

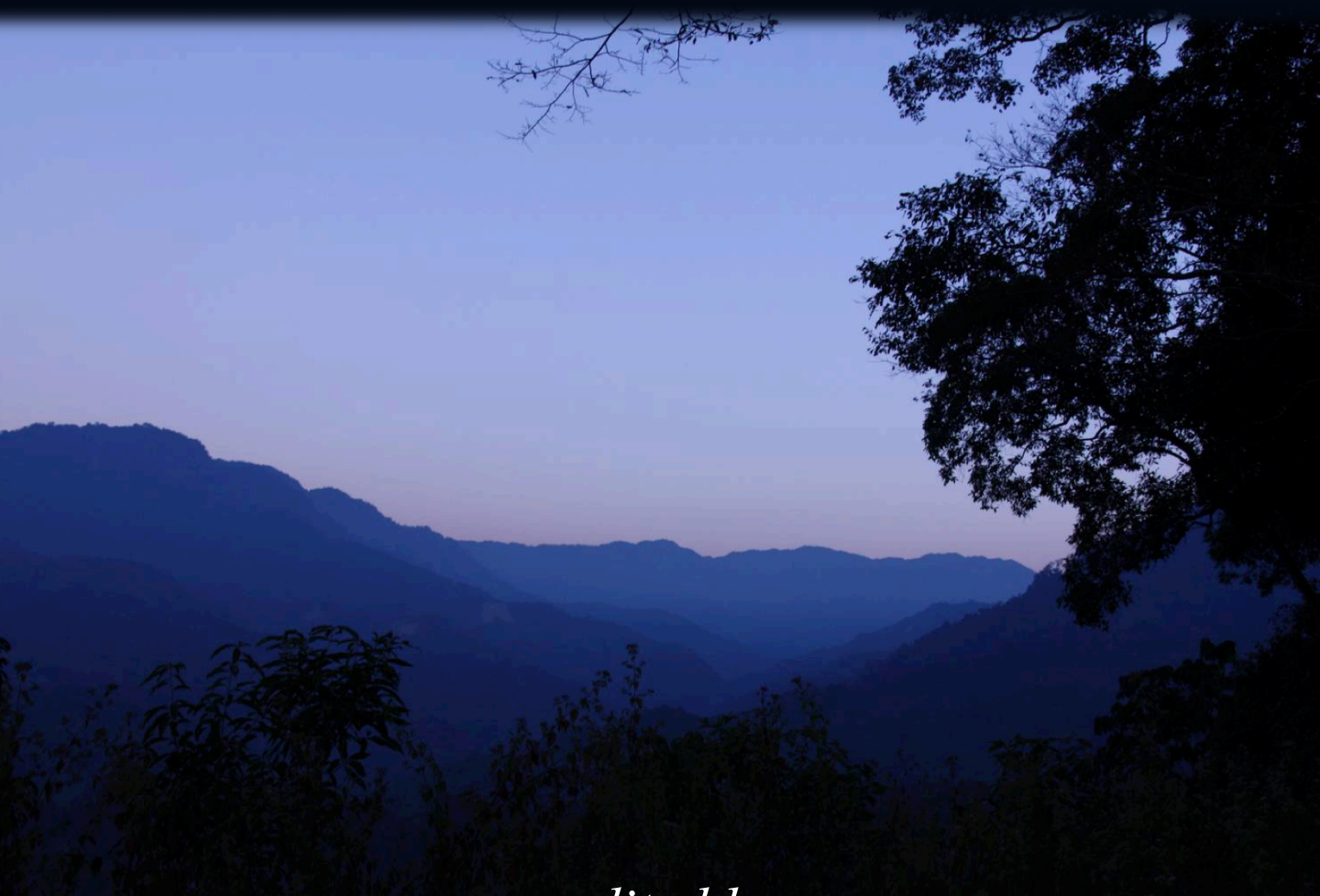


Asia-Pacific Linguistics

Open Access

College of Asia and the Pacific
The Australian National University

North East Indian Linguistics 8 (NEIL 8)



edited by

Linda Konnerth, Stephen Morey
and Amos Teo

**North East Indian Linguistics 8
(NEIL 8)**

edited by

Linda Konnerth, Stephen Morey and Amos Teo

This is the eighth volume of North East Indian Linguistics, a series of volumes for publishing current research on the languages of North East India, the first volume of which was published in 2008. The papers in this volume were presented at the 9th conference of the North East Indian Linguistics Society (NEILS), held at Tezpur University in February 2016 to mark the 10th anniversary of NEILS. The papers for this anniversary volume continue the NEILS tradition of research by both local and international scholars on a wide range of languages and topics. This eighth volume includes papers on small community languages and larger regional languages from across North East India, and presents detailed phonological, semantic and morphosyntactic studies of structures that are characteristic of particular languages or language groups, alongside sociolinguistic studies that explore language attitudes in contexts of language shift.



Asia-Pacific Linguistics

Open Access

North East Indian Linguistics 8 (NEIL 8)

edited by

Linda Konnerth, Stephen Morey and Amos Teo



Asia-Pacific Linguistics

Open Access

EDITORIAL BOARD: I Wayan Arka, Danielle Barth, Don Daniels, Bethwyn Evans, Nicholas Evans, Simon Greenhill, Gwendolyn Hyslop, David Nash, Bruno Olsson, Bill Palmer, Andrew Pawley, Malcolm Ross, Hannah Sarvasy, Jane Simpson.

Published by Asia-Pacific Linguistics
College of Asia and the Pacific
The Australian National University
Canberra ACT 2600
Australia

Copyright in this edition is vested with the author(s)
Released under Creative Commons License (Attribution 4.0 International)

First published: 2018

ISBN: 978-1-922185-41-9

URL: <http://hdl.handle.net/1885/155280>



A catalogue record for this
book is available from the
National Library of Australia

Cover photo: Patkai range at Manmau, Arunachal Pradesh.
Photographer Stephen Morey.

Contents

About the Contributors
A Note from the Editors

Field reports

1. Typological characteristics of Mnar1
Ruth Rymbai
2. Case marking in Liangmai17
WichamDinbo Mataina

Languages in Arunachal Pradesh

3. High vowel devoicing in Hrusso Aka33
Vijay A. D'Souza
4. A cross-varietal description of modifiers of basic colour terms47
in Tangsa-Nocte
Kellen Parker van Dam
5. Language attitude among Noctes of Bordumsa under Changlang district65
of Arunachal Pradesh
Trisha Wangno

Sino-Tibetan/Trans-Himalayan morphosyntax

6. The inclusive-exclusive distinction in Kuki-Chin and Naga Belt languages75
Scott DeLancey
7. Verb stem alternation in Pangwa Tangsa86
Stephen Morey
8. Nominalizations in Bodo110
Prafulla Basumatary
9. Reflexives and reciprocals in Dimasa126
Monali Longmailai

Verbal person indexation in languages at the Indo-Myanmar border

10. Hierarchical indexing and the inverse in Muklom Tangsa141
Mijke Mulder
11. Verb-pronominal agreement in Mara156
Jyoti Sharma
12. An introduction to the Nocte verb172
Iftiqar Rahman

Assamese studies

13. On the anatomy of an entry in an Assamese-Assamese-English dictionary186
Mouchumi Handique
14. Indefinite and definite quantifiers in Assamese: Some observations196
Gautam K. Borah
15. Changing forms and their usage in Assamese among Assamese youths208
Pushpa Renu Bhattacharyya

About the Contributors

Editors

Linda KONNERTH (linda.konnerth@mail.huji.ac.il) is a Postdoctoral Research Associate at the Martin Buber Society of Fellows at the Hebrew University in Jerusalem, Israel, and is also affiliated with the University of Oregon, USA. For her PhD dissertation, she has written a grammar of Karbi, and since then she has been working on the Northwestern Kuki-Chin (or “Northwestern South-Central”) languages of Chandel District in Manipur. She has mostly published on topics in historical-comparative linguistics and morphosyntax in Trans-Himalayan languages.

Stephen MOREY (s.morey@latrobe.edu.au) is a Senior Lecturer in the Department of Languages and Linguistics, La Trobe University. He is the author of two books (and multiple articles) on tribal languages in Assam, from both Tai-Kadai and Tibeto-Burman families. He is the secretary of the North East Indian Linguistics Society and has also written on the Aboriginal languages of Victoria, Australia.

Amos TEO (amosbteo@gmail.com) is a PhD candidate in the Department of Linguistics at the University of Oregon. He also completed an MA in Linguistics at the University of Melbourne and has been working on Tibeto-Burman languages of Northeast India and Nepal since 2007. In addition to language documentation, his current research projects range from experimental work on the production and perception of linguistic tone in various Tibeto-Burman languages to a corpus study of semantically/pragmatically-motivated case markers in Sümi, a Tibeto-Burman language of Nagaland.

Authors

Prafulla BASUMATARY (prafulla@gauhati.ac.in) is a Postdoctoral Research Fellow in the Department of Linguistics at Gauhati University, Assam. He is the author of a book *Verbal semantics in a Tibeto-Burman language: The Bodo verb*. He has recently started working on initial literacy measures for Bodo learners as part of the Bodo mother-tongue based multilingual education (MTB-MLE) programme.

Pushpa Renu BHATTACHARYYA (dr.pushparb@gmail.com) is a Research Associate in the Centre for Endangered Languages (CFEL) Tezpur University, Assam, India. Her research interests include language documentation and description.

Gautam K. BORAH (gkbtez@gmail.com) is Professor at the Dept of English and Foreign Languages, Tezpur University, Assam. His areas of interest are linguistic categorization, construction grammar, and the grammars of languages spoken in North East India. He has widely published on classifiers and other grammatical aspects of Assamese.

Vijay A. D'SOUZA (vijay.dsouza@campion.ox.ac.uk) is a PhD student at University of Oxford. He is working on Hrusso Aka, an endangered language spoken in West Kameng District of Arunachal Pradesh. Along with his PhD studies he is also documenting Hrusso Aka with the help of an ELDP (Endangered Languages Documentation Programme) scholarship and is involved in orthography and curriculum development for Hrusso Aka. His areas of interest include phonology, language documentation, description and revitalization.

Scott DELANCEY (delancey@uoregon.edu) is Professor of Linguistics at the University of Oregon. He has published extensively in the areas of Tibeto-Burman, Southeast Asian, and North American linguistics and on functional syntax and grammaticalization. He also works with the Northwest Indian Language Institute in Oregon.

Mouchumi HANDIQUE (handiquemouchumi@gmail.com) is a PhD research scholar in the Department of Linguistics, Gauhati University, Assam. She is writing her PhD dissertation on compiling an Assamese-Assamese-English Bilingual dictionary.

Monali LONGMAILAI (monalilong@gmail.com) is an Assistant Professor at the Centre for Endangered Languages, Tezpur University, Assam. She received her PhD degree in 2014 from the department of Linguistics, North-Eastern Hill University, Shillong, with research on the morphosyntax of Dimasa. Her main interests are language documentation, language typology, historical linguistics and Northeast Indian languages.

WichamDinbo MATAINA (wicham.mataina@gmail.com) is with the Centre for Endangered Languages, Sikkim University as an Assistant Professor. His PhD study, entitled "A Descriptive Grammar of Liangmai" was completed at Jawaharlal Nehru University, New Delhi. His research interest dwells on cultural linguistics, language documentation and language description. He is currently leading descriptive work on Sherpa and Gurung spoken in India under the Centre for Endangered Languages, Sikkim University, Gangtok.

Mijke MULDER (mijke.mulder@gmail.com) is a PhD student at the Centre for Research on Language Diversity, based at La Trobe University, Australia. Her research interests include Tibeto-Burman, Semitic, and language documentation. She is currently writing a descriptive grammar of Muklom Tangsa, a Tibeto-Burman language spoken in Northeast India.

Iftiqar RAHMAN (iftiqar.rahman@gmail.com) is a Linguistics graduate from the Gauhati University, Assam. He also did a second M.A. in Linguistics from the Payap University of Thailand. He is doing research on the Nocte language, which comes under the northern Naga branch of the Tibeto-Burman. His main areas of interests are language documentation, syntax and historical linguistics. His M.A. thesis has been published in the year 2016 titled "Verbal Agreement and Grammatical Description of Ha?wa Nocte".

Ruth RYMBAI (rymbairuth231986@gmail.com) is a Research Scholar at the Department of Linguistics, North Eastern Hill University, Shillong. Her areas of interest are phonetics and phonology, prosodic typology and language typology. She is currently doing research on prosody and intonation of Khasi.

Jyoti SHARMA (jyotitilakrajsharma@gmail.com) is a PhD research scholar at the Centre for Advance Studies in Linguistics, University of Delhi. Her research interests include language description and formal syntax.

Kellen Parker VAN DAM (kellenparker@gmail.com) is a graduate researcher in the Department of Languages and Linguistics, La Trobe University. His main areas of interest are tone, phonology, historical linguistics and minority language rights. In addition to Northern Naga, he has done work on Sinitic varieties within the Hakka and Wu language groups.

Trisha WANGNO (wangno.trisha@gmail.com) is a Research Scholar in the Department of English and Foreign Languages, Tezpur University. Her topic of research is “Socio-linguistic survey of Nocte in Changlang district of Arunachal Pradesh.”

A Note from the Editors

This is the eighth volume of *North East Indian Linguistics*, a series of volumes for publishing current research on the languages of North East India, the first volume of which was published in 2008. The papers in this volume were presented at the 9th conference of the North East Indian Linguistics Society (NEILS), held at Tezpur University in February 2016.

The North East Indian Linguistics Society was formed in 2005 by Jyotiprakash Tamuli (Gauhati University) and Mark Post (then at La Trobe University, Australia), subsequently joined by Stephen Morey (also from La Trobe University). From the time of the inaugural international conference in February 2006 at Gauhati University, up to the most recent conference at Assam University Silchar in January 2018, NEILS has been providing an opportunity for scholars within India, in the region around, and internationally to meet, to present their recent research, to interact and to help build a community of researchers working with common aim to understand the astonishing linguistic diversity of North East India.

Since 2008, a total of 116 papers have been published, all peer-reviewed by leading international specialists in the relevant subfields.

The papers for this anniversary volume are arranged into several sections, commencing with two field reports covering an overview sketch of Mnar, a Khasian language, written by Ruth Rymbai, and a more detailed study of case marking Liangmai, a “Naga” language of the Zeme group, presented by WichamDinbo Mataina.

The second section of the volume deals with detailed work on three languages spoken in Arunachal Pradesh. Vijay D’Souza presents an analysis of the phenomenon of devoicing high vowels in Hrusso Aka, suggesting that feature may explain why the language appears to have a complex and divergent phonology compared to its neighbours. Kellen Parker van Dam has undertaken a study of basic colour terms and their wide range modifiers in the Tangsa languages of Changlang district, while Trisha Wangno offers a sociolinguistic study of language attitudes in a group of Nocte people who have shifted from their ancestral villages to Bordumsa town, where Singpho and Assamese are the lingua francas.

Our next section deals with morphosyntax in the language family now known both by its older name – Sino-Tibetan, and the more recently coined Trans-Himalayan. Readers will note that both terms are found in this volume, referring to a family or phylum of languages that are related to each other. Scott DeLancey discusses an example of a linguistic process, the marking of distinctions between 1st person inclusive and exclusive forms in verbal agreement, that is found in the Kuki-Chin languages and shared by at least some of the languages in the ‘Naga Belt’. Stephen Morey’s paper takes an in-depth look at the phenomenon of Verb Stem Alternation, well studied and reported for the Kuki-Chin languages, but not much reported for the Northern Naga languages, and demonstrates the diversity of this phenomenon even within closely related linguistic varieties. Prafulla Basumatary presents a detailed study of the phenomenon of nominalization in his native language Bodo, adding to the considerable literature on the topic in Sino-Tibetan/Trans-Himalayan, observing that both derivation and clausal nominalization is found. This section is rounded off by another paper by another linguist working on her native language, Monali Longmailai with a comprehensive account of reflexives and reciprocals in another Bodo-Garo language, Dimasa.

A very exciting feature of the 9th NEILS conference was the number of papers on the topic of person indexation (agreement) on verbs in a range of languages, particularly from Northern Naga and Kuki-Chin. These two groups share both the feature of verb stem alternation and person indexation, and the three papers in this section present rich data that profoundly expand on knowledge on this topic. The first paper, by Mijke Mulder, documents both hierarchical and inverse participant marking in Muklom Tangsa. Jyoti Sharma, writing on Mara (Kuki-Chin) offers a careful study of another agreement system with hierarchical

features and Iftiqar Rahman investigates hierarchical marking in Nocte, a Northern Naga language.

Our final section covers the study of Assamese, the major language of the area that is an Indo-European language with a long written history. Mouchumi Handique launches this section with a detailed study of dictionary making for this important language of education and government, reporting on a long term project that led to the publication of a comprehensive new dictionary. Gautam Borah builds on a number of papers of recent work on classifiers and quantifiers in Assamese, some of which has been published in earlier NEIL volumes, discussing an area of syntax in Indic languages that is very special to the Eastern Indic area. Our final paper, by Pushpa Renu Bhattacharyya documents language attitudes and changing linguistic forms in the speech of younger Assamese speakers, rounding off this diverse and innovative volume.

This book represents the third volume published with Asia-Pacific Open Access. *North East Indian Linguistics* thus continues to be available for free download and easily accessible to everyone. We wish to thank everyone who helped bring this book into fruition, including the authors and peer reviewers, as well as our publisher.

Linda Konnerth

Jerusalem, Israel

Stephen Morey

Melbourne, Australia

Amos Teo

Eugene, Oregon, USA

Field reports

Typological characteristics of Mnar¹

Ruth Rymbai

North Eastern Hill University, Shillong

Abstract

This paper is an overture introduction to the typological features of Mnar, a Khasian language, which is classified under the Mon-Khmer branch of the Austro-Asiatic language family. Mnar inhabits the area known as Jirang, which falls under the Ri Bhoi district of Meghalaya. Although, Mnar shares common structural traits with other Khasian languages, it still exhibits unique properties, providing a great opportunity for in depth investigation of the intricacies of the typological features of Mnar. The paper includes discussion on phonological typology (consonant and vowel inventory, syllable structure and consonant clusters); morphological typology (word formation processes) and word order typology as found to exist in Mnar.

Citation	Rymbai, Ruth. 2018. <i>Typological characteristics of Mnar</i> . North East Indian Linguistics (NEIL) 8, 1-16. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.
Volume Editors	Linda Konnerth, Stephen Morey, Amos Teo
Copyright	© 2018, the author(s), release under Creative Commons Attribution license
URL	http://hdl.handle.net/1885/155280

1. Introduction

Mnar is a member of the Khasian group, a sub-branch of the Northern Mon-Khmer languages. Diffloth (2005) coined the term ‘Khasian’ to refer to the varieties related to Khasi spoken in the Khasi and Jaintia hills of Meghalaya. Mnar is spoken by the inhabitants residing in Jirang, which is situated in the Ri Bhoi district of Meghalaya. As per the Census of India (2011)², the total number of inhabitants in Jirang is approximately 30919. Mnar is not well studied linguistically and there is hardly any account of studies pertaining to Mnar. There is, however, a description of Mnar morpho-syntax in Koshy and Wahlang (2011). Further, Gurdon (1914) has listed Mnar as one of the Khasi dialects and mentioned briefly about the numeral system in Mnar. Although, he listed Mnar as one of the Khasi dialects, he still mentions that Mnar is affiliated to Synteng, Lakadong and Amwi rather than to Standard Khasi. There is no mutual intelligibility between Khasi and Mnar.

The purpose of this paper is to provide a typological classification of major constructions selected from the principle components: phonology, morphology and syntax; with a view of arriving at an understanding of the nature of such constructions in Mnar. Since there is very little work available that compares and contrasts the rich assortment of data found in Mnar, the empirical evidence is drawn from primary sources, i.e. field work and language consultants.

1.1. Data

This paper is an outcome of an MA dissertation conducted every year for fourth semester students of the Department of Linguistics, North Eastern Hill University, Shillong. The

¹ I would like to acknowledge the NEILS editorial board and the reviewers of this paper for their valuable remarks and suggestions.

² Census of India 2011, Meghalaya. District Census Handbook (Ri Bhoi) published by Directorate of Census Operations, Meghalaya.

students are required to study a previously unknown or lesser known dialect or language within the North East Region of India. Thus, confirming the fact that the author is not a native speaker of Mnar. The data collection for this dissertation was undertaken in the village of Jirang. The data analysed comes from fieldwork transcription in the form of word list, derived from Swadesh List. For sentences the consultants were asked to translate Khasi texts into Mnar, so that the investigators could note down and transcribe the spoken data. The drawback of the present study is that it only presents data elicited by way of translation. Prospective work will need to include naturally occurring data related through folktales, narratives and communicative events that people regard as important parts of their cultural heritage, occurring in oratory, song, dance recitals of poetry or story-telling, rituals, and so on. Unfortunately, no audio or video recordings were made in this fieldtrip. The following is a list of the informants involved in the fieldtrip:

Table 1 – List of Informants

Name	Age	Gender	Occupation	Languages Spoken
Nodis Thangkhiew	72	M	Farmer	Mnar, Khasi, Maram, Assamese, Bengali
Enias Rani	70	F	Farmer	Mnar, Khasi, Assamese
Dorathy Rani	40	F	Farmer	Mnar, Khasi, English
Convener Rani	35	M	Businessman	Mnar, Khasi
Overseas Rani	28	M	B.D. (research)	Mnar, Khasi, Bhoi, Pnar,
Nikulcy Rani	22	F	Farmer	Mnar
Twaston Thangkhiew	45	M	Govt. Servant	Mnar, Khasi, English
Ocarmis Wahlang	38	F	Govt. Servant	Mnar, Khasi, English
Daphisabet Wahlang	18	F	Student	Mnar, Khasi, English
Willina Wahlang	32	F	Teacher	Mnar, Khasi, English
Banlambha Wahlang	20	M	Student	Mnar, Khasi, English
Justice Syiem	35	M	Guard	Mnar, Khasi, Hindi, English
Hamberly Wahlang	66	M	Chief	Mnar, Khasi, Assamese, English
Chesteroy Khyrniam	30	M	Teacher	Mnar, Pnar, Khasi, English, Hindi

2. Phonological Typology

The phonology of Khasi and few other dialects including Pnar, Lyngngam and Bhoi has been presented adequately by Nagaraja (1985, 1993 1996, and 2014) and Ring (2012, 2014, and 2015). However, as mentioned by Diffloth and Zide (1992), the “so called Khasi dialects, such as Synteng (Pnar), Lyngngam and Amwi (War) are clearly distinct but related languages” (cited in Koshy and Wahlang 2011). Daladier (2011) in an extensive research on Pnar, War, Lyngngam and Standard Khasi also recognizes the fact that the aforesaid varieties are mutually unintelligible and therefore adduced that these varieties are four disparate languages. In keeping with the typical typological phonological profile of Austro-Asiatic (AA) languages, Mnar exhibits the following characteristics.

2.1. Consonants

Mnar has five places of articulation of stops (bilabial, alveolar, palatal, velar and glottal) and four for nasals (bilabial, alveolar, palatal and velar). The fricatives available are /s/ and /h/. Furthermore, Mnar has one rhotic /r/, one lateral approximant /l/, and two semi vowels /w/ and /j/.

Additionally, Mnar has two stop series, usually distinguishing between voiceless (unaspirated) and voiced stops. Aspiration is phonemic, thereby giving way to three way voicing contrast in oral stops. Peculiarly, there is also the presence of voiceless aspirated bilabial /b^h/ without voiced aspirated alveolar /d^h/ in Mnar (eg. *b^hep* ‘many’; *b^hin* ‘another’; *b^hiʔ* ‘good’ and *b^hew* ‘spade’). Nagaraja (1996) found the occurrence of voiced velar stop /g/ in his study of Lyngngam, but in Mnar its occurrence is limited and at the same time suspicious. First and foremost, it occurs as an alternate variant of *ka* ‘3 singular feminine’ *ka~ga* and *ki* ‘3 plural’ *ki~gi*. Secondly, there is no phonemic contrast for /g/ and its occurrence is limited to the initial position only as in *gana* ‘this’; *gatai* ‘that’; *gitai* ‘those’; *gira* ‘relative marker’ etc.

Table 2 – Consonant Inventory of Mnar

	Bilabial	Alveolar	Palatal	Velar	Glottal
Stop					
Unaspirated	p b	t d	c ɟ	k	ʔ
Aspirated	p ^h b ^h	t ^h	c ^h	k ^h	
Nasal	m	n	ɲ	ŋ	
Trill		r			
Fricative		s			h
Lateral		l			
Approximant	w		j		

The following minimal and sub minimal sets illustrate the contrasts between Mnar consonants:

Table 3 – Consonant Phoneme Oppositions in Mnar

Stops	
/p/ <i>pɔʔ</i> ‘belly’	/b/ <i>bɔʔ</i> ‘to tie’
/t/ <i>tɛw</i> ‘tree’	/d/ <i>dɛw</i> ‘cloth’
/c/ <i>cut</i> ‘choose’	/ɟ/ <i>ɟut</i> ‘count’
/k/ <i>ksei</i> ‘dance’	/p/ <i>psei</i> ‘sister-in-law’

Aspirated vs. Unaspirated	
/p/ <i>pi</i> ‘2PL’	/p ^h / <i>p^hi</i> ‘ten’
/b/ <i>biʔ</i> ‘poison’	/b ^h / <i>b^hiʔ</i> ‘good’
/t/ <i>tai</i> ‘that’	/t ^h / <i>t^hai</i> ‘herd’
/c/ <i>ci</i> ‘see’	/c ^h / <i>c^him</i> ‘blood’
/k/ <i>kam</i> ‘work’	/k ^h / <i>k^ham</i> ‘handful’
Trill vs. Lateral	
/r/ <i>reʔ</i> ‘auxiliary verb’	/l/ <i>leʔ</i> ‘past tense marker’
Nasals	
/m/ <i>mar</i> ‘each’	/n/ <i>nar</i> ‘iron’
/n/ <i>nar</i> ‘iron’	/ɲ/ <i>naʔ</i> ‘drive’
/ɲ/ <i>t^haɲ</i> ‘red’	/ŋ/ <i>t^haŋ</i> ‘burn’
Fricatives	
/s/ <i>saʔ</i> ‘stay’	/h/ <i>ha</i> ‘locative case marker’
Approximants	
/w/ <i>waʔ</i> ‘also’	/j/ <i>jaʔ</i> ‘black’

As already mentioned, Mnar has four nasals; /m n ɲ ŋ/. They can occur in both onset and coda positions as illustrated below:

Table 4 – Distribution of Nasals

Phoneme	Onset	Coda
m	<i>mat</i> ‘eye’	<i>crɲam</i> ‘green’
n	<i>nɔ</i> ‘leg’	<i>san</i> ‘five’
ɲ	<i>ɲim</i> ‘cry’	<i>t^haɲ</i> ‘red’
ŋ	<i>ŋa</i> ‘1SG’	<i>tɾlaŋ</i> ‘ear’

Mnar has stop and nasal palatals such as /c/ and /ɲ/ in the final position of a word as shown below.

Table 5 – Palatal Nasal and Stop in word final position

Word	Gloss
<i>ninaic</i>	‘small’
<i>sɲeic</i>	‘stout’
<i>k^hlain</i>	‘strong’
<i>t^haɲ</i>	‘red’
<i>aɲ</i>	‘law’

The inventory of syllable-final consonants is smaller than that of onset consonants in Mnar in that there is only one series of stops, which are always voiceless and typically unreleased, i.e. accompanied by glottal restriction which stops the airflow such as *skep* ‘ribcage’ and *sut* ‘veins’. Other consonants that can occur in the coda position are /k/ (*wak* ‘many’); /ʔ/ (*sniəʔ* ‘skin’); /m/ (*chim* ‘blood’); /n/ (*lmen* ‘tooth’) and /ŋ/ (*tʰaŋ* ‘red’); /ŋ/ (*kŋrɔŋ* ‘long’); /r/ (*i: r* ‘two’).

2.2. Vowels

The richness of Mnar vocalic system is typically Khasian (Nagaraja 1985; 1996 and Ring 2012), with a minimum of three degrees of vowel height. Nagaraja (1985) listed 11 vowels in Khasi while Ring (2012) presented a total of 7 vowels in Pnar. Lyngngam has 9 vowels as per Nagaraja’s findings (Nagaraja 1996). Mnar has a total of 8 monophthong vowels: one high front vowel /i/, two mid front vowel series /e ε/, one low front vowel /a/, three back vowels /u o ɔ/ and one central vowel /ɨ/.

Table 6 – Vowel Inventory of Mnar

Height	Front	Central	Back
High	i	ɨ	u
Mid	e		o
Low-Mid	ε		ɔ
Low	a		

Table 7 – Vowel Contrast Pairs in Mnar

/i/ <i>im</i> ‘sibling’	/a/ <i>am</i> ‘progressive marker’
/a/ <i>sam</i> ‘take’	/e/ <i>sem</i> ‘bathe’
/a/ <i>ma</i> ‘2sm’	/u/ <i>mu</i> ‘mother’
/ε/ <i>εʔ</i> ‘loud’	/a/ <i>aʔ</i> ‘auxiliary verb’
/e/ <i>reʔ</i> ‘auxiliary verb’	/ε/ <i>εʔ</i> ‘loud’
/e/ <i>hen</i> ‘measuring unit’	/u/ <i>hun</i> ‘child’
/ε/ <i>et</i> ‘stick’	/ɔ/ <i>ɔt</i> ‘cut’
/i/ <i>liʔ</i> ‘go’	/e/ <i>leʔ</i> ‘past tense marker’
/a/ <i>saʔ</i> ‘stay’	/ɔ/ <i>sɔʔ</i> ‘fruit’
/u/ <i>tʰuʔ</i> ‘search’	/ɔ/ <i>tʰɔʔ</i> ‘write’
/i/ <i>wiʔ</i> ‘earthworm’	/ɔ/ <i>wɔʔ</i> ‘grandfather’
/o/ <i>noʔ</i> ‘throw’	/ɔ/ <i>nɔ</i> ‘leg’

Keeping in line with the Khasian languages (Nagaraja 2014: Ring 2014 and 2015), vowel length contrast is irregular and restricted to a subgroup of vowel in Mnar. Only /a:/ shows phonemic contrast in length. The long vowel /i:/ is inconsistent as well. Other examples with long vowels include: *pla:m* ‘cloud’; *sta:t* ‘wise’; *pa:m* ‘slice’; *i:r* ‘two’; *bi:m* ‘eat’; *ji:* ‘rice’ and *ki:ʔ* ‘climb’.

Table 8 – Vowel Length in Mnar

Short	Word with Gloss	Long	Word with Gloss
/a/	<i>sam</i> ‘take’	/a:/	<i>sa:m</i> ‘buy’
/a/	<i>am</i> ‘progressive marker’	/a:/	<i>a:m</i> ‘water’

Ring (2012) describes Pnar as having two types of diphthongs: (a) two vowel elements that constitute a single syllable and (b) a vowel sound and a glide which operate as a secondary vowel component. Mnar is similar to Pnar (Ring, 2012 and 2015) in its inventory of diphthongs. There are only two phonemic diphthong /ia/ and /ɔu/. The diphthong /ia/ is found in closed syllables as in *sniəʔ* ‘skin’; *sʔiaŋ* ‘seed’; *tʰiaʔ* ‘sleep’. While, /ɔu/ is found in open syllables as in *ksɔu* ‘dog’; *smɔu* ‘stone’; *hʌɔu* ‘door’. Mnar also shows the presence of a vowel sound with either a labial /w/ or palatal /j/ glide, that either precedes (an on-glide) or follows (an off-glide) the main vowel.

Table 9 – Mnar Diphthongs with on- and off-glides

Diphthongs	Description	Word	Gloss
/ei/	a front mid vowel with palatal off-glide	<i>tei</i>	‘hand’
/ai/	an open vowel with palatal off-glide	<i>mai</i>	‘face’
/oi/	a back mid vowel with palatal off-glide	<i>moi</i>	‘buffalo’
/ie/	a front mid vowel with palatal on-glide	<i>biet</i>	‘fool’
/ui/	a front close vowel with labial on-glide	<i>kʰuic</i>	‘clean’

2.3. Syllable Structure

The phonological word in Mnar consists of monosyllabic, sesquisyllabic and multisyllabic roots. Disyllabic words are most frequent and phonologically they are merely the composite of two syllables, each of which follows the same phonological requisites. Trisyllabic words are rare, if found to prevail they are an outcome of word formation (compounding). A syllable in Mnar may consist of an initial consonant cluster (CC), an obligatory vowel nuclei V, and an optional final consonant phoneme (C). The formulaic structure of the syllable is (C₁) (C₂) V (C₃). The smallest word shape allowed in Mnar is a single vowel nuclei and the largest word shape consists of a complex onset of two consonants, a diphthong nucleus and a coda consonant.

Table 10 – Structure of Syllables in Mnar

Monosyllables	Word	Gloss
V	<i>i</i>	‘1PL’
VC	<i>im</i>	‘niece’
CV	<i>mu</i>	‘mother’
CVC	<i>san</i>	‘five’
CCV	<i>sŋi</i>	‘house’
CVV	<i>lei</i>	‘three’
CCVC	<i>ksaŋ</i>	‘bitter’

2.3.1. Consonant Clusters

The combination of consonants in the initial position is rich in Mnar, usually all consonants can appear in the onset slot; however, sequences of the same place of articulation are avoided. Mnar onset clusters include the sequences of stop plus nasal, stop plus liquid, fricative plus stop, fricative plus nasal, liquid plus nasal etc.

Table 11 – Permissible Onset Clusters

Cluster	Word	Gloss
pn	<i>pnuʔ</i>	‘salt’
ps	<i>psen</i>	‘snake’
pr	<i>praʔ</i>	‘know’
pl	<i>pla:m</i>	‘cloud’
bl	<i>blaŋ</i>	‘goat’
tŋ	<i>tŋam</i>	‘cold weather’
tl	<i>tlɔt</i>	‘weak’
t ^h r	<i>t^hrei</i>	‘six’
t ^h m	<i>t^hmi</i>	‘war’
c ^h l	<i>c^hlam</i>	‘cold water’
kt ^h	<i>kt^har</i>	‘axe’
kl	<i>klɔŋsnam</i>	‘heart’
ks	<i>ksaŋb^hi</i>	‘good’
km	<i>kmen</i>	‘happy’
k ^h l	<i>k^hlɔu</i>	‘head’
sn	<i>sniɑʔ</i>	‘bone’
sk	<i>skɔu</i>	‘sit’
sm	<i>smɔu</i>	‘stone’
sŋ	<i>sŋeic</i>	‘stout’
st	<i>sta:t</i>	‘clever’

2.3.2. Sesquisyllabic Structure

A distinct syllable type termed ‘sesquisyllabic’ (Matisoff 1973:86) is found to exist in many AA languages (Jenny et al. 2014) wherein a disyllabic word consists of an initial unstressed syllable often called a minor (Henderson 1952) or presyllable followed by a stressed full syllable (main syllable). In this minor or presyllable the nucleus of the syllable is occupied by either of these sonorants (l,r,m,n, ŋ), in lieu of certain weak vowels and as such they carry the main weight of the first syllable in a disyllabic word. This phenomenon can be explained with the disyllabic word *ŋyndaŋ* ‘neck’. The presyllable in this example is *ŋyn* which is orthographically spelled as *ŋyn* and transcribed phonetically as *ŋŋ* or *ŋin*. Here the vowel is eliminated or rendered weak and the following nasal /n/ occupies the nucleus position, thereby giving rise to a sesquisyllabic structure. Other examples are listed in Table 12 below.

Table 12 – Sesquisyllabic Structure

Orthography	Phonemic Transcription	Gloss
ɣyndaŋ	/ɣndaŋ/	‘neck
tyrlaŋ	/trlaŋ/	‘nose’
tylliaŋ	/tlliaŋ/	‘shoulder’
pynjip	/pnjip/	‘cause to kill’

2.3.3. Suprasegmental

Since most AA languages are sesquisyllabic, they tend to be strongly iambic, wherein a weak and unstressed syllable is followed by a strong and full stressed syllable (Jenny et al., 2014). Likewise, the stress in Mnar is generally on the last syllable of the word.

3. Morphological Typology

Mnar is isolating, in that it is extremely analytic with words consisting of a single morpheme constituting a separate word and independent grammatical words. However, no language is purely or predominantly of one type. Thus, Mnar is also characterized by some degree of agglutination, mainly, in its technique of employing affixes to be juxtaposed to root words. The addition causes no significant changes in the root and the different affixes are readily identifiable and easily segmented from the root and from one another as illustrated in the examples below.

- (1) [u pitar] [bi:m sɔʔ=u]
 3SM PN eat fruit=3SM
 ‘Peter eats fruit’
- (2) [ga sap^hi] [pŋ-rɔʔ=ka ha-rara]
 3SF PN CAUS-praise=3SF ACC-REFL
 ‘Saphi praised herself’

Sentence (1) illustrates how Mnar comes close to being an isolating type. Each word in the sentence consists of just a single and free morpheme, except for the clitic =u ‘3SM’. Each morpheme is invariable in that the words are strung together in a sentence but without change; thus, *bi:m* ‘eat’ does not inflect to show person, number or tense.

In sentence (2) the word *pŋ-rɔʔ* ‘cause to praise’ consists of two morphemes *pŋ-*, a causative prefix, and *rɔʔ* meaning ‘praise’. The boundary between these two morphemes in the word is clear-cut; moreover, the identification of morpheme in terms of their phonetic shape is also straightforward. The anaphoric expression *ha-rara* is a combination of two morphemes *ha-* ‘accusative case’ and *rara* ‘personal reflexive’, here too the boundary is clear-cut.

The lack of inflectional categories as attested in Mnar is compensated by the extensive derivational morphology, including class changing derivational prefixes, compounding and reduplication. Compounding is common, and can consist of two or more nouns, a sequence of a noun plus verb, or a combination of two or more verbs.

Table 13 – Compound words

Combination	Compounds	Gloss	Gloss 2
Noun+Noun	<i>sʔiaŋ+tɲpɔŋ</i>	back+bone	‘backbone’
Verb+Noun	<i>dɔn+burɔm</i>	be+honour	‘honourable’
Verb+Adjective	<i>ksaŋ+kmen</i>	feel+happy	‘elated’
Noun+Verb	<i>a:m+biaʔ</i>	water+spit	‘saliva’

Derivational morphology in Mnar is operational through affixation attaining nominalization and causativization processes. Normally verbs are nominalised by adding prefixes as in the following table. Additionally, Mnar also has an agentive nominalising prefix *men-* as well as a causativizing prefix containing the element *pŋ-* is attached to a simplex verb *jip* meaning ‘die’ as in *pŋ-jip* ‘cause to kill’.

Table 14 – Word formation with derivational prefixes in Mnar

Prefixes	Word	Gloss 1	Gloss2
<i>t^hei-</i> (its function is to derive noun stems from verb and adjectives)	<i>t^hei+sali</i> <i>t^hei+t^huʔ</i> <i>t^hei+leʔkai</i>	Nomlz+lazy Nomlz+search Nomlz+play	‘sloth’ ‘survey’ ‘sports’
<i>ŋa-</i> (its function is to derive abstract nouns from verbs)	<i>ŋa+jɲŋar</i> <i>ŋa+praʔ</i> <i>ŋa+kɛt</i>	Nomlz+trouble Nomlz+know Nomlz+carry	‘problem’ ‘information’ ‘load’
<i>men-</i> (derives agentive nominals from verbs)	<i>men+anla:m</i> <i>men+cni</i> <i>men+stuʔ</i>	Nomlz+lead Nomlz+construct Nomlz+steal	‘leader’ ‘builder’ ‘thief’
<i>pŋ-</i> (derives causative stems from simple verb roots)	<i>pŋ+pʔem</i> <i>pŋ+sem</i> <i>pŋ+rwei</i>	Caus+live Caus+bathe Caus+fall	‘cause to live’ ‘cause/allow to take bath’ ‘cause to fall’

Reduplication is yet another common word formation process attested in almost all AA languages. It appears in various manifestations either as full or partial reduplication. Adjectives, question words and adverbs all have reduplicated forms in the various Khasian varieties (Ring 2014, Nagaraja 2014). This also holds true for Mnar, in that it uses reduplication to mark emphasis and intensify the meaning through complete repetition of adverbials and deictic words. Full reduplication is constructed by repeating the element in an identical form whereas in partial reduplication a consonant is changed in the initial position in the repeated element. It is also significant to mention that interrogative pronouns are not always reduplicated. The root morpheme also imparts meaning of their own as in these examples: *ju* ‘what’ and *ʔɔkro* ‘who’.

Table 15 – Examples of reduplication in Mnar

Reduplicating Elements	Part of speech	Function	Gloss
Full Reduplication			
<i>imra-imra</i>	Adverb	Deictic (Intensification)	‘there’
<i>mən-mən</i>	Adverb	Manner (Intensification)	‘slowly’
<i>ɟkro-ɟkro</i>	Interrogative Pronoun	Grammatical Function	‘someone’
Partial Reduplication			
<i>ɟk^ha-mək^ha</i>	Indefinite Pronoun	Distributive	‘some’

4. Word Order Typology

Obligatory morphological marking³ of tense, agreement, number or any other morpho-syntactic category generally expressed by inflection, is absent in Mnar. As in Pnar (Ring 2015) syntactic and grammatical relationships in Mnar are primarily shown by word order. Having said so, on the surface level Mnar exhibits three types of word order (SVO/VSO/VSO) and is characterized by a fairly rigid word order, wherein constituents cannot be freely moved from one position to the other. It has a rich set of functional words which mark the grammatical properties of phrases and clauses. It is head initial displaying modified-modifier ordering and has a noun classifier system. Before describing the word order typology of Mnar it is important to discuss the pronominal and gender system.

Table 16 – Pronominal chart of Mnar

Singular	Nominative	Accusative	Genitive/Possessive
First Person	<i>ɟa</i> ‘I’	<i>ha-o</i> ‘me’	<i>ɟo-o</i> ‘my’
Second Person	<i>ma(M) pa(F)</i> ‘you’	<i>ma(M) pa(F)</i> ‘you’	<i>ɟo-ma(M)</i> <i>ɟo-pa (F)</i> ‘your’
Third Person	<i>u / ka~ga</i> ‘he’ / ‘she’	<i>wei / kai</i> ‘him’ / ‘her’	<i>ɟo-wei/ɟo-kai</i> ‘his’ / ‘her’
Plural			
First Person	<i>wi</i> ‘we’	<i>wi</i> ‘us’	<i>ɟo-wi</i> ‘our’
Second Person	<i>pi</i> ‘you’	<i>pi</i> ‘you’	<i>ɟo-pi</i> ‘your’
Third Person	<i>ki~gi</i> ‘they’	<i>kei~gei</i> ‘them’	<i>ɟo-kei</i> ‘their’

³ Case is morphologically marked and indicated by prepositions in Mnar.

Mnar has a rich pronominal system with two number distinctions: singular and plural. Feminine and masculine gender distinction is seen in second person singular *ma* (M) and *pa* (F). Further it must be pointed out that Mnar has a nominative-accusative form of alignment in terms of case marking and/or verb agreement, wherein the subject (S) of an intransitive verb has the same case as the subject or agent (A) of a transitive verb, contrasting with the patient (P) / object (O) of a transitive verb which gets coded differently. Third person pronoun *u* and *ka* can function as gender clitics, personal pronouns as well as agreement markers in Mnar.

Gender distinction in Mnar is seen not only in pronouns but also in lexical nouns (nominals). In Mnar nouns cannot function without gender, it is obligatorily attached before a noun. Some animate nouns can be divided into female and male. The following are the gender clitics: *ka~ga* (feminine); *u* (masculine) and *ηa* (neuter).

(3) *ga sara*
 F PN
 ‘Sara’

(4) *u jɔn*
 M PN
 ‘John’

(5) *ηa hai*
 N thing
 ‘thing’

It may be also added that first, second and third person plural appears to have evolved from the feminine singular counterparts as evidenced in the following example:

(6) *o* ‘first person singular (non-nominative)’ → *wi* ‘first person plural’
pa ‘second person feminine singular’ → *pi* ‘second person plural’
ka~ga ‘third person feminine singular’ → *ki~gi* ‘third person plural’

4.1. Verbal Agreement

In Mnar the verb obligatorily agrees with 3rd person subjects NP in terms of person and number and gender but there is no agreement for non-3rd person subject. The subject is marked by the pronominal clitics placed after the verb or predicate. The pronominal agreement clitic have the same shape as personal pronouns, the third person is marked by *u* (masculine) and *ka* (feminine) respectively in the singular, and by *ki* in the plural; when an animate noun stands as the subject NP, it agrees with the verb by its clitic form. Here, it must be mentioned that post verbal clitic form which I consider to be verb agreement is predominantly used rather than its free independent preverbal base. Furthermore, Mnar is also a pro-drop language, where agreement enables the independent subject pronoun to be dropped.

(7) [*u jɔn*] [*ieid = u*] [*ha ga meri*] (agreement pattern)
 3SM PN love=3SM ACC F PN
 ‘John loves Mary.’

In sentence (7) the highlighted **u** marks agreement of the verb *ieid* ‘love’ with its subject *u jɔn* ‘John’. The subject is third person singular masculine, which is also marked by the *u* directly preceding the head noun *jɔn* ‘John’. Further, the agreement marker is located after the verb.

- (8) *sa ji=ki*
eat rice=3PL
‘They eat rice.’

In an alternate way sentence (8), can be seen as an example where the pronominal marker *ki* occurs as a bound pronominal rather than displaying pro-drop.

Further analysis show that if the object is animate as in (9), it does not form a grammatical unit with the Verb Phrase and the pronominal marker occurs immediately after the verb, on the other hand if the object is inanimate as in (10), then it gets incorporated in the Verb Phrase and the pronominal marker follows the object.

- (9) [*u jɔn*] [*ieid = u*] [*ha ga meri*]
3SM PN love=3SM ACC F PN
‘John loves Mary.’

- (10) [*ga meri*] [*tiej haicbi:m=ka*]
3SF PN cook food=3SF
‘Mary cooks food.’

Authors, who studied subject-verb agreement in Khasi and other varieties related to it, differ (a) in their analyses of what constitutes subject-verb agreement and (b) the choice of terminology employed in the description of such patterns. For instance, Nagaraja (1993) claims Khasi has a feature of concordial agreement where the pronominal marker which occurs in the subject (Noun Phrase) gets repeated in the verb (Verb Phrase). Ring (2015) on the other hand says there is no agreement on verbs in Pnar. However, contrary to his view is Koshy (2007) who states that in Pnar the agreement of the subject with the verb appears as an enclitic on the verb/verb phrase.

4.2. Basic word order typology

The word order characteristics discussed in this section will be based on the writings and research of Greenberg (1963) and Dryer (1992). At the surface level, Mnar seems to exhibit three different types of word order. They are: (a) SVO, (b) VSO and (c) VOS.

The following is an illustration to show the basic word order of Mnar (SVO) in a simple declarative sentence.

- (11) [*ga meri*] [*tiej*] [*haicbi:m=ka*]
3SF PN cook food=3SF
‘Mary cooks food.’

Sentence (11) exhibits S-V-O word order where the subject/agent precedes the main verb; and the object/patient follows the verb. The post verbal subject clitic *ka* ‘3SF’ occurs after the object *haicbi:m* ‘food’. However, possible alternate word orders also exist in Mnar as in the following examples. A further research on word order variations is needed as no conclusion has been arrived at as yet.

- (12) *ieid=ki wei (VSO)*
 love=3PL 3SM (ACC)
 ‘They love him.’
- (13) *deʔ a:m wi i:rwat si siŋ (VOS)*
 drink water 1PL twice QUANT day
 ‘We drink water twice a day.’

Mnar has prepositions such as *ha, i, dimaʔ* and as predicted by the implicational universal (Greenberg 1963) for verb medial language the genitive follows the head noun.

- (14) *am saʔ=ka i s^hillɔŋ*
 PROG live=3SF LOC PN
 ‘She still lives in Shillong.’
- (15) *[ga im] [jɔ ga meri]*
 3SF niece GEN 3SF PN
 ‘Mary’s niece’

4.3. Main Clause S-V-O order in Mnar

Word order in intransitive, transitive, ditransitive, complement and relative clauses are listed below with appropriate examples.

- (16) *[u ban] [lhɔʔsari=u]*
 3SM PN laugh=3SM
 ‘Ban laughs.’ (S-V)
- (17) *[u jɔp] [leʔ ai] [ga bɔl] [ha u ban]*
 3SM PN PST give F ball PR.OBJ 3SM PN
 ‘Jop gave the ball to Ban.’ (S-V-DO-IO)
- (18) *[u jɔn] [leʔ oŋ=u] [ba ŋat liʔ ŋa i siŋ]*
 3SM PN PST say=3SM COMP AUX go 1SG LOC house
 ‘John said that I should go to the house.’ (S-V-Comp)
- (19) *[gana kai ga nandiʔ][gara ga laŋai ŋa sur]*
 this COP 3SF girl REL 3SF good N voice
 ‘She is the girl whose voice is good.’ (S-V-Rel)

4.4. Prenominal modifiers

Demonstratives precede the head noun.

- (20) *gana ga ksem*
 this F bird
 ‘this bird’

Classifiers and numerals precede the head noun.

(21) *i:r* *k^hlen* *gi* *kət*
 two CLF PL book
 ‘two books’

(22) *lei* *bei* *gi* *ant^hei*
 three CLF PL girls
 ‘three girls’

Quantifiers precede the head noun.

(23) *bat* *k^hlen* *gi* *kət*
 all CLF PL book
 ‘all the books’

Qualifiers precede the head noun and follow the adjective.

(24) *he?* *palat* *ga* *dəwsla*
 big QUAL(too) F shirt
 ‘The shirt is too big.’

4.5. Postnominal modifiers

Adjectives follow the head noun.

(25) *ga* *ant^hei* *simat*
 F girl beautiful
 ‘beautiful girl’

The comparative and superlative constructions in Mnar use the particles *hap* and *tam*.

(26) [*u* *jən*] [*hap* *kɲaŋ=u*] [*ha* *ka* *meri*]
 3SM PN COMPR tall=3SM ACC 3SF PN
 ‘John is taller than Mary.’

(27) [*kɲaŋ tam*] [*u* *jən*] [*pa?ei* *ka* *klas*]
 tall COMPR 3SM PN in F class
 ‘John is the tallest in the class.’

4.6. Order of Auxiliary verb

The auxiliary verb precedes the main verb.

(28) *ɲat* *pule=ma* *mən-mən*
 AUX read=2SM slow-INTENS
 ‘You should read slowly.’

4.7. Position of the interrogative

Question words occur initially whereas Yes/No question clauses have two types of constructions. The main verb or auxiliary verb *lep* ‘can’ is either put sentence initially as in (28) or question can be expressed with or without a rising intonation.

(29) *mnə ləŋ ma?*
 how COP 2SM
 ‘How are you?’

(30) *lep ma ŋa li? i agra duŋkni*
 AUX 2SM INF go LOC PN tomorrow
 ‘Can you go to Agra tomorrow?’

4.8. Position of the Negative particle

Mnar makes use of double negation to negate affirmative sentences. The negative particle *ba* meaning ‘not’ is the standard preverbal negative in Mnar. Following the verb is a negative marker *ra-* which is prefixed to third person pronouns such as *wei/kai/kei/wi* (him/her/them/we). The paper does not state any conclusive argument about the nature of the negative particles found in Mnar, hence further analysis is required in the near future.

(31) *[u karo] [pŋjip][ha ga blaŋ]* (affirmative)
 3SM man kill ACC F goat
 ‘The man killed a goat.’

(32) *[u karo] [ba a? pŋjip] [ha ra-wei ga blaŋ]* (negative)
 3SM man NEG AUX kill ACC NEG-3SM F goat
 ‘The man did not kill a goat.’

5. Conclusion

This work is an attempt to make few pertinent observations on the typological features of Mnar in general and to give a brief account of some important features as a means to classify the language in question. The findings presented in this paper do validate the fact that Mnar does share a number of features with other Austro-Asiatic languages. However, as matter of course, Mnar has some peculiar characteristics of its own, thus demonstrating that each language retains its identity in spite of intense contact with other languages.

Abbreviation

1SG	1st person singular	FEM	feminine
2SM	2nd person singular	DAT	dative
	masculine	GEN	genitive
3SM	3rd pers singular masculine	INF	infinitive
3SF	3rd pers singular feminine	LOC	locative
ACC	accusative	NOMLZ	nominalizer
AUX	auxiliary	O	object
CLF	classifier	PL	plural
COMP	complementizer	PR.OBJ	primary object
COMPR	comparative	PREP	prep

PROG	progressive marker	QUAL	qualifier
PST	past	QUANT	quantifier
RED	reduplication	V	verb
REL	relative	=	clitic boundary
S	subject		

References

- Comrie, Bernard. (1981). *Language Universals and Linguistic Typology*. Chicago, Chicago University Press.
- Croft, William. (1990). *Typology and Universals*. Cambridge, Cambridge University Press.
- Daladier, Anne. (2011). "The group Pnaric-War-Lyngngam and Khasi as a branch of Pnaric." *Journal of the Southeast Asian Linguistics Society* 4(2): 169-206.
- Diffloth, Gerard. (2005). "The Contribution of Linguistic Palaentology and Austroasiatic." In Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas, Eds. *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. London, Routledge: 77–81.
- Dryer, Matthew. (1992). "The Greenbergian Word Order Correlations." *Language* 68: 81-138.
- Greenberg, Joseph. (1963). "Some Universal of Grammar with Particular Reference to the Order of Meaningful Elements." In Joseph Greenberg, Ed. *Universals of Human Language*. Cambridge, MA, MIT Press: 73-113.
- Gurdon, Phillip R. T. (1914). *The Khasis*. London, Macmillan and Co.
- Henderson, Eugénie J. A. (1952). "The Main Features of Cambodian Pronunciation." *Bulletin of the School of Oriental and African Studies* 14: 159-174.
- Hogg, Richard and C. B. McCully. (1987). *Metrical Phonology: A course book*. Cambridge, Cambridge University Press.
- Jenny, Mathias, Tobias Weber and Rachel Weymuth. (2014). "The Austroasiatic Languages: A Typological Overview." In Mathias Jenny and Paul Sidwell, Eds. *The Handbook of Austroasiatic Languages* (2 vols). Leiden, Brill: 13–143.
- Matisoff, James. (1973). "Tonogenesis in Southeast Asia." In, Larry Hyman, Ed. *Consonant Types and Tone (Southern California Occasional Papers in Linguistics 1)*. Los Angeles, UCLA.
- Nagaraja, K. S. (1985). *Khasi, A Descriptive Analysis*. PhD thesis, Deccan College, Pune.
- _____. (1993). "Khasi dialects, a typological consideration." *The Mon-Khmer Studies Journal* 23: 1–10.
- _____. (1993). "Agreement in Khasi and Munda Languages." *Bulletin of the Deccan College Research Institute* 53: 271-276
- _____. (1996). "The status of Lyngngam." *Mon-Khmer Studies Journal* 26: 37-50.
- _____. (2014). "Standard Khasi." In Mathias Jenny and Paul Sidwell, Eds. *The Handbook of Austroasiatic Languages* (2 vols), Leiden, Brill: 1143–1185.
- Koshy, Anish. (2007). *Pronominal clitics in Khasi and Pnar*. MPhil Dissertation. Jawaharlal Nehru University, New Delhi.
- Koshy, Anish and Maranatha G.T. Wahlang. (2011). "Mnar Morpho-Syntax: A Preliminary Study." *Indian Linguistics* 72: 153–171
- Ring, Hiram. (2012). A phonetic description and phonemic analysis of Jowai-Pnar. *Mon-Khmer Studies*, 40: 133–175.
- _____. (2014). "Pnar." In Mathias Jenny and Paul Sidwell, Eds. *The Handbook of Austroasiatic Languages* (2 vols), Leiden, Brill: 1186–1226.
- _____. (2015). *A Grammar of Pnar*. PhD thesis. Singapore, Nanyang Technical University.

Case marking in Liangmai¹

WichamDinbo Mataina
Sikkim University

Abstract

The case markers in Liangmai, a Tibeto-Burman member spoken in Northeast India, demonstrate the typical distributional pattern of affixes/clitics occurring after the last constituent of a noun phrase (NP). This NP may be a noun, a pronoun, a personal noun or a demonstrative noun. The present study describes the main case markers of Liangmai. The language shows semantically/pragmatically motivated case marking. The paper discusses the individual case morphemes with their semantic and grammatical roles.

Citation	Mataina, Wichamdinbo. 2018. <i>Case marking in Liangmai</i> . North East Indian Linguistics (NEIL) 8, 17-31. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.
Volume Editors	Linda Konnerth, Stephen Morey, Amos Teo
Copyright	© 2018, the author(s), release under Creative Commons Attribution license
URL	http://hdl.handle.net/1885/155280

1. Introduction

Liangmai is a Tibeto-Burman language spoken in Northeast India. The term *Liangmai* refers to both the language and the speakers of the language. It is one of the Naga tribes living in contiguous areas of Manipur and Nagaland. There are 49,811 speakers, according to the 2011 census of India, who are concentrated in and around Manipur's Tamei Sub-Division of Tamenglong district. Others are found in Senapati and Imphal West districts in Manipur, and Peren district in Nagaland.

It is common for many tribes in Northeast India to bear the name given by outsiders. Liangmai too was known in the literature and Government documents by several other names until 2012: in Manipur it was known officially as "Kacha-Naga"². Other names that are known in the literature referring to Liangmai are "Lyangmai", "Liangmei" or "Quoireng". The Constitution (Scheduled Tribes) Order (Amendment) Act, 2011 of the Parliament of India formally corrected the misnomer "Kacha-Naga" and constitutionally recognized the tribe by its correct name. Linguistically, the Liangmai language is classified as part of Burling's (2007) "Zeme" group. In the social-political sphere, Liangmai is grouped with its cognate tribes Zeme and Rongmei within an umbrella body known as "Zeliangrong"³. However, in Nagaland state government documents, Liangmai is still called "Zeliang", together with the kindred Zeme tribe.

In this paper, I present a detailed description of the case markers and their functions in Liangmai.

¹ I am grateful to the anonymous reviewer and editor of this paper for their priceless comments and suggestions and for constantly prodding me into the completion of this work. Any shortcomings, if any, in this study will be mine alone.

² This term is considered to be derogatory by some people.

³ Zeliangrong, a civil group, is a conglomeration of Zeme, Liangmai and Rongmei, each with separate, but related, languages. They live in a contiguous area across the Indian states of Assam, Nagaland and Manipur. It is believed that they have strong blood relations which are inseparable; and that these three tribes are descendants of the three brothers of the same parents from Makuilongdi, to which these cognate tribes hold firmly even today. For more on Zeliangrong history, see Kamei (2004).

2. The Liangmai case markers

A case marker in Liangmai is a type of analytic morpheme that is typically attached to the rightmost element of the noun phrase it marks. This element may be a noun, a personal noun, or a demonstrative pronoun. The case marker can take the form of an affix or postposition: if the head noun is a pronoun, the case marker is morpho-syntactically bound to their head noun, e.g. *pə̀-gù* ‘3SG-GEN’. However, if the head noun is a proper noun or common noun, the case markers are not morpho-syntactically bound to them.

The Liangmai case markers are listed in Table 1 below. The tonal value of each case marker is influenced by the tone of the preceding noun. Throughout the paper, IPA tone diacritics are used for transcribing the four register tones.⁴ As Butt (2006: 7) demonstrates with examples from German and Korean, case markers can also be used to mark adjuncts in the world’s languages, a situation that we also find in Liangmai.

Table 1 – Overview of Liangmai case markers

Case name	Markers	Semantic function
Agentive	<i>niu</i>	agent; contrast
Accusative/Dative	<i>tu</i>	affected patient / recipient
Genitive	<i>gu</i>	possessor
Instrumental	<i>niu</i> <i>k^huiluziu</i>	instrument
Comitative	<i>nai</i>	accompany
Benefactive	<i>leng</i>	benefit
Locative	<i>ga</i>	location
Ablative	<i>gasu</i> <i>ganiu</i>	source, source of action
Allative	<i>lam</i>	movement towards a place
Terminative	<i>kə̀taŋ</i>	movement towards a time/place

Charengna (2014: 399) describes *niu* as a nominative marker in Liangmai. However, I analyse *niu* as an agentive marker rather than a nominative, with support from the examples I provide in the next section. Similarly, in analyses of Meitei, another Tibeto-Burman language, Bhat and Ningomba (1997: 52) describe *nə* as a nominative case marker, but Chelliah (1997: 107-113) analyses this as an agentive marker, since it is used when the actor is volitional.

2.1. Agentive⁵

The agentive case marker in Liangmai has the form *niu*. It is homophonous with an instrumental case marker, which will be discussed in section 2.4.⁶ The agentive marker is typically found on subjects of transitive clauses, as in (1) through (3); and on subjects of intransitive subjects, as in (4), that act wilfully and/or have control over an action.

⁴ The tonal analysis is based on the author’s work and is marked as follows: *é* = high, *ē* = mid, *è* = low, *ě* = extra low.

⁵ The term “ergative” is avoided here, following LaPolla (1995)’s reservation of the term “ergative” for instances where transitive subject marking follows a “systematic” or “paradigmatic” pattern.

⁶ Although one could analyse the instrumental use of *niu* as an extension of agentive *niu* to inanimate nouns, instrumental *niu* will be treated as a separate case marker because it can be used interchangeably with another instrumental case marker *k^huiluziu*, which cannot be used interchangeably with all instances of agentive *niu*.

- (1) *ī-niu kə-bàk tu gīsàt-nei*
 1SG-AGT NRL-pig ACC kill-FUT
 ‘I will slay the pig.’
- (2) *ī-nīu tsə-rǎ-sī tìu lù-dei*
 1SG-AGT NRL-fruit-CL eat PERF-DECL
 ‘I have eaten the fruit.’
- (3) *ī-nīu hǎi-bau nā tū kám məkàp-peì*
 1SG-AGT PROX-NMLZ kid ACC do cry-DECL
 ‘I (intentionally) made this child cried (because he was disobedient).’
- (4) *lunthonbau niu mək^hiu-jei*
 PN AGT cough-DECL
 ‘Lunthonbou (intentionally) coughed.’

In contrast, in the following examples without the ergative marker *niu*, none of the referents has control over the respective action. In (5), there is an experiencer subject, while in (6) and (7), the subject has the semantic role of theme.

- (5) *ǎ-wán t^hiu-jei*
 1SG-stomach pain-DECL
 ‘My stomach aches.’
- (6) *lunthonbau tsə-k^hau ga kəlùì wəŋ mí-dei*
 PN NRL-cleft LOC roll go PERF-DECL
 ‘Lunthonbou has fallen on the cleft.’
- (7) *tə-dùì zāu bām-meì*
 NRL-water drip PROG-DECL
 ‘Water is dripping (from the roof).’

However, this analysis is complicated by the fact that *niu* can sometimes be omitted from highly agentive A arguments and can even be used with non-agentive A/S arguments. For example, in (8) and (9), *niu* is not used even though the referent of the A argument has acted wilfully and/or is in control of the action. In these examples, it is difficult to explain what the difference in meaning is when *niu* is present vs. when it is absent.

- (8) *ī ŋtsòm biu bam k^ha-jei⁷*
 1SG drum beat PROG AP-DECL
 ‘I am still beating the drum.’
- (9) *lunthonbau ǎ-tū tsə-fāi pi-jei*
 PN 1SG-ACC NRL-shawl give-DECL
 ‘Lunthonbou gave me a shawl (yesterday).’

Furthermore, in (10) and (11), the agentive marker occurs with non-intentional subjects, though these statements sound less natural compared to ones where the agentive is not used.

⁷ An alternative analysis might gloss this particle as non-future.

Once again, it is difficult to explain how exactly the presence of the agentive changes the meaning of these sentences.

- (10) *pə̀-ñiù* *nə̀-tū* *pə̀-tsūn* *sà* *bām-mei*
 3SG-AGT 2-ACC 3SG-heart sad PROG-DECL
 ‘He is upset with you.’
- (11) *hāi-ga* *ī-ñiù* *kəsə̀ŋ-ziu* *sái-ra*
 PROX-LOC 1SG-AGT hot-CPM die-PURP
- kúm* *bām-bau-gā* *dekám tsə̀-mì* *t^hū-ziu* *lau?*
 like PROG-NMLZ-CONJ why NRL-fire make-CPM QPTCL
 ‘Here, I am about to die due to heat and why are (you) making fire?’

There are, nevertheless, contexts in which the agentive marker in Liangmai is obligatory. For instance, it is obligatory in *wh*- questions like “Who (did something)?”, and in the answer statements to these questions. An example of such a question is (12); and its corresponding answer statement, which is given in (13). However, not every sentence with *niù* needs to be interpreted as a response to a *wh*- question, and it is possible for (13) to also be uttered as one of the statements in a speaker’s narration of an incident.

- (12) *sāu* *niù* *nə̀-tū* *kə̀fi-ziu* *lau*
 who AGT 2SG-ACC slap-COMP Q
 ‘Who slaps/slapped you?’
- (13) *lun^honbau* *niù* *ə̀-tū* *kə̀fi-jei*
 PN AGT 1SG-ACC slap-DECL
 ‘Lunthonbou slapped me.’

There is another construction which necessitates the use of the agentive marker. They are the kind of sentences which state something about one product resulting from some cause, which is marked by *niù*. The genitive marker *gu* always follows a nominalized predicate in such sentences, as seen in (14) and (15).

- (14) *hai-bau* *ŋam-si* *lun^honbau* *niù* *ka-kám-báu* *gu* *jei*
 PROX-NMLZ work-DET Lungthonbou AGT NLR-do-NMLZ GEN DECL
 ‘This work is done by Lungthonbou.’
- (15) *hai-si* *t^hŋkāi* *niù* *hún-pák* *mí-bau* *gu* *jei*
 PROX-DET wind AGT blow-run PRF-NMLZ GEN DECL
 ‘This was blown away by the wind.’

But if the statement is about a product resulting from a source material, then the source is marked with the ablative *gasu* or *ganiù*, as in (16).

- (16) *tə-sāŋ* *hai-si*
 NRL-fermented.soya.bean PROX-DET
- ə-sĩo* *gasu* *kám* *kə-pət-bāu* *gū* *jei*
 NRL-soyabean ABL do NLR-go-NMLZ GEN DECL
 ‘Fermented soya beans are made from soya beans.’

Outside of these aforementioned constructions, the agentive is “optional”, though there are contexts where it is more natural to use it and contexts where it is more natural to omit it. For example, in both reported speech, as in (17), and quoted speech, as in (18), it is more natural to use the agentive marker, though it is still optional in these contexts.

- (17) *ápē* *niu* *dín-zīu* *nə-tū* *kīu-lau* *sɔ:*
 grandma AGT din-COMP 2SG-ACC come-IMP RP
 ‘Grandma told you to come down (to her place).’
- (18) *zisu* *niu* *din-nei*, *lūŋ-lí-lau*
 Jesus AGT said-DECL heart-change-IMP
 ‘Jesus said, “Repent.”’

The “optionality” of the agentive (and as we shall see, accusative case marking) makes it difficult to talk about “alignment” in Liangmai, and more work is needed to identify other syntactic/semantic/pragmatic factors that condition the use of the agentive *niu*. In any case, such semantically and pragmatically-motivated agentive marking in Liangmai is in consonance with systems found in other languages of the region (see Chelliah and Hyslop 2011). For example, Coupe (2007: 173) notes that in Mongsen Ao, a Tibeto-Burman language of Nagaland, the agentive *nə* is “most typically used to mark the A argument of a transitive clause in which the referent is an independently motivated, willfully acting animate participant.”

2.1.1. Contrastive *niu*

Additionally, the marker *niu* can be found on subjects of verbless clauses, as illustrated in sentences (19) and (20). Here, the marker *niu* is used to set up a contrast between two different subjects.

