ALT.IT light-ALT.IT
 tət li-ì-ù? tət sa
 thus be-IRR-DEC thus say.PST
 [she] said “In the world, humankind will have dark, light, dark, light
 infinitely alternating – it will be thus”.
- 23 *tətʃhaɪ, “u atʃu sitakù?”*
 tətʃhà-əɪ u a-tʃu sitak-ù?
 thus-do-SEQ EXCLM NRL-DIST be.correct-DEC
 And then [they said] “Oh! That’s right!”

- 24 *“íjǎ? puŋə̀ù? atʃu”*
 íjǎ? puŋ-ə̀-ù? a-tʃu
 much be.good-PRES-DEC NRL-DIST
 “That’s very nice!”
- 25 *“á nàŋ ta.rú.pàw? ni.”*
 a nàŋ tə-a.rú-pà?-ù? ni
 EXCLM 2SG NZP-good.SUP-NR-DEC PTCL
 “Ah! You’re really the best”.
- 26 *tətʃhə̀i wàzà? təmáŋ tʃu nə pa thàku mijmsipa?, tələm tʃu ku,*
 tə-tʃhə̀-ə̀i wàzà? təmáŋ tʃu nə pa thak ku
 thus-do-SEQ bird all DIST AGT 3SG PLACE LOC
 mijim-si-pà? tə-ləm tʃu ku
 stroke-RPET-NR RL-head DIST LOC
 And then, because all the birds stroked her on the head,
- 27 *“a.rú.ù? ni,” tə sapa?,*
 a.rú-ə̀-ù? ni tə sa-pà?
 good-PRES PTCL thus say-NR
 saying “It’s good!”
- 28 *pa tasù? à táŋ kəmtʃuk.*
 pa tə-asu? à taŋ kəm-tʃuk
 3SG small one just become-PFV.PST
 she became a very small [bird].
- 29 *tə̀i thakuj pa wàzà? tə̀pə̀ti à nə tasù? à taŋ kəmə̀i li.*
 tə̀-ə̀i thaku i pa wàzà? tə̀-pə̀ti à nə
 thus-SEQ now+LOC PROX 3SG bird NZP-be.big one AGT
 tə̀-asu? à taŋ kəm-ə̀i li-ə̀i
 small one just become-SEQ be-PRES
 And now she, a big bird, has become just a small one.
- 30 *tə̀i pùkphula tʃu wàzà? tipuŋ ta.rú à tʃhàpa?,*
 tə̀-ə̀i pùkphu-la tʃu
 thus-SEQ owl-F DIST
 wàzà? tə̀-puŋ tə̀-a.rú à tʃhà-pà?
 bird NZP-be.good NZP-be.good one do-NR
 And that Owl, the bird that was a beautiful one’,

- 31 *pa nə mə.ɹu.la aji.mpaʔ tʃu nə,*
 pa nə mə-ɹu-la aji.m-pàʔ tʃu nə
 3SG AGT NEG-be.good-NEG.CV cry.out-NR DIST AGT
 ‘she who made the bad proclamation,
- 32 *wàzàʔ təmáŋ nə jak-əɪ,*
 wàzàʔ təmáŋ nə jak-əɪ
 bird all AGT beat-SEQ
 ‘after all the birds beat [her]’,
- 33 *pa tələm tʃu apakə kəm-əɪ,*
 pa tə-ləm tʃu àpak-apak-kə kəm-əɪ
 3SG RL-head DIST be.flat-RED-SIM become-SEQ
 ‘and her head became flattened’,
- 34 *pa thàku i sù.ɹu məpuŋla lì.*
 pa thaku i sù.ɹu mə-puŋ-la lì-əɪ
 3SG now+LOC PROX visage NEG-be.good-NEG.CV be-PRES
 now she is now no longer beautiful.
- 35 *tə hjutsə wa.*
 tə hjutsə wa
 thus story PTCL
 Thus, that’s the the story as I remember it.

colour səntsəʔ
copper, brass jùŋpənjin
country, land lima (‘earth’ + FACE) (<PTB *mliy)
egg həntsə
dirt, rubbish tʃànàʔ
dust uphuŋ
fire məzəʔ (variant: mi)
flame mi-lak (‘fire’ + TERM)
footprint tʃaŋ-sə (‘foot/leg’ + ?)
hail [a]-zu
hole kupaʔ
iron [a]-jin
parting and rejoining of a stream around island or other obstruction [a]-sət
jungle [a]-uŋ
lightning thunderbolt tsəŋtəphəlaʔ
lower range aliŋ
kind i
marsh tsə-thàʔ (‘water’ + ?)
moon làtà (<PTB *s-la)
mountain tónəm
mud līsəʔ
oil thu-tsə (? + ‘water’) (probably <PTB *sa-w)
path ənti (variant: ìnti)
place tʃākà (<Nagamese jagā)
plain intən
pond tsə-pàʔ (‘water’ + ?)
power azən
precipice khəm (<PTB *r-ka[.]m)
rain tsəŋi
rainbow puŋjaŋsən
ravine taluk
river [a]-juŋ
river bank juŋ-khəm (‘river’ + ‘precipice’)
rust [a]-tshəŋ
sky [a]-niŋ
slope wàpət
smoke mukhuli (<PTB *kuw)

star pətinu
stench of a rotting carcass [tə]-sə-nəm (‘die’ + ‘smell’)
stone [a]-luŋ (<PTB *r-luŋ)
storm ipu
stream [a]-juŋ-za (‘river’ + DIM)
sun tsəhŋi
sunray tsəhŋi-phu (‘sun’ + ‘expose to sun’)
thunder tsəŋmuk
valley tələmuk
water [a]-tsə
water, drop of tsə-tʃaŋ (‘water’ + ‘seed’)
wind məpuŋ
world, earth [a]-lí-má (‘earth, soil’ + FACE) (<PTB *mliy)

3. Time

ancient (times) akhì
century phusù
cool season tsəŋkəm
day niŋhuŋ
one day tsəŋikhu, níʔ
every day, always tətthì
from that day onwards tʃu-ni-phi- <tʃu> nə (DIST + ‘day’ + <DIST> ABL)
June/July tʃalipaŋ
night [a]-ja
now thaku
October tʃak-hən-paʔ làtà (‘paddy’ + ‘carry’ + NR + ‘month’)
olden (times) aləmi
summer, monsoon season [a]-ləm-jim (‘be warm’ + ‘season’)
time hmapaŋ
winter [a]-tsək-jim (‘be cold’ + ‘season’)
year [a]-kəm

See Tables 4.13 and 4.14 for
additional time words

4. Animals – land

animal sà.ɪ.à.ɪə
bat aləptì
bear ɪ.əm
buffalo (water) [a]-paŋ (<PTB
 *braŋ)
bull puŋtʃa
calf mäsə-za ('ox' + DIM)
cat (domestic) mùni
deer (generic term) mətshə
dog (domestic), [a]-ji
dhole (*Cuon alpinus*) tsóhŋì
duck phàtāk (<Hindi *batak*)
elephant sà-ti (probably 'animal'
 + 'big')
fox, Indian (*Vulpes bengalensis*)
 tshəluŋ
gaur tsəmuŋ
goat nəpùŋ
hoolook gibbon sùmi
horned wild animal, male jaŋ.ɹu
horned wild animal, female tətsəʔ,
 tsəla
house gecko kì-tsəŋ-ti ('house' +
 INSIDE + 'older sibling')
Indian pangolin kələp
leopard tʃuŋraŋkhù
leopard cat hən-sələj ('chicken' +
 'take stealthily')
lion mulutsəŋi
lizard sàŋphila
mithun (*Bos frontalis*) [a]-tshə
macaque (Assamese, pig-tailed)
 sàŋatʃa
macaque, rhesus tʃaŋti
ox (domestic) mäsəʔ
pig, domestic [a]-úk

pig, domestic boar [a]-úk təhla
pig, domestic sow [a]-úk [tə]-ji
 ('pig' + 'mother')
pig wild pùŋì
python alisàʔ
rabbit inisaʔ
rat, bamboo hitʃaʔ
rat [a]-hiʔ
rhinoceros, Indian uŋtuk
slow loris sùwì
snake [a]-sə
squirrel (tree shrew) thaŋtsə
tiger [a]-khu

5. Animals – water

catfish (*Macropteronatus magur*)
 kutilà
crab tʃàŋpàŋ
crocodile (gharial) tshə.ɪ.áŋ
eel kùŋsija
fish (generic term)[a]-hŋáʔ (<PTB
 *ŋya)
fish scale hŋa-kəp ('fish' + 'skin')
frog [a]-khi
mahseer tʃùŋtʃàŋ
otter si.əm (<PTB *s-ram)
prawn [a]-kuŋ
tadpole hŋa-tsə-laj ('fish' +
 'water' + ?)
toad tsəŋi rakhì

6. Birds

bird wàzàʔ (<PTB *wa)
blue-throated barbet ulutʃa
cuckoo wakù wakù (probably
 onomatopoeic)
crow, raven wah.ɹu
domestic fowl, chicken [a]-hən

domestic fowl chick [a]-hən-za

('chicken' + DIM)

dove ahlútəmtəm**dove, spotted** kìmətshə**eagle** [a]-hmu**great pied hornbill** tóhnəm**great pied hornbill cock** tóhnəm

wapaŋ

great pied hornbill hen tóhnəm

wàtsə

long-tailed broadbill khəliŋli**owl** pùkphu**parakeet** uh.ru?**peacock** tʃùti**pigeon** təplulu**raquet-tailed drongo** ùjim**red junglefowl** ùphəla**rooster, cock** [a]-hən təpuŋ**woodpecker** ulən**yellow-backed sunbird** tʃhəmpaŋ

mətʃəmpùŋ

7. Insects, arachnids and invertebrates

ant sp., small hjuŋza**ant sp., large** mi.ri**bed bug** məna?**bee** (generic term) [a]-na?**bee, honey** nuŋaŋ**borer insect** (possibly a species of*Dermestes*) hùtsə?**cicada** tʃaŋtʃaŋ zíŋzínlā (probably

'insect' + ONOM + RED + ?)

centipede sp. jàwŋ**cockroach** ləph.u**earthworm, intestinal worm** tsəlà?**flea** [a]-hli (<PTB *s-*liy*)**firefly** [a]-ja tsək tha ('night' +

'ignite' + ?)

fly maŋaphu**grasshopper** tʃhaŋkú**insect** (generic term) tʃaŋtʃaŋ**leech, land** hməntsə?**leech, water** mənət (<PTB*(*m-*)*li-t*)**louse**, [a]-tshək (<PTB *s-*rik*)**maggot, lava** məsən**mosquito** mə.ila**nit** hlutsə**scorpion** tʃaŋpaŋ khu**silkworm** lùtha**spider** sù.rāk**stick insect** sə.rùŋi**termite ant** tʃhàmà?**termite, winged** [a]-hluŋ