- (19) *hài-si* *nīu* *ránkáj mədái* , *wúi-si* *nīu* *ránkáj məŋíiu*
 PROX-DET CONTR money four DIST-DET CONTR money five
 ‘This is four rupees, that is five rupees.’
- (20) *luŋ^honbau* *nīu* *pə-lūŋ* *ke-wī-pūi-nā* ,
 PN CONTR 3SG-heart NZP-good-CL-child
- əsemliu* *nīu* *pə-lūŋ* *ke-sà-pūi-nā* *jei*
 PN CONTR 3SG-heart NZP-bad-CL-child DECL
 ‘Lunthonbou is the child of a compassionate mother, Asemliu is the child of a cruel mother.’

At the present time, it is unclear whether to treat this as either an example of a polysemous agentive *niu*, or a separate but homophonous marker. In other Tibeto-Burman languages of the region, a separate contrastive marker that is homophonous with the agentive is reported, as in Meitei (Chelliah 1997: 114-116), although the boundary between an agentive and a contrastive marker is not always clear, as in Sumi (Teo 2012).

2.2. Accusative/Dative

Liangmai has an accusative/dative marker *tù*, which marks either the object of a transitive clause, as in (21) and (22), or the recipient of a ditransitive clause, as in (23). Note that the theme argument in a ditransitive clause does not receive a case marker.

(21) *i mə-rūi tù gūi bam-mei*
 1SG NRL-chicken ACC feed PROG-DECL
 ‘I am feeding the chicken.’

(22) *nāŋ ə-kī ŋéonā tù ŋau mak ma hāi-lām*
 2SG 1SG-house cat ACC see NEG QPTCL PROX-side
 ‘Did you see my cat this side?’

(23) *pǎ ə-tū tū-pí pì-jei*
 3SG 1SG-ACC bean-CL give-DECL
 ‘He gave me beans (yesterday).’

Although AOV word order is most common in Liangmai, the position of the A and O arguments is interchangeable and does not influence the case marking pattern. This is shown in examples (24) and (25).

(24) *pǎ ə-tū fūi-jei*
 3SG 1SG-ACC look-DECL
 ‘He looked at me (yesterday).’

(25) *ə-tū pǎ fūi-jei*
 1SG-ACC 3SG look-DECL
 ‘He looked at me (yesterday).’

There are certain transitive verbs whose object arguments are unmarked. Such verbs include ‘like’, as seen in (26), but also ‘eat’, ‘drink’, ‘smell’, ‘taste’ etc, as in (27). Further, the accusative case marker is usually not used when the affected object is an inanimate object, as in (28).

(26) *ī pen məsēn-nei*
 1SG pen like-DECL
 ‘I like a pen.’

(27) *i tə-dūi tʰi-nei*
 1SG NRL-water drink-FUT
 ‘I will drink water.’

- (28) *awaŋbau te-tú-káŋ kədī bam-mei*
 PN NRL-stone-CL hit PROG-DECL
 ‘Awangbou is chiseling the stone.’

In certain contexts, accusative marking is used to express object argument focus, i.e. usually if a statement is a response to a question such as “What are you cutting?”, as in (29); or when someone does not know what she/he is supposed to do, as in (30). Outside of these contexts, it is still grammatical when these types of object arguments do not take the accusative marker.

- (29) *tə-sìŋ-bāŋ tu hūi-lau*
 NRL-wood-CL ACC cut-IMP
 ‘(I) chop the tree (with an axe).’ (response to: “What are you cutting?”)

- (30) *te-tú-káŋ tu kədī-lau*
 NRL-stone-CL ACC hit-IMP
 ‘You chisel the stone.’

Borrowed nouns typically do not take the accusative case, as in (31).

- (31) *kampiuter kək^htu-pāŋ-mi-lau*
 computer beat-break-PERF-IMP
 ‘Break the computer.’

The dative marker *tū* can be replaced by a locative marker *k^hu-ga* (literally, ‘side-LOC’) in some contexts. This type of construction occurs usually when one makes a request or supplication, especially before God for abstract nouns such as peace, love, wisdom, etc. This is illustrated in (32). Otherwise, using the locative marker *k^huga* would imply that whatever the recipient received would not be meant for himself/herself, as shown in example (33). More examples of *k^huga* are given in section 2.7.

- (32) *nāŋ pə-k^hu-ga nīŋtíŋ pì k^hāi-lau*
 2SG 3-side-LOC wisdom give POL-IMP
 ‘(God,) May you grant him your wisdom.’

- (33) *nāŋ pə-k^hu-ga rāŋkáj pì k^hāi-lau*
 2SG 3-side-LOC money give POL-IMP
 ‘You leave the money with him.’

2.3. Genitive

There are two genitive constructions in Liangmai: one that uses an overt genitive case marker *gu*, and one that does not. We will first examine the construction that does not use the overt case marker.

2.3.1. Genitive Construction Type 1

In the first kind of genitive construction, which I call “Type 1”, the possessor is not marked with the genitive case marker, while the possessed noun appears only as a bound root, without the non-relational prefix that it would have in citation form, as illustrated in (34) through (40).

The construction is always used when the possessed noun is a kinship term, body part, or represents a part-whole relationship.

(34) *pənäi-kí*
 3DU-house
 ‘their (two) house’ (cf. *tsə-kí* ‘NRL-house’ in citation form)

(35) *nə̀-ṗüi*
 2SG-mother
 ‘your mother’

(36) *nə̀-ṗüi kî*
 2SG-mother house
 ‘your mother’s house’

(37) *ə̀-tíŋ*
 1SG-back
 ‘my back’

(38) *tse-ṗüaŋ hīn*
 NRL-elephant tusk
 ‘the tusk of an elephant’

(39) *tsə-gan tsəlî gɔm*
 NRL-curry pot lid
 ‘the lid of a curry pot’

(40) *pə̀-lü fɔn*
 3SG-field corner
 ‘The border area of his (paddy) field.’

2.3.2. Genitive Construction Type 2

In the second type of genitive construction, which I will call “Type 2”, the possessor takes the genitive marker *gu*, while the full citation form of the possessed noun is retained. This is illustrated in (41) through (43).

(41) *ə-gu tsə-fāi*
 1SG-GEN NRL-shawl
 ‘My shawl.’

(42) *nə-gu tsə-láŋ*
 2SG-GEN NRL-cotton
 ‘Your cotton’

(43) *mizoram gū tìŋkái*
 Mizoram GEN wind
 ‘Mizoram’s weather.’ (lit. ‘Mizoram’s wind’)

When the possessed noun is a borrowed word, *gū* must be used regardless of what or who the possessor is.

- (44) *ə-gu* *tiket*
 1SG-GEN ticket
 ‘My ticket’
- (45) **ə-tiket*
- (46) *nə-gu* *kəm'piutar*
 2SG-GEN computer
 ‘Your computer’

2.4. Instrumental

There are two instrumental case markers: *niu* and *k^huiluziu*, which mark the instrument that is used by an agent noun, as in (47) through (49). The instrumental marker *niu* is homophonous with the agentive marker. Both instrumental case markers can be used interchangeably, but the marker *k^huiluzū* is more common. The marker *k^huiluzū* is a compound form derived from the verb *k^hui* ‘take’, the aspectual marker *lu* and the complementizer *ziu*.

- (47) *təmei gāsū tsə-fī* *nīu* *tat* *wəŋ-ŋei*
 Tamei ABL NRL-leg INST go come-DECL
 ‘(Yesterday, we) came from Tamei by foot.’
- (48) *tənū k^huiluziu* *tsə-fāi-tə-nī* *kət^hà-jei*
 herb INST NRL-shawl-NRL-pants wash-DECL
 ‘(We) wash clothes with a herbal plant.’
- (49) *nə-ben* *k^huiluzū* *tū-lau*
 2SG-hand INST eat-IMP
 ‘Eat with your hand.’

2.5. Comitative

The comitative case is used to denote “accompaniment” (Stolz et al. 2006: 17-23). In Liangmai, the comitative has the form *nai*, as shown in (50) and (51).

- (50) *pä* *tsə-rā* *pə-nái* *ŋpāu* *bām-mei*
 3SG NRL-spirit 3SG-COM talk PROG-DECL
 ‘He is talking with a spirit.’
- (51) *əliu-nái* *pəliu wāŋ* *ŋtsám-mei*
 1PL-COM 3PL come same-DECL
 ‘We and they came together.’ (literally: ‘With us they came together.’)

However, it is optional when the first element is a pronoun, as in (52), or a singular noun, as in (53).

(52) *pǎ əliu-(nái) wāŋ ɲtsám-mei*
 3SG 1PL-(COM) come same-DECL
 ‘He and us come together.’

(53) *minister-píu əliu tàt ɲtsám-néi*
 Minister-CL 1PL go same-FUT
 ‘Minister and we will go together.’

2.6. Benefactive

As with the other case markers, the benefactive marker *lēŋ* follows the noun phrase it marks. It marks the semantic role of beneficiary, as in (54) and (55).

(54) *kə-bàk ləŋ kə-bì-nūi fúŋ wāŋ-lau*
 NRL-pig BEN NRL-yam-leaf carry come-IMP
 ‘Bring yam leaves for the pig.’

(55) *zisu nəl̩i lēŋ sái-khái-jei*
 Jesus 2PL BEN die-POL-DECL
 ‘Jesus died for you (all).’

2.7. Locative

Liangmai has a rich inventory of constructions that express static locational semantics. A list of these is given in Table 2, along with their main functions. In these constructions, a general locative marker *ga* is used either by itself or with another morpheme. The latter will be called “locative markers” in this section, since it is not always possible to decide if they should be treated as stacked locative case markers, or as relator noun constructions, though the origins of some of these are clearly nominal, e.g. *-k^hu* ‘side’.

Table 2 – Locative expressions in Liangmai

Form	Translation	Usage
<i>ga</i>	at/in	general spatial location
<i>k^hu/k^hu-ga</i>	beside	position close to animate nouns and pronouns
<i>luŋ-ga</i>	inside	location inside something
<i>soŋ-ga</i>	in front of	location outside something
<i>ri-ga</i>	above	location above something
<i>haŋ-ga</i>	under	location under something
<i>tai-ga</i>	nearby	location near by
<i>muŋ-ga</i>	nearby	location near by

The general locative marker *ga* is used to indicate an unspecific location in space, as in (56) through (60). As with other case markers, it is a suffix when attached to a pronoun.

(56) *hai-ga*
 PROX-LOC
 ‘Here.’

- (57) *kədĩ gā*
 earth LOC
 ‘On the ground.’
- (58) *ī təmei gā lúŋ-né*
 1SG Tamei LOC live-FUT
 ‘I will live in Tamei.’
- (59) *i tə-duit ga ɲùt piŋ-ŋei*
 1SG NRL-water LOC enter afraid-DECL
 ‘I am afraid to enter the water.’
- (60) *wàntsák kəriu gā wāŋ-lāu*
 time ten LOC come-IMP
 ‘Come at 10 o’clock.’

The locative marker *k^huga* (literally, ‘side-LOC’) is used if the marked noun, typically a personal pronoun or an animate noun, is physically close to another animate noun, as in (61), or in possession of something, as in (62) and (63).

- (61) *ŋgiāk-pūi hāi tsè-pùan k^hú-gā bām-mei*
 crow-CL DET NRL-elephant side-LOC exist-DECL
 ‘The crow is with an elephant.’ (i.e. ‘The crow is staying with an elephant.’)
- (62) *pə-k^hu-ga wahina bam-mei*
 3SG-side-LOC everything exist-DECL
 ‘He has everything.’
- (63) *nə-pau k^hú-gā tsə-k^hā-tèn bām-mei*
 2SG-grandfather side-LOC NRL-fish-rotten exist-DECL
 ‘Your grandfather has fermented fish.’ (i.e. ‘Fermented fish is available with your grandfather.’)

The locative marker *lūŋga* is used to indicate a location inside of the marked noun.

- (64) *pāirīan lūŋ-ga tətú s̄an mi-dei*
 pipe inside-LOC stone fix PERF-DECL
 ‘A pebble is stuck inside the pipe.’

The locative marker *sōŋga* is used to indicate a location in front of the marked noun.

- (65) *ə-sōŋ-gā tsàp tù-lau*
 1SG-front-LOC stand NEG-IMP
 ‘Do not stand in front of me.’
- (66) *tsə-kì sōŋ-ga*
 NRL-house front-LOC
 ‘in front of the house’

The locative marker *riga* marks referents that are situated above something in an elevated location, as in (67) through (69). Note that if the main locative marker is used without *ri*, as in (70), this would mean that the pigeon is inside the house, not on top of it.

- (67) *mālúan ri-gā*
 mountain above-LOC
 ‘on the mountain top’
- (68) *tə-zī rì-gā tsə-fāi bam-mei*
 NRL-bed above-LOC NRL-shawl exist-DECL
 ‘The shawl is on the bed.’
- (69) *tsə-kì ri-gā mǎk^háu-pui ñtáu bam-mei*
 NRL-house above-LOC pigeon-CL sit present-DECL
 ‘The pigeon is sitting on top of the house.’
- (70) *tsə-kì gā mǎk^háu-pui ñtáu bam-mei*
 NRL-house LOC pigeon-CL sit present-DECL
 ‘The pigeon is sitting in the house.’

The locative marker *hāngā* indicates a location beneath something.

- (71) *tə-sìŋ-bāŋ hāŋ-gā tsap-lau*
 NRL-wood-CL below-LOC stand-IMP
 ‘Stand under the (shade of the) tree.’
- (72) *məguilūan nam si kùilūan nám hāŋ-gā jei*
 Maguilong village EMP Kuilong village below-LOC DECL
 ‘Maguilong village is situated below the village of Kuilong.’

Two locative markers can be used to indicate a location that is ‘nearby’: *tāiga* and *mūngá*, as in (73) and (74) respectively. The latter literally means ‘edge-LOC’, and is used when referring to some location that has a clear edge/boundary, such as a pond, table, terrace, etc. For that reason, (75) sounds unusual.

- (73) *tsə-kì tāi-gā bóm ñpök tu-lau*
 NRL-house nearby-LOC cracker break NEG-IMP
 ‘Do not burst crackers near the house.’
- (74) *əzái múŋ-ga*
 sea edge-LOC
 ‘by the sea’
- (75) *? ə-muŋ-ga tsàp tù-lau*
 1SG-edge-LOC stand NEG-IMP
 ‘Do not stand near me.’

2.8. Ablative

The ablative marker *gasu* indicates the source of a movement, as in (76) and (77). In these examples, *gasu* can be used interchangeably with *ganiu*.⁸

- (76) *tə-zà-pūi pə-kì gāsū búan̄ ɲpat mi-dei*
 NRL-rat-CL 3SG-house ABL crawl go PERF-DECL
 ‘The rat came out of its hole.’ (lit. ‘The rat came out of its house.’)

- (77) *təmei gāsū tsə-ft̄ n̄iū tat waŋ-ŋei*
 Tamei ABL NRL-leg INST go come-DECL
 ‘(Yesterday, we) came from Tamei by foot.’

As noted in section 2.1, the ablative is also used in statements about a product resulting from a source material, as in (78).

- (78) *tə-sāŋ hai-si*
 NRL-fermented.soya.bean PROX-DET
- ə-siə gasu kám kə-pət-bāu gū jei*
 NRL-soyabean ABL do NLR-go-NMLZ GEN DECL
 ‘Fermented soya beans are made from soya beans.’

2.9. Allative

The allative marker *lam*, literally ‘side’, expresses movement towards a location. Unlike locative *ga*, which is used to indicate a static location, the marker *lam* is used to express movement, i.e. ‘going to’.

- (79) *hāi-làm*
 PROX-ALL
 ‘(Come) this side.’
- (80) *tsə-kì lām wāŋ-lāu*
 NRL-house ALL come-IMP
 ‘Come home.’
- (81) *ī təmei lām tət-né*
 1SG Tamei ALL go-FUT
 ‘I will go to Tamei.’

2.10. Terminative

The terminative marker *kətán̄* is used to mark a temporal or spatial endpoint, as illustrated by (82) and (83).

⁸ These two markers are also homophonous with Liangmai conjunctive markers *gāsū* and *ganiu*, both meaning ‘because’.

- (82) *təmei kətáj i nə-tū kəpāu-k^hai pì-néi*
 Tamei TERM 1SG 2SG-ACC drop.off-POL give-FUT
 ‘I will take you till the village of Tamei.’
- (83) *pəliu púŋ kəriu kətáj nə-tū k^hōn-nei*
 3PL clock ten TERM 2SG-ACC wait-DECL
 ‘They waited for you until 10 o’clock (yesterday).’

3. Conclusion

In this paper, I have described the forms and functions of case markers in Liangmai. The use of the agentive marker appears to be semantically/pragmatically motivated. The agentive marker commonly occurs when the agent argument is in control of the action, and it is obligatory when the subject is in focus. Further, accusative marking is also optional when the experiencer or patient is inanimate, but obligatory in statements where the affected patient is in focus. However, there are numerous contexts where agentive or accusative marker can be used with no apparent change in meaning, though what semantic/pragmatic factors condition their distribution in these “optional” contexts is left to be ascertained in further studies. In addition, there are two genitive constructions: one in which the genitive marker *gu* is obligatory, and one where a pronominal prefix is added to the root of a possessed noun. Finally, I described various case markers that provide both static and dynamic spatial information.

Abbreviations

ABL	Ablative	LOC	Locative
ACC	Accusative/Dative	NLR	Non-relational prefix
AGT	Agentive	NMLZ	Nominalising suffix
AP	Adverbial particle	NZP	Nominalising prefix
BEN	Benefactive	PL	Plural
CONTR	Contrastive	PN	Proper name
CL	Classifier	POL	Politeness marker
DECL	Declarative	PRF	Perfective
DET	Determiner	PROG	Progressive
DIST	Distal	PROX	Proximate
DU	Dual	PURP	Purposive
FUT	Future	QPTCL	Question particle
GEN	Genitive	RP	Reportative particle
IMP	Imperative	SG	Singular
INST	Instrumental		

References

- Bhat, D. N. Shankara and M. S. Ningompa. (1997). *Manipuri grammar*. München, Lincom Europa.
- Burling, Robbins. (2007). “The Tibeto-Burman languages of Northeastern India.” In Graham Thurgood and Randy J. LaPolla, Eds. *The Sino-Tibetan languages*. New York and London, Routledge: 160–91.
- Butt, Miriam. (2006). *Theories of Case*. Cambridge, Cambridge University Press.
- Charengna, Widinibou. (2014). “Case markers in Liangmai.” *Language in India* 14(7): 396–407. (<http://www.languageinindia.com/july2014/liangmaicasemarkers.html>)

- Chelliah, Shobhana, L. (1997). *A grammar of Meitei*. Berlin, Mouton de Gruyter.
- Chelliah, Shobhana L. and Gwendolyn Hyslop. (2011). "Introduction to special issue on optional case marking in Tibeto-Burman." *Linguistics of the Tibeto-Burman Area* 34(2): 1–7.
- Coupe, Alexander R. (2007). *A grammar of Mongsen Ao*. Berlin, Mouton de Gruyter. .
- Kamei, Gangmumei. (2004). *A history of Zeliangrong Nagas from Makhel to Rani Gaidinliu*. Guwahati, Spectrum Publications.
- LaPolla, J. Randy. (1995). "'Ergative' marking in Tibeto-Burman." In Yoshio Nishi, James Matisoff and Yasuhiko Nagano, Eds. *New horizons in Tibeto-Burman morphosyntax*, Osaka, National Museum of Ethnology: 189–228.
- Pamei, Namthiubuiyang. (2001). "The trial from Makuilongdi: The continuing saga of the Zeliangrong people." Tamenglong, Girona Charitable Foundation.
- Stolz, Thomas, Cornelia Stroh and Aina Urdze. (2006). *On comitatives and related categories: A typological study with special focus on the languages of Europe*. (Empirical Approaches to Language Typology 33). Berlin and New York, Mouton de Gruyter.
- Teo, Amos. (2012). "Sumi agentive and topic markers: *no* and *ye*". *Linguistics of the Tibeto-Burman Area* 35(1): 49–74.

Languages in Arunachal Pradesh

High vowel devoicing in Hrusso Aka¹

Vijay A. D'Souza
University of Oxford

Abstract

High vowel devoicing is an important phonetic and phonological process in Hrusso Aka, a minority language of Arunachal Pradesh. The pervasive nature of high vowel devoicing is perhaps one of the reasons why the language appears to have a complex and divergent phonology compared to its neighbours. This paper gives a preliminary description of the process, presenting evidence for it, and discussing its significance for Hrusso Aka phonology. High vowel devoicing occurs when any of the high vowels /i u w/ is preceded by a voiceless consonant in a non-stressed syllable. The process involves the loss of vocal fold vibration during vowel production, giving an impression that the vowels are being whispered. Devoiced vowels are also often severely truncated and overlapped with the preceding consonants, and usually appear to be deleted, especially in rapid speech. However, we have clear evidence of their existence in slow speech and syllable-inverting language games. Understanding the process of high vowel devoicing is crucial to the study of Hrusso Aka phonology, especially in the context of its highly complex surface consonantal sequences. Further, a careful distinction must be made between high vowel devoicing and two other processes, namely, high vowel deletion that occurs during re-syllabification, and sibilant-obstruent cluster formation. A good grasp of high vowel devoicing is also necessary for clear phonemic and orthographic transcriptions.

Citation	D'Souza, Vijay A. 2018. <i>High vowel devoicing in Hrusso Aka</i> . North East Indian Linguistics (NEIL) 8, 33-46. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.
Volume Editors	Linda Konnerth, Stephen Morey, Amos Teo
Copyright	© 2018, the author(s), release under Creative Commons Attribution license
URL	http://hdl.handle.net/1885/155280

1. Introduction: People and the language

Hrusso Aka is a little studied minority language spoken by the members of the indigenous community who call themselves *Ġusso* [yus:o].² They live in 21 villages of the Jamiri, Thrizino, and Bhalukpong circles³ of the West Kameng District of Arunachal Pradesh. According to the 2011 Census of India, their total population is around 6,700.⁴ The number of people who actually speak Hrusso Aka is not known, but can be estimated to be around 4,000,

¹ This paper is a revised version of a section in my MPhil Thesis (D'Souza, 2015). I am grateful to my academic supervisors Profs. Mary Dalrymple and Aditi Lahiri, and Prof. John Coleman for their support and guidance during the process of writing this paper. Any mistakes and inconsistencies are mine alone. Fieldwork for data collection was made possible by the Individual Graduate Scholarship (IGS) from ELDP, SOAS, London. Thanks to Mr. Dorjee K. Jebisow, Mr. Govardhan Nimasow, Mrs. Anu Jebisow, Mr. Delu Sapon, Mr. Düva Dususow, Mr. Babu Regisow and many other native Hrusso Aka speakers for their patience and excellent support during the collection of data used in this paper. Thanks also to the anonymous reviewer and editor for their insightful comments and suggestions.

² This is the autonym. For official purposes, the community calls themselves "Aka (Hrusso)". Alternative exonymous spellings are "Hruso Aka" and "Hruso".

³ Circles are sub-district level administrative units in Arunachal Pradesh.

⁴ According to the 2011 census, the population of "Aka" is 8,011. However, this figure includes the Koro Aka of the neighbouring East Kameng district, who speak a different language (see Anderson and Murmu 2010). See the "ST Appendix" on the Census India website <http://www.censusindia.gov.in/2011census/PCA/ST.html> (accessed on 15/09/2016) for the population figures in the West Kameng District, which has two entries relevant for our purposes: One, the number of "Aka" (6718). The other, the number of "Hrusso" (56). This is a mistaken distinction. Both should be read as referring to Hrusso Akas.

given the rapid language shift towards Hindi in several villages, and especially in towns like Bhalukpong, Tippi and Thrizino.

Hrusso Aka is a fascinating language because of its unique lexicon and peculiar phonology. This “fabulously complex language full of wicked tongue twisters”, as Harrison (2010: 121) calls it, has eluded classification so far. Although there is some evidence to support the claim that it belongs to the “Hrusish” subgroup of the Tibeto-Burman family along with Sajolong⁵ (Miji/Dhammai) and Bangru (Levai), the evidence is not conclusive (Thurgood and LaPolla 2007, Bodt and Lieberherr 2015). Further, the status of the “Hrusish” group in the Tibeto-Burman family itself is unclear. Bradley (1997) places it under the “West Arunachal” language group. There have been also suggestions that Hrusso Aka might be a language isolate, or a direct descendent of the Proto-Sino-Tibetan with no relatives in the Tibeto-Burman family (Blench and Post 2011, 2014).

There is some confusion about the dialects of the language. For example, Shafer’s (1947) misleading identification of “Hruso A” and “Hruso B” dialects, based on his study of limited data sets provided by Anderson (1896), Campbell (1874), Hesselmeyer (1868) and Payne (1909). These “dialects of Hruso” were in fact two different languages: Sajolong (Miji/Dhammai) and Hrusso Aka, respectively. Koro Aka, a neighbouring language, has been called a dialect of Hrusso Aka (Sinha 1962). However, Koro is a different language having no genetic affiliation with the latter (Anderson and Murmu 2010). In fact, there are only two “varieties”⁶ of the language that the native speakers generally recognize and there is linguistic evidence for one variety spoken by people whose villages constitute the Thrizino “circle”; and the other, spoken by the people living in Bhalukpong and Jamiri circles. These varieties and the people who speak them are informally referred to as *Hushüñ* [xʊʂúŋ]⁷ and *Ĝusso* [ɣússɔ]⁸ respectively, but the terminology is contentious. Therefore, although the difference between the two varieties is widely recognized, they have not been officially named by the Hrusso Aka community. The two varieties are very close to each other in lexicon, phonology, grammar and prosody, and have almost complete mutual intelligibility. One noticeable difference is the fronting of the fricatives /ʃ ʒ x ɣ/ in the so-called *Hushüñ* variety. The two varieties have further sub-varieties with minor differences, mainly due to village-level lexical and prosodic peculiarities. High vowel devoicing, the process under discussion here, is found in all varieties of Hrusso Aka.

Interestingly, in the scanty and inconsistent literature on the language, a regularly noted fact is the complexity of its sound inventory, which makes the language strikingly different from its neighbours (see Grierson 1909, Harrison 2010: 121-22). There seem to be two reasons for the difference: first, its fricative-rich sound inventory, a fact already noted since Grierson (1909) but not yet described well, and second, devoicing of high vowels, a ubiquitous but hitherto undescribed process, that gives an impression that the language has highly complex consonant sequences. This paper is a preliminary discussion on the latter. I shall present data on high vowel devoicing in Hrusso Aka; discuss its importance for the study of the Hrusso Aka phonology; and distinguish between three processes: high vowel deletion; cluster formation; and high vowel devoicing, with the help of articulatory, acoustic and phonological evidence.

⁵ The present autonym of the erstwhile Miji/Dhammai community.

⁶ The term “dialect” is somewhat problematic in Northeast India. For the time being, I shall use the term “varieties” instead, to denote the two versions of Hrusso Aka.

⁷ *Hushüñ* connotes ‘people across the river’, referring to the people and the variety spoken on the northern side of the Bichom river in the West Kameng district.

⁸ *Ĝusso* [ɣusso], is the word from which the name “Hrusso” is derived. Apart from being the autonym for the entire tribe, informally it refers to the clans that live on the southern side of the Bichom river.

2. Data and methodology

The data for this study were primarily obtained from two word lists collected between 2014-16, uttered by five native speakers, three male and two female. A word list of 800 words was collected from two male speakers. One was aged 51 years and a speaker of the *Ĝusso* variety from Buragaon village, and the other aged 32 years, a native speaker of the *Hushüñ* variety, and a resident of Bhalukpong town. A subset list of 200 words was collected from one male and two female speakers, aged 30-35 years. Word lists were recorded as three isolated utterances of each word followed by two utterances of a sentence containing the word. Besides the word lists, natural texts and sentence lists of around 50 hours, recorded between 2012-16, were used as supplementary data. All recordings of the word lists were done by the author, in collaboration with consultants Dorjee Jebisow and Delu Sapon using a Marantz PMD 661 MKII recorder and two Audio Technica AT 803B omni-directional lapel microphones in relatively quiet village surroundings. Recording settings were: 32 bit stereo with 48 kHz sampling rate.

Relevant tokens were then selected from the above data and were analysed acoustically using Praat (Boersma and Weenink 2017). Spectral slices were obtained by FFT analysis using 0-8 kHz range and a window length of 5 ms. All the tokens shown here are from the speech of DS, aged 51, speaking the *Ĝusso* variety of the language.

Transcriptions without brackets are orthographic. Slanted brackets and square brackets indicate phonemic and phonetic transcription (broad) respectively. For simplicity, vowel devoicing (a small circle beneath a syllable nucleus) is indicated directly on phonemic transcription. It is important to remember however, that vowel devoicing is not phonemic in Hrusso Aka. The acute accent marks word stress.

3. Vowels and consonants of Hrusso Aka

3.1. Vowels

Hrusso Aka has six vowels, represented in Figure 1, distinguishable from each other on the basis of place of articulation, tongue height and roundedness. The three high vowels are: the high front unrounded vowel /i/; the high back unrounded vowel /ɯ/; and the high back rounded vowel /u/. These are the vowels most relevant to this paper, since vowel devoicing primarily involves them. The mid vowels are the mid front unrounded vowel /e/ and the mid back rounded vowel /o/. The phonemic status of the mid back unrounded vowel /ɤ/ is not clear as yet, and is being studied. The low vowel is the low central unrounded vowel /a/.

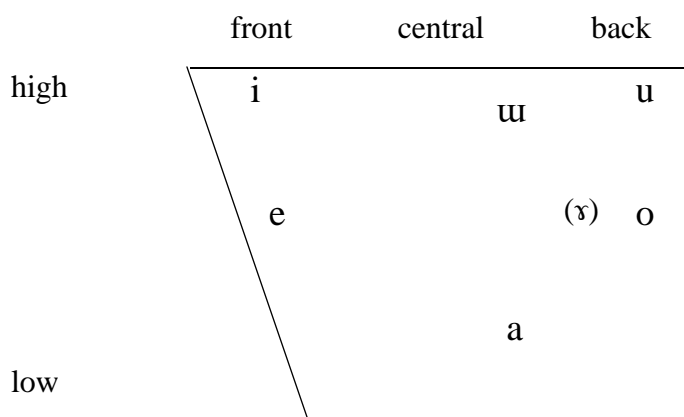


Figure 1. Hrusso Aka vowel chart

3.1.1. Consonants:

Hrusso Aka has 26 consonants, represented in Table 1. Among the consonants only the voiceless obstruents, that is, the voiceless stops /p t k/, the voiceless fricatives /f s ʃ ɣ x/ and the affricate /tʃ/, are involved in the vowel-devoicing process.

Table 1: Hrusso Aka consonant chart

	<i>bilabial</i>	<i>labio-dental</i>	<i>dental</i>	<i>alveopalatal</i>	<i>retroflex</i>	<i>velar</i>			
stop	p	b	t	d		k g			
affricate				tʃ	ɖʒ				
fricative		f	v	s	ʃ	ɣ	ɣ̣	x	ɣ̥
nasal		m		n	ɲ				ŋ
lateral				l	ɭ				
trill				r					
glide		w			j				

3.2. Vowel devoicing

Vowel devoicing is the production of vowels without default sonorant voicing (Varden 2010). It occurs when the vocal fold vibration typically associated with vowel production is absent (Laver 1994: 189). The resulting vowels are called “voiceless vowels” or “whispered vowels”.⁹ Vowel devoicing, a “non-modal” type of vowel production (Gordon 1998), is not a cross-linguistically common phenomenon, presumably because languages generally disfavour reduced perceptual salience, which vowel devoicing entails.¹⁰

Languages that have voiceless vowels are of two types: first, those with a phonemic contrast between voiced and voiceless vowels, and second, those in which voiceless vowels are allophones of their voiced counterparts. Languages of the first type are quite rare (Gordon 1998, Ladefoged and Maddieson 1996: 315). Among Northeast Indian languages, Nyishi has been described as having a phonemic contrast in voicing for the high vowels /i/ and /u/ (Ray 1967), although this has not been corroborated by subsequent studies.

Hrusso Aka belongs to the second, more common, type, of which Japanese is perhaps the most extensively studied example (Shibatani 1990, Kondo 1997, Varden 1998, Tsuchida 2001, *inter alia*). Some other languages known for allophonic vowel devoicing include Korean (Jun and Beckman 1993), French (Torreira and Ernestus 2010), Greek (Dauer 1980),

⁹ Among the vowels that lack vocal fold vibration, Laver (1994: 189) makes a distinction between “voiceless” and “whispered” vowels depending on the nature of phonation. In the present analysis, no attempt is made to categorise the Hrusso Aka devoiced vowels one way or the other. This is a topic for further study.

¹⁰ See Jaeger (1978). Of the 221 languages in the Stanford Phonology Archive, only 44 had voiceless vowels, and of these, 20 only preferentially devoiced high vowels, and the rest devoiced other vowels in addition to the high vowels. No language was found to devoice non-high vowels without also devoicing high vowels.

Turkish (Jannedy 1995), Oneida and Blackfoot (Gick et al. 2012). In vowel-devoicing languages, high vowels are more prone to de-voicing than others, which has been attributed to aerodynamic reasons (Jaeger 1978).

4. High vowel devoicing in Hrusso Aka

High vowel devoicing is a common phenomenon in Hrusso Aka. It is found across all age groups, social strata, across the genders and the two varieties of Hrusso Aka. While in most languages cited above, high vowel devoicing is a marginal process with minor significance, in Hrusso Aka it is a pervasive and important phenomenon, perhaps more so than in Japanese, which is often cited as the prototypical example of high vowel devoicing.

4.1. Evidence

Devoiced vowels in Hrusso Aka can be recognized by their whispery nature which, as already noted, is a frequent and familiar phenomenon in the language. On spectrograms, they either totally lack the typical formant bands that are characteristic of modal vowels or have only faint formant bands. They are also often truncated and overlapped with the preceding consonant. These facts are illustrated in Figure 2:

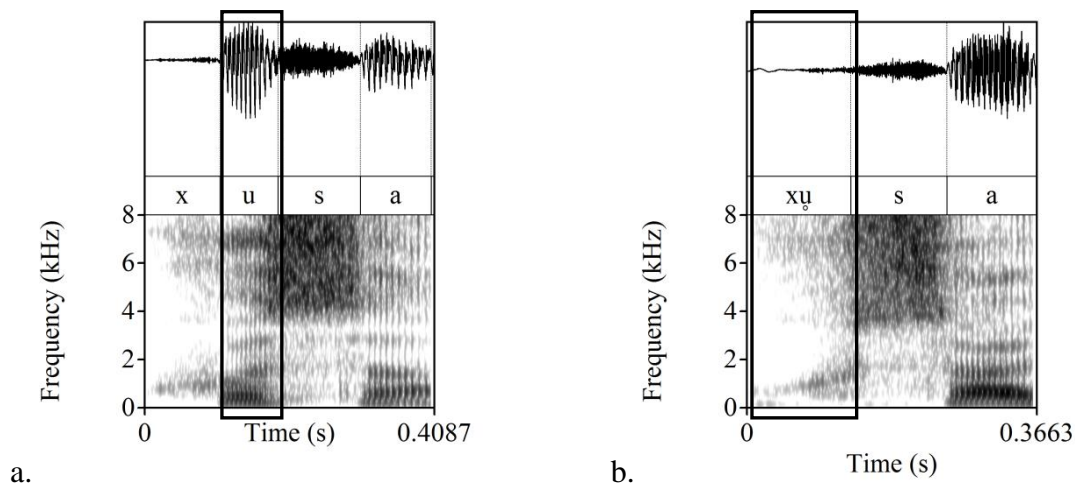


Figure 2: Waveform and spectrograms of voiced and devoiced utterances of /u/

The spectrograms in Figure 2 show an instance each of voiced and devoiced high vowel /u/. In both instances, the vowel is a part of the morpheme /xusa/ ‘streamlet’. Figure 2a. shows the morpheme uttered as an independent word. The vowel is voiced and the formants are clearly noticeable as shown in the boxed portion. In Figure 2b the same morpheme is uttered as a part of the larger word /xusadzjo/ ‘waterhen’ (lit. ‘hen of the streamlet’). Here the vowel is devoiced and consequently the vowel formants are only faintly noticed as shown in the boxed area. Also, it is impossible to segment the two phonemes separately in the syllable /xu/.

The two instances of the vowel /u/ have the same phonemic environment between /x/ and /s/, and are part of the same morpheme, but are prosodically different. The former is the nucleus of a stressed syllable in /xú.sa/ (Figure 2a), and the latter is unstressed in /xu.sá.dzjo/ (Figure 2b). This is seen in Figure 3, showing the pitch analysis of the two instances of the morpheme /xusa/. The high pitch shown on /u/ in Figure 3a. correlates to the stressed unit, in keeping with the stress rules in the language. The pitch line is not visible in 3b since the vowel is devoiced:

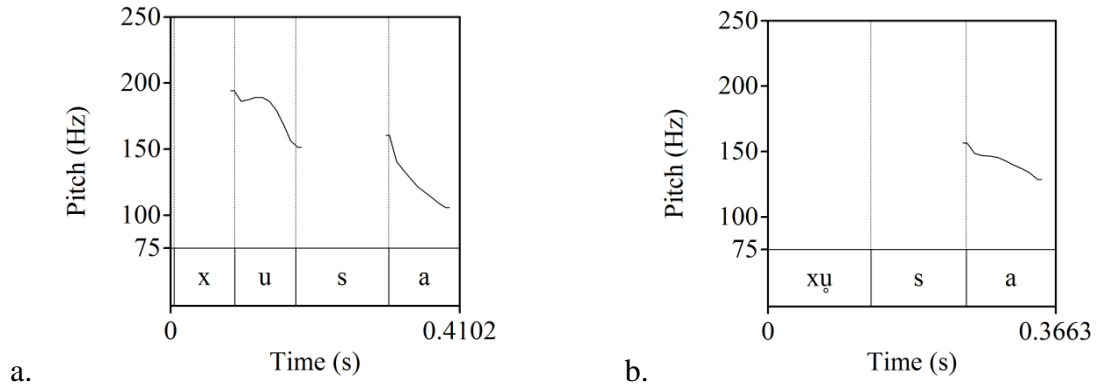


Figure 3: Pitch traces of /u/ uttered in stressed and unstressed syllables.

4.2. The rule

The devoicing of the vowel /u/ in the example above has occurred in a non-stressed C_C environment. On examining examples in the data set (1) below, more facts about high vowel devoicing emerge.

- (1)
- | | | |
|----|------------------------|---------------------------|
| a. | /xúmtʂu/ ¹¹ | ‘the <i>Bichom</i> river’ |
| b. | /ákʲi/ | ‘father’s elder brother’ |
| c. | /xuʂtʂú/ | ‘tiger’ |
| d. | /sútʂaw/ | ‘hang’ |
| e. | /xuʂów/ | ‘big river’ |

Both voiced and devoiced high vowels are exemplified in the data set. As was observed in the case of /xúsa/ above, no high vowel that is the nucleus of a stress-bearing syllable is devoiced. This also applies to (1)d. On the other hand, in an unstressed syllable, devoicing takes place in the following environments:

- (2)
- a. $\left[\begin{smallmatrix} +CONS \\ -VOICE \end{smallmatrix} \right] \text{ — } \left[\begin{smallmatrix} +CONS \\ -VOICE \end{smallmatrix} \right]$ as in (1) c.
- b. $\left[\begin{smallmatrix} +CONS \\ -VOICE \end{smallmatrix} \right] \text{ — } \left[\begin{smallmatrix} +CONS \\ +VOICE \end{smallmatrix} \right]$ as in (1) e.
- c. $\left[\begin{smallmatrix} +CONS \\ -VOICE \end{smallmatrix} \right] \text{ — } \#$ as in (1) a & b.

It is clear from the above examples that there are two crucial conditions for high vowel devoicing: first, the high vowel must follow a voiceless obstruent and second, it must not be stressed. The environment following the vowel does not seem to matter.

Based on these observations the high vowel devoicing rule in Hrusso Aka can be formulated as:

¹¹ For the sake of simplicity, I have marked devoicing directly on the phonemic transcription in most of the examples.

- (3) Devoice a high vowel following a voiceless consonant, provided it is not the nucleus of a stress-bearing syllable.

$$\begin{bmatrix} +\text{VOC} \\ +\text{HIGH} \\ +\text{VOICE} \\ -\text{STRESS} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{VOC} \\ +\text{HIGH} \\ -\text{VOICE} \\ -\text{STRESS} \end{bmatrix} / \begin{bmatrix} +\text{CONS} \\ -\text{VOICE} \end{bmatrix} \text{---}$$

4.3. High vowel devoicing and other similar processes

Understanding high vowel devoicing is important not only because of its ubiquity in Hrusso Aka, but also because there are two other processes that need to be carefully distinguished from it. These processes are: high vowel deletion that occurs during re-syllabification; and sibilant-obstruent cluster formation. Confusing high vowel devoicing with high vowel deletion yields implausible consonant sequences. The problem of potential confusion between sibilant-obstruent cluster formation and high vowel devoicing can be illustrated with the examples in (4) below. A possible phonetic transcription of the surface output of these words uttered in normal speech is given in square brackets:

- | | | | | |
|-----|----|------------|------------------------------------|-------------|
| (4) | a. | /fʊɾta/ | [ft ^h a] | ‘fence’ |
| | b. | /sʊɾto/ | [st ^h o] | ‘attic’ |
| | c. | /kʃʊw/ | [kʃʊ ^w] | ‘untamed’ |
| | d. | /kʊʃʊw/ | [k ^h ʃ:ʊ ^w] | ‘seer’ |
| | e. | /kʊʦʊkʃʊŋ/ | [k ^h ʦkʃiŋ] | ‘Yame clan’ |

Although according to the phonetic transcriptions all the examples above seem to have consonant clusters, there are only three true clusters among them, namely, [kʃ] in (4)c., and [ts] and [kʃ] in (4)e. However, this is not so obvious on the surface. For example, in Figure 4, which shows spectral and waveform representations of example (4)b. [st^ho] ‘attic’, no vowel is visible between /s/ and /t/. However, phonological evidence discussed below in section 4.4 suggests that this is not a true cluster in the language because /s/ has a “hidden” vowel /ʊ/ following it. This is a case where, due to devoicing, the vowel is completely obscured and overlapped with the preceding consonant, leading to the impression that it is deleted (Coleman 1992, 2001). This in turn can result in a mistaken impression that the language has clusters such as /st/ and /ft/ in words like [sʊɾto] ‘attic’ and [fʊɾta] ‘fence’.

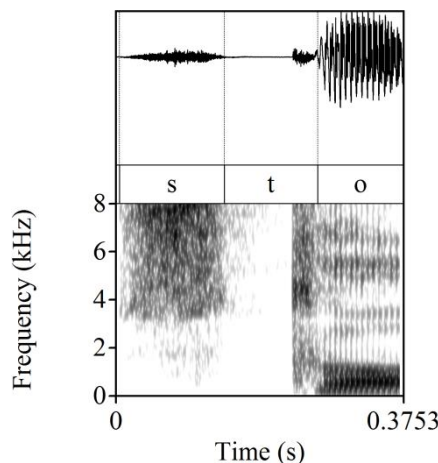


Figure 4: Waveform and spectrogram of [st^ho] ‘attic’

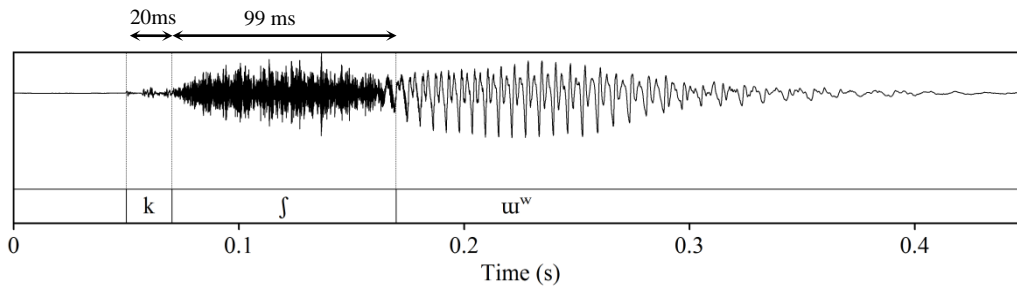
Failure to distinguish between the true and pseudo clusters in the language can lead to considerable confusion in phonemic and orthographic transcription of words like (4)d. /kʃuʷ/ ‘untamed’, and (4)c. /kʷʃuʷ/ ‘seer’. Moreover, in words like /kʷtsuʷkʃuʷ/ ‘Yame clan’, with a greatly reduced vowel /uʷ/ in the first two syllables, it would be difficult to draw syllable boundaries and understand the consonant sequences if one considers it to be a case of vowel deletion, which yields an implausible cluster /kʷtskʃ/. Although the examples given here involve only the vowel /uʷ/, similar confusion can arise in the case of the other two high vowels, /i/ and /u/ as well.

4.4. Devoicing or deletion?

From the acoustic data, whether words in isolation or normal speech, it is not always easy to determine whether the high vowels in the devoicing environments described above are devoiced or deleted. The difficulty is greater when high vowels follow voiceless fricatives or the vowel /uʷ/ is involved. This is amply illustrated in the case of /sʷto/ ‘attic’, a plausible phonetic transcription of which is [st^ho] in Figure 4, where vowel formant bands between /s/ and /t/ are completely absent.

Figure 5 illustrates the difficulty further. The figure depicts the waveforms of examples (4)c. [kʃuʷ] ‘untamed’ & (4)d. [k^hʃ:uʷ] ‘seer’ discussed in section 4.3 above. For ease of visual comparison, the words are represented in a 450 ms time frame, with focus on the sequence /kʃ/ and disregarding the final vowel /uʷ/. What is immediately obvious from the figures is the difference in the length of the segments. The time taken from the initial burst of /k/ to the onset of /ʃ/ is found to be 18 ms longer in 4(d). However, in normal speech when tokens are uttered with variable speed, such a small difference may not be a reliable perceptual cue. More striking is the difference in the length of 131 ms between the instances of /ʃ/ in the two words (99 ms and 230 ms respectively). Moreover, /k/ is aspirated in the second word but not in the first word. Neither length nor aspiration are phonemically contrastive in the language. Lengthening of certain singleton consonants preceding a high vowel is a phonotactic rule in the language, therefore lengthening of /ʃ/ in (4)d. [k^hʃ:uʷ] indicates that /ʃ/ is singleton. Moreover, aspiration is observed in all singleton voiceless stops due to the [SPREAD GLOTTIS] feature. This further indicates that /k/ in (4)d. [k^hʃ:uʷ] is singleton. On the contrary, the fact that in (4)c. [kʃuʷ], neither lengthening of /ʃ/ nor aspiration of /k/ are observed indicates that they are not singleton consonants. Furthermore, [k^hʃ:uʷ] is clearly a disyllabic word. Evidence for its disyllabicity comes from the syllable inversion game discussed in section 4.4.3; see example (7)d. If there is no vowel between /k/ and /ʃ/ in (4)d. [k^hʃ:uʷ], then the only way it could be a disyllabic word is by /k/ being a syllabic consonant, as in [k^hʃ:uʷ]. However, if extrapolated to other words, such a conclusion would yield quite unusual syllable sequences as in [k^h.ts.kʃiŋ] ‘Yame clan’ and put Hrusso Aka in the category of languages with vowel-less syllables like Tshlhyit Berber (Ridouane 2008). This is neither necessary nor true, as I shall show in the following discussion.

a.



b.

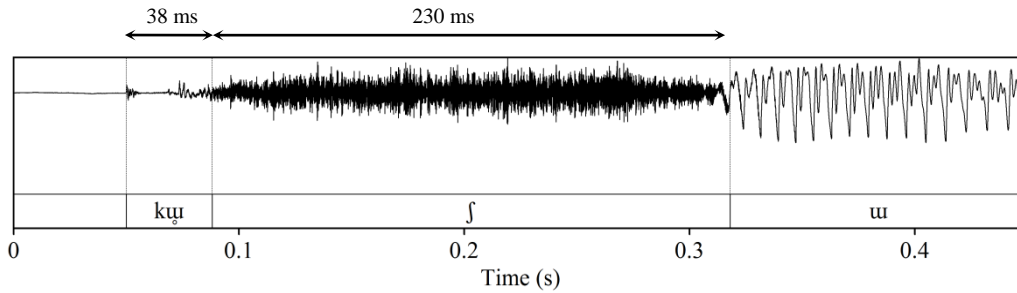


Figure 5: Waveform representations of (4)c. [kʃuʷ] ‘untamed’ and (4)d. [kʰʃuʷ] ‘seer’

4.4.1. Evidence from articulatory gestures.

Preservation of articulatory gestures in the case of obscured vowels can be an important clue to their presence. For example, Gick et al. (2012) have shown that in the case of the “soundless” vowels of Blackfoot and Oneida languages some vowels in certain environments have nil acoustic output, but their articulatory correlates like lip rounding and tongue position are preserved. In a similar manner, the presence of devoiced and obscured Hrusso Aka vowels can be detected by lip gestures even when auditory and acoustic data fail to indicate their presence. Figure 6 below shows the lip gestures during the pronunciation of /xʷiʃu/ ‘tiger’ and /xʷsádʒio/ ‘waterhen’.



Figure 6: Lip gestures in the pronunciation of [ʷi] and [ʷu]

The pictures show both the front and the side view simultaneously.¹² The picture on the left shows spread lips during the pronunciation of /xʷi/ and the picture on the right shows rounded lips during the pronunciation of /xʷu/, proving that neither vowel is deleted. An

¹² The side view was captured using a mirror.

important fact to note here is that devoiced vowels leave their articulatory marks on preceding consonants because of gestural overlap (Coleman 1992, Varden 2010). This would lead to “colouring” of the consonant by the obscured vowel, which gives insight into native speakers’ perception of obscured vowels. Such consonant colouring is a matter for further investigation, and will not be analysed in this paper.

4.4.2. Evidence from slow speech

Devoiced vowels always surface in slow and careful speech and are voiced, a fact observed also in Japanese (Martin et al. 2014). This is a good test to distinguish between true and pseudo clusters in words such as (4)c. /kʃuw/ ‘untamed’ and (4)d. /kʷʃuw/ ‘seer’. In a slow speech experiment, it was found that no matter how slowly it was uttered, the sequence /kʃ/ in /kʃuw/ was never broken, and the word was pronounced as [kʃu:.....w].¹³ In contrast to this, /kʷʃuw/ was pronounced always as [k^hu:.....ʃu:.....w] indicating that the native speakers perceive a vowel between /k/ and /ʃ/.

4.4.3. Evidence from syllable games.

A language game played by children in some Hrusso Aka villages involves mirror imaging words by syllable inversion as follows:

- | | | | | |
|-----|----|-----------|-----------|---------------|
| (5) | a. | /sundade/ | /dedasun/ | ‘therefore’ |
| | b. | /kameje/ | /nemeka/ | ‘having come’ |

True clusters are never broken up either by syllable boundaries or by vowel insertion in these games, as the following examples demonstrate:

- | | | | | |
|-----|----|----------|-----------|------------|
| (6) | a. | /labzuŋ/ | /bzuŋ.la/ | ‘lentil’ |
| | b. | /utʃu/ | /tʃu.u/ | ‘skin’ |
| | c. | /puksʃu/ | /kʃu.pu/ | ‘outsider’ |
| | d. | /dzubzu/ | /bzu.dzu/ | ‘pumpkin’ |

However, when devoiced high vowels are present, they invariably surface and participate in syllable inversion. The examples from (4) yield the following results:

- | | | | | | |
|-----|----|---------|---------|----------|-----------|
| (7) | a. | /fʊta/ | */fta/ | /ta.fu/ | ‘fence’ |
| | b. | /sʊto/ | */sto/ | /to.su/ | ‘attic’ |
| | c. | /kʃuw/ | | /kʃuw/ | ‘untamed’ |
| | d. | /kʷʃuw/ | */kʃuw/ | /ʃuw.ku/ | ‘seer’ |

Note that (7)c. remains unchanged since it is a monosyllabic word containing a true cluster, and there is no hidden vowel that can surface. In all the other words, the hidden vowels surface. Similar results are obtained in the secret language of the elders, Zekulyu, which has the same syllable inversion strategy as in the language game discussed above.

¹³ Repeated length marks “.....” indicate extremely slow speech.

5. Conclusion

This preliminary discussion raises many issues for further investigation. First, pervasiveness of high vowel devoicing needs to be studied with a larger and more variable corpus. Vowel devoicing has been noted to be quite a variable phenomenon even within a single language, ranging from partial devoicing to complete deletion depending on not only phonetic environments; and also within speakers, across different types of speech and communicative contexts (see Varden 2010: 38, Martin et al. 2014). The frequency of devoicing when the necessary conditions are present, and the extent of devoicing in the presence of different triggers such as the voiceless stops and fricatives, are both interesting areas for further exploration. One related issue here concerns the devoicing environment when high vowels are followed by a sonorant, where the data show a somewhat irregular pattern of devoicing, as in the following example:

- (8) a. /kífiw/ ‘able to touch’
b. /sɯɲá/ or /sɯɲá/ ‘that much’

In the case of the word /kífiw/, no devoicing of the vowel /i/ in the second syllable was observed. However, the vowel /u/ in /sɯɲá/ was voiced in all the isolated utterances, but consistently devoiced in a sentence. The data suggests that occasional irregularity is found in all environments, but high vowels are more prone to resist devoicing when followed by a sonorant. This too needs further study.

As for communicative contexts, it is also not clear as to what extent high vowel devoicing is present during formal speech, shamanic chanting and child-directed speech where slow and deliberate speaking is the norm. Preliminary investigation indicates that high vowel devoicing is almost completely absent in songs, but this needs to be corroborated with further evidence.

Another question is about the nature of influence that the high vowels exert on the preceding voiceless consonants during devoicing. This is particularly important in the case of the high vowels that are completely obscured during devoicing. We have already seen that the articulatory gestures are retained in the case of obscured vowels, giving visual cues to the listener. Are there also acoustic cues present? In the case of Japanese high vowel devoicing, it has been observed in the acoustic studies by Beckman and Shoji (1984) and Varden (2010) that fricatives are sufficiently “coloured” by the devoiced vowels, which makes the perception of the vowels possible even when they are deleted. Further investigation in this direction is needed to understand the nature and extent of such acoustic cues in Hrusso Aka. Also as regards perceptual cues, the heavy affrication of voiceless stops preceding devoiced high vowels needs further study. Although such affrication is common cross-linguistically (Lin 2011), it is noticeably heavier in Hrusso Aka, perhaps due to the need to increase perceptual salience in the context of devoiced vowels, and contributes to the peculiarity of the language.

An elaborate phonological analysis of high vowel devoicing is not attempted in this descriptive study – see D’Souza (2015), where the process is analysed as resulting from spreading of the [SPREAD GLOTTIS] feature to vowels from the preceding voiceless obstruents.

Recognizing high vowel devoicing and distinguishing it from the other processes as vowel deletion and cluster formation gives insight into and clarity on many phonological processes in the language, and helps to understand the complexities of its consonantal structure. More importantly, from the point of view of the Hrusso Aka native speaker community, the crucial importance of the study of high vowel devoicing lies in how it influences the orthography of this endangered language. Representing the devoiced vowels in orthography is important to reflect the psychological reality of these vowels in the mind of the native speaker and to avoid implausible consonant sequences that can render orthography difficult to read.

Finally, one suspects that there are more languages in the Tibeto-Burman area that have vowel devoicing.¹⁴ For example, van Breugel (2012) briefly describes vowel devoicing in Atong. Also, Teo (2014: 41–42) notes a similar phenomenon in Sumi, but treats it as high vowel syncope. It will be interesting to see if, with further articulatory investigation, this turns out to be a case of high vowel devoicing.

A wider investigation of this hitherto under-described phenomenon will not only help in better understanding the phonologies of individual languages, but also enrich research in the areal typology of the region.

References:

- Anderson, Gregory and Ganesh Murmu. (2010). “Preliminary notes on Koro, a “hidden” language of Arunachal Pradesh.” *Journal of Indian Linguistics* 71: 1–32.
- Anderson, James D. (1896). *A short vocabulary of the Aka language*. Shillong.
- Beckman, Mary and Atsuko Shoji. (1984). “Spectral and perceptual evidence for CV coarticulation in devoiced /si/ and /syu/ in Japanese.” *Phonetica* 41: 61–71.
- Blench, Roger and Mark W. Post. (2011). *(De)Classifying Arunachal Languages: Reconsidering the Evidence*. Unpublished manuscript.
- _____. (2014). “Rethinking Sino-Tibetan Phylogeny from the Perspective of North East Indian Languages.” In Nathan Hill and Thomas Owen-Smith, Ed. *Trans-Himalayan Linguistics*. Boston, De Gruyter Mouton: 71–104.
- Bodt, Timotheus A. and Ismael Lieberherr. (2015). “First notes on the phonology and classification of the Bangru language of India.” *Linguistics of the Tibeto-Burman Area* 38(1): 66–123.
- Boersma, Paul and David Weenink. (2017). *Praat: doing phonetics by computer* (Version 6.0.29). Retrieved from <http://www.fon.hum.uva.nl/praat>
- Bradley, David. (1997). “Tibeto-Burman languages and classification.” In David Bradley, Ed. *Papers in Southeast Asian linguistics No. 14: Tibeto-Burman languages of the Himalayas*. Canberra, Pacific Linguistics Research School of Pacific and Asian Studies, Australian National University: 1–72.
- Campbell, George. (1874). *Specimens of Languages of India, including those of the Aboriginal Tribes of Bengal, the Central Provinces, and the Eastern Frontier, Calcutta*. Calcutta, Bengal Secretariat Press.
- Coleman, John. (1992). “The phonetic interpretation of headed phonological structures containing overlapping constituents.” *Phonology* 9: 1–44.
- _____. (2001). “The phonetics and phonology of Tashlhiyt Bereber syllabic consonants.” *Transactions of the Philological Society* 99: 29–64.
- Dauer, Rebecca M. (1980). “The reduction of unstressed high vowels in Modern Greek.” *Journal of the International Phonetic Association* 10(1–2): 17–27.
- D’Souza, Vijay A. (2015). *Towards a Phonology of Hrusso Aka*. MPhil thesis. Oxford, University of Oxford.
- Gick, Bryan, Heather Bliss, Karin Michelson and Bosko Radanov. (2012). “Articulation without acoustics: ‘Soundless’ vowels in Oneida and Blackfoot.” *Journal of Phonetics* 40(1): 46–53.
- Gordon, Matthew. (1998). “The phonetics and phonology of non-modal vowels: A cross-linguistic perspective.” *Proceedings of the Twenty-Fourth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Phonetics and Phonological Universals*. Berkeley, CA, Berkeley Linguistics Society: 93–105.

¹⁴ Thanks to Prof. Shakuntala Mahanta, IIT, Guwahati for sharing her observations on this matter.

- Grierson, George A., Ed. (1909). *Linguistic Survey of India* (Vol. III,1. Tibeto-Burman family). Calcutta, Government of India, Central Publication Branch.
- Harrison, K. David. (2010). *The last speakers: the quest to save the world's most endangered languages*. Washington, D.C., National Geographic.
- Hesselmeyer, C. H. (1868). "The Hill Tribes of the Northern Frontier of Assam." *Journal of the Asiatic Society of Bengal* 27: 192–208.
- Jaeger, Jeri J. (1978). "Speech aerodynamics and phonological universals." *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA, University of California: 311–329.
- Jannedy, Stefanie. (1995). "Gestural phasing as an explanation for vowel devoicing in Turkish." *Ohio State University Working Papers in Linguistics* 45: 56–84.
- Jun, Sun-Ah and Mary E. Beckman. (1993). "A gestural-overlap analysis of vowel devoicing in Japanese and Korean." *Paper presented at the 67th Annual Meeting of the Linguistic Society of America, Los Angeles, CA, January 7–10*.
- Kondo, Mariko. (1997). *Mechanisms of Vowel Devoicing in Japanese*. Edinburgh, University of Edinburgh.
- Ladefoged, Peter and Ian Maddieson. (1996). *The Sounds of the World's Languages*. Oxford, Blackwell Publishers.
- Laver, John. (1994). *Principles of phonetics*. Cambridge, Cambridge University Press.
- Lin, Yen-Hwei. (2011). "Affricates." In Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume and Keren Rice, Eds. *The Blackwell Companion to Phonology Vol 1*. Malden, MA, Wiley-Blackwell: 367–390.
- Martin, Andrew, Akira Utsugi and Reiko Mazuka. (2014). "The multidimensional nature of hyperspeech: Evidence from Japanese vowel devoicing." *Cognition* 132: 216–228.
- Payne, Russel. (1909). "Aka text, vocabulary, and phrases." In *Linguistic Survey of India* (Vol. 3.1). Calcutta, Office of the Superintendent of Government Printing: 582–583, 622–640.
- Ray, Punya Sloka. (1967). "Dafla phonology and morphology." *Anthropological Linguistics* 9(8): 9–14.
- Ridouane, Rachid. (2008). "Syllables without vowels: Phonetic and phonological evidence from Tashlhiyt Berber." *Phonology* 25(2): 321–359.
- Shafer, Robert. (1947). "Hruso." *Bulletin of the School of Oriental and African Studies* 12(1): 184–196.
- Shibatani, Masayoshi. (1990). *The languages of Japan*. Cambridge, Cambridge University Press.
- Sinha, Raghuvir. (1962). *The Akas, the people of NEFA*. Shillong, Research Department, Adviser's Secretariat.
- Teo, Amos. (2014). *A phonological and phonetic description of Sumi, a Tibeto-Burman language of Nagaland*. Canberra, Asia-Pacific Linguistics.
- Thurgood, Graham and Randy J. LaPolla. (2007). *The Sino-Tibetan languages*. London, Routledge.
- Torreira, Francisco and Mirjam Ernestus. (2010). "Phrase-medial vowel devoicing in spontaneous French." *Proceedings of the 11th Annual Conference of the International Speech Communication Association (Interspeech 2010)*. Makuhari, Japan: 2006-2009.
- Tsuchida, Ayako. (2001). "Japanese vowel devoicing: Cases of consecutive devoicing environments." *Journal of East Asian Linguistics* 10(3): 225–245.
- Van Breugel, Seino. (2014). *A Grammar of Atong*. Leiden, Brill.
- Varden, J. Kevin. (1998). *On High Vowel Devoicing in Standard Modern Japanese: Implications for current phonological Theory*. PhD dissertation. Department of Linguistics, University of Washington.

_____. (2010). "Acoustic correlates of devoiced Japanese vowels: Velar context." *The Journal of English & American Literature and Linguistics* 135: 35–49.

A cross-varietal description of modifiers of basic colour terms in Tangsa-Nocte¹

Kellen Parker van Dam

Centre for Research on Linguistic Diversity, La Trobe University, Australia

Abstract

In many varieties of the Tangsa-Nocte language group, there are a handful of unique suffixes that modify simple adjectives. These terms include but are not limited to the majority of basic colour terms, with which they show up with a high frequency cross-varietally. Such terms have not yet been addressed in any depth, aside from brief mention in a single unpublished grammatical sketch. This paper remedies that by providing a description of the various forms the modifiers take, and by addressing cross-varietal similarities and differences in both the affixes and lexicon of basic colour terms. In the past these modifiers have been referred to as intensifiers (Morey, p.c.). However, it is now clear that this is only one function, and that these affixes provide a much greater range of nuance in the majority of Tangsa-Nocte varieties studied to date. With this in mind, this paper addresses a number of questions on the topic of affixes. First, to what extent do basic colour terms show consistency cross-varietally? Second, to what extent are the modifiers known across speech communities? And finally, to what degree may they be considered cognate?

Citation	van Dam, Kellen Parker. 2018. <i>A cross-varietal description of modifiers of basic colour terms in Tangsa-Nocte</i> . North East Indian Linguistics (NEIL) 8, 47-64. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.
Volume Editors	Linda Konnerth, Stephen Morey, Amos Teo
Copyright	© 2018, the author(s), release under Creative Commons Attribution license
URL	http://hdl.handle.net/1885/155280

1. Introduction

Tangsa² is a Tibeto-Burman language, classified under Northern Naga. It is spoken primarily in the Patkai mountain range and the surrounding lowlands in Northeast India and Northwest Myanmar. In many varieties of the Tangsa-Nocte language group, there are a handful of suffixes which modify simple adjectives. They appear not to occur with compound or prefixed stems. However, they can occur on the majority of basic colour terms and common adjectives.

¹ Work for this paper was undertaken as part of a doctoral research scholarship provided by the Australia Research Council. Thank you to my colleagues Niharika Dutta, Deepjyoti Goswami, Mijke Mulder and Syed Iftiqar Rahman, who've helped considerably by sharing their time and observations, and to my graduate advisor Stephen Morey. The speakers who assisted in this study and made it possible were Ninshom Chena, Sunumi Changmi, John Darok Chamkok, Dungmon Champang, Nubong Gaqlun, Ngopong Hahcheng, Kamtong Kamba, Bin Khämlan, Kejung Kevang, Datho Langching, Andrew Longri, Chonsam Longri, Pongshan Lungkhi, Nepo Maitai, Wanglung Kelim Mossang, Renman Kelim Mossang and family, R.B. Mungre, Wangkui Ngaimong, Machow Ngaimong, Sanpe Rangsi, Linkam Rekhu, Wamyong Ronrang, Longrem Ronrang, Namshap Ronrang, Mohem Ronrang, Simon Ronrang, Tipwun Shangthi, Nokrang Shangthi, Shoqreu Shecyü, Khulkam Taikam, Takjung Taikam and family, Nong Yongkuk, Wangdan Yongkuk, Walong Yongkuk and Songkam Yongkuk. I am grateful to Kheshem of Zeephyugone for his time and patience.

² Tangsa is also referred to as Tangshang in Myanmar, a recent label which includes Wancho and Tutsa as well as Nocte, the most closely related language group of varieties. Within groups referred to as Tangsa in India, Burmese speakers may instead identify as Rangpang or Heimi, two of the larger subgroups. There is considerable overlap between Nocte and Tangsa varieties. For this reason, the joint term “Tangsa-Nocte” is used here, in order to be most inclusive and less restricted by divisions, which are more sociopolitical than linguistic.

This paper provides a description of the various forms and uses of these modifiers, as well as addressing similarities of these terms across the different varieties³ of Tangsa and varieties of closely related language varieties which may fall outside of the usual “Tangsa” classification. In the past, these modifiers have been referred to as intensifiers (Morey, p.c.). However, it is now clear that intensification is only one function of these affixes.

A typical example of how these modifiers are used is given in the following examples. The basic term for RED is /ə.ʃaŋ/, as in example (1). A modifier /de/ exists which can be considered an intensifier. When attached to the term for RED, the nominal prefix /ə-/ is dropped, and the translation changes from ‘red’ to ‘truly red’, example (2). The modifiers may also be reduplicated as in (3), which does not necessarily increase the intensity of the modifier. In many cases, the resulting meaning is less strong than if the modifier were not reduplicated. Modifiers also exist, such as /ro/, which lessen the intensity of the colour which they modify, as in example (4).

	prefix	root	modifier	gloss
1)	ə-	ʃaŋ		‘red’
2)		ʃaŋ	de	‘truly red’
3)		ʃaŋ	de de	‘more red than /ə.ʃaŋ/’
4)		ʃaŋ	ro ro	‘less red than /ə.ʃaŋ/’

The following will be addressed in this paper through examples from a wide range of Tangsa-Nocte varieties:

1. To what extent do basic colour terms show lexical consistency cross-varietally?
2. To what extent are the modifiers shared across speech communities?
3. To what degree may the modifiers be considered cognate cross-varietally?

1.1. Field Sites & Fieldwork

Modifiers of basic colour terms were collected over the course of three months of fieldwork in the winter of 2015-2016. They were elicited from speakers of various varieties of the Tangsa language, including Chamchang (Kimsing), Hahcheng, Joglei, Moklum, Mueshaung (Mossang, Mueshaungx), Mungre (Mawrang), Ngaimong, Phong (Pontai), Rera (Ronrang) and Shangwal in Arunachal Pradesh and Upper Assam, India. Additional work was performed in March 2016 in Sagaing Region, Myanmar. Data from Wancho, a related Northern Naga language spoken in Myanmar, India and possibly Bhutan, have also been included in some of the following discussion for cases where they may be relevant for comparative purposes.

For the sake of brevity and in order to make the current study more manageable, the only modifiers which were elicited were ones which attach to colour terms that could potentially be considered basic. Such modifiers are also attested for a much larger number of adjectives. However, due to their greater inconsistency between dialects and the difficulty in consistent elicitation across varieties, as well as the overwhelming number of terms that would need to be addressed in the short time spent in the field, the focus here has been reduced to deal only with the modifiers on colour terms and not the larger class occurring across common

³ It remains to be seen how the various language varieties which fall under the classification “Tangsa” and “Nocte” are actually related to each other. Much more work needs to be done to be able to provide a good model of relationships.

adjectives. Terms were elicited through conversations with speakers, rather than with colour stimuli such as Munsell colour chips.

In citation form, colour terms will take a nominalising prefix (most typically /ə-/ or /e-/). Cognates of this nominalising prefix are found in Tibeto-Burman languages of the region (Matisoff 2003: 106). In Tangsa, the prefix itself is toneless and unstressed, but it may trigger changes in the tone of the following stem.

In the examples given in this paper, only the root form is provided, without the use of the /ə-/ prefix. The nominalising prefix is typically lacking when a modifier is affixed, although there are exceptional varieties for which this pattern is absent. For the sake of space and clarity, it has been omitted here.

Subscript numerals refer to phonemic tone categories. All syllables are marked where I am confident about the toneme. For many lexemes, approximate reconstructions have been offered for Proto-Tangsa-Nocte (PTN). These are provided for the sake of having a way to refer to the lexemes as they occur in the following pages. It is not an attempt to present an overall reconstruction of the proto-forms. I have also made an effort to suggest possible connections to other Tibeto-Burman languages and have provided Proto-Tibeto-Burman (PTB) reconstructions where connections are likely.

2. Berlin & Kay's Hierarchy of Basic Colour Terms

In order to address the role of modifiers, it is necessary to first address the colour system as it is found in the Tangsa languages surveyed here.

What constitutes a basic colour term is here defined following Berlin & Kay (1969). In this definition, the following criteria must be met. First, the term must be known and consistently used by members of the speech community. Second, it must not be considered a subset of some other colour term (e.g. English *pink* treated as a subset of RED). Third, the word's meaning must not be apparent from the component parts, for example *sky blue* in English.

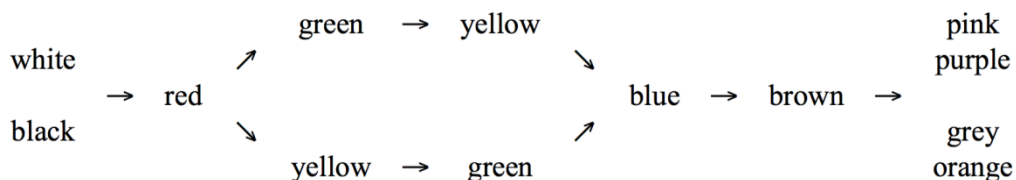


Figure 1 – Hierarchy of basic colour terms (Berlin & Kay 1969)

It should be stated that it would be inaccurate to describe a four-colour system as consisting of only those four colours as imagined by modern English speakers, namely black, white, red and green. In many cases a better way to label the categories may be DARK, LIGHT, WARM and COOL to better reflect which colours tend to fall into these categories.

This will differ for three-category systems, for which 'warm' no longer applies as accurately, or five-colour systems where what falls under YELLOW would otherwise be considered either WARM or COOL depending on context in a four-category system. Since the labels DARK, LIGHT, WARM and COOL are really only best suited for a four-category system, the English colour names are used here instead. While the basic colour terms of English are used in this paper to represent the categories as presented by Berlin & Kay, it is important to recognise their use not as reflecting an accurate one-to-one translation between the two.

2.1. A Note on Terms

In what follows, I employ the use of small capital letters when referring to colours as categories. As an example, GREEN refers not to the concept as an English speaker might translate it, but rather the category in the colour system as described in Berlin & Kay (see Figure 1). When in quotations, ‘green’ is a gloss in English. Thus, to say that ‘yellow’ is classified within RED in Tangsa, what I mean is that the colour hue that might be called *yellow* in English is contained within the same category in Tangsa as what in English we might call ‘red’ in English, RED being the more fundamental category in the hierarchy.

For systems which are here identified as falling under Stage IIIa, containing BLACK, WHITE, RED and GREEN, the labels are equivalent to the labels DARK, LIGHT, WARM and COOL respectively. I have used both labels where appropriate as an attempt to make it clear for the reader that RED does not strictly speaking mean ‘red’ translated into English.

3. The Basic Tangsa-Nocte Colour System

After discussion with members of a wide range of Tangsa dialects, it is clear that the basic colour system is currently undergoing a change where the term for YELLOW, **mjən*, has been added as a basic term widely accepted by younger speakers (30s and below) but rejected as basic by the oldest speakers (70s and above) who deny that there is a word for YELLOW, stating instead that it is referred to as “turmeric colour” if it needs a special label. In the current study, younger speakers have often claimed ignorance of any recent origin of **mjən* and treat it as a basic colour word. However, as will be shown further down, it still lacks an important morphosyntactic feature true of other basic colour terms in Tangsa – the ability to co-occur with modifiers.

In fact, much discussion was prompted by what were at the time thought by myself as straightforward, and eliciting colour terms to be applied to objects treated as GREY, PURPLE, BLUE among others showed no real consensus between varieties or sometimes between speakers. In other words, an object which we would consistently refer to as ‘blue’ in English may be given any number of labels in Tangsa. The same goes for what we would consistency label ‘grey’ or ‘purple’ in English.

More surprising is that, among the oldest speakers consulted, items which might be described with YELLOW terms were alternately described with those normally representing RED or WHITE. In other words, the colour of a specific item might be called one word one time, for example /men/, and another five minutes later, for example /ʃaŋ/, where the two words given refer to two different basic colour terms YELLOW and RED respectively.

What in English may be called ‘blue’ were most often described with terms covering the GREEN category, but sometimes also those for BLACK or GREY (when the situation is interfered with and GREY is given as an option). Hues named ‘purple’ in English may be considered BLACK or GREY, or given a new label which will be discussed further below.

The older speakers’ system may be labelled a Stage IIIa system (Kay & McDaniel 1978). Younger speakers are now floating somewhere between Stages IV and V, depending on which community is being consulted. An additional change which is ongoing is the addition of other terms being treated as basic, but with far less consistency cross-varietally as is found with YELLOW **mjən*. One of the more common examples is **k^huai₂* for GREY. While GREY was originally proposed to occur quite late in Berlin & Kay’s (1969) hierarchy, later revisions allowed for more flexibility with regard to its appearance (Kay 1975, Kay, Berlin & Merrifield 1991). This too is not so clear-cut, and **k^huai₂* may occur just as often referring to BLUE as GREY.

A number of Tangsa-Nocte varieties have disyllabic basic colour terms, for which the behaviour of the term when affixed by a modifier changes. The Rera variety found in the villages of Balinong and Phulbari is one such case. Many of the basic colour terms are made up of two morphemes. The category for BLACK is no exception.

The basic term for BLACK in isolation is /ɐ.nak₄-k^hɛm₁/. It was also given as /ɐ.nak-man/; the distinction between the meanings of /k^hɛm₁/ and /man/ was unclear, other than that /man/ was less frequently used than /k^hɛm₁/. This lexeme /nak₄/ is derived from the same etymon as BLACK (section 5.1), while /k^ham₁/ is the lexeme for ‘burn’, likely of the sense of ‘charred’. This same /k^ham₁/ lexeme also occurs with RED in Ronrang. A minority of varieties addressed in this paper have polysyllabic forms for some basic colour terms, but those varieties in which they appear do not have polysyllabic terms for all basic colour terms. Overall there seems to be a larger inventory of modifiers and greater likelihood of polysyllabic terms for basic colour terms near the start of Berlin & Kay’s hierarchy. It remains to be seen how well this observation holds up with a much larger sample of Tangsa-Nocte varieties.

4. The Role of Modifiers

While not all those consulted for the present study agree, or were able to articulate the specific details of the colour terms, the overall consensus was that there are two main types of what were previously called “intensifiers” (Morey, p.c.), which I refer to as “modifiers”. The first are used to strengthen the meaning of the term used, while the second set of common affixes tends to lessen the implied intensity. In nearly all cases, the modified colour term can have the structure AB, where A is the basic colour term and B is the suffix. Alternatively, it can have the structure ABB with the suffix being reduplicated. This too causes a shift in meaning. Alone, the colour term is given the nominal prefix, most typically /ə-/ , thus əA. When suffixed with the modifier, this /ə-/ prefix is dropped. A construction of əABB would be ungrammatical in the varieties investigated here.

With this in mind, the basic system is as follows. əA is the basic elicitation form of the colour term. N əA means ‘the N is A-coloured’ and is a sentence on its own. Without the /ə-/ prefix — N A — the result is a noun phrase. Taking examples from Mueshaung, /man əʃɔ/ should be glossed ‘the cow is red’, while /man ʃɔ/ is ‘the red cow’, as seen in examples (1) and (2).

(1) *man* ə-ʃɔ
 cow NOMZ-red
 ‘The cow is red.’

(2) *man* ʃɔ
 cow red
 ‘The red cow...’

As mentioned, there are two common general categories of modifying suffixes, one which increases intensity and one which lessens it. These are not the only modifiers, but are the two most common categories cross-varietally. In the former category, where the modifiers are intensifiers, the AB form is often translated as the “true” colour by consultants. Meanwhile, ABB is often translated by consultants as ‘A-ish’. In the other common category where the modifiers lessen the intensity of the colour, both AB and ABB signal a ‘less A than A’ meaning, and the ABB pattern seems to be more typical.

There is also the question as to whether the modifiers discussed here are suffixes or not. It is not yet clear to me if all of these really should be treated as suffixes. Many of these do exist

as words on their own, and many others do not. For the sake of simplicity, they have been treated as suffixes here. More work needs to be done to work to determine which modifiers are more like words that can stand on their own and which cannot.

5. Overview of Modifiers

The following sections present the inventories of all modifiers as collected from consultants up to the time of writing, organised by colour category.

In many – if not most – cases, the modifiers create a clear change in meaning which can be agreed upon by all surveyed members of the speech community. In some cases, the meaning has been clearly explained by consultants, and in a way which is consistent across varieties. The modifier /put/, for example, when applied to WHITE will always mean ‘spotted’ or ‘speckled’. However, due to limitations in the ability to communicate such nuance between myself and the speakers, I have been unable to ascertain the exact meaning intended by the modifier for many of the modifiers given. Thus, no translation can be provided for many of the modifiers, at the time of publication. I have done my best to indicate meanings where they are known.

In each section below, a table is given showing the varieties being addressed, the basic colour terms for those varieties, and then the modifiers which have been attested.

5.1. Black, Dark

5.1.1. Basic colour names

There are two likely related etyma for BLACK in Tibeto-Burman following Matisoff’s (2003) Proto-Tibeto-Burman (PTB) reconstruction. One is PTB **s-nak*⁴ > Northern Naga **njak* BLACK and another PTB **s-ma(ŋ/k)* > PTN **mak* DARK. The distinction is retained in certain varieties, such as Moklum and Phong, but in many other dialects **njak* would translate as both ‘dark’ and ‘black’. Not all varieties are the same in their treatment of what in English would be called *dark*. In varieties where both **mak* and **njak* are used, there is disagreement between speakers of different varieties as to which means ‘dark’ and which means ‘black’ as a basic colour term. For example, in Moklum, /mak/ is treated as the basic colour term while /nak/ has been given as ‘dark’, as in describing the night. This is contrary to the etymologies given above.

In Central Naga, **njak* is attested in Ao as /anjak/ (Marrison 1967). In reconstructions for Proto-Tibeto-Burman, there is a reconstructed **s-* prefix for this etymon making it **s-nak*. It has been argued elsewhere that the glide is likely a result of metathesis from the **s-* prefix – see Matisoff (2003: 101) regarding this change. This is fairly uncontroversial as a basic colour term.

An additional form **k^ha* is found in varieties spoken in Myanmar, notably Rinkhu, Lungkhi and Shangthi. In the case of Rinkhu, no modifiers are attested.

5.1.2. Modifiers

Table 1 shows the most common modifiers of BLACK elicited for a number of varieties. Not all speakers of a single variety provided the same modifiers. However, for varieties in which **duk* and **rieC* were found, they were provided consistently from speaker to speaker. In both cases, the meaning given was something like “exceptionally black”.

⁴ Etymon #2483 from the Sino-Tibetan Etymological Dictionary and Thesaurus (STEDT) (<http://stedt.berkeley.edu>)

I have attempted to arrange the following tables so that the first column contains the most typical modifier. The last column shows cognates of **ro*. This is the only modifier which appears cross varieties as well as applying to the majority of the colour terms in a given variety. The modifier **ro* has the meaning that the colour is less intense than normal. This does not necessarily mean ‘faded’, as there is another modifier (**mɔt*) which may be used in some varieties and also perhaps co-occurs with more than one basic colour term. Where used, **ro* can generally be considered an opposite to the term given in the first column.⁵

Table 1 – Modifiers for BLACK in Tangsa-Nocte varieties

	<i>*njak₄</i>	<i>*duk₄</i>	<i>*rieC₄</i>	<i>*sim</i>	<i>*rum₃</i>	<i>*ham⁵</i>	<i>*htaun₂</i>	<i>*tʃaŋ</i>		<i>*ro₁</i>
Chamkok	n ^h ak ₄	duk	rik	sim	rum					
Hahcheng	n ^h ɛk ₄	duk ₄	r ^h iet ₄			ham ₂				
Joglei	n ^h ak ₄	duk ₄		sim ₂					k ^h liʔ ₄	ro ₁
Chamchang	n ^h aʔ ₄	dok ₄			rum ₃					ro ₁
Lungkhi	mak								yip	
Lungphi	n ^h ɔk ₄	duk ₄	riet	sim	rim ₃	ham	htaun ₂			ruu ₂
Maitai	n ^h aʔ									
Moklum	nak/mak	dik		sam ₂	rim					
Mueshaung	n ^h auk ₄	duk ₄			rum ₃		htaun ₂			ruu ₂
Mungre	n ^h aʔ ₄		riyʔ ₄		rum ₃				deɣʔ	ro ₁
Ngaimong	n ^h ɛk ₄	duk ₄	riek ₄							
Phong	mak ₄	dik	rik		ram ₂				t ^h um	
Pinkhu	mak								diŋ	
Rangsi	nak/mak									
Rera	nak ₄ .k ^h am		riit ₄	ʃip				tʃaŋ		
Shangthi	nak*			ʃik						
Shangwal	n ^h ak ₄	duk ₄	rit ₄							
Tikhak	n ^h ak ₄									
Yongkuk	n ^h ak ₄	dok	riek	sim ₂	rum ₃	hum				

⁵ See section 5.6.2 for **ham* as a basic colour term.

	*k ^h a	*duk ₄	*rieC ₄	*sim	*rum ₃	*ham	*htaun ₂	*tʃaŋ		*ro ₁
Lungkhi	k ^h a					gam			mut, pu?	
Rinkhu	k ^h a									
Shangthi	ŋ.k ^h e							tʃaŋ		

In Table 1, there are two modifiers to note that have other meanings cross-varietally. In the case of **ham*, in addition to being a modifier of BLACK, it is also the word for ‘bruise’. Rera is one variety where /ham/ exists on its own with the meaning of ‘bruise’, but did not come up as a modifier. It may still exist as one, but did not come up in conversation and was not actively elicited.

Another is **sim*. As speech communities are developing words for what in English could be glossed as ‘blue’ or ‘purple’, **sim*-like words have been adopted in to cover these meanings in a number of varieties. The presence of a cognate of **sim* in Yongkuk was first proposed and then later rejected by the consultants with whom this was being discussed. It is uncertain if this is actually used in this variety or if the speaker was influenced by other varieties.

Likewise, it is yet to be determined if the presence of this modifier in both Moklum and Joglei is the result of contact. The majority of consulted speakers for these dialects live in Khasan and Kuttom, two villages an easy 5 minute walk from one another on the outskirts of Kharsang, Arunachal Pradesh. Speakers in both villages admit to borrowings elsewhere in the lexicon. This same situation arises in a great many villages in the Kharsang-Miao area where many different language communities have been in close contact for multiple generations.

It must also be noted that Rera /ʃip/ and Shangthi /ʃik/ may or may not be related to **sim* in other varieties. Its inclusion in this column is purely speculative. A similar form exists as a modifier on RED seen in section 5.3 below, and given the high incidence of crossover between RED and BLACK modifiers, it is possible these are a result of borrowings of modifiers between BLACK and RED.

5.2. White, Bright

5.2.1. Basic colour names

In most varieties WHITE is simply /loŋ₁/ as a single syllable. However, as with other colour terms, Rera has an additional way of handling WHITE. While WHITE alone is /v.lo₁ba/, when modifiers are involved the base colour term changes from /lo/ to /mu/, an alternate term for WHITE which is not used on its own. Thus, instead of **/lo tu tu/*, we find */mu tu tu/* in Rera. This would translate as ‘whiteish’ in English. Given modifiers were */mu tu tu/*, */mu tuu tuu/* and */mu tʃop tʃop/*, though no explanation was discovered. One exception that came up was */lo₁ put₄ put₄/*, likely cognate with the Joglei */loŋ₁ pet₄ pet₄/* for ‘off white’. A cognate of */mu/* is not found elsewhere in the languages surveyed.

Hahcheng likewise has an additional variant: */poŋ₁/*, not included in Table 2. There is a slight difference in meaning between */loŋ₁/* and */poŋ₁/*, the latter having some additional sense of ‘shining’ or ‘pearlescent’. Note that there is another word for ‘shining’ in other dialects, */peŋ/*. Additionally, in Hahcheng there is the word */ə.ha₁/* meaning ‘light/bright’ and can also be used with the ‘lightning’ modifier (see next section).

The tone category of /loŋ₁/ is significant here since /loŋ₃/ means ‘afame’ and may be related to Matisoff’s PTB *ploŋ ‘burn’ and thus cognate with Jingpho /prōŋ/. These may be cognate with PTN *loŋ.

5.2.2. Modifiers

Table 4 gives the modifiers that are found on terms for WHITE.

Table 2 – Modifiers for WHITE in Tangsa-Nocte varieties

	*loŋ ₁	*raŋ	*ljep	*tak	*teiaŋ	*put		*ro
Chamchang	lu ₁	ruu ₂						ro ₁
Chamkok	loŋ	reŋ	lep	tak		put		
Hahcheng	loŋ ₁	reŋ ₂	ljep ₄					
Joglei	loŋ ₁		lep ₄	tak ₄		pet ₄		ro ₁
Mueshaung	lu ₁	ra ₂						ruu ₂
Muŋre	loŋ ₁	re ₂		teʔ ₄				ro ₁
Ngaimong	loŋ ₁	reŋ ₂						
Rera	lo ₁ .ba					put ₄		
	mu						tu, tuu, tʃop	
Shangwal	loŋ ₁	raŋ ₂						
Tikhak	loŋ ₁		lep					
Yongkuk	loŋ ₁	raŋ ₂	ljep	teʔ ₄		pa ₄	raʔ	

Rangsi	voan						ʃoŋ	
--------	------	--	--	--	--	--	-----	--

Rera	mu						tʃop, t ^h y	
------	----	--	--	--	--	--	---------------------------	--

	*poŋ ₁	*raŋ	*ljep	*tak	*teiaŋ	*put		*ro
Moklum	puŋ ₁				teiaŋ ₂	poŋ	tʃaʔ	
Phong	poŋ		lep	tuk ₄	teiaŋ ₂		ɬ	

Column III **ljep* means ‘flash, shine’. Lightning is literally ‘sky flash’, but /loŋ/ alone is generally sufficient for clean metal which is not otherwise catching the light. Matisoff (2003: 338) has reconstructed this as **lyap* in PTB. There are cases where WHITE has the form /loŋ ljep ljep/, defined as describing the glistening of water when fish move near the surface on a sunny day. The use of /ljep/ is perhaps less a modifier – in the sense of others discussed here – as more a case of poetic description. For many modifiers, an original meaning, if it existed at all, is unknown. The form /loŋ ljep ljep/ is attested for this context in a number of Pangwa varieties.

Column V, **pət*, has a clearer meaning of ‘speckled’. Something which is /loŋ.pət.pət/ is mostly another colour, but with flecks of white. This same **pət* modifier has been described as possible occurring with other colour terms. For example, in Yongkuk, /seŋ.pət.pət/ can be used to describe an animal which is mostly one colour but has areas of red, and at least one Yongkuk consultant said it can be used with the other colour roots as well. Button (2011) gives /pəʔ/ for Thado, a nearby Kuki-Chin language, with the meaning ‘white’ or ‘spotted’, which may be cognate.

The bottom rows, showing Moklum and Phong, will be discussed further in section 5.6.3.

5.3. Red, Warm

5.3.1. Basic colour names

The Tangsa term for RED has been tentatively reconstructed as **faŋ₂* for Proto-Tangsa. The full standalone word for RED in Rera is /ɤʃe₂khəm₁/. Both RED and BLACK include /khəm₁/ in isolation. Marrison (1967) gives *kaishan* as the Mueshaung word for RED, *akal* for ‘hot’, and *ashang* for Joglei RED. The Mueshaung form given in Marrison was not repeated in 2015 among Mueshaung speakers in Arunachal Pradesh, however it is possible that this shows a connection to some RED terms in other varieties

5.3.2. Modifiers

The first consultant for Yongkuk provided a number of modifiers that were unlike those seen elsewhere. Presented without hesitation were /ti/, /hu/ and /ro₁/. However, it was made clear that /ro₁/ here functioned the same as **ro₁* in other varieties; but unlike in other varieties, it was only possible to apply it to RED and not other colours. However, Yongkuk speakers who were consulted in another location at a later date accepted it as universal. /seŋ₂ tək/ was additionally given for ‘too red’, and /rum₃/ was rejected for ‘red’ but acceptable for ‘black’.

The /koŋ/ form given for Ngaimong was stated to be rare. It may occur in other varieties as well but was not offered up. Chamkok /put/ and Yongkuk /pʌk₄/ is the same as ‘spotted’ and can work for any basic colour.

In at least some varieties, such as Mueshaung, the forms under **t^haN* have the clear meaning of ‘less red than red’, while consistently across dialects the forms in Column I can be translated as ‘truly red’ when the suffix is not reduplicated, and ‘reddish’ when it is.

Table 3 – Modifiers for RED in Tangsa-Nocte varieties

	* <i>faj</i> ₂	* <i>din</i>	* <i>rjeC</i>	* <i>t^haN</i>	* <i>pən</i>	* <i>rəm</i>	* <i>rɤk</i>	* <i>ʃip</i>	* <i>vət</i>	* <i>put</i>		* <i>ro</i>
Chamchang	ʃɤ ₂	dəŋ ₁		t ^h am ₂	pən	rəm ₂	rɤk					ro ₁
Chamkok	ʒaŋ	din				rum				put	ki	
Hahcheng	seŋ ₂	din ₁	riɛt	t ^h əŋ ₂				ʃip ₄				
Joglei	ʃaŋ ₂	dəŋ ₁										ro ₁
Lungki	ʃa.rek		rek					ʃip			p ^h a	
Lungphi	ʃɔ			t ^h auŋ	pən	rum	ruk		vət		te	ruu ₂
Maitai	ʃa											
Moklum	ʃaŋ ₂			t ^h əŋ ₂					vat		pok	
Mueshaung	ʃɔ ₂	de ₁		t ^h auŋ ₂		rum ₂	ruk					ruu ₂
Mungre	ʃɤ ₂	dəŋ ₁			pən	rəm ₂						ro ₁
Ngaimong	seŋ ₂		riɛt ₄	t ^h əm ₂	pən		rɤk				koŋ ₁	
Phong	saŋ ₂	dik	ret				rɤk	tʃip	wət		t ^h u	
Pinkhu	nen	din										
Rangsi	sa		rjep									ru
Rera	ʃe ₂ .k ^h am		rit					ʃip ₄			ʃu	
Shangthi	ʃa		rjet									
Shangwal	ʃaŋ ₂											
Yongkuk	seŋ ₂						rɤk	sip		pət ₄		ro ₁

5.4. Green, Cool

5.4.1. Basic colour names

Note that throughout Tangsa varieties, the realisation of /β/ can be [β], [v], [ʋ] or [w], and the realisation of /ʒ/ can be [j], [ʒ] or, less frequently, [dʒ]. Thus Mueshaung /βul₁ ʒut₄/ is likely cognate with Ngaimong /βil₁ teit₄/, especially given a clear correspondence between nuclear /u/ in Mueshaung with nuclear /i/ in Ngaimong (compare βul₁ and βil₁ for GREEN). There are instances where a given lexeme will occur with only a single and consistent realisation of onset, but these seem to be a minority.

5.4.2. Modifiers

Table 4 – Modifiers for GREEN in Tangsa-Nocte varieties

	<i>*BəL₁</i>	<i>*tʃit₄</i>	<i>*diŋ₃</i>	<i>*suk₄</i>	<i>*soŋ</i>	<i>*teŋ</i>	<i>*fu</i>	<i>*lik₄</i>		<i>*ro₁</i>
Chamkok	vəŋ					t ^h aŋ			re, jaŋ	
Hahcheng	pin ₁									
Joglei	pil ₁	teit ₄	diŋ ₃							ro ₁
Chamchang	βai	teət ₄								ro ₁
Lungphi	vul ₁	ʃit	diŋ					lik		ruu
Lunkhi	v ^w li [?]						ʃu		ki [?]	
Maitai	pai [?]									
Moklum	pil / pul ₁	tʃit ₄	diŋ ₃							
Mueshaung	βul ₁	ʒut ₄								ruu ₂
Mungre	βar	teət ₄								ro ₁
Ngaimong	βil ₁	teit ₄								
Phong	vi ₁		diŋ ₃	suk ₄		teŋ		lik/liŋ		
Pinkhu	vin								sam	
Rangsi	vin			suk	soŋ				nam	
Rera	βim ₁	tʃit ₄		ʃok ₄		tʃaŋ				
Shangthi	vin						ʃɔ			
Shangwal	pil ₁		diŋ							
Tikhak	pəi [?]	tʃit ₄								
Yongkuk	pəi [?] ₁	tʃit ₄								

	<i>*diŋ</i>	<i>*tʃit₄</i>	<i>*diŋ₃</i>	<i>*suk₄</i>	<i>*soŋ</i>	<i>*teŋ</i>	<i>*fu</i>	<i>*lik₄</i>		<i>ro₁</i>
Hahcheng	diŋ									
Wancho	diŋ									

Chamkok modifier /t^haŋ/ was described as having some meaning of ‘glossy’ in the sense of vegetation. It was unclear if **teŋ* takes this meaning more generally. Rera /tʃaŋ/ also

occurs with GREEN, and /βin.tʃaŋ.tʃaŋ/ was described by one consultant as ‘pure deep green like plants’. A possible cognate exists in Phong as well: /teaŋ/, which occurs on a number of other colours including WHITE but not BLACK.

Column II /diŋ/ is of particular note for its potential to provide a possible origin for these terms. In Wancho, a related Northern Naga language group, the basic term for GREEN is /diŋ/. This was also the basic colour term given by a Hahcheng speaker in Myanmar, rather than /pin₁/ given by Hahcheng speakers in India. Wancho does not otherwise have such modifiers as found in the other languages addressed here. French (1983) reconstructs **criŋ* for ‘grass’, ‘alive’, which is one possible origin.

5.5. Yellow

5.5.1. Basic colour names

There was no word given for YELLOW in Rera by one consultant who was 99 years old. This was also claimed by an elderly (88 years old) speaker of Haidley and Hahcheng, who said no such word exists to his knowledge. In Rera, the term for YELLOW was instead described as ‘turmeric colour’, though the word for ‘turmeric’ was later identified as a loan, likely from Singpho. Joglei and Muklom speakers also pointed to a kind of orchid of which the bark is used as colouring. The term for turmeric is /k^ho.men/, and for the orchid it is /min/. The term /k^ho.men/ in Tangsa-Nocte is a loan, but the second syllable /men/ is also identified as meaning YELLOW in some varieties.

PTB **min* ‘ripe’ may be a candidate for the origin of the term. Reflexes of **min* are found in Jingpho as /myîn/, Mizo as /hmin/, Karbi (formerly Mikir) as /men/, Lahu as /mɛ/ and Tangsa as /min₂/. An alternative, and I believe a more likely one, is that the colour term is borrowed from Singpho/Jingpho.

Yongkuk and Tikhak have a different word for YELLOW, /ʒuŋ₁/. It is more widely used even among older people as compared to **men* in other varieties. However, /ʒuŋ₁/ does not take modifier suffixes. There is no cognate of **men* in Yongkuk and Tikhak. One was briefly provided by a consultant, but then quickly rejected by others. The distinction was stated by the first consultant that /ʒuŋ₁/ was a little bit yellow, but truly yellow was /men/, however other consultants believed the first was simply mistaken, influenced by other varieties in which /men/ is found. No modifiers were given by this consultant for /men/, and for /ʒuŋ₁/ there was only one, /hut/, not found in other varieties.

5.5.2. Modifiers

YELLOW has fewer modifiers than other basic terms. This is most likely due to its relative newness among speakers. I had been told by numerous informants that in Rera there are no modifiers for YELLOW. Some varieties lack modifiers for YELLOW, or they may exist but were not known to those consulted. They are however found in other varieties.

Yongkuk, despite having a long-used term for YELLOW, also lacks many modifiers. In numerous conversations with a range of consultants only one modifying suffix was attested. However, **put* which exists as a modifier for WHITE was later shown to be applicable to RED in an un-elicited utterance, after which point consultants said it could apply to “most colours”. There was disagreement as to whether this was extended to YELLOW.

Table 5 – Modifiers for YELLOW **mjen*₂ in Tangsa-Nocte varieties

	* <i>mjen</i>	* <i>t^ho</i>	* <i>t^həŋ</i>	* <i>t^hi</i>	* <i>tV_k</i>	* <i>lik</i>	* <i>teəŋ</i>	* <i>naʔ</i>	* <i>ʒu</i>	* <i>ʒeŋ</i> ₁	* <i>kuɹi</i> ₁	* <i>ro</i>
Chamkok	men							naʔ				
Joglei	men ₂		t ^h əŋ ₁	t ^h i								ro ₁
Kimsing	min	tʁu										ro ₁
Lungphi	mjen	t ^h o		t ^h i	tək	lik						ruu ₂
Lunkhi	men								ʒu			
Moklum	min ₂		t ^h əŋ	t ^h i	tak ₄		teəŋ					
Mueshaung	men ₂	t ^h ɔ ₁				luək ₄						ruu ₂
Mungre	men			t ^h e								ro ₁
Ngaimong	m ⁱ ɛn									ʒeŋ ₁		
Phong	men ₂				dik		teəŋ				kuɹi ₁	

The Rera term [nò.dí] was provided by a Rera informant in Phulbari, India as the term for YELLOW, however with uncertain origins. It was later explained by a Shangthi informant in Hkamti Town, Myanmar, that this in fact means ‘baby’s faeces’. This has yet to be verified. Regardless of the origin, the term does permit modification in Rera through reduplication of the second syllable.

Chamkok has /men/ but not with modifiers, at least according to my Chamkok consultant from Pangsau.

5.6. BLUE, GREY, PURPLE and everything else

For BLUE and GREY there is much more variability in terms used. The following are exceptions which should not be considered basic colour terms at this point in time but which may be headed in that direction.

There were a number of cases where a single word was given for BLUE, but no modifiers existed in that variety. One example is in Pinkhu, for which the term was given as /mɔ/. This was provided by a speaker in Hkamti, Myanmar, but one who was not comfortable giving much more. This is a common occurrence with younger speakers who may feel their speech is “inauthentic”. It is possible this /mɔ/ relates back to an alternate form of WHITE given by a Rera speaker, /mu/. Most Rera migrated to India some time in the past century, and those that remained in Myanmar are said to have joined the communities that are now called Pinkhu.

5.6.1. **k^huai*₂

In a number of dialects, **k^huai*₂ has been given for both BLUE and GREY, as well as in some varieties as a non-basic term for a mixture of black, white, tan. In those cases, such as in the Tikhak/Yongkuk dialects where the term is more properly translated as ‘tan’, it does not take

modifiers. In Wancho, a related language in the Northern Naga family, the term is also present with this sense of ‘grey/tan’.

Modifiers are few in number and appear in only a fraction of the varieties which have this colour term.

Table 6 – Modifiers on **k^huai₂* in Tangsa-Nocte

	<i>*k^huai₂</i>						
Joglei	k ^h oi ₂		mok ₄	dun ₂			ro ₁
Moklum	k ^h i				boŋ	tsək/tsəŋ	
Mueshaung	k ^h oi ₂	t ^h ok ₄					ruu ₂
Wancho	k ^h uai						
Yongkuk	k ^h uai ₂						

5.6.2. **ham*

There is a reconstructed **ha:ŋ* BLACK for PTB⁶, but any connection to modern Tangsa-Nocte is unclear. In many varieties, *ham* does come up in the word for ‘bruise’, and at least among the Joglei speakers I had discussed this research with, they were in general agreement that **ham* is the colour of a bruise.

Table 7 – Modifiers on **ham* in Tangsa-Nocte

	<i>*ham</i>	<i>*tat</i>	<i>*diŋ</i>	<i>*lik</i>	<i>*baK</i>	teəŋ		<i>*ro₁</i>
Joglei	ham ₃	tʌt ₄	diŋ ₁					ro ₁
Moklum	hʌm	tat	diŋ	lik			təŋ	
Mueshaung	ɣəm ₂				baʔ ₄			ruu ₂
Phong	ham			lik/liŋ	bəŋ ₂	teəŋ		
Yongkuk	ham			lik/liŋ	bəŋ ₂	teəŋ		

However, *ham* has proven to be somewhat controversial; **ham* exists in Yongkuk as a deep blue, being accepted by speakers in one community but rejected by those in another in which it exists only as a modifier for **njak*. In the community where **ham* is acceptable, lighter blues will be **pil*, while **ham* appears to be limited to only a narrow range of darker blue hues, but not those dark enough to be identified as **njak*.

It is unclear if /təŋ/ here should be the same as /təŋ/ for other colours. There was some difficulty in eliciting clear responses.

⁶ STEDT #315 (<http://stedt.berkeley.edu>)

Table 8 – Modifiers on *sim₁

	*sim ₁	*lik
Ngaimong	ʃim ₁	
Moklum	sim ₁	lik ₄

An additional colour term *sim was provided for Moklum and Ngaimong. Only one modifier was found for reflexes of *sim in Moklum. It is worth noting that the Ngaimong consultant was originally from a village which is mostly inhabited by Moklum speakers. According to this informant, Ngaimong has a similar word to Moklum /sim/, but no modifiers have yet been elicited. However, *sim does have good representation in other sub-groups, and links back to PTB, reconstructed as *syim₄ for PTB (Matisoff 2003) and present in Bodo-Garo as /sim/ and Jingpho as /sin/, both with the meanings of ‘blue’ ‘dark’ or ‘black’. This is the most likely origin of the term in Tangsa.

The /lik/ modifier for *sim appears in Phong for *ham, and for Mueshaung YELLOW as well as being used for Phong ‘green’.

5.6.3. *poŋ

Ngaimong has an additional word, /əpoŋ₂/. More generally /la₃poŋ₂/ means ‘ashes’, which is likely the source of the colour term and not the other way around. There is a modifier suffix for /əpoŋ₂/, /reŋ₂/, which functions the same as the other modifiers. In Phong, the same word ‘ashes’ /boŋ/ is used as a modifier on ‘black’ as well as having a set of modifiers, seen in Table 4.6, and as a modifier for *k^huai₂ in Moklum. While this might possibly be related to Tani *puŋ ‘white’ (Sun 1993), such a connection has yet to be shown.

Table 9 – Modifiers on *poŋ₂

	*poŋ ₂	*reŋ ₂	*lep	*teaŋ	*tuk ₄	*tɣ	*t ^h eŋ ₁
Ngaimong	poŋ ₂	reŋ ₂					
Moklum	pəŋ ₂			teaŋ	tuk ₄	tɣ	t ^h eŋ ₁
Phong	pəŋ ₂		lep	teaŋ	tuk ₄	tɣ	

Another issue to be resolved is that in Moklum, /puŋ₁/ has been given for ‘white’ with a modifier of /pəŋ/, and different Phong speaking consultants have given different meanings and modifiers for /pəŋ/. The Phong term /pəŋ₂/ in the table above has been more generally treated as WHITE.

6. Conclusion

There is one more colour term which is worth bringing up. In Joglei, /dʒom.tʃ^ha/ is a term which has only come up in discussion with Joglei speakers. It is a term used for PURPLE, specifically the colour appearing on the ubiquitous *longyi*, a plaid-patterned garment worn by older Tangsa men. It is not a basic colour term in any real sense, as /tʃ^ha/ is understood to

mean ‘to paint’ and /dzom/ refers to an insect which produces a waxy substance used as a purple pigment.

Of course /dzom.tʰa/ fails to meet the criteria of what constitutes a basic colour term, as it is not typically used outside of the context of the colour of the longyi, it’s primarily only used by older women and many men do not provide it in elicitation, and the component parts are easily identified by speakers. However, it is generally accepted as being a “real” term in the language by Joglei speakers, when it is presented. This helps to illustrate a point: things are changing rapidly for speakers of Tangsa-Nocte. The high degree of proficiency in either Hindi, Assamese or Burmese by many speakers and the rapid modernisation of the regions where the languages are spoken, due in part both to development in the more remote regions as well as large scale migration to places like Kharsang and Hkamti, resulting in continued contact between various speech communities, means that we may see an increase in both basic colour terms and standardisation of these terms cross-varietally. This is already happening with **min*, which now is nearly universal among the younger speakers of Tangsa-Nocte consulted for this research.

From the discussions had with speakers and the data shown above, the following conclusions can be drawn:

1. It has been previously suggested that this AB/ABB pattern of modifiers originated from the traditional poetic Wihu song structure of the Rangpang group of Tangsa varieties. This is unlikely since it is found in a wide range of Tangsa and Nocte varieties, and not just those few which have the Wihu tradition. While certain affixes may have originated in a single song, many others clearly exist outside the tradition.

2. Additionally, the forms are claimed by some speakers to occur in much smaller number among the varieties lacking Wihu song, such as the Yongkuk-Tikhak group, though this has so far been unconfirmed. It may be the case then that the song traditions have helped maintain these patterns of modifiers.

3. There is enough consistency in terms between varieties to suggest that these are not simply loans into other varieties — as some consultants have suggested — but are rather an older system of which individual parts have fallen out of use. A larger set of data is required to work out the correspondences.

4. Gaps in the tables above, where a given suffix is otherwise well attested, are thus likely the result of loss. This was made clear by a number of speakers who stated that a given suffix existed, but now is not used by that community (though others are). One such example was /ʃaŋ₂ thaŋ₂/ as a modified form of RED in Moklum, which in at least one village is no longer used.

5. In some clear cases, such as **lje*₄ ‘glistening’ or **poŋ*₂ ‘ashen’, modifiers are derived from other lexemes with meanings that are transparent to the speech community. It is also likely that other modifiers once had independent meanings, such as /pət/ for ‘speckled’; or that they still have these meanings but speakers were unable to convey them to the author due to the barriers inherent when relying on the contact language.

6. The colour system of a century ago, as presented by the centenarian consultants, consisted of at most four categories which may be labelled as DARK/BLACK, LIGHT/WHITE, WARM/RED, and COOL/GREEN. It should be no surprise that green and blue typically fall into the same category.

7. Future Research

The current study has been limited to the subject of colour. This is not to say that that is the end of the discussion. The type of modifiers discussed above are also widespread beyond just the basic colour terms.

It is the author's hope that this brief study has laid the groundwork for a much more detailed investigation into both basic colour terms and modifiers in Tangsa-Nocte varieties. In doing so, it is my hope that some of the limiting factors of this first round of investigation can be better navigated at a future date.

One major difficulty in this research was the elicitation context. It was not uncommon that a speaker would say with confidence that they had only a single modifier for a given colour, but when prompted with the form from another dialect, the consultant would immediately recall their own cognate form. When possible, investigation for a given variety was done over the course of days involving many consultants working together to brainstorm in the hopes that it would provide a fuller and more consistent inventory. This of course did not always work, and there were many instances where only a single modifier was given and claimed to exist. See for example Table 4.3 for which it was stated repeatedly that Hahcheng does not have these words. However, over the course of multiple visits quite a few were provided which show clear cognacy to other varieties.

It is thus my hope to return to the topic in during a future trip to the area in order to refine and improve upon the data and analysis above.

Gaps need to be filled in the above tables where possible. The likelihood that a given consultant recalled all of the possible forms actually in use in the language is small, and consultants were generally not helped along with a list of choices. Transcriptions are also approximate and will need refinement in order to better determine the likelihood of cognacy between terms, as not enough time was available with any of the above communities to accurately determine phonemic inventories, something I hope to see resolved in the coming years.

Abbreviations

NOMZ nominaliser

References

- Berlin, Brent and Paul Kay. (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley, University of California Press.
- Button, Christopher (2011). *Proto-Northern Chin*. STEDT Monograph 10. Berkeley, University of California.
- Kay, Paul (1975). "Synchronic Variability and Diachronic Change in Basic Color Terms." *Language in Society* 4: 257–270.
- Kay, Paul, Brent Berlin and William Merrifield. (1991). "Biocultural implications of systems of color naming." *Journal of Linguistic Anthropology* 1(1): 12–25.
- Kay, Paul and Chad K. McDaniel (1978). "The Linguistic Significance of the Meanings of Basic Color Terms." *Language* 54(3): 610–646.
- Matisoff, James A. (2003). *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction*. Berkeley, University of California Press.
- Marrison, Geoffrey E. (1967). *The Classification of the Naga Languages of North-East India*. PhD Dissertation. Department of Phonetics and Linguistics. London, School of Oriental and African Studies.
- Sun, Jackson Tianshin. (1993). *A historical-comparative study of the Tani (Mirish) branch in Tibeto-Burman*. PhD Dissertation. Berkeley, University of California.

Language attitude among Noctes of Bordumsa under Changlang district of Arunachal Pradesh

Trisha Wangno
Tezpur University

Abstract

This paper studies the attitude of the Nocte people of Bordumsa in Changlang district, Arunachal Pradesh, India, towards their mother tongue (Nocte) as well as towards other languages spoken in the area. Assamese is retained as a market language and Hindi has become the language of the educated' in other words a 'prestige language'.

A survey on attitude towards mother tongue and Assamese / Hindi was conducted during my field work. In this survey 40 people were interviewed based on different gender, age, occupation and educational qualifications. The methods used for the interview were: personal interviews, telephone conversations and group discussions. The consultants were asked a series of questions concerning their language, culture, traditions and also their views regarding other languages used in and around them. The use of language was examined through seven domains: home, market, educational institutions, religion, government offices, friends'peer groups and strangers. When it came to offering an opinion, the women were reluctant and usually mimicked what their men-folk had to say. Another interesting fact observed was that when the speakers use their mother tongue in public, it is to maintain secrecy from outsiders. In a way mother tongue serves as a form of code language.

Citation	Wangno, Trisha. 2018. <i>Language attitude among Noctes of Bordumsa under Changlang district of Arunachal Pradesh</i> . North East Indian Linguistics (NEIL) 8, 65-73. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.
Volume Editors	Linda Konnerth, Stephen Morey, Amos Teo
Copyright	© 2018, the author(s), release under Creative Commons Attribution license
URL	http://hdl.handle.net/1885/155280

1. Introduction

Language attitude refers to the views of the speakers towards their mother tongue. This is one of the parameters in the study of languages especially the endangered languages. Baker (1988: 112) states that "Attitudes are crucial in language growth or decay, restoration or destruction". He further adds that "the status and importance of a language in society and within an individual is derived largely from adopted or learnt attitudes" which suggests that the sustainability of an endangered language depends on the active use and positive attitude of the speakers towards their mother tongue.

In the case of a multi-lingual nation, it is very difficult to adopt a language policy which is suitable for all the linguistic communities present within its territory as is generally seen in India. According to the National Policy Resolution 1968 formulated by the Ministry of Indian Government a three-language formula has to implemented in every government educational institution (www.mhrd.gov.in). According to the three-language formula, instructions to children are to be imparted in English-Hindi-Local language; a local language which is a major language or the state language of the concerned states. India has 22 languages Scheduled Languages and the rest of the languages are grouped under non-scheduled languages of India. The languages whose population fall under 10,000 are not given the status of a language. In the state of Arunachal Pradesh Nocte is considered a major language with a population of 33,680 along with Nyishi, Apatani Monpa, Wancho, Tangsa, Mishmi and others. Thus it becomes harder for the state to choose a local major language. The present study is of the language attitude of the speakers of Nocte code ISO 639-3 njb (Ethnologue).

Under the UNESCO list of endangered languages list, the degree of endangerment of Nocte is listed as Vulnerable. The field of study for this research are the Nocte villages under Changlang district of Arunachal Pradesh.

1.1 Nocte

Nocte is a Tibeto-Burman language, generally spoken in the Tirap and some parts of Changlang districts of Arunachal Pradesh. Burling (1983, 2003) classified Nocte together with a group of other Naga languages (the Northern Naga) which is closely related to Bodo-Garo languages and also to Jinghpaw. This group is also referred to as Sal languages. Benedict (1977) groups them as Bodo-Konyak-Jinghpaw. Dutta (1978) stated that Nocte has nine dialects but the dialects/varieties seem to be regional as most are of them are intelligible to the speakers. It is also observed that when two Nocte speakers converse with each other, they speak in their own varieties. Nocte and Tangsa share many similarities, one of them is the presence of tones. Weidert (1979) had stated three tones and Rahman (2016: 12-16) also finds three tones on open syllables in the Haʔwa Nocte variety.

According to the 2011 census published by the Government of Arunachal Pradesh the Nocte population is around 33,680. The literacy rate (in Assamese, English and Hindi) is 47.2% (male 60% and female 34.6%). Table 1 shows the Nocte varieties as listed in Dutta (1978):

Table 1 – Nocte varieties

SL No	Nocte varieties	Villages
1.	<i>Hakhun</i>	Khonsa, Kheti, Thinsa
2.	<i>Khapa</i>	Noksa, Tupi, Polung
3.	<i>Haʔwa</i>	Borduria, Paniduria, Namsang
4.	<i>Dadom</i>	Dadom, Chingkoi, Muktowa, Lahu, Kothin, Bera, Kapu, Hoakan, Rajanagar
5.	<i>Chaniyak</i>	Kolagaon
6.	<i>Jope</i>	Tut, Nokna, Thingja, Kothong, Lianwang, Khatang, Kolam, Lamsa, Lamlo, Bunting
7.	<i>Phonthing</i>	Soha, Turet, Kenoi, Mapuiya, Dongroan
8.	<i>Domlak</i>	Doidam
9.	<i>Laju</i>	Laju, Noklo, Sinham, Sannu, Rahu Longliang, Ponkong, Senliyam, Longbu

1.1. The Language Scenario

After India's Independence in 1947, Arunachal Pradesh was known as North-Eastern Frontier Agency (NEFA). It was placed under the administration of Assam until 1972. Assamese was the dominant language in Arunachal because of its long-standing historical relations from the Ahom ages, Assam's huge population, literature and the use of Assamese in the market areas. In accordance with the three-language policy of the Indian State, Assamese was the medium of instruction in educational institutions. After 1972, the Union Territory of Arunachal

Pradesh, having been divided from the greater Assam, gradually saw change in the language scenario. Since Arunachal Pradesh does not have a dominant language of its own Hindi replaced Assamese as the medium of instruction in all educational institutions and it also became the language to be used in government offices. Even though English was made the official language of Arunachal Pradesh, Hindi was the *lingua-franca*. Assamese lost its status in Arunachal once replaced by English and Hindi. However, Assamese is still preferred by the older generations while the younger generation do not consider it as prestigious as Hindi and English. However, in the Southern districts of Arunachal Pradesh and other districts sharing boundaries with Assam, Assamese is still the only medium of communication.

1.2. Migration and the present area; Bordumsa

In 1976 with the approval of the Government of Arunachal Pradesh, a large group of people numbering approximately sixty households migrated from Dadom village in Khonsa, Tirap District to the plain areas of Bordumsa in Changlang district, Arunachal Pradesh. According to the elders of the community the main reason for migration was economy. The hilly and arid topography of Tirap was unable to sustain an ever growing population which led to migration to fertile plains areas. According to the official the present population of the three Nocte villages namely; Rajanagar I, Rajanagar II and Rajanagar III is 1,200¹.

Bordumsa town (27.51821° N, 95.88079° E) is a border area between Arunachal Pradesh and Assam. Originally Bordumsa was a Singpho settlement, the word 'bor' is Assamese for big and 'dumsa' is the name of a Singpho clan and also a traditional Singpho priest. The Bordumsa area falls under the jurisdiction of both the states which is why Assamese is still the dominant language there; therefore, most of the Arunachalis living here communicate in Assamese. Other major communities apart from Nocte and Singpho are the Tangsa: Ponthai, Mossang and Kimsing (Morey 2017), Khampati and recently the Chakmas migrating from the Chittagong hill tracts of Bangladesh. Even though Hindi is the *lingua franca* of the state of Arunachal Pradesh, Bordumsa is an Assamese-speaking region. With these many ethnic communities in and around the Nocte community has become divided in its opinions. Some people have a developing sense of fear and insecurity for their own identity while some others believe that the development of the community is only through the learning and the use of the job providing languages, which in this case are Hindi and English.

2. Language Attitude

For an endangered and small language community such as Nocte to survive, it is crucial that the speakers have a positive attitude towards their language. But due to many economic and logistical problems which include under-representation of the community in the government institutions as well as the education sector, a minority language speaker tends to view their language as not so important and not prestigious enough. The attitudes towards one's mother tongue vary according to each native speaker's individual experience. The present study will try to focus and determine the factors for the attitude of the Nocte speakers of Bordumsa of Arunachal Pradesh towards their mother tongue.

2.1. Initial findings

In this study the interview and questionnaire method was mostly used. But since most consultants would rather talk than fill up the questionnaire, the researcher had to do it while

¹ The Office of the Additional Deputy Commissioner Bordumsa has records of the population of every village and the number of voters.

still engaging in conversations with them. Apart from this, unplanned and unstructured group discussions and silent observations were also conducted. Following these methods several observations were made, which are:

- The Noctes have divided themselves into two groups: The Upper Noctes and the Lower Noctes. The Noctes of Bordumsa come from Dadom and Chingkoi villages which is the Upper Nocte group.
- The Upper Noctes are believed to have originated from the Wanchos (another language community of Tirap and Longding), which is why Nocte spoken by them is not considered ‘perfect’ or totally acceptable by the Lower Noctes. The Lower Noctes consider the language form of Upper Noctes as having been acquired over some period of time. This form is seen as rustic and different than the one spoken by them i.e. ‘real Noctes’
- Most of the times it was seen that there was a slight hesitation to give language data among the Upper Nocte group and instead the researcher was directed to go the Lower Nocte areas like Borduria to “learn the language in its true/best form”. On forcing to provide data, there was constant comparison with ‘what they say’ and ‘how we say’. This kind of reaction have been observed among the migrated Noctes of Bordumsa, this could have been probably due to feeling of being alienated from the bigger Nocte group.
- The most commonly heard phrase was; it is very easy to learn Nocte as there are very few words, meaning it is easy for a non-Nocte to pick-up the language who might want to learn the language. This statement is contradictory as none of the other communities have learned Nocte and neither have the Noctes of Bordumsa learned other neighbouring languages like Singpho or Tangsa. The reluctance to provide language data and the fear of being ‘proven wrong’ by ‘other Noctes’ thus is very visible, the clear statement is: learn from hearing us speak rather than asking which also indicates the insecurity of the speakers themselves.

The above statements indicate the pressure of the Bordumsa Noctes to be identified as a single group with the Noctes of other districts while still maintaining their differences. What is also interesting is that these ‘narratives’ have been passed to the younger generations as well. Edwards (1982) describes this kind of phenomenon as “Group solidarity and social status as one dimension along which views towards a language may vary. In-group solidarity or language loyalty, reflects the social pressures to maintain languages/language varieties, even one without social prestige”

2.2. Interview Participants

A field study was conducted in the three villages of Rajanagar in Bordumsa to determine the attitudes of the people towards their language. Forty (40) native speakers of different age group, qualifications and gender were interviewed. There were 13 Females and 27 Males ranging from the age-group 15 to 60. Table 2 shows the different age groups and the number of participants of each group while Table 3 shows their educational qualifications so to show the diversity of the participants and their move-ability mostly in case of men due to pursuing higher education, business purpose etc. Table 4 shows their occupations.

Table 2 – Different age groups and the number of people

Age-group	15 - 24	25 - 34	35 - 44	45 – 59	60 and above
No of Male	10	05	06	04	02
No of Female	05	03	03	02	00

Table 3 – Educational qualifications of the participants

Qualifications	No of people
Uneducated	7
Studied till high school	14
College students	5
School students	10
Graduates	4

Table 4 – Various occupations of the participants

Occupation	No of people
Involved in Agriculture	4
Defence personnel	9
Housewives	3
Business	8
Students	13
Teachers	2
Bank clerk	1

2.3. The Questionnaire

Keeping in mind the varied backgrounds of the speakers shown in the above tables, the speakers were asked to state their preferred language in the following domains given below:

Question 1. What language they regularly use in the following domains:

- Home
 - Market
 - School or Colleges
 - Church
 - Government Offices
 - Among friends
 - With strangers
-

To answer this question, a chart has been listed and is indicated in the Figure 1 given below:

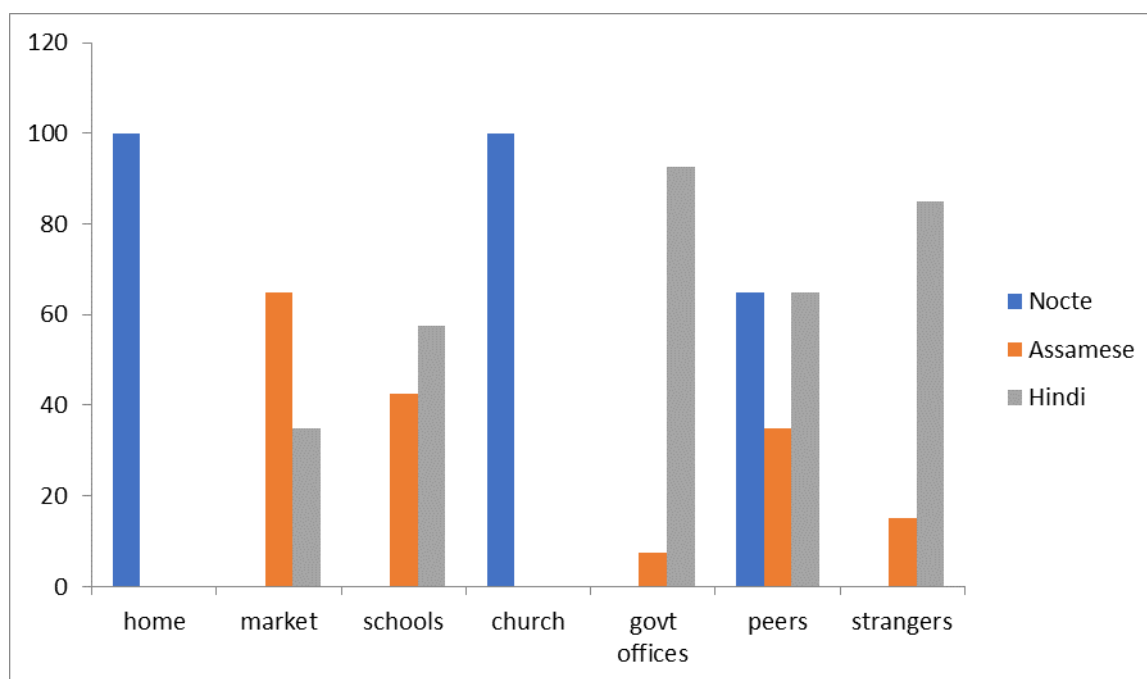


Figure 1 – Preference of language in various domains

In this chart, it is clearly indicated that native speakers prefer to use their mother tongue in the home domain and their comfort zones; speaking with their peers or while praying, shown here by church. The church is an interesting domain as observed by Morey (2017) in Tangsa language context where multiple languages are used. In the case of Bordumsa Nocte, all the speakers said that they were comfortable with Nocte but it was observed that in most churches the Bible reading was done in Hindi, hymns were sung mostly in Hindi sometimes in Assamese/Nagamese and English but rarely in Nocte. The Catholic church would sometimes use the English to Nocte translated hymns and prayer books. The Nocte bible published by the Bible Society of India is in circulation, but they have not been in use in the churches yet. The language used to deliver sermons would depend on the priest whether he/she was a Nocte speaker or an outsider. Mass prayers would be left to the individuals' choice of language and Assamese sermons would be used only in the presence of other non-Nocte speakers.

2.3.1. Additional Observations

While the questionnaire determined the use of different languages in different domains according to the interview participants, several other observations were made:

- It was noted that young people below the age of 35 tended to speak in Hindi among themselves whenever they had to explain topics like politics, diseases, technological information etc.
- Speaking grammatically correct Hindi was considered as a sign of being educated and knowledgeable whereas speaking grammatically correct Assamese (by Nocte speakers) was usually mimicked and made fun of.
- A significant number of people happen to be employed in the military services and are stationed across India, where knowledge of Hindi is the unspoken compulsory requirement/criteria. Therefore, some of the personnel's children learn Hindi as their first language.
- The medium of instructions in 'English' schools and colleges is Hindi. A huge number of people in the teaching sector and other government offices are mostly Hindi speakers from Northern India further contribute in high level of Hindi usage.
- Sanskrit as the third language in school curriculum has replaced Assamese. This has resulted in school students being unable to read in Assamese thus widening the gap of illiteracy between the older and younger generations. The older generation who had some knowledge of Assamese were considered literate but the situation changed in recent times.
- Growing popularity of Hindi serials among women, movies and news agencies in general show the preference for this popular culture.

Question 2. How many languages do you speak?

Out of the 40 native speakers, there was one monolingual, who only spoke Nocte, but the person did understand Assamese enough to communicate or gesture while four spoke both Nocte and Assamese and 35 people used Nocte, Assamese and Hindi on a regular basis. Thus, it can safely be said that the Nocte community in Bordumsa is a multi-lingual one. As shown in Figure 1 above, the preference of different languages in certain domains lies not just in the speakers' choice but also in the feasibility: i.e. the presence of other speech communities.

3. The Interviewees

To test the speakers and record their opinion regarding the preservation of their language few questions were prepared. The answers mostly varied depending on the age group of the speakers. The women did not have separate opinion and usually agreed with men.

Q 1. Are you a fluent speaker of Nocte?

The speakers above the age group of 45 replied in affirmation whereas the speakers below the age of 45 replied in negative. The younger generation seemed not so confident. The terms "pure" and "original" was used many times. Most statements would be like "we do not know pure Nocte". "The language we use is not so pure", "original Nocte is spoken by Borduria people".

Q 2. Will you teach your children Nocte?

To this question every age-group said yes. “The reality on the ground, however, was different.” There were several young people who had been married to people from other communities and their children spoke in Hindi or Assamese. In some cases, children spoke to their grandparents in Hindi even though both set of parents were Noctes, this was mostly the case with families who lived outside the village/community mostly the ones from military personnel. When it was pointed out to these families that their children were not speaking in their mother tongue, the parents casually replied that ‘they (the children) would learn when they grow up’.

Q 3. Should Nocte be taught in schools? (in the future)

This question had varied responses. People above the age of 45 could not imagine the possibility of Nocte being taught in schools. They mostly scoffed and questioned back asking ‘who will learn Nocte in schools?’ They would also say why should Nocte be learned in the schools and why can’t they learn at home. However, the speakers between the age-group of 44 to 25 questioned the economic benefits saying that learning Nocte in schools was unprofitable, as it would not provide jobs. The people below the age of 25 were not sure.

Q 4. Why should we know Nocte?

To this question mostly women from all the age group and also some young people replied ‘so that we can speak in private or secretly when in a hostile situation’ while others said since we call ourselves Nocte, therefore we should know Nocte.

Q 5. Which language is good for jobs?

All the speakers agreed that knowing English was beneficial for jobs. Some people also pointed the fact that knowledge of Hindi was an added bonus, especially for other parts of Arunachal Pradesh.

4. Analysis and Conclusion

From the above presented answers we find that there is a mixed attitude among the speakers of Nocte in Bordumsa area regarding their language. No conclusion can be made about the attitudes of the Noctes from the Tirap district as the field study was limited to Bordumsa. The insecure feeling that one is not speaking the ‘pure or original’ language tends to elevate one regional variety above the others. The younger generation with who lies the burden of preserving the language understand the implications but do not act accordingly. For the young people it is more necessary to maintain prestige and work towards their economic survival than maintain their language. While maintaining identity is considered necessary, speaking the language and teaching young children is not actually being done. Community loyalty is appreciated but not used outside the home domain. The vitality of the language might not be threatened for the present but it will remain questionable in the future.

References

- Baker, Colin (1988). *Key issues in bilingualism and bilingual education*. Clevedon, UK and Philadelphia, Multilingual Matters.
- Benedict, Paul K. (1977). "Where it all began: memories of Robert Shafer and the Sino-Tibetan Linguistics Project." *Linguistics of the Tibeto-Burman Area* 2(1): 81–92.
- Burling, Robbins. (1983). "The Sal Languages." *Linguistics of the Tibeto-Burman Area* 7(2): 1-32.
- _____. (2003). "The Tibeto-Burman languages of Northeastern India." In Graham Thurgood and Randy J. LaPolla, Eds. *The Sino-Tibetan languages*. London, Routledge: 178-182.
- Census of India. (2010-11). <http://censusindia.gov.in>. Accessed on 10th Feb 2016.
- Dutta, Parul Chandra (1978). *The Noctes*. Shillong, Directorate of Research Government of Arunachal Pradesh.
- Edwards John R. (1982). "Language attitudes and their implications among English speakers." In Ellen B. Ryan and Howard Giles, Eds. *Attitudes Toward Language Variation: Social and Applied Contexts*. London, Edward Arnold.
- Morey Stephen. (2017a). "The Sociolinguistic context of the Tangsa languages." In Picus Sizhi Ding and Jamin Pelkey, Eds. *Sociohistorical Linguistics in Southeast Asia: New Horizons for Tibeto-Burman Studies in honor of David Bradley*. Boston and Leiden, Brill: 169-188.
- _____. (2017b). "Tangsa." In Graham Thurgood and Randy LaPolla, Eds. *The Sino-Tibetan Languages* (2nd Edition). London and New York, Routledge: 350-368.
- Rahman, Syed Iftiqar (2016). *Verbal Agreement and Grammatical Description in Ha?wa Nocte*. Masters thesis. Chiangmai, Payap University.
- Weidert, Alfons. (1979). "The Sino-Tibetan Tonogenetic Laryngeal Reconstruction Theory." *Linguistics of the Tibeto-Burman Area* 5(1): 49-127.
- _____. (1987). *Tibeto-Burman tonology: A comparative account*. Amsterdam, John Benjamins.

Sino-Tibetan/Trans-Himalayan morphosyntax

The inclusive-exclusive distinction in Kuki-Chin and Naga Belt Languages

Scott DeLancey
University of Oregon

Abstract

In some Tibeto-Burman languages we find a distinction, in the pronouns and/or in verbal agreement forms, between exclusive and inclusive 1st person plural forms – a category which has come to be known as CLUSIVITY (Filiminova 2005). This distinction is absent in many other languages of the family (LaPolla 2005). Until recently we had very few examples of this distinction in the “Naga” and Kuki-Chin languages of the Indo-Myanmar border region. Recent research has uncovered the exclusive/inclusive distinct in several more of these languages, and correspondences between the forms make it clear that we must reconstruct this distinction for the common ancestor at least of Kuki-Chin and the Ao group, and probably of the other languages of the Naga Belt. This then offers a likely explanation for the innovative 1st person singular pronouns found in most of these languages, as we will see in Section 2. The Kuki-Chin and Naga forms further correspond to forms in Kiranti languages, suggesting that they can be reconstructed for Proto-Trans-Himalayan (Sino-Tibetan). Nevertheless, the similarities among the exclusive, inclusive, and 1st person singular forms in the various Kuki-Chin and Naga Belt languages provide an argument for a Kuki-Naga branch of the family.

Citation DeLancey, Scott 2018. *The inclusive-exclusive distinction in Kuki-Chin and Naga Belt Languages*. North East Indian Linguistics (NEIL) 8, 75-85. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.

Volume Editors Linda Konnerth, Stephen Morey, Amos Teo

Copyright © 2018, the author(s), release under Creative Commons Attribution license

URL <http://hdl.handle.net/1885/155280>

1. Clusivity in Kuki-Chin and Naga Belt languages

Many languages of the world distinguish two categories of 1st person non-singular pronominal reference: inclusive, which includes the addressee as well as the speaker, and exclusive, which includes the speaker and others, but not the addressee. In recent typological work this distinction is referred to as *clusivity* (Filiminova 2005). Until recently this distinction was reported only in very few languages of Northeast India, but recent descriptions show that it is more common than we had thought. Moreover, we can now see that Kuki-Chin and at least some of the Naga languages share the same forms, showing that clusivity must have been a feature of their common ancestor.

In Section 1.1 I will describe the Sino-Tibetan context for the question of clusivity in languages of Northeast India. Section 1.2 will introduce the Kuki-Naga languages which are the main topic of this paper. Section 1.3 will give examples of Kuki-Chin languages which have the distinction; Section 1.4 will present data from Mongsen Ao which shows the same distinction with corresponding forms. In Section 1.5 I will summarize the comparative argument for reconstructing this distinction and these forms to the common ancestor of Kuki-Chin, Ao, and perhaps other languages.

1.1. Clusivity in Sino-Tibetan

The inclusive-exclusive distinction is distributed inconsistently across the Trans-Himalayan (Sino-Tibetan) family: many languages and groups have it, many others do not. Even very closely related languages may differ in whether they distinguish clusivity. For example, the

distinction is found in some (but not all) varieties of Mandarin (Norman 1988: 121), although it is not a characteristic feature of Sinitic.

Where it is attested, the forms which mark it usually do not correspond across even low-level branches. This is illustrated in Table 1 (sources for data in the tables are given at the end of the paper):

Table 1 – Inclusive and exclusive pronouns across Tibeto-Burman

Branch	Language	1SG	EXC	INC
E. Kiranti	Athpare	<i>a-ŋa</i>	<i>an-ni-ŋa</i>	<i>a-ni</i>
W. Kiranti	Khaling	<i>uŋ</i>	<i>ok</i>	<i>ik</i>
Burmish	Zaiwa	<i>ngo³¹</i>	<i>nga³⁵-moq¹</i>	<i>nga³⁵-nvung³⁵</i>
Nungish	Trung (Dulong)	<i>ŋà</i>	<i>ŋà-maʔ</i>	<i>ŋuŋ-maʔ</i>
Tani	Milang	<i>ŋa</i>	<i>ŋa-zi</i>	<i>ŋa-a</i>

It is immediately obvious that the 1SG forms have something in common. Since we know that these languages are related, we assume that the velar nasal shared by all the forms in the different languages is shared inheritance from their common ancestor. And they are; all these forms reflect Proto-Sino-Tibetan **ŋa*, in the Kiranti languages compounded with a possessive proclitic.

So it is particularly striking to see that the plural forms cannot possibly be shared inheritance. Zaiwa *-moq¹* and Trung *-maʔ* look cognate, and the 1st person **ŋa* root is recognizable in several forms, but otherwise there is no apparent correspondence between any two forms in any two languages. And not only are the formatives not cognate, the formation of the pronouns varies. In Athpare the Exclusive is the Inclusive plus an extra element related to the singular pronoun. In Khaling Inclusive and Exclusive are distinct, possibly related forms. In Milang and Zaiwa there are distinct Inclusive and Exclusive plural formants added to a 1st person stem. In Trung the same plural formative is attached to distinct Inclusive and Exclusive stems. (For other examples and extended discussion see LaPolla 2005). This tells us that all (or, hypothetically, all but one) of these clusivity systems must be secondary developments in the individual languages or low-level branches rather than inheritance from the common ancestor of all.