8. Plants

aspidistra.sp am**bamboo sp.** má?**bamboo sp.** [a]-u**bamboo sp.** [a]-hwá?**banboo sp.** luŋmi?**bamboo sap** hwà-tsə ('bamboo' + 'water')**banana** maŋu**bark** səŋ-kəp ('wood' + 'skin')**bean, poor man's** (*Dolichos lab lab*) làpàki**branch** khuŋsaŋ**bud** [tə]-mətsə**cane, rattan** [a]-hù (<PTB **rey*)**cane strip** (used for tying) [a]-liŋ**cotton** khəmphá**flower** hna.u**fruit** səŋ-tʃaŋ ('wood' + 'seed')**gnarly wood** səŋ-ji ('wood' + ?)**gourd sp.** aku?**grass** [a]-za**gum** [a]-hŋi**husk** wana

knot in wood səŋ-nik ('wood' + 'eye')
leaf səŋàwà
moss, lichen, mould maŋɹaŋ
mushroom kùŋhɹə
paddy [a]-tʃak
root [tə]-zəŋ
sap səŋ-tsə ('wood' + 'water')
seed tʃaŋ
seedling (rice) hmatsəʔ
sesame təsu
sprout niksə̀m
stem, trunk [tə]-tuŋ
stick [a]-tʃhuk
stump of tree hmə̀mpaŋ
sword bean creeper (*Entada scandens*) khəlɪŋ
thorn or splinter ahiŋ
tree səŋtuŋ ('wood' + 'stem, trunk') (< PTB **siŋ*)
vine hli
wood [a]-səŋ
yam [a]-mi

9. Human relationships

daughter tʃàla ('offspring' + F)
family kì-tsəŋ ('house' + INSIDE)
father [tə]-paʔ (< PTB **pa*)
father-in-law [tə]-ku
father's younger brother [tə]-pá-nu-za-paʔ ('father' + 'younger sibling' + DIM + M)
father's older brother [tə]-pá-zəm-paʔ ('father' + 'be senior' + M)
grandchild səmtʃàɪ
grandfather (maternal) [tə]-u
grandmother (maternal/paternal) [tə]-zə

husband [tə]-pá-záʔ ('father' + DIM?)
maternal uncle [tə]-ku
mother (common) [tə]-ja
mother (varies according to clan) [tə]-jiʔ, ùtʃa
mother-in-law (different phratry) [tə]-jánuzaj
mother-in-law (same phratry) [tə]-tì
mother's younger sister [tə]-jánuzaj
mother's older sister [tə]-jázəni
older (brother, uncle) [tə]-zəmpaʔ ('be senior' + M)
relative kì-tsəŋ-əɪ ('house' + INSIDE + ANOM)
son tʃà-(ə)ɪ ('offspring' + ANOM)
son-in-law [t(ə)]-əmhnaŋ
sibling – younger [tə]-nu
sibling – older [tə]-tì
wife təni
youngest female sibling [tə]-nuzajlà
youngest male sibling [tə]-nu-za-paʔ ('younger sibling' + DIM + M)

10. Types of people

ancestor, forebear nùŋsə̀ʔ niŋ-əɪ ('ancient' + 'name' + ANOM)
Alang clan aláŋ
bachelor tʃhaŋati
boy anu-tʃhaŋ ('child' + MAN)
bride kijəmlà
child [a]-nu
children [tə]-za-[tə]-nu
Chungli tʃùŋlì
clan kìn
friend jìmlà

friend, male [tə]-təm-paʔ ('like' + NR)

fool tʃhúwáɾə

humankind mi-jim-tʃàɪ (possibly 'person' + 'village' + 'child')

Imchen clan imtʃən

infant [a]-nu-za ('child' + DIM)

Jamir clan tʃàmìɪ

king tʃuphà (< Tai *caw faa* 'prince, ruler')

Kuki tribe luŋkhàj

Lemtu clan ləm̩tu

Lotha tribesman tsən̩.ɾə (pejorative); hlùtha niŋ-əɪ ('Lotha' + 'name' + ANOM)

lover hjaŋaj

man apanʃtʃhaŋ-əɪ ('man' + ANOM)

Mongsen mùŋsən

neighbour ki-taŋ-əɪ ('house' + SIDE + ANOM)

old person [a]-zənti

other tùhŋàɪ

owner kiphuɪ, kijupàʔ

person [a]-miʔ

plains-dweller, Indian tsəmàʔ niŋ-əɪ ('Assam Plain' + 'name' + ANOM)

Pongen clan puŋən

population mi-tʃàŋ ('person' + 'seed')

servant kì-li-(ə)ɪ ('house' + 'stay' + ANOM)

slave alàɪ

soldier sipaj (< Hindi *sipāhī*)

taxi driver taksi nì-tʃhaŋ ('taxi' + 'lead' + MAN)

thief awəɪʔ

Tsəti clan tsəti

unmarried woman, maiden lazati

village headman kawnpuaɪ (< Nagamese *gaun bura*)

villager [a]-jim ku li-(ə)ɪ ('village' + LOC + 'stay' + ANOM)

villager jim-əɪ ('village' + ANOM)

woman ánítì

11. Body parts and fluids – human and animal

abdomen, belly, stomach [tə]-puk (< PTB **puk*)

arm, hand [tə]-khát

armpit [tə]-tʃha-pùk ('wing' + STOMACH)

back [tə]-sin

bile [tə]-sə

blood [a]-jíʔ (< PTB **s-hwiy*)

body [tə]-maŋ

bone [tə]-læt (< PTB **rus*)

bowels [tə]-hrə

brain [tə]-kuluk

breast mama

breath [tə]-saka (< PTB **sak*)

calf [tə]-mu

chest [tə]-kùʔ

chin [tə]-mákháʔ (< PTB **(m)-kha*)

corpse [a]-maŋ (cf. 'body')

ear [tə]-hna.ruŋ (< PTB **r-na~*g-na*)

elbow [tə]-tsəsəlàj

eye [tə]-nik (< PTB **mik~*myak*)

face [tə]-ma

faeces [a]-tsóʔ

feather [tə]-tʃha

finger [tə]-mijunʃ (< PTB **[m]-yuuŋ*)

flesh, meat [a]-sáʔ (< PTB **sya*)

foot, dorsal aspect [tə]-məpha

forehead [tə]-nik khəm ('eye' + 'precipice')

fork of legs kàjà (probably < PTB
**kap*)
gills hɣàpà?
goitre [tə]-wəŋ
gums phá-t[hám ('tooth' + ?)
hair (body) phì.lùwà
hair (head) luma
hand, arm [tə]-khát
head [tə]-ləm
heart [tə]-məluŋ-tʃəŋ ('be willing'
+ 'seed')
horn [tə]-ji
intestines [tə]-h.ɪə
jaw (molar teeth) pha-ɪət ('tooth'
+ 'bone')
joint [tə]-luktəp
kidney [tə]-mækɪtʃəŋ (probably
< PTB **m-kal*)
knee [tə]-mækuk (< PTB **m-kurk*)
left hand ajtʃa?
leg, foot [tə]-tʃəŋ
lip paŋ-kəp ('mouth' + 'skin')
liver [tə]-məsən (< PTB **m-sin*)
lungs [tə]-məsə
marrow [tə]-liŋla (< PYB **kliŋ*)
mole (on skin) ləph.lu-tsó?
('cockroach' + 'faeces')
mouth, beak [tə]-paŋ
name [tə]-niŋ
nail (finger, toe), claw [tə]-məzəŋ
(< PTB **m-tsyen*)
navel [tə]-phəla (< PTB **la-y*)
neck [tə]-khuŋ
nose [tə]-na? (< PTB **s-na*)
nostril na-ku? ('nose' + 'hole')
pair (of hands, eyes, etc) anə-zə
(probably < 'two' + ?)
palm [tə]-mija
penis [tə]-ku
phlegm hɣaksák
right hand atʃaj
rump [tə]-sa-mi ('meat' + 'tail')

shadow, reflection [tə]-hmila
shin [tə]-məzən
shoulder [tə]-uphaŋ
sinew [tə]-hlu
skin [tə]-kəp
snot, watery na-tsə ('nose' +
'water')
snot, dried na-tsó? ('nose' +
'faeces')
sole of foot tʃəŋ-pùk ('foot' +
STOMACH)
spittle mətsə (< PTB **m-tʃriɪ*)
stomach (the visceral organ)
[tə]-puk làŋ
tail [tə]-mi (< PTB **r-may*)
thigh [tə]-phi
thumb [tə]-jì-pà? ('mother' + NR)
tongue [tə]-məli (< PTB
**m-lay~*s-lay*)
tooth [tə]-pha (< PTB **s-wa*)
urine hjuŋ-tsə (? + 'water')
vein [tə]-hli
vulva [tə]-ka
waist, middle (specifically, section
between lower ribs and hips)
[tə]-ɪəm
wing [tə]-tʃha
womb [a]-nu mən-tʃən ('child' +
'sit' + LNOM)

12. Life, death, illness and injury

birth [tə]-sù-pà? (NZP + 'bear' +
NR)
blind [tə]-nik [tə]-puk ('eye' +
'stomach')
boil, carbuncle khuŋmətʃəp
corpse [a]-maŋ, [tə]-sə-maŋ ('die'
+ 'body')

deaf [tə]-hna:uŋ tə-tʃak ('ear' + NZP + 'break')

death tə-sə (NZP + 'die')

diarrhoea [tə]-puk wà-pàʔ ('stomach' + 'go' + NR)

disease [tə]-tʃha-tʃən ('be hot' + LNOM)

drug, medicine muli

influenza jimhrà, maŋa

illness tʃha:u-paʔ ('be sick' + NR)

lame [tə]-tʃaŋ tə-tʃak ('leg' + 'NZP + break')

life tə-kəm (NZP + 'be alive')

mute [tə]-məli tə-tʃak ('tongue' + NZP + 'break')

poison məzəm

prickly heat mósák.ɿà

pus ahna-tsə (? + 'water')

malaria h.iamən

sting (n.) [tə]-mətak-paʔ ('sting' (v.) + NR)

wound, sore on skin khùma (probably < PTB **r-ma*)

13. Human activity, experience and belief

afterlife village, village of the

dead [a]-sə-jim ('die' + 'village')

accident ləntuŋ

beauty tuɿ

dream aja-maŋ ('night' + 'believe')

fortune, luck thija

god, deity tsəŋh.ɿəm

god of creation lítʃápáʔ

god of the dead mutsuŋ

heaven kúták

heavenly being, angel kúták-əɿ ('heaven' + ANOM)

manner of doing mətəm

sacrifice h.ɿəntən

soul [tə]-hmila

were-tiger, lycanthropist [a]-khu-hmila ('tiger' + 'soul')

worship khələm

14. Food

bamboo shoot [a]-zi

bamboo shoot, fermented itʃak

bean (generic term) hluli

beef mäsə-sàʔ ('ox' + 'meat')

bird meat wàzà-sàʔ ('bird' + 'meat')

biscuit, roti (incl. any flour-based food item) [a]-hjim

bitter gourd kha.ɿa

capsicum mə.ɿətsə khəmùŋlà

chicken meat [a]-hən-sàʔ ('chicken' + 'meat')

chilli mə.ɿətsə (< Hindi *mirich*)

corn cob mənti

corn seed mənti-tʃaŋ

crop sənaj

cucumber hmatshə

curry əntʃhu

dog meat [a]-ji-sàʔ ('dog' + 'meat')

duck meat phàtək-saʔ ('duck' + 'meat')

fern, edible sp. khuŋà

food (generic term) tʃuŋi

food, taboo tʃuŋi tə-mənən ('food' + NZP + 'be dirty')

garlic, white (*Allium sativum*)

tsəmà-ɿ lasəŋ ('Assam Plain' + ANOM + 'garlic')

garlic, red saŋtəm lasəŋ ('Sangtam tribe' + 'garlic')

ginger səŋmúk

guava mutiɾəm
honey na-kù-tsə (probably ‘bee’ + ‘hole’ + ‘water’)
jackfruit pùluŋ
milk mama-tsə (‘breast’ + ‘water’)
mustard (the plant) an
oil thu-tsə (? + ‘water’)
papaya taləpɦu
pineapple tɦuɾu
pork áwk-sà? (‘pig’ + ‘meat’)
passionfruit hənsə-làsì (probably ‘egg’ + ?)
potato alu (< Hindi *alū*)
pumpkin máphú?
rice(cooked) [a]-tʃa
rice(husked, uncooked) [a]-tʃaŋ
rice beer [a]-ji
rice wine mətʃəmtsə
rind (skin) [tə]-kəp
salt mətsə
soyabean, fermented tʃàmisə?
sticky rice (uncooked) masutʃaŋ
sugar cane mùtsə
tea sàŋà
tiger meat [a]-kɦu-sà?
water, boiled tsə-məlu (‘water’ + ‘boil’)
venison mətshə-sà? (‘deer’ + ‘meat’)
yam, fermented mijasə?