There are two striking exceptions to this generalization. The first is the well-established status of clusivity in the Kiranti languages, which we will see in Section 3.1. The second is the topic of this paper, a shared inclusive/exclusive opposition in Kuki-Chin and Ao marked by #*kai* EXC and #(a)*i* INC, which must represent shared inheritance from their nearest common ancestor. In Section 2 we will see that these are the source of innovative 1st person singular pronouns in almost all of the Kuki-Chin and Naga Belt languages. Finally in Section 3 we will consider the implications of these facts for the genetic classification of the Kuki-Chin and Naga Belt languages with respect to one another and to the rest of the family.

1.2. Kuki-Chin and Naga Belt languages

We are concerned in this paper with two sets of Tibeto-Burman languages: the Kuki-Chin languages of Chin State in Myanmar and Mizoram, southern Manipur, and adjacent parts of India and Bangladesh, and the languages of Nagaland, central and northern Manipur, and adjacent parts of Myanmar which have traditionally been grouped together under the heading “Naga”. At present most linguists do not see strong evidence that the latter set of languages constitute a genetic unit, and so argue against recognizing a Naga branch of the family (Burling 2003, Coupe 2007). Nevertheless, a group of contiguous languages, including all of

Angami, Ao, Tangkhul, Zeme, and Meithei, as well as some Northern Naga (Konyak, Phom, Chang, Wancho) and a few NW Kuki-Chin (Sorbung, Sorte) languages, share a grammatical profile which distinguishes them from the more conservative Northern Naga languages to the north and almost all of Kuki-Chin to the south, and, more importantly, most of them share the same innovative 1st person pronoun, as we will discuss in Section 2. The same grammatical profile is found in some Northern Naga languages, and in some languages in Ukhrul District of Manipur which may be related to Kuki-Chin, so this seems to be an areal rather than a genetic feature. All of these languages occupy a geographically contiguous area; I will refer to this area as the Naga Belt, and the languages in question as Naga Belt languages, with the understanding that this is an areal grouping and not a distinct branch of the family.

I will summarize what we know about clusivity in Kuki-Chin in Section 1.3, in the Naga Belt in Section 1.4, and the historical implications of these facts in Section 1.5. In Section 2 I will suggest that this historical reconstruction offers an explanation for the innovative 1st person pronouns which set all of these languages apart from the rest of the family.

1.3. Inclusive and exclusive in Kuki-Chin

Clusivity is not marked in most Kuki-Chin languages. But Stern (1963) and Henderson (1965) long ago documented the distinction in two closely-related languages, Sizang and Tedim, both belonging to the Northeast (or “Northern Chin”) subbranch of Kuki-Chin. Baruah and Bapui (1996) report clusivity in Hmar, which probably belongs to the Northwest subbranch (called “Old Kuki” in the Linguistic Survey of India). (Some sources list Hmar in the Central subbranch, but based on the available data – in particular these pronominal forms – I consider it a NWKC language which has been heavily influenced by Mizo.). More recent work documents the same forms in the NE language Paite (N.S. Singh 2006) and the NW languages Monsang (Konnerth and Wanglar 2014, DeLancey et al. 2015) and Hrangkhoh (Haokip ms.); the data are summarized in Table 2.

Table 2 – Inclusive and exclusive forms in Kuki-Chin

Branch	Language	1SG	EXC	INC
NEKC	Tedim	<i>kei</i>	<i>ko-mah</i>	<i>ei-mah</i>
	Paite	<i>key</i>		<i>ey</i>
NWKC	Monsang	<i>kʰ</i>	<i>kí-nn^wú</i>	<i>ì-nḡ</i>
	Hrangkhoh	<i>kei</i>	<i>kei-ni</i>	<i>ei-ni</i>
	Hmar	<i>kéi</i>	<i>kéi-ni</i>	<i>ei-ni</i>

The Tedim and NW paradigms correspond closely. They have innovated different plural markers: NE *-ma(?)*, NW *-ni*. But they share *#k-* in both singular and exclusive in contrast with *(e)i* inclusive. Such a close paradigmatic correspondence can only be shared inheritance. We do not see this paradigm in the other subbranches of Kuki-Chin, Central, Southern, or Maraic, which have for the most part lost the inclusive form, although it is preserved as a 1st person proclitic in Mara, and a 2nd person proclitic in Mizo (DeLancey 2013). But NE and NW are independent subbranches, and do not form a larger subgroup within the branch, which means that their nearest common ancestor is Proto-Kuki-Chin. So we must reconstruct exclusive *#kV*, inclusive *#i ~ ei* for Proto-Kuki-Chin, even though the opposition has been lost in the other subbranches. (I use *#* to indicate a provisional representation of a form which must be reconstructed to a proto-language, but has not yet been systematically reconstructed).

1.4. Clusivity in the Naga Belt

We have even less attestation of clusivity in the Naga Belt languages than in Kuki-Chin. This may to some degree simply reflect the fact that we have in-depth descriptions for only a few of these languages, but it seems to be the case that the inclusive-exclusive distinction is rare in this area. But Coupe (2007) describes an exclusive/inclusive distinction in Mongsen Ao, in both dual and plural. The dual forms correspond to the Exclusive $\#kV$, Inclusive $\#(e)i$ which we find in Kuki-Chin, as shown in Table 3.

Table 3 – Mongsen Ao pronouns (Coupe 2007: 89)

	SG	DU	PL
1(EXC)	<i>ni</i>	<i>kə-nət</i>	<i>ì-la</i>
INC		<i>i-nət</i>	<i>ì-sa</i>
2	<i>nàŋ</i>	<i>nəŋ-ət</i>	<i>nàŋ-la</i>
3	<i>pa</i>	<i>pa-nət</i>	<i>túŋ-la</i>

Interestingly, the plural forms use only *i-* as a pronominal root, and distinguish clusivity in the plural marker rather than in the root.

1.5. Implications for reconstruction

The Exclusive and Inclusive forms discussed in the previous sections are summarized in Table 4.

Table 4 – Summary of Exclusive and Inclusive forms in Kuki-Chin and Ao

	Language	1SG	EXC	INC
NEKC	Tedim	<i>kei</i>	<i>ko-mah</i>	<i>ei-mah</i>
	Paite	<i>key</i>		<i>ey</i>
NWKC	Monsang	<i>kʰs</i>	<i>kí-nn^wú</i>	<i>ì-nə̀</i>
	Hrangkhoh	<i>kei</i>	<i>kei-ni</i>	<i>ei-ni</i>
	Hmar	<i>kéi</i>	<i>kéi-ni</i>	<i>ei-ni</i>
Ao	Mongsen (dual)	<i>ni</i>	<i>kə-nət</i>	<i>i-nət</i>

Since the NE and NW Kuki-Chin languages and Mongsen Ao all have the same distinction, marked by corresponding forms, in the same syntagmatic position (i.e. as possessive proclitics on the noun), the $\#k-$ / $\#(e)i-$ exclusive/inclusive opposition must be reconstructed for the nearest common ancestor of Kuki-Chin and Ao. The next question is, what is that nearest common ancestor?

There is no other evidence for a linkage specifically between Kuki-Chin and Ao, excluding other neighboring languages. But many scholars have suggested that Kuki-Chin and all the “Naga” languages form a subgroup within Tibeto-Burman, largely on grounds of shared lexicon. If we reject the Naga hypothesis and thus do not consider Ao and the other Naga Belt groups to constitute a branch, then we cannot automatically conclude that this common ancestor is also ancestral to the other Naga Belt languages, but it seems likely that the nearest common ancestor of Kuki-Chin and Ao is probably also ancestral to some of the other languages of the region. In Section 2 we will see that most of the other Naga Belt

languages have an innovative 1SG pronoun which appears to be cognate to the Ao (and thus also the Kuki-Chin) inclusive form, and that this can be interpreted as evidence supporting a Kuki-Naga branch, but probably not supporting a genetic Naga grouping within it.

2. Innovative first person pronouns in Kuki-Chin and Naga Belt languages

Most of the Kuki-Chin and Naga Belt languages are distinguished from the rest of the family by their innovative 1st person singular pronoun forms. All of these languages retain the original **naŋ* ‘2SG’ root, but almost all have replaced **ŋa* ‘1SG’ with one of two other forms: Kuki-Chin languages with *#kai*, and the Naga Belt languages, including Meithei, with *#i* ~ *#ai*. (VanBik (2009) reconstructs **kay* for PKC, and Bruhn (2014) reconstructs **aj* for Proto-Ao. I write the offglide in both as a vowel to reflect the origins of the forms: *#ai* is *#i* ‘inclusive’ plus some kind of augment, and *#kai* is the pronominal exclusive root *#ka* plus a contrastive focus marker *#i* < *#ya*.) From a comparative perspective these are two distinct innovations, the former considered one of the characteristic shared innovations that define Kuki-Chin (but see Section 3.2), the latter arguably relevant to the question of the “Naga” hypothesis, that most or all of Angami-Pochury, Ao, Zeme, Tangkhul and possibly Meithei are one another’s closest relatives. In Section 2.1 we will quickly survey some evidence for the Proto-Trans-Himalayan forms. In Section 2.2 we will describe the innovative forms in Kuki-Chin and the Naga Belt. The historical interpretation of these facts will be presented in Section 2.4, showing that the innovative forms are reinterpretations of originally plural exclusive and inclusive forms, and 2.5, where we will present a provisional reconstruction of the original pronominal system from which all of the attested paradigms descend.

2.1. Pronouns in other Tibeto-Burman languages

Innovative pronoun forms are found in scattered languages, including the West Himalayan branch, but the vast majority of the family, including Sinitic, have reflexes of **ŋa* ‘1SG’ and **naŋ* ‘2SG’. Table 5 gives a sampling from across the family.

Table 5 – Reflexes of **ŋa* and **naŋ* in a sample of Tibeto-Burman languages

Branch	Language	1SG	1PL	2SG	2PL
rGyalrongic	Situ	<i>ŋa</i>	<i>jo</i>	<i>no</i>	<i>ŋo</i>
Magaric	Syangja	<i>ŋa</i>	<i>kaŋ-ko</i>	<i>naŋ</i>	<i>nak-ko</i>
Nungish	Rawang	<i>ŋa</i>	<i>ŋa-niŋ</i>	<i>na</i>	<i>na-niŋ</i>
Lolo-Burmese	Burmese	<i>ŋa</i>	<i>ŋa-dui.</i>	<i>naŋ</i>	<i>naŋ-dui.</i>
Tani	Galo	<i>ŋó</i>	<i>ŋu-nù</i>	<i>nó</i>	<i>nu-nù</i>
Sal	Jinghpaw	<i>ŋai</i>	<i>an-the</i>	<i>naŋ</i>	<i>naŋ-the</i>
	Nocte	<i>ŋa</i>	<i>ni</i>	<i>naŋ</i>	<i>ne(-khu)</i>
	Boro	<i>aŋ</i>	<i>zwaŋ</i>	<i>nwaŋ</i>	<i>nwaŋ-swr</i>

In the commonest pattern in this sample and across the family, the plural forms are the singular root plus some plural formative, the same for both persons. (In this sample the exceptions are the rGyalrong and Northern Naga forms and the Bodo-Garo 1PL. Rgyalrong preserves the original pronominal roots; 1PL *jo* is cognate to the *#i* inclusive which we saw in the previous section, and thus to the innovative 1SG form which we will see below.

2.2. Innovative 1SG pronouns in Kuki-Chin and Naga Belt languages

All of the languages of the Kuki-Chin and Naga Belt groups retain the original PST 2nd person pronoun **naŋ* ‘2SG’, but with a small handful of exceptions almost all have replaced the 1SG pronoun. The exceptional cases will be discussed in Section 3.2. The languages show two distinct innovations, one characteristic of Kuki-Chin, and another which is found in most of the Naga Belt languages. We will see that these correspond neatly to the exclusive and inclusive forms which we have seen previously.

In almost all the Kuki-Chin languages, the 1st person pronouns reflect a root which VanBik (2009) reconstructs as **kay*, and which I will write as #*kai*. Table 6 gives a sample of the evidence, including languages from each of the five subbranches of Kuki-Chin.

Table 6 – 1st and 2nd person pronouns in Kuki-Chin

Branch	Language	1SG	1PL.EX	2SG	2PL
NEKC	Sizang	<i>key</i>	<i>kə-w</i>	<i>nəŋ</i>	<i>nə-w</i>
NWKC	Koireng	<i>kəy</i>	<i>kəy-ni</i>	<i>nəŋ</i>	<i>nəŋ-ni</i>
Central	Mizo	<i>kéy</i>	<i>kei-ni</i>	<i>nəŋ</i>	<i>nəŋ-ni</i>
Maraic	Mara	<i>kéi</i>	<i>kéi-mo</i>	<i>na</i>	<i>na-mo</i>
Southern	Hyow	<i>key</i>	<i>key-ni?</i>	<i>naŋ</i>	<i>naŋ-ni?</i>

The history of the various plural formatives is unclear. The most widespread is #*ni*, found in NW, Central, and Southern; this may be reconstructable to PKC, although its absence in the NE subbranch is problematic. The *-w* plural in NE is found more broadly across the family as a postverbal particle #*u* indexing a plural argument; it may have originated as a plural word in PKC, although there are possible correspondences outside of the branch. Likewise Maraic *-mo* is of uncertain origin, but may have connections to plural forms in Kiranti and Jinghpaw.

As we saw in Section 1.2, the languages of northern Manipur and Nagaland belong to five very low-level genetic units, Angami-Pochuri, Zeme, Ao, Tangkhulic and Meithei. At present there is no consensus on any higher order affiliations among these. But most of the languages of all of these groups share the innovation of a new 1st person form related to #*i* or #*ai*, as shown in Table 7.

Table 7 – 1st and 2nd person pronouns in some languages of the Naga Belt

	Language	1SG	1PL	2SG	2PL
Angami-Pochuri	Angami	<i>a</i>	<i>hie-ko</i>	<i>no</i>	<i>na-ko</i>
Zeme	Zeme	<i>i</i>	<i>a-nui</i>	<i>nang</i>	<i>ning-nui</i>
Ao	Yimchungrü	<i>i</i>	<i>i-rü</i>	<i>nü</i>	<i>nü-sa</i>
Tangkhulic	Tangkhul	<i>i</i>	<i>i-thum</i>	<i>nA</i>	<i>nA-thum</i>
Meithei	Meithei	<i>əy</i>	<i>əy-khoy</i>	<i>nəŋ</i>	<i>nə-khoy</i>

Notice that, while there is some correspondence in plural formatives across the Kuki-Chin subbranches, there is none across the Naga Belt. All of the languages have the same structure for plural pronouns, the singular root plus a plural marker, but each low-level group has its own plural form, with no apparent resemblants either in other Naga Belt languages or elsewhere in the family. This suggests that all of these plural pronouns are quite recent developments.

2.3. Innovative pronouns as reanalyzed exclusive and inclusive forms

The resemblance between these innovative 1st person forms and the inclusive/exclusive forms discussed in Section 1 is unmistakable, and it is obvious that they must be historically connected. The use of 1PL forms for 1SG reference is common throughout the world. In English, for example, *we* is used with singular reference in a wide variety of functions: for example the canonical “royal *we*” and “editorial *we*” (which the alert reader will notice being used in this paper), or the common use of the plural pronoun with singular reference in contexts of real or imitation intimacy, e.g. a nurse asking a patient: *How are we feeling this morning?* There are many examples around the world of the use of specifically inclusive or exclusive plural forms for 1SG reference. Cysouw (2005a) gives examples of exclusive for polite, and inclusive for “humble”, “bonding”, and impolite singular reference. Since we know, on the basis of the Mongsen evidence, that the ancestral pronominal paradigm distinguished three separate forms, a nasal-initial 1SG, a velar stop-initial 1EXC, and a front vocalic 1INC, we can easily derive the systems of the modern languages by positing that in Kuki-Chin a pervasive (perhaps polite) use of the #*k*- ‘exclusive’ with singular reference was grammaticalized, leading to the loss of the original 1SG form, and that in the Naga Belt languages the same thing happened with the inclusive.

The resulting systems fit into common typological patterns (Cysouw 2005b, Bickel and Nichols 2005). In Kuki-Chin, the 1SG and Exclusive forms are based on the same root, while the Inclusive form is distinct (see Table 1). Cysouw notes that this is a common phenomenon worldwide. However, the other pattern, where 1SG and inclusive are based on the same root and the exclusive form is distinct, is rare or nonexistent. And the development of the Naga Belt languages conforms to this pattern: as far as we can tell from the available data, all of the languages with 1SG #*i* ~ *ai* forms have lost the clusivity distinction altogether, and have only a unified 1PL form, based on the 1st person #*i* root plus a plural formative.

3. Clusivity in Kuki-Naga in Tibeto-Burman context

In the previous sections we examined evidence which at first glance supports the traditional classification of these languages into a Kuki-Chin and a Naga branch, and the likewise traditional, but more speculative hypothesis that these groups are joined together in a larger Kuki-Naga branch. In this section we will see that the evidence is not so clear; we will need a more complex account of how these innovations arose and spread.

The innovative 1st person pronouns discussed in Section 2 seem to support the idea of distinct Kuki-Chin and Naga branches. The innovative #*kai* pronoun has been cited as one of the shared innovations which define the Kuki-Chin branch (Thurgood 1985, VanBik 2009). And if the similarly innovative #(a)*i* were universally shared by all of the Naga Belt languages, it would constitute a strong argument for the traditional concept of a Naga branch of Tibeto-Burman. But in fact, as we will see in Section 3.2, neither innovation is universal across either branch.

The paradigm of exclusive #*kai*, inclusive #(a)*i* described in Section 1 suggests a genetic connection between Kuki-Chin and at least one Naga group, and the prevalence of #(a)*i* as a singular pronoun throughout the Naga Belt suggests that the rest of the languages once shared the same clusivity system. But we will see in Section 3.1 that this system probably has deeper roots than that, and thus does not automatically support for the Kuki-Naga hypothesis. In Section 4 I will try to reconcile these seemingly contradictory facts and offer an account of the origins of both innovative pronouns which, will argue, does at least weakly support the concept of Kuki-Naga as a genetic unit.

3.1. The Kiranti connection

As we noted in Section 1.1, most of the attested inclusive/exclusive systems found in Sino-Tibetan represent relatively recent developments. The clearest exception is Kiranti, where there is substantial evidence for Proto-Kiranti pronoun and verb agreement paradigms that include clusivity (Bauman 1975, van Driem 1991, LaPolla 2005). Strikingly, the most likely reconstructed forms match those which we have seen in Kuki-Chin and Ao. A clear example illustrating the relevant roots in independent pronominal forms is Thulung, as shown in Table 8.

Table 8 – 1st person pronouns and agreement suffixes in Thulung (Lahaussais 2003)

Thulung	1SG	EXC	INC
pronoun	<i>go</i>	<i>gu-ku</i>	<i>gu-i</i>
verb agreement	<i>-ŋu</i>	<i>-ku</i>	<i>-i</i>

The correspondence between Thulung exclusive *-ku*, inclusive *-i* and the *#ka* and *#i* forms which we have been examining is compelling. This correspondence between the Kuki-Chin, Ao, and Kiranti paradigms suggests that the clusivity distinction and the *#k-* and *#i* forms may be reconstructable for their nearest common ancestor, which would have to be Proto-Tibeto-Burman. Thus the fact that Ao and Kuki-Chin share the same forms is not itself evidence for a lower-level Kuki-Naga branch; it is possible in principle that their nearest common ancestor is PTB. However, we have seen that the inclusive and exclusive forms which we have seen above provide an explanation for innovative pronoun forms in both Kuki-Chin and Naga languages, and in Section 4 I will suggest that the hypothesis of a Kuki-Naga branch then helps to explain the overall variation which we find in 1st person pronouns across all of the languages in question.

3.2. Relict pronominal forms

Although the great majority of Naga Belt languages have a 1SG form relatable to *#ai*, and almost all Kuki-Chin languages have a reflex of *#kai*, there are exceptions to both. These exceptional forms are so far reported only in a handful of Ao and Angami-Pochuri languages, and in Lamkang and Anal, two very closely related languages of the NW branch of Kuki-Chin. In all cases the non-corresponding form is something like *#ni*, which Bruhn (2014), in his reconstruction of Proto-Ao, derives from **ŋay*, relatable to the attested Jinghpaw pronoun *ŋai*. I will take this reconstruction as applicable to all the similar 1SG forms which we find, since they correspond well, and there is no other apparent source for them. The relevant data are summarized in Table 9.

Table 9 – Conservative 1st person singular forms

		1SG	1PL	2SG
Angami-Pochuri	Sumi	<i>ni</i>	<i>ni-ŋù</i>	<i>no</i>
	Kheza	<i>nye</i>	<i>a-ro</i>	<i>nò</i>
Ao	Mongsen Ao	<i>ni</i>	<i>i-sa</i>	<i>na</i>
NW Kuki-Chin	Lamkang	<i>nəy</i>	<i>nəy-n</i>	<i>nəŋ</i>
	Anal	<i>ni</i>	<i>ni-hin</i>	<i>nang</i>

These languages are very important witnesses for our reconstruction of the history of pronominal forms in these branches. If Anal *ni* and Lamkang *nəy* are cognate to Sumi and Mongsen *ni*, as they evidently are, then this form is reconstructable to their nearest common ancestor. Since Sumi and Mongsen are clearly not Kuki-Chin languages, that nearest common ancestor must be older than Proto-Kuki-Chin. If these forms are reflexes of the Proto-Tibeto-Burman root **ŋa*, as Bruhn suggests, then the irregular *#nəy* reflex must have been an innovation in an intermediate proto-language, ancestral to Proto-Kuki-Chin, Proto-Ao, and Proto-Angami-Pochuri, but not to Northern Naga or Jinghpaw. And that in turn means that the wholesale replacement of this root by the erstwhile plural forms is not a shared innovation in Proto-Kuki-Chin, or even in Proto-Ao or Proto-Angami-Pochuri, much less at any deeper level. Thus not only do the innovative 1SG forms in the Naga Belt not argue for a Naga branch; even the innovative form in Kuki-Chin turns out not to be a truly shared innovation for that branch.

4. A case for Kuki-Naga?

We usually make a case for two languages or groups being related on the basis of shared innovation. We do not have such a case here. In Section 3.1 we saw that the inclusive and exclusive morphemes shared by Ao and NE and NW Kuki-Chin are also found in Kiranti, and therefore may not represent shared innovation. And while Kuki-Chin and Naga Belt languages both have innovative 1SG pronouns, these represent two distinct innovations: *#kai* in Kuki-Chin, and *#(a)i* in Angami-Pochuri, Ao, Zeme, Tangkhul and Meithei. And neither innovation is even shared completely within its group – we saw in Section Section 3.2 that we find NW Kuki-Chin languages that do not share the innovation of *#kai*, and Ao and Angami-Pochuri languages that do not share the innovation of *#(a)i*.

But it is still possible to make some inferences about the relationship of these languages from these facts. First recall that, as discussed in Section 3.2, the retention of the original nasal 1SG root in Lamkang, Anal, Mongsen and Sumi tells us that the nearest common ancestor of all these languages must have had the same pronominal system: 1SG *#ŋai*, exclusive *#kai*, inclusive *#(a)i*. Now note that a shift in the use of a form as a plural pronoun to singular reference can only come about from a situation in which a basically plural form is used in certain circumstances with singular reference – as is the case with English *we*. So the different development of the Kuki-Chin and Naga Belt languages implies that in Kuki-Chin the use of exclusive with singular reference was predominant, while the Naga Belt languages preferred to use the inclusive. Here we see the common Kuki-Chin and common Naga Belt factors which divide the two sets of languages.

This is still two separate innovations, but not necessarily independent. It is easier to imagine this shift in different directions if it is occurring in languages which are mutually intelligible and belong to the same larger speech community. There is ample evidence that the Kuki-Chin languages form a genetic unit (VanBik 2009), but no conclusive evidence that the Naga Belt languages do. But if all these languages together form a genetic unit, then there was a time when all of their ancestors comprised a speech community of closely-related, mutually-intelligible languages. Assume that there was a prevalent use of inclusive forms for singular reference, as we find reflected in all of the daughters except for Kuki-Chin. Then the genesis of Proto-Kuki-Chin would require that the ancestral PKC community at some point distinguished itself from the rest; perhaps this self-distinction involved a shift toward general use of the exclusive rather than the inclusive for 1st person singular reference.

Obviously this argument is very speculative. But I have shown in this paper that the innovative pronouns of both Kuki-Chin and the Naga Belt must derived from an inclusive-exclusive distinction that was once shared by the ancestors of all of these languages.

Data sources

Anal: Onahring 2016; Angami: Marrison 1967; Athpare: Ebert 1997; Hmar: Baruah and Bapui 1996; Hrangkhoh: P. Haokip 2016; Hyow: Peterson 2003; Khaling: Toba 1979; Koireng: C. Y. Singh 2010; Lamkang: Khular and Khular 2016; Mara: Arden 2010; Meithei: Chelliah 1997; Milang: Post and Modi 2011; Mizo: Chhange 1993; Mongsen Ao: Coupe 2007; Monsang: Konnerth and Wanglar 2014; Paite: N. S. Singh 2006; Sizang: Stern 1963, Sarangthem 2010; Sumi: Sreedhar 1980; Tangkhul: Arokianathan 1987; Tedim: Henderson 1965; Thulung: Lahaussais 2003; Yimchungrü, Zeme: Marrison 1967; Zaiwa: Lustig 2010.

References

- Arden, Michelle. (2010). *A phonetic, phonological, and morphosyntactic analysis of the Mara language*. MA Thesis. San Diego, San Diego State University. Accessed 22/01/2012 at http://scholarworks.sjsu.edu/etd_theses/3744.
- Arokianathan, S. (1987). *Tangkhul Naga Grammar*. Mysore, Central Institute of Indian Languages.
- Baruah, P.N. Dutta and V.L.T. Bapui. (1996). *Hmar Grammar*. Mysore, Central Institute of Indian Languages.
- Bauman, James. (1975). *Pronouns and Pronominal Morphology in Tibeto-Burman*. PhD Dissertation. Department of Linguistics. Berkeley, University of California.
- Bickel, Balthasar and Johanna Nichols. 2005. "Inclusive-exclusive as person vs. number categories worldwide." In Elena Filimonova, Ed. *Clusivity: Typology and Case Studies of the Inclusive-Exclusive Distinction*. Amsterdam, Benjamins: 49-72.
- Bruhn, Daniel. (2014). "A Phonological Reconstruction of Proto-Central Naga." PhD Dissertation. Department of Linguistics. Berkeley, University of California.
- Chhange, Lalnunthangi. (1993). *Mizo syntax*. PhD Dissertation. Department of Linguistics. Eugene, University of Oregon.
- Chelliah, Shobhana. (1997). *A Grammar of Meithei*. Berlin, de Gruyter.
- Coupe, Alexander R. (2007). *A Grammar of Mongsen Ao*. Berlin, de Gruyter.
- Cysouw, Michael. 2005a. "A typology of honorific uses of clusivity." In Elena Filimonova, Ed. *Clusivity: Typology and Case Studies of the Inclusive-Exclusive Distinction*. Amsterdam, Benjamins: 213-230.
- _____. 2005b. "Syncretisms involving clusivity." In Elena Filimonova, Ed. *Clusivity: Typology and Case Studies of the Inclusive-Exclusive Distinction*. Amsterdam, Benjamins: 73-111.
- DeLancey, Scott. (2013). "Argument indexation (verb agreement) in Kuki-Chin." *Paper presented at the 46th International Conference on Sino-Tibetan Languages and Linguistics, Dartmouth College, Hanover, New Hampshire, USA, August 7-10*.
- _____. (2015). "Morphological evidence for a Central branch of Trans-Himalayan (Sino-Tibetan)." *Cahiers de Linguistique – Asie Orientale* 44(2): 122-149.
- DeLancey, Scott, Krishna Boro, Linda Konnerth, and Amos Teo. (2015). "Tibeto-Burman languages of the Indo-Myanmar borderland." Presented at the 31st South Asian Languages Analysis Roundtable, Lancaster, 14 May 2015.
- Driem, George van. 1991. "Bahing and the Proto-Kiranti verb." *Bulletin of the School of Oriental and African Studies* 54(2): 336-356.
- Ebert, Karen. 1997. *A Grammar of Athpare*. München: Lincom Europa.
- Filimonova, Elena, Ed. (2005). *Clusivity: Typology and Case Studies of the Inclusive-Exclusive Distinction*. Amsterdam, Benjamins.
- Haokip, Paothang. (2016) *Four Kuki-Chin Languages of the Barak Valley*. Delhi, Jawaharlal Nehru University. Unpublished manuscript.

- Henderson, Eugénie J. A. (1965). *Tiddim Chin: A descriptive analysis of two texts*. London, Oxford University Press.
- Khular, Sumshot and Tony Khular. (2016). "Lamkang." In G. N. Devy and K. Nipuni Mao, Eds. *The Languages of Manipur* (People's Linguistic Survey of India, Volume 18, part 2). Hyderabad, Orient BlackSwan: 186-202.
- Konnerth, Linda and Koninglee Wanglar. (2014). "Monsang, a Northwestern Kuki-Chin language from Manipur." Paper presented at the 47th International Conference on Sino-Tibetan Languages and Linguistics. Yunnan Normal University, Kunming.
- Lahaussais, Aimee. 2003. "Thulung Rai." *Himalayan Linguistics Archive* 1: 1-25
- LaPolla, Randy. (2005). "The inclusive-exclusive distinction in Tibeto-Burman languages." In Elena Filimonova, Ed. *Clusivity: Typology and Case Studies of the Inclusive-Exclusive Distinction*. Amsterdam, Benjamins: 291-311.
- Lustig, Anton. (2010). *A Grammar and Dictionary of Zaiwa*. Leiden, Brill.
- Marrison, Geoffrey E. (1967). *The Classification of the Naga Languages of North-East India*. PhD Dissertation. Department of Phonetics and Linguistics. London, School of Oriental and African Studies.
- Norman, Jerry. (1988). *Chinese*. Cambridge, Cambridge University Press.
- Onahring, B. D. (2016). "Anal." In G. N. Devy and K. Nipuni Mao, Eds. *The Languages of Manipur* (People's Linguistic Survey of India, Volume 18, part 2). Hyderabad, Orient BlackSwan: 42-49.
- Peterson, David. (2003). "Agreement and grammatical relations in Hyow." In David Bradley, Randy LaPolla, Boyd Michailovsky and Graham Thurgood, Eds. *Language Variation: Papers on Variation and Change in the Sinosphere in honour of James A. Matisoff*. Canberra, Department of Linguistics, Australian National University: 173-183.
- Post, Mark W. and Yankee Modi. (2011). "Language contact and the genetic position of Milang (Eastern Himalaya)." *Anthropological Linguistics* 53(3): 215-258.
- Sarangthem, Bobita. (2010). *Sizang (Siyin) Grammar*. PhD Dissertation. Department of Linguistics. Imphal, Manipur University.
- Singh, Chungkham Yashawanta. (2010). *Koireng grammar*. New Delhi, Akansha.
- Singh, Naorem Saratchandra. (2006). *A Grammar of Paite*. New Delhi, Mittal.
- Sreedhar, M. V. (1980). *A Sema Grammar*. Mysore, Central Institute of Indian Languages.
- Stern, Theodore. (1963). "A provisional sketch of Sizang (Siyin) Chin." *Asia Major* 10: 222-78.
- Toba, Sueyoshi. 1979. *Khaling*. [Asian & African Grammatical Manual 13d]. Tokyo: Tokyo Gaikokugo Daigaku.
- Thurgood, Graham. (1985). "Pronouns, verb agreement systems, and the subgrouping of Tibeto-Burman." In Graham Thurgood, James A. Matisoff, and David Bradley, Eds. *Linguistics of the Sino-Tibetan Area: The State of the Art*. [Pacific Linguistics Series C, #87]. Canberra, Australian National University: 376-400.
- VanBik, Kenneth (2009). *Proto-Kuki-Chin: A Reconstructed Ancestor of the Kuki-Chin Languages*. (STEDT Monograph 89). Berkeley, Sino-Tibetan Etymological Dictionary and Thesaurus Project, University of California.

Verb stem alternation in Pangwa Tangsa¹

Stephen Morey

Centre for Research on Language Diversity, La Trobe University

Abstract

Tangsa is a very linguistically diverse group spoken on the India-Myanmar border, with around 80 distinct and named varieties, many of which are mutually intelligible but many of which are not. In Myanmar, the Tangsa sub-tribes are grouped under Tangshang Naga.

This paper examines and exemplifies the phenomenon of verb stem alternation, whereby a single verb has two stems, a verbal stem, which is demonstrated to be the 'base' or underived form, and a nominal stem which has different form, most frequently a different tone category.

As with other Tibeto-Burman languages, both the forms and the functions of the stem alternations in the Tangsa varieties show considerable diversity. In the Tangsa varieties treated here, one of the stems, which we term the verbal form, is clearly the underived root and the alternate stem is derived from it. The most frequent way of forming an alternate stem is a change of tone category, keeping the vowel and any final consonants the same. In multiple cases in one Tangsa variety there is an alternate stem carrying a different form, and in a related variety there is no alternation between the stems.

The verbal and nominal stems for 151 roots in the Mueshaung and Ngaimong varieties will be compared, and the findings of this comparison will be enhanced by observations of stem alternation from other Tangsa varieties.

Citation	Morey, Stephen. 2018. <i>Verb stem alternation in Pangwa Tangsa</i> . North East Indian Linguistics (NEIL) 8, 86-109. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.
Volume Editors	Linda Konnerth, Stephen Morey, Amos Teo
Copyright	© 2018, the author(s), release under Creative Commons Attribution license
URL	http://hdl.handle.net/1885/155280

1 Introduction

In many of the varieties of the Tangsa languages, there are two main forms of the verb, a verbal form, which combines with the person marked agreement particles (Morey 2011a, 2017), and a nominalized form, often preceded by a nominalizer. In many cases these two forms are identical, but they may vary, as shown with the verb 'to weep', the 'verbal form' of which is exemplified in (1) and the 'nominal form' in (2) in the Chamchang variety of Tangsa².

¹ The research that has led to this paper was undertaken as part of an Australian Research Council funded Future Fellowship (ARC FT100100614), taken up at the Centre for Research on Language Diversity, La Trobe University, and a subsequent ARC Discovery Project grant to study the traditional Tangsa Wihu song (ARC DP160103061). Much of the data was collected during an earlier project funded by the DoBeS program of the Volkswagen Foundation, Germany. I am very grateful to all my linguistic colleagues, particularly Scott DeLancey, Linda Konnerth, Pavel Ozerov, Mijke Mulder, Kellen Parker van Dam, Iftiqar Rahman, Niharika Dutta and Deepjyoti Goswami. My greatest debt is to the many Tangsa / Tangshang consultants who have provided me with much information, in particular Wanglung Mossang, Renman Mossang (Mueshuang), Wangkui Ngaimong, Shinyung Ngaimong (Ngaimong), Rev. Yanger Thungwa, Ninshom Chena (Chamchang), Bynn Kham Lann (Shecyü and Mungre), Shapwon Wonti (Rinkhu), Lukam Cholim, Yan Linn Aung (Cholim), John Innga (Yvngban Wvng), Kheshem (Shangti), Satum Ronrang (Rera).

² We have spelled the names of the Tangsa varieties or subtribes using a Romanization that represents the way each group pronounces their own name (autonym). (See Morey 2017 for a further discussion). Examples are presented with the Chamchang orthography devised by Rev. Longkhap Yanger Thungwa, where the letter <q> stands for a glottal stop, <f> for the 3rd tone, realised as high falling in Chamchang, and <x> for the 1st tone, realised as low falling in Chamchang. The second tone, which is realised as mid, is unmarked in Rev. Yanger's

- (1) *hoq* *tof*
hoʔ *to³*
 weep PST.3
 ‘(He/she/it/they) wept’
- (2) *aho*
ə-ho²
 NOMZ-weep
 ‘weeping’ / ‘to weep’

We can view the two forms of ‘weep’ in (1) and (2) as alternate verb stems. The nominal stem, in this case *ho²*, is here preceded by a nominalizer that is written in the Chamchang orthography³ with the letter <a>. The form of the nominalizer differs from one variety to the next, usually *ə-* or *ɪ-*. In the Shangti variety, however, the nominalizer is a nasal that is homorganically assimilated to the place of articulation of the initial consonant of the verb. (See below Table 4 for examples of the different prefix forms). Other prefixes, such as causative, middle and reflexive, also combine with the nominal stem (see below section 3).

Tangsa is the name given to a group of ‘sub tribes’ in India, each of which speaks a distinct linguistic variety. Some of these varieties are fully mutually intelligible with each other, while others are fully unintelligible, but all belong to the Northern Naga group of languages. In Myanmar (Burma), the languages that are grouped under Tangsa are now referred to as Tangshang, a term which also includes Nocte and Wancho, both of which are listed as separate languages in India. Both Tangsa and Tangshang have the ISO 639-3 code nst. (See Statezni 2013, Morey 2015, 2017, Khan Lann 2017 for further discussion of the diversity within Tangsa and the many names used to refer to them.) Within Tangsa the biggest group of varieties are termed Pangwa⁴ in India, and it is the verb stem alternation in Pangwa that we will explore in this paper.

After reviewing the situation of verb stem alternation more generally in the Tibeto-Burman languages (section 2), we will deal in more detail with the functions of the alternate stems (section 3), and present a comparison of forms and functions in two varieties, Mueshaung and Ngaimong (section 4).

As will be seen in section 4, there are significant differences, not only in the form of the stem alternations, but also in which verbs exhibit this feature. The paper will therefore seek to explore the following questions:

- 1) Can we establish that the ‘root’ or ‘underived’ form is the same as the verbal stem?
- 2) What differences have been observed between varieties in the forms and functions of stem alternation?
- 3) Can we suggest that verb stem alternation is a feature of an earlier stage of Tangsa language?

system. Discussions with Rev. Yanger suggested that the nominalizer prefix should be analysed as having Tone 1. However, comparison with other varieties suggests that it is toneless and should be written as /ə-/. In our phonemic representation here, and throughout this paper, tone categories are shown as superscripted numbers. Tone, which plays an important role in stem alternation, is discussed briefly in Section 3.

³ In the Tangsa Naga Unified Orthography developed in Myanmar, this is written <ä-> (Khan Lann 2017); in the Mueshaung orthography developed by Rev. Gam Win, it is written <v->. Each of these represents an unstressed vowel that we will write as schwa /ə/.

⁴ Khan Lann (2017) spells this as Pangva. His thesis uses <v> throughout for the sound usually written as <w> by Tangsa groups in India. The sound represented is variously [w], [β] and [v] in different varieties and for different speakers.

2 Verb stem alternation in a Sino-Tibetan perspective

Most of the Northern Naga languages have still not been described in any depth, and many of our observations are, by necessity, still rather preliminary. Verb stem alternation was first recorded by Weidert (1979, 1987) for the Joglei variety of Tangsa. He listed several verbs as having two stems, with the second stem described as “derived/non-finite”. Weidert also recorded some Nocte words, which also showed stem alternation.

A closely related language, Wancho, does not appear to have verb stem alternation, both as reported in Burling and Wangsu (1999) and in our own more recent work with the Wancho consultant Banwang Losu (work in preparation).

More widely, verb stem alternation is found in many language groups within the Tibeto-Burman (Trans-Himalayan) languages: in Kuki-Chin (VanBik 2009); in Gyalrongic (Sun 2004, Jacques 2017); in Tibetan, where there are four stem forms usually termed ‘Present’, ‘Past’, ‘Future’ and ‘Imperative’ (DeLancey 2017)); in Tangut (Gong forthcoming) and also in the Kiranti languages, for example Khaling (Jacques et. al. 2012: 1101).

For Kuki-Chin, VanBik (2009: 9) describes Verb Stem Alternation as an innovation in Proto-Kuki-Chin, suggesting that it was one of the features “that separate Kuki-Chin languages from the rest of the Tibeto-Burman family”. In view of the widespread nature of verb stem alternation, as mentioned earlier, this observation seems no longer justifiable, although there could be particular kinds of verb stem alternation that are innovations at the level of Proto-Kuki-Chin. In the India-Myanmar border area, verb stem alternation is a feature, like person agreement/indexation, found in two separate areas, the Tangsa area in the Northeast and the Kuki-Chin area to the Southwest (see DeLancey 2004). The presence of both verb stem alternation and person agreement in languages in two distinct areas suggests a connection between the two.

VanBik (2009: 10) observed that “This alternation is arguably not linked in any straightforward way to a single parameter of variation such as tense, aspect, or transitivity.” In Kuki-Chin the two stems are “usually called Form I / Form II”, noting that in many of the languages researched, Form I is used in main clauses and Form II in subordinate clauses.

VanBik also observed that the morphophonemic alternations between the forms are not regular across languages (2009: 12), concluding that “the basic distinction between Form I and Form II needs to be determined by the criterion of whether they appear in main or subordinate clauses, regardless of their morphophonemic features.” As an example of this, in Hakha Lai, intransitive verbs use the Form I and transitive verbs use Form II, except that in negation, imperative and questions Form I is always required.

The terms for the different verb stems vary in different analyses. Bedell (2013)⁵, for example, used the terms ‘base’ and ‘derived’, while So-Hartmann (2009) uses Stem A and Stem B. The difference in these two terminologies goes to one of our key questions, is there are root or underived form? The terminology of VanBik (2009) and So-Hartmann (2009), and many others, does not make a claim that one form is derived from the other, whereas Bedell (2013) does. In a number of Kuki-Chin languages the stems vary for indicative clauses, such as in Daai Chin where intransitive indicative clauses exhibit Stem B and transitive indicative clauses exhibit Stem A. Stem A is also the stem for nominalization and non-final adverbials. (A fuller list of the different functions of the two stems is found in So-Hartmann 2009: 71). This type of functional distinction, where indicative sentences vary in stem depending on transitivity would suggest that neither stem is the ‘base’. However in Lai, both transitive and intransitive constructions exhibit Stem I, but if a transitive sentence is clearly agentive, then Stem 2 is used. In other words, a sentence ‘the man is catching fish’ would take Stem I, but

⁵ Cited in Bawi Tawng 2017.

‘the man caught the fish’, where the fish has been affected, would take Stem 2 (Bawi Tawng 2017: 41). In these circumstances it might be possible to argue the So-Hartmann’s Stem B and Bawi Tawng’s Stem 1 are both functionally more ‘base’. In the Pangwa varieties of Tangsa, it appears clearer that one form is ‘base’ and the other ‘derived’, and that the verbal form is the base form. This will be discussed further below.

As with the Pangwa varieties of Tangsa discussed in this paper, the verbs that have alternate stems are not always cognate in each of the Kuki-Chin languages. For example, in Daai Chin around 20% of verbs have alternate stems (So-Hartmann 2009: 71), but in related languages like Mizo, almost all verbs have alternating stems (Chhangte 1993), and in Tedim, all verbs have two stems, (Henderson 1965: 72). The Monsang language has a number of verbs that do not have alternate forms but for which in proto-Kuki-Chin VanBik reconstructs two forms. Konnerth (pers. comm.) has suggested that there may have been a common core of alternating stems, but in some languages, alternation was extended to a larger group of verbs.

Unlike in the Tangsa data presented in this paper, where most of the alternate stems are formed by tonal alternation, in many other Tibeto-Burman languages alternations are more frequently formed by a change to the final consonant. For Kuki-Chin, VanBik (2009), suggests that tone alternation was a feature of the proto-language, positing a number of processes, which most frequently involve the alternate form (II) having a final stop, as shown by examples of some of the more frequent alternations presented in Table 1.

Table 1: Examples of Verb Stem Alternation in Proto-Kuki-Chin (van Bik 2009)

Van Bik’s No.	Semantics	Proto stems	Process
9	BOIL (v.) ₁	*but-I, *buʔ-II	*-t-I, *-ʔ-II
93	LIFT ₁ / SUPPORT ₁	*dom-I, *domʔ-II	*-m-I, *-mʔ-II
39	OWE / BORROW ₁	*baa-I, *bat-II	*-VV-I, *-Vt-II
108	STAND / STOP ₁	*dɪŋ-I, *dɪn-II	*-ŋ-I, *-n-II

Typical patterns for Stem II in the Proto Kuki-Chin are a change from a stop *-t* or *-k* to *-ʔ*, or the addition of a glottal stop to an existing final nasal or lateral, or the change from a vowel final to a final stop, such as *-t*, usually involving vowel shortening. A minority of examples reconstructed by VanBik have the change *-ŋ* (Stem I) to *-n* (Stem II), but none show tone alteration. This is due in part to VanBik not suggesting tonal reconstruction for verbs. While I have not been able to examine all of the data in Kuki-Chin, it is clear that only a minority of verbs in the Kuki-Chin languages show tonal alternation as the form of marking the alternate stems. Examples include Hakha Lai *bùay-I, bŭay-II* ‘be busy, be disordered’ and *bùl-I, bŭl-II* ‘be maimed (e.g. lost limbs, etc.)’.

For Showu rGyalrong, Sun (2004) suggested that there are four processes of ‘stem formation’ which he named as vocalic alternation, consonantal alternation, tonal alternation and accentual alternation. Those showing tonal alternation only are somewhat less frequent than the vocalic or consonantal alternations.

A fuller comparison of the forms of alternation across the range of languages that exhibit it will have to await a more complete collection of data, but we can perhaps suggest some possible cognate processes⁶. In Table 18 below, we see the following forms for the verb ‘go’, (given in the order verbal then nominalized stem), Mueshaung: *ka^l, əkai^l*; Ngaimong *ka^l, əkal^l*; Rinkhu *ka^l, ikat*. This correspondence of final *-i, -l* and *-t* is also found for the verbs

⁶ I am very grateful to Guillaume Jacques for suggesting this possibility.

‘wash clothes’ and ‘laugh’⁷. It may be that this correspondence set arises from an earlier *-t suffix, a correspondence of which is found in a Gyalrongic language, shown for instance by Ngaimong *mul*^l, Mueshaung *mui*^l ‘blow’, cognate with Japhug *ɣɾmut* ‘blow’. This suggests that perhaps the alternate stem for ‘go’ is indeed a reflex of an earlier proto final *-t suffix that was one documented means of forming an alternate stem. It is found in this function in Japhug as *ti*, *tut* ‘say’ (Jacques 2017: 615)⁸ and also in VanBik’s (2009) proto Kuki-Chin reconstruction for the word ‘oweborrow’ in Table 1 above. King (2009:143) suggests that this goes back to a nominalizer, noting that it is a well-attested Proto-Tibeto-Burman morpheme, see for example Matisoff (2003: 454).

3 Functions of stem alternation in Tangsa

The basic distinction of verbal and nominal stems, the latter of which is combined with a nominalizer, presented in the introduction as examples (1) and (2), is a somewhat oversimplified model. The functions of the two stems varies considerably in the Pangwa Tangsa varieties, and this variation will be explored in this section.

Before proceeding to discuss the functions in detail, it will first be necessary to give some information about the tone systems in the Pangwa Tangsa varieties. There are three tone categories for words with final vowels or sonorants (nasals and in some case laterals and rhotics). These categories are here numbered 1, 2 and 3 and are shown with superscripted numbers. The form of the tone varies considerably from one variety to the next, so Tone 2 in Mueshaung is high falling, but in Chamchang is mid level. (See Morey 2015, 2017, Khan Lann 2017 for more detail on the tones in a range of varieties.)

Consider the Mueshaung example presented as (3). In this example, the Mueshaung language is presented first in the script developed by Lakhum Mossang, then on the second line in the Roman based orthography developed by Rev. Gam Win, and the third line is a phonemic transcription.

- (3)
- | | | | | | | |
|------------------|-----------------|-------------------|--------------------|--------------------------------|---------------------|-----------------|
| ႃ႗႗႗ | ႃ႗႗႗ | ႗႗႗႗ | ႗႗႗႗ | ႗႗႗ | ႗႗႗႗႗ | ႗႗႗ |
| <i>lvmez</i> | <i>laq</i> | <i>vtueq</i> | <i>vshoayz</i> | <i>phuiuz</i> | <i>rvghayz</i> | <i>tueq.</i> |
| ləm ¹ | la [?] | ə-tɾ [?] | ə-foi ¹ | p ^h uu ¹ | rə-ɣai ¹ | tɾ [?] |
| bring | after | NOMZ-PST.3 | NOMZ-roast | eat.NOMZ | AUX-DISCARD | PST.3 |
- ‘And then they were given the writing on animal skin, and after that, having roasted it, it was eaten away.’

Writing Story, nst-mos_20140304_05_SM_H4n_Wanglung_WritingStory, No (7)⁹

In (3) there are two nominalized forms, *ə-foi*¹ ‘roasting’ (verbal form is *fu*¹) and *p^huu*¹ ‘eating’ (verbal form is *p^hauk*). Only the first of these is found in combination with the Mueshaung nominalizer *ə-*, and that same nominalizer is also found in combination with the person agreement marker *tɾ*[?]. One view of example (3) is that the two forms *foi*¹ and *p^huu*¹

⁷ There is also a relationship between this correspondence and the ‘Marked Tone Correspondence’ set listed in Table below.

⁸ I am grateful to Guillaume Jacques for pointing out that this final *-t could also be related to the nominalizing -d suffix in Tibetan (Hill 2014: 625). Since the primary function of stem alternation in the Tangsa languages is nominalization, this connection is hardly surprising.

⁹ I have given the names of the recordings from which these examples are taken, as in this case nst-mos_20140304_05_SM_H4n_Wanglung_WritingStory which refers to a .wav file. This, along with all the recordings of the Tangsa materials, is in the process of being archived in The Language Archive (<https://tla.mpi.nl/>, then click on ‘Access the Archive’ and select ‘DoBeS Archive’ then ‘Tangsa, Tai, Singpho in North East India’).

are both nominalized, and that the main verb is the light verb *rə-*, literally ‘do’ in combination with an auxiliary *ɣai¹*, which literally means ‘discard’ and it is this combination that carries the agreement marker, the final element in the example.

Thus one pattern for the use of the nominalized stem is that shown in (4), one or more nominalized verb stems preceded by the nominalizer, with the main verb being a light verb *rə-* in combination with an agreement marker:

- (4) *ə*-V *rə*-AGR
 NOMZ-V.NOMZ LV-AGR

In the Mueshaung variety at least, the place of the light verb can be taken by an auxiliary, unmarked with agreement, such as *raunɣ¹* ‘need’, a verb that does not mark agreement with the actor *ŋa¹* ‘I’. This is shown in (5), where the verb ‘catch’ takes the nominalized stem, its verbal form being *lu¹*. When the nominalized stem is used, it will henceforth be marked .NOMZ as in (5).

- (5) *ngaz* *wuz* *vloyz* *raungz*
 ηa^1 βu^1 ∂ - loi^1 $rau\eta^1$
 1SG chicken NOMZ-catch.NOMZ need
 ‘I have to catch a chicken.’

However, in Mueshaung, at least, it is not the presence of the prefix *ə-* that triggers the stem alternation. Mueshaung has a present continuous construction that employs *ə-* but does not trigger the alternate stem. This is demonstrated in (6):

- (6) *ngaz* *wuz* *vluz*
 ηa^1 βu^1 ∂ - lu^1
 1SG chicken NOMZ-catch
 ‘I am catching chickens.’

Despite the fact that the verbal stem is used in (6) we have nevertheless glossed *ə-* as a nominalizer, although it may be the case that it is being reanalyzed as present continuous prefix when combined with the verbal stem¹⁰. In Mueshaung at least the prefix *ə-* seems to have three principal functions, to nominalize, as in (5), to mark the present progressive, with a verbal stem, as in (6) and as a prefix on nouns, roots that are not derived from verbs, as with *ə-sa¹* ‘child’ and *ə-ɲuu¹* ‘mother’ in (7):

- (7) *vsaz* *rueq* *vnyuiuz* *muex* *ngoc* *vkhoc* *ngoꝛ* *tueq*
 ∂ - sa^1 $r\gamma?$ ∂ - ηuu^1 $m\gamma^2$ ηo^3 $\partial k^h o^3$ ηo^1 $t\gamma?$
 PF-child AG PF-mother A.AG say.NOMZ thus say PST.3
 ‘The child said like this to (its) mother ...’

Yam Story (nst-mos_20140304_03_SM_H4n_Wanglung_YamStory), told by Wanglung Mossang, No (7)

¹⁰ I am grateful to Linda Konnerth for suggesting this. It may well be more appropriate to gloss this as a continuous prefix, from the synchronic perspective, but until further research is done I have chosen to leave the gloss as NOMZ.

This nominal prefix is thought to have been historically a 3rd person singular possessive marker, but is now somewhat bleached of its meaning, used if a noun is not otherwise modified¹¹.

A clear and unambiguous single gloss for these three functions of ə- in Mueshuang is elusive, but we will retain the NOMZ- gloss for both constructions involving the two verb stems in (5) and (6) because across the various other Tangsa varieties, the present continuous function of ə- is not generally found. In Chamchang, for example (where the nominalizer is written in the orthography as <a->, the present continuous is expressed by the verbal stem of the verb and a continuous/habitual marker which in the 2nd person singular is *təlaw*³, as demonstrated in (8).

- (8) *ju-se dang mii-wa shi- tiilai nak ngu ki.*
 ʒu²-se¹ daŋ² mə-βa² ʃi²-təlaw³ nak ŋu² ki¹
 who-child voice A.AG-RL sing-CONT.2SG PROH say GO
 ‘Don’t say with whose voice you are singing.’

Explanation of the Love Song

(SDM13-20111101-02_SM_JVC_Kamshey_LoveSong), No 5.1

Just as it is possible to have a verbal stem in combination with the prefix ə- (‘nominalizer’), it is also possible to have nominal stem that is not preceded by ə-. Both of these possibilities are demonstrated in (9):

- (9) *ngaz ghurc vghurx*
 ŋa¹ ɣur³ ə-ɣur²
 1SG load.NOMZ NOMZ-load
 ‘I am carrying a load’

The verbal stem of ‘carry’ is *ɣur*², and the nominal stem shows a change to the 3rd tone *ɣur*³. The nominal stem used alone without the prefix is in this case a concrete noun, the thing carried, hence ‘load’.

Nominalized stems are also used frequently in subordinate clauses, as in the following example from Chamchang, presented in (10). The verbs preceding *ʒai²fe²* ‘after’ are both nominalized.

- (10) *ekra chin kaix jaishe kra rai ka lam raiila ...*
 e³kəra² cin²-kai¹ ʒai²fe² kəra² rai² ka² lam² rau²la² ...
 that soak-go.NOMZ after that thread that remove if ...
 ‘After going to soak it, if the thread is taken from the water ...’

Story of Traditional Thread

(nst-kim_20121015_02_SM_T_Yanger_TraditionalThreads), No 12

The nominal stem is also employed with the various valency changing prefixes: causative, middle, reflexive, reciprocal. In Mueshaung, for example, there are three such derivational

¹¹ I am grateful to Guillaume Jacques for pointing out that this prefix could be cognate to Japhug *u-* and Bantawa *u-*, both 3rd person possessive prefixes. Another possible source for the Tangsa ə- prefix could be the Lolo-Burmese *?a-* nominalization prefix. Konnerth (2014: 202) points out that the *a-* prefix on the noun gives some sort of definiteness reading.

prefixes are given in (11), each shown in combination with the verb *k^{hi}1* ‘see’, the nominal stem of which is *k^{hi}3*. These prefixes are productive.

- (11) *tə-* ‘causative’ as in *tək^{hi}3* ‘show, cause to see’
kə- ‘reflexive’ as in *kək^{hi}3* ‘reveal itself/oneself, show itself’
rə- ‘reciprocal’, as in *rək^{hi}3* ‘see each other’

Not all varieties of Pangwa Tangsa have these three forms. Cholim, for example, has only a causative *tə-* and a middle form *rə-* that subsumes the reflexive and reciprocal functions (Morey 2011b). In Ngaimong, the causative is *i³-* and the middle form is *kə-*. A comparison of forms for the verb ‘lift / rise’, in Ngaimong and Mueshaung is presented in Table 2:

Table 2: Mueshaung and Ngaimong forms of the verb for ‘lift’ or ‘rise’

	Mueshaung	Ngaimong
verbal stem	ten ¹	tuən ¹
nominal stem	əten ³	ətuən ³
middle reflexive	kəten ³	kətuən ³

In both varieties, the verbal stem carries Tone 1, but the nominal stem, which is also the form found in combination with the middle / reflexive prefix *kə-* carries Tone 3.

The derived forms can combine with each other and with the nominalizer *ə-*. In Chamchang, for example, there are two derivational prefixes *tə-* ‘causative’ and *rə-* ‘middle’. The full list of verb prefixes that trigger the alternative form are given in (12):

- (12) *ə-* ‘nominalizer’
rə- ‘middle’
ə-rə- ‘nominalizer-middle’
tə- ‘causative’
ə-tə- ‘nominalizer-causative’
tə-rə- ‘causative-middle’
ə-tə-rə- ‘nominalizer-causative-middle’

With the verb *bom²* ‘to speak’, which has the alternate stem *bom³* in combination with these prefixes, the following meanings were advanced by Rev. Yanger Thungwa. In some cases, these are combined with the (singular) imperative particle *la²*, and in others the form *teəŋ¹*, which was translated as ‘being’, ‘should be’ was employed.

- (13) *ə-bom³* ‘speaking’
rə-bom³ la² ‘have a conversation!’¹²
ə-rə-bom³ ‘something to be discussed’
tə-bom³ la² ‘let him speak!’
ə-tə-bom³ teəŋ¹ ‘let there be someone speaking’
tə-rə-bom³ la² ‘let them talk (together)!’
ə-tə-rə-bom³ teəŋ¹ ‘let there be a talking over’

Some verbs are found only in derived forms, such as Ngaimong *kək^{hi}l³* ‘shiver’. We were assured that people do not say *ək^{hi}l³*.

¹² Guillaume Jacques has suggested that the *rə-* reciprocal prefix appears to have a antipassive function in this example. Jacques (2014) is a detailed discussion of the antipassive in Japhug.

In many, but not all, Tangsa varieties, the nominalizing prefix also has an attributive function when used with verbal roots that convey the meaning of property terms, that is to say adjectives. In the case of these words, the non-prefixed form may not be used. Thus in Ngaimong *ə^hiŋ²* means ‘raw, unripe’ but the unprefixed **t^hiŋ²* is not used.

Both Ngaimong and Mueshaung have two verbs meaning ‘wash’ that are clearly cognate, but with slightly different meanings¹³. These are presented in Table 3:

Table 3: Ngaimong verbs for ‘wash’

English	Ngaimong		Mueshaung	
	Verbal	Nominal	Verbal	Nominal
wash (hand, body, dishes)	huɔl ¹	əhuɔl ³	hol ¹	əhol ³
wash (clothes)	hul ¹	əhul ¹	hui ¹	əhui ¹

The verbal forms are clearly cognate, but interestingly, there is no stem alternation with *hul¹~hui¹*, but there is with *huɔl¹~hol¹*.

In Table 4 we present forms for the word ‘give’ in 13 Pangwa varieties, in both verbal and nominal form. In Mungre, Mueshaung, Chamchang, Shecyü, Yvngban Wvng and Rera, there is stem alternation, but the forms of it differ.

Table 4: Comparison of verbs for ‘give’

Language	Verbal	Nominal
Ngaimong	koʔ	əkoʔ
Hahcheng	koʔ	əkoʔ
Mungre	kuʔ	əkoi ¹
Mueshuang	kuʔ	əkui ¹
Cholim	kuʔ	ikʷʔ
Chamchang	kuʔ	əkɛ ²
Shecyü	kuʔ	əkɛ ²
Lauchang	kau ¹	ikau ¹
Lungkhi	kuʔ	ikuʔ
Yvngban Wvng	koʔ	ikut
Shangti	k ^h yʔ	ŋk ^h yʔ
Rinkhu	kuʔ	ikuʔ
Rera	k ^h uʔ	ik ^h ut

The verbal forms are almost all identical, with an initial voiceless velar stop, a high back vowel and a final glottal stop. We suggest that except in the case of Lauchang, these verbal forms are all clearly cognate. In both Shangti and Rera¹⁴, the initial is shown as aspirated, but aspiration is not contrastive with unaspirated /k/ in those varieties, whereas aspiration is contrastive in all the other varieties listed here. Thus the initials are all cognate. Shangti typically has a front rounded vowel where most other languages have back rounded, and Cholim often has back unrounded in the same position. A reconstructed proto Pangwa form of

¹³ Note that Mueshaung has a third form for washing face /mi¹/.

¹⁴ Goswami (2017) includes a phonological sketch of Rera, but writes all the voiceless stops as unaspirated. I have written /k^huʔ/ which is a more phonetic transcription for this verb, whereas Goswami writes /kuʔ/.

ku? is then easily posited. This set suggests that the verbal stem is indeed the ‘base’ form and the nominal stem is ‘derived’.

The nominal forms on the other hand, are divergent. Although the table does not show it, in Ngaimong and Hahcheng, prefixes cause a regular pitch change to the form of the tone wherever the syllable carries Tone 1, or a final stop. Thus the nominalized *ə-ko?* is realised on a significantly higher pitch than the verbal *ko?* (see further below for more discussion of this). Since this process is entirely regular, it is not marked as a different tone category and not treated as verb stem alternation. If it were to be so treated, then almost all verbs in Ngaimong would show stem alternation. The verbal and nominal stems in Cholim, Lungkhi and Rinkhu are identical.

In Mungre and Mueshaung, however, the nominal stem does not have a final glottal stop, but is a diphthong and the syllable carries Tone 1. Lauchang, which we will discuss further below, also has a nominal stem carrying Tone 1. In Chamchang and Shecyü, on the other hand, the nominal stem has an open vowel carrying Tone 2. In Rera and Yvngvan Wvng, there is an oral stop *-t* as the coda of the nominal stem. This pattern is remarkably similar to the patterning of a marked set of tone correspondences for the following roots, as shown in Table 5:

Table 5: Examples of marked Tone Correspondences between 8 Tangsa varieties

English	Ngaimong	Mueshaungx	Mungre	Cholim	Chamchang	Shecyü	Lauchang	Rera
Tone:	1	1	1	?	2	2	1	stop
blow	mul ¹	əmui ¹	moi ¹	mo?	mei ²	me ²	əmau ¹	
fall	dəl ¹	dəi ¹	dai ¹	djɾ?	dəə ²	diə ²	ədə ¹	dat
ill	ða ¹	ətuu ¹	atsaa. ¹	ədə?	tsi ²	ədzi ²	ədi ¹	rak
cloth	k ^h əl ¹	k ^h əi ¹	khai ¹	k ^h jɾ?	k ^h εə ²	khiə ²	khe ¹	
trample	na ¹	ənuu ¹	na ¹	ne?	ɲi ² / ni ²	ni ²		

The Lauchang form for the verbal stem of ‘give’ in Table is divergent, with a final diphthong carrying tone 1, the same form as is found in the nominal stem. This looks to be a case of backformation, where at some previous stage in the history of Lauchang, there may have been a verbal stem *ku?* and a derived nominal stem *kau¹* but the nominal stem has been re-interpreted as the verbal stem. Note that this would also correspond with the pattern of tones exhibited in Table 5.

A final function of verb stem alternation is in compounding. There are two words for ‘eat’ in Tangsa languages, to eat what are sometimes grouped as ‘soft foods’, cooked rice, bananas, eggs and fish, and to eat other foods. The forms for these two verbs, with both stems, are given in Table 6. As can be seen, there is no stem alternation in Ngaimong (save the change of pitch that is predictable, discussed above), but there are alternate stems in Mueshaung.

Table 6: Verb stems for ‘eat’ in both Mueshaung and Ngaimong

English	Ngaimong		Mueshaung	
	Verbal	Nominal	Verbal	Nominal
eat (soft foods)	sa?	əsa?	sa?	əsəi ¹
eat (hard foods)	p ^h ək	əp ^h ək	p ^h auk	əp ^h uu ¹

Both varieties have a word for ‘food, edible things’, which in Mueshaung is *p^huu¹ səi¹*, a form that is made up of the two nominal stems seen in Table 6. The Ngaimong form is *p^hək səi¹*, where the root form of ‘eat (hard foods)’ is employed, in combination with an alternate

stem of the word for ‘eat (soft foods)’, *səl¹*, which is a cognate form with the Mueshaung *səl¹*. We thus see that stem alternation applies in these noun compounds even when the nominal stem preceded by the nominalizer *ə-* does not show alternation.

One further function of stem alternation is in noun-verb compounds where the object noun is the expected construction (a ‘cognate object’). Two examples from Mueshaung are shown in (14):

- (14) *pu²-sur³* ‘to milk’ literally ‘breast-squeeze.NOM’
tsəm¹-yən³ ‘to reap’ literally ‘paddy-reap.NOM’

In both of these examples the nominal stem is used, the verbal stems of these two verbs being *sur¹* ‘squeeze’ and *yən¹* ‘mow, reap.’

Examples comparing the use of the verb *sur¹* ‘squeeze’ and the nominal compound *pu²sur³* ‘milking’ are given in (15), (16) and (17), three sentences elicited from Wanglung Mossang¹⁵.

- (15) *yalxjuik* *saz* *rueq* *manx* *puix* *surz* *tueq.*
 jal²təuk sa¹ rɿʔ man² pu² sur¹ tɿʔ
 girl child AG cow breast squeeze PST.3
 ‘The girl milked the cow.’

- (16) *ngaz* *manx* *puixsurc* *kayz.*
 ŋa¹ man² pu²sur³ kai¹
 1SG cow milking go.NOMZ
 ‘I am going to milk the cow.’

- (17) *manx* *puixsurc* *ghvrc* *hvlc*
 man² pu²sur³ yər³ həl³
 cow milking difficult.NOMZ good
 ‘Milking the cow is very difficult.’

4 A comparison of verbal stems in two varieties: Ngaimong and Mueshaung

In Appendix 1, below, we present 151 verb roots that are found with both verbal forms, and nominal forms, preceded by the nominalizer *ə-*, in two Tangsa varieties, Ngaimong and Mueshaung. These were elicited as part of a larger word list and the tone categories of every word were checked with the two main consultants, Wangkui Ngaimong (Ngaimong) and Wanglung Mossang (Mueshaung).

As mentioned earlier, there is a regular process in Ngaimong of a significant pitch rise, where any root with a final stop or carrying Tone 1, whether a noun or a verb, is preceded by a prefix. This was first pointed out to me in 2011 by Wangkui Ngaimong, and at the time he regarded this as a ‘fourth tone’¹⁶. In this paper this variation in pitch is not treated as a change of tone category, and therefore not treated as verb stem alternation.

¹⁵ I cannot help but acknowledge the great help that my Facebook account has been in keeping contact with consultants in India and Myanmar, and being able to ask questions and get answers promptly! However there is no substitute for face to face fieldwork and further research will be needed on the syntactic processes being illustrated here.

¹⁶ Two examples of this will suffice. Wangkui Ngaimong was recorded pronouncing the words /p^hal¹/ ‘vomit’ and /pəiʔ/ ‘3rd person singular pronoun. In both cases the pitch falls, but in the case of the prefixed words the

The statistical data relating to the 151 verbs is presented in Table 7:

Table 7: Comparison of Verb Stem Alternation in Mueshaung and Ngaimong

Type of alternation	Number of examples	Percentage
No change in either variety	95	62.9%
Cognate change in both varieties	34	22.5%
Change in Mueshaung only, no change in Ngaimong	11	7.3%
Change in Ngaimong only, no change in Mueshaung	5	3.3%
Changes in both varieties, different processes	6	4.0%

As we can see in Table 7, nearly two thirds of verbs in the two varieties undergo no variation in the stems. These are listed below in Appendix 1, Table 13. Note that this number would be much lower if we included as verb stem alternation the pitch variation in prefixed words carrying Tone 1 or final stop (including glottal stop) in Ngaimong. No fewer than 56 of the verbs in the ‘no change’ category carry either Tone 1 or final stop. This means that most Ngaimong verbs have a perceptually different sound in the nominal stem, even though these are not here analysed as stem alternations. As far as we know, since this pitch raising is a phonological process, there are no raised pitches on un-prefixed stems, whereas, as we saw above in (9) and (10), true alternate stems can occur un-prefixed.