15. Weapons and warfare

arrow litʃak-tʃaŋ (‘catapult/bow’ + ‘seed’)
bullet in-tʃaŋ (‘iron’ + ‘seed’)
catapult/bow litʃak (< PTB **d-liy*)
enemy aɾiɾə
fight (n.) hũŋtəp-pa? (‘fight’ + NR)

firearm pəɾiəŋ pəɾiəŋ (ONOM + ECHO)
human head trophy maŋku
machete (*dao* in Nagamese)
 [a]-nuk
panji [a]-tʃɦu (< PTB **tsow*)
shield [a]-tʃɦuŋ
spear [a]-mi
spent cartridge shell intʃaŋ pùŋ
sword nuk-láŋ-nùk (‘machete’ + ‘be long’ + ‘machete’)
war tuŋpaŋ

16. Cultural artefacts, possessions

attire supùj sùlum (attire + ECHO)
ivory armlet sa-ɾət khəmpaŋ
 (‘animal’ + ‘bone’ + ?)
axe aw
bag hjaŋkɦu
bamboo container tatsəŋ
bamboo cup mà.rùk
bamboo thong, tie [a]-liŋ
bamboo wall or matting, woven
 tsəŋti
bamboo water vessel takzi
basket məlùk
basket, loosely woven bamboo
 (with square base, used to carry wood, water, produce etc.) aku
basket, tightly woven bamboo
 (conical with pointed base, used for carrying grain) akɦi
bed jip-tʃəŋ (‘sleep’ + LNOM)
bedstead ləm-taŋ (‘head’ + ?)
boat [a]-ɾuŋ (< PTB **(m-)loŋ*)
cage kɦaŋ (< PTB **kru-ŋ*)
car kà.ɾi (< Hindi *kāṛī*)
chair mən-tʃəŋ (‘sit’ + LNOM)

clock, watch tsəhŋi-[tə]-ji ('sun' + 'mother')

clothing sùsi

comb ləmsìrù

crab trap tʃəŋpəŋ khaŋ

cover tə-nəm-phəŋ-pà? (NZP + 'press' + COVER + NR)

flute pənpa?

fork phəŋtʃiə

gift səmpət

gold hun (< Assamese *hon*)

hat, head-dress kùh.iəŋ

hoe [a]-tʃəŋ

horizontal beam of house pə.ɹə?

knife kuta.ɹi (< Nagamese *kotari*)

ladle tatʃhət

log drum tuŋ-tən ('trunk' + 'sing')

money [a]-tʃhən

needle impən

paddy hulling table tʃhəmku

pestle mitəm

plank səŋ-pak ('wood' + 'be flat')

plate jəpu

razor tə-tʃhi-pa? (NZP + 'shave' + NR)

rice bin [a]-tʃha

sarong, woman's (*mekhela* in Assamese) [a]-ni?

shawl [a]-sə

shoe tʃəŋ-tʃhəm ('leg/foot' + 'wear')

slab luŋ-pak ('stone' + 'be flat')

spindle pəŋthəm.ɹu

stuff usət

sword, weaver's ahnəm

thing, possession inthaŋ

tobacco pipe mu-khuŋ ('smoke' + ?)

torch, firebrand mìlən

tongs mǐjǐptʃu.ɹu

wages thi.jà?

wages, contract hjajaŋ

water dipper [a]-tsə [t]-ətʃhət-pà?
mà.ɹùk ('water' + 'ladle' + NR
+ 'bamboo cup')

well tsə-pà? ('water' + NR)

whetstone luŋzak

17. Buildings, structures and locations

bachelors' house (*morung* in Nagamese) ah.ɹitʃu

boundary, border aɹ?

corner (of house, room) kìnìk

courtyard kì-mà ('house' + FACE)

door ki-tʃhi? ('house' + 'close')

east tsəhŋi tʃhuwa-tʃəŋ ('sun' + 'emerge' + LNOM)

eave of house kiləm

fence ikhu

field [a]-hlú

field hut [a]-thi

fireplace mǎzǎ-phu-tʃəŋ ('fire' + 'blow' + LNOM)

fireplace shelf kusu

floor [a]-təm

garden ikhu

granary [a]-thi

house [a]-ki

hulling room of house thijun

market pazaɹ (< Hindi *bāzār*)

path wa-tʃəŋ ('go' + LNOM)

pit phaŋ-tsəŋ ('cover' + INSIDE)

path ənti (variant: ìnti)

road jìmaŋ

road (main) im-tuŋ ('village' + 'stem/trunk')

path (middle of) ənti-ɹəm ('road' + 'waist/middle')

tea garden səŋàni

village [a]-jim

village gate səŋkhəm (probably
‘wood’ + ‘precipice’)

village name (Kabza) khapza

village name (Kiniunger)

kinìjuŋəɪ

village name (Lungkum)

lùŋ-khəm (‘stone’ +
‘precipice’)

village name (Mangmetong) maŋ-
mətuŋ (‘body’ + ‘be erect’)

village name (Ritu) .ritu

village name (Chungliyimti)

tʃùŋlì-jim-ti (‘Chungli’ +
‘village’ + ‘be big’)

ward (*khel* in Nagamese) məphù

ward, lower im-laŋ məphu
(‘village’ + ‘bottom’ +
‘ward’)

ward, upper im-paŋ məphu

(‘village’ + ‘mouth’ + ‘ward’)

west tsəhŋi-wa-tʃən (‘sun’ + go +
LNOM)

18. Cultural terms – community, tradition

blessing tə-muwaì (NRP + ‘bless’)

conference, meeting muŋthaŋ

custom jim sə

fine (of a pig) jim-hju-thaŋ-la
(‘village’ + ‘word’ + ‘sever’
+ ?)

festival pəntʃùŋ

folkdance jaɪə

generational age group hjaŋə

Greetings! saləm (< Arabic *salām*)

judgement pəntaŋ

law thəmzəŋ

price tʃhən-tʃàŋ (‘money’ +
‘seed’)

share saləm

tattoo [a]-jak

wedding sati-pa? (‘wedding’ +
NR) (< Hindi *sādī*)

work hməpa

19. Language and communication

book kákát (< Hindi *kāgaz*
‘paper’)

example tatisì

language hju-sì (‘word’ + ?)

letter sīti (< Hindi *ciṭṭhī*)

letter of alphabet hju-tʃaŋ (‘word’
+ ‘seed’)

lie (n.) tə-məɪak (NZP + ‘lie’)

meaning tətəzə

no mə?

okay hàj?

song [a]-tʃhu

tale, story hju-tsə (‘word’ + ?)

voice hju-la (‘word’ + ?)

word [a]-hju

yes hàw?

20. Verbs – stative

All stative verbs with (NZP) preceding the root require the nominalizing prefix *t(ə)-* to derive a deverbal adjective.

be, exist, stay li

be (past tense) tʃhà (variant: litʃhà)

be abundant pulu

be accidental phələk

be angry jaja

be arrogant ŋasaŋtəm

be awake sisa

be bad tamá.ru (< NZP- + NEG + 'be good')
be bald ləm-tshè? ('head' + 'pull')
be boastful atshəkhəm
be big (NZP-) pəti (possibly < PTB *tay)
be bitter (NZP-) khá? (< PTB *ka)
be bitterly afraid kha-tsəpha ('be bitter' + 'fear')
be black (NZP-) nák
be blue, be green (NZP-) nən
be brown (NZP-) nák-ɪəm ('be black' + 'be red')
be careful kəmtəŋ
be central (NZP-) juŋ
be certain (NZP-) tʃəŋtʃa
be clean (NZP-) məɪək
be clear (NRP) tʃəŋtʃà
be cold (bodily sensation) atshək
be cold (NZP-) məkuŋ
be complete pən
be correct sitak
be cunning (NZP-) kələk
be dangerous (NZP-) tʃhatsə
be dark (NZP-) məŋ (< PTB *mu-ŋ)
be deep (NZP-) aruk (possibly < PTB *tu-k)
be deliberate (NZP-) asá?
be distressed (NZP-) tʃàsi
be dirty (NZP-) mənən
be dry (NZP-) kuŋ
be enough (NZP-) pə.ɪ
be erect, upright (NZP-) mətuŋ
be excessive (NZP-) ali
be fast (NZP-) azəŋ
be fitting (NZP-) jái
be flat (NZP-) apak
be full səŋ
be good (NZP-) a.ru, (NZP-) puŋ
be happy (NZP-) pəla

be hard, be stiff (NZP-) məɪəŋ
be heavy (NZP-) atʃəŋ
be hot (NZP-) tʃha (< PTB *tsa)
be humorous (NZP-) məni-mì ('laugh' + DESID)
be hungry hjim
be illuminated (NZP-) asaŋ
be intoxicated tʃhà?
be late mənu
be light (illuminated) (NZP-) saŋa, saŋwa
be like mətəm
be long (NZP-) hləŋ
be lovable (NZP-) mijim-mì ('love' + DESID)
be near áhná (possibly < PTB *ney)
be old (NZP-) tʃán
be one's fill (NZP-) jah.ɪ
be pale (NZP-) pakla
be plump (NZP-) amət
be pure (NZP-) mətʃa
be putrefied, rotten, decayed (NZP-) əp
be quick aju
be quiet atʃhəm, (NZP-) səŋtʃəm
be ready ɪənəm
be red (NZP-) məɪəm
be ripe (NZP-) hniŋ (< PTB *s-min)
be rough (NZP-) ɪaɪa?
be round (NZP-) luŋluŋ (< PTB *zlum)
be satisfied nəm
be senior (NZP-) zəm
be shy, ashamed ahjak (< PTB *g-yak)
be short (NZP-) tshə
be sick tʃha.ru (possibly < PTB *tsa)
be small (NZP-) asu?
be smooth (NZP-) səsə

be soft (NZP-) nəp
be sour (NZP-) sən
be strong (NZP-) tsəʔ
be sweet (NZP-) mijəŋ
be swollen wək
be talented (NZP-) tʃhatsə
be tender (NZP-) anik
be thick (NZP-) mələm
be thirsty tsəh.ɪə
be tired (NZP-) phələm
be tired nì
be true, correct sitak
be warm (NZP-) ləm (<PTB **lum*)
be wet (NZP-) mətsə
be white (NZP-) məsəŋ
be willing (NZP-) məluŋ
be wrong (NZP-) aj

21. Verbs – sense and cognition

believed məŋ
deceived asiʔ
feared tsəpha
felt pain ɪək
forgot alak (possibly <PTB **b-lap*)
heard jàʔ
knew (something, someone etc)
 mətət
knew (how to do something) si
learned si-lak (probably <‘know’
 + TERM)
listened hɲa (<PTB **g-na*)
longed for juŋja
looked atsə
remembered, realized philəm-
 tʃhət (‘think’ + ABIL)
saw hùŋ
smelled məhnəm (<PTB **m-nam*)
thought philəm
touched taŋtʃhàʔ

understood hɲa-tʃhət (‘listen’ +
 ABIL)
wondered tsəməŋ

22. Verbs – activity

abandoned, left tʃuk (variant: tʃak)
adhered nəp
affixed mənaʔ
angled hju
argued laŋli-təp (‘reply’ + RECIP)
arose sisa
arranged zaŋ
asked səmtsə
assembled athənsi
ate meal tʃuŋ (variant: tʃijun)
awaited ata
baked li
barked tsəŋ
beat jək
became kəm
bit mùʔ
blessed muwa
blew phu
blocked, obstructed jùŋthəŋ
blocked, stopped up thəŋ
bloomed puŋ
boasted atshəkhəm
boiled məlu
bore (gave birth) sùʔ
bore, withstood ələm
borrowed tʃhəluʔ
bought hlì (<PTB **b-rey*)
bound nəj
bound (with a bamboo strip) ɪək
breathed sənsi
broke ɪək.səʔ
brooded (eggs) muk
brought hənə.ɪà (<‘take’ + SEQ
 + ‘come’)
built zaŋlu

buried ɹəm
burned ɹuŋ (<PTB *ploŋ)
called tʃa
came ɹà
came across at same level hiɹa
came down ləɹa
came in za-ɹà ('enter' + 'come')
came up kəɹa
carried ju
carried on back pù? (<PTB *buw)
carried on shoulder pu
 (possibly <PTB *ba)
carved tsə
caught phàɹa? (variant: phà?)
challenged, dared azə
changed mələn (<PTB *lay)
chased əɹsi
chewed mətʃək
choked (strangled) niŋ
choked on food ahnaŋ
chopped (with axe) sə
cleaned məɹək
cleaned (by rubbing, burnishing)
 məzu
cleaned (pond or spring of
 undergrowth) məsəm
cleared phi
cleared (a field, with fire) tsə
closed tʃhì? (<PTB *ts(y)ip)
clouded over màn-phàŋ ('be dark'
 + 'cover' (<PTB *muŋ)
coerced məɹən (<PTB *nyen)
collected tʃhaɹa ('pick.up' +
 'come')
collided, bumped tshək
concealed mijim
congealed hìlù?
consumed tʃà? (<PTB *tsa)
convalesced puŋtu
cooked tʃhuɹu (<PTB *tsyow)
copulated, joined mətəm

coughed akhət
counted zəŋ
covered nəm-phəŋ ('compress' +
 'cover')
cracked kaŋ
crawled aw
cried tʃəp
cried out, screamed ajim
crossed tən
crossed over, passed səj
crowed khuŋ
crumpled məzək
crushed, digested məhnən
cultivated jìm
cured nə
cursed ŋaŋsi
cut into pieces ləm-sa? ('divide' +
 SEPARATE)
cut with hacking action laŋ
cut with sawing or slicing action
 wà?
cut, cut off ləp (<PTB *lep)
danced tsəŋsaŋ
dangled saŋ
daubed tì
defied ahnəm
demanded məsà?
departed sənsa
did, made tʃhà
died sə (<PTB *siy)
disappeared, lost tʃhàmà? (<PTB
 *ma-t)
discarded, threw away həntʃuk
discussed thəpsi
dispersed, broke up pak
distributed ləm-sa ('divide' +
 SPREAD)
divided ləm-thaŋ ('divide' +
 SEVER), hiasà
drank tʃəm
drowned ahɹək
dug tù? (<PTB *tu)

dug out tù-tʃhət ('dig' + ABIL)
elongated, extended tshə-hlaŋ
 ('pull' + 'be long')
emerged tʃhuwa
encircled lùŋ-phaŋ (probably 'be
 round' + 'cover')
ended, finished thəm
entered za
exposed to the sun phu
extinguished phu-nik ('blow' + ?)
fanned hìp-si ('fan' + RPET)
fell down athà
fell down (of fruit) atsə
fell over luŋa
filled səŋ
fled tʃhən
flew, flowed jim
followed, tracked hni (possibly
 < PTB **naŋ*)
fought hùŋtəp
gathered together (of people)
 səntəp
gave khì?
gnawed məzəp
got up tʃhàktʃhà
grazed tʃhìlak
ground teeth tʃhitəp
guarded hnàk
happened tʃhà
heaped naŋ
held hèmət
honed h.ruŋ
hopped jaŋ
howled aŋu
hung hitʃak
ignited tsək
interfered tʃusi.luk
itched məsak (< PTB **m-sak*)
jested asaja
joined luk
judged, passed judgement ləpthàŋ
jumped puŋ

kept, left tʃhàwk (variant: tʃuk)
killed hən-sət (possibly 'by hand'
 + 'kill'); təp-sət ('hit' + 'kill')
 (< PTB **g-sat*)
kissed mətʃhəp (< PTB **dzop*)
kneaded atʃət
knocked khù-sì ('knock' + RPET)
laughed, smiled mənì (< PTB
 **m-nwi*[y])
leaked sən
leaned back nəntak
led nì
licked məlak (< PTB *(*m-*)*lak*)
loved mijim
made tʃhà
met tsə
moved to-and-fro kəp-tən ('move'
 + TO.AND.FRO)
nodded ŋam
obtained, saw hùŋ
opened tʃhəpà?
opened mouth ka (< PTB **m-ka*)
oppressed khaŋsi
overflowed t(ə)-ali-jim (NZP + 'be
 in excess' + 'flow')
painted tàp
panted sənma?
passed by hi?
pecked tsək
peeped hnihi
perforated, sewed, bored (a hole)
 h.ɿà
permitted məla
picked up tʃhà
picked up with fingers hlù
pinched mətshək
pitied min
placed, put tʃhìjàk
plaited mətsə
planted liŋ, tʃhəm
played tʃhaj (< PTB **tsya-y*)
plucked tsək

plugged thàn
poisoned məzəm
poked tù?
possessed kət, khà
pounced upon puŋ-tshəŋ (probably
 ‘jump’ + ATTACH)
pounded tshè
poured ilak, səwk
praised anak
prepared .ənəm
pressed nəm
pressed nənət
proclaimed lətʃhət, phiju
pulled tshè?
pulled towards tshəɹa (‘pull’ +
 ‘come’)
**punched, struck with fist or
 elbow, speared** tsəŋ
punished mə.ɹənsi
put into hən-uk (‘carry’ + INTO)
put into mouth nəm-uk (probably
 ‘press’ + INTO)
quarrelled laŋli-təp (‘reply’ +
 RECIP)
ran səm tsə
reached thùŋ
read (cf. ‘counted’) zəŋ
reaped zì
relaxed anina?
released atizək
remained at, rested, observed
 (anniversary, sabbath etc.) mùŋ
repeated tʃhà-tʃən-tʃhà (‘do’ +
 DUR + ‘do’)
replied laŋli
requested məpisi
returned, turned around mijip
revolved səpsi
roared ah.ɹə
roasted (by burying in coals) tsək
roasted, toasted (over coals) hwaŋ
 (possibly <PTB *kaŋ)

rolled luŋ
rooted (in earth) mukuŋ
said sa
sang tən
sat mən
saturated za
scattered by casting thəp-sa
 (‘throw’ + SPREAD)
scattered (carron pips, snooker
 balls etc.) kəp-sa (‘shoot’ +
 SPREAD)
scattered (rice while eating) thùksa
scolded hwaŋsi
scooped (grain etc) jətʃhət
scraped tʃhətùk
scratched hnək
scratched (in dirt, of a chicken)
 sùsa
screamed, cried out, announced
 ajim
scrubbed tʃhət
searched pasi
selected tʃhàmʃhət
sent zək
separated phi
set hmən
severed thàn
severed in one blow laŋ
shaved tʃhì
shined jà
shook athèn
shot, struck kəp
shouted asà
showed, demonstrated si
shunned .ijuk
slapped mà?
slept jip-li (possibly ‘sleep’ + COP
 ‘be’) (variant: jip) (<PTB *ip)
slid .ɹə
slipped (when walking) sakti?
smashed to pieces thəp-sa?
 (‘throw’ + SEPARATE)

smeard, **spread** zə
snapped in two hjù-thàn ('snap' + SEVER)
snapped off hjù
snatched lukɿa
sneezed atʃhi
sold jùk
sowed, broadcast phəɹuk
sowed, planted thən
spat məthá
split səmà
spoke sana
sprinkled phà
squeezed atʃət
stabbed, poked thən
stalked, crept ɹə
stamped on tsə
started tənɹak
stayed là?, li
stepped mətʃak
stole phù?
stood hùŋli
stood on tshət
stopped hmənsə
strangled niŋ
stretched tshə-hlaj ('pull' + 'be long')
stripped (off clothing) tʃhəmlak
stripped (off skin, bark etc) sàk
struck tətəp
struck (of lightning) là; phəɹatsəŋ
stung mətak
succeeded, won kùk
sucked mətʃhəp (<PTB *dzop)
surrounded məkətpħaŋ
swaggered kàli
swallowed mi (probably (<PTB *mlyuw))
swam atsə wà? ('water' + 'slice')
swayed səmsì
sweated asən-tsə (? + 'water')
swelled wàk

swept ùk
taught saja
threaded ɹuk, ɹə
threatened məɹən
threshed tsə
threw thətəp
threw away thùktʃuk
tickled kəlìk (<PTB *g-li)
tied mətshə
tipped over khəp
took tshə
took out əntʃhət
took, carried hən
tore səsə?
trampled tʃəm
trapped, caught tʃhà?
turned si
twirled, spun məŋa
twisted mətʃhən
unravelled phèlà
untied sèla
urinated hjuŋtsə hjuŋ
used həmsi
visited tsəŋtsə
vomited áthú?
waited là?
walked tʃəlì
wandered ili
washed (body) ju
washed (clothes, car, legs, kitchen utensils), rinsed mouth tshəj
washed (face) mənə?
washed (hair) luk
washed (hands) mətsək
waved jipsi
weeded ɹaktʃhət
weighed hju
went wa (<PTB *s-wa)
went across at same level hja
went down hlà (variant: lala)
went up kəwa
wept tʃətəp

whispered nəm̀ǹəm

wore (clothes) tʃhəm, tshəŋ

worked, laboured tʃhə̀nɪŋ

worried sisàj

wove tàk (<PTB **tak*)

wrapped in shawl mùk

wrapped hləp (<PTB **klup*)

wrapped around khàŋ

wringed mətʃhən (variant: tʃhən)

wrinkled kùŋrəp

wrote zəlu

yawned hwàməsa

References

- Aikhenvald, Alexandra Y., and Robert Malcom Ward Dixon
1998 Dependencies between grammatical systems. *Language* 74 (1): 56–80.
- Anderson, Stephen R.
1985 Typological distinctions in word formation. In *Language typology and syntactic description. Vol. III: Grammatical categories and the lexicon*, Timothy Shopen (ed.), 3–56. Cambridge: Cambridge University Press.
- Anderson, Neil, and Martha Wade
1988 Ergativity and control in Foloapa. *Language and linguistics in Melanesia* 19: 1–16.
- Andrews, Avery
1985 The major functions of the noun phrase. In *Language typology and syntactic description. Vol. I: Clause structure*, Timothy Shopen (ed.), 62–154. Cambridge: Cambridge University Press.
- Ao, Bendangyabang A.
2001 Bible translation in Ao Naga. *Language in India*. Vol. 1, 5 September. <http://www.languageinindia.com/sep2001/bibleaonaga.html> (Accessed 13 November 2003).
- Austin, Peter
1981 Switch-reference in Australia. *Language* 57.2: 309–334.
2005 Causative and applicative constructions in Australian Aboriginal languages. Unpublished ms, Department of Linguistics, SOAS.
- Austin, Peter (ed.)
1988 *Complex Constructions in Australian languages*. Amsterdam/Philadelphia: John Benjamins.
- Barz, Richard and Yogendra Yadav
1993 *An introduction to Hindi and Urdu*, 5th ed. New Delhi: Munshiram Manoharlal.
- Bauer, Laurie
1983 *English word-formation*. Cambridge/New York: Cambridge University Press.
1988 *Introducing linguistic morphology*. Edinburgh: Edinburgh University Press.
- Benedict, Paul K.
1972 *Sino-Tibetan: a conspectus* (contributing editor: James A. Matisoff). Cambridge: Cambridge University Press.
- Beyer, Stephan V.
1992 *The classical Tibetan language*. Albany: State University of New York Press.
- Bhat, D. N. Shankara and M. S. Ningompa
1997 *Manipuri grammar*. München: Lincom Europa.