Note also that there were cases where there was no tone change in either variety for the verbal stem and its nominal counterpart, but there was tone change for the causatives. One example of this is the verb meaning ‘to be hot’. The data concerning this verb is presented in Table 8, where it is clear that both the nominal and verbal stems for the verb ‘be hot’ are *kal*² in both varieties, but that the causative forms in both varieties have an alternate stem carrying Tone 3, *kal*³.

Table 8: Verb stems for ‘to be hot’ in both Mueshaung and Ngaimong

English	Ngaimong		Mueshaung	
	Verbal	Nominal	Verbal	Nominal
be hot	<i>kal</i> ²	<i>ək</i> <i>al</i> ²	<i>kal</i> ²	<i>ək</i> <i>al</i> ²
make hot	<i>i</i> ³ <i>kal</i> ³		<i>i</i> ³ <i>kal</i> ³	

We have not been able to collect causative and/or middle, reciprocal and reflexive forms for all of the verbs that show no change in both varieties for the nominal and verbal stems, but there are clearly more of that pattern like *kal*². Other examples include the Ngaimong form *lim*² ‘be warm’, the nominal stem of which is *əlim*², showing no change of form, but the causative is *i*³*lim*³, showing a change to Tone 3, exactly parallel to the situation with ‘be hot’. However, in Mueshaung, only the phrase *rauŋ*² *lum*³, literally ‘sky warm’, meaning ‘warm

initial pitch is significantly higher. Hz measurements were made of his pronunciation of these words with and without prefixes, as follows:

<i>p</i> ^h <i>al</i> ¹	130-97Hz
<i>əp</i> ^h <i>al</i> ¹	170-102Hz
<i>pəi</i> [?]	133-110 Hz
<i>əpəi</i> [?]	193-173Hz

weather' has been recorded for this verb, and we are not able to say if the verbal root form carried Tone 2 as with the cognate in Ngaimong.

Another verb that has similar behavior is 'wake up', which has the verbal stem *səm'* and nominal stem *əsem'* in both Mueshaung and Ngaimong, again showing no stem alternation. However in Mueshaung, there are both causative and reflexive forms recorded, *təsəm³* and *kəsəm³* respectively, which show alternation of the stem to Tone 3.

Of the 34 verb roots which showed a cognate change in both varieties, listed in Appendix 1, Table 14, 28 involved the alternate, nominal stem, differing only in Tone category, with the nominal stem carrying Tone 3, and the verbal root carrying Tone 1 in 13 cases and Tone 2 in 15 cases. Tone 3 is the least frequently encountered tone in the Pangwa Tangsa varieties. As an example, of the 151 verb stems examined in this study, the numbers of verb stems carrying each of the tones is presented as (18):

- (18) Tone 1 – 56 tokens
 Tone 2 – 39 tokens
 Tone 3 – 28 tokens
 stop final – 28 tokens

Tone 3 is significantly more frequently encountered in the nominal forms of verbs, with, in Ngaimong, 63 nominal stems carrying Tone 3.

A small number of verbs have a cognate change which involves an alternation to whole coda. One such example is the verb 'to roast, to fry', as seen in Table 9. Note that final *-l* in Ngaimong often corresponds with final *-i* in Mueshaung.

Table 9: Verb stems for 'to roast, to fry' in both Mueshaung and Ngaimong

English	Ngaimong		Mueshaung	
	Verbal	Nominal	Verbal	Nominal
roast, fry	ʃu ¹	əʃuəl ¹	ʃu ¹	əʃoi ¹

In both varieties, the verbal stem is *ʃu¹*. In each case the nominal stem has a different, but cognate coda, both of which also carry Tone 1.

As already shown in Table 7, over 85% of the verb stems studied in the two varieties show the same behaviour in each, either no change to the stem, or a cognate change, usually a change of tone. However there were 22 verbs, a total of 14.6% of those which we recorded, where the patterns were different between the two varieties, either showing a change of stem in only one of the two varieties (see below Appendix 1, Tables 15 and 16) or a change in both but those changes being not cognate (see below Appendix 1, Table 17).

An example of this last type is the verb for 'cry' or 'weep', presented in Table 10:

Table 10: Verb stems for 'to cry, to weep' in both Mueshaung and Ngaimong

English	Ngaimong		Mueshaung	
	Verbal	Nominal	Verbal	Nominal
cry, weep	k ^h ok	ək ^h o ¹	k ^h oar ¹	ək ^h oar ³

Here we can see that two verbs, which appear to be cognate, behave quite differently. The tone alternation in Mueshaung is from a verbal stem carrying Tone 1 to a nominal stem

carrying Tone 3, a change found in many other Mueshaung verbs. The Ngaimong pattern is quite different, changing both the final consonant and the tone category.

Many examples of alternation were recorded in only one variety with a non-cognate word in the other. These were not included in the 151 roots mentioned earlier. While it is by no means a comprehensive list, Table 11 lists some of those meanings for which there are non-cognate stems in the two varieties, at least one of which shows stem alternation.

Table 11: Non cognate verb stems, in Mueshaung and Ngaimong, at least one of which shows verb stem alternations

English	Ngaimong		Mueshaung	
	Verbal	Nominal	Verbal	Nominal
split	p ^h əl ²	əp ^h əl ³	sal ²	əsəl ³
pull	huəm ²	əhuəm ³	het	əhet
go down	jəu ³	əjo ¹	sət	əset
touch	i ³ nəp	i ³ nəp	pem ¹	əpem ³
feel	ʃjen ²	kəʃjen ³	hel ³	əhel ³
embrace	pəm ²	əpəm ³	yoat	əyoat
know	ʃjen ²	əʃjen ³	tuu ¹	ətuu ¹

An interesting pair of related verbs, one showing no stem alternation and the other showing the same stem alternation pattern in both varieties, is presented in Table 12, comparing verbs meaning ‘have’ and ‘live, dwell’. In both varieties the verbal form is the same, carrying Tone 2. With the meaning of ‘have’, there is no stem alternation, but with the meaning of ‘live, dwell’, the nominal stem in both varieties carries Tone 1. This pattern of Tone 2 verbal stem and Tone 1 nominal stem is only recorded on one other verb, ‘laugh’ (see Appendix 1).

Table 12: Verb stems for ‘have’ and ‘live, dwell’ in both Mueshaung and Ngaimong

English	Ngaimong		Mueshaung	
	Verbal	Nominal	Verbal	Nominal
have	t ^h a ²	ət ^h a ²	t ^h a ²	ət ^h a ²
live, dwell	t ^h a ²	ət ^h a ¹	t ^h a ²	ət ^h a ¹

5 Verb Stem Alternation – a wider Pangwa Perspective

The kind of detailed comparison that has been possible between Mueshaung and Ngaimong needs to be done for each of the 30 or so Pangwa varieties, but at present that level of documentation has not been done for most of the varieties.

In Table 18 in Appendix 2 below, we present 13 cognate verbs across 8 varieties, including Mueshaung and Ngaimong. The verbs are arranged into four groups, and we can advance the following generalization about the patterns of verb stem alternation that we observed.

Group 1 consists of verbs that have a verbal stem that is either glottal stop final or carry Tone 1. The nominal stems are different in most cases, with the following patterns:

- Chamchang/Shecyü: Tone 1 / glottal stop verbal stems → Tone 2 nominal stems
- Mueshaung: Tone 1 / glottal stop verbal steps → Tone 1 nominal stems
- Lauchang: No stem alternation
- Rinkhu, Yvngban Wvng and Shangti: Tone 1 / glottal stop verbal stems → final stops in some of the nominal stems, usually *-t*.

This pattern for the nominal stems, is similar to the marked set of tone correspondences shown in Table 5 above.

Group 2 consists of verbs that all have verb roots carrying Tone 1 in Chamchang, Shecyü, Mueshaung, Lauchang and Ngaimong. In all of these varieties, and also for some of the verbs in Shangti, the nominal stem changes to Tone 3, except in Lauchang where it changes to Tone 2.

We cannot advance a clear explanation of why Lauchang is divergent, but it is interesting to note that the Tone 2 in Lauchang is a high falling tone, where Tone 3 in Chamchang and Shecyü is a high falling tone.

Group 3 consists of verbs that all have verb roots carrying Tone 3 in Chamchang, Shecyü, Mueshaung, Lauchang and Ngaimong. There is no stem alternation in Chamchang, Shecyü, Mueshaung and Ngaimong, but there is alternation to Tone 2 in Lauchang. There verbs also show Stem alternation in Rinkhu, but the pattern is reverse, the verbal stem carrying Tone 2 and the nominal (alternate) stem carrying Tone 3.

The verbs in Group 4 are quite variable, with no stem alternation for any of them in Chamchang, Shecyü and Lauchang, but some alternation on different roots in each of the other varieties.

The data presented in Table 18 shows that Chamchang and Shecyü are almost identical. These two varieties are fully mutually intelligible with only minor differences. Morey (2017) points out that many of the varieties are ‘paired’ with another very similar variety; Chamchang and Shecyü is an example of one such pair. On the basis of this set of 13 verbs, however, no other varieties show identical patterns to those two, and some show very different patterns. It is noteworthy that some of the verbs that show no alternation in Ngaimong and Mueshaung, do have it in other varieties, such as those in Group 3 in Lauchang and Rinkhu. Certainly much more detailed information about all these Pangwa varieties is needed to fully understand the variety of verb stem alternation.

6 Conclusion

Verb stem alternation is present on many stems in a range of the Tangsa / Tangshang varieties. Probably the most basic functional distinction between the two stems is that of verbal and nominal, with the verbal stem being demonstrated to be ‘base’ or underived form, at least in most cases.

In most cases the alternation is marked by a change of tone category; there are however examples in which the whole coda is changed, or the vowel and coda.

As with the Kuki-Chin languages, for which verb stem alternation has been well described, there are multiple cases where in one Tangsa variety there is an alternate stem carrying a different form, and in a related variety there is no alternation between the stems. In some cases it seems that alternation has been lost by means of re-analysis of the verbal form to take the same form as the derived nominal, as in the Lauchang form for ‘give’.

Some preliminary observations about the relationship between the stem alternation processes in the Tangsa varieties with other Tibeto-Burman languages can be made, as discussed in Section 2 above. This is also an area requiring much more research.

Our understanding of the Tangsa languages is still at a very early stage and further research in future years is likely to considerably advance the preliminary conclusions presented here.

References

- Bawi Tawng. (2017). *A typology of subordinate constructions in Lai*. MA Thesis. Department of Linguistics. Chiangmai, Payap University.
- Bedell, George. (2013). “The morphosyntax of verb stem alternation.” *Paper presented at the 46th International Conference on Sino-Tibetan Languages and Linguistics, Dartmouth College, Hanover, New Hampshire, USA, August 7-10.*
- Burling, Robbins and Mankai Wangsu. (1999). “Wancho phonology and word list.” *Linguistics of the Tibeto-Burman Area*. 21(2): 43–71.
- Chhange, Lalnunthangi. (1993). *Mizo Syntax*. PhD dissertation, Department of Linguistics. Eugene, University of Oregon.
- DeLancey, Scott. (2014). “What happened in Nagalang.” *Paper presented at the 1st meeting of the International Consortium for Eastern Himalayan Ethnolinguistic Prehistory. University of New England, Armidale, New South Wales, Australia. September 6-7.*
- _____. (2017). “Classical Tibetan.” In Graham Thurgood and Randy LaPolla, Eds. *The Sino-Tibetan Languages*, 2nd edition. London, Routledge: 369–384.
- Gong, Xun. (forthcoming). “Verb stems in Tangut and their orthography.” (Downloaded from http://www.academia.edu/33277842/Verb_stems_in_Tangut_and_their_orthography, accessed 07/11/2017.)
- Goswami, Dipjyoti. (2017). *Lexical nominalization and the extended functions of the demonstrative kəra² in Rera, a Northern Naga Language*. MA Thesis. Department of Linguistics. Chiang Mai, Payap University.
- Henderson, Eugénie J. A. (1965). *Tiddim Chin: a descriptive analysis of two texts*. (London Oriental Series #15). London, Oxford University Press.
- Hill, Nathan W. (2014). “Tibetan.” In Rochelle Lieber and Pavol Štekauer, Eds. *The Oxford Handbook of Derivational Morphology*. Oxford, Oxford University Press: 620–630.
- Jacques, Guillaume. (2014). “Denominal affixes as sources of antipassive markers in Japhug Rgyalrong.” *Lingua* 138: 1–22.
- _____. (2017). “Japhug.” In Graham Thurgood and Randy LaPolla, Eds. *The Sino-Tibetan Languages*, 2nd edition. London, Routledge: 614–634.
- Jacques, Guillaume, Aimee Lahaussais, Boyd Michailovsky and Dhan Bahadur Rai. (2012). “An overview of Khaling verbal morphology.” *Language and Linguistics* 13(6): 1095–1170.
- Khan Lann. (2017). *A phonological comparison of Shecyü and Mungre and its contribution to a common Tangshang Naga orthography*. MA Dissertation. Department of Linguistics. Chiang Mai, Payap University.
- King, Deborah. (2009). “Structural and pragmatic functions of Kuki-Chin verbal stem alternations.” *Journal of the Southeast Asian Linguistic Society (JSEALS)* 1: 141–157.
- Konnerth, Linda. (2014). *A grammar of Karbi*. PhD Dissertation. Department of Linguistics. Eugene, University of Oregon.
- Morey, Stephen. (2011a). “Tangsa agreement markers.” In Gwendolyn Hyslop, Stephen Morey and Mark W. Post, Eds. *North East Indian Linguistics, Volume 3*. Delhi, Cambridge University Press, India: 76–103.

- _____. (2011b). “Transitivity in Cholim Tangsa.” [Special issue]. *Studies in Language* 35(3): 676–701.
- _____. (2015). “The internal diversity of Tangsa: Vocabulary and morphosyntax.” In Mark W. Post, Stephen Morey and Scott DeLancey, Eds. *Language and culture in Northeast India and beyond: In honour of Robbins Burling*. Canberra, Australia-Pacific Linguistics: 23–40.
- _____. (2017). “Tangsa.” In Graham Thurgood and Randy LaPolla, Eds. *The Sino-Tibetan Languages*, 2nd edition. London and New York, Routledge: 350–368.
- Statezni, Nathan. (2013). *Fifty-five dialects and growing: Literacy and comprehension of vernacular literature among the Tangshang Naga in Myanmar*. MA Thesis. Dallas, Graduate Institute of Applied Linguistics.
- So-Hartmann, Helga. (2009). *A descriptive grammar of Daai Chin*. (STEDT Monograph 7). Berkeley, Sino-Tibetan Etymological Dictionary and Thesaurus Project, University of California.
- Sun, Jackson T.-S. (2004). “Verb-stem variations in Showu rGyalrong.” In Ying-chin Lin, Fang-min Hsu Hsu, Chun-chih Lee, Jackson T.-S. Sun, Hsiu-fang Yang Yang and Dah-an Ho, Eds. *Studies on Sino-Tibetan languages: Papers in honor of Professor Hwang-cherng Gong on his seventieth birthday*. (Language and Linguistics Monograph Series W4). Taipei, Institute of Linguistics, Academia Sinica: 269–296.
- VanBik, Kenneth (2009). *Proto-Kuki-Chin: A Reconstructed Ancestor of the Kuki-Chin Languages*. (STEDT Monograph 89). Berkeley, Sino-Tibetan Etymological Dictionary and Thesaurus Project, University of California.
- Weidert, Alfons. (1979). “The Sino-Tibetan tonogenetic laryngeal reconstruction theory.” *Linguistics of the Tibeto-Burman Area* 5(1): 49–127.
- _____. (1987). *Tibeto-Burman tonology*. Amsterdam, John Benjamins.

Appendix 1 – Comparison of verb stems in Ngaimong and Mueshaung

Table 13: No stem alternation in either variety

English	Ngaimong		Mueshaung	
	Verbal	Nominal	Verbal	Nominal
to fly	ba ¹	əba ¹	ba ¹	əba ¹
wait	bəm ³	əbəm ³	bəm ³	əbəm ³
beat with heavy object	bəu [?]	əbəu [?]	buu [?]	əbuu [?]
pain, ill	ða ¹	əða ¹	t̪uu ¹	ət̪uu ¹
untie	da ³	əda ³	dɔ ³	ədɔ ³
dead, die	ðəi ³	əðəi ³	t̪ui ³	ət̪ui ³
fall	dɛl ¹	ədɛl ¹	dəi ¹	ədəi ¹
jump, leap	dok	ədok	dok	ədok
kick	hap ¹	əhap ¹	hʉp	əhʉp
sour, acid	həi ²	əhəi ²	hi ²	əhi ²
damp, wet	həm ²	əhəm ²	hem ²	əhem ²
find	him ³	əhim ³	hum ³	əhum ³
groan	hin ³	əhin ³	hun ³	əhun ³
build	huɔn ³	əhuɔn ³	ɣur ³	əɣur ³
cooked	jəŋ ²	əjəŋ ²	jaŋ ²	əjaŋ ²
fan (vb)	jəp	əjəp	jəp	əjəp
come back	jiŋ ¹	əjiŋ ¹	jek	əjek
sow (seed), with stick	jiŋ ³	əjiŋ ³	juŋ ³	əjuŋ ³
run	ju ¹	əju ¹	ju ¹	əju ¹
flow	juɔl ¹	əjuɔl ¹	juŋ ¹	rəjuŋ ¹
swallow	juk	əjuk	juk	əjuk
pursue	jul ³	əjul ³	jur ³	əjur ³
hot	kal ²	əkəl ²	kal ²	əkəl ²
cool	kəi [?]	əkəi [?]	kui [?]	əkui [?]
believe	kəm ²	əkəm ²	kəm ²	əkəm ²
plant seedling	k ^h ai ³	ək ^h ai ³	k ^h i ³	ək ^h i ³
burn (intr.)	k ^h am ³	ək ^h am ³	k ^h am ³	ək ^h am ³
bitter	k ^h a [?]	ək ^h a [?]	k ^h ʉp	ək ^h ʉp
land (vb)	k ^h iŋ ²	ək ^h iŋ ²	k ^h iŋ ²	ək ^h iŋ ²
flay, skin	kok	əkok	k ^h ʉ ²	ək ^h ʉ ²

heavy	ləiʔ	ələiʔ	liʔ	əliʔ
look, look at	leŋ³	ələŋ³	lə³	ələ³
get, obtain	ləp	ələp	ləp	ələp
cook	lim³	əlim³	lum³	əlum³
bury	əlip¹	əlip¹	əlɯp	əlɯp
white	loŋ¹	əloŋ¹	lu¹	əlu¹
full	mel²	əmel²	mər²	əmər²
ripe	min²	əmin²	mɯn²	əmɯn²
yellow	mjen²	əmjen²	men²	əmen²
blow	mul¹	əmul¹	mui¹	əmui¹
extinguish	i³mut	i³mut	ətəmut	ətəmut
to thresh, stamping	na¹	əna¹	nɯu¹	ənɯu¹
black	ŋək	əŋək	ŋauk	əŋauk
smooth	ŋəl¹	əŋəl¹	ŋel¹	əŋel¹
grasp, seize	ŋəl³	əŋəl³	ŋəl³	əŋəl³
spoiled, rotten	nəm²	ənəm²	nəm²	ənəm²
soft (to the touch)	ŋen¹	əŋen¹	ŋen¹	əŋen¹
sit	ŋəŋ³	əŋəŋ³	ŋaŋ³	əŋaŋ³
enter	ŋip	əŋip	ŋɯp	əŋɯp
grow	pa¹	əpa¹	pe¹	əpe¹
vomit	pʰal¹	əpʰal¹	pʰai¹	əpʰai¹
float	pʰoŋ¹	əpʰoŋ¹	pʰu¹	əpʰu¹
full (from eating)	pʰuəl¹	əpʰuəl¹	pʰoai¹	əpʰoai¹
drive	rip	ərip	rɯp	ərɯp
turn around	kəro¹	kəro¹	təro¹	ətəro¹
clean	sa³	əsə³	sə³	əsə³
dry	san²	əsən²	ʃən²	əʃən²
hire	ʃap	əʃap	ʃap²	əʃap²
pour	səl¹	əsəl¹	ʃəi¹	əʃəi¹
drip	səl³	əsəl³	sər³	əsər³
wake up	səm¹	əsəm¹	səm¹	əsəm¹
red	seŋ²	əseŋ²	ʃə²	əʃə²
pierce, stab	səuʔ	əsəuʔ	əʃuʔ	əʃuʔ
sharp	ʃi²	əʃi²	tʰe¹	ətʰe¹

urinate	ʃil ¹	əʃil ¹	ʃi ¹	əʃi ¹
work (v.)	ʃim ³	əʃim ³	ʃum ³	əʃum ³
suck, kiss	ʃip	əʃip	ʃup	əʃup
dye	ʃjem ³	əʃjem ³	tʃa ¹	ətʃa ³
smell (vb trans)	ʃjeŋ ³	əʃjeŋ ³	ʃe ³	əʃe ³
cough	so ¹	əso ¹	sau ¹	əsau ¹
cut with knife	ʃuɔl ³	əʃuɔl ³	soal ³	əsoal ³
dirty, soiled	sul ¹	əsul ¹	ʃui ¹	əʃui ¹
hear	tal ¹	ətal ¹	tai ¹	ətai ¹
hard	tean ¹	ətean ¹	tsan ¹	ətsan ¹
winnow (up and down)	teap	əteap	teap	əteap
weave	tək	ətək	tauk	ətauk
bite	tən ³	ətən ³	tən ³	ətən ³
light (in weight)	teŋ ¹	əteŋ ¹	teɔ ¹	əteɔ ¹
carry-underarm	teɐp	əteɐp	tsəp	ətseɐp
stand	teəp	əteəp	teəp	əteəp
blunt, dull	təu ¹	ətəu ¹	tuu ¹	ətuu ¹
measure	t ^h a ¹	ət ^h a ¹	t ^h a ¹	ət ^h a ¹
have	t ^h a ²	ət ^h a ²	t ^h a ²	ət ^h a ²
put	t ^h eŋ ³	ət ^h eŋ ³	t ^h ɔ ³	ət ^h ɔ ³
raw, unripe	t ^h iŋ ²	ət ^h iŋ ²	t ^h uŋ ²	ət ^h uŋ ²
drowned	t ^h ip	ət ^h ip	t ^h uɐp	ət ^h uɐp
saliva, spit	t ^h o ¹	ət ^h o ¹	t ^h o ¹	ət ^h o ¹
to dig	t ^h uɔl ¹	ət ^h uɔl ¹	t ^h oai ¹	ət ^h oai ¹
wrap with cloth or leaf	t ^h uɔp	ət ^h uɔp	t ^h op	ət ^h op
swim	vai ³	əvai ³	βəi ³	əβəi ³
plait, weave basket	val ¹	əval ¹	βai ¹	əβai ¹
bore a hole, carve	vət	əvət	βət	əβət
green	vil ¹	əvil ¹	βul ¹	əβul ¹
hold (in mouth)	vim ³	əvim ³	βum ³	əβum ³

Table 14: Cognate stem alternation process in both varieties

English	Ngaimong		Mueshaung	
	Verbal	Nominal	Verbal	Nominal
shake	bil ¹	əbil ³	bul ¹	əbul ³
smell (vb intrans)	ðiŋ ²	əðiŋ ³	ʈuŋ ²	əʈuŋ ³
climb	doŋ ¹	ədoŋ ³	du ¹	ədu ³
wash (hand, body, dishes)	huɔl ¹	əhuɔl ³	hoal ¹	əhoal ³
carry (bear)	hul ²	əhul ³	ɣur ²	əɣur ³
go	ka ¹	əkai ¹	ka ¹	əkai ¹
mow, reap	k ^h ən ¹	ək ^h ən ³	ɣən ¹	əɣən ³
walk	k ^h uɔm ²	ək ^h uɔm ³	k ^h ɣm ²	ək ^h ɣm ³
see	k ^h ui ¹	ək ^h ui ³	k ^h i ¹	ək ^h i ³
roast, fry	kuɔl ²	əkuɔl ³	koar ²	əkoar ³
catch (ball)	lu ¹	əluɔl ¹	lu ¹	əloj ¹
laugh	nəi ²	ənii ¹	nui ²	ənui ¹
drink	niŋ ²	əniŋ ³	nuŋ ²	ənuŋ ³
tired, weary	njeŋ ²	ənjeŋ ³	je ²	əje ³
carry-on-shoulder	pai ²	əpai ³	pi ²	əpi ³
drunk	peŋ ²	əpeŋ ³	pɔ ²	əpɔ ³
sweep with broom	peŋ ²	əpeŋ ³	pe ²	əpe ³
follow	p ^h ən ²	əp ^h ən ³	p ^h ən ²	əp ^h ən ³
distribute, share	p ^h əl ²	əp ^h əl ³	p ^h əl ²	əp ^h əl ³
flood (v.)	piŋ ²	əpiŋ ³	puŋ ²	əpuŋ ³
chew	sau ¹	əsau ³	ʃo ¹	əʃo ³
lead (vb)	səl ²	əsəl ³	ʃəl ²	əʃəl ³
sell	səŋ ¹	əsəŋ ³	ʃauŋ ¹	əʃauŋ ³
roast, fry	ʃu ¹	əʃuɔl ¹	ʃu ¹	əʃoai ¹
squeeze, wring	sul ¹	əsul ³	sur ¹	əsur ³
taste	tal ²	ətal ³	tal ²	ətal ³
beat lightly	təl ²	ətəl ³	təl ²	ətəl ³
tbirth, born (to be)	teŋ ²	əteŋ ³	teɔ ²	əteɔ ³

live, dwell	t ^h a ²	ət ^h a ¹	t ^h a ²	ət ^h a ¹
tear (vb)	tɛi ¹	ətɛi ³	tɛur ¹	ətɛur ³
lift, rise	tuɔn ¹	ətɔn ³	ten ¹	əten ³
pick up	tɛuɔn ¹	ətɛuɔn ³	tsɤn ¹	ətsɤn ³
count	vəl ¹	əvəl ³	βəl ¹	əβəl ³
go up ascend	vəŋ ¹	əvəŋ ³	βauŋ ¹	əβauŋ ³

Table 15: Stem alternation in Mueshaung but not Ngaimong

English	Ngaimong		Mueshaung	
	Verbal	Nominal	Verbal	Nominal
crouch	bim ³	əbim ³	bum ²	əbum ³
bite	kək	əkək	kauk	əkɯu ¹
give	koʔ	əkoʔ	kuʔ	əkui ¹
roll (intr.)	kuɔl ¹	əkɯɔl ¹	ku ²	əkɯ ³
eat	p ^h ək	əp ^h ək	p ^h auk	əp ^h ɯu ¹
eat	saʔ	əsəʔ	saʔ	əsəi ¹
close, shut	sək	əsək	ʃauk	əʃɯu ¹
pound with fist	t ^h əu ¹	ət ^h əu ¹	t ^h ɯu ¹	ət ^h ui ¹
plow	t ^h uɔl ¹	ət ^h uɔl ¹	t ^h ɛi ³	ət ^h ɛi ³
fragrant	tɛim ²	ətɛim ²	tum ²	ətum ³
push, shove	t ^h u ²	ət ^h u ²	t ^h uʔ	ət ^h u ²

Table 16: Stem alternation in Ngaimong but not Mueshaung

English	Ngaimong		Mueshaung	
	Verbal	Nominal	Verbal	Nominal
pull	huɔm ²	əhuɔm ³	het	əhet
lazy	joŋ ³	əjoŋ ¹	jo ¹	əjo ¹
lick	mjel ¹	əmjel ³	mel ¹	əmel ¹
comb (v.)	ʃil ¹	əʃil ³	t ^h i ¹	ət ^h i ¹
arrive, reach	t ^h əu ³	ət ^h o ¹	t ^h ɯu ³	ət ^h ɯu ³

Table 17: Cognate words with different stem alternation processes in both languages

English	Ngaimong		Mueshaung	
	Verbal	Nominal	Verbal	Nominal
kill	ðəi ³	i ³ ði ¹ , i ³ ðəi ³	t̪ət ɣai ¹	ət̪əi ¹ ɣai ¹
cry, weep	k ^h ok	ək ^h o ¹	k ^h oar ¹	ək ^h oar ³
bring	ləm ²	ələm ³	ləm ¹	ələm ³
take (M: 'take and go')	ləm ²	ələm ³	laʔ	ələi ¹
flaming	loŋ ¹	əloŋ ³	lu ³	
buy	ri ²	ərjel ¹	rui ²	ərə ¹
burn (tr.)	ʃu ¹	əʃuəl ¹	təuk	ətɛru ¹

Appendix 2 – Comparison of verb stems in 8 Tangsa varieties

Table 18: Examples comparing verbal and nominal stems in 8 Tangsa varieties

	Chamchang		Shecyü		Mueshaung		Lauchang		Ngaimong		Rinkhu		Yangban Vang		Shangti	
	verbal	nominal	verbal	nominal	verbal	nominal	verbal	nominal	verbal	nominal	verbal	nominal	verbal	nominal	verbal	nominal
1)																
roast			fu ¹	əfe ²	fu ¹	əfoi ¹	fau ¹	i-fau ¹	fu ¹	əfə ¹	fu ²	ɪfot			fy ¹	nfot
give	ku?	əke ²	ku?	əke ²	ku?	əkui ¹	kau ¹	i-kau ¹	ko?	əko?	ku?	iku?	ko?	ikut	k ^h y?	nk ^h y?
eat (rice)	si?	əsɛə ²	si?	əsiə ²	sa?	əsəi ¹	sə?	i-sə?	sa?	əsa? / əsəl ¹	sa?	ɪsa?	sa?	isa?	se?	nse?
go	ki ¹	əkai ²	ki ¹	əkai ²	ka ¹	əkai ¹	ki ¹	i-ki ¹	ka ¹	əkai ¹	ka ¹	ɪkat	ka ¹	ika?	k ^h e ¹	nk ^h et
2)																
see	tɛ ^h i ¹	ətɛ ^h i ³	tɛ ^h i ¹	ətɛ ^h i ³	k ^h i ¹	ək ^h i ³	k ^h u ¹	i-k ^h u ²	k ^h ui ¹	ək ^h ui ³		ɪk ^h i ³	koa ³		k ^h i ¹	nk ^h i ³
wash	he ¹	əhe ³	he ¹	əhe ³	həl ¹	əhəl ³	hau ¹	i-hau ²	həl ¹	əhəl ³	hu ²	ihu ³	hut	ihut	hən ³	nhən ³
wash clothes					hui ¹	əhui ¹			hul ¹	əhul ¹	het	ihet			hat	nhat
3)																
sit	ŋa ³	əŋa ³	ŋa ³	əŋa ³	ŋauŋ ³	əŋauŋ ³	ŋa ³	i-ŋa ²	ŋəŋ ³	əŋəŋ ³	ŋaŋ ²	ŋaŋ ³	təo?	iteo?	ŋaŋ ³	ŋŋaŋ ³
look	lɪ ³	əlɪ ³	lɪ ³	əlɪ ³	lɔ ³	əlɔ ³	lau ³	i-lau ²	leŋ ³	ələŋ ³	la ²	ila ³	la ³	ila ³	laa ³	nlaa ³
4)																
steal	huk	əhuk	huk	əhuk	ɡuu?	əɡuu?	yauk		hauk	əhauk	ɡu?	ɪɡu?	ɡut / ɡu?	igut	ɡu?	ŋɡu?
bite	ka?	əka?	ka?	əka?	kauk	əkau ¹	ka?	i-ka?	kək	əkək	kak	ɪkak			k ^h ak	nk ^h ak
laugh	nai ²	ənai ²	nai ²	ənai ²	nvi ¹	ənvi ¹	ŋai ¹	i-ŋai ¹	nai ²	ənai ¹	ni ²	mit			ni ²	nnik
run / flee	ʒu ¹	əʒu ¹	ju ¹		ju ¹	əju ¹	ʒau ¹	i-ʒau ¹	ʒu ¹	əʒu ¹	jok	ɪjok	jo ¹	ijo ³	jwa?	njwa?

Nominalizations in Bodo¹

Prafulla Basumatary
Gauhati University

Abstract

This paper discusses the behavior of nominal structures in Bodo. It will demonstrate that Bodo has two nominalization processes, namely, derivational and clausal nominalization. In derivational nominalization, it will be shown that a lexical noun is derived mainly from verbs and to some extent from nouns and adjectives by suffixation of one of -t'ai, -t'i, -sula/suli, -zalu/-zali, -giri, -gra, -ari, -ma, and -sa. In clausal nominalization, a noun phrase is derived from a clause, often by suffixation of -nai with the verb. In terms of distribution and function, it will be shown that clausal nominalizations are found in attributive phrases, complementation, relative clauses, adverbial clauses and independent clauses.

Citation	Basumatary, Prafulla. 2018. <i>Nominalizations in Bodo</i> . North East Indian Linguistics (NEIL) 8, 110-125. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.
Volume Editors	Linda Konnerth, Stephen Morey, Amos Teo
Copyright	© 2018, the author(s), release under Creative Commons Attribution license
URL	http://hdl.handle.net/1885/155280

1. Introduction

Nominalization is a prominent characteristic of Tibeto-Burman languages (Genetti 2011, DeLancey 2011). It refers to the process by which we derive nominal expressions (Comrie & Thompson 1985), for example, from verbs or adjectives. Two types of nominalization are attested in world's languages – lexical or derivational and clausal. Derivational nominalization takes as its domain the verb root and works at the morphological level to derive lexical nouns, whereas clausal nominalization takes its domain the clause or clause combination and works at the syntactic level to allow a grammatical clause to be treated as a noun phrase within a broader syntactic context (Genetti 2010). Nominalized clauses cover a wide range of functions in most Tibeto-Burman languages, such as complements of a verb, dependents in possessive constructions, relative clauses, adverbial clauses, and so on. Nominalization is also a source for new finite morphosyntax in Tibeto-Burman languages. Nominalized clauses often replace a whole finite system, and they start functioning as finite clauses over time (DeLancey 2011). Bodo, a Tibeto-Burman language spoken in Assam in the North-East India exhibits most of these functions found in other Tibeto-Burman languages.

Both of these nominalization processes are found in Bodo. We will describe and discuss the common issues in both the nominalization processes as they are reflected in Bodo as well as semantic and morphosyntactic features which are peculiar to Bodo.

The current paper is organized into the following sections. §2 gives a typological overview of the language. §3 talks about the data and language variety on which this study is based. §4 will define and describe the subject matter, i.e. the lexical nominalization as

¹ This paper would not have been possible without the help and guidance of Krishna Boro, Gauhati University, Guwahati. I also sincerely convey my thanks to the anonymous reviewer whose suggestions were immensely helpful in improving this paper. I am also thankful to the editor Stephen Morey, La Trobe University, Australia for his valuable suggestions during the whole stage of the revision of the paper. I owe a lot to Prof. Jyotiprakash Tamuli, Gauhati University and Prof. Scott DeLancey, University of Oregon for their constant support and encouragement while preparing this paper. All errors are my own.

exhibited in Bodo followed by a detailed discussion on clausal nominalization in §5 and §6. §7 will conclude the paper.

2. Typological overview

Bodo is a verb final language with a clause chaining structure. Several clauses marked with a non-final or medial morpheme on the verb can be concatenated along with a final clause which contains the finite verb. It has serial-verb constructions, which we often see in clause chaining languages (Boro 2012). The dependent clauses generally precede the matrix clause. However, the word-order is flexible to a great extent. For instance, the subject is very often post-posed to the end of the sentence. The dependent clause may also follow the matrix clause. Similarly, the nominal modifiers such as numerals and adjectives can either precede or follow the head noun easily. However, demonstratives always precede the head noun, and so do relative clauses. The word-order variations are governed by pragmatic principles, which are not yet systematically described in the language. Bodo has postpositions, which are mostly case markers. The case marking of core arguments is pragmatically conditioned, i.e. the case markers carry pragmatic interpretations of various sorts besides carrying the information of grammatical relations. For example, overt case marking often codes ‘identifiability’ or ‘referentiality’ whereas a lack of overt case marking expresses a non-referential participant. The postpositions are cliticized to the final morpheme of the noun phrases. The subordinators also follow the clause which they mark as dependent clauses. They are phonologically bound to their dependents.

3. Data and the language variety

Bodo has several distinct regional varieties, but these are intelligible to a great extent among the speakers of different dialects. One of these dialects is regarded as standard and it is taught at school, used in the media, and in literature. This dialect is mainly spoken in the district of Kokrajhar though its speakers are found to be spread in other districts as well. I am a speaker of this dialect and this is the variety I am working on for this paper.

The current study is based on a 1.2 million word corpora comprising of different genres both written and spoken. The linguistic data were collected from Udalguri and Kokrajhar districts during 2009-2015 partly for my Ph.D. work (Basumatary 2015) and partly for the project called ‘Shallow Parser Tools for Indian Languages’ at Gauhati University. The recorded materials will be ultimately archived at the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC, see <http://paradisec.org.au>)². The analysis is done using the SIL FLEx programme which allows concordance search of texts. The author draws most of the illustrative examples from this corpus. The author also uses constructed examples where natural data lack instances of certain constructions. The examples drawn from the corpus have tags in parentheses which identify them in the corpus. The constructed examples do not have such tags. In Bodo, tone has a very low functional load, so we are marking only the high tone on the words which have a low tone counterpart. In other words, words with low tone and words that are not found to have minimal pair counterparts are not marked for tone.

² 23 sound files have already been archived with PARADISEC so far.

4. Derivational or lexical nominalization

In Bodo, a lexical noun is derived mainly from verbs, adjectives and nouns by addition of one of around nine or ten nominalizing suffixes. The derived nouns can head an NP and they can take the usual nominal morphology. Some of these nominalizing suffixes such as *-giri*, *-gra* and *-nai* are very productive whereas *-t^hai*, *-t^hi*, *-sula*, *-zalu*, *-ari* and *-sali* are not productive. Table 1 shows the nominalizing suffixes that exist in the language, type of nominalization and their sources.

Table 1 – Nominalizers, type of nominalization and their source

Nominalizer	Type of nominalization	Derives from
<i>-t^hai</i>	Product/result	Verb
<i>-t^hi</i>	Product/result	Verb, adjective, noun
<i>-sula/-suli</i>	Agent/patient	Verb
<i>-zalu/-zali</i>	Agent/patient	Verb
<i>-giri</i>	Agent/patient	Verb, noun
<i>-gra</i>	Agent/patient/instrument	Verb
<i>-ari</i>	Agent/patient	Verb
<i>-sali</i>	Location	Verb, adjective, noun
<i>-nai</i>	Action	Verb

These nominalizers will be discussed one by one in the following sections.

Nominalizer *-t^hai*

In Bodo, the suffix *-t^hai* mostly derives abstract nouns that can head an NP. For example, the word *nut^hai* ‘scene/sight’ is derived from the word *nu* ‘see’ by suffixation of nominalizer *-t^hai* as in (1).

- (1) ...*unau* *bauk^hungri* *hazuu-ni* *nu-t^hai-k^hou* *nu-nu*
 ...later PN hill-GEN see-NMLZ-OBJ see-INF
 mun-u.
 get-AFF
 ‘... later (we) could see the scene of the Baukhungri hill.’
 (Sandw Baodia, 145:2.11)

Nominalizer *-t^hi*

Like *-t^hai*, *-t^hi* also mostly derives abstract nouns that can head an NP. In (2), the word *buhuit^hi* ‘flow’ is derived from the word *buhui* ‘to flow’.

- (2) *bazui dodere-ni* *solo-ni* *buhui-t^hi-ja* *ese* *zuda-mun.*
 PN PN-GEN novel-GEN flow-NMLZ-SUB a little different-PAST
 The flow of the novel of ‘Bajwi Dodere’ was a little different.’
 (Kharokhandai, 1044:116.1)

-t^hi additionally derives abstract nouns from adjectives that can head an NP. In (3), *sanṅraṅt^hi* ‘awareness’ is derived from the adjective *sanṅraṅ* ‘aware/conscious’.

- (3) *abhawa sangray-t^{hi}-ni suluwɛt^{hai} boinibu*
 environment aware-NMLZ-GEN education everyone
gunanɛ-t^{har}.
 necessary-definitely
 ‘The education of awareness on environment is necessary for all.’
 (Baukhungri, 1136: 17.2)

Further, *-t^{hi}* sometimes derives abstract nouns from other nouns. For example, in (4), the word *mohort^{hi}* ‘image’ is derived from noun *mohor* ‘look/appearance’.

- (4) *be shib mohadeb-ni mohor-t^{hi}-a-nuu sijou.*
 this PN PN-GEN image-NMLZ-SUB-FOC plant.type
 ‘The Sijou (plant) is the actual image of Shiv Mahadeb.’
 (Bathou Dwhwrwmni Sayao, 422: 2.47)

Nominalizer *-giri*

The nominalizing suffix *-giri* derives agent nominals from verbs and they can head NPs. In (5), the word *phut^huigiri* ‘one who is responsible for death’ is derived from *p^hu-t^hui* ‘cause to die’.

- (5) *boro bizuwɛ-ni gidir p^hu-t^hui-giri-ja boro-a-nuu.*
 PN-SUB medium-GEN big CAUS-die-NMLZ-SUB PN-SUB-FOC
 ‘The Bodos (themselves) are mainly responsible for the death of Bodo medium (education).’ (Boro Thunlai Afad, 540:12.1)

In (5), *phut^huigiri* has the sense of ‘killer’ and the whole sentence conveys that the killers of Bodo medium education are none other than the Bodos themselves³.

-giri also derives agent nominals from other nouns illustrated below. In (6) the word *mohorgiri* is derived from *mohor* ‘look/appearance’.

- (6) *be ... aida-jau jiu zuub-zasim nanɛt^hab-nanui*
 this topic-LOC life end-till engage-NF
hanthi-lan-nai gedema mohor-giri-a-nuu za-duwɛ
 live.through-DIST-NMLZ renowned image-NMLZ-SUB-FOC happen-REAL
doctor s^hob^ha brohmo.
 doctor PN PN
 ‘In this regard the one who really lived through his life (as a dedicated artist) was the renowned sculptor Dr. Shobha Brahma.’ (Lalit Kala Ratna Dr. Shobha Brahma, 1021: 1.2)

This nominalizing suffix is diachronically related to *-gra* discussed below.

Nominalizer *-gra*

Like *-giri*, *-gra* also derives nouns from verbs. It derives both agent and patient nominals. The two suffixes seem to be historically related. Synchronically, *-giri* is used with limited verbs and has a formal sense and seems to have developed later for academic purposes. *-gra*, on the

³ Many Bodo intellectuals and parents did not value Bodo medium education and chose to send their children to English medium schools.

other hand, is very productive and has a less formal sense. In (7), *luṅgra* ‘one who drinks’ is derived from *luṅ* ‘to drink’.

- (7) *gubaṅ saha luṅ-gra-p^hur hab-duṅ aru uṅk^har-duṅ.*
 many tea drink-NMLZ-PL enter-REAL and come.out-REAL
 ‘Many tea drinkers are entering and coming out.’
 (Dangduli Saha Gola, 100:14.4)

- (8) *zó-gra-ja k^hibu k^heb-u.*
 sit-NMLZ-SUB hip squeeze-AFF
 ‘The stool squeezes (my) hip.’
 (Gwmanai Diary, 323: 2.92)

In (8), *zógra* ‘a seat’ is derived from *zó* ‘to sit’.

Nominalizer *-ari*

-ari mostly derives adjectives from nouns, for example, *gamiari* ‘of or relating to a village’ is derived from *gami* ‘village’ and *somazari* ‘of or relating to society’ is derived from *somaz* ‘society’. Some of the derived adjectives such as *gamiari* ‘people belonging to a village’ derived from *gami* ‘village’ can head an NP as shown in example (9)

- (9) *suluṅt^hai abhawa-k^hou gami-p^hur-au gami-ari-p^hur-a-nu*
 education environment-OBJ village-PL-LOC village-NMLZ-PL-SUB-FOC
t^hi ruk^hum-ui lak^hi-naṅ-gou.
 stable kind-ADVZ keep-need-AFF
 ‘The villagers (themselves) need to keep the educational atmosphere healthy in the village.’ (Swlwngthai Abhawa, 1025: 1.11)

-ari is also sometimes used to derive nouns from verbs. For example in (10), *bibai-ari* ‘beggar’ is derived from *bibai* ‘to beg’.

- (10) *bibai-ari-p^hur-nu dan hu-juu-bla burdan mún-p^hin-u.*
 beg-NMLZ-PL-DAT alms give-AFF-if blessing get-again-AFF
 ‘If (you) give alms to the beggars (you) get blessings in return.’
 (Bibayari, 517: 1.1)

Nominalizer *-sula/-suli*

-sula derives a noun that has masculine properties from a verb. Similarly *-suli* derives a noun that has feminine properties⁴ from a verb. In (11), *sibi* means ‘to worship or practice’; *sibi-sula* has the sense of ‘a worshipper/practitioner (male)’.

⁴ The final *-i* marking feminine in this case is not a borrowing from Assamese, though there are a few instances where feminine is formed in Bodo in a similar manner to Assamese with final *-i* as in *p^hant^ha* ‘male goat’, *p^hant^hi* ‘female goat’, *k^hana* ‘blind-male’, *k^hani* ‘blind-female’, etc.

- (11) *harimu sibi-sula-p^hur-ni biban-a-bu baŋ-bai.*
 culture practice-NMLZ-PL-GEN responsibility-SUB-FOC be.more-PERF
 ‘The responsibility of the cultural practitioners has become more.’
 (Boro Thunlai Jouganayau, 797:2.50)

The nominals derived through this process are mostly used as adjectives syntactically.

Nominalizer *-zalu/-zali*

Like *-sula* and *-suli* above, *-zalu* and *-zali* also derive nouns of masculine and feminine properties respectively from verbs. The difference between *-sula/-suli* and *-zalu/-zali* is that the use of the latter is found to be very restricted such that we have discovered its use only with *ón* ‘to love’ in our corpus. For example, in (12), *ónzalu* ‘lover’ is derived from *ón* ‘to love’.

- (12) ... *gao-ha gubaŋ ón-zalu-p^hur doŋ.*
 herself-LOC many love-NMLZ-PL exist
 ‘(She says) she has many lovers.’ (Gelethi, 354:15.2)

Nominalizer *-sali*

-sali derives a locational noun from verbs, adjectives and other nouns. In (13) the word *baósali* ‘a place of worship’ is derived from *baó* ‘to offer’, *gut^huisali* ‘cemetery’ is derived from *gut^hui* ‘dead’ in (14) and *habasali* ‘place of wedding’ is derived from *haba* ‘wedding’ in (15).

- (13) *raizuu-ni baó-sali-ao aroz gab-k^haŋ-nai-ni un-ao ...*
 public-GEN offer-NMLZ-LOC prayer cry-after-NMLZ-GEN later-LOC
 ‘After offering prayers at the village temple...’
 (Bhutanni Nareng Khomla, 776:2.20)
- (14) *boro-ni nem-zuŋ maoria-k^hou gut^hui-sali-sim gar-hui-a-sui.*
 PN-GEN ritual-INSTR PN-OBJ dead-NMLZ-ALL throw-DIST-AFF-CS
 ‘They disposed of Maoria at the cemetery following the rituals of the Boros.’
 (Abwi Abouni Solo, 6:2.148)
- (15) ... *gami-ni gasui-bu haba-sali-au nuza-t^hi-p^hui-duŋ.*
 village-GEN all-FOC wedding-NMLZ-LOC appear-clearly-PROX-REAL
 ‘All the villagers showed up at the wedding place.’ (Swrjini Maya, 495:2.121)

If we observe, we will notice that all the derived nominals in the above examples function like any other nouns in that they all have case markers attached to them following the nominalizing suffixes.

Predicate nominalization:

Bodo, as mentioned in the beginning, exhibits predicate nominalization (V+Object) of the derived nominals and it normally takes the order Object+V as exemplified by (16).

- (16) *zuŋ [taigar hil-ni somaina san raiziŋ nu-t^hai-k^hou*
 1PL PN PN-GEN beautiful sun rising see-NMLZ-OBJ
nu-duŋ-muun].
 see-REAL-PAST
 ‘We saw the beautiful sun-rising scenery of the Tiger Hill.’
 (Hajw Rani Darjeeling, 743:2.53)

In (16), the predicate construction is within the square bracket in which the object *nu-t^hai-k^hou* ‘scenery’ is followed by the verb *nu* ‘see’.

5. Clausal nominalization and nominalizers

A nominalized clause is defined as a clause or clauses which are somehow turned into a noun phrase, both functionally as well as morpho-syntactically.

Let us consider a prototypical nominalized clause embedded in a sentence, to begin with. The following sentence is made up by the author mainly because examples found in the natural texts are often very long and contains lots of complex structures.

- (17) *muider-a p^hui-bai.*
 elephant-SUB come-PERF
 ‘The elephant has come.’
- (18) [*muider-a p^hui-nai-k^hou*] *raizuu-p^hur-a mit^hi-ja.*
 elephant-SUB come-NMLZ-OBJ public-PL-SUB know-NEG
 ‘The villagers do not know that the elephant has come.’

Here we are talking about a particular elephant, and the hearer knows which elephant we are talking about, i.e. the elephant is identifiable. The villagers are not aware of the fact that the elephant has come. In (17) we have an intransitive sentence, which contains the subject argument *muider* ‘elephant’ and the finite verb *p^hui* ‘come’ marked with a perfective marker *-bai*. In (18) the same intransitive clause in (17) is embedded in a larger sentence and the intransitive clause functions as the object argument of the verb *mit^hi* ‘know’. This results in a few changes in the intransitive clause found in (17). First, the whole clause is marked with an object marker *-k^hou*. Second, the intransitive clause loses its finiteness, and acquires a new morpheme *-nai*, which we will call nominalizer. In other words, we have turned the intransitive clause in (17) into an NP and used it as an object of the verb *mit^hi* ‘know’. However, nominalized clauses do several other functions besides that of complementation, as mentioned in the beginning.

There are two clause nominalizers in Bodo, namely, *-nai* and *-nuu*. Nominalization with *-nai* is very productive, whereas nominalization with the other nominalizer *-nuu* is very restricted in terms of both structural and functional diversity. It is restricted to same subject clauses and it mostly functions as purpose clause. For example,

- (19) *goiti-a na gur-nuu p^hui-duŋ.*
 PN-SUB fish catch-NMLZ come-REAL
 ‘Goiti has come to catch fish.’
 (Gwdan Phwisali, 809:3.4)

In (19), the nominalized clause *gurnuu* ‘to catch’ is expressing the purpose of the event of ‘coming’ denoted by the matrix verb *p^hui* ‘come’, i.e. she has moved herself in order to catch fish. Both the clauses *gur-nuu* ‘catch-NMLZ’ and *p^hui* ‘come’ have the same subject *goiti*.

Nominalized clauses with *-nai* contain most of the variations in structure and function. We will enumerate all the different structures and functions involved in the *-nai* nominalized clauses.

5.1. Structure of Nominalized Clauses with *-nai*

The wide range of structural variations found in the clauses with this nominalizer ranges from a highly finite clause structure to a highly noun-phrase like structure (Boro, unpublished). We will identify three different structures within the nominalized clauses with *-nai*, namely, ‘highly finite structure’, ‘raised object structure’, and ‘NP like structure’. All the three structures will be discussed one by one below.

5.1.1. Nominalized clause with ‘highly finite structure’

In its most finite form, the nominalized clause with the nominalizer *-nai* contains a fully coded argument structure. All arguments within the relative clause are assigned their respective case roles with appropriate postpositions. However, unlike an independent clause the nominalized clauses lack finite verb morphology, such as tense/aspect-polarity markers. Arguments are often dropped in narratives where they can be easily retrieved from the discourse. Consider the following example.

- (20) ... *dok^hinapat* *sotro-ni* *sotrad^hikar-a* [*godad^hor* *siyho-a*
 PN PN-GEN chief-SUB PN PN-SUB
razulap^had-au *gak^hu-nai-k^hou*] *sab-za-jak^hui-mun.*
 throne-LOC climb-NMLZ-OBJ like-PASS-NEG.PERF-PAST
 ‘The chief of the Dakhinapat Satra didn’t like the coronation of Godadhar Singha as a king.’ (Moamoria Bigrainai, 1014: 5.4)

5.1.2. Nominalized clause with ‘subject to object raising’

The above examples show us that the internal structure of a clause remains intact when the clause is nominalized and embedded in another clause. However, modification in the internal structure of the nominalized clause is also often found. The phenomenon that we are going to look at next is often called ‘raising to object’, in which an agent participant of an embedded clause is coded as the patient of a higher clause. In terms of finiteness, the nominalized clauses with ‘subject to object raising’ are less finite than the one above in that it lacks subject argument. Consider the following example.

- (21) *at^hik^hal-au* *bi-k^hou* *p^hui-nai* *nu-lia.*
 nowadays-LOC he-OBJ come-NMLZ see-DICON
 ‘Nowadays he is not seen coming (here).’
 (Shah Jahanni Mwsou Gari, 672:2.45)

In (21), the agent of the event of coming in the boldface nominalized clause above is *bi* ‘he’. However, it is coded as the object of the matrix verb *nu* ‘see’. Thus, we have a raising of the subject of a nominalized clause to the object status of the matrix clause. In the above

sentence, it is also possible to use a subject marker on the word for ‘he’, as shown below. Both of the sentences have the same meaning.

- (22) *at^hik^hal-au* *bi-u* *p^hui-nai-k^hou* *nu-lia*.
 nowadays-LOC 3SG-SUB come-NMLZ-OBJ see-DICON
 ‘Nowadays he is not seen coming (here).’

In (22) *bi* ‘he’ is marked with a subject marker, thus morphosyntactically a subject of the nominalized verb *nu* ‘see’, as well as it is a subject semantically. In (21), on the other hand, *bi* ‘he’ is morphosyntactically an object, although it is the subject of the nominalized verb *p^hui* ‘come’ semantically.

5.1.3. Nominalized clause with ‘NP like structure’

In addition to the above, we also get to see a much more noun-phrase like structure in the nominalized clause. The subject argument turns into a genitive modifier of the nominalized verb, and the object argument is mostly unmarked for case. Consider the following example.

- (23) [*t^həŋp^hak^hri-ni* *buŋ-nai-k^hou*] *k^huna-nanui* *mansi-p^hr-a*
 PN-GEN say-NMLZ-OBJ hear-NF man-PL-SUB
gi-zub-bai.
 be.scared-completely-PERF
 ‘All the people were scared to hear Thengphakhri’s words.’
 (Birgwsrini Thungri, 257:25.1)

- (24) [*aŋ-ni* *buŋ-nai-k^hou*] *raza* *aru* *p^hurza* *boibu*
 1SG-GEN say-NMLZ-OBJ king and subject all
gonai-zub-u.
 obey-completely-AFF
 ‘All the kings and subjects (of my kingdom) obey whatever I say.’ (*lit.* All the kings and subjects follow my sayings.)
 (Jwhwlaow Dwimalu, 93:14.2)

In example (23), the agent participant *t^həŋp^hak^hri* ‘PN’ of the nominalized verb *buŋ* ‘say’ in the bracketed clause is marked in the genitive marker *-ni*. Similarly in example (24), the agent participant of the event of saying expressed in the nominalized verb is marked with the genitive marker, i.e. *aŋ-ni*. Thus, the nominalized verbs have genitive modifiers instead of subject arguments. It is possible to add the object argument of the nominalized verb in (23) and we will get a sentence like (25) below.

- (25) [*t^həŋp^hak^hri-ni* *bat^hra* *buŋ-nai-k^hou*] *k^huna-nanui*
 PN-GEN word say-NMLZ-OBJ hear-NF
mansi-p^hr-a *gi-zub-bai*.
 man-PL-SUB be.scared-completely-PERF
 ‘The people were scared to hear Thengphakhri’s words.’

The above sentence is well formed with the object argument *bat^hra* ‘word’ without any case marker.

So far we have talked about the internal structure of the nominalized clauses. We have seen that the structure of the nominalized clause may vary from a more finite structure to a

more noun-phrase like structure. In the following section, we will talk about the various functions of the nominalized clauses.

6. The distribution and functions of *-nai* clauses

In this section we will talk about the distribution and function of the nominalized clauses with *-nai*. The functions of the nominalized clauses with the nominalizer *-nai* include the following: independent clauses, complements, adverbials, and relative clauses (Boro, K. unpublished). We will talk about each function below.

6.1. Nominalized clause as an independent clause

Nominalized clauses are found to occur all by themselves, without being embedded in another clause in some Tibeto-Burman languages, such as Dolokha Newar (Genetti 2010). This function is variously referred to as ‘non-embedded nominalization’, ‘main-clause’ or ‘stand-alone’ nominalization (Matisoff 1972). We see this phenomenon in Bodo as well. For example,

- (26) *binip^hrainu got^ho-ni muŋ-k^hou alari dambra duun-nai.*
 from.this child-GEN name-OBJ PN PN keep-NMLZ
 ‘From that time, the child’s name is kept Alari Dambra.’
 (Alari Dambra, 1034:21.12)

The sentence does not have a finite verb instead we have a nominalized verb *duun* ‘keep’ in (26). Similarly, we can also find a sentence like the one below.

- (27) *bisur-ni onzima-ja-bu asi k^horo-au-nu sán-nu há-nai.*
 3PL-GEN number-SUB-also finger head-LOC-FOC count-INF can-NMLZ
 ‘Their number can also be counted on the tip of (one’s) finger.’
 (Boro Aizw arw Baihra Mulug, 917:2.14)

Genetti (2010) argues that it is possible to interpret nominalized clauses like the ones above as the subject of an intransitive copula in Dolokha Newar (even though the copula is not overt). The above hypothesis may hold historically, but synchronically it is not completely true in Bodo. Even though it is possible to put a copula with a negative marker, it is not possible at all to put an affirmative copula, as shown below.

- (28) **zuy [zá-nu há-nai] noŋ-gou.*
 1PL eat-NMLZ can-NMLZ COP-AFF
 ‘We will be able to eat.’
- (29) *zuy [zá-nu há-nai] noŋ-a.*
 1PL eat-NMLZ can-NMLZ COP-NEG
 ‘We won’t be able to eat.’

In this respect, the nominalized clauses which can occur as independent clauses are not functionally identical to ordinary NPs. NPs can always be followed by the copula, but the nominalized clauses that occur as independent clauses can never have a copula unless they are negative. Thus, the nominalized clauses which can function as independent clauses stand apart from copular clauses.

A more convincing example can be found in interrogative sentences which contain just the nominalized clause. Consider the following example.

- (30) *nuŋ manu tʰáŋ-nai?*
 2SG why go-NMLZ
 ‘Why did you go?’

We have nominalized interrogative clause above. But, it is not possible to analyse it as a copula construction.

In short, although ‘non-embedded’ nominalized clauses may have derived from copular constructions historically, they are a distinct construction in the synchronic syntax of the language. The productivity of the above nominalized construction is very limited in Bodo. This construction is mainly found in interrogative sentences as in (30) and in sentences expressing uncertainty which contain the verb *há* ‘can’ as in (29). Sentences like (28) are rare and they seem to be a version of a ‘copulative sentence’ with an ellipsis of a copula-like verb. Example (31) has a copula-like verb *za* ‘be/happen’ which (32) lacks. Both express the ‘same’ meaning.

- (31) *binip^hrainnu [got^ho-ni muŋ-k^hou alari dambra*
 from.this child-GEN name-OBJ PN PN
dun-nai] za-ju.
 keep-NMLZ happen-AFF
 ‘From that time, the child’s name is kept Alari Dambra.’
 (Alari Dambra, 1034:21.12)

- (32) *binip^hrainnu got^ho-ni muŋ-k^hou alari dambra dun-nai-sui.*
 from.this child-GEN name-OBJ PN PN keep-NMLZ-CS
 ‘From that time, the child’s name is kept Alari Dambra.’
 (Alari Dambra, 1034:21.12)

6.2. Nominalized clauses as sentence predicates

We often see in the Tibeto-Burman literature that languages create whole finite systems out of nominalized constructions over time. Modern Tibetan is a convincing case of this kind of evolution, in which the tense/aspect system consists of a nominalized verb plus existential copula (DeLancey 2011). Another example is Kham, a Tibeto-Burman language spoken in Nepal, where a nominalizer is reanalyzed as a past tense marker (Watters 2002). Bodo also has a copula verb which functions like the auxiliary verbs, which do not add much semantic content to the proposition expressed in the sentence, and can occur with nominalized clauses. Even though this construction is not highly grammaticalized, it is at least starting to form a distinct construction. As we will see, some of these verbs are undergoing broadening/bleaching in this construction.

- (33) ...*k^hohab-gunaŋ royza-nai mún-nu k^hela-p^hur-k^hou*
 ...taste-bearing be.happy-NMLZ get-NMLZ game-PL-OBJ
gele-ju.
 play-AFF
 ‘(They) play games to get the taste of happiness.’
 (Borophwrni Gwdwni Khela, 1018:1.3)

- (34) ...*aŋ-ni* *gusuu-a* *guuzun-nai* *noŋ-a.*
 1SG-GEN mind-SUB be.satisfied-NMLZ COP-NEG
 ‘I will not be happy.’ (*lit.* My mind will not be satisfied.)
 (Jwhwlao Dwimalu, 102:31.3)
- (35) ...*be* *zaiga-jao-nuu* [*k^herai-k^hou* *hu-bai* *t^ha-nai* *za-ju.*]
 ...this place-LOC-FOC ritual-OBJ give-AUX stay-NMLZ happen-AFF
 ‘Every time the ritual is held in this place.’
 (Borophwrni Rongjathai Phwrwb, 1048:2.4)

The underlined verbs, *mún* ‘get’, *noŋ* ‘cop’, *za* ‘happen’ are the auxiliary-like verbs in the above sentences. They do not add much meaning to the proposition expressed in the clauses, thus they do not have discourse prominence. It is the nominalized verbs which add the main meaning to the sentences. The underlined verbs have the kinds of meaning that can be easily grammaticalized. Since the nominalized verbs add the main meaning to the sentences they are functioning as the predicates of the sentences.

6.3. Nominalized clauses as arguments

Nominalized clauses with both ‘highly finite structure’ and ‘noun-phrase like structure’ can function as arguments of verbal as well as non-verbal predicates. They can function as subject, object, locative, or instrument argument. Consider the following examples.

Subject

- (36) [*sulug* *la-ja* *lasinu* *p^hui-nai-ja*] *nuŋsi-ni*
 information take-NEG without come-NMLZ-SUB 2PL-GEN
gurunt^hi *za-duŋ.*
 wrong happen-REAL
 ‘Coming here without taking (prior) information is your mistake.’ (Bigrai arw
 Dwisrai, 245: 22.11)
- (37) [*bit^haŋ-ni* *lir-nai-ja*] *da-ni* *nuizizou*
 3SG.HON-GEN write-NMLZ-SUB now-GEN twenty.hundred
múga-ni *bwt^hur-zuŋ* *ose* *sumundu* *doŋ.*
 era-GEN season-INSTR close relation exist
 ‘His writing has a close relation with the times of twentieth century.’
 (Manjulani Siddaunai Thwi, 324: 4.2)

Object

- (38) ... [*bit^haŋ-ni* *buŋ-nai-k^hou*] *aŋ* *zубut* *gusuu* *hu-nanui*
 ... 3SG.HON-GEN say-NMLZ-OBJ 1SG very.much mind give-NF
k^hunason-laŋ-duŋ-mun.
 listen-DIST-REAL-PAST
 ‘I kept listening to his speech very attentively.’ (Mahanta, 1030:4.1)

- (39) *p^horai-gra-ja* *p^horai-nanui* [*lirigiri-ni* *buŋ-nai-k^hou*]
 read-NMLZ-SUB read-NF writer-GEN say-NMLZ-OBJ
hom-nuu *há-naŋ-gun.*
 catch-INF can-need-FUT
 ‘The reader should be able to understand the writer’s message after reading (his/her writing).’ (Boro Rebgraphwrni Jengna, 1123:5.8)

Location

- (40) ... [*zirai-nai-au-nuu*] *dublaseo* *undu-nanui* *doŋ-mun.*
 ... rest-NMLZ-LOC-FOC like.dead sleep-NF exist-PAST
 ‘(He) was asleep like a dead (man) while resting.’
 (Abwi Abouni Solo, 6:2.8)

Instrument

- (41) *aŋ-ni* *gusu-a* [*royza-nai-zuuŋ*] *bazlo-k^haŋ-duŋ.*
 1SG-GEN mind-SUB be.happy-NMLZ-INSTR jump-upward-REAL
 ‘My mind is dancing in happiness.’ (Anju, 218: 6.8)

The nominalized clauses are functioning as subjects in (36) and (37), and they are marked with a subject marker *-a*⁵. The nominalized clauses are functioning as object arguments in (38) and (39), and they are marked with an object marker *-k^hou*. The nominalized clause is functioning as a location in (40), and it is marked with the location marker *-au*. The nominalized clause in (41) is functioning as an instrument argument, and it is marked with the instrument marker *-zuuŋ*.

6.4. Nominalized clauses as nominal modifiers

Nominalized clauses also function as relative clauses, which modify a head noun in an NP. We can find three different kinds of relative clause structures in Bodo, namely, head external, head internal and headless. Consider the following examples.

- (42) *mohadeb-a-buu* [*be* *p^he-nai*] *mua-k^hou* *luŋ-buu-duŋ.*
 PN-SUB-FOC this be.drunk-NMLZ thing-OBJ drink-PROX-REAL
 ‘Mahadeb has also been drinking wine (since long).’ (Phenai Besad, 425: 2.43)
- (43) [*diŋa-zuuŋ* *mansi* *bár-hu-nai-a-nuu*] *birgusri-ni*
 boat-INSTR man cross-CAUS-NMLZ-SUB-FOC PN-GEN
gahai *k^hamani* *za-bai.*
 main work happen-PERF
 ‘To get people across (the river) with a boat became the main profession of Birgwsri.’
 (Birgwsrini Thungri, 271: 2.2)

⁵ The subject marker *-a* is realized as *-ja* following non-back vowels /i/, /e/, /a/; *-wa* following back vowel /u/; *-a* following all consonants and following /o/ and /u/.

- (44) *[duk^huram-ni dusi-nai-ja] suk^husri-ni gusuu-au*
 PN-GEN accuse-NMLZ-SUB PN-GEN mind-LOC
*z**u**but duk^hu huu-juu.*
 very.much sorrow give-AFF
 ‘The accusation of Dukharam (to someone) hurts Sukhusri a lot.’
 (*lit.* Dukharam’s accusation gives a lot of pain in Sukusri’s mind.)
 (Phisa Hinzau, 808:2.40)

Example (42) illustrates a head external relative clause in which the bracketed nominalized clause modifies the noun *mua* ‘thing’ in underline. The relative clause always precedes the head noun. Example (43) illustrates a head internal relative clause in which the head noun *mansi* ‘man’ occurs within the nominalized clause. The nominalized clause modifies the head noun and thus functions as a relative clause. The bracketed clause in Example (44) is a headless relative clause which modifies a noun which is unspecified in this context. It may be that Dukharam accused someone who is close to Sukhusri.

6.5. Nominalized clauses as Adverbials

In Bodo, nominalized clauses are found in different kinds of adverbial clauses, such as counterfactual conditional adverbials, reason adverbials, etc. These are discussed with the help of examples below.