Bickel, Balthasar

- 1993 Belhare subordination and the theory of topic. In *Studies in clause linkage: papers from the first Köln-Zürich workshop*, Karen H. Ebert (ed.), 23–55. Zürich: Arbeiten des Seminars für Allgemeine Sprachwissenschaft.
- 1998 Review article: Converbs in cross-linguistic perspective. *Linguistic typology* 2 (3): 381–397.

Bradley, David

- 1979 *Proto-Loloish*. Scandinavian Institute of Asian Studies Monograph Series No. 39. London and Malmo: Curzon Press.
- 1992 Tone alternations in Ugong. In *Papers in Tai Languages, Linguistics and Literatures in Honor of William J. Gedney on his 77th Birthday*, Carol Compton and John Hartmann (eds.), 55–64. De Kalb, IL: Northern Illinois University.
- 1997 Tibeto-Burman languages and classification. In *Papers in Southeast Asian linguistics No. 14: Tibeto-Burman languages of the Himalayas*, David Bradley (ed.), 1–71. Canberra: Pacific Linguistics.

Burling, Robbins

- 1961 *A Garo grammar*. Poona: Deccan College Postgraduate and Research Institute.
- 2003 The Tibeto-Burman languages of north-east India. In *The Sino-Tibetan languages*, Graham Thurgood and Randy J. LaPolla (eds.), 169–191. London/New York: Routledge.
- 2004 *The language of the Modhupur Mandi (Garo)* Vol. I: grammar. New Delhi and New Jersey: Bibliophile South Asia.

Bybee, Joan L.

- 1985 *Morphology: a study of the relation between meaning and form*. Amsterdam/Philadelphia: John Benjamins.

Bybee, Joan L, Revere Perkins and William Pagliuca

- 1994 *The evolution of grammar: tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.

Census of India. 1991. <http://www.censusindia.net>. Accessed 11 November 2003.

Chappell, Hilary and William McGregor (eds.)

- 1996 *The grammar of inalienability*. Berlin/New York: Mouton.

Chelliah, Shobhana

- 1997 *A grammar of Meithei*. Berlin/New York: Mouton de Gruyter.

Chung, Sandra, and Allan Timberlake

- 1985 Tense, aspect and mood. In *Language typology and syntactic description. Vol. III: Grammatical categories and the lexicon*, Timothy Shopen (ed.), 202–258. Cambridge: Cambridge University Press.

Chao, Yuen Ren

- 1930 A system of tone letters. *Le Maître Phonétique* 30: 24–27.

Clark, Mary Mead

- 1978 Reprint. *A corner in India*. Gauhati: Christian Literature Centre. Original edition, 1907.

- 2002 Reprint. *The Ao Naga grammar*. Delhi: Gian Publications. Original edition, 1893.
- Clark, Edward Winter
1990 Reprint. *Ao Naga dictionary*. Mokokchung (Nagaland): R. Lisen Ao. Original edition, 1911.
- Comrie, Bernard
1985 Causative verb formation and other verb-deriving morphology. In *Language typology and syntactic description. Vol. III: Grammatical categories and the lexicon*, Timothy Shopen (ed.), 309–348. Cambridge: Cambridge University Press.
- Comrie, Bernard and Sandra A. Thompson
1985 Lexical nominalization, . In *Language typology and syntactic description. Vol. III: Grammatical categories and the lexicon*, Timothy Shopen (ed.), 349–398. Cambridge: Cambridge University Press.
- Coupe, Alexander Robertson
1998 The acoustic and perceptual features of tone in the Tibeto-Burman language Ao Naga. In *Proceedings of the 5th international conference on spoken language processing. Vol. 2*, Robert Mannell and Jordi Robert-Ribes (eds.), 57–60. Canberra: The Australian Speech Science and Technology Association (also produced on CD-ROM ISBN 1 876346 17 5).
- 1999 A phonetic and phonological description of Ao: a language of Nagaland, north-east India. MA thesis, The Australian National University.
- 2001 Thai. In *Facts about the world's languages: an encyclopedia of the world's major languages, past and present*, Jane Garry and Carl Rubino (eds.), 733–739. New York/Dublin: H.W. Wilson.
- 2002 Temporal deixis in Ao. *Melbourne Papers in Linguistics and Applied Linguistics* 1(1): 5–20.
- 2003a *A phonetic and phonological description of Ao: a Tibeto-Burman language of Nagaland, north-east India*. Canberra: Pacific Linguistics.
- 2003b Verbal lexical suffixes in the Mongsen dialect of Ao. Paper read at the 36th International Conference on Sino-Tibetan Languages and Linguistics, La Trobe University, Melbourne 28th–30th November.
- 2003c The Mongsen dialect of Ao: a language of Nagaland. PhD diss., La Trobe University.
- 2004 Recounting the numeral systems of Nagaland: old evidence for a new subgrouping. Paper read at the Xth Himalayan Languages Symposium, Thimphu, Bhutan 1st–3rd December.
- 2005 Converbs. In *Encyclopedia of languages and linguistics* 2nd ed., Vol. 3, Keith Brown (ed.), 145–152. Oxford: Elsevier.
- 2007 Converging patterns of clause linkage in Nagaland. In *New challenges in typology: broadening the horizons and redefining the foundations*, Matti Miestamo and Bernhard Wälchli (eds.) 341–363. Berlin/New York: Mouton de Gruyter.

Craig, Colette G.

- 1991 Ways to go in Rama. In *Approaches to Grammaticalization*. Vol. 2, Elizabeth Closs Traugott and Bernd Heine (eds.), 455–492. Amsterdam/Philadelphia: John Benjamins.

DeLancey, Scott

- 1986 Relativization as nominalization in Tibetan and Newari. Paper presented at the 19th International Conference on Sino-Tibetan Languages and Linguistics.
- 1997a Grammaticalization and the gradience of categories. In *Essays on language function and language type*, Joan Bybee, John Haiman, and Sandra A. Thompson (eds.), 51–69. Amsterdam/Philadelphia: John Benjamins.
- 1997b Mirativity: the grammatical marking of unexpected information. *Linguistic Typology* 1: 33–52.
- 1999 Relativization in Tibetan. In *Topics in Nepalese Linguistics*, Yogendra P. Yadava and Warren W. Glover (eds.), 231–249. Kathmandu: Royal Nepal Academy.
- 2001 The mirative and evidentiality. *Journal of Pragmatics* 33: 369–382.

Dench, Alan

- 1988 Complex Sentences in Martuthunira. In *Complex Constructions in Australian languages*, Peter Austin (ed.), 97–139. Amsterdam/Philadelphia: John Benjamins.

Dixon, Robert Malcom Ward

- 1979 Ergativity. *Language* 55: 59–138.
- 1982 *Where have all the adjectives gone? and other essays in semantics and syntax*. Berlin: Mouton.
- 1994 *Ergativity*. Cambridge: Cambridge University Press.
- 2004 Adjective classes in typological perspective. In *Adjective classes: a cross-linguistic typology*, R.M.W. Dixon and Alexandra Y. Aikhenvald (eds.), 1–49. Oxford/New York: Oxford University Press.

Dixon, Robert Malcom Ward and Alexandra Y. Aikhenvald

- 2002 Word: a typological framework. In *Word, a cross-linguistic typology*, Robert Malcom Ward Dixon and Alexandra Y. Aikhenvald (eds.), 1–41. Cambridge/ New York: Cambridge University Press.

Dryer, Matthew

- 2003 Word order in Sino-Tibetan languages from a typological and geographical perspective. In *The Sino-Tibetan languages*, Graham Thurgood and Randy J. LaPolla (eds.), 43–55. London/New York: Routledge.

Durie, Mark

- 1985 *A grammar of Acehnese, on the basis of a dialect of North Aceh*. Dordrecht: Foris.

Ebert, Karen

- 1997 *A grammar of Athpare*. LINCOM studies in Asian linguistics; 01. München: LINCOM Europa.

- 1999 Non-finite verbs in Kiranti languages – an areal perspective. In *Topics in Nepalese linguistics*, Y.P. Yadava and W.W. Glover (eds.), 371–400. Kathmandu: Royal Nepal Academy.
- Foley, William A.
 1986 *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.
 1991 *The Yimas language of New Guinea*. Stanford: Stanford University Press.
- Foley, William A. and Mike Olson
 1985 Clausehood and verb serialization. In *Grammar inside and outside the clause*, J. Nichols and A.C. Woodbury (eds.), 17–60. Cambridge: Cambridge University Press.
- Foley, William A. and Robert D. Van Valin
 1984 *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Fortesque, Michael
 1984 *West Greenlandic*. London/Sydney/Dover NH: Croom Helm.
- Gait, Edward
 2003 Reprint. *A history of Assam*. New Delhi: Surjeet Publications. Original edition, 1906.
- Gasser, Michael
 1983 Topic continuity in written Amharic discourse. In *Topic continuity in discourse*, T. Givón (ed.), 95–139. Amsterdam/Philadelphia: John Benjamins.
- Gedney, William J.
 1972 A checklist for determining tones in Tai dialects. In *Studies in linguistics in honor of George L. Trager*, M. Estellie Smith, (ed.), 423–437. The Hague: Mouton.
- Genetti, Carol
 1986 The Development of Subordinators from Postpositions in Bodic Languages. *BLS* Vol. 12: 387–400.
 1991 From Postposition to Subordinator in Newari. In *Approaches to Grammaticalization* Vol. 2, Elizabeth Closs Traugott and Bernd Heine (eds.), 227–255. Amsterdam/Philadelphia: John Benjamins.
 1992 Semantic and grammatical categories of relative clause morphology in the languages of Nepal. *Studies in Language* 16 (2): 405–427.
 1993 On the morphological status of case markers in Dolakha Newari. *Linguistics of the Tibeto-Burman Area* 16 (1): 57–73.
 1994 *A descriptive and historical account of the Dolakha Newari dialect*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
 2005 The participial construction of Dolakā Newar: syntactic implications of an Asian converb. *Studies in Language* 29 (1): 35–87.
- Giridhar, P.P.
 1980 *Angami grammar*. Mysore: Central Institute of Indian Languages.
 1994 *Mao Naga grammar*. Mysore: Central Institute of Indian Languages.

- Givón, Talmy
1979 *On understanding grammar*. New York /San Francisco: Academic Press.
- Goldsmith, John A.
1990 *Autosegmental and metrical phonology*. Oxford/Cambridge MA: Blackwell.
- Gowda, K.S. Gurubasave
1972 *Ao-Naga phonetic reader*. Mysore: Central Institute of Indian Languages.
1975 *Ao grammar*. Mysore: Central Institute of Indian Languages.
- Grierson, George A. (ed.)
1967 Reprint. *Linguistic survey of India*. Vol. 3, Part 2. Delhi: Motilal Banarsidass. Original series, 1903–1928.
- Haiman, John
1978 Conditionals are topics. *Language* 54: 564–589.
1980 *Hua: a Papuan language of the Eastern Highlands of New Guinea*. Amsterdam/Philadelphia: John Benjamins.
1985 Symmetry. In *Iconicity in syntax*, John Haiman (ed.), 73–95. Amsterdam/Philadelphia: John Benjamins.
- Haiman, John and Sandra A. Thompson
1984 ‘Subordination’ in universal grammar. In *Proceedings of the tenth annual meeting of the Berkeley Linguistics Society*, Claudia Brugmann and Monica Macauley (eds.), 510–523. Berkeley: University of California Press.
- Hargreaves, David
1996 From interrogation to topicalization: proto-Tibeto-Burman **la* in Kathmandu Newar. *Linguistics of the Tibeto-Burman Area* 19 (2): 31–44.
- Haspelmath, Martin
1995 The converb as a cross-linguistically valid category. In *Converbs in Cross-Linguistic Perspective*, Martin Haspelmath and Ekkehard König (eds.), 1–55. Berlin/New York: Mouton de Gruyter.
1997 From space to time: temporal adverbials in the world’s languages. München: LINCOM Europa.
- Haspelmath, Martin and Ekkehard König (eds.)
1995 *Converbs in Cross-Linguistic Perspective: structure and meaning of adverbial verb forms – adverbial participles, gerunds*. Berlin/New York: Mouton de Gruyter.
- Heine, Bernd and Tanya Kuteva
2002 *World lexicon of grammaticalization*. Cambridge/New York: Cambridge University Press.
- Henderson, Eugénie J.A.
1965 *Tiddim Chin: a descriptive analysis of two texts*. London: Oxford University Press.
- Hockett, Charles Francis
1958 Two models of grammatical description. In *Readings in linguistics*. 2nd ed., Martin Joos (ed.), 386–399. Chicago: University of Chicago Press.

- Holisky, Dee Ann
 1987 The case of the intransitive subject in Tsova-Tush (Batsbi). *Lingua* 71: 103–132.
- Hopper, Paul J.
 1991 On some principles of grammaticalization. In *Approaches to Grammaticalization* Vol. 1, Elizabeth Closs Traugott and Bernd Heine (eds.), 17–35. Amsterdam/Philadelphia: John Benjamins.
- Hutton, John Henry
 1921a *The Angami Nagas*. London: Macmillan and Co.
 1921b *The Sema Nagas*. London: Macmillan and Co.
 1987 Reprint. *Chang language*. Delhi: Gian Publishing House. Original edition, 1929.
- Johanson, Lars
 1995 On Turkic converb clauses. In *Converbs in Cross-Linguistic Perspective*, Martin Haspelmath and Ekkehard König (eds.), 313–347. Berlin/New York: Mouton de Gruyter.
- Kachru, Yamuna
 1968 The copula in Hindi. In *The verb 'be' and its synonyms* Part 2, John W.M Verhaar (ed.), 35–59. Dordrecht: D. Reidel Publishing Company.
- Kari, James
 1989 Affix positions and zones in the Athapascan verb complex: Ahtna and Navajo. *International Journal of American Linguistics* 55 (4): 424–454.
- Kapfo, Kedutso
 1989 Tones in Khezha noun constructions. *Linguistics of the Tibeto-Burman Area* 12(2): 67–78.
- Keenan, Edward L.
 1985 Relative clauses. In *Language typology and syntactic description: Vol. II Complex constructions*, Timothy Shopen (ed.), 141–170. Cambridge: Cambridge University Press.
- Keenan, Edward L. and Bernard Comrie
 1977 Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8 (1): 63–100.
- Klavans, Judith L.
 1982 *Some problems in a theory of clitics*. Bloomington: Indiana Linguistics club.
 1985 The independence of syntax and phonology in cliticization. *Language* 61: 95–120.
- Koptjevskaja-Tamm, Maria
 1994 Finiteness. In *The encyclopedia of languages and linguistics*. 1st ed., Ronald E. Asher, and J. M. Y. Simpson (eds.), 1245–1248. Oxford/New York: Pergamon Press.
- Ladefoged, Peter and Peri Bhaskararao
 1991 Two types of voiceless nasals. *Journal of the International Phonetic Association* 21 (2): 80–88.

- Ladefoged, Peter and Ian Maddieson
1996 *The sounds of the world's languages*. Oxford/Cambridge MA.: Blackwell.
- LaPolla, Randy
1994 Parallel grammaticalisations 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. *Senri Ethnological Studies* 41: 189–228.
- Lamberti, Marcello and Roberto Sottile
1997 *The Wolaytta language*. Köhn: Rüdiger Köppe Verlag.
- Lass, Roger
1984 *Phonology: an introduction to basic concepts*. Cambridge/New York: Cambridge University Press.
- Lazard, Gilbert
1999 Mirativity, evidentiality, mediativity, or other? *Linguistic Typology* 3: 91–109.
- Lehmann, Christian
1988 Towards a typology of clause linkage. In *Clause combining in grammar and discourse*, John Haiman and Sandra A. Thompson (eds.), 181–225. Amsterdam/Philadelphia: John Benjamins.
- Mackenzie, Alexander
1995 Reprint. *The north-east frontier of India*. New Delhi: Mittal Publications. Original edition, 1884.
- Maddieson, Ian
1978 Universals of tone. In *Universals of human language, Vol. 2: Phonology*, Joseph H. Greenberg (ed.), 337–365. Stanford: Stanford University Press.
- Masica, Colin P.
1976 *On defining a linguistic area*. Chicago/London: University of Chicago Press.
1991 *The Indo-Aryan languages*. Cambridge: Cambridge University Press.
- Marrison, Geoffrey Edward
1967 The classification of the Naga languages of north-east India. Vol. I (Texts) and Vol. II (Appendices). Unpublished PhD diss., SOAS, University of London.
- Matisoff, James A.
1972 Lahu nominalization, relativization and genitivization. In *Syntax and Semantics I*, J. Kimball (ed.), 237–257. New York: Seminar Press.
1973 *The grammar of Lahu*. Berkeley, Los Angeles/London: University of California Press.
1976 Lahu causative constructions. In *Syntax and semantics, Vol. 6: the grammar of causative constructions*, Masayoshi Shibitani (ed.), 413–442. New York/San Francisco/London: Academic Press.
1985 God and the Sino-Tibetan copula, with some good news concerning selected Tibeto-Burman rhymes. *Journal of Asian and African Studies* 29: 1–81.

- 1986 The languages and dialects of Tibeto-Burman: an alphabetic/genetic listing, with some prefatory remarks on ethnonymic and glossonymic complications. In *Contributions to Sino-Tibetan Studies*, John McCoy and Timothy Light (eds.), 3–75. Leiden: E. J. Brill.
- 1988 *The dictionary of Lahu*. University of California Publications in Linguistics 11. Berkeley: University of California Press.
- 1992 The mother of all morphemes: Augmentatives and diminutives in areal and universal perspective. In *Papers from the first annual meeting of the Southeast Asian Linguistics Society 1991*, Martha Ratliff and Eric Schiller (eds.), 293–349. Arizona: Arizona State University, Program for Southeast Asian Studies.
- 1997 *Sino-Tibetan numeral systems: prefixes, protoforms and problems*. Canberra: Pacific Linguistics.
- 1998 *The dictionary of Lahu*. Berkeley: University of California Press.
- Matthews, P.H.
1997 *The concise Oxford dictionary of linguistics*. Oxford/New York: Oxford University Press.
- Mazaudon, Martine
1977 Tibeto-Burman tonogenetics. *Linguistics of the Tibeto-Burman Area* 3 (2): 1–123.
- McCabe, R.B.
1887 *Outline grammar of the Angami Naga language*. Calcutta: Supt. of Govt. Printing.
- Menninger, Karl
1969 *Number Words and number symbols: a cultural history of numbers*. [Zahlwort und Ziffer]. Translated by Paul Broneer from the revised German edition. Cambridge, MS: MIT Press.
- Miller, Marion
1999 *Desano grammar: Studies in the languages of Columbia* 6. Arlington: Summer Institute of Linguistics and the University of Texas.
- Mills, James Phillip
1926 *The Ao Nagas*. London: Macmillan and Co.
1937 *The Rengma Nagas*. London: Macmillan and Co.
n.d. Typescript, [reported to be ca.1927 by Marrison 1967]. Mongsen Ao word list. Library of SOAS, University of London.
- Mithun, Marianne
1991 Active/agentive case marking and its motivations. *Language* 67 (3): 510–546.
2000 Valency changing derivation in Central Alaskan Yup'ik. In *Changing valency: case studies in transitivity*, Robert Malcom Ward Dixon and Alexandra Y. Aikhenvald (eds.), 84–114. Cambridge/New York: Oxford University Press.
- Nadkarni, Mangesh V.
1975 Bilingualism and syntactic change in Konkani. *Language* 51 (3): 672–683.

- Nagano, Yasuhiko
1985 *A historical study of the rGyarong verb system*. Tokyo: Seishido.
- Nedjalkov, Igor' V.
1995 Converbs in Evenki. In *Converbs in Cross-Linguistic Perspective*, Martin Haspelmath and Ekkehard König (eds.), 441–463. Berlin/New York: Mouton de Gruyter.
1998 Converbs in the languages of eastern Siberia. *Language Sciences* 20 (3): 339–351.
- Nedjalkov, Vladimir P.
1995 Some typological parameters of converbs. In *Converbs in Cross-Linguistic Perspective*, Martin Haspelmath and Ekkehard König (eds.), 97–136. Berlin/New York: Mouton de Gruyter.
- Newman, Paul and Martha Ratliff (eds.)
2001 *Linguistic fieldwork*. Cambridge: Cambridge University Press.
- Nichols, Joanna
1986 Head-marking and dependent-marking grammar. *Language* 62 (1): 56–119.
- Noonan, Michael
1997 Versatile nominalizations. In *Essays on language function and language type*, Joan Bybee, John Haiman and Sandra A. Thompson (eds.), 373–394. Amsterdam/Philadelphia: John Benjamins.
- Okell, John
1989 Notes on tone alternation in Maru verbs. In *Prosodic analysis and Asian linguistics: to honour R.K. Sprigg*, David Bradley, Eugénie J.A. Henderson and Martine Mazaudon (eds.), 109–114. Canberra: Pacific Linguistics.
- Payne, Doris L.
1990 *The pragmatics of word order*. Berlin/New York: Mouton de Gruyter.
- Ramstedt, Gustaf John
1903 *Über die Konjugation des Khalka-Mongolischen*. Mémoires de la Société Finno-Ougrienne 19. Helsingfors: Société Finno-Ougrienne.
- Reid, Sir Robert
1997 Reprint. *History of the frontier areas bordering on Assam from 1883 to 1941*. Guwahati/Delhi: Spectrum Publications. Original edition, 1942.
- Rice, Keren
2000 *Morpheme order and semantic scope: word formation in the Athapaskan verb*. Cambridge: Cambridge University Press.
- Rutgers, Roland
1998 *Yamphu: grammar, text & lexicon*. Languages of the Greater Himalayan Region, Vol. 2. Leiden: Leiden University.
- Sapir, Edward
1921 *Language: an introduction to the study of speech*. San Diego/New York/London: Harcourt Brace Jovanovich.