Counterfactual conditional adverbials:

- (45) *[abo-a dinui t^hab p^hui-nai-ba] ham-si-gou-hai.*
 sister-SUB today quick come-NMLZ-IF be.good-IMMED-AFF-INFL
 ‘It would have been nice if (my) sister was here (by now).’
 (Thajjou Zanai, 1053: 116.3)

Reason adverbials:

- (46) *[nungou-k^hou buŋ-nai-k^hai] lazi mún-duŋ.*
 true-OBJ say-NMLZ-BECAUSE be.ashamed get-REAL
 ‘(S/he) is ashamed because the truth has been said.’ (*lit.* S/he got ashamed because of saying the truth.) (Swrjini Maya, 495:2.97)

Manner adverbials:

- (47) *[bi liŋhor-nai-lek^ha] daut^hu-p^hur-a p^hui-nanui*
 3SG call-NMLZ-ACCORDINGLY dove-PL-SUB come-NF
besor-k^hou k^honk^haŋ-p^hin-u.
 mustard.seed-OBJ pick.up-again-AFF
 ‘In response to his invitation the doves came and picked the mustard seeds again.’
 (Alari Dambra, 1034: 45.3)

Time adverbials:

- (48) *aŋ* *tʰiari-mun-bla-bu* *liŋ-nai-kʰali* *natʰai* ...
 1SG be.prepared-PAST-if-even call-NMLZ-when but ...
tʰurtʰur *burbur* *za-duŋ-mun.*
 ONOM ONOM happen-REAL-PAST
 ‘I was shaky on the day I was called (for interview) even though I was prepared.’
 (Manju Bala Devi, 106:1.492)

In (45) we have a nominalized clause functioning as a counterfactual conditional adverbial marked another adverbializer *-ba*. Similarly, in (46) we have a nominalized clause functioning as a reason adverbial marked with the adverbializer *-kʰai*. Again we have another nominalized clause functioning as a manner adverbial marked with the adverbializer *-lekʰa* in (47). Further, in (48), we have a nominalized clause functioning as a time adverbial marked with the adverbializer *-kʰali*.

7. Summary

This paper has described the range of formal and functional variations of nominalized constructions exhibited in Bodo. The lexical nominalizations are simple functionally and structurally. Some of the lexical nominalizers are found to be less productive in the synchronic grammar except in a few instances. Clausal nominalization, on the other hand, is quite productive. Structurally, the *-nai* clauses range from a fairly finite structure to a noun-phrase like structure. In its finite form, the *-nai* clause has all its arguments fully coded for their grammatical role. In its noun-phrase like form, the subject is coded as a genitive modifier of the nominalized verb. The *-nai* clause can occur in a wide range of syntactic functions and distributions. It can function as an independent clause all by itself. It can also function as different kinds of arguments of the verb such as subject, object, location, and instrument. It can also function as a relative clause and modify a head noun in an NP. It can also function as a verbal predicate in a sentence. Finally, the *-nai* clause also functions as an adverbial clause, and takes adverbial markers.

Abbreviations

1	First person	GEN	Genitive
2	Second person	HON	Honorific
3	Third person	IMMED	Immediative
ADV	Adverbial suffix	INF	Infinitive
ADVZ	Adverbializer	INFL	Informal
AFF	Affirmative	INSTR	Instrumental case
ALL	Allative case	LOC	Locative case
CAUS	Causative	NEG	Negative
CONTR	Contrastive suffix	NF	Non-finite
COP	Copulative verb	NMLZ	Nominalizer
CS	Change of state	OBJ	Object
DAT	Dative case	ONOM	Onomatopoeia
DISCON	Discontinuative	PASS	Passive
DIST	Distal motion	PAST	Past tense
FOC	Focus particle	PERF	Perfect aspect
FUT	Future	PL	Plural

PN	Proper noun	SG	Singular
PROX	Proximal motion	SUB	Subject
REAL	Realis		

References

- Basumatary, Prafulla. (2015): *The Bodo verb: A descriptive-pedagogic study*. Ph.D. Dissertation. Guwahati, Gauhati University.
- _____. (2017): *Verbal semantics in a Tibeto-Burman language: The Bodo verb*. (*Contemporary Studies in Descriptive Linguistics, Vol. 42*). Oxford, Peter Lang.
- Boro, Krishna. (2012). "Serialization in Boro." In Gwendolyn Hyslop, Stephen Morey and Mark W. Post, Eds. *North East Indian Linguistics Vol. 4*. New Delhi, Cambridge University Press, India: 83–103.
- _____. (in preparation). *Nominalized Clauses in Boro*. Guwahati, Gauhati University. Unpublished manuscript.
- Comrie, Bernard and Sandra A. Thompson. (1985). "Lexical nominalization." In Timothy Shopen, Ed. *Language typology and Syntactic description, Volume III: Grammatical Categories and the Lexicon*. Cambridge, Cambridge University Press: 349-398.
- DeLancey, Scott. (2011). "Finite structures from clausal nominalization in Tibeto-Burman." In Foong Ha Yap, Karen Grunow-Hårsta and Janick Wrona, Eds. *Nominalization in Asian Languages: Diachronic and typological perspectives*. Amsterdam, John Benjamins: 343–362.
- Genetti, Carol. (2011). "Nominalization in Tibeto-Burman languages of the Himalayan area: A typological perspective." In Foong Ha Yap, Karen Grunow-Hårsta and Janick Wrona, Eds. *Nominalization in Asian Languages: Diachronic and typological perspectives*. Amsterdam, John Benjamins: 163–193.
- Matisoff, James. (1972). "Lahu nominalization, relativization, and genitivization." In John Kimball, Ed. *Syntax and Semantics (Vol.1)*. New York, NY, Academic Press: 237–257.
- Morey, Stephen. (2011). "Nominalization in Numhpuk Singpho." In Foong Ha Yap, Karen Grunow-Hårsta and Janick Wrona, Eds. *Nominalization in Asian Languages: Diachronic and typological perspectives*. Amsterdam, John Benjamins: 289–311.
- Watters, David. (2008). "Nominalization in the East and Central Himalayish Languages of Nepal." *Linguistics of the Tibeto-Burman Area* 31(2): 1–43.

Reflexives and reciprocals in Dimasa¹

Monali Longmailai
Tezpur University, Assam

Abstract This paper aims to describe the kinds of reflexives and reciprocals found in Dimasa, a Tibeto-Burman language from the Bodo-Garo sub-group. It attempts to discuss some of the morphosyntactic features of the nominal and verbal reflexives and the verbal reciprocal, case-marking in nominal reflexives and reduplication. It also discusses the detransitivizing features, co-referentiality and functions of reflexives as benefactives, activities pertaining to body parts and emphasis.

Citation Longmailai, Monali. 2018. *Reflexives and reciprocals in Dimasa*. North East Indian Linguistics (NEIL) 8, 126-139. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.

Volume Editors Linda Konnerth, Stephen Morey, Amos Teo

Copyright © 2018, the author(s), release under Creative Commons Attribution license

URL <http://hdl.handle.net/1885/155280>

1. Introduction

Dimasa (639-3: dis) belongs to the Sal group of languages (Boro-Garo) from the Tibeto-Burman language family as mentioned in Post and Burling (2017).² It has several dialects namely, Hasao, Hawar, Dembra, Dijuwa, Humri, Semsá, Walgong among which, Hasao is adopted as the standard form since 2004. It is spoken mainly in Assam with 110, 976 speakers according to 2001 census (Dima Hasao, Nagaon, Karbi Anglong and Cachar districts) and in parts of Nagaland, Meghalaya, Mizoram and Manipur.³ Its alternate names are the exonyms Dimasa Kachari and Hills Kachari as given in the Ethnologue although, it has different exonyms: Hills Kachari (Dima Hasao and Karbi Anglong), Barman (Cachar), Hojai (Nagaon) and Kachari Naga (Dimapur in Nagaland).⁴ The Dimasa community refers to themselves and the language as Dimasa, which is an autonym.

It is an agglutinating language with SOV word order, and has tonal features. Honorificity is expressed with the second person plural *nisi* since it is not encoded by bound morphemes in the language like Assamese. Dimasas are first learners of the language and they are multilingual speakers presently assimilated with Bengali in the Cachar plains of Assam, Assamese in Karbi Anglong and Nagaon districts of Assam and Nagamese in Nagaland, Shillong Hindi and few bilinguals in Khasi in Shillong, Meghalaya. The standard script is

¹ I am deeply indebted to Robbins Burling for helping me with valuable comments and suggestions during my entire research on Dimasa, Kh. Dhiren Singha for the several discussions on the linguistic features present in this language, Francois Jacquesson for his suggestions in writing this paper and the title of the paper, Umarani Pappuswamy for guiding me throughout my doctoral research, where this paper is a short section of a chapter on grammatical relations, the Dimasa community for their constant help and support throughout the years of working on Dimasa, especially Dimasa Sahitya Sabha (Spelling and Glossary Committee) for sharing orthographic conventions of Dimasa and for their encouragement and support in my researches on the language. I am also thankful to my colleague, Daimalu Brahma for listening to my queries on reflexives and reciprocals and my family for clarifying my doubts on the data. I am finally thankful to the reviewers for their comments and suggestions in writing the paper.

² The language code of Dimasa is *dis* and it is identified in the list of ISO 639-3 in the Ethnologue.

³ The approximate population of Dimasa in the entire north-eastern region of India could be 1-2 lakhs. The population 110, 976 in the Census of India consists of only the twin hill districts of Assam, Dima Hasao and Karbi Anglong.

⁴ The Eighth Schedule of the Constitution of India, 1950 represents only Hills Kachari as Dimasa.

Roman, although the vernacular script, Bengali, remains to be used only by the older generations.

Few linguists have contributed towards the descriptive study of Dimasa in recent years, such as: Singha has done research on the phonology and morphology of Dimasa (2001), and later published his works on Dimasa phonology (2007) and a Dimasa word book (2010), Jacquesson has written an outline grammar of the Dimasa language and made a Dimasa-English lexical dictionary (2008), Sarmah's work (2009) is based on the acoustic study of tones of Dimasa and Rabha, both of which belong to the same group of languages. Longmailai also has studied few morphosyntactic aspects of Dimasa: personal pronouns (2012), adjectives (2014a), passivization (2017) and the unpublished work on morphosyntax (2014b). Other literary writers such as Dundas (1908) is the earliest documented outline grammar and dictionary of Dimasa language, Thaosen (1994) written in Bengali, is the first attempt to introduce the phonology and morphological features such as word formation in Dimasa and Nampui and Ngirsim (1996) is a multi-lingual dictionary with Dimasa, Biata, Assamese, Hindi and English languages. There is very scarce documentation done on reflexives and reciprocals in the language. The present paper, therefore, brings forward the morphosyntactic properties of reflexives and reciprocals in Dimasa, since the language is highly rich in noun and verb morphology.

Reflexives are constructions in which the subject and the object refer to the same entity. Reciprocals are again constructions in which two participants, the subject and the object equally are related to the same event or action (Payne 1997). Among the Bodo-Garo languages, the reflexives and reciprocals are mostly unidentical in forms: Garo (Burling 2004) has the nominal reflexive *-tang/an-tang* meaning 'self' and the verbal reciprocal *grik* 'mutually, each other'. van Breugel (2008) describes *-than* as a possessive enclitic (nominal reflexive) in Atong similar in form to Garo, and *-ruk* as a verbal reciprocal while *-maran* is a nominal reciprocal attached with kinship and interpersonal relations. In Bodo (Brahma 2013), *gao* is the nominal reflexive while it is reduplicated to function as nominal reciprocal *gaozwng gao*. The nominal reflexive in Rabha (Joseph 2007) is derived by partial reduplication, from *kai* 'person' as *kakai*. Dimasa has the nominal reflexive *ḍḍàr* 'self', the verbal reflexive *-lā*, and the verbal reciprocal *-leī*. A nominal reciprocal is absent in this language.

This paper, brings forth a preliminary overview of these reflexives and reciprocals present in the language (§2). The case-marking properties and reduplication of the nominal reflexive are outlined in §2.1. The verbal reflexives and verbal reciprocals and their morpho-syntactic functions are discussed in §2.2 and §2.3. In §3, the omission of reflexives and reciprocals in some constructions are briefly discussed. In §4, the detransitivizing properties of these reflexives and reciprocals as well as the co-referentiality besides other morpho-syntactic functions are highlighted. The data used in the paper are self-generated, since the author is a Dimasa-speaking linguist.⁵

2. An overview of reflexives and reciprocals

Among the three types of reflexives, namely, lexical, morphological and analytic, the Dimasa language has both morphological and analytic reflexives. The nominal reflexive *ḍḍàr* is analytic, as a free root, whereas the verbal reflexive *-lā* is morphological, as a bound suffix, in this language. Dimasa has only a verbal reciprocal *-leī*, which is morphological in nature. The reflexive *-lā* and the reciprocal *-leī* in Dimasa are identical in form and structure since they are derived from the verb stem *lā* 'take'. This section identifies the nominal (§2.1) and verbal

⁵ Some of the data are sample illustrations extracted from the author's doctoral research work on the morphosyntax of Dimasa.

reflexives (§2.2) and the verbal reciprocals (§2.3) in Dimasa and discusses their morphosyntactic features.

2.1. Nominal reflexive

The nominal reflexive is morphologically a reflexive particle *dʒàr* ‘self’, which is a free lexeme.⁶ *dʒàr* is a nominal reflexive since it has case-marking properties which are prototypical features of nouns and nominals which is discussed in §2.1.1. It lexically operates in a reflexive construction distinct from the verbal counterpart *-lā*⁷, which is morphologically expressed with a verb (Payne 1997: 200). The occurrences of nominal and verbal reflexives are separately illustrated in (1) and (2):

- (1) *āŋ dʒàr-kʰē bāū-dū*
 1SG REFL-ACC think-PRES
 ‘I do think of myself’.

- (2) *waī ftaī-lā*
 light/fire douse-REFL
 ‘Turn off the light by yourself/on your own’.

However, Dimasa also allows double reflexives in order to topicalize the anaphoric reference of the subject, as shown in (3), *dʒàr* with the accusative case *-kʰē* and *-lā* with the transitive verb *ftaī* ‘kill’ as shown in the following labelled brackets:

- (3) *bō [dʒàr-kʰē] mūli dʒī-hī*
 3SG REFL-ACC medicine eat-NF

[/ftaī-lā-kʰā] bākʰálī-bā gđà-hà
 kill.CAUS-REFL-PFV when-INDF past-LOC
 ‘He killed himself consuming the medicine, long time back’.

dʒàr can also function as an emphatic particle when used as the adjunct/oblique argument of a clause, since it is cross linguistically common for a reflexive form to have an emphatic counterpart, as shown in (4), where it emphasizes the subject/agent *bō* ‘third person singular’ (Payne 1997):

- (4) *bō [dʒàr] ākʰī-bā ālú-kʰē*
 3SG EMPH draw-PST cat-ACC
 ‘She drew the cat by herself’.

⁶ Only two Indo-Aryan languages from the Eastern sub-group use the similar nominal reflexive as *nidʒe/nidʒor* in Assamese and *niʒe* in Bangla, since their related languages as illustrated by Subbārão (2012) are: Hindi *apne aap* and *khud*, Marathi *apan*, Maithili *apan*, Kashmiri *pan*, to express the same, which indicates *dʒàr* not to be of Indo-Aryan, but that of Tibeto-Burman origin. Assamese also uses *apun* ‘own’ as a nominal reflexive with kinship and close relations as in *apun manuh* besides *nidʒor manuh* to mean ‘own people’ and in Bangla *apna/apnar* ‘own/self’, based on the knowledge of the author on these two languages.

⁷ The verbal reflexive *-la* bears a similarity with the Khasi nominal reflexive *lade*, mentioned in Subbārão (2012: 56), a neighbouring Austro-Asiatic language. Longmailai (2015) has compared some of the Dimasa and Khasi lexemes which bear common similarities due to areal contact and borrowings.

Dimasa has a flexible word order construction. Hence, both the reflexive *džàr* and the emphatic *džàr* can occur in any part of the clause.

Some other reflexives recorded in Tibeto-Burman languages include Sino-Tibetan **guŋ*, Tibeto-Burman **goŋ* and Lolo-Burmese **daw* (Bradley 1995), as well as the Bodo *gou* derived from the etymological meaning of ‘body’, which however, are not identical with Dimasa *džàr*, though the word *faū-džāŋ* means ‘alone’, a combination of *faū* ‘body’, which seems to be a cognate of Bodo *gou*, and the comitative case marker *-džāŋ*, which can sometimes alternatively occur as an emphatic *džàr*, as in (5) and (6):

(5) *bō mjā faū-džāŋ p^haì-raù-k^ha*
 3SG boy body-COM come-strong-PFV
 ‘The boy dared to come alone/ on his own/ by himself’.

(6) *bō mjā džàr p^haì-raù-k^hà*
 3SG boy EMPH come-strong-PFV
 ‘He dared to come on his own/ by himself’.

faū-džāŋ, however, is not a reflexive since it does not take case-marking as the object in a transitive construction as shown in (9) whereas *džàr* is not adjectival in form unlike *faū-džāŋ* in (7) and hence cannot occur in intransitive sentences as in (8). *faū-džāŋ* is used either for emphasis in (5) or attribute in (7) whereas *džàr* is used for co-referentiality between the subject and the object and thus occurs usually with transitive verbs as in (10). Hence, the nominal reflexive *džàr* and the adjectival/adverbial *faū-džāŋ* are not related to each other in Dimasa.

(7) *bō mjā faū-džāŋ-fè*
 3SG boy body-COM-EMPH
 ‘The boy is alone/ is on his own/ himself’.

(8) **bō mjā džàr-fè*
 3SG boy REFL-EMPH
 ‘The boy is on his own/ himself’.

(9) **bō mjā faū-džāŋ-nē fōŋ-bā*
 3SG boy body-COM-BEN COOK-PST
 ‘The boy cooked for alone’.

(10) *bō mjā džàr-nē fōŋ-bā*
 3SG boy REFL-BEN cook-PST
 ‘The boy cooked for himself’.

The origin of *džàr* is unknown since it does not resemble the forms found in related Sino-Tibetan languages, including Bodo-Garo. It is worth mentioning here that Jacques (2010: 3) in his discussion on the origin of the reflexive prefix in the Rgyalrong languages of the Sino-Tibetan language family, shows the presence of the reflexive prefix *zyr-* in Japhug as derived from the pronouns, among other forms from other Sino-Tibetan languages besides *jɛ-* in Tshobdun. *zyr-* in Japhug which is a verbal reflexive distantly resembles phonologically with the nominal *džàr* in Dimasa.

2.1.1. Case marking properties of *d̥ʒàr*

The nominal reflexive *d̥ʒàr* carries case-marking and functions as an anaphor to the nouns and pronouns. It never carries the nominative case since the subject is always the anaphoric nouns/pronoun which can be omitted as shown in the bracket ‘{}’ in (11) *d̥ʒìŋ* ‘we’ and the object is *d̥ʒàr-nì dārā dīfà* ‘one’s traditions’ even though it can be fronted in the sentence:

- (11) {*d̥ʒìŋ-ø*} [*d̥ʒàr-nì*] *dārā dīfà* *mt^haū-mā* *nāŋ-dū*
 1PL-NOM REFL-GEN rules ways remain-INF need-HABG
 ‘One should keep one’s own traditions’.

d̥ʒàr is seen to occur with the accusative, genitive, locative and comitative cases as the object anaphors from (12) to (14):

- (12) *bō d̥ʒàr-nīŋ* [*d̥ʒàr-k^hè*] *nai-fō* *blaī-dū*
 3SG EMPH-only REFL-ACC see-complete efficient-HABG
 ‘She, herself, can watch herself’.

- (13) *bō-nī* [*d̥ʒàr-nī-hā*] *k^hrib* *d̥ʒādī* *rī-ni*
 3SG-GEN REFL-GEN-LOC all type cloth-GEN

k^hùn *maī-dū*
 thread get-HABG
 ‘She has all kinds of threads available with her’.

- (14) [*d̥ʒàr-d̥ʒāŋ*] *k^hèd̥ʒā-là-bā-fē* *bō-dē*
 REFL-COM happy-REFL-HABG-EMPH 3SG-TOP
 ‘He is happy with himself’.

The emphatic *d̥ʒàr-nīŋ* ‘only herself’ in (12) functions as an adjunct while *-nīŋ* in the reflexive *d̥ʒàr-k^hè-nīŋ* ‘only herself’ in (15) topicalizes the object compliment.

- (15) *bō* [*d̥ʒàr-k^hè-nīŋ*] *nai-fō* *blaī-dū*
 3SG REFL-ACC-only see-complete efficient-HABG
 ‘She can/is able to watch only herself (not any other person than her)’.

2.1.2. Reduplication of *d̥ʒàr*

Reduplication of reflexives to function as plurality is found across several languages. Among the related languages from northeast India, the possessive enclitic *-than* in Atong (van Breugel 2008: 292) is used for the nominal reflexive form that is reduplicated for plurality in S/A argument besides its use in kinship terms. Burling (2004: 181), in Garo, mentions of *-tang*, as doubled to create a plural distributive meaning. Rabha undergoes a partial reduplication of the singular *kakai* to *kakakai* for the plural and distributive meanings besides the use of the plural and distributive *kakroŋ* (Joseph 2007: 342-343). Similarly, *d̥ʒàr* in Dimasa can be reduplicated to indicate plurality and distributive meaning in the subject in a noun phrase as shown in (16), where it occurs with the accusative case *-k^hè*.

- (16) *k^hrib-bò* [*d̥ʒàr-d̥ʒàr-k^hè-fē*] *mìt^hi-bà*
 all-also REFL-REFL-ACC-EMPH know-HABG
 ‘Everyone knows only themselves/each one of them’.

dʒàr cannot be reduplicated on its own as *dʒàr-dʒàr* as a nominal reflexive, but as an emphatic particle as in (17):

- (17) *nīfī* *dʒàr-dʒàr* *tʰāŋ-leī-lā*
 2PL EMPH-EMPH go-RECIP-REFL
 ‘(You) go by yourselves/ on your own’.

2.2. Verbal reflexive

The verbal reflexive is constructed with a verbal suffix *-lā*. This suffix seems to derive from the verb stem *lā* ‘take’ as a verbal reflexive. *-lā* is also homophonous to the emphatic and the adversative particles which nearly behave like reflexives (Jacquesson 2008). In Dimasa, both the verbs *lāŋ* and *lā* mean ‘take’ and they can be compounded as *lāŋ-lā* ‘take away’, where the second stem is delexicalized and transforms into a verbal reflexive. While *lāŋ* means ‘take as in, from one destination to another, not necessarily having immediate physical movement’, *lā* means ‘take requiring physical involvement in one destination’. Hence, *lāŋ-lā* literally translates as ‘take it away/ take by yourselves’.

Dimasa verbs can be derived by *lāŋ* ‘take’ into a middle construction (LaPolla 1996) while *-lā* ‘take’ is a verbal reflexive as shown in (18) and (19).

- (18) *waī* *nàŋ-làŋ-kʰà*
 fire catch-MID-PFV
 ‘Fire started catching/ Fire caught on its own’.

- (19) *waī* *nàŋ-là-kʰà*
 fire catch-REFL-PFV
 ‘Fire caught on its own’.

The contextual information in (18) indicates that the fire started catching without any possible cause or reason to trace as in, a house burning suddenly, while (19) indicates the fire-catching situation in general, i.e. the subject, *waī* ‘fire’ in (18) undergoes an experience or state beyond its control, while (19) is the controller of the action as in, firewood used to cook rice.

-lā as a verbal reflexive cannot be reduplicated unlike the nominal reflexive *dʒàr*, though it has the form of reduplication in a verbal clause where the main verb is *lā* ‘take’ such as (20), with the finite verb *lā* means ‘take’, as compared with *lāŋ* in (21).

- (20) *kʰāfrīŋ* *laīfī* *lā-lā-kʰà*
 Khasring book take-REFL-PFV
 ‘Khasring has taken the book (simply taken it by himself)’.

- (21) *kʰāfrīŋ* *laīfī* *lāŋ-lā-kʰà*
 Khasring book take-take-PFV
 ‘Khasring has taken away the book (carried and taken away with him)’.

In (21), *lāŋ-lā* is a compound word (sometimes considered to be a serial verb) meaning ‘he came to take the book and left with it altogether’, whereas (20) is a reflexive construction, ‘he has taken it by himself’. The reflexive *-lā*, therefore, cannot occur when both the verb stems *lāŋ* and *lā* meaning ‘take’ co-occur as in (21), since the second verb is already

delexicalised to become a reflexive when it occurs with other verbs, and hence, it is ungrammatical in (22):

- (22) **kʰāfrīŋ* *laīfī* *lāŋ-lā-lā-kʰà*
 Khasring book take-take (REFL)-REFL-PFV
 ‘Khasring has taken the book (by himself)’.

The functions of the verb stems, *lāŋ* and *lā* are contrasted in the following examples, (23) to (26):

- (23) *jaū* *fū-lā*
 hand wash-REFL
 ‘Wash your own hands’.

- (24) *jaū* *fū-lāŋ*
 hand wash-MID
 ‘Go and wash your own hands and then come back’.

- (25) *kʰūfī* *dāŋ-là*
 work do-REFL
 ‘Do your own work’.

- (26) *kʰūfī* *dāŋ-lāŋ*
 work do-MID
 ‘Go and do your own work and keep doing it since it started’.

Both *lāŋ* and *lā* can also occur with verb-like adjectives pertaining to states, which are illustrated with *lēŋ* ‘be tired’ in intransitive constructions in (27) and (28).

- (27) *àŋ* *lēŋ-lā-kʰā*
 1SG be.tired-REFL-PFV
 ‘I have become tired (I was doing something)’.

- (28) *àŋ* *lēŋ-lāŋ-kʰā*
 1SG be.tired-MID-PFV
 ‘I have become tired (by either doing something or nothing)’.

2.3. Verbal reciprocal

The reciprocal particle is morphologically expressed by the post-verbal *-leī*, for expressing the meaning of ‘each other. Besides the verbal reflexive *-lā* and the middle reflexive *-lāŋ*, the verbal reciprocal *-leī* possibly is derived from the verb *lā* meaning ‘take’ as well. *-leī* is homophonous to the adverbial *-leī* to indicate similarity in comparison. The example in (29) illustrates the use of the verbal reciprocal *-leī*:

- (29) *rùpòl-dzāŋ* *džoīdzīt* *fū-leī-dù*
 Rupal-COM Joyjit hit-RECIP-PRES
 ‘Rupal and Joyjit are hitting each other’.

-leī indicates the action of event as not occurring with the speaker or the observer, but with a group of people, that is, it indicates plurality. Jacquesson (2008) has mentioned *-leī* as a sociative-reciprocal form in his list of actancy suffixes⁸, since it associates participants in terms of performing an action with each other. The example in (30) shows the association or togetherness of a group of people in doing a common work, i.e. writing.

- (30) *būnfi* [*rēp-leī-dù*]
 3PL write-RECIP-HABG
 ‘They are writing for each other’.

The reciprocal *-leī*, cannot be reduplicated like the verbal reflexive *-lā*. Besides, it occurs either with *lāŋ* as *lāŋ-leī*, or with *lā* as *lā-leī* to indicate the position of ‘taking’ from within the domain or outside, but never together with *lāŋ-lā* as *lāŋ-lā-leī* as shown in (31), (32) and (33):

- (31) *ḍḍiŋ* *ṣāmlai* [*lāŋ-leī-nàŋ*]
 1PL food take-RECIP-FUT.DIST
 ‘We will go take the food along (from one place to another)’.

- (32) *ḍḍiŋ* *ṣāmlai* [*lā-leī-nàŋ*]
 1PL food take-RECIP-FUT.DIST
 ‘We will take the food among ourselves/from each other (within one place)’.

- (33) **ḍḍiŋ* *ṣāmlai* [*lāŋ-lā-leī-nàŋ*]
 1PL food take-TAKE.REFL-RECIP-FUT.DIST
 ‘We will go take away the food from each other’.

3. Deletion in reflexives and reciprocals

The subject, can be alternately removed in reflexive construction in which case, the object, that is the reflexive, is fronted to the subject position as topic since the reflexive refers back to the subject anaphorically, as highlighted in §2, in example (11). Besides the object deletion in case of plural subject in reciprocal construction as seen in previous examples, singular subjects can also undergo object deletion in certain cases. The verb determines if the action is required to be ‘with another entity’ rather than ‘with oneself’.

The example in (34) illustrates the subject *ḍḍiŋ* ‘we’ as optionally deleted in case of reflexive construction due to co-referentiality with the objective noun phrase *ḍḍàr-nì dārā diṣà* ‘one’s own traditions’.

- (34) {*ḍḍiŋ*} [*ḍḍàr-nì*] *dārā diṣà* *mt^haū-mā* *nāŋ-dū*
 1PL REFL-GEN rules ways stop-INF need-HABG
 ‘One should keep one’s own traditions’.

In case of reciprocals, the verbs like *nām-ḍḍā* ‘quarrel’ in (35) and *hām-ḍḍaù* ‘love’ in (36) indicate that there are two entities needed to perform one single action, that of ‘quarrelling’ and ‘loving (someone)’, for which reason, the object undergoes deletion from the construction. This is unlike the semantic function of the verb *ḍḍi* ‘eat’ in (37), which does not

⁸ Actancy suffixes are listed in Jacquesson (2008) as the factitive *-ri*, the passive *-žao* and the reflexive *-la* based on the combination of these verb suffixes and the number of nominal arguments as three-actant, two-actant and one-actant verbs respectively.

require two persons, but the agent can perform the action alone. Hence, (35) and (36) are co-referential subjects with a covert object while (36) is an ungrammatical reciprocal construction which requires an oblique object.

(35) [bō] {/rèbà-dʒāŋ} [nām-dʒā-leī-dū]
 3SG someone-COM quarrel-be-RECIP-HABG
 ‘She quarrels (with somebody)’.

(36) bō hām-dʒaù-leī-dù
 3SG good-PASS-RECIP-HABG
 ‘He is in love (Lit: He is having good feeling felt with someone)’.

(37) *adaī mək^hàm dʒī-leī-dū
 grandmother rice eat-RECIP-PRES
 ‘Grandmother is eating rice with each other’.

4. Detransitivizing agents

Transitivity in a language can be described in terms of the number of nominal arguments or participants in a syntactic construction and their relations (grammatical) with the verb (Payne 1997). Thus, the valence or the number of arguments in a sentence such as, ‘John hit the car’ increases from two (‘John’ and ‘car’) to three (‘Mary’, ‘John’ and ‘car’) in a causative construction as in, ‘Mary made John hit the car’. Reflexives and reciprocals, on the other hand, decreases it to two or one, as in ‘John and Mary hit each other’ and ‘John hit himself’ due to co-referentiality of the subject and the object as one entity.

This section, thus, discusses the detransitivizing or anti-causative property of the nominal and verbal reflexives and the verbal reciprocals in Dimasa. It also briefly shows the co-referentiality of these valency decreasing categories and their other morphosyntactic functions.

4.1. Reflexives

Both nominal and verbal reflexives have only one argument, that is, the subject and hence, it functions as a detransitivizing agent. They are illustrated from (38) to (40):

(38) āŋ_i dʒàr_i-k^hè nai-p^hū-dū
 1SG REFL-ACC see-POT-HABG
 ‘He can look after himself’.

(39) bō_i dʒàr_i-k^hè dʒàr ftaī-lā_i-k^hā
 3SG REFL-ACC EMPH kill-REFL-PFV
 ‘He killed himself on his own’.

(40) bō_i jaū_j fū-lā_i-dù
 3SG hand wash-REFL-HABG
 ‘He is washing his own hands’.

The antecedents here are the pronominal subjects while the nominal anaphor is *dʒàr* ‘self’ in (38) and (39) and the verbal anaphor is *-lā* in (39) and (40). Co-referentiality is a

morphosyntactic property of the reflexives and either the nominal or verbal or both can occur together in the construction.

The post-verbal suffix *-lā* in transitive constructions having two or three arguments in (41) and (42) is identical to the verbal reflexive *-lā* in that it refers back to the subject, although it is an adverbial particle mostly used for emphasizing on the distance in near space/direction from the speaker/observer. Brahma (2017) has compared the verbal deixis in Bodo and Dimasa, with the Dimasa *-ha/-hi* verbal suffixes which have similar deictic functions with *-la* as well as with the combination of *lāŋ-leī* and *lā-leī* in the previous examples (31) and (32) in §2.3:

- (41) *bō_i bɔ̌ʒāŋ-k^hē_j k^hāfaú-lā_i-dū*
 3SG younger sibling-ACC care-DIST-HABG
 ‘He cares for his younger sister/brother (by himself)’.
- (42) *bō_i ā-nē_j mk^hām fōŋ-hī ɔ̌tī-rī-lā_i-bā*
 3SG 1SG-DAT rice cook-NF eat-CAUS-DIST-PST
 ‘He treated me with rice (by himself)’.
- (43) *bō_i āŋ-ɔ̌ʒāŋ_j ɔ̌ʒàr_i t^hāŋ-mā t^hī-hī t^hī-bā*
 3SG 1SG-COM EMPH go-INF say-NF say-PST
 ‘He, himself, planned to go with me’.

4.2. Reciprocity

Verbal reciprocal constructions are detransitivizing in that they decrease the number of arguments from two to one in (44), three to two in (45) and two to one in (46). Hence, the verbal reciprocal *-leī* is illustrated with the anaphoric subjects as a pronoun in (44), proper name in (45) and a noun phrase in (46):

- (44) *būnfī dōŋ-p^hā-leī_i-dū*
 3PL stay-COM-RECIP-HABG
 ‘They stay with each other’.
- (45) [*hāmring-ɔ̌ʒāŋ nairīŋ_j p^hūtòl_j mlaū-leī_i-du*]
 Hamring-COM Nairing football play-RECIP-HABG
 ‘Hamring and Nairing play football with each other’.
- (46) [*àɔ̌ʒāŋ-ɔ̌ʒāŋ bō-nī lūgù-dè_i mlaū-leī_i-dū*]
 younger sibling-COM 3SG-GEN friend-TOP play- RECIP-PRES
 ‘(My younger) sister/brother is playing with her/his friend’.

4.3. Co-referentiality of reflexives and reciprocals

Co-referentiality of reflexives and reciprocals with the anaphoric subjects can co-occur in the same construction as exemplified in (47):

- (47) *būnfī ɔ̌ʒàr-ɔ̌ʒāŋ k^həɔ̌ʒā-leī-lā_i-dū*
 3PL REFL-COM (be)happy-RECIP-REFL-HABG
 ‘They are happy along with each other, among themselves’.

In terms of co-referentiality of reflexives and reciprocals, the verb codas *lāŋ* and *lā* cannot be compounded together, without the intervention of the verbal reciprocal *-leí* as seen in (48) in comparison to (49) and (50), where *-leí* occurs between *lāŋ* and *lā*.

- (48) *dʒīŋ fāmlai lāŋ-leí-lā-nāŋ*
 1PL food take-RECIP-REFL-FUT.DIST
 ‘We will take the food from each other by ourselves (from one place to another)’.
- (49) *dʒīŋ fāmlai lā-leí-lā-nāŋ*
 1PL food take-RECIP-REFL-FUT.DIST
 ‘We will take the food from each other by ourselves (within the same place)’.
- (50) **dʒīŋ fāmlai lāŋ-lā-leí-lā-nāŋ*
 1PL food take-take-RECIP-REFL-FUT.DIST
 ‘We will take the food from each other by ourselves’.

4.4. Other morphosyntactic functions

Verbal reflexives and verbal reciprocals in Dimasa can function as a self-benefactive, for activities pertaining to body parts and for emphasis which is similar to Subbārāo’s analysis of several South Asian languages such as, Kannada and Telugu (Dravidian), Ho (Munda), and Mizo and Hmar (Tibeto-Burman) (2012: 62-64).

In (51), the nominal reflexive *dʒàr* is marked for the benefactive *-nè* referring to the subject *ʃaūrīŋ* ‘Shaoring’, while in (52), the verbal reflexive functions as a benefactive with the verb-like adjective *kʰə̀dʒà* ‘be happy’:

- (51) *ʃaūrīŋ dʒàr-nè mkʰàm pʰāntʰā dèŋ-dù*
 Shaoring REFL-BEN rice share keep-PRES
 ‘Shaoring is keeping his share of rice/ the share of rice for himself’.
- (52) *bō ānʃā kʰə̀dʒà-là-kʰà*
 3SG child be.happy-REFL-PFV
 ‘The child is happy for himself’.

In (53), *-lā* is suffixed with the physical activity related to the body, that is, *dùgùr* ‘take bath’ and the future marker *-nāŋ* here indicates for asking question with a rising intonation:

- (53) *dùgùr-là-nāŋ*
 take.bath-REFL-FUT.DIST
 ‘Why don’t you take bath?’

Both *dʒàr* and *-lā* co-occur in (54) and (55) for emphasis:

- (54) *nīŋ dʒàr baū-lā*
 2SG EMPH think-REFL
 ‘(You) think about it (by yourself)’.
- (55) *nīŋ dʒàr-kʰè baū-lā*
 2SG REFL-ACC think-REFL
 ‘(You) think for yourself’.

5. Conclusion

To sum up the paper, *d̥àr* as a nominal reflexive in Dimasa has a flexible order in a clause construction since it is inflected for case while the emphatic reflexive is unmarked and almost always occurs next to the subject. The nominal reflexive *d̥àr* carries case except for the nominative since it is anaphoric to the antecedent subject. Both nominal and emphatic reflexives can have total reduplication as *d̥àr-d̥àr* though the nominal use requires case marking. *-lā* is a verbal reflexive while *-lāŋ* (both literally meaning ‘take’), is used in middle construction which is very similar to that of reflexives. *-leí* is the verbal reciprocal and plural in form, while nominal reciprocal is absent in the language. *-lā* and *-leí* are verbal suffixes probably derived from the verb *lā* meaning ‘take’, while *lā* and *lāŋ* as free roots are coda variations. Verbal reflexive and reciprocal occur with both transitive and intransitive verbs and they cannot be reduplicated. However, either *lāŋ* or *lā* can occur with *-leí* but not together as *lāŋ-lā-leí*, including when the verbal reflexive *-lā* follows the order of these morphemes.

There is subject deletion for reflexives while object deletion for reciprocals, since the reflexives and reciprocals are co-referential to their anaphoric subjects. Both the categories can co-occur and reflexives can function as benefactives, pertain to activities related to body parts and emphasis. Since the present study is a preliminary work on the reflexives and reciprocals in Dimasa, further research would be required to throw light into the areal typology of these reflexives and reciprocals in the Eastern Indo-Aryan, Khasi (Austro-Asiatic) and the neighbouring Tibeto-Burman languages, including that of Bodo-Garo, since Dimasa is found to have different forms for reflexives and reciprocals as compared to these related languages.

Abbreviations

1	First person	n	Noun
2	Second person	NF	Non-finite
3	Third person	NMZ	Nominalizer
ACC	Accusative	NOM	Nominative
BEN	Benefactive	PFV	Perfective
CAUS	Causative	PL	Plural
COM	Comitative	POT	Potential
CNT	Contrast	PRED	Predicative
DAT	Dative	PRES	Present
DIST	Distal	PST	Past
EMPH	Emphatic	RECIP	Reciprocal
FUT	Future	REDUP	Reduplication
GEN	Genitive	REFL	Reflexive
HABG	Habitual Generic	SG	Singular
INDF	Indefinite	TOP	Topicalizer
INF	Infinitive	v	Verb
LOC	Locative		

References

- Bradley, David. (1995). “Reflexives in Burmese” In David Bradley, Ed. *Papers from South East Asian Linguistics No. 13: Studies in Burmese languages*. [Pacific Linguistics Series A, #83]. Canberra, Australian National University: 139-172.

- Brahma, D. (2017). "A study on the verbal deixis in Bodo and Dimasa." In Monali Longmailai, Ed. *Studies on Dimasa history, language and culture, Vol. 1, DYF Selected Papers*. Guwahati, DVS Publishers.
- Brahma, O. (2013). *Modern Bodo grammar. Morphology: Bases and Affixes*. (1st edition, Vol. 1). Guwahati, N. L. Publications.
- Burling, Robbins. (2004). *The language of the Modhupur Mandi (Garo): Vol. 1: Grammar*. New Delhi, Bibliophile South Asia.
- _____. (2009). *Dimasa Tones*. Unpublished manuscript.
- Census of India (2001). (Downloaded from www.censusindia.gov.in, accessed 30/05/2016).
- Dundas, W. C. M. (1908). *An outline grammar and dictionary of the Kachari (Dimasa) Language*. Shillong, Eastern Bengal and Assam Government.
- Jacques, Guillaume. (2010). "The origin of the reflexive prefix in Rgyalrong languages." *Bulletin of the School of Oriental and African Studies* 73(2): 261–268.
- Jacquesson, François. (2008a). "A Dimasa Grammar". (Downloaded from http://brahmaputra.ceh.cnrs.fr/bdd/IMG/pdf/Dimasa_Grammar-2.pdf, accessed 05/05/2016).
- _____. (2008b). "Dimasa-English Lexicon". (Downloaded from http://brahmaputra.ceh.cnrs.fr/bdd/IMG/pdf/Dimasa_English-2.pdf, accessed 05/05/2016).
- Joseph, U. V. (2007). *Rabha*. Leiden, Brill.
- LaPolla, Randy J. (1996). "Middle voice marking in Tibeto-Burman." In *Pacific-Asian Linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics, Vol. V, Mahidol University, Thailand: 1940–1954*.
- Longmailai, Monali. (2012). "Personal Pronouns in Dimasa." In Gwendolyn Hyslop, Stephen Morey & Mark W. Post, Eds. *North East Indian Linguistics: Volume 4*. New Delhi, Cambridge University Press: 104–118.
- _____. (2014a). *The Morphosyntax of Dimasa*. PhD Thesis. Department of Linguistics, North-Eastern Hill University, Shillong.
- _____. (2014b). "Adjectives in Dimasa." In J. War, S.K. Singh, S.A. Lyngdoh and B. Khyriem, Eds. *Tibeto Burman Linguistics of North East India: Volume 3*. Guwahati, Eastern Book House Publishers.
- _____. (2015) "Lexical and morphological resemblances of Khasi and Dimasa." In Mark W. Post, Stephen Morey and Scott DeLancey, Eds. *Language and culture in Northeast India and beyond: In honour of Robbins Burling*. Canberra, Australia-Pacific Linguistics: 126–139.
- _____. (2017). "Passivization in Dimasa." In S. K. Singh and S. A. Lyngdoh, Eds. *Syntactic Typology: Language Contact and Convergence*. Guwahati, EBH Publishers.
- Nampui, R. & Ngirsim, S. (1996). *Multi lingual learner book*. Haflong, Nampui, R. & Ngirsim, S.
- Payne, Thomas E. (1997). *Describing morphosyntax: A guide for field linguists*. Cambridge, Cambridge University Press.
- Post, Mark W. and Robbins Burling. (2017). "The Tibeto-Burman languages of Northeast India." In Randy J. LaPolla and Graham Thurgood, Eds. *The Sino-Tibetan languages* (2nd edition). United Kingdom, Routledge: 213–242.
- Sarmah, Priyankoo. (2009). *Tone Systems of Dimasa and Rabha: A phonetic and phonological study*. PhD Dissertation. Gainesville, University of Florida.
- Simons, Gary F. and Charles D. Fennig, Eds. 2017. *Ethnologue: Languages of the world, Twentieth edition*. Dallas, Texas: SIL International. (Online version: <http://www.ethnologue.com>, accessed on 18/08/2017).

- Singha, Dhiren Kh. (2001). *The phonology & morphology of Dimasa*. MPhil Dissertation. Silchar, Assam University.
- _____. (2007). *Introduction to Dimasa phonology*. Guwahati, DVS Publishers.
- _____. (2010). *Dimasa word book: A classified vocabulary*. Guwahati, DVS Publishers.
- Subbārão, K. V. (2012). *South Asian languages: A syntactic typology*. New Delhi, Cambridge University Press.
- Thaosen, S. (1994). *Dimasa Bhasa Bigyan Porichoy*. Haflong, Haflong Art Society.
- van Breugel, Seino. (2008). *Grammar of Atong*. PhD Thesis. Melbourne, La Trobe University Research Centre for Linguistic Typology.

Verbal person indexation in languages
at the Indo-Myanmar border

Hierarchical indexing and the inverse in Muklom Tangsa¹

Mijke Mulder
La Trobe University

Abstract

Muklom Tangsa is a Tibeto-Burman language variety spoken in Northeast India. This article discusses verb inflection in this variety, in specific person-number indexing and inverse marking. So far, very little data had been available for Muklom, but based on newly collected data we can obtain a better understanding of its complex verb system. I propose that Muklom exhibits the following two features in the domain of the verb: 1) hierarchical person-number indexing, and 2) inverse, or non-A speech act participant marking. The first, hierarchical person indexing, specifically for Muklom entails that indexing of a first or second person argument is preferred over third person indexing in the mixed domain, regardless of the syntactical function of that first or second person argument. P and R arguments in the mixed domain may trigger indexing on the verb when they are higher on the so-called person hierarchy than the A argument. The second feature, the inverse, consists of an invariable suffix on the verb. In Muklom, the inverse signals the presence of a non-A speech act participant. The inverse is not restricted to the core arguments P and R. Non-arguments, at least possessors but perhaps including others, may also trigger non-A speech act participant marking on the verb. The two features have in common that they show how NP and verb marking in Muklom are not neatly aligned, but disintegrated, as is not uncommon for a Tibeto-Burman language.

Citation	Mulder, Mijke. 2018. <i>Hierarchical indexing and the inverse in Muklom Tangsa</i> . North East Indian Linguistics (NEIL) 8, 141-155. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.
Volume Editors	Linda Konnerth, Stephen Morey, Amos Teo
Copyright	© 2018, the author(s), release under Creative Commons Attribution license
URL	http://hdl.handle.net/1885/155280

1. Introduction

The Tangsa population is spread out over parts of the states Arunachal Pradesh and Assam in India, but also over the northern part of the Sagaing Region in Myanmar; they dwell on both sides of the India-Myanmar border. The Tangsa people can be divided into tens of so-called sub-tribes, one of which are the Muklom people. While some sub-tribes span the border, the Muklom Tangsa are confined to Indian territory, specifically the Changlang district of the state Arunachal Pradesh. The Tangsa sub-tribes speak a group of language varieties that belong to the Tibeto-Burman (TB) family. In their attempts to classify the TB languages of Northeast India, scholars have placed Tangsa varieties in various lower level and higher level super groups, but they did consistently assign Tangsa to the Northern Naga clade. I will not explore the topic of genetic affiliation further (for the development of classifications see

¹ I would like to thank the Muklom Tangsa community, all members that have acted as language consultants, and specifically the Changmi and Khimhun families from Kuttom and New Chingsa, who exhibited an astonishing degree of hospitality and who diligently helped me to record and summarise recorded stories and discussions. Mr Nongmai Changmi was the main consultant in elicitation sessions regarding verb inflection. Tone determination would not have been possible without the hard work of consultants, especially Ms Tchithan Techi and Mr Nongmai Changmi, but also Mr Chupan Changmi, Mr Phanglong Sajung, ‘i³ βΛŋ²’ Mr Ngulang Changmi and ‘i³ tək’ Misses Changmi, Ms Bina Changmi, Ms Sunumi Changmi, Mr Munkap Khimhun, and Mr Munson Khimhun. I am very grateful to Stephen Morey, Linda Konnerth, and an anonymous reviewer for providing useful comments on this paper. The research reported in this paper was funded by La Trobe University Postgraduate Research Scholarships (LTUPRS), La Trobe University Full Fee Research Scholarships (LTUFFRS), and the Internal Research Grant Scheme of the School of Humanities and Social sciences, La Trobe University.

among others Shafer 1955, Benedict 1972, Bradley 1997, Burling 2003, Blench & Post 2015). However, for the current discussion it may be relevant to know that different scholars have suggested Nocte as the closest relative of Tangsa (e.g. Shafer 1953: 228, French 1983: 726, Burling 1983: 17, Bradley 1997: 21, and Burling 2003: 175). In fact, sometimes the same language variety is called Tangsa if its speakers live within Tangsa areas in the Changlang district of Arunachal Pradesh or in Assam, but Nocte if its speakers are located within a Nocte area in the Tirap district of Arunachal Pradesh. This is happening for example with Hakhun and Phong/Ponthai (Morey 2017). The fact that some varieties show fluidity in sub-tribe-association should by no means be taken as conclusive proof for the lack of an actual linguistic boundary between Tangsa and Nocte, even though this is an interesting hypothesis to explore. The flexibility in association does however suggest a high degree of linguistic similarity between the two.

It should be noted that the Tangsa variety that is the topic of discussion in the current article was most often referred to as ‘Moklum’ in previous literature (e.g. Das Gupta 1980²). However, the native pronunciation of the language and sub-tribe name is [muklɔm]. The spelling ‘Muklom’ is therefore considered preferable and is adopted in the current article.

While research into Tangsa varieties fortunately is slowly increasing, published original linguistic literature on Muklom Tangsa in specific is still extremely limited. Even when amateur-linguistic work is included only the following four sources are available³, as far as I am aware: 1) a sixty-six page description of Muklom including remarks on grammar, phrases with English translation and a Muklom-English wordlist by a former language officer of Changlang in Arunachal Pradesh, Ngemu (1977), 2) a six-page sketch by the former director of research in Arunachal Pradesh, Das Gupta (1980), 3) a thirteen-page wordlist by Bandyopadhyay (1989)⁴, and 4) more recent work done by Morey, who in addition to re-interpreting data from the previous two sources (Morey 2011a), discusses a few lexical items and verbal inflection in Muklom based on newly collected data (Morey 2014: 667, Morey 2015: passim).

Muklom Tangsa has a rich verb system. We will discuss two elements of it, 1) person-number (P-N) indexing, which shows hierarchical alignment, and 2) and the inverse, or non-A speech act participant (SAP) marking. Muklom Tangsa indexes P-N on the verb, and no gender. The first and second person indexes are portmanteau morphemes, encoding number and person. However, the third person index is neutral for number, i.e. the same form is employed with single and plural referents. Indexation is straightforward in intransitive clauses, but things become more interesting when we look at indexation on transitive verbs. We find that the agent-like argument of the transitive verb (A) is not by rule indexed on the verb. Instead, in some configurations, the first or second person argument is indexed on the verb, regardless of its syntactic or semantic role in the sentence. Giving preference to first and second person over third person is a type of indexing that has been termed hierarchical in the literature (see Siewierska 2013 for a comparison of the different verbal alignment systems). ‘Hierarchical’ refers to the ranking of arguments on a referential hierarchy. Different variables may be included in this referential hierarchy, depending on the language. For Muklom the variable ‘person’ suffices. On the Muklom person hierarchy, first and second person have a higher rank than third person, meaning they are given preference over third person referents,

² Throughout this booklet Das Gupta uses the spelling ‘Moklum’, but in the index he does provide the word ‘Muklom’ between brackets as an alternative spelling, which better reflects the native pronunciation.

³ There is in fact one more published source which could contain original linguistic data: Das Gupta, Shri K. (1971). “A few concepts in the Moklum dialect.” *Pines*, 1.1, Shillong. This article is mentioned by Ngemu (1977: 7), but I have not been able to locate a copy. This article perhaps forms a preparation for the short book *The Tangsa language. A synopsis*, which the author published later (Das Gupta 1980).

⁴ This wordlist is an abstract from the following unpublished manuscript: Bandyopadhyay, Swapon Kumar (1980). *A linguistics study of Tangsa. A Tibeto-Burman language*. Dissertation. University of Delhi.

1/2>3. When both a low ranked and a high ranked argument are encoded in a clause, the high ranked argument is given priority with respect to indexing on the verb, irrespective of syntactic role.

Muklom Tangsa exhibits another feature in the domain of verb inflection that is cross-linguistically rare: inverse marking. A suffix *-p^h* attaches to verbs in certain sentences but must remain absent in others. This small morpheme is the second topic of discussion in this article, and the main issue that will be addressed is the trigger behind the appearance of this suffix on the verb, i.e. the function of the verbal marker *-p^h*. The marker has previously been noticed by scholars (e.g. Das Gupta 1980: 9, Morey 2011: 90, 96) but a detailed analysis of its function was not possible, since the amount of original data at their disposal was still limited. In this article, I will present new data that I collected during recent fieldtrips, which will help gain a better understanding of the function of the marker.

The complex system of P-N indexing and the inverse in Muklom Tangsa will be broken up and discussed piece by piece in the body of this article. I present the following simplified twofold proposal, which will be expanded in the discussion: 1) P-N indexing on transitive verbs is hierarchical in part of the paradigm, but is ruled by the A argument in the remainder, 2) the inverse suffix *-p^h* signals the presence of a non-A speech act participant.

The next section will describe how, where and with what equipment new data presented in this paper were collected. After that, we will discuss P-N indexing and the inverse marker *-p^h* in Muklom Tangsa, including a brief comparison to three related language varieties. The last section summarizes the findings and makes suggestions for further research.

2. Methods

New Muklom data reported in this article have been collected by the author during a fieldtrip from October 2015 to March 2016. Audio recordings have been made with a Zoom H4n, analysed using Praat (Boersma and Weenink 2017), and glossed in FieldWorks Language Explorer (SIL). In addition to audio recordings, field notes have been taken with pen and notebook.

All native speaker Muklom examples provided by the author in this article derive from controlled elicitation session, except (19) which is a spontaneous utterance. The examples were uttered by residents of Kuttom village near Kharsang Town and recorded and/or noted down.

There is currently no standardized spelling for Muklom and therefore IPA transcriptions are used. Tones are represented by a superscript number. Muklom exhibits a three-way tone system, as is typical for a Tangsa language. The tone categories consist of a bundle of features: pitch height and contour, phonation, and duration.

3. Hierarchical person-number indexing

Not all TB languages of Northeast India exhibit P-N indexing. Indeed, Burling (2003: 173) states that many lack verb agreement entirely. Although a small minority of the Tangsa varieties lack indexing, it is abundantly present in the majority. Earlier it was assumed that all Tangsa varieties exhibited P-N indexing (Morey 2011a: 98), but with the collection of new data from the region it has become apparent that there are in fact at least three Tangsa varieties that do not exhibit P-N indexing on the verb in any form. These are Champang, Pinkhu and Yasawa (Morey 2015: 32). That Champang does not show P-N indexing is quite remarkable because its closest relative, Haqchum, does. This shows again that the Tangsa cluster exhibits a high degree of linguistic diversity. Not surprisingly, many Tangsa varieties are not mutually intelligible.

Muklom Tangsa is among the Tangsa varieties that do show P-N indexing. As mentioned in the introduction, person is always expressed in the index, and number is expressed for SAP argument, i.e. first and second person. Non-SAP indexation is neutral for number, i.e. the same morpheme is employed for third person singular and plural referents. Number-neutral third person indexation seems to be a feature that all Tangsa indexing systems share (Morey 2011: 94). Table 1 lists all indexes of the intransitive paradigm in the past, present, and future tenses. The phonological form of the indexes differs slightly between the tenses. Changes in tone (all persons), as well as changes in vowel and coda (first singular and third person) can be observed. Since the shape of the indexes also changes after negation, the negative paradigm is provided in the last column. The sigma symbol, Σ , represents the verb stem, and the superscript numbers represent the tone categories (in Muklom, 1 is a creaky voice, mid-to-low falling tone, 2 is a long, rising-falling tone, 3 is a breathy, mid or low, level tone).

Table 1 – P-N indexes in Muklom Tangsa (intransitive)

P-N	Form			
	PST	PRS	FUT	NEG
1SG	Σ -t- $\Lambda\eta^1$	Σ - $\Lambda\eta^3$	Σ -n- $\Lambda\eta^3$	Σ -m- $\text{ɔ}\eta^1$
1PL	Σ -t- i^1	Σ - i^3	Σ -n- i^3	Σ -m- i^1
2SG	Σ -t- u^1	Σ - u^3	Σ -n- u^3	Σ -m- u^1
2PL	Σ -t- m^1	Σ - m^2	Σ -n- m^3	Σ -m- m^1
3	Σ -t- a^1	Σ - a^3	Σ -n- a^3	Σ -m- $\text{ɔ}^?$

Certain Tangsa varieties exhibit two sets of indexes, one with plosive codas and one with nasal or no codas, where tense and polarity determine the selection of the set (Morey 2015: 33). Although Muklom does show variation in tone, vowel, and even coda, the latter is limited to the third person negative $-m\text{ɔ}^?$. Examples of third person indexes (in bold) are provided in (1) and (2), the first without negation and the second with negation.

- (1) pi^2 - a^3 tl^2 sa^2 - a^3 .
 3SG-ERG rice eat.soft-3
 ‘He eats rice.’ (20121103_04)

- (2) pi^2 - a^3 tl^2 sa^2 - m - $\text{ɔ}^?$.
 3SG-ERG rice eat.soft-NEG-3
 ‘He does not eat rice.’ (20121103_04)

Vowel changes are observed only in the first person singular index, $\Lambda \rightarrow \text{ɔ}$, and third person index, $a \rightarrow \text{ɔ}$, after the negator $-m$. The result is negative indexes $- \text{ɔ}\eta^1$ and $- \text{ɔ}^?$, which have rounded, low back vowels. Examples (3) and (4) show a non-negated and a negated first person singular index.

- (3) ηa^1 - a^3 tl^2 sa^2 - $\Lambda\eta^3$.
 1SG-ERG rice eat.soft-1SG
 ‘I eat rice.’ (2009Tascam-001)

- (4) ηa^1 - a^3 tl^2 sa^2 - m - $\text{ɔ}\eta^1$.
 1SG-ERG rice eat.soft-NEG-1SG
 ‘I do not eat rice.’ (20100124-112656)

Note that Table 1 does not provide tense distinctions for the negative, and this is for a reason: negation and tense marking cannot co-occur on a verb stem. When a verb is negated, tense marking moves to an auxiliary. Interestingly, both the negated main verb and the tense-bearing auxiliary receive P-N indexes. An example of this redundant indexing in future tense is provided in (5), and past tense in (6).

(5) ηa^1-a^3 ti^2 $sa^2-m-\mathfrak{a}\eta^1$ $\lambda\eta^2-n-\lambda\eta^3$.
 1SG-ERG rice eat.soft-NEG-1SG be-FUT-1SG
 ‘I will not eat rice.’ (20121103_04)

(6) pi^2-a^3 $k\mathfrak{a}^2-m-\mathfrak{a}^2$ $\lambda\eta^2-t-a^1$.
 3SG-ERG give-NEG-3 be-PST-3
 ‘He did not give it (to him).’ (20121103_04)

In the indexation paradigm, we see variation in vowel quality, presence or absence of a coda, but also in tone category. Each column in Table 1 is associated with a tone category. Past tense always comes with tone 1, present tense indexes mostly carry tone 3, future tense is exclusively tone 3, and negation is associated with tone 1. Roughly, the paradigm falls apart into two groups: past-negation with tone 1, present-future with tone 3. Please note that tone 2 marking in Table 1 on 2PL index $-m^2$ is not a typo: it has been confirmed with consultants more than once that the present tense 2PL index carries tone 2 and not the expected tone 3. With respect to the clustering of past and negation, it is interesting to point out that this pattern is found in a substantial amount of the Tangsa varieties. Morey (2016) reports that a number of varieties have index series with plosive coda for the past and negated forms (as opposed to non-stopped in the future tense). In his list of 25 (Morey 2016: 2), six show this pattern: Ngaimong, Joglei, Mueshaung, Mungre, Haqchum, and Hakhun. With respect to associated tones we can make further comparisons between Muklom and Morey’s sample of 25 Tangsa varieties. Although past and negative are sometimes associated with tone 1, more varieties show an association with tone 3. Similarly, the future tense is more often associated with tone 2, though future indexes with tone 3 do also occur in the sample. For present tense, we cannot make a comparison, as it is not included in Morey’s sample data (the reason for not including present tense is the presence of non-cognate constructions, Morey, p.c.).

As a final remark before moving on to hierarchical indexing, I would like to point out that Muklom and a few other Tangsa varieties (e.g. Yongkuk, Tikhak, and Phong) show P-N paradigms remarkably similar to that of Nocte as discussed by DeLancey (2010: 15)⁵: $-\lambda\eta$ ‘1SG’, $-\mathfrak{a}^2$ ‘2SG’, $-i^2$ ‘1PL’, $-\lambda n$ ‘2PL’, and $-a/-a^2$ ‘3PL’.

Sentences (1) to (6) above show indexes on the verbs that corresponded to the A argument of the clause. In examples (1), (2), and (6), a third person A argument acts upon a third person P argument, 3→3, while in examples (3), (4), and (5) a first person A argument acts upon a third person P argument, 1→3. In these configurations we do find indexation of A. However, agreement is not always with the A argument. If for example third person A acts upon first person P, 3→1, the index aligns with the P argument. Example (7) provides an example of the latter configuration.⁶

⁵ DeLancey quotes this paradigm from unpublished notes by Alfons Weidert, who re-elicited a Nocte paradigm published by Das Gupta (1971: 16).

⁶ A few remarks on the verb form in example (7): 1) the element $-p^h$ as in the verb $watta^2p^h\lambda\eta^2$ ‘they hit me’ is the inverse marker and will be discussed further in Section 4, 2) the past tense marker has two allomorphs, $-ta^2$ in front of the inverse marker as in this example, and $-t$ elsewhere, 3) the tone of the 1SG marker changes into 2, $-\lambda\eta^2$, after the inverse marker $-p^h$, see Table 2.

- (7) $n\eta^2-a^3$ ηa^1 $wat-ta^2-p^h-\mathcal{A}\eta^2$.
 3PL-ERG 1SG hit-PST-INV-1SG
 ‘They have hit me.’ (B1.86)

Notice how $-\mathcal{A}\eta^2$ on the verb corresponds to the P argument, ηa^1 ‘I’, and not to the A argument $n\eta^2$ ‘they’. The indexing pattern of Muklom is described as follows: indexes align with the S/A argument of the clause, except when a third person A acts upon a first or second person P argument, in which case indexes align with the ‘higher ranked’ P argument. To be precise, SAPs trigger indexing not only in the syntactic role of P, but also as R in a ditransitive clause, or even as a possessor. The latter two, less typical constructions will be discussed later. Overall, SAPs have a higher status in verbal indexing. This phenomenon can be visualised by means of a hierarchy such as in Figure 1.

SAP		non-SAP
1^{st} person	/	2^{nd} person > 3^{rd} person

Figure 1 – A person hierarchy for Muklom Tangsa P-N indexing

With respect to SAP arguments, first person does not have priority over second, and neither does second person have priority over first. Rather, the semantic role of the two SAPs determines alignment of the index in configurations where they co-occur: the A argument is given priority. Not all languages that exhibit hierarchical indexing assign the same rank to first and second person. Some rank second person higher than first (e.g. the Nakh-Dagestanian language Icar Darci, Jacques and Antonov 2014: 309, based on data from Sumbatova and Mutalov 2003), and others rank first person higher than second (e.g. Hakhun Tangsa, DeLancey 2017: 90, based on data from Boro 2012). Neither do languages automatically index A when semantic role determines indexing in SAP→SAP configurations. On the contrary, DeLancey (2017: 87) states that Sino-Tibetan hierarchical systems most often index the P in clauses that contain two SAP arguments.

Combinations of two SAPs in the transitive paradigm, i.e. $1 \rightarrow 2$ and $2 \rightarrow 1$, are often referred to as the ‘local domain’ (Jacques and Antonov 2014: 302). Configurations consisting of non-SAPs, i.e. $3 \rightarrow 3$, are called the ‘non-local’ domain of the transitive paradigm, and mixed configurations, i.e. $1/2 \rightarrow 3$ and $3 \rightarrow 1/2$, are known as the ‘mixed’ domain. Using this terminology, we can say that semantic role determines index alignment in the local domain, while position on the person hierarchy regulates the mixed domain (we could add the non-local domain, but since number is not encoded, we can never determine whether third person indexes align with A or P).

Person-based or hierarchical indexing is certainly not uncommon among TB language languages. Some argue that hierarchical indexing can be reconstructed as far back as the level of Proto-Tibeto-Burman (e.g. DeLancey 1981, Bauman 1975) and that S/A-agreement like in Kuki-Chin language varieties is an innovation. It seems most Tangsa varieties lack hierarchical indexing (Morey 2011a: 96), but at least one other variety has been reported to show hierarchical indexing, namely Hakhun (Morey 2011b: 677). It also occurs in two related language groups, Jinghpaw and Nocte (DeLancey 2011). The Tangsa people consider their sub-tribes and their language varieties to be divided into two groups: 1) the so-called Pangwa varieties, and 2) the rest or the ‘non-Pangwa’ varieties. The non-Pangwa groups mostly arrived earlier than the Pangwa groups in what is now Arunachal Pradesh (Morey 2015: 25). It remains to be tested whether this subgrouping also reflects a linguistic classification split.

Muklom and Hakhun both fall under non-Pangwa (Morey 2015: 34) and it would be interesting to investigate whether hierarchical indexing is restricted to this subgroup.

The Muklom indexation paradigm will be illustrated with more examples, after which an overview is given in Table 2. The first two examples are of a P argument that is lower on the person hierarchy than the A argument, 1→3 in the first example (8), and 2→3 in the second example (9). The first and second person A arguments trigger indexation on the verb.

- (8) ηa^1-a^3 mi^2 nu^1 $wat-t-\lambda\eta^1$.
 1SG-ERG person female hit-PST-1SG
 ‘I have hit the woman.’ (B1.85)

- (9) $n\lambda\eta^1-a^3$ mi^2 nu^1 $wat-t-u^1$.
 2SG-ERG person female hit-PST-2SG
 ‘You(sg) have hit the woman.’ (B1.85)

In example (10) and (11) the roles are reversed. The non-SAPs are now acting upon the SAPs, 3→1 and 3→2 respectively. These configurations fall within the mixed domain, where rank on the person hierarchy determines indexation on the verb. As you can see in example (10) and (11), the higher ranked first and second person P arguments trigger indexation.

- (10) $n\eta^2-a^3$ ηa^1 $wat-ta^2-p^h-\lambda\eta^2$.
 3PL-ERG 1SG hit-PST-INV-1SG
 ‘They have hit me.’ (B1.86)

- (11) $n\eta^2-a^3$ $n\lambda\eta^1$ $wat-ta^2-p^h-u^3$.
 3PL-ERG 2SG hit-PST-INV-2SG
 ‘They have hit you(sg).’ (B1.86)

We move on to the local domain, with example (12) illustrating the configuration 1→2, and example (13) illustrating the reverse, 2→1. In SAP→SAP configurations, semantic role determines indexation, A is given preference.

- (12) ηa^1-a^3 $n\lambda\eta^1$ $wat-ta^2-p^h-\lambda\eta^2$.
 1SG-ERG 2SG hit-PST-INV-1SG
 ‘I have hit you(sg).’ (B1.86)

- (13) $n\lambda\eta^1-a^3$ ηa^1 $wat-ta^2-p^h-u^3$.
 2SG-ERG 1SG hit-PST-INV-2SG
 ‘You(sg) have hit me.’ (B1.86)

Consultants indicate that indexing of the P is not acceptable when first and second person are acting upon one another, and deem example (14) ungrammatical, 1→2 with A indexing, and example (15), 2→1 with A indexing.

- (14) * ηa^1-a^3 $n\lambda\eta^1$ $wat-ta^2-p^h-u^3$.
 1SG-ERG 2SG hit-PST-INV-2SG (B1.86)

- (15) * $n\lambda\eta^1-a^3$ ηa^1 $wat-ta^2-p^h-\lambda\eta^2$.
 2SG-ERG 1SG hit-PST-INV-2SG (B1.86)

Not only the A or P argument of a transitive clause can trigger indexation, the R argument of a ditransitive clause can as well. The R argument noun phrase is marked by the suffix *-ma³* which has been labelled dative case. Although P arguments in all above examples do not carry case marking, it is possible for the P to be marked with an absolutive suffix. Topicality seems to play a role in absolutive marking of the P argument. Differential P marking or differential absolutive marking in Muklom would be an interesting topic for future research.

Example (16) illustrates indexing of the R argument. Notice that the T argument is a non-SAP in this sentence. It would be interesting to investigate whether an R can still be indexed on the verb when the T argument is also an SAP.

- (16) *niŋ²-a³ ŋa¹-ma³ bɔ³-ti¹ kɔɔ²-taɔ²-p^h-aŋ².*
 3PL-ERG 1SG-DAT papaya-CLASS give-PST-INV-1SG
 ‘They gave papaya to me.’ (B1.86)

In the example sentences above I have provided a few of the possible and impossible patterns of indexing, and Table 2 gives an overview of indexes with all configurations of A argument and P/R argument. For easy comparison, the indexes of the intransitive (-trans) verb are listed in the last column. The overview is nearly complete; there is one other verbal suffix, *-p^ha³*, which serves as an alternative for all 3→SAP forms listed in Table 2. This variation will be discussed after Table 3. The A/S arguments are listed on the left side of the table and the P/R arguments are listed at the top of the table. The indexes that combine with inverse marker *-p^h* (in all tenses) are identical to the indexes found on the intransitive present tense verb, with one exception: 1SG carries tone 2, *-aŋ²*, instead of the expected tone 3. This deviation in tone has been confirmed with different consultants.