- Saxena, Anju
 1988 On syntactic convergence: the case of the verb 'say' in Tibeto-Burman. *Berkeley Linguistics Society* 14: 375–388.
- Schachter, Paul
 1985 Parts-of-speech systems. In *Language typology and syntactic description: Vol. I: Clause structure*, Timothy Shopen (ed.), 1–61. Cambridge: Cambridge University Press.
- Shafer, Robert
 1955 The classification of the Sino-Tibetan languages. *Word* 11:94–111.
 1974 *Introduction to Sino-Tibetan*. Weisbaden: Otto Harrassowitz.
- Shakespeare, Leslie Waterfield
 1980 Reprint. *History of the Assam Rifles*. Gauhati: Spectrum Publications. Original edition, 1929.
- Shopen, Timothy (ed.)
 1985 *Language typology and syntactic description: Vol. I: Clause structure, Vol. II: Complex constructions, and Vol. III: Grammatical categories and the lexicon*. Cambridge: Cambridge University Press.
- Simpson, Jane
 1988 Case and complementiser suffixes in Warlpiri. In *Complex Constructions in Australian languages*, Peter Austin (ed.), 205–218. Amsterdam/Philadelphia: John Benjamins.
- Solnit, David
 1997 *Eastern Kayah Li; grammar, texts, glossary*. Honolulu: University of Hawai'i Press.
- Sony Corporation
 1997 *Sony Portable MiniDisc Recorder operating instructions*.
- Sreedhar, M.V.
 1974 *Naga Pidgin: a sociolinguistic study of inter-lingual communication pattern in Nagaland*. Mysore: Central Institute of Indian Languages.
- Thompson, Sandra A. and Robert E. Longacre
 1985 Adverbial clauses. In *Language typology and syntactic description: Vol. II Complex constructions*, Timothy Shopen (ed.), 171–234. Cambridge: Cambridge University Press.
- Thurgood, Graham
 2003 A sub-grouping of the Sino-Tibetan languages: the interaction between language contact, change, and inheritance. In *The Sino-Tibetan languages*, Graham Thurgood and Randy J. LaPolla (eds.), 3–21. London/New York: Routledge.
- Watters, David E.
 2002 *A grammar of Kham*. Cambridge: Cambridge University Press.
- Weirisma, Grace
 2003 Yunnan Bai in Brief. In *The Sino-Tibetan languages*, Graham Thurgood and Randy J. LaPolla (eds.), 651–673. London/New York: Routledge.

Witter, W.E.

1888 *Outline grammar of the Lhota Naga language: with vocabulary and illustrative sentences.* Calcutta: Supt of Govt. Printing.

Wolfenden, Stuart Norris

1929 *Outlines of Tibeto-Burman linguistic morphology.* London: Royal Asiatic Society.

Zwicky, Arnold M.

1985 Clitics and particles. *Language* 61 (2): 283–305.

Index of subjects

- acoustic analysis
 - central approximants, 39–43
 - lateral approximants, 43–44
 - nasals, 32–38
- adjective of degree derivations
 - as a type of nominalization, 261–263
 - comparative, 60–63, 261–262
 - superlative, 61–63, 262–263
- admonitive mood, 291, 292–393
- adverbs, 124, *also see lexical suffixes and converbs*
- affricates, 31
- afterthought, 152, 359
- agentive case marking
 - and control, 163
 - and disambiguation of semantic roles, 154, 159
 - and habitual activity, 165
 - and socially-marked behaviour, 159–160
 - in intransitive clauses, 160–165
 - not conditioned by other grammatical categories, 160
- agentive nominalizer
 - genitivizing, relativizing and nominalizing functions, 252–255, 263–264
 - diachronic source of sequential converb, 264, 344, 446–447
 - diachronic source of present tense marker, 344–345
 - masculine semantic gender marking function, 264
- ambitransitive verbs, 153, 168–169, 285–287
- Andrews, 150
- Ao
 - autonym, 9–10
 - and language endangerment, 6
 - colonial history, 11–12
 - genetic classification, 7
 - literature, 18
 - moieties of tribe, 10
 - number of speakers, 1
 - villages, 8
 - village life, 13–14
- Ao dialects
 - bidialectalism, 16
 - Changki, 16, 115, 121
 - Chungli, 9–10, 16–18, 29–30, 45–46, 91, 115, 119–120, 127, 172, 186, 455–458
 - Longla, 16, 118
 - Longsa, 5
 - overview of, 16–17
 - previous research on, 17
 - segmental differences between Mongsen varieties, 47–49
 - Yacham, 5, 118, 119
 - Yong-Yacham, 5, 16
- Ao religion, 14–16
- approximants, 38–44
- aspectual suffixes
 - completive, 327–328
 - continuative, 325–326
 - description of**, 325–329
 - habitual past, 328–329
 - repetitive, 326–327
- augmentative, grammaticalization of, 272
- basic clause types
 - description of**, 145–150
 - declarative, 147–148

- interrogative, 148–150
- imperative, 150, 387–389
- Benedict, 59, 90, 113, 115
- Burling, 4, 308
- case markers
 - ablative, 177–178
 - allative, 176–177
 - and morphological variation across Mongsen varieties, 172
 - agentive, 173–174
 - agentive case signalled by a rising tone, 64–65
 - association with noun phrases, 169–170, 201, 213–214
 - benefactive, 172, 181
 - comitative, 183–184
 - comitative in comparative derivations, 261–262
 - dative, 171, 180
 - core, 173
 - core versus oblique marking
 - functions of, 121, 170–171
 - description of**, 169–189
 - instrumental, 178–179
 - isomorphism of agentive, allative and instrumental, 105, 171–172
 - locative, 182–183
 - non-canonical semantically motivated dative marking in causative derivations, 191–194, 197
 - oblique, description of, 175–189
- causal particle, 450–451
- causative derivation
 - applied reading of, 193–194
 - analytic, 197–198
 - and semantically motivated case marking, 190–193
 - description of**, 190–198
 - encoded by the directional suffix, 324
 - morphological, 190–197
 - of transitive stems, 194–197
 - of intransitive stems, 190–194
 - causative suffix, 340–341
- Chang, 4, 164fn., 346fn., 407–408
- Chang “imperative reflexive”, 407–408
- Chelliah, 128, 155–156
- Christianity, 12, 14–16
- Clark, Edward Winter, 15, 127, 172, 181, 186, 320
- Clark, Mary Mead, 15, 117–118
- clitics
 - characteristics of, 51–52
 - case marking clitics, 213–214
 - declarative, 142–144
- collective derivation, 198–200, 321–322
- complements
 - embedded direct speech, 452–453
 - expressing onomatopoeia, 130, 453
 - in purposive clauses, 265–266
 - types of complement-taking predicates, 453
- compound nouns
 - description of**, 241–245
 - nominalized verb-noun, 245
 - noun-noun, 243–244
 - noun-nominalized verb, 245
 - noun-verb, 244
- conjunctive elements, 127
- consonant phonemes, 27–44
- constituent order
 - description of**, 151–153
 - pragmatic influences on, 152–153
 - expressing the functional equivalent of passive voice, 153
- converbs
 - alternating iterative converb, 428–431
 - and discourse focus particles, 419–420
 - and matrix clause operator scope, 397, 415, 418–419, 434

- and the notion of subordination, 412
- and summary-head linkage, 125–127, 421, 439–440, 444–445
- and tail-head linkage, 381–382, 421
- and the subordinate/cosubordinate distinction, 416–419
- causal converb, 442–445
- characteristics of the Asian converb, 418
- circumstantial converb, 437–439
- concessive converb, 439–440
- conditional converb, 435–437
- contemporative converb, 440–441
- definitions of, 410–413
- description of**, 409–447
- diachronic origins of, 446–447
- diachronic origin of sequential converb, 425
- as non-finite verb forms, 343–344
- diachronic sources of suffixes, 344,
- disjunctive expression, 448–449
- durative converb, 441–442
- effect of negation on narrative chaining converbs, 432–434
- forms of suffixes, 343, 421–422
- grammaticalization pathway for causal converb, 443–444
- locative converb, 445–446
- modifying and non-modifying functions of simultaneous converbs, 426–427
- narrative chaining, 422–424
- overview of types, 421–422
- sequential converb, 424–425
- simultaneous converb, 425–428
- specialized converbs, 435–446
- and their syntactic status in Mongsén, 413–421
- copula
 - diachronic origin of, 386
 - function in converb clauses, 378–379
 - function in imperative clauses, 379
 - grammaticalized functions, 231–232, 349–350, 361, 371–374, 375–376, 382–383, 431, 438
- copula clauses and expression of TAM, 371–374
- core grammatical marking
 - and disambiguation of semantic roles, 159
 - and the influence of derivational morphology, 161–162
 - and socially marked/wilful behaviour of referents, 159–164
 - description of**, 154–165
 - in intransitive clauses, 162–164
 - in transitive clauses, 156–160
 - not assigned by predicate argument structure, 173
 - pragmatically/semantically determined, in other languages of the region, 286
 - formally unmarked core arguments, 174
 - pragmatic/semantic motivations for, 156–163
- creaky voice vowel, 45, 46–47
- declarative mood clitic, 142–144
- DeLancey, 189
- demonstrative pronouns, 97–98
- deverbal adjectives and nouns, 256
- diminutive suffix, 82, 272–273, 305
- diphthongs, 25–26, 49–50, 55, 142
- directional suffix, 322–325
- discourse connectives, 124–127
- discourse focus particles, 137–138
- Dixon, 154, 155–156, 159
- Durie, 155

ellipsis of core arguments, 285–287
 exclamations, 129

fluid-S marking pattern,
 in other languages, 154–156,
 in Mongsen Ao, 160–165
 fricatives, 31–32

Garo, 308

general nominalizer

 and its relationship to the emphatic demonstrative, 108
 and tone sandhi, 62–64,
 as a semantic gender marker, 270–271
 in adjective of degree derivations, 260–263
 in compound nouns, 245
 in instrument nominalizations, 265
 in purposive complement clauses, 265–266
 in relative clause derivation, 208–209, 219–229,
 in the derivation of numeral adverbs, 277–278
 in the derivation of ordinal numerals, 277
 in the expression of deontic modality, 249–350
 in the nominalization of words, phrases and clauses, 236–239

generic pronoun, 105

glottal stop as a word prosody, 23, 24–25, 28, 50–51, 77–78

Gowda, 17, 30, 91

grammatical functions distinguished from grammatical relations, 150–151

grammaticalization of nouns as temporal markers, 449–450

Haspelmath, 410

Hutton, 407

illocutionary force particles, 138–142

imperative clauses

 admonitive, 392–393
 and compatible grammatical categories, 399–402
 and evidentiality, 402
 and semantic verb classes, 393–395
 as a distinct clause type, 387–389
 autobeneficiary imperative, 398, 409
 causative marking in, 397–399
 curses and blessings, 404–405
 dependencies with other grammatical categories, 395–399
 description of, 387–408
 incompatibility with continuative aspect, 395–396
 marking of core arguments in, 387–389
 meanings of particles in, 403–404
 origins of imperative morphology, 406–408
 positive imperative, 342, 390–391
 prohibitive, 290–291, 391–392
 topic marking in, 388
 use in discourse, 405–406

indefinite article, reanalysis of numeral ‘one’, 114

indefinite pronouns

 negative, 101–102
 non-specific, 103–104
 overview of, 100–101,
 reduplication of, 104–105

intensifier suffix, 270

interjections, 129

interrogative pronouns

description of, 98–100,
 use in content questions, 150, 348

use in relative-correlative constructions, 99, 134
 interrogative particles, 139
 interrogative clauses
 content questions, 98–100, 150, 348
 description of, 148–150
 disjunctive questions, 148
 polar questions, 148–149