Table 2 – Hierarchical P-N indexing in Muklom Tangsa (present tense)

A \ P/R	1SG	1PL	2SG	2PL	3	S
1SG			Σ -p ^h -aŋ ²	Σ -p ^h -aŋ ²	Σ -aŋ ³	Σ -aŋ ³
1PL			Σ -p ^h -i ³	Σ -p ^h -i ³	Σ -i ³	Σ -i ³
2SG	Σ -p ^h -u ³	Σ -p ^h -u ³			Σ -u ³	Σ -u ³
2PL	Σ -p ^h -m ²	Σ -p ^h -m ²			Σ -m ²	Σ -m ²
3	Σ -p ^h -aŋ ²	Σ -p ^h -i ³	Σ -p ^h -u ³	Σ -p ^h -m ²	Σ -a ³	Σ -a ³

Notice that the transitive indexes in the first four rows correspond to the A argument listed on the left, while in the last row, the first four indexes correspond to the P/R argument listed at the top (for the fifth index, *-a³*, we cannot tell whether the correspondence is with A or P/R since both are third person and third person does not distinguish number). The shaded area in Table 3 shows the area in which P/R indexing occurs.

Table 3 – The three domains and non-A indexing in Muklom Tangsa

A \ P/R	1SG	1PL	2SG	2PL	3	S
1SG			<i>local</i>		<i>mixed</i>	
1PL						
2SG	<i>local</i>					
2PL						
3	<i>mixed</i>				<i>non-local</i>	

The collected data indicate that there is one possible index per combination of A and P/R in the white area of Table 3. However, the situation is more complex for combinations where the P/R outranks the A (3→SAP), the shaded area in Table 3. Indexing of the higher ranked first or second person P/R is the most frequent option, but we also find indexation of the lower ranked third person by means of the verb ending $-p^h a^3$. Also, one consultant indicated that indexing of a lower ranked third person A without inverse marker $-p^h$ is acceptable on the condition that ‘you add something to make a complete sentence’. When these options are added, the shaded area of Table 3 will look like Table 4.

Table 4 – Variation in hierarchical P-N indexing (present tense)

A \ P/R	1SG	1PL	2SG	2PL
3	$\Sigma-p^h-\Delta j^2$ $\Sigma-p^h-a^3$ $\Sigma-a^3$	$\Sigma-p^h-i^3$ $\Sigma-p^h-a^3$ $\Sigma-a^3$	$\Sigma-p^h-u^3$ $\Sigma-p^h-a^3$ $\Sigma-a^3$	$\Sigma-p^h-m^2$ $\Sigma-p^h-a^3$ $\Sigma-a^3$

What could explain this deviation from the person hierarchy $SAP > 3$? It should be mentioned that the use of the index $-a^3$ was at first deemed ungrammatical by one of my younger consultants, but after other (older) consultants objected, they agreed in the end that indexing of a lower ranked third person A is acceptable ‘in the right sentence’. What exactly this right environment is that allows the person hierarchy to be overruled, has yet to be established, but pragmatics is one factor that could be playing a role. Perhaps indexing of a third person A argument in 3→SAP configurations is allowed when this third person is a topic in the preceding discourse. Alternatively, syntax could be playing a role. For example, perhaps $\Sigma-p^h-a^3$ and $\Sigma-a^3$ are allowed in subordinate clauses, but not in main clauses. This is a topic for further investigation.

4. The inverse

This section discusses the function and distribution of the inverse suffix $-p^h$ in Muklom. The last section showed how Muklom takes first and second person together and treats them as a separate category when it comes to P-N indexing. It has long been observed that various constructions in TB languages recognize first and second person, i.e. speaker and addressee, as a separate category. This category has been given different names, for example ‘speech participant category’ (Bauman: 1975) and ‘speech act participants’ (DeLancey 1980). In this article, the latter term is employed. In Muklom Tangsa, SAPs not only play a role in P-N indexing, but also trigger a specific invariable marker $-p^h$ on the verb. The marker signals the presence of an SAP P/R argument in a transitive clause, or even an SAP possessor. The shaded area in Table 5 shows the distribution of $-p^h$.

Table 5 – The three domains and inverse marking in Muklom Tangsa

A \ NON-A	1SG	1PL	2SG	2PL	3	S
1SG			<i>local</i>		<i>mixed</i>	
1PL						
2SG	<i>local</i>					
2PL						
3	<i>mixed</i>				<i>non-local</i>	

Based on still very limited sources, Morey (2011: 90, 96) briefly discusses the function of Muklom $-p^h$. He suggests that the verbal suffix indicates the presence of first person P with imperatives or the presence of first or second person P in transitive or R in ditransitive clauses. Based on new data collected during recent fieldwork we can now explore the function of $-p^h$ in more depth and it will be shown that the marker can be used in a wider variety of situations than has so far been assumed. With the help of examples, I will map out the functional range of the inverse in Muklom.

The proposal that will be put forward, is: $-p^h$ marks specifically non-A SAPs. After this proposal is discussed in the following two sections, a third section will compare the Muklom inverse marker $-p^h$ to morphemes, perhaps cognate, with a similar function in a few related language varieties.

4.1. Inverse marking is for non-A arguments

While inverse marking is cross-linguistically rare, it is not uncommon in Tibeto-Burman, as has been observed already by Ebert (1987: 477). We can divide inverse systems into those that are canonical, and those that are not. A canonical inverse system marks verbs when P outranks A on a referential hierarchy (Jacques and Antonov 2014, Table 2). In Muklom, the marker $-p^h$ extends beyond that domain. If inverse marking would be triggered by a person hierarchy $1/2 > 3$, we would not expect the inverse to show up in the local domain, but it does. Similarly, we could propose a person hierarchy $1 > 2 > 3$, but this would not explain the presence of the inverse marker in the $1 \rightarrow 2$ configuration. And again, if we proposed a hierarchy $2 > 1 > 3$, this would not explain the presence of the inverse marker in the $2 \rightarrow 1$ configuration. In other words, we cannot account for the distribution of the inverse marker in Muklom by evoking a person hierarchy; the language exhibits a non-canonical inverse system. This is not a rare feature among languages that possess an inverse marker. Highly uncanonical systems, which extend the use of the inverse in different directions, are not uncommon (Jacques and Antonov 2014). The function of the inverse in Muklom, as shown in Table 5 above, can be described as follows: $-p^h$ serves to mark non-A SAPs.

When participants in a conversation are encoded in an utterance with a semantic role that is non-A, namely P, R, possessor or perhaps even other roles, then the verb must be marked with $-p^h$. Examples will be provided for each of these three non-A roles. First, a sentence with an SAP P argument that triggers the appearance of $-p^h$ on the verb:

- (17) $n\eta^2-a^3$?i $wat-ta\text{?}-p^h-i^3$.
 3PL-ERG 1PL hit-PST-SAP-1PL
 ‘They have hit us’ (B1.86)

Second, an example of an SAP encoded as R that triggers the appearance of *-p^h* on the verb. This is the same sentence that was used before to illustrate that R arguments can trigger P-N indexing in Muklom.

- (18) *nɪŋ²-a³* *ŋa¹-ma³* *bɔ³-ti¹* *kɔ²-ta²-p^h-aŋ².*
 3PL-ERG 1SG-DAT papaya-CLASS give-PST-SAP-1SG
 ‘They gave papaya to me.’ (B1.86)

The third category that will be illustrated here, is different from the first two in that it does not represent a core argument. The following sentence shows how even an SAP encoded as possessor, which is not a clausal argument, can trigger the appearance of the suffix *-p^h* on the verb.

- (19) *ba³* *k^hak-waŋ³-na²* *ŋm²* *t^hΛ-lat* *ka¹-p^h-a³.*
 2SG.POSS bag-ABL-LOC money MED-pull.out go-SAP-3
 ‘Money is sticking/coming out of your bag.’ (B1.133)

In the above sentence, the first word *ba³* ‘your’, referring to the addressee as possessor, triggers SAP marking on the verb. It is not for the first time that non-A SAPs are found to trigger marking on the verb in a TB language of Northeast India. For example, Konnerth (2014: 223) describes how the morpheme *nang=*, which attaches to the verb, marks non-A SAPs in the Karbi language. Moreover, just as in Muklom possessors may trigger this SAP marker (Konnerth 2014: 228).

R arguments which trigger P-N indexing and possessors which make the verb carry an inverse or non-A SAP marker bring into mind what Bickel (2000) refers to as ‘associative agreement’. In Muklom, there is a disintegration of NP and verb marking. The system of expressing semantic roles on NPs by case suffixes does not align nicely with the system of P-N indexing and inverse marking on the verb. A language that aligns the two systems is said to exhibit ‘integrative agreement’ (Bickel 2000). Integrative verb inflection is, for example, often found among Indo-European languages. Tibeto-Burman languages commonly show the pattern of associative verb inflection as defined by Bickel.

4.2. Remote SAPs

It should be mentioned that even though SAPs are typically physically present at the moment of speaking, they need not always be. For example, when a speaker addresses an entity in the spiritual world, the addressee functions as a full SAP. Whether the conversation partner is located in the physical or spiritual realm does not affect SAP status. Nor does separation in space seem to affect SAP status, as long as there is an audial connection, such as through a telephone. In telephone conversations, the addressee functions as a full SAP. Similarly, separation in time does not have an effect on SAP status, as long as there is a delayed audial connection. Examples of the latter are recorded messages that are after some time heard by the addressee, such as a voice mail message.

4.3. Potentially cognate inverse markers in related language varieties

The inverse marker *-p^h* present in Muklom is unlikely to have been an independent innovation restricted to this language variety. At least some related varieties have been reported to have similar looking markers, perhaps cognate, with related functions. Three of those will be

discussed here: 1) the marker *pha* in Cholim (Tangsa), 2) the marker *-h* in Phong (Tangsa/Nocte), 3) a Nocte variety described by Das Gupta (1971).

Morey (2011a: 90) discusses the marker *pha* in Cholim Tangsa. He suggests this Cholim form and Muklom *-p^h* may be cognates. There is only example of *pha* in the texts that he recorded, which is copied below (original glosses).

- (20) *a²-βe¹ miŋ² chuŋ¹ βɾɳ² pha²...*
 1SG-father name tell COS IMP
 ‘Tell us our father’s name ...’ (Morey 2011a: 90)

Morey analyses the morpheme as a first person inclusive marker, then glosses it as imperative as the above example shows, but the actual description he provides of its function may in fact suggest that also Cholim *pha* is an SAP marker: “It seems that *pha* is always inclusive of everyone present, either as actors or undergoers” (Morey 2011a: 90). The word ‘present’ indicates that the conversation, who are participants in the speech act (SAP) and who are not, in some way controls the use of *pha*.

Phong is a Tangsa or Nocte variety for which more information is becoming available. In this language variety, some verbs must carry the suffix *-h*. This morpheme is labelled ‘inverse marker’ by Dutta (2016). The morpheme attaches to the verb to signal that P outranks the A on the person hierarchy. For inverse marking in Phong, first and second person are not treated as of equal rank, but instead first is considered higher on the hierarchy than second person, as illustrated by the following figure.

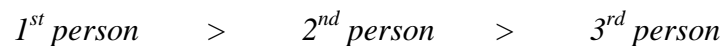


Figure 4 – A person hierarchy for the inverse in Phong (Nocte/Tangsa)

The examples that Dutta (2016) shares, show that at least in the past tense the suffix *-h* occurs whenever second person acts upon first person (2→1), or third person acts upon first or second person (3→1; 3→2). Those are combinations in which the event expressed by the verb is directed towards the higher end of the hierarchy in Figure 4. No marker is present when the A outranks the P (1→2, 1→3 and 2→3).

The Nocte variety described by Das Gupta (1971) exhibits an inverse marker phonologically identical to Phong. The examples provided by Das Gupta show that constructions with P outranking A on the person hierarchy (here 2→1, 3→1, 3→2) require the presence of the inverse suffix *-h* that on the verb. The following sentence is an example of the combination third person A, first person R (glosses by current author). Notice how in this Nocte example, just as in Muklom, a first person R is allowed to trigger indexing on the verb.

- (21) *ate-ma nga-nang ko-t-h-ang.*
 3SG-CASE 1SG-CASE give-PST-INV-1SG
 ‘He gave to me.’ (Das Gupta 1971: 20)

To conclude, the markers *-p^h* in Muklom, *pha* in Cholim, *-h* in Phong and the Nocte variety reported by Das Gupta (1971) show striking similarities in form and function. It has been suggested before that the Muklom and other Tangsa markers could be related (Morey 2011a: 90). Whether these forms are indeed cognate or not has to be tested through thorough comparison and historical reconstruction. It will be interesting to see whether these markers constitute a shared innovation within (certain) Nocte and Tangsa varieties and if so, whether this innovation is restricted to these groups or whether cognates can be found in other TB

branches. To reach this point of knowledge, we will need to increase our linguistic understanding of the diverse Tangsa varieties.

5. Conclusion

Based on newly collected data, it has been shown that Muklom Tangsa exhibits P-N indexing with hierarchical alignment in the mixed domain. The underlying hierarchy can be referred to as a person hierarchy and takes the following shape: 1/2>3 or SAP>3. P and R arguments may trigger indexing on the verb if they outrank the A argument of the clause. First and second person are assigned equal status on the hierarchy. Although higher ranked arguments normally trigger indexing in the mixed domain, we do also encounter indexing of the lower-ranked argument. Further investigation must determine which factors, perhaps topicality or subordination, allow third person to be indexed in these configurations. We have also shown that semantic role determines index alignment in the local domain: the A argument must be indexed in the local domain.

Besides hierarchical indexing, the verb paradigm includes a morpheme that can be called an inverse marker. The distribution of the inverse suffix $-p^h$ is such, that it can also be called a non-A SAP marker. The inverse is not restricted to the core arguments P and R. Non-arguments, at least possessors but possibly other roles as well, may trigger the appearance of $-p^h$ on the verb. It has also been shown that inverse marker may be dropped under certain circumstances in 3→SAP configurations. Further investigation must determine which factor allows the inverse to be dropped. Perhaps this occurs when the verb is located in a subordinate clause.

Syntactic and pragmatic factors need to be explored to explain the variation found in hierarchical indexing and the inverse, but also the possibility of language change has to be taken into account. Anecdotal evidence suggests older speakers are more prone to use the inverse in controlled elicitation than younger speakers. To arrive at a fuller understanding of the mechanisms behind indexing and inverse marking, the current dataset should be extended to include a considerable amount of less controlled elicitation such as storytelling or free discussions.

Muklom P-N indexing and inverse marking have been compared to a few related varieties and one cannot escape the impression that at least certain Nocte varieties show a high degree of similarity to Muklom with respect to verb inflection. It would be interesting to investigate how closely related the different Tangsa and Nocte varieties are.

Abbreviations

1	first person	LOC	locative
2	second person	MED	middle voice
3	third person	NEG	negation
ABL	ablative	PL	plural
CASE	case marker, undefined	P-N	person-number indexing
COS	change of state	POSS	possessive
DAT	dative	PST	past tense
ERG	ergative	SAP	speech act participant
IMP	imperative	SG	singular
INV	inverse	TB	Tibeto-Burman

References

- Bandyopadhyay, Swapon Kumar. (1989). "A Tangsa wordlist." *Linguistics of the Tibeto-Burman Area* 12(2): 79–91.
- Bauman, James. (1975). *Pronouns and Pronominal Morphology in Tibeto-Burman*. PhD Dissertation. University of California, Berkeley.
- Benedict, Paul K. (1972). *Sino Tibetan. A conspectus*. Cambridge, Cambridge University Press.
- Bickel, Balthasar. 2000. "On the syntax of agreement in Tibeto-Burman." *Studies in Language* 24(3): 583–610.
- Blench, Roger and Mark W. Post. (2013). "Rethinking Sino-Tibetan phylogeny from the perspective of North East Indian languages." In Nathan W. Hill and Thomas Owen-Smith, Eds. *Trans-Himalayan Linguistics*. Berlin: Mouton de Gruyter, 71–104.
- Boersma, Paul and David Weenink. (2017). *Praat: doing phonetics by computer* [Computer program]. Retrieved from <http://www.praat.org/>.
- Boro, Krishna. (2012). "Verb agreement in Hakhun Tangsa." *Paper presented at the 18th Himalayan Languages Symposium, Banaras Hindu University, Varanasi, September 10-12*.
- Bradley, David. (1997). "Tibeto-Burman languages and classification." In David Bradley, Ed. *Papers in Southeast Asian linguistics No. 14: Tibeto-Burman languages of the Himalayas*. Canberra, Pacific Linguistics Research School of Pacific and Asian Studies, Australian National University: 1–72.
- Burling, Robbins. (1983). "The Sal languages." *Linguistics of the Tibeto-Burman Area* 7(2): 1–31.
- _____. (2003). "The Tibeto-Burman languages of Northeastern India." In Graham Thurgood and Randy J. LaPolla, Eds. *The Sino-Tibetan Languages*. London, Routledge: 169–191.
- Das Gupta, K. (1971). *An introduction to the Nocte language*. Shillong: Northeast Frontier Agency.
- _____. (1980). *The Tangsa language. A synopsis*. Shillong: North East Indian Frontier Agency.
- DeLancey, Scott. (1980). *Deictic Categories in the Tibeto-Burman verb*. PhD Dissertation. Bloomington, Indiana University.
- _____. (1981). "The Category of Direction in Tibeto-Burman." *Linguistics of the Tibeto-Burman Area* 6(1): 83–101.
- _____. (2010). "Towards a history of verb agreement in Tibeto-Burman." *Himalayan Linguistics* 9: 1–38.
- _____. (2011). "Nocte and Jinghpaw. Morphological correspondences." In Gwendolyn Hyslop, Stephen D. Morey and Mark W. Post, Eds. *North East Indian Linguistics. Volume 3*. New Delhi, Cambridge University Press India: 61–75.
- _____. (2017). "Hierarchical and accusative alignment of verbal person marking in Trans-Himalayan." *Journal of South Asian Languages and Linguistics* 4(1): 85-105.
- Dutta, Niharika. (2016). "Tense-aspect marking in Phong." *Paper presented at the 9th International Conference of North-east Indian Linguistic Society, Tezpur University. Tezpur, Assam, India, February 5-7*.
- Ebert, Karen H. (1987). "Grammatical marking of speech act participants in Tibeto-Burman." *Journal of Pragmatics* 11: 473-482.
- French, Walter T. (1983). *Northern Naga. A Tibeto-Burman mesolanguage*. PhD Dissertation. City University of New York.

- Jacques, Guillaume and Anton Antonov. (2014). "Direct/inverse systems." *Languages and Linguistics Compass* 8(7): 301-318.
- Konnerth, Linda A. (2014). *A grammar of Karbi*. PhD Dissertation. Eugene, University of Oregon.
- _____. (2015). "A new type of convergence at the deictic center. Second person and cislocative in Karbi (Tibeto-Burman)." *Studies in Language* 39(1): 24-45.
- Morey, Stephen D. (2011a). "Tangsa agreement markers." In Gwendolyn Hyslop, Stephen Morey and Morey Post, Eds. *North East Indian Linguistics Vol. 3*. New Delhi, Cambridge University Press, India: 76-101.
- _____. (2011b). "Transitivity in Cholim Tangsa." *Studies in Language* 35(3): 676-701.
- _____. (2014). "Studying tones in North East India. Tai, Singpho and Tangsa." *Language Documentation and Conservation* 8: 637-671.
- _____. (2015). "The internal diversity of Tangsa: vocabulary and morphosyntax." In Mark W. Post, Stephen Morey and Scott DeLancey, Eds. *Language and culture in Northeast India and beyond: In honour of Robbins Burling*. Canberra, Australia-Pacific Linguistics: 23-40.
- _____. (2016). "Verbal agreement marking in Tangsa/Tangshang Languages." *Paper presented at the New Fields in Morphology Workshop, Centre for Excellence on the Dynamics of Language, Melbourne University. November 29*.
- _____. (2017). "Tangsa". In Graham Thurgood and Randy LaPolla, Eds. *The Sino-Tibetan Languages*. London and New York, Routledge: 350-368.
- Ngemu, T. (1977). *Moklum language guide*. Shillong, Government of Arunachal Pradesh.
- Shafer, Robert. (1953). "Classification of the northern-most Naga languages." *Journal of the Bihar Research Society* 39(3): 225-264.
- _____. (1955). "Classification of the Sino-Tibetan languages." *Word: Journal of the Linguistic Circle of New York* 11: 94-111.
- Siewierska, Anna. (2013). "Alignment of verbal person marking." In Matthew S. Dryer and Martin Haspelmath, Eds. *The world atlas of language structures online*. Leipzig, Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/100>, Accessed on 2017-06-01.)
- Sumbatova, Nina R. and Rasul O. Matulov. (2003). *A grammar of Icari Dargwa*. München, Lincom Europa.

Verb-pronominal agreement in Mara

Jyoti Sharma
University of Delhi

Abstract *This paper aims at giving a detailed description of pronominal agreement marking on verbs in Mara, a Chin language of the Tibeto-Burman family. In the Mara agreement system, the verb agrees with both the pronominal participants of the clause. It also shows a person hierarchy of 1st>2nd>3rd when marking agreement. There are certain features in the Mara system that make it complex and interesting. All these features will be discussed in this paper in detail.*

Citation Sharma, Jyoti. 2018. *Verb-pronominal agreement in Mara*. North East Indian Linguistics (NEIL) 8, 156-171. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.

Volume Editors Linda Konnerth, Stephen Morey, Amos Teo

Copyright © 2018, the author(s), release under Creative Commons Attribution license

URL <http://hdl.handle.net/1885/155280>

1. Introduction

Mara is one of the Maraic languages, a sub-group of the Kuki-Chin branch, spoken primarily in the south-eastern part of the state of Mizoram (India) and in the adjacent area of the Chin state of Myanmar. The language is named after the name of the community itself. The Mara tribe is said to be sub-grouped into eight tribes, out of which five tribes are found in India, namely *Tlosai*, *Hawthai*, *Zyhno*, *Chapi* and *Vytu*; and other three tribes are found in Myanmar (Peter Beita, p.c.). ‘Mara’ is used as an umbrella term for all these sub-tribes. Though the variety of Mara spoken by these tribes shows some variations, it is mutually intelligible among them. In Indian side Mara spoken by Tlosai tribe is regarded as the standard form and is used as the language of text and teaching in the schools. Earlier, Maras were also known by different names. Some of them were Lakher, Zo, Sendhu, Maring, Zyu etc. One widely used name by the neighbors to acknowledge the Mara community was Lakher (mainly given by Lushai/Mizo), which was also adopted by British authorities.

Mara is spoken by Mara people living in 60 villages of Chhimtuipui district (one of the three original districts of Mizoram) and the whole area is commonly referred as Maraland. Under the sixth schedule of Constitution of India an Autonomous Council was established for Mara people. It came to be known as Lakher Autonomous District council. After 1988 the name was changed to Mara Autonomous District council. It is one of the three autonomous councils of Mizoram; the other two are: Lai Autonomous District Council (LADC) and Chakma Autonomous District Council (CADC). Like LADC and CADC, MADC enjoys vast legislative, executive and financial power within their jurisdiction as it enacts various Acts, Rules and regulations. It covers around 1445 square kilometres.¹ According to 2012 census report conducted by MADC the population of Mara speaker is 60,453 out of which 31,025 are males and 28,996 are females.

Several classifications are given by the linguists for Tibeto-Burman group and Kuki-Chin branch. Some previous studies put Mara under the Central Chin. However VanBik (2009) has divided Kuki-Chin branch into Northern Chin, Old Chin, Central Chin, Southern Chin and Maraic as shown in Figure 1. He has put Mara under Maraic.

¹ <https://madc.mizoram.gov.in/>

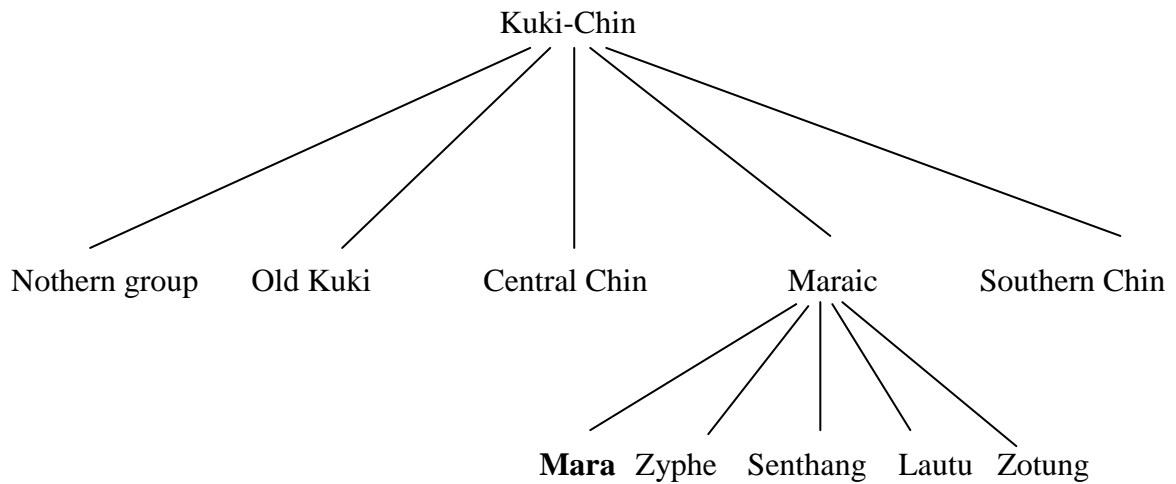


Figure 1: The classification of Kuki-Chin branch

Though very few, there are some work done on Mara. Early studies of Mara (Lakher) were conducted by the British scholars. The first Mara Grammar book, ‘*A Grammar and dictionary of Lakher Language*’ was written by Fred W. Savidge and published in 1908. N.E. Parry published the monograph in 1932, which is still in use. Further R.A Lorrain, with the release of his book ‘*Grammar and Dictionary of Lakher or Mara Language*’ in 1951, contributed to the development of the writing system of Mara. Lorrain (1951) and Savidge (1908) both have documented Mara, but according to Loffler (2002) their works emphasized more on the phonology and phonetics of the language; however they have given partial description of the agreement system. Arden (2010) has worked on *Sabeu*, a variety of Mara that is spoken in Burman side.² Another recent work is by Ngun TinPar (2016), who has worked on Senthang, one of the Maraic languages spoken in the Chin state of Myanmar. With this I would also like to mention that Tlosai Mara was also taken up for the study by the M.A Field-methods class (2014) conducted in Delhi University. There are 23 M.A dissertations submitted by the students on the various aspects of Mara grammar.³

This paper focuses on Mara spoken by the Tlosai tribe. This paper is limited to the elicited data. Data is gathered from the language consultant Delson Notlia (26) from Siaha district of Mizoram who was available in Delhi. He is a native speaker of Tlosai Mara. He is also well-versed in Mizo and English.

2. Preliminary information of grammar

Mara is a predicate-final language and the word order is SV for intransitive clause and APV for transitive clauses as shown in sentences 1 and 2. According to the typological study of agreement patterns in South Asian Languages (SALs) by K.V Subbarao (2012), some Tibeto-Burman languages show an agreement system in which various argument are embodied in the verb and the presence or absence of case markers in noun phrase does not influence the

² The pronominal marking system in *Sabeu* variety, as described by Arden (2010), resembles to the Tlosai variety where S/A and P are marked on the verb and follows the person hierarchy of 1st > 2nd > 3rd. However there are some clear differences between her and my analysis. She considers the agreement markers as ‘agreement words’; but I regard them as the clitic form of the independent pronouns on the basis of two factors: a) Most of the Clitic forms are identical to their respective independent forms. b) They cannot occur independently; they need a host to attach on. However this issue needs further investigation. For this paper I will consider agreement markers as the clitic forms of the independent pronominals.

³ I would like to acknowledge my M.phil Supervisor Dr. Gail Coelho, who encouraged me to take up the language for further study and helped me understand the concepts and nature of the Kuki-Chin languages.

marking of arguments on the verb. He considers this pattern as *Type C*. Mara also shows these two traits of the *Type C* agreement pattern. Mara shows the agreement of participants' number and person feature with the verb and presence or absence of the case markers does not stop participant from showing/marking agreement on the verb.

Mara exhibits ergative-absolutive case marking pattern. Ergative case is marked overtly and absolutive case is not marked. Therefore in sentences 1 and 2, S of intransitive verb and P of transitive verb are unmarked, whereas A is marked with the ergative case marker *-tə*.

(1) *lyna ə=pəli*
Lyuna 3=vomit
'Lyuna vomited.'

(2) *lyna-tə rapha ə=ŋəʊ*
Lyuna-ERG Rapha 3.S=see
'Lyuna saw Rapha.'

Mara only makes two-ways distinction in marking tense, i.e. future and non-future. The future tense is morphologically marked in the verb-complex by *=ə* as shown in (5). Mara doesn't distinguish between past and present as in (3) and (4). Note that for the present paper all the examples provided are in past tense form.

(3) *zəŋia rap^ha-tə pəti ə=ne*
yesterday Rapha-ERG rice 3.A=eat
'Rapha ate rice yesterday.'

(4) *ənaʊ tʃərejtə rap^ha-tə pəti ə=ne=ty*
every day Rapha-ERG rice 3.A=eat=HAB
'Rapha eats rice everyday.'

(5) *rap^ha-tə pəti ə=ne=ə*
Rapha-ERG rice 3.A=eat=FUT
'Rapha will eat rice.'

Mara is one of the Kuki-Chin languages that has gone through the process of pronominalization, in which the independent pronouns of the Proto form have been cliticized and used as the agreement elements on the verb by the language. So there are two major kinds of person pronouns in Mara, i.e., independent person pronouns and clitic person pronouns. The table 1 shows the set independent pronouns.

Table 1: Independent person pronouns

	SG	PL
1 st person	kej	kejmaʊ
2 nd person	nənaʊ	nəmaʊ
3 rd person	ənaʊ	əmaʊ

Clitic forms of these pronouns, the semantics they encode and the order in which they occur in the verb-complex to mark the agreement are discussed in details in the upcoming sections.

The main aim of this paper is to provide detailed description of verb-pronominal agreement in the Tlosai variety of Mara. Like other Kuki-Chin languages, Mara shows an agreement system where verb agrees with the subject and /or object in number and person. It also shows the person hierarchy, i.e. 1st person is higher than 2nd and 3rd, and 2nd person is higher than 3rd person (1st>2nd>3rd). The number distinction is only between singular and plural; and the inclusive/exclusive distinction is also not present in the language.

3. Verb-pronominal agreement in Mara

There are two different agreement paradigms for positive and negative verbs. Thus the agreement marking and the order in which agreement clitics occur in a verb complex plays an important role in signaling the affirmation or negation of the proposition.

3.1. Positive paradigm

Mara positive agreement paradigm is predominantly pre-verbal with some exceptions. This section presents agreement marking in the positive paradigm. Section 3.1.1 gives a detailed description of singular and plural intransitive paradigm, followed by section 3.1.2 which gives a detailed description of singular and plural transitive paradigm.

3.1.1. Intransitive paradigm

S is marked on verb by pre-verbal pronominal clitics. The pronominal agreement clitics for S in the intransitive clause are: *ej=* for '1st person' *nə=* for '2nd person' and *ə=* for '3rd person', as shown in Table 2 and sentences (6)-(8).

Table 2: Positive paradigm for intransitive verbs: Singular pronominal argument

1 st S	<i>ej=</i> Σ
2 nd S	<i>nə=</i> Σ
3 rd S	<i>ə=</i> Σ

(6) *zəŋija* *ej=pəŋej*
 yesterday 1.S=laugh
 'I laughed yesterday.'

(7) *zəŋija* *nə=pəŋej*
 yesterday 2.S=laugh
 'You laughed yesterday.'

(8) *zəŋija* *ə=pəŋej*
 yesterday 3.S=laugh
 'He laughed yesterday.'

When S is plural, plurality is marked by *mə=* immediately after the S clitic. The intransitive verbs with plural pronominal making are shown in Table 3 and sentences (9)-(11).

Table 3: Positive paradigm for intransitive verbs: Plural pronominal argument

1 st S	ej=mə=Σ
2 nd S	nə=mə=Σ
3 rd S	ə=mə=Σ

- (9) *ej=mə=pəŋej*
 1.S=PL=laugh
 ‘We laughed.’
- (10) *nə=mə=pəŋej*
 2.S=PL=laugh
 ‘You (PL) laughed.’
- (11) *ə=mə=pəŋej*
 3.S=PL=laugh
 ‘They laughed.’

3.1.2. Transitive paradigm

3.1.2.1. Transitive verb agreement with singular pronominal

In the transitive sentences, agreement on verb shows complexities because both A and P (except 3rd P) are marked and in varied ways. In a positive singular paradigm the agreement for both A and P is marked pre-verbally except for 2nd person A when the P is 1st person, as shown in Table 4. Unlike intransitive verb, 2nd person A has two markers; they are: the regular 2nd person clitic *nə=* which is marked pre-verbally when P is 3rd person and *=tʃi* which is marked post-verbally when P of the clause is 1st person. Pronominal clitics for 1st A and 3rd A are same as the 1st S and 3rd S in the intransitive clause. The P clitics are: *ej=* for ‘1st person’ and *tʃə=* for ‘2nd person’, while 3rd person is unmarked. The singular transitive paradigm of Lakher (present time Mara) recorded by Weidert (1985:929) given in Delancey (1989:330) is similar to the one shown in Table 4.

Table 4: Positive paradigm for transitive verbs: Singular pronominal arguments

	1 st P	2 nd P	3 rd P
1 st A		ej=tʃə=Σ	ej=Σ
2 nd A	ej=nə=Σ=tʃi		nə=Σ
3 rd A	ej=nə=Σ	ə=tʃə=Σ	ə=Σ

Based on columns 2 and 3, the generalized pattern of agreement marking on verbs in the positive paradigm is: **A=(P)=V**, where 3rd person P is never marked.⁴ Sentences (12) - (16) show the order of agreement marking in the verb complex.

⁴ This is the case with 3rd singular objects only – plural objects show certain differences in agreement marking.

- (12) *zəŋija* *ej=ɸə=ŋɔɔ*
yesterday 1.A=2.P=see
‘I saw you yesterday.’
- (13) *zəŋija* *ej=ŋɔɔ*
yesterday 1.A=see
‘I saw him/her/somebody yesterday.’
- (14) *zəŋija* *nə=ŋɔɔ*
yesterday 2.A=see
‘You saw him/her/somebody yesterday.’
- (15) *zəŋija* *ə=ɸə=ŋɔɔ*
yesterday 3.A=2.P=see
‘He saw you yesterday.’
- (16) *zəŋija* *ə=ŋɔɔ*
yesterday 3.A=see
‘He saw him/her/somebody yesterday.’

Column 1, however, shows a different ordering. When the P is 1st person, the order changes such that P alone is marked before the verb, while 2nd A is marked after the verb; the 3rd A is left unmarked. Thus, the order with 1st person P is **P= V= (A)** as evident in the following sentences (17) and (18). Further, when the order is **P= V= (A)**, a particle *nə=* appears immediately following the 1st person agreement clitic *ej=*. Given that *nə=* occurs only when the normal order is violated, it seems to me that it could be analysed as patient marker for 1st person, because its occurrence semantically points out that the *ej=*, which is occurring in the A position of agreement marking is not A, rather it is P of the clause. However this marker does not otherwise appear in the Mara agreement system (not even for plural pronominals).

- (17) *zəŋija* *ej=nə=ŋɔɔ=ɸi*
yesterday 1=1.P=see=2.S
‘You saw me yesterday.’
- (18) *zəŋija* *ej=nə=ŋɔɔ*
yesterday 1=1.P =see
‘S/He saw me yesterday.’

3.1.2.2. Transitive verb agreement with plural pronominals

Like singular paradigm the A agreement clitics are *ej=* for 1st person, *nə=* (pre-verbal) and *=ɸi* (post-verbal) for 2nd person, *ə=* for 3rd person; and P clitics are *ej=* ‘1st person’, *ɸə=* ‘2nd person’ and 3rd P is never marked. The plurality of A is marked by the pre-verbal plural clitic *mə=* that occurs immediately after the respective A agreement clitic as shown in (19) – (23). Table 5 shows the order of agreement marking of plural A in the verb complex. Based on Column 2 and 3 the ordering is:

Agent= Plural=(Patient)=Verb

However there is another plural marker =ej that is post-verbal which comes into action when A of the clause is 2nd plural or 3rd plural and P is 1st person as evident in column 1 and sentences (24) and (25). Though 3rd person A is not marked on the verb when the P is 1st person, but the plurality of 3rd A is always marked by the plural clitic =ej. So for column 1 where A is 2nd plural or 3rd plural and the P is 1st the ordering is:

$$Patient=Verb=Plural=(Agent)$$

Table 5: Positive paradigm for transitive verbs: Plural agent and singular patient

	1 st SG P	2 nd SG P	3 rd SG P
1 st PL A		ej=mə=ʈə=Σ	ej=mə=Σ
2 nd PL A	ej=nə=Σ=ej=ʈi		nə=mə=Σ
3 rd PL A	ej=nə=Σ=ej	ə=mə=ʈə=Σ	ə=mə=Σ

- (19) *ej=mə=ʈə=ŋɔʋ*
 1.A=PL.A=2.P=see
 ‘We saw you.’
- (20) *ej=mə=ŋɔʋ*
 1.A=PL.A=see
 ‘We saw him/her/something.’
- (21) *nə=mə=ŋɔʋ*
 2.A=PL.A=see
 ‘You (pl) saw him/her/something.’
- (22) *ə=mə=ʈə=ŋɔʋ*
 3.A=PL.A=2.P=see
 ‘They saw you.’
- (23) *ə=mə=ŋɔʋ*
 3.A.PL.A=see
 ‘They saw him/her/something.’
- (24) *ej=nə=ŋɔʋ=ej=ʈi*
 1=1.P=see.V=PL.A=2.A
 ‘You (pl) saw me.’
- (25) *ej=nə=ŋɔʋ=ej*
 1=1.P=see.V=PL.A
 ‘They saw me.’

In the case when both participants of the clause are plural then plurality of A is marked by the pre-verbal plural clitic mə= and plurality of P is marked by the post-verbal plural clitic =ej except for 1st person because Mara has a different clitic form for 1st plural patient i.e, mənijə=. However the plurality of 2nd A and 3rd A is also marked by =ej, but only when the P

of clause is 1st person (also evident in table 5). The plural pronominal agreement on verb is shown in the Table 6.

Table 6: Positive paradigm for transitive verbs: Plural agent and plural patient

	1 st PL P	2 nd PL P	3 rd PL P
1 st PL A		ej=mə=ʈʃə=Σ=ej	ej=mə=Σ=ej
2 nd PL A	mənijə=Σ=ej=ʈʃi mənijə=nə=mə=Σ(=ej=ʈʃi)		nə=mə=Σ=ej
3 rd PL A	mənijə=ə=mə=Σ	ə=mə=ʈʃə=Σ=ej	ə=mə=Σ=ej

Based on columns 2 and 3 of Table 6, the pattern of plural marking in the positive transitive clauses is:

$$Agent = Agent\ Plural = (Patient) = Verb = Patient\ Plural$$

Though the 3rd P is not marked, plurality of 3rd P is always marked on the verb. Sentences (26) – (28) show the ordering of person and number agreement clitics on the verbs.

(26) *ej=mə=ηɔʊ = ej*
1.A=PL.A=see.V=PL.P
'We saw them.'

(27) *nə=mə=ηɔʊ = ej*
2.A=PL.A=see.V=PL.P
'You (PL) saw them.'

(28) *ə=mə=ηɔʊ = ej*
3.A=PL.A=see.V=PL.P
'They saw them.'

Column 1 of Table 6, however shows some interesting differences between singular paradigm and plural paradigm. First, when the A is 2nd person and P is 1st person plural the ordering of agreement clitics on verb is flexible:

- It may be **P=A=V**, in which case the 2nd A agreement clitic is *nə=* that is marked pre-verbally, as in (29),
- or **P=V=A**, in which case the 2nd A agreement clitic is *=ʈʃi* that is marked post-verbally, as in (30),
- or both of these clitics for 2nd A can be marked in the same verb complex, evident in (31).

(29) *mənijə=nə=mə=ηɔʊ*
1.PL.P=2.A=PL.A=see
'You (PL) saw us.'

(30) *mənijə=ηɔʊ=ej=ʈʃi*
1.PL.P=see=PL.A=2.A
'You (PL) saw us.'

- (31) *mənijə=nə=mə=ŋɔʋ=ej=ʈfi*
 1.PL.P=2.A=PL.A=see=PL.A=2.A
 ‘You (PL) saw us.’

Second, when P is 1st person PL, the order of agreement marking is **P=A=V**. It means when the **A is 3rd person** (SG or PL) and **P is 1st PL**, as in (32) and (33), A is not left unmarked on the verb complex unlike in singular paradigm.

- (32) *mənijə=ə=mə=ŋɔʋ*
 1.PL.P=3.A=PL.A=see
 ‘They saw us.’

- (33) *mənijə=ə=ŋɔʋ*
 1.PL.P=3.A=see
 ‘He saw us.’

3.1.3. Summary: Positive paradigm

The intransitive and transitive paradigm above show that Mara, like many other TB languages, displays the person hierarchy of 1st > 2nd > 3rd. 1st person regardless of A or P is marked first and precedes all the other agreement clitics. The marker *nə=* immediately follows the 1st person clitic *=ej* if it is an under-goer of the action. 2nd person is always marked (*nə=* and *=ʈfi* for A and *ʈə=* for the P); it is marked postverbally if the P is 1st person. 3rd A is marked only if the P is not 1st person; further, 3rd person P is never marked. However plurality of the pronominals for both A and P are always marked on the transitive verb, even for the 3rd person.

3.2. Negative paradigm

DeLancey (2013a, b) reports several Kuki-Chin languages like Rangkhol, Hallam and Pūrūm (Grierson 1904) of North-Western group (Old Kuki-Chin group), Tedim (Hederson 1957, 1965) and Sizang (Stern 1963) of North-Eastern group and Hyow Chin (Peterson 2003) and Daai Chin (So-Hartmann 2009) of Southern group, that have special negative agreement paradigm. It means these languages employ two different agreement paradigms designed for positive clause and negative clause. This system is also evident in Mara, a Maraic language.

Mara exhibits two agreement paradigms: (a) for positive verbs (as discussed in section 3.1) and (b) for the negative verb. The agreement shown by Mara in negative clauses differs from that of positive clauses as they differ in some agreement clitics for persons and the ordering of these clitics in the verb complex also differs which will be thoroughly discussed below.

Negative agreement paradigm is a simpler system without the ordering differences and alternation in 2nd A pronominal clitic, as in the positive paradigm. Unlike positive agreement paradigm where agreement marking is predominantly pre-verbal (except for 2nd A *=ʈfi* and P PL marker *=ej* that occurs post-verbally), in negative paradigm A is marked post-verbally and P is marked pre-verbally. Negative paradigm shows person hierarchy in a limited way as compared to the positive paradigm. This section presents the pronominal agreement marking in the negative paradigm. Section 3.2.1 gives a detailed description of singular and plural intransitive paradigm, followed by section 3.2.2 which gives a description of singular and plural transitive paradigm.

3.2.1. Intransitive negative paradigm

Table 7 and Table 8 (repetition of Table 2) show the agreement for intransitive verbs in negative paradigm and positive paradigm respectively. Unlike in positive paradigm S agreement clitics for 1st and 2nd are marked post-verbally. Further, the 1st person clitic in negative paradigm is =nə and the 2nd person is =tʃi (like the 2nd S of the positive transitive paradigm), as evident in (34) and (35). The 3rd person appears to be marked with =vej in (36), but as I will argue below, =vej is a negative particle and not an agreement clitic. Its status will be clear after looking at the plural paradigm. Preceding the verb, there is a default marker ə=, which has no discernible meaning – except perhaps to mark negation itself.

Table 7: Negative paradigm for intransitive verbs: Singular pronominal argument

1 st S	ə=Σ=nə
2 nd S	ə=Σ=tʃi
3 rd S	ə=Σ=vej

Table 8: Positive paradigm for intransitive verbs: Singular pronominal argument

1 st S	ej=Σ
2 nd S	nə=Σ
3 rd S	ə=Σ

(34) *zəŋija* ə=pəŋej=nə
yesterday ??=laugh=1.S
'I did not laugh yesterday.'

(35) *zəŋija* ə=pəŋej=tʃi
yesterday ??=laugh=2.S
'You did not laugh yesterday.'

(36) *zəŋija* ə=pəŋej=vej
yesterday ??=laugh=NEGP
'He did not laugh yesterday.'

=vej which looks like 3rd person S agreement clitic in the singular negative paradigm also occurs when S of the clause is 2nd person plural, as it is evident in Table 9 and sentences (38) - (39). So it can be considered as the negative particle and it occurs immediately after the verb. The plurality in negative paradigm is marked by =ej, except for 1st person. The 1st person plural S negative marker is =məpi as in (37).

Table 9: Negative paradigm for intransitive verbs: Plural pronominal argument

1 st PL S	ə=Σ=məpi
2 nd PL S	ə=Σ=vej=ej=ɬi
3 rd PL S	ə=Σ=vej=ej

Table 10 (repetition of Table 2) shows the intransitive verb agreement with plural pronominal subjects in the positive paradigm.

Table 10: Positive paradigm for intransitive verbs: Plural pronominal argument

1 st S	ej= mə=Σ
2 nd S	nə=mə=Σ
3 rd S	ə= mə=Σ

- (37) *zəŋija* ə=pəŋej=məpi
yesterday ??=laugh=1.S.PL
‘We did not laugh yesterday.’
- (38) *zəŋija* ə=pəŋej=vej=ej=ɬi
yesterday ??=laugh=NEGP=PL=2.S
‘You (PL) did not laugh yesterday.’
- (39) *zəŋija* ə=pəŋej=vej=ej
yesterday ?? =laugh= NEG=PL
‘They did not laugh yesterday.’

3.2.2. Transitive negative paradigm

3.2.2.1. Transitive verb agreement with singular pronominal

Pronominal agreement on the verb in the negative transitive clauses is not as complicated as in the positive paradigm. In negative verb complex the order of marking doesn’t change when P is 1st person, as it happens in the positive paradigm. It means the ordering of agreement markers in negative sentence is consistent and does not show any variation as explained below.

The agreement marking clitics are shown in Table 11. The agreement clitics for 1st, 2nd and 3rd A are same as the S clitics of negative intransitive sentences and occur post-verbally. Unlike in the negative intransitive clauses, marker ə= doesn’t occur before the verb in transitive clauses. The P clitics in the negative paradigm are: nə= for 1st person, ɬə= for 2nd person, while 3rd person is unmarked. Like positive paradigm P agreement clitic always occur pre-verbally. The pronominal agreement marking in positive paradigm is again shown in Table 12 (same as Table 4) for the sake of comparison and convenience.

The ordering of markers around the verb in the negative paradigm is (P)=V=A. Order differentiation between the positive and negative clause plays an important role in signalling the negation of the proposition. The hierarchy followed in marking the agreement is 1st and 2nd are greater than 3rd person (1st, 2nd > 3rd), as it is evident in the following examples. Only

1st and 2nd person A and P are marked on the verb, whereas neither 3rd person A nor P is marked on the verb.

Table 11: Negative paradigm for transitive verbs: Singular pronominal arguments

	1 st P	2 nd P	3 rd P
1 st A		tʃə=Σ= nə	Σ=nə
2 nd A	nə=Σ= tʃi		Σ=tʃi
3 rd A	nə=Σ=vej	tʃə=Σ=vej	Σ=vej

Table 12: Positive paradigm for transitive verbs: Singular pronominal arguments

	1 st P	2 nd P	3 rd P
1 st A		ej=tʃə=Σ	ej=Σ
2 nd A	ej=nə=Σ=tʃi		nə=Σ
3 rd A	ej=nə=Σ	ə=tʃə=Σ	ə=Σ

Sentences (40) - (46) show the markers and ordering of markers in the negative paradigm.

- (40) *zanjija* (*kej-tə*) **tʃə=ḡɔɔ=nə**
 yesterday (I-ERG) 2.P=see.=1.A
 ‘I did not see you yesterday.’
- (41) *zanjija* (*kej-tə*) (*anau*) **ḡɔɔ=nə**
 yesterday (I-ERG) (he) see.V=1.A
 ‘I did not see him/her yesterday.’
- (42) *zanjija* (*nanau-tə*) **nə=ḡɔɔ=tʃi**
 yesterday (You-ERG) 1.P=see=2.A
 ‘You did not see me yesterday.’
- (43) *zanjija* (*nanau-tə*) **ḡɔɔ=tʃi**
 yesterday (I-ERG) see=2.A
 ‘You did not see him/her yesterday.’
- (44) *zanjija* (*anau-tə*) (*kej-φ*) **nə=ḡɔɔ =vej**
 yesterday (He-ERG) (I-ABS) 1.P=see=NEG
 ‘S/He did not see me yesterday.’
- (45) *zanjija* (*anau-tə*) **tʃə=ḡɔɔ =vej**
 yesterday (He-ERG) 2.P=see=NEG
 ‘S/He did not see me yesterday.’
- (46) *zanjija* (*anau-tə*) **ḡɔɔ=vej**
 yesterday (He-ERG) see=NEG
 ‘S/He did not see him/her yesterday.’

3.2.2.2. Transitive verb agreement with plural pronominal

The agreement clitics and the order of marking of these clitics on transitive verb by plural participants are shown in Table 13, 15 and 16. 1st A PL agreement clitic is *məpi* and the 1st P PL agreement clitic is *mənijə*. Unlike positive paradigm, plurality of both A and P is marked by the same post-verbal plural clitic =*ej* except for the 1st person.

As evident from Table 13, when both the participants are plural, the plural marking is ambiguous, except when one of the participants is 1st person, as in (47) - (50). In positive paradigm there are two positions for the plural marking in a verb complex when both the participants are plural; the A is marked by *mə=* pre-verbally (except for 2nd person A when the P is 1st) and P is marked by =*ej* post-verbally as evident from Table 14. In the negative paradigm there is only one position for the plural marker even when both the participants are plural; both A and P are marked by same marker =*ej* post-verbally, hence creating the ambiguity as the marker =*ej* in the verb complex may be the number feature of A or P, as evident in (51) - (53).

Table 13: Negative paradigm for transitive verbs: Plural pronominal arguments

	1 st PL P	2 nd PL P	3 rd PL P
1 st PL A		tʃə=Σ=ej=məpi	Σ=ej= məpi
2 nd PL A	mənijə=Σ=vej=ej=tʃi		Σ=vej=ej=tʃi
3 rd PL A	mənijə=Σ=vej=ej	tʃə=Σ=vej=ej	Σ=vej=ej

Table 14 (repetition of table 6) shows the transitive verb agreement with plural pronominals in the positive paradigm.

Table 14: Positive paradigm for transitive verbs: Plural pronominal arguments

	1 st PL P	2 nd PL P	3 rd PL P
1 st PL A		ejmə= tʃə=Σ =ej	ej= mə=Σ=ej
2 nd PL A	mənijə=Σ=ej=tʃi mənijə=nə=mə=Σ(=ej=tʃi)		nə= mə=Σ=ej
3 rd PL A	mənijə=ə=mə=Σ	ə=mə=tʃə=Σ=ej	ə=mə=Σ=ej

The ordering in negative paradigm is consistent, i.e. **P=V=A**. Unlike positive plural paradigm, negative plural paradigm doesn't show variation from the singular paradigm. The pattern of plural marking in the transitive clauses is:

$$Patient=Verb=(NEG)=Plural=Agent$$

- (47) *zəɲija* *tʃə=ŋəʋ=ej=məpi*
 yesterday 2.P=see.PL=1.A.PL
 'We did not see you (pl) yesterday.'

- (48) *zaŋija* $\eta\omega\upsilon=ej=m\grave{a}pi$
yesterday see= P.PL=1.A.PL
‘We did not see them yesterday.’
- (49) *zaŋija* $m\grave{a}nij\grave{a}=\eta\omega\upsilon=vej=ej=t\acute{f}i$
yesterday 1.P.PL =see=NEG=PL=2.A
‘You (pl) did not see us yesterday.’
- (50) *zaŋija* $m\grave{a}nij\grave{a}=\eta\omega\upsilon=vej=ej$
yesterday 1.P.PL = see = NEG=PL
‘They did not see us yesterday.’
- (51) *zaŋija* $\eta\omega\upsilon=vej=ej=t\acute{f}i$
yesterday see= NEG=PL=2.A
‘You did not see them yesterday.’
- (52) *zaŋija* $t\acute{f}\grave{a}=\eta\omega\upsilon=vej=ej$
yesterday 2.P =see= NEG=PL
‘They did not see you (pl).’
- (53) *zaŋija* $\eta\omega\upsilon=vej=ej$
yesterday see= NEG=PL
‘They did not see them.’

Table 15 and 16 however do not show any ambiguity in marking of plural, as one of the participants is singular and other is plural. So the plural clitic =*ej* in these verb complexes is the number agreement feature of the plural participant.

Table 15: Negative paradigm for transitive verbs: Plural agent and singular patient

	1 st SG P	2 nd SG P	3 rd SG P
1 st PL A		$t\acute{f}\grave{a}=\Sigma=m\grave{a}pi$	$\Sigma=m\grave{a}pi$
2 nd PL A	$n\grave{a}=\Sigma=vej=ej=t\acute{f}i$		$\Sigma=vej=ej=t\acute{f}i$
3 rd PL A	$n\grave{a}=\Sigma=vej=ej$	$t\acute{f}\grave{a}=\Sigma=vej=ej$	$\Sigma=vej=ej$

Table 16: Negative paradigm for transitive verbs: Singular agent and plural patient

	1 st PL P	2 nd PL P	3 rd PL P
1 st SG A		$t\acute{f}\grave{a}=\Sigma=vej=ej=n\grave{a}$	$\Sigma=vej=ej=n\grave{a}$
2 nd SG A	$m\grave{a}nij\grave{a}=\Sigma=vej=t\acute{f}i$		$\Sigma=vej=ej=t\acute{f}i$
3 rd SG A	$m\grave{a}nij\grave{a}=\Sigma=vej$	$t\acute{f}\grave{a}=\Sigma=vej=ej$	$\Sigma=vej=ej$

3.2.3. Summary: Negative Paradigm

The person hierarchy is limited to the positive paradigm in Mara; negative paradigm doesn't show person hierarchy. The ordering of agreement particles in the negative verb complex is consistent, i.e. P=V=A, as opposed to the positive clauses in which the ordering is A=P=V (except when the P is 1st person). The A is always marked post-verbally and P is always marked pre-verbally. The markers used in the negative paradigm are mostly similar to the positive paradigm, except for the 1st person. In negative paradigm, 1st person marker is *nə* (which functions as the 1st person P marker in positive paradigm) for both A and P. It can occur pre-verbally or post-verbally depending upon its grammatical function. 2nd A is marked only by post-verbal clitic *=tʃi* and 3rd person (both A and P) is never marked.

Plural marking seems ambiguous in negative paradigm as the plural marker is same, i.e. *=ej*, for A and P except for 1st person. But as illustrated in the above sections Mara shows the person hierarchy in marking the pronominals on the verb, it is expected from the language that plural marking also follows some hierarchy. It means if both the A and P are plural then the presence of plural clitic *=ej* in verb complex denotes the number feature of A (with respect to person hierarchy). Therefore in the above sentences (51) - (53) *=ej* is the number agreement for A. If the A is singular and P is plural then the presence of *=ej* signifies the number agreement of P, as evident in (54) - (56). If A is 1st person PL and P is 2nd or 3rd person PL as evident in sentences (47) - (50), presence of *=ej* is for P rather than A as 1st person plural A and P are marked with *məpi=* and *məniʃə=* respectively.

(54) *zəŋija* *ŋəʋ=vej=ej=tʃi*
 yesterday see= NEG=PL=2.A
 'You did not see them yesterday.'

(55) *zəŋija* *tʃə=ŋəʋ=vej=ej*
 yesterday 2.P=see= NEG=PL
 'He did not see you (all) yesterday.'

(56) *zəŋija* *ŋəʋ=vej=ej*
 yesterday see= NEG=PL
 'He did not see them yesterday.'

Abbreviations

1	First person
2	Second person
3	Third person
S	Single argument of an intransitive verb
A	Agent-like argument of transitive verb
P	Patient-like argument of transitive verb
Σ	Verb
FUT	Future tense
HAB	Habitual aspect
ERG	Ergative case
SG	Singular
PL	Plural
NEG	Negative particle

References

- Arden, Michelle. (2010). *A phonetic, phonological and morphosyntactic analysis of the Mara language*. MA Thesis. San Jose State University.
- DeLancey, Scott. (1989). "Verb agreement in Proto-Tibeto-Burman." *Bulletin of the School of Oriental and African Studies* 52: 315–33.
- _____. (2013a). "The history of postverbal agreement in Kuki-Chin." *Journal of the Southeast Asia Linguistics Society (JSEALS)* 6: 1–17
- _____. (2013b). "Argument indexation (verb agreement) in Kuki-Chin." *Paper presented at the 46th International Conference on Sino-Tibetan Languages and Linguistics, Dartmouth College, Hanover, New Hampshire, USA, August 7-10.*
- Grierson, George, Ed. (1904). *Linguistic Survey of India, Vol. 3, Part 3: Specimens of the Kuki-Chin and Burma Groups*.
- Henderson, Eugénie J. A. (1957). "Colloquial Chin as a pronominalized language." *Bulletin of the School of Oriental and African Studies* 20(2): 323–327.
- _____. (1965). *Tiddim Chin: A descriptive analysis of two texts*. London, Oxford University Press.
- Lorrain, Reginald A. (1951). *Grammar and dictionary of the Lakher language*. Guwahati, Department of Historical and Antiquarian Studies, Government of Assam.
- Löffler, Lorenz. (2002). "Some notes on Mara". *Linguistics of Tibeto-Burman Area* 25: 123–136.
- Ngun Tin Par. (2016). *Agreement and verb stem alternation in Senchang Chin*. MA thesis. Chiangmai, Payap University.
- Parry, Nevill Edward. (1932). *The Lakhers*. Firma KLM Private Limited.
- Peterson, David. (2003). "Agreement and grammatical relations in Hyow." In David Bradley, Randy LaPolla, Boyd Michailovsky and Graham Thurgood, Eds. *Language variation: Papers on variation and change in the Sinosphere in honour of James A. Matisoff*. Canberra, Department of Linguistics, Australian National University: 173–183.
- Savidge, Fred. (1908). *A grammar and dictionary of the Lakher language*. Allahabad: Pioneer Press.
- So-Hartmann, Helga. (2009). *A descriptive grammar of Daai Chin*. (STEDT Monograph 7). Berkeley, Sino-Tibetan Etymological Dictionary and Thesaurus Project, University of California.
- Stern, Theodore. (1963). "A provisional sketch of Sizang (Siyin) Chin." *Asia Major*, n.s., x, 2: 22–78.
- _____. (1984). "Sizang (Siyin) Chin texts." *Linguistics of the Tibeto-Burman Area*, 8(1): 43–58.
- Subbārāo, K. V. (2012). *South Asian languages: A syntactic typology*. New Delhi, Cambridge University Press.
- Thurgood, Graham. (2003). "A subgrouping of the Sino-Tibetan Languages: The interaction between language contact, change, and inheritance." In Graham Thurgood and Randy J. LaPolla, Eds. *The Sino-Tibetan Languages*. London, Routledge: 503–517.
- VanBik, Kenneth (2009). *Proto-Kuki-Chin: A reconstructed ancestor of the Kuki-Chin languages*. (STEDT Monograph 89). Berkeley, Sino-Tibetan Etymological Dictionary and Thesaurus Project, University of California.
- Weidert, Alfons. (1985). "Paradigmatic typology and its application to verb agreement analysis." In Ursula Pieper and Gerhard Stickel, Eds. *Studia Linguistica et Diachronica*. Berlin, Mouton de Gruyter, 903–963.

An introduction to the Nocte verb

Syed Iftiqar Rahman
Payap University

Abstract *Nocte belongs to the Northern Naga subgroup within the Tibeto-Burman language family. Nocte, like Jingphaw and few other Tibeto-Burman languages, shows a hierarchy in the verb agreement marking. Irrespective of the subject or the object argument, in Nocte the verb agrees with the person that is higher in the hierarchy. In Nocte, it is seen that 1st person and 2nd person are higher than the 3rd person in the hierarchy. However, in some cases, it is also possible that apart from tense, person and aspectual information the agreement markings can also carry some pragmatic and contextual information.*

Citation Rahman, Syed Iftiqar. 2018. *An introduction to the Nocte verb*. North East Indian Linguistics (NEIL) 8, 172-184. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.

Volume Editors Linda Konnerth, Stephen Morey, Amos Teo

Copyright © 2018, the author(s), release under Creative Commons Attribution license

URL <http://hdl.handle.net/1885/155280>

1. Introduction

This paper is a description of the verbal agreement system and a brief phonology sketch of the HaꞤwa variety of Nocte, a language variety spoken by the Noctes settled in the Borduria village of Tirap district in Arunachal Pradesh, India.

My paper is mainly based on the data collected from Borduria village. The neighboring Nocte villages are Kaimai, Paniduria, Khonsa, Polung, Laptang, Pansumthong. There are also two Wancho Naga villages in the area namely Lapnan and Lokthong. Wancho is a closely related language variety of the Nocte and this community resides in the Longding district. In Nocte, *nok* means ‘village’ and *te* stands for the people, so *Nocte* refers to the people living in the village. According to a 1971 census, there were 58 Nocte villages with a population of 21,853. And according to a 2001 census the population is about 33,000. There are approximately 29 Nocte villages under the Khonsa division, the district head quarter of Tirap district. The Nocte villages I have visited so far are, Borduria, Kheti, Paltan, Dadam, Thinsa, Polung, New Tupi, Deomali, Hokan and Khonsa.

The Borduria people speak the HaꞤwa variety of Nocte. Hindi is their second language. However, the old generation can speak Assamese very fluently as Assamese was once the medium of instruction in schools. Now Hindi and English are widely used in schools.

2. Linguistic Background

Tirap district is the home to the greater Nocte community. Tirap shares a district border with Changlang and Longding, a state border with Assam and Nagaland, and an international border with Myanmar. There is one Nocte village in the Tinsukia district of Assam called the Paltan Basti or Dihing Kinar Nocte Gaon. Recently, I learned of a few more Nocte villages in the Changlang district as well.

Nocte belongs to the Tibeto-Burman with the code ISO 639:3 njb Naga Nocte. The Ethnologue records the population of the Nocte community to be 33,000 (2001 census) and the language status as 6a (vigorous). Burling (2003) categorizes Nocte together with Bodo-Garo, Koch, Konyak and Jingphaw languages into *Sal* subgroup of Tibeto-Burman languages.

3. Phonology

This paper is an introduction to the Nocte verb. Therefore, a brief phonology sketch is provided here. Tones are not marked throughout the wordlist or in the stories. However, some tonal minimal pairs were elicited during the field work in the Paltan village which is expected to be similar to the Haʔwa variety. A detailed description of the tones needs further research.

3.1. Syllable structure

Syllable structure in Nocte is very simple. The Nocte syllable canon is (C)V(C). A Nocte syllable can consist of just a rhyme with a vowel or an onset and a rhyme with an optional coda. Table 1 below shows the possible syllable types in the language:

Table 1 – Nocte syllable structure

Syllable type	Structure	Example
Type 1	V	<i>i</i> ‘1SG.POSS’
Type 2	CV	<i>ka</i> ‘go’
Type 3	CVC	<i>vet</i> ‘hit’

In Nocte, only the stops [p,t,k,ʔ] and nasals [m,n,ŋ] can occur in the coda. However, I also encountered a complex syllable type [CCVC]. The only two words that I elicited are *dvək* ‘to cut’ and *tvək* ‘to dive’. These two words are the only ones with this syllable structure out of the CALMSEA wordlist of 250 words. But due to the lack of enough data I did not analyze it to be a separate type of syllable structure for now. Therefore, I am concluding the syllable formula for Nocte to be [(C)V(C)].

3.2. Phonemes

This section presents the consonants and vowels charts. Nocte has 20 consonants, 9 monophthong vowels and 5 diphthong vowels. Table 2, 3 and 4 below present the consonants and vowels respectively.

Table 2 – Consonants

	Bilabial	Labio-dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Plosives	p, p ^h , b		t, t ^h , d			k, k ^h , g	ʔ
Nasals	m		n		ɲ	ŋ	
Fricatives Affricates			s, z	tʃ			
Approximants		v	ɹ		j	w	
Lateral approximant			l				

Table 3 – Monophthong vowels

	Front	Back	
	Unrounded	Rounded	
Close	i	ɯ	u
Mid-close	e	ə	o
Mid-open	ɛ		ɔ
Open		a	

Table 4 – Diphthong vowels

Diphthongs	Example
ia	<i>zian</i> ‘be able’
ie	<i>tfien</i> ‘send’
ei	<i>natfei</i> ‘be born’
oi	<i>koi</i> ‘climb’
ua	<i>tfuaŋ</i> ‘tall’
ue	<i>tfuen</i> ‘run’

3.3. Tones

Pitch is contrastive in Nocte. Pitch variation was studied with the help of PRAAT software and also by eliciting tonal minimal sets. There are three tones on smooth syllables. Smooth syllables are syllables with vowel or nasal endings. There is only one tone on closed syllables with /p, t, k, ʔ/ in the coda. Tables 5 and 6 below present the tonal categories along with illustrative examples and Table 7 lists some tonal minimal sets.

Table 5 – Tones on smooth syllables

Tone no.	Word	Transcription	Description of Tone
1	‘teeth’	<i>pa</i>	low-falling
2	‘snake’	<i>pu</i>	high-falling
3	‘sell’	<i>saŋ</i>	mid-level

Table 6 – Tone on checked syllables

Tone no.	Word	Transcription	Description of Tone
4	‘tiger’	<i>saʔ</i>	level with glottal constriction and final stops (p, t, k)
	‘belly’	<i>vok</i>	
	‘sleep’	<i>zup</i>	

Table 7 – Tonal minimal sets

Tone				
1	<i>pa</i> ‘teeth’	<i>˧saŋ</i> ‘wing’	<i>kʰo</i> ‘nose’	<i>sa</i> ‘urine’
2		<i>˨saŋ</i> ‘sky’	<i>kʰo</i> ‘head’	
3	<i>pa</i> ‘mad’	<i>˨saŋ</i> ‘appreciate’	<i>kʰo</i> ‘path’	<i>sa</i> ‘to clean’
4	<i>paʔ</i> ‘spear’		<i>kʰoʔ</i> ‘top’	<i>saʔ</i> ‘tiger’

4. Verb predicate and agreement

This section discusses the Haʔwa verb predicate showing the different constituents that occur within the predicate. An approximate position class diagram of the Nocte verb predicate is also presented below to give the readers an overview.

Table 8 – Position classes in the Nocte predicate

1	2	3	4	5	6
NEG	Root	ASP	NEG	INVERSE	AG
PROHIBITIVE		NOM	T		QP
ASP		NEG	M		
		CISLOCATIVE			

Section 4.1 gives an introduction to the Nocte verb stem alternation. Section 4.2 discusses the auxiliary verb. 4.3 is focused on the tense-aspect markers and modality is discussed in section 4.4. In section 4.5, hierarchical agreement system is discussed in detail and section 4.6 and 4.7 talk about inverse and cislocative marking in the Haʔwa variety of Nocte.

4.1. Verb stem alternation

Stephen Morey, in his unpublished grammar sketch of Mueshaung Tangsa, gives an analysis of verb stem alternation in that variety. According to Morey, verb stem alternation is a common feature of the Tangsa varieties. These language varieties normally have two verb forms: one is the verbal and the other is the nominalized form that is often but not always preceded by a vowel /ə/. He recorded a total of 64 different verbs with stem alternation in Mueshaung and found that around ninety percent of the nominalized roots differ from the verbal root in tone category. In my data I have come across a few verb stem alternations. One hypothesis is that tone plays a crucial role in verb stem alternation. In addition, stem alternation also involves segmental changes. Tone has not been marked yet on the text I have collected. Therefore, it is difficult to present a detailed discussion on this and the topic requires further study.