Konyak languages, 4, 5, 122

lexical quantifiers, 111–112

lexical suffixes

‘AFFECT’, 316–319
 ‘ATTACH’, 314–316
 ‘ADDIT’, 319–320
 ‘ASCEND’, 301–302
 ‘ASSIST’, 320
 ‘COVER’, 320–321
 ‘DESCEND’, 302–303
description of, 298–321
 diachronic sources of, 298–301
 ‘END’, 310
 ‘INTO’, 303–304
 ‘REMAIN’, 321
 ‘SEPARATE’, 305–306
 ‘SEVER’, 307–308
 ‘SPLIT’, 308–309
 ‘SPREAD’, 306–307
 table of, 299
 ‘TERMINATIVE’, 310–312
 ‘TO AND FRO’, 304–305
 ‘TOGETHER’, 312–314

lycanthropy, 14

Marrison, 7, 29–30, 120

Matisoff, 113, 116, 118, 255

Meithei, 92, 128, 116, 128, 155–156,
 161fn, 302, 302, 319

Mills, 9, 16, 17, 119

Mithun, 155, 168

modality suffixes

abilitive, 329–330
 chance marker, 334–336
 conative, 332–333
description of, 329–336
 desiderative, 333–334
 frustrative, 330–332

mood and modality

deontic modality, 349–350
 epistemic modality, 349–350
 irrealis, 349–351

Nagaland

geographical location, 7
 Mokokchung district of, 8
 indigenous minorities of, 4

“Naga” languages, 4

Nagamese, 5–6, 17

nasals, 32–38

nascent postpositions

and categorial shift, 189
 and relationship to possessors, 175,
 182–183

description of, 184–189

diachronic origins of, 185, 187

Nedjalkov, V.P., 410

negative particle, 141

negative verbal prefix, 291–292

negative verbal suffix, 341–342

neighbouring languages, 5

also see Map 2

nominal deictics

anaphoric demonstrative, 110–111

as noun phrase heads, 97

description of, 106–111

distal, 97–98, 109–110

emphatic demonstrative, 106–108

“infixd” position of, 177, 202

overview, 106

proximate demonstrative, 97–98,

108–109

nominal derivation

- adjectives of degree, 260–263
- by prefixation, 256–259, 265, 280, 363
- by the agentive nominalizer, 263–264
- by the general nominalizer, 259–260, 260–263
- by the locational suffix, 274–275
- by the MAN nominalizer, 268–269
- deictic nominalizations, 268
- instrument nominalization, 265
- locative nominalizer, 265
- patientive nominalization, 256–258
- purposive nominalization, 265–266
- temporal nominalization, 258–259, 267

nominal prefixes

- atonal status of, 66–67
- description of**, 246–248
- nominalizing prefix, 248, 292–294
- non-relational, 83, 86, 248
- relational, 83, 84–85, 247
- vocative, 85–86, 246–247

nominalizations

- as clausal arguments, 237
- description of**, 236–239
- types of nominalizing morphemes, 219

non-embedded nominalization, 349–350

noun phrases

- and case marking clitics, 204, 213–214
- and intonation, 74–75, 205
- and pre-head quantifiers, 202–203
- appositive, 210
- description of**, 201–217
- determiners of, 212–213
- discourse focus and topic particles in, 214–216
- functions of, 204–205

headless, 211–212

- linear structure of, 201–202
- modifiers of, 208–213
- pre-head constituents of, 106, 203–204
- quantifiers in, 211–212
- recursive embedding in, 216–217
- types of attributes of, 208–211
- types of heads of, 205–208
- with complex heads, 207–208

nouns

- kinship, 85–86
- non-relational bound, 86–87
- personal names as nominalizations, 270–271
- relational bound, 84–86, 87, 248–249
- semantic characteristics of, 84
- unbound, 87–88,
- also see Tables 4.1 and 4.2*, 82–83

numerals

- as attributes in verbless clause complements, 370–371
- decimalization of the overcounting numeral system, 112–113, 117–120
- description of** cardinal, 112–117
- distributive, 54, 275–277
- numeral adverbs, 277–278
- ordinal, 277
- overcounting numeral systems in other languages of Nagaland, 119–120

oblique case marking, 175–189

onomatopoeia, 129–130

particles

- description of**, 131–144
- occurring within the noun phrase, 202

passive voice,

- absence of, 168–169

- functional equivalent of, 153
- pauses identifying postposed constituents, 152
- perfective aspect marker, 338–340
- personal pronouns
 - allomorphy in dual series, 91
 - allomorphy in plural series, 92
 - description of**, 89–94
 - diachronic source of third person singular form, 90
 - determined by nominal deictics, 93
 - dual, 89, 91
 - inclusive/exclusive contrast, 89
 - in core and oblique noun phrases, 84, 86–87, 175–176, 181,
 - plural series, 92–93
 - possessive marking function of, 90, 249
 - syncretism with possessive pronouns, 90
- Phom, 4, 5
- phonological processes,
 - consonant assimilation, 54–55
 - formation of phonetic diphthongs, 25–26
 - glide assimilation, 54
 - resyllabification, 55
 - schwa deletion in prefixes, 52
 - sporadic vowel harmony, 53
 - vowel deletions, 55
 - vowel dissimilation, 53–54
- phonotactics, 23–24
- phrasal conjunction
 - description of**, 127–129
 - diachronic origin of, 128
- pivot, absence of, 165–168, 419–421
- possession
 - absence of recursive, 252
 - collective, 96, 251
 - description of**, 248–255
 - encoded by juxtaposed nouns, 251–252
 - encoded by pronouns, 90, 249–251
 - encoded by the agentive nominalizer, 252–255
 - encoding strategies, 249
 - of semantic classes of body parts and kinship nouns, 253–254
- possessive pronouns
 - as heads of oblique noun phrases, 94–95, 96, 175, 181, 186, 189
 - comparison with personal pronouns, 94
 - use with reflexive marker, 273–274
 - description of**, 94–97
- possessive verbal clauses, 383–385
- postlexical intonation
 - description of**, 73–77
 - in converb clauses, 75–76, 431–432
 - in matrix clauses, 76–77, 143–144
 - in noun phrases, 74–75
- pragmatic influences
 - on constituent order, 153
 - on core case marking, 158, 389
 - on core case marking in other languages, 155
- indefinite pronouns
 - description of**, 100–105
 - negative indefinite, 101–102
 - non-specific indefinite, 103–104
 - reduplicated indefinite pronouns, 105–105
 - specific indefinite, 102–103
- purposive clauses, 265–266
- quotative particle
 - complement marking function, 453
 - description of**, 131–132
 - in discourse connectives, 125
- reciprocal derivation, 198–200, 321–322

reduplication

- as the functional equivalent of continuative aspect in imperative clauses, 397–397
- in distributive numerals, 275–276
- in durative converb constructions, 441–442
- of negative indefinite pronouns, 104–105
- of simultaneous converb stems, 279–280, 287, 414
- of verb roots in ascriptive clauses, 366
- of verb stems encoding frustrative modality, 331

reflexive suffix, 273–274

relative clauses

- and similarity to adjective of degree derivations, 262
- and negation, 233–234
- aspect and mood marking in, 231–234

description of, 217–236

differentiated from nominalizations, 217–218

formed with the agentive nominalizer, 229–230

formed with the locative nominalizer, 230–231

headed, 219–227

headless, 227–229

historical origin of, 218

internally-headed, 223–225

pre- and post-head positions of occurrence in, 208–209, 220–221

pronominal heads of, 221–222

relativizable positions of noun phrases, 226–227

restrictive versus non-restrictive meanings of, 220–221

relative-correlative constructions, 99, 134–135, 234–236

reportative particle, 132–133

resultant state suffix, 336–338

sal languages, 118, fn. 8

Saxena, 132, 161

segmental free variation

between [u] and [o], 45

in the vowel of the resultant state marker, 44, 337

in the vowel of the perfective aspect marker, 44, 338

semantic gender markers, 82, 270–271

stops, 30

subject and object, evidence against, 169

syllable

description of, 23–25

dialectal differences in structure, 25

domain of tone, 67

minimum number in a word, 49

permitted coda consonants, 24

temporal antecedence, expression of, 447

tense and modality

anterior, 353–354

immediate future, 351–353

immediate future and epistemic authority, 352

past, 346–348

present, 348–349

description of, 344–354

summary of TAM categories, 354

terminative nominal suffix, 273

Tibeto-Burman languages of Nagaland, 4

time words, 122–123

traditional divisions of the day, 123

tonally unspecified prefixes, 65

tone

autosegmental analysis of, 58–59, 62–64

- coalescence, 64–65, 173–174
- floating high, 62–63
- overview of grammatical, 60–65
- combinations in phonological words, 68,
- methodology for analysis of, 59–60
- phonemic contrasts, 23, 26–27, 57–58,
- copying by tonally unspecified prefixes, 60, 65–67
- tone sandhi
 - across grammatical word boundaries, 50, 71
 - description of**, 67–73
 - generated by morpheme concatenation, 68
 - in adjectives of degree, 60–61
 - in compound nouns, 242
 - in nominal compounds, 72–73
 - triggered by the general nominalizer, 58–59
- topic continuity and intonation, 77
- topic marking
 - attenuating function in imperative clauses, 388
 - contrastive function of, 135
 - in dependent clauses, 133–134
 - in relative-correlative constructions, 134–135, 234–236
- topic particle
 - description of**, 133–136
 - marking headless noun phrases, 212
 - marking preposed arguments, 153
 - occurring within the noun phrase, 201–202
 - pairing with time words, 133
- transitivity, 81, 273, 279–287, 323
- typological overview of Mongsén, 2–4
- valency
 - in applicative derivation, 193–194
 - in causative derivations of intransitives, 190–191
 - in causative derivations of transitives, 194–196
 - in causativized imperative clauses, 398
 - in reciprocal/collective derivations, 198–200,
 - in transitive and intransitive clauses, 81
- verbal existential clauses
 - declarative, 375–379
 - description of**, 374–381
 - expression of ‘happening’, 382–383
 - expression of possession, 383–385
 - interrogative, 379–381
- verbless clauses
 - ascriptive, 362–371
 - constituent order in, 358
 - description of**, 357–371
 - identity, 358–360
 - interrogative identity, 361–362
 - interrogative ascriptive, 369
 - negated ascriptive, 368–369
 - negated identity, 360–361
 - unspecified for tense, 358
- verbs
 - ambitransitive, 81, 153, 168, 285–287
 - bare roots, 294
 - compound, 294–295
 - description of**, 279–354
 - ditransitive, evidence against, 283–285
 - expressing directional motion, 280–282
 - expression of property concepts, 80
 - formal criteria for word class, 79
 - negated stems, 291–292
 - nominalization of roots, 292–294
 - predicate structure, 287–290
 - prefixes of, 290–294

reduplicated roots, 279–280, 287
stative intransitive, 280
transitive, 282
transitivity classes, 279
voice onset time, 2, 28–31 32–38,
39–44
voiceless sonorants, 29, 38, 47
vowel phonemes, 45–47
vowel sequences, 25–26

word

criteria for, 49–51

word classes, closed

description of, 88–144

nominal modifiers, 106

case marking clitics, 121

*also see nominal deictics, lexical
quantifiers, adverbs, numerals, time
words, and phrasal conjunction*

word classes, open

criteria for nouns, 81–84

description of, 79–88

overview of, 79

criteria for verbs, 79–81

Yimchungrü, 10

zero anaphora, 153, 168–169