Although it is difficult to strictly categorize the verb stems on the basis of their occurrence in a particular construction here, I exemplified a few instances and try to give an idea about the possible environments for alternation.

Table 9 – Nocte verb stem alternations

Gloss	Intransitive realis	Imperative	Interrogative/Negative
‘to go’	<i>ka</i>	<i>ka</i>	<i>ket/keʔ</i>
‘to kill’	<i>.iit</i>	<i>.iit</i>	<i>.iʔ</i>
‘to stand’	<i>tfəp</i>	<i>tfəp</i>	<i>səp</i>
‘to buy’	<i>.iiaak</i>	<i>.i</i>	<i>.iiaak</i>
‘to laugh’	<i>ŋe</i>	<i>əŋit</i>	<i>ŋe</i>

Table 9 above, shows some Nocte verbs and their alternate forms in different constructions. All of these verbs except *ka* ‘to go’ have two forms but do not alternate in identical constructions in a regular pattern. It is possible to argue that the verb *ka* ‘to go’ has three stems, stem 1 *ka* occurs in intransitive realis and imperative constructions and, stem 2 and 3 *ket and keʔ* occurs in interrogative and negative constructions. But then, this is the only

verb in my data that shows three stem alternations. On the other hand for the verb *əsəp* ‘to stand’ stem 1 *səp* occurs in intransitive realis, interrogative and negative constructions and stem 2 *əsəp* occurs only in imperative constructions. This is summarized in Table 10:

Table 10 – Stem specific constructions

Gloss	Stem 1	Stem 2
‘to go’	Intransitive realis & Imperative	Interrogative & Negative
‘to kill’	Intransitive realis & Imperative	Interrogative & Negative
‘to stand’	Intransitive realis, Interrogative and negative	Imperative
‘to buy’	Intransitive realis, Interrogative and negative	Imperative
‘to laugh’	Intransitive realis, Interrogative and negative	Imperative

From the examples I have presented below it is clear that verb stem alternation happens depending on the construction it occurs in like imperative, negative or interrogative. I have used the verb *ka* ‘to go’ to demonstrate this.

Future intransitive (stem 1)

- (1) *ŋa pit-naŋ ka aŋ*
 1SG field-LOC go:I 1SG:FUT
 ‘I will go to the field’

Question (stem 2)

- (2) *naŋ pit-naŋ kɛ? min ne?*
 2SG go-LOC go:II FUT QP
 ‘will you go to the field?’

Negative (stem 3)

- (3) *ŋa ket ho*
 1SG go:II not.have
 ‘I don’t have to go’

As we can see from the above example sentences stem 1 occurs in future, stem 2 occurs in interrogative constructions and stem 3 occurs in negative constructions.

4.2. Structure of the verbal system

According to DeLancey (2014), in a number of languages of the northern Myanmar and Northeast India agreement markers do not attach to the verb stem and rather occur as phonologically independent words which he refers to as the agreement words. These agreement markers can also get attached to certain morphemes carrying information about tense, aspect and other verbal categories. It is in this sense as suggested by DeLancey these agreement words can be referred to as verbal operators. In Nocte verbal operators give additional information to the main verb like tense, aspect, mood, negation etc. This section of the paper gives an overview of how the verbal operators work in Nocte. Morey (2011)

recorded some of the Tangsa varieties having agreement as verbal suffixes. Das Gupta (1971) records Nocte as having verbal suffixes to mark tense and aspect. However in the Ha?wa variety of Nocte, I have recorded agreement words together with verbal operators as separate phonological words. Also, Alfons Weidert in his unpublished notes on Nocte records agreement, TAM, negation as separate words.

Verbal operators can be of two types: inflected and uninflected. For example, Nocte tense-aspect markers can be both inflected and uninflected. Nocte future tense can be expressed in two different ways: one with the agreement words that inflect for person and the other with an uninflected verbal operator that does not change according to person. This will be discussed later in section 4.2.2.1 while discussing the future tense.

Verbal operators give additional information to the main verb like tense, aspect, mood, negation etc. Nocte however, has a very complicated post-verbal morphology. Here, DeLancey (2011) discusses the complex SFWs (Sentence Final Word, coined by Dai and Diehl, 2003) in Jingphaw and Nocte. Another important paper by DeLancey (2014) talks about the presence of similar SFW in archaic Kuki-Chin language varieties which he considers as ‘transparently gramaticalized auxiliary verbs’.

The following examples show verbal operators and agreement word with no operator:

- a. V+ uninflected operator
- (4) *zɔn hu-kɔ vɛt tʰu*
 John dog-LOC hit PRS.PROG
 ‘John is hitting the dog’
- b. V+ inflected operator
- (5) *zɔn hu-kɔ vɛt k-a*
 John dog-LOC hit PRS.PROG.3
 ‘John is hitting the dog’
- c. V+ agreement word with no operator
- (6) *ate pit-naŋ ka a*
 3SG field-LOC go FUT:3
 ‘he will go to the field’

Besides tense, aspect and negation, verbal operators also include cislocative and inverse marking morphemes.

4.2.1. Person, tense and aspect markers

Nocte distinguishes tense on the basis of past and non-past whereas aspect shows whether an event has an end point, i.e., perfective, or imperfective that marks it as progressive or inceptive. Non-past in Nocte includes future and habitual. Moreover, the present tense progressive prefixes recorded by Das Gupta (1971) also finds description in this section.

Table 11 – Person, tense and aspect markers

Person	Future	Habitual (Present)	Past	Progressive			Imperative	Prohibitive
				Past	Present	Future		
1SG	-aŋ	ˌaŋ-v-aŋ	-t-ak	-ka-t-ak	-k-aŋ	-ka-aŋ		
1PL	-ε	ˌaŋ-v-ε	-t-iʔ	-ka-t-iʔ	-k-i	-ka-i		
2SG	-ɔ	ˌaŋ-v-ɔ	-t-ɔʔ	-ka-t-ɔʔ	-k-ɔ	-ka-ɔ	-ɔ	-nak-V-ɔ
2PL	-εn	ˌaŋ-v-εn	-t-εt	-ka-t-εt	-k-εn	-ka-εn	-εn	-nak-V-εn
3SG	-a	ˌaŋ-v-a	-t-aʔ	-ka-t-aʔ	-k-ε	-ka-a		

In the future progressive forms the two adjacent vowels are pronounced as two syllables. As we can see from the above table, there are two sets of verbal operators: one, those take stop final agreement markers and two, have nasal ending agreement markers. Throughout this paper we will see in what different constructions these forms can occur.

4.2.2. Future, habitual present and past

4.2.2.1. Future

As introduced in section 4.2.1, in Haʔwa Nocte, one way of expressing future tense is by using portmanteau forms. The second way of expressing future tense is to have uninflected verbal operators *min* or *ɛ min* after the verb. Table 11 below shows the intransitive future tense markers in Nocte.

Table 12 – Future tense markers

Person	Future marker		
	Inceptive/ about to	Time not bound	Certainty
1SG	aŋ	<i>min</i>	<i>ɛ min</i>
1PL	ε		
2SG	ɔ		
2PL	εn		
3SG	a		
3PL			

Example sentences below show some future tense marking in the language. Example sentence (7) shows person-indexed future marking whereas (8) and (9) exemplifies invariant *min* and *ɛ min*.

- (7) ate pit-naŋ ka a
 3SG field-LOC go FUT:3
 ‘he/she will go to the field’

Here in (7), the speaker knows that the subject is about to leave.

- (8) ate ˌaŋsuamtʰin ŋin tʰo ko min
 3SG church about tell give FUT
 ‘he/she will talk about church’

But here in (8), the speaker is not aware of the time when the subject is going to talk about church so *min* is used.

- (9) *ate-mε pit-nay kεʔ .ie min*
 3SG-ERG field-LOC go MOOD FUT
 ‘he will certainly go to the field’

In (9) however, the speaker has an attitude of certainty that he will surely go to the field.

4.2.2.2. Habitual present

In Haʔwa the habitual present is expressed through a morpheme *.nay* that precedes the verb root and agreement particles. The following examples are cited from Das Gupta (1971). I have confirmed the usage of *.nay* ‘habitual present tense marker’ with my informant Mr. Nawang Lowang Medam.

- (10) *ηa skul-nay .ioantay .nay ka aη*
 1SG school-LOC always HAB go 1SG
 ‘I always go to school’

- (11) *ate skul-nay .ioantay .nay ka a*
 3SG school-LOC always HAB go 3
 ‘he always goes to school’

4.2.2.3. Past

Past tense in Nocte is marked by an inflected verbal operator *t-* which takes a stop final agreement marker. The following table shows the simple past tense in Haʔwa without the interference of hierarchy or inverse marking.

Table 13 – Past tense markers

Person	Past
1SG	<i>t-ak</i>
1PL	<i>t-iʔ</i>
2SG	<i>t-ɔʔ</i>
2PL	<i>t-εt</i>
3	<i>t-a</i>

Some examples are as follows.

- (12) *ηa tfam tfaʔ t-ak*
 1SG rice eat PAST-1SG
 ‘I ate rice’

- (13) *hu-pe .i t-aʔ*
 dog-ABS die PAST-3
 ‘the dog died’

In Haʔwa there is one more past tense marking morpheme *wa* which is invariant for any person. However, this morpheme has more sense of completion of an event therefore it is included with the aspect markers.

4.2.3. Aspect (habitual, completive, progressive, inceptive)

In the text I have collected for this paper writing I have come across the following aspect markers in Haʔwa Nocte.

Table 14 – Aspect markers

Aspect		Morpheme
Completive		<i>kε</i>
Progressive	Past	<i>wa</i>
	Present	k-AGREEMENT
		<i>e/i-AG</i>
		<i>t^hu</i>

The following examples will illustrate the aspect markers in Nocte.

- (14) *John hu-kɔ wεt kε*
 John dog-LOC hit COMPL
 'John hit the dog'

According to my informant Mr. Nawang the action of hitting in sentence (14) is completed in the recent past.

- (15) *John hu-kɔ wεt wa*
 John dog-LOC hit PAST-PROG
 'John was hitting the dog'

The action of hitting was ongoing in (15).

- (16) *John hu-kɔ wεt k-a*
 John dog-LOC hit PRS-PROG-3
 'John is hitting the dog'

The action of hitting is ongoing in (16). The aspect is similar in the following example (17):

- (17) *John hu-kɔ wεt t^hu*
 John dog-ERG hit PRS-PROG
 'John is hitting the dog'

4.2.4. Cislocative

Nocte has a cislocative marker *ɹ-* that gets prefixed to the agreement morpheme. It occurs only with motion verbs to show that the motion is towards the deictic center. DeLancey (2011) describes the cislocative marker and gives a comparison between Nocte and Jinghpaw. DeLancey also refers to Weidert's claim that the cislocative marker is possible only in the present tense. Table 14 below presents the Nocte cislocative paradigm from DeLancey (2011)

Table 15 – Cislocative paradigm (Weidert)

Person	Singular	Plural
1	<i>ɹ-ɹŋ</i>	<i>ɹ-iɹʔ</i>
2	<i>ɹ-ɔʔ</i>	<i>ɹ-ɹn</i>
3	<i>ɹa</i> (non-past) <i>ɹaʔ</i> (past)	

Some examples are appended here:

(18) *naŋ ka ɔ*
2SG go IMP.2SG
'you go!'

(19) *naŋ ka ɹ-ɔ*
2SG go CIS-IMP.2SG
'you come'

The only difference in between examples (18) and (19) is the cislocative marker *-ɹ* which occurs with motion verbs in Nocte to show that the movement is toward the deictic centre. In (20), the cislocative marker *-ɹ* occurs with the verb *ka* 'go' to mean 'come' instead. Another example is (20) below:

(20) *naŋ-mɛ ɲa-ɹaŋ kɔlɔm wən ɹ-ɔ aʔ*
2SG-ERG 1SG-for pen take CIS-IMP.2SG QP
'will you bring a pen for me?'

In (20), the usual meaning for the verb *wən* is 'take' but when cislocative *-ɹ* occurs the meaning changes to 'bring' that is the motion of the action is towards the deictic centre or the speaker.

5. Hierarchical agreement and inverse marking

An unusual aspect of Nocte is having a hierarchical agreement system. Similar hierarchical systems are also found in Rawang (LaPolla 2010), Jinghpaw (DeLancey 2011) and some other Northern Naga varieties like Muklom Tangsa (Morey 2011). Hierarchical agreement is a kind of special agreement system, where the transitive verb agrees with a person that is higher in the hierarchy irrespective of its semantic and grammatical role. In Nocte person hierarchy, first person is higher than the second person and both first person and the second person are higher than the third person i.e. 1>2>3.

Apart from hierarchy in the agreement system Nocte also has inverse marking that functions similar to the person indexation system as it also concentrates on marking the hierarchy of a person in a speech event. However, its main function is to show whether the object argument outranks the agent or not. Nocte has an inverse marker *h-* that occurs after the main verb. But since Nocte has already a hierarchical agreement system to depict the person hierarchy, the overt inverse marking merely serves to make the understanding of the agreement unambiguous.

The following table shows the person hierarchy and inverse marking in Haʔwa verbal agreement of transitive verbs:

Table 16 – Hierarchical agreement and inverse marking

		Patient					
Agent		1SG	1PL	2SG	2PL	3SG	3PL
	1SG			ε	$a\eta$	$a\eta$	$a\eta$
	1PL			ε	ε	ε	ε
	2SG	$h-a\eta$	$h-i$			ɔ	ɔ
	2PL	$h-a\eta$	$h-i$			εn	εn
	3SG	$h-a\eta$	$h-i$	$h-\text{ɔ}$	$h-\varepsilon n$	a	a
	3PL	$h-a\eta$	$h-i$	$h-\text{ɔ}$	$h-\varepsilon n$	a	a

Some examples showing hierarchical agreement are as follows:

- (21) *ate-ma* *dihjaʔnja-naŋ* *ko* *a*
 3SG-ERG girl-LOC give FUT-3
 ‘He will give to the girl’

Example (21) does not show any hierarchy in the agreement marking as both the agent and the recipient are 3SG. Whereas, in example (22) below the recipient 1SG is higher in the hierarchy than 3SG therefore, the verb agrees with the 1SG.

- (22) *ate-ma* *ŋa-naŋ* *ko* *h-aŋ*
 3SG-ERG 1SG-LOC give INV-FUT:1SG
 ‘He will give to me’

- (23) *ŋa-mɛ* *ate-naŋ* *ko* *aŋ*
 1SG-ERG 3SG-ABS give FUT-1SG
 ‘I will give to him’

Comparing examples (22) and (23), it is clear that the inverse marker *h* appears when the hierarchy forces the agreement with the P argument instead of the A argument. In (22), the agreement is with the P argument as 1P>3A therefore we also see the inverse marker *h* attached to the agreement word. However example (23) does not show the inverse marking as the agent itself (1A) is higher in the hierarchy and therefore the verb also agrees with it and there is no need to have the inverse marker.

5.1. Agreement in focus

In elicitation, we always get the hierarchy in the agreement system where the verb agrees to the argument that is higher in the hierarchy. However, a transitive sentence can be marked for either of the arguments to mark focus. In other words, it is also possible that apart from tense, person and aspectual information the agreement markings can also carry some pragmatic and contextual information. This inference is required for the understanding of the apparently irregular agreement marking of the subject and objects in Nocte. Two such situations are described below:

Firstly, in Nocte there is a special agreement marking between 1SG subject and 2SG object, where the verb instead of agreeing with the 1SG subject and 2SG object rather is marked with 1PL agreement. One example is:

- (24) *ɲa-mɛ* *naŋ-naŋ* *ko* *ɛ*
 1SG-ERG 2SG-LOC give FUT-1PL
 ‘I will give to you’

Here in (24), the verb instead of being marked for 1SG or 2SG is marked for 1PL which suggests that the action of giving is seen as a process that involves both the giver *ɲa* (1SG) and the receiver *naŋ* (2SG) achieving the meaning of *we ni* (1PL).

Secondly, the agreement marking can also suggest whether the subject or the object is marked. For example, in the following two sentences the pragmatics and context will help the listener to understand the speaker’s selective use of 3rd person agreement marking over the 1st person agreement or vice versa.

- (25) *ate-mɛ* *ɲa-naŋ* *ko* *t-a?*
 3SG-ERG 1SG-LOC give PAST-3SG
 ‘**He** gave it to me’

- (26) *ate-mɛ* *ɲa-naŋ* *ko* *t-h-aŋ*
 3SG-ERG 1SG-LOCABS give PAST-INV-1SG
 ‘He gave it to **me**’

In example (25), the verb agrees with the 3SG and thus emphasizes the 3rd person. Pragmatically this sentence answers to the question ‘Who gave it to me?’. Whereas, in sentence (26) the verb agrees with the 1SG object and thus emphasizes 1SG. And this answers to the question, ‘Whom did he give it to?’

6. Conclusion

Nocte post-verbal morphology is very complex with the presence of phonologically independent agreement words that can also get attached with other verbal categories like tense, aspect, mood and negation. In addition, Nocte has other fascinating features like the hierarchy in the agreement marking system or the presence of the inverse marker etc. Further studies will allow to explain the complexity of the Nocte verbal morphology with greater details.

Abbreviations

1	1 st person	INV	Inverse
2	2 nd person	LOC	Locative
3	3 rd person	M	Mood
ABS	Absolutive	NOM	Nominalizer
AG	Agreement	PL	Plural
ASP	Aspect	PROG	Progressive
CIS	Cislocative	PRS	Present
ERG	Ergative	QP	Question particle
FUT	Future	SG	Singular
HAB	Habitual	T	Tense

References

- Burling, Robbins. (1983). "The Sal languages." *Linguistics of the Tibeto-Burman Area* 7(2): 1–31.
- Das Gupta, K. (1971). *An introduction to the Nocte language*. Shillong, North-East Frontier Agency.
- DeLancey, Scott. (2004). "The blue bird of ergativity." In Francesca Queixalós, Ed. *Ergativity in Amazonia III*. Paris, CELIA-CNRS: 1–37.
- _____. (2011). "Nocte and Jingphaw: Morphological correspondences." In Gwendolyn Hyslop, Stephen Morey and Mark W. Post, Eds. *North East Indian Linguistics Vol. 3*. New Delhi, Cambridge University Press, India: 61–75.
- LaPolla, Randy. (2010). "Hierarchical person marking in the Rawang Language". In Dai Zhaoming, Ed. *Forty years of Sino-Tibetan language studies: Proceedings of ICSTLL-40*. Heilongjiang, Heilongjiang University Press: 107–113.
- Morey, Stephen. (2011). "Tangsa agreement markers." In Gwendolyn Hyslop, Stephen Morey and Mark W. Post, Eds. *North East Indian Linguistics Vol. 3*. New Delhi, Cambridge University Press, India: 76–104.

Assamese studies

On the anatomy of an entry in an Assamese-Assamese-English dictionary¹

Mouchumi Handique
Gauhati University

Abstract

Conceived as a part of the research work on the implementation of an Assamese-Assamese-English (AAE) bilingual dictionary, this paper focuses on some select aspects of existing practices on one hand, and good practices of dictionary-making by incorporating current linguistic and lexicographic principles as well as computational techniques on the other. Most practices of existing AAE dictionaries appear to have serious shortcomings that affect the clarity in the exposition of the form-meaning in relevant contexts. An entry is the nucleus of a dictionary, and this paper deals with the priority issues that can contribute to the clarity of the dictionary and make it a useful tool for the intended users. One such priority is the need to include the pronunciation of entries in an AAE bilingual dictionary because of the complex mapping between speech and writing in Assamese. In the areas of grammar, the distinctive features of Assamese morphology are yet another aspect to be kept in mind in deciding the entries for the AAE dictionary. Similarly, the entries for compound verbs and conjunct verbs have to be planned in terms of their lexical affiliation, constituency and translation equivalence with regard to English.

Citation	Handique, Mouchumi. 2018. <i>On the anatomy of an entry in an Assamese-Assamese-English dictionary</i> . North East Indian Linguistics (NEIL) 8, 186-195. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.
Volume Editors	Linda Konnerth, Stephen Morey, Amos Teo
Copyright	© 2018, the author(s), release under Creative Commons Attribution license
URL	http://hdl.handle.net/1885/155280

1. Introduction

A lexical unit is grammatically and semantically connected to the other units of a sentence where it can appear. To present an entry with as much relevant phonological, morphological, graphological, etymological, syntactic, semantic and pragmatic information is challenging but crucial, especially in a bilingual dictionary. Regarding the compilation of an entire dictionary, the following quote seems rather appropriate:

There can be no doubt that lexicography is a very difficult sphere of linguistic activity. Many lexicographers have given vent to their feelings in this respect. Perhaps the most colourful of these opinions based on a lexicographer's long experience is that of J.J Scaliger (16th -17th cent.) who says in fine Latin verses that the worst criminals should neither be executed nor sentenced to forced labour, but should be condemned to compile dictionaries, because all the tortures are included in this work. (Zgusta, 1971: 15).

A dictionary not only presents linguistic features such as pronunciation, spelling and inflectional paradigms; but also diachronic information such as changes it has undergone over time; and extra-linguistic features such as the entry's socio-cultural significance. In planning a project of this kind, it is necessary to first understand the needs for the users of the dictionary. Users typically do not read the whole dictionary from the beginning to the end, but will look up a certain entry to confirm its meaning; check how it is spelt; how it is used in context; and sometimes how to pronounce it. Meaning is the pivotal and the trickiest of issues a lexicographer has to confront. Conceptualising "meaning" as independent of context, as seen

¹ I am privileged to thank Prof. Jyotiprakash Tamuli, Head of the Department of Linguistics, Gauhati University for suggesting the topic and offering constant help to me.

in some existing regional language dictionaries, including Assamese ones, is one of the major hindrances to being able to meet users' needs.

The first-ever dictionary in Assamese was Miles Bronson's (1867) *A Dictionary in Assamese and English*. It was published by The American Baptist Mission Press at Sibsagar. This work ushered in a new awareness about learning and preserving the language through the production of teaching-learning materials. Bronson wrote in his preface to the book,

After thirty years of familiar acquaintance with the people, I am fully persuaded that it is a mistake to ignore their language. It ought rather to be cultivated. If suitable encouragement were given, the educated Assamese would soon supply vernacular School books and a new impulse in favor of education would manifest itself among the masses (Bronson, 1867: iv).

This work pioneered the art and craft of dictionary making in Assam. Today we have a good number of dictionaries, including some available in electronic form. To name a few: *Hemkosha* by Hemchandra Barua, (1998, 10th edition), (Assamese-Assamese-English), *Jyoti Dwibhaxik Abhidhan* by Dr Pranabjyoti Deka, (1990), (English-Assamese), *Saraighat Abhidhan* by Dr Dinesh Chandra Goswami (2013) (Assamese-Assamese-English), Xobdo online multilingual Dictionary (www.xobdo.org) etc. These dictionaries also keep a record of the changes Assamese has undergone over time, especially in terms of spelling. For example, *loud* 'sunshine' was written as *ৰৱ* in *Hemkosha* (Barua, 1998: 855) and now it is written as *ৰ'ৱ*. Older dictionaries also include words that are very rarely used today. Today, we also have a number of useful electronic dictionaries that are compiled and edited by multiple authors, i.e. dictionaries that anyone can contribute to. Yet, there are issues to be resolved that users of these dictionaries are rarely aware of and that compilers of these dictionaries either do not consider or have no answer for.

The idea for this work was conceived during the period of my association with the *Oxford Compact English-English-Assamese Dictionary* as a translator-cum-editor. The impact of that work on me in developing an insight about the art and craft of dictionary making has been immense. While doing this work, I had to consult with most of the existing Assamese dictionaries (both monolingual and bilingual) in both print and electronic forms. To my surprise, all of them were found to have shortcomings that did not meet users' needs.

This paper therefore discusses some of the major issues being overlooked in present Assamese lexicographic practices; and proposes a best practice model for compiling a good Assamese-Assamese-English Learners Bilingual Dictionary. This discussion will help to summarise the present lexicographic scenario in Assamese and the issues that are not being addressed in current practices.

2. Existing versus proposed practice; a comparative study

2.1. Existing practices

There are a fairly good number of Assamese dictionaries (both monolingual and bilingual) in both print and electronic forms that are quite elaborate, well organized and, hence, useful. Yet, they still somewhat fail to meet users' needs. The discussion to follow is about certain shortcomings of current lexicographic practices.

To begin with, a guide to pronunciation is a must in a bilingual dictionary, as it is designed for users of both the source and the target language(s). Sometimes, even native users would like a guide to standard pronunciation. Most of the available practices of Assamese-Assamese-English Dictionaries, including some recent ones, do not have such an aid.

Secondly, most currently available dictionaries are treated as a mere collection of words, with no measure to show how to use them. In contrast, lexical units are context dependent, i.e.

a word is best understood only when it is placed in a context. Meanings and grammatical functions differ from context to context. This fact is not acknowledged in most of the existing practices. The importance of context is emphasised in this quote:

Meaning is the product of context. In order to represent meaning in a reasonable way we need to recognize the importance of context and avoid unnecessary and unsatisfactory distinctions in meaning. The use of evidence and of formal criteria in analyzing meaning should enable us to achieve the desideratum of reasonableness. (Sinclair, 1987: 87)

In other words, what matters is not the size of a dictionary, but how much can be learned about an entry in terms of its semantic and grammatical functions. This can be attained only by putting them in context, by using the words in a sentence or a phrase. However, examples of these are found in only a few available Assamese dictionaries.

The existing dictionaries in Assamese, (both monolingual and bilingual) are found to classify words into a handful of word classes, namely, *bihesjɔ* ‘noun’, *bihehɔn* ‘adjective’, *hɔnbɔnam* ‘pronoun’, *ɔbjɔj* ‘the indeclinable’, *kɔija* ‘verb’ etc. However, this categorisation is not consistent across dictionaries. In addition, more fine-grained classifications that posit word classes such as postpositions, demonstratives etc. are not found.

Furthermore, most of the conjunct and compound verbs in Assamese with enormous semantic and contextual variation are yet to find a place in many of the available dictionaries, including recent ones. For instance, the two components of a phrasal verb are introduced as separate entries, thereby allowing the users learn their individual sense and use. But when they form a phrase, they stand for a completely different sense. The following example will help to illustrate this.

- (1) পেলা *pela* ‘to throw/cast/cause something to fall’
- (2) দে *de* ‘to give’
- (3) থ *tʰɔ* ‘to keep something (somewhere)’
- (4) পেলাই দে *pela-i de* ‘to throw/cast/cause something to fall’
- (5) পেলাই থ *pela-i tʰɔ* ‘to keep something to do later’

In (4) and (5), *pela* is suffixed with the non-finite *-i* marker and used with *de* and *tʰɔ* respectively. Although *pela-i de* and *pela* carry almost the same sense, there is a difference between them in terms of volitionality, which can be illustrated by examples (6) and (7).

- (6) *hi sola-to-t sah pela-i di-l-e*
3SG.M shirt-CLS-LOC tea to cause to fall-NF give-PST-3
‘He caused the tea fall on his shirt.’ (intentionally).
- (7) *hi sola-to-t sah pela-l-e*
3SG.M shirt-CLS-LOC tea let.fall-PST-3
‘He caused the tea fall on his shirt.’ (unintentionally)

In (6), *pela-i de* is used to express an intentional involvement of the doer in the action whereas in (7), *pela* is used to express accidental involvement. This, however, does not imply that the agent of *pela* is completely non-volitional, but rather that it is less volitional than that of *pela-i de*.

In many currently available dictionaries, only *pela*, *de* and *tʰɔ* are given separate entries, whereas *pela-i de* and *pela-i tʰɔ* are not. However, phrasal verbs like these two are self-contained elements, and we must rely on their use in contexts, not simply on the definitions of

their component parts. Moreover, without a separate entry for (5), a non-native learner or a new learner is likely to do a literal translation of *pela-i tʰɔ* as ‘to keep by throwing, casting something and causing something to fall’, which is not consistent with how this phrasal verb is used in context.

Another issue is the inconsistency in the identification and selection of the root or bare forms of certain lexical units. Verbs and nouns are the most inflected word classes in Assamese. For instance, the 1st person plural form is আমি *ami* ‘we’, and আমাৰ *ama* and আমাক *amak* are its genitive and accusative forms respectively. Most dictionaries have identified the first two forms as independent entries, whereas the third one is not mentioned. If two case-marked forms of the same root can be introduced as independent entries, it is unclear why the third should be left un-introduced.

Some verbs also display special features such as suppletion, default person marking, variation in spelling and pronunciation (when inflected for tense, aspect and person) etc. For instance, যা *za* ‘go’ has the suppletive stem *gɔ-* when it is inflected for the simple past, past imperfective and present imperfective. There are also inflected forms such as তিষ্ঠিবলৈ *tistʰ-ib-ɔbi* ‘to sustain’, জন্মিছে *zɔnm-is-e* ‘being/is born’ etc. that never occur as uninflected verb roots, i.e. তিষ্ঠ *tistʰ* and জন্ম *zɔnm*. But some Assamese dictionaries include these uninflected roots as lexical entries (e.g. Pathak 2017: 681), while others introduce neither the inflected forms nor the bare root forms as entries. Another example is the word *hɔltʰek*, which might be translated as ‘trouble’, but does not have an independent meaning outside of the construction শলঠেকত পৰ *hɔltʰekɔt pɔ* ‘to be in trouble’ (*hɔltʰek-* ‘trouble’, *ɔt-* ‘LOC’, *pɔ-* ‘to fall’). Nevertheless, some dictionaries treat *hɔltʰek* as its own entry with an independent meaning.

Discourse particles are unavoidable elements of Assamese vocabulary. However, in almost all the currently available Assamese dictionaries, they are either left out or are introduced in awkwardly. These particles include দেখোন *dekʰon*, জানোচা *zanosa*, কিজানি *kizani*, হবলা *hɔbɔla* etc. – no attempt is made to gloss these here, since they cannot be explained without a lot more contextual information. There are also words such as আৰু *au* ‘and’, the coordinating conjunction and আকৌ *akou* ‘again’, which can be used as discourse particles too. However, in most current dictionaries, no mention is made of these separate senses.

2.2. The proposed practice

In this section, we propose a best practice model to deal with the issues discussed in the previous section. Although these measures are designed to apply mostly in an Assamese-Assamese-English Bilingual Dictionary, they can also be applied to monolingual dictionaries.

2.2.1. Pronunciation, example sentences and phrases etc:

As mentioned, pronunciation guides and examples sentences are not yet adopted in many of the currently available dictionaries, including some recently launched ones. Some dictionaries use Roman script for pronunciation and some comparatively newer ones use the International Phonetic Alphabet (IPA). The following examples will illustrate when the use of IPA is particularly useful in a bilingual dictionary.

For discussion, we can classify Assamese words into three categories:

- (a) words with the same spelling, but different pronunciation;
- (b) words with the same pronunciation, but different spelling; and

- (c) words with the same spelling, same pronunciation but different meanings and word classes.

Examples from Category (a) are shown below:

- (8) পাৰ [pa:ɔ] *noun* ‘pigeon’
(9) পাৰ [pa:] *verb* ‘can; be able to’

The words in (8) and (9) are homographic, i.e. they have the same written form, but are not homophonous, i.e. they do not have the same pronunciation. This is because, the Assamese script cannot represent a syllable-final [ɔ] vowel sound (a rounded mid-low back vowel). Literate native speakers will be able to rely on the context to identify the words and therefore know whether or not to produce the vowel. But for non-native users, this is where a pronunciation guide would be most helpful.

Examples from Category (b) are given here:

- (10) শাহ [ħah] *noun* ‘the kernel of a fruit’
(11) সাহ [ħah] *noun* ‘courage, bravery’

The words in examples (10) and (11) are homophonous, but not homographic. This is because the Assamese writing system has more than one letter to represent a single sound. For example, the pharyngeal fricative [ħ] is represented by three letters in Assamese, স (*dantya xa*), ষ (*murdhanya xa*), শ (*talabya xa*). Similarly, there are two sets of letters to represent the sounds [t^h d^h n].

Finally, words that fall under the third type are identical in spelling and pronunciation, but are semantically and/or syntactically different. This is where pronunciation guides and example sentences are most useful, as shown here:

- (12) কলা [kɔla] *বিশেষণ adjective* কাণেৰে কম শুনা বা একেবাৰে নুশুনা লোক বা জন্তু
[kane.ɛ kɔm ħuna ba ekeba.ɛ nuħuna lok ba zɔtu] a person or an animal with no or
poor hearing ability; deaf
example: মানুহজন কলা [manuhzɔn kɔla] (the person is deaf)
- (13) কলা [kɔla] *বিশেষ্য noun* জীৱনটো আৰু বেছি ধুনীয়া কৰিবলৈ মানুহে কৰা কিছুমান
কাম যেনে, নাচ-গান, ছবি অঁকা, নাটক আদি [zibɔtɔ a.ɪu besi ħunija kɔ.ɪbɔloi manuhe
kɔ.ɪa kisuman kam zene nas gan sɔbi ĩka adi] activities such as dance and music,
drawing, acting etc which people do to make life more enjoyable; art
example: কলা-সংস্কৃতিত অসম চহকী [kɔla ħɔŋskritit ɔħɔm sɔħɔki] (Assam is rich in
art and culture).

2.2.2. Special aids (notes, reference tables etc.) for special features

In this section, examples of entries that require an additional illustrative note about some of their special features will be presented.

The words উৎসাহ *utsah* ‘enthusiasm’ and চিকিৎসা *sikitsa* ‘medical treatment’ contain the grapheme sequence ৎস. This sequence of graphemes is read as the sound sequence [ts] or [tħ]. However, although উৎসাহ *utsah* ‘enthusiasm’ can be pronounced as either [utsah] or [utħah], চিকিৎসা *sikitsa* ‘treatment’ is pronounced only as [sikitsa], not as *[sikitħa]. A

special note about pronunciation for entries that contain this grapheme sequence would therefore be needed. Such notes can be placed inside a separate box with a different colour and/or font size to show that it is not an entry by itself, but an illustrative note to the entry it is with.

The pedagogic principle of a dictionary asserts that there must be some guidance to the use of the words. However, introducing both lexical and syntactic information together in a dictionary is a big challenge and a lexicographer's task is to make informed decisions about how much of each kind of information to include. Teaching a bit of grammar can be included in a dictionary, using special aids such as referential notes and/or tables. For example, in section 2.1, we looked at the verb forms *তিষ্ঠিবলৈ* *tist^hibɔloi* (*tist^h-ibɔloi*) 'to sustain' and *জন্মিছে* *zɔnmise* (*zɔnm-is-e*) 'being/is born'. *-ibɔloi* is the non-finite marker in *tist^hibɔloi* and in *zɔnmise*, *-is* is the aspect marker and *-e* is the third person marker. These two forms do not have bare forms such as *তিষ্ঠ* *tist^h* and *জন্ম* *zɔnm*. However, these forms are used only in conjunct and compound verb constructions, i.e. with another verb such as *হও* 'be' or *তাক* 'stay, continue'. These features need to be introduced in special notes.

In addition, a reference table is another aid that users can follow to get a broader idea about the different inflectional forms of a word. Further discussion about the use of reference tables with examples will follow in Section 3.

Certain idiomatic expressions can be illustrated by a special note that explains their meanings and shows how they are used. These include metaphorical expressions: *শহকণীয়া* *ḥḥa kɔnija* (*ḥḥa* 'rabbit', *kɔnija* 'eared') which means 'a person with very sharp hearing', not 'a person with ears shaped like those of a rabbit'; *দীঘল ঠেঙীয়া* *digɔl t^henija* (*digɔl* 'long', *t^henija* 'legged') which means 'rain', not 'someone with long legs'; *গেলা গৰম* *gela gɔɔm* 'suffocating heat' (literally, 'rotten heat'); and idioms such as *নাদৰ ভেকুলী* *nadɔ ɓekuli* (literally, 'a frog in a well') which refers to 'a person who thinks they know everything, whereas in reality they do not'.

Similarly, culturally specific concepts like *ছুৱা লাগ* *suwa lag* 'to be impure'; *তোলনী বিয়া* *tolɔni bija* 'a ceremony organized when a girl attains puberty' need some explanation to be able to be understood by non-native users. Explanations for cultural concepts such as rituals and customs, festivals, costumes, plants and herbs, beliefs etc. can be supported by special notes.

2.2.3. Synonyms, antonyms and pictures

The use of synonyms, antonyms and pictures can enhance the understanding of a concept and help in distinguishing one concept from another. For example, *টান* *tan* as an adjective can be translated into English as 'hard', 'tight' or 'difficult'. The use of synonyms, antonyms and collocations can therefore help users understand these various senses of the lexical entry.

(14) *টান* [*tan*] বিশেষণ. *adjective*

Sense 1: *কঠিন*; যিটো কটা বা ভঙা সহজ নহয়। [*kɔt^hin; zito kɔta ba ɓɔŋa ḥḥɔz nɔɓɔ]*
hard; difficult to cut or break.

example *টান মাটি* [*tan mati*] hard soil

antonym *কোমল* [*kɔmɔl*] soft.

Sense 2: (সাজ-পাৰ আদিৰ ক্ষেত্ৰত ব্যৱহাৰ হয় [*ḥaz paɔ adiɔ k^het.ɔt ɓjɔɓɔhaɔ ḥɔj*])

জোখতকৈ সৰু হোৱা বাবে গাত ভালকৈ নোসোমোৱা [*zɔk^hɔtkoi ḥɔɔu ḥɔwa babe gat ɓalkoi nɔḥɔmɔwa*] (used about clothes etc.) not fitting well in one's body for being smaller than required; tight

example টান চোলা [tan sola] a tight shirt

antonym ঢিলা [dila] loose.

Sense 3: (কিবা কাম, পঢ়া বিষয়, পৰীক্ষা আদিৰ ক্ষেত্ৰত ব্যৱহাৰ হয় [kiba kam, pɔ.ɪha bihɔj adia k^hetrɔt bɔbɔha.ɪ hɔj]) জটিল; যিটো কৰা, বুজা, সমাধান কৰা বা যিটোত উত্তীৰ্ণ হোৱা সহজ নহয় [zɔtil; zitɔ kɔ.ɪa, buja, hɔmaɔdan kɔ.ɪa ba zitɔt utti.ɪnɔ howa hɔhɔz nɔhɔj] (used about works, subject of study, examination etc.) difficult to do, understand or pass

example টান অংক [tan ɔŋkɔ] a difficult sum

antonym উজু [uzu]; সহজ [hɔhɔz] easy.

Some currently available dictionaries also provide pictorial representation of certain items, especially culturally specific items such as ঢোল *dol* ‘a drum like instrument’, চেপা *sepa* ‘an item used to catch fish’ and মাকো *mako* ‘an item used in weaving’. However, a simple pictorial depiction is not enough to explain culturally specific objects and concepts. A bilingual dictionary not only deals with two languages, but also with different cultures, and the description of such entries must also give some idea about the cultural relevance of these items.

3. The inflected forms

When compiling a dictionary, questions naturally arise when deciding what constitutes a “word”. An orthographic word is generally a meaningful grouping of letters with empty spaces on either side. A dictionary often deals with bare forms, not the paradigmatic forms of a word, which is why a dictionary is not a list of words, but rather a list of root words. Nevertheless, popular dictionaries such as the *Oxford English Dictionary* (Simpson and Weiner 2012) provide some inflected forms of roots, especially verb roots, using brackets. This practice can be applied in Assamese dictionaries too. However, nouns in Assamese inflect for case and verbal paradigms in Assamese are also much larger and more complicated than that of those in English. In particular, within second person agreement, there are three honorific divisions: non-honorific, mid-honorific and high-honorific.

The inflected forms of the most productive word classes can be shown in at least two ways: (i) by showing all the paradigmatic forms of a word along with the main lexical entry; or (ii) by using tables and figures which users can look up in order to learn about word formation strategies. Each of these ways has its advantages and disadvantages.

The first strategy can introduce the various forms of the most productive lexical items as part of the description of the entry. Example (15) presents the noun মা *ma* ‘mother’, followed by with all its case forms; and example (16) presents the entry of the verb যা *za* ‘go’, followed by its various forms inflected for tense, aspect and person.

(15) মা [ma] বিশেষ্য *noun* ‘mother’
(মাৰ [ma.ɔ], মাৰা [ma.ɪa], মাক [mak], মাক [mak])

(16) যা [za] ক্ৰিয়া *verb* ‘go’
(যাওঁ [zaɔ̃] যাৰ [zawɔ] যোৱা [zɔwa] যায় [zaj] যায় [zaj]
গৈছোঁ [goisɔ̃], গৈছ [goisɔ], গৈছা [goisa], গৈছে [goise], গৈছে [goise],
গ'লোঁ [golɔ̃], গ'লি [goli], গ'লা [gola], গ'ল [gol], গ'ল [gol],
গৈছিলোঁ [goisilɔ̃], গৈছিলি [goisili], গৈছিলি [goisila], গৈছিল [goisil], গৈছিল [goisil])

As a second strategy, reference tables can be used to introduce the various forms of a word. These tables can be kept in a separate section towards the end of the dictionary. They can be looked up for learning various word formation strategies. Instructions to refer to these tables can also be placed under a dictionary entry, where necessary. Two such tables are presented here. As examples, Table 1 presents the case markers for nouns and pronouns, with their morphophonemically variant final sounds of the root and the non-finite forms of verbs respectively. Table 2 presents the non-finite forms of verbs.

Table 1 - Case markers for nouns

Final letter/sound of the Root	Ergative	Accusative	Locative	Dative	Genitive	Instrumental/Commutative	Comparative
-া (-a ending)	-ই -i/	-ক /-k/	-ত /-t/	-লৈ /-loi/	-ৰ /-r/	-ৰে /-re/	-তকৈ /-tkoi/
-ই, -ঈ, -এ, -ি, -ী, -ে, -ৈ (-i, -e, -oi ending)/	-য়ে /-je/	-ক /-k/	-ত /-t/	-লৈ /-loi/	-ৰ /-r/	-ৰে /-re/	-তকৈ /-tkoi/
-উ, -ঊ, -ঋ, -ৌ, -ঔ, -ৌ (-u, -o, -oo ending)	-ৱে /-we/	-ক /-k/	-ত /-t/	-লৈ /-loi/	-ৰ /-r/	-ৰে /-re/	-তকৈ /-tkoi/
-ও (-o ending)	-ৱে /-we/	-ক /-ok/, -ৱক /-wɔk/	-ত /-ot/, -ৱত /-wɔt/	-লৈ /-loi/, -ৱলৈ /-wɔloi/	-ৰ /-r/, -ৱৰ /-wɔr/	-ৰে /-re/, -ৱেৰে /-we:re/	-তকৈ /-tkoi/, -ৱতকৈ /-wɔtkoi/
Consonant ending & -য় (-j ending/)	-ে /-e/	-ক /-ok/	-ত /-ot/	-লৈ /-ɔloi/	-ৰ /-ɔr/	-েৰে /-e:re/	-তকৈ /-ɔtkoi/

Table 2 - Non-finite forms of verbs

Word final sound	Nominal	Time and action referent	Purposive	Resultant	Conditional	Connective
-o	-ৱা /-wa/	-ওঁতে /-ɔte/	-ব /-bo/, -বলৈ /-bo:loi/	-ৱাত /-wat/	-লে /-le/	-ই /-i/
-a	-োৱা /-owa/	-ওঁতে /-ɔte/	-ব /-bo/, -বলৈ /-bo:loi/	-োৱাত /-owat/	-লে /-le/	-ই /-i/
-i	-য়া /-ja/	-ওঁতে /-ɔte/	-ব /-bo/, -বলৈ /-bo:loi/	-য়াত /-jat/	-লে /-le/	-∅
-ɔ	-োৱা /-owa/	-ওঁতে /-ɔte/	-'ব /-obɔ/, -'বলৈ /-obɔ:loi/	-োৱাত /-owat/	'লে /-ole/	-ৈ /-i/
Consonant	-া /-a/	-োতে /-ɔte/	-িব /-ibo/, -িবলৈ /-ibo:loi/	-াত /-at/	-িলে /-ile/	-ি /-i/

The first of these two methods shows the user the various forms of a word all at once, so the user does not need to generate the required forms by themselves. On the other hand, this method requires a lot more space and entries can look repetitive. The second of these measures can save a lot of space, since the user can consult a single table for many words, but it does require some additional effort on the user's end, since they are supposed to learn and apply word formation strategies by themselves. This is a task that might be difficult, as well as tedious for them. Information about irregular verbs will also need to be provided. For example, যা 'go' shows suppletion, with some word forms having the stem *za-* and others *go-*. In general, for a print dictionary, the second method would be more appropriate. However, an electronic dictionary would be better equipped to show all the variations under each entry. Even so, some explanation about word formation should be provided for users to understand the various different word forms.

In addition to these two methods, there is a third strategy: introducing the frequently used forms of a word in appropriate contexts to help users learn the syntactic, semantic and pragmatic variations of that word. This method is more appropriate for an electronic dictionary, where space is not an issue, but examples can still be added even in a print dictionary. A profile of an entry with this method applied is presented here:

- (17) নৈ [noi] বিশেষ্য *noun*
 পানীৰ প্ৰাকৃতিক সঁতি যি বৈ গৈ ডাঙৰ নৈ বা সাগৰত মিলি যায়। [*pani.ɪ pɾak.ɪtik h̥iti zi boi goi daŋɔ noi ba h̥agɔ.ɔt mili zaj*] *River; a large natural stream of water that flows in a channel to the sea or another river.*
 নদী [*nɔdi*] (formal) নদ-নদী [*nɔd nɔdi*] (rivers in general)

Examples of different case-marked forms of the lexical entry can be embedded within phrases and presented in a separate box like this:

নৈ বৈ যায় [*noi boi zaj*] 'a river flows'
 নৈয়ে কথা কয় [*noije kɔtʰa kɔj*] 'a river speaks' (poetic)
 নৈৰ সোঁত [*noi.ɪ h̥ɔt*] 'the current of the river'
 নৈলৈ বহুদূৰ [*noiloi bɔhudu.ɪ*] 'the river is far away'
 নৈত চলা নাও [*noit sɔla naɔ*] 'a boat that runs in a river'

Frequent collocations can also be included in a separate box:

ভৰা নৈ [*bɔɾa noi*] 'a river full of water' (usually referred to a river in the summer)

Adding such information can therefore enhance the user's knowledge of the language by providing more context regarding the use of a lexical entry and its various forms.

4. Conclusion

This paper presents some of the best practices for compiling an Assamese-Assamese-English Dictionary. Although the focus here was on developing a bilingual dictionary, it can serve as a good model for designing a monolingual dictionary too. A dictionary is a compilation of words of a language in a certain order, defining the meanings of those words in the same language and/or another or more with or without the other aids such as pronunciation,

examples etc. When compiling a dictionary, a number of questions come up: How to define a word? Should the dictionary be only about words and their meanings with no reference to their syntactic use? Is meaning of a word free of context, i.e. uninterrupted by the elements preceding and following? To be useful for language pedagogy, a dictionary must have some space dedicated to explaining the use of words and their different forms. However, introducing both lexicon and syntax of a language together is a big challenge. A lexicographer's challenge is to understand where to mark the separation of these two and how much of each to present.

Abbreviations

CLS	Classifier
FUT	Future
HH	High-honorific
LOC	Locative
M	Male
MH	Mid-honorific
NF	Non-finite
NH	Non-honorific
PST	Past

References

- Barua, Hemchandra. (1998). *Hemkosha; 10th Edition*. Guwahati, Hemkosha Prakashan.
- Bronson, Miles. (1867). *A Dictionary in Assamese and English*. Sibsagor, American Baptist Mission Press.
- Deka, Pranabjyoti. (1990). *Jyoti Dwibhaxik Abhidhan*. Guwahati, Assam Book Depot.
- Goswami, Dinesh Chandra. (2013). *Saraighat Abhidhan: An Assamese-Assamese-English Dictionary*. Guwahati, Saraighat Prakashan.
- Pathak, D. R. (2017). *Samagrik Axomiya Sabdakosh*. Guwahati, Ashok Book Stall.
- Simpson, John and Edmund Weiner. (2012). *The Oxford English dictionary*, 2nd edition. Oxford, Oxford University Press.
- Sinclair, John M., Ed. (1987). *Looking up: An account of the COBUILD project in lexical computing*. London, Collins ELT.
- Zgusta, Ladislav. (1971). *Manual of lexicography*. Prague, Walter de Gruyter.

Indefinite and definite quantifiers in Assamese: Some observations

Gautam K Borah
Tezpur University

Abstract

The present chapter is an attempt at a description of the grammatical ways Assamese offers for scalar quantification, which refers to a magnitude along a scale. Usually scalar quantification is either individuating, often called number quantification (e.g. many/three oranges), or non-individuating, often called amount quantification (e.g. a lot of water). In Assamese, the mass-count distinction is a-grammatical so that most scalar quantifiers (except for the numerals) quantify both a mass noun and a countable noun (e.g. *alop pāni/kamalā* 'a little water/a few oranges'). However, when the scalar quantifier is a numeral, Assamese uses a classifier with the countable noun, distinguishing it from the mass noun, which cannot be numerically counted (e.g. *tini-tā kamalā* 'three oranges', where *tā* is a classifier which classifies predominantly three-dimensional objects). Thus, the numerically quantified noun phrase in Assamese is a classifier phrase, which, depending on the order of the numeral, the classifier and the noun in the noun phrase, is either indefinite or definite. On the other hand, the non-numerical scalar quantifier in Assamese is either inherently indefinite or inherently definite (e.g. *alop* and *bor* in *alop kamalā* 'a few oranges', and *kamalā-bor* 'the oranges'). In other words, Assamese has a set of inherently indefinite and a set of inherently definite quantifiers for non-numerical scalar quantification. The present chapter mainly focuses on the grammar and meaning of these scalar quantifiers. The data used in the present study comes mainly from the author's native speaker's competence in Assamese.

Citation	Borah, Gautam K. 2018. <i>Indefinite and definite quantifiers in Assamese: Some observations</i> . North East Indian Linguistics (NEIL) 8, 196-207. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.
Volume Editors	Linda Konnerth, Stephen Morey, Amos Teo
Copyright	© 2018, the author(s), release under Creative Commons Attribution license
URL	http://hdl.handle.net/1885/155280

1. Numerical scalar quantification in Assamese

A noun in Assamese may have a generic reference in a generic statement; depending on the semantic predicate involved, it may also have an individuating reference, i.e. it may mean one or more than one indefinite or definite instance of the kind involved. The examples in (1) below exemplify this¹:

(1)

a. *kitāp abināxi*
kitāp abināxi
book immortal
'Books are immortal.'

b. *xi kitāp loise*
xi kitāp lo-is-e
he book borrow-PROG-3
'He is borrowing a book/some books.'

¹ All Assamese examples cited in the current article are transcribed by using letters of the Roman alphabet. A diacritical mark is used above the letter *a* to represent the IPA [ʌ]; otherwise, *a* represents the IPA [ɔ].

- c. *moloi kitāp kinibā?*
 mo.loi kitāp kin-ib-ā
 me-DAT book bring-FUT-2
 ‘Are you going to buy a book/some books for me?’
- d. *kitāp pālā-ne?*
 kitāp pā-l-ā-ne?
 book find-PERF-2-INTER
 ‘Have you found the book/the books?’

In (a) of (1) above, the noun *kitāp* ‘book’ has a generic reference. In (b), (c), and (d) it has, by contrast, an individuating reference; the individuating reference is indefinite in (b) and (c); in (d) it is definite. In respect to number, (b), (c) and (d) above are either singular or plural.

Thus, nouns in Assamese are NPs at the same time (see e.g. Borah 2006, 2012, 2014). In the remainder of the paper we will call such a noun a *bare noun*.²

1.1. The Classifier Phrase

When subjected to numerical quantification the bare noun emerges into a classifier phrase, i.e. a noun phrase that consists basically made of a numeral, a classifier, and a noun.

Such a classifier, often called a numeral classifier (i.e. a classifier that is used with a numeral), classifies the noun referents *basically* in terms of shape and animacy. Now, only objects have a shape and are, therefore, countable; a mass, by contrast, does not have a shape, and is, therefore, uncountable. Thus, numeral classifiers classify only countable nouns, i.e. nouns that refer to objects³ (see Borah 2012: 294). Turning to shape, objectively, all objects of the world are three-dimensional, but in our ordinary perception, objects may appear as one, two, or three-dimensional, i.e. typically as long, broad and round. Thus, long objects are those that extend in space *predominantly* in one dimension (e.g. a pencil); broad ones are those that extend *predominantly* in two dimensions (e.g. a leaf); on the other hand, round ones are those that extend in space *equally* in three dimensions (e.g. a ball).

Thus, numeral classifiers have at least two functions: (a) they separate countable noun referents, i.e. objects, from uncountable noun referents, i.e. masses; (b) at the same time they mention the particular shape of the concerned countable noun referent (i.e. whether the object is long, broad, or round). In other words, numeral classifiers, as classificatory elements, mark our mental focus on a particular property of things, i.e. shape, which, as noted, has to do with countability of things. This is to say that numeral classifiers, by focusing on shape, facilitate counting.⁴ Thus, unless a classifier classifies a noun in Assamese, it resists counting even though the noun is a countable noun. Thus, the Assamese version of ‘three books’ is *tini-khan*

² From a Cognitive Linguistics point of view, a noun (e.g. *dog*) refers to an abstract concept of a kind of thing; an NP (e.g. *a dog*) refers to a particular instance of the kind (see e.g. Taylor 2002). Given this, an NP implies a conceptual move from the abstract world of the concept (i.e. the general), to the concrete world of the particulars. However, in Assamese, this move, unlike in a language like English, is not necessarily marked so that the *bare* noun itself may refer to the particular. This move is marked in English either by the presence of the indefinite article (which picks up at least one instance from a kind) or the plural marker /s/ (which picks up a more than one instance). Thus, we say in English ‘*He’s going to buy a book*’ or ‘*He’s going to buy books*’, not *‘*He’s going to buy book*’. Thus, the noun *book*, unlike Assamese *kitāp* ‘book’, is not an NP in its bare form.

³ Note that objects are either animate or inanimate; humans or animals. Classifiers thus classify objects on the basis of animacy as well.

⁴ A kind of thing (e.g. *dog*) has properties other than shape, but when it comes to counting of the kind, it is the property of shape (i.e. that the kind of thing comes in similar replicable individuated units, e.g. dogs) becomes immediately relevant. Thus, a numeral classifier *individuates* a kind.

kitāp, not *tini kitāp*, where the classifier *khan*, which classifies two dimensional objects like *kitāp* ‘book’, is missing. By way of another example, it is *sāri-zan mānuh* ‘four-CLF-man’, not just *sāri-mānuh* ‘four-man’, where the human classifier *zan*⁵ is missing.

The expansion of the Assamese bare noun into classifier phrases is schematized in Figure 1 below (where NUM stands for Numeral; CLF for classifier) (see also Borah 2012: 297).

Figure 1: Indefinite and Definite Classifier Phrase

Bare NP level:	<i>Noun</i> (refers to a kind or one/more than one (in)definite instance of the kind)
Numerical quantification of the Bare Noun:	<i>Classifier phrases</i>
	Indefinite
	[NUM-CLF] <i>Noun</i> (refers to one/more than one indefinite instance)
	Alternatively, <i>Noun</i> [NUM-CLF] (refers to the same above with a change in focus)
	Definite
	[<i>Noun</i> -CLF] (refers to one definite instance)
	[<i>Noun</i> -NUM* -CLF] (refers to more than one definite instance)
	[*numeral greater than one]

1.1.1. The indefinite classifier phrases

In the following examples, the indefinite classifier phrase is exemplified:

(2)

- a. *sāri-khan kitāp*
 sāri-khan *kitāp*
 sāri-CLF:2dmnsl *book*
 ‘four books’
- b. *kitāp sāri-khan*
 kitāp *sāri-khan*
 book *sāri-CLF:2dmnsl*
 ‘four books’
- c. *tini-tā mekuri*
 tini-tā *mekuri*
 three-CLF:3dmnsl *cat*
 ‘three cats’

⁵ *Zan* classifies objects that are male humans.

- d. *mekuri tini-tā*
 mekuri tini-tā
 cat three-CLF: 3dmnsl
 ‘three cats’

In (a) and (b) of (2) above, the classifier *khan* classifies the kind as referred to by *kitāp* ‘book’⁶, which is then quantified by the numeral *sāri* ‘four’. It is clear from these examples that the indefinite classifier phrase in Assamese comes in two different orders⁷. In terms of reference, both are indefinite, either specific or non-specific, depending on the context involved (see Borah 2012). Examples (c) and (d) above further exemplify the two orders of the indefinite classifier phrase in Assamese.

1.1.2. The definite classifier phrases

In (3)-(4) below, the definite classifier phrases are exemplified. While the structure underlying (3) is [*Noun-CLF*], the structure underlying (4) is the following: [*Noun-NUM⁸-CLF*].

- (3) *kitāp-khan*
 book-CLF:2dmnsl
 ‘the one book’
- (4) *kitāp-sāri-khan*
 book-sāri-CLF:2dmnsl
 ‘the four books’

It is clear from the discussion above that the indefiniteness/definiteness in Assamese is indicated by the position of the classifier relative to the noun and the numeral, and also the way these three elements (i.e., the noun, the numeral, and the classifier) are combined for production, i.e. in the case of indefiniteness, the [noun] and the [numeral-classifier] are produced as two separate units as in the examples in (2); in the case of definiteness, these units are produced as a contiguous sound unit as in (4) (see Borah 2012).

1.1.3. Blurring the precision: the use of ‘-mān’ and ‘-kei’

The bound morpheme *mān* meaning ‘approximately’ is used in an indefinite classifier phrase as in (5) below:

5. *tini-khan-mān sidi ān*
 tini-khan-mān sidi ān
 three-CLF:2dmnsl-mān CD bring.IMP
 ‘Please bring three or four CDs.’

⁶ The classifier *khan* typically classifies broad objects such as papers and leaves; but by way of metonymic extension, *khan* also classifies objects such as tables and books. In the extended classification, the *broad* surface of the book or the table metonymically stands for the whole object (see Borah 2012). The classifier *khan* thus classifies objects which *appear* or are predominately two-dimensional in our ordinary perception.

⁷ While the order underlying (a) above is neutral, the order underlying (b), i.e. *Noun [NUM-CLF]* is usually used to focus on the referent rather than its quantity; or to make the referent more prominent (see Borah 2012).

⁸ numeral bigger than one.

With the bound morpheme, *kei*, precise number can be wholly suppressed in the classifier phrase. *kei* thus replaces the numeral in the indefinite classifier phrase. The morpheme *mān* meaning ‘approximately’ then follows the classifier. This quasi-numerical quantification, as we term it, is illustrated in (6) below.

6. *kei-khan-mān sidi tār*
 kei-khan-mān sidi tār
 kei- CLF:2dmnsl-mān CD his
 ‘Some of the CDs are his.’

In Borah (2016), a study of the discourse functions of classifiers in Assamese, the present authore has shown that blurring of precision with *kei* and *mān* is in fact a politeness strategy in Assamese⁹.

2. Non-numerical scalar quantification in Assamese

We now turn to non-numerical scalar quantification of the bare noun in Assamese. We will try to show that Assamese has a set of inherently indefinite and a set of inherently definite quantifiers for such a quantification.

As noted, absolute quantifiers, i.e., numerals, do not quantify a noun in Assamese unless a classifier classifies it, i.e. focuses on its countability, explicitly stating the shape of its referent.

However, Assamese has a good number of non-numeral scalar quantifiers which quantify the noun without a classifier.

These quantifiers are mostly: *alop*, *akonmān*, *naomān* ‘a little/a few’; *bahut* ‘a lot; ‘*axangxya* ‘many’; *aganan* ‘innumerable; *bhāleman* ‘quite a lot/few’; *gotāserek*, *gotādiek*, *khanserek*, *zandiek* ‘a few’.

When they quantify a noun we have an *indefinite* noun phrase. This is illustrated in the examples below. The usual word order of such a noun phrase is: *Quantifier Noun* (however, it is also possible to reverse the order to affect a change in focus: in one, the object is backgrounded; in the other, the quantity is foregrounded).

- (7) *alop/akonmān/naomān pāni/seni/bix/xahāy/mānuh/kamalā*
 alop/akonmān/naomān pāni/seni/bix/xahāy/mānuh/kamalā
 a.little/a.few water/sugar/pain/help/man/orange
 ‘a little water/sugar/pain/help/a few people/oranges’

⁹ Speakers of Japanese, a numeral classifier language, have a similar tendency to blur the precision as expressed in a classifier phrase, as is reported in Ikegami (1993: 810): “[W]hen the number is to be mentioned for the thing to be talked about, the Japanese speaker characteristically tends to represent it in ambiguous, approximate terms rather than in clear, definite terms. An utterance like *May I have about two ten p. stamps?* would certainly sound odd in English, but the equivalent sentence can be heard quite commonly in Japanese. The psychology behind the preference for this rather oddly ambiguous way of saying things is that by avoiding the mentioning of a definite number, one will sound less demanding, leaving the ultimate decision to the other party, and therefore more polite. The preference is especially characteristic of the speech of elderly ladies. Saying things in manner contributes to quantitatively blurring the extension of the reference.”

- (8) *bahut pāni/seni/bix/xahāy/mānuh/kamalā*
 bahut pāni/seni/bix/xahāy/mānuh/kamalā
 a lot water/sugar/pain/help/man/orange
 ‘a lot of water/sugar/pain/help/people/oranges’
- (9) *bhālemān pāni/mānuh*
 bhālemān pāni/mānuh
 quite.a.lot water/man
 ‘quite of lot of water/people’
- (10) *axangkhyā/aganan paruā/sarā*
 axangkhyā/aganan paruā/sarā
 innumerable ant/bird
 ‘innumerable ants/birds’
- (11) *gotāserek/gotādiek kamalā/kalam*
 gotāserek/gotādiek kamalā/kalam
 a.few orange/pen
 ‘a few oranges/pens’
- (12) *khanserek kitāp/saki*
 khanserek kitāp/saki
 a.few book/chair
 ‘a few books/chairs’
- (13) *zandiek adhyāpak*
 zandiek adhyāpak
 a.few professor
 ‘a few professors’

A closer look at the examples above will reveal that most non-numeral quantifiers in these examples are fuzzy by nature: they quantify both countable and uncountable (i.e. mass) nouns (note once again that we are using the terms *countable* and *uncountable* here only notionally). As noted, in a language like English, non-numerical scalar quantifiers are usually either individuating or non-individuating (i.e. *many* vs. *much*). While the individuating quantifier goes with the countable noun, i.e. the noun that refers to a group of some similar individuated entities (e.g. *many birds*), the non-individuating quantifier goes with the noun that refers to a mass, which is something homogeneous and continuous, i.e. unindividuated (e.g. *much water*). The individuating-non-individuating distinction between scalar quantifiers in English seems to reflect the grammatical nature of the mass-count distinction in the language. By contrast, this distinction in Assamese is *a*-grammatical (see Borah 2006) so that scalar quantifiers in the language (other than the numerals) are mostly non-individuating, i.e. they go with both countable and mass nouns.¹⁰

¹⁰ Note further that a plurality of individuated entities is often perceived or understood as a mass (see e.g. Talmy 1988: 128ff; Lakoff 1987: 440-444; 1988: 144-149). Thus, *This pond is well stocked with fish* rather than *fishes*. Or, (to borrow a line from Crystal 1995: 201), the professional hunter goes shooting *duck*, never *ducks*. This explains why English, even though it systematically distinguishes between individuating and non-individuating scalar quantifiers, has a few fuzzy quantifiers (e.g. *a lot of*, *some*, *lots of*) which quantify both a count and a mass noun (see e.g. Allan 1980: 554ff; Langacker 1987: 77-78).

Some of the non-numeral scalar quantifiers in Assamese are, however, individuating: for example, *axangkhyā* and *aganaṅ* in (10). This is because they are derived from the words, respectively, *xangkhyā* meaning ‘number’, and *aganaṅ* meaning ‘counting’ (the morpheme *a* that is prefixed to them is a negative prefix).

The quantifiers *khanserek*, and *zandiek* in (12)-(13) are also individuating. In each of them we have a classifier and a numeral: in *khanserek* in (12) the classifier *khan* follows *serek* meaning ‘around four’. Likewise, in *zandiek* we have the classifier *zan* followed by *diek* meaning ‘around two’. In *gotaserek* in (13), on the other hand, *got*, i.e. the morpheme that precedes *serek*, means ‘unit’.

The indefiniteness of an indefinite non-classifier phrase in Assamese cannot be overridden. The example in (14) below illustrates this:

- (14) **xei alop mānuh*
 xei alop mānuh
 DEM a.few man

2.1. The definite non-classifier phrase

When a noun is suffixed by a bound morpheme like *bor* (*bor* is one of several such morphemes in Assamese which are discussed below), the resulting phrase is a definite expression referring to more than one definite instance, however, without committing to any precise number, which is exemplified in (15) below:

- (15)
 a. *ālahi-bor*
 guest-bor
 ‘the guests’
 b. *kamalā-bor*
 orange-bor
 ‘the oranges’

2.1.1. *bor*: a collective marker rather than a plural marker

Note that *bor* is not just a plural marker; it is rather a *definite* plural marker. For example, *ālahibor* in (15a) does not mean what ‘guests’ mean; it means what ‘the guests’ means in English.

In other words, the Assamese plural form refers to a definite delimited kind. It is not like the English plural form, which can give a generic reading (e.g. *Dogs bark*); or can also refer to an *indefinite* delimited kind (e.g. *Dogs are barking outside*).

Note also that, unlike the English plural, it cannot be subjected to quantification, either by a numeral or any other scalar quantifier; nor can its inherent definiteness be overridden. Thus, both (16) and (17) are ungrammatical:

- (16) **tini/xat mānuh-bor*
 three/seven man-bor
 ‘three/seven the men’

- (17) **alop mānuh-bor*
 some man-*bor*
 ‘some the men’

All this shows that the construction *Noun-bor* never refers to an abstract class or to the indefinite, but invariably to a well-defined, definite group. We, therefore, designate *bor* as a *definite collective marker* rather than a plural marker.

The other definite collective markers in Assamese are *bilāk*; *lok*, *hat*, *xakal*, *brinda*, and *mandali*.

2.1.2. *bor/bilāk*

These two are used with inanimate objects, or non-human animates; humans of lower social rank, or humans that are non-adult:

- (18)
- | | |
|--------------------------|---|
| Inanimate objects | |
| <i>gāri-bor/bilāk</i> | <i>the cars</i> |
| <i>pāni-bor/bilāk</i> | <i>the measures (e.g. glasses) of water</i> |
| Animals | |
| <i>bāgh-bor/bilāk</i> | <i>the tigers</i> |
| <i>sarāi-bor/bilāk</i> | <i>the birds</i> |
| Humans | |
| <i>dokāni-bor/bilāk</i> | <i>the shopkeepers</i> |
| <i>lorā-bor/bilāk</i> | <i>the boys</i> |

While *bor* and *bilāk* are used with to refer to humans, animals, and inanimate objects, the other collective markers are used to refer to only humans.

Note that the definite collective is always viewed by the speaker relative to a given locator. In (19) below, the group is viewed relative to the locator, *tāk* ‘him’ so that the NP *bandhu-bor* means ‘his friends’. In (20) and (21), however, no locator surfaces, which implies that it is the speaker himself.

Consequently, *bandhu-bor* in (19), and *lorā-bor* in (21) will be interpreted respectively as ‘my/our friends’ and ‘our young soldiers’. Thus, *bor* is an *empathy marker* in these examples.

- (19) *bandhu-bor-e tāk thag-il-e*
 bandhu-bor-e tāk *thag-il-e*
 friend-DCM-NOM him betray-PERF-3
 ‘His friends have betrayed him.’

- (20) *bandhu-bor āji kot?*
 bandhu-bor āji kot?
 friend-DCM today where
 ‘Where are my friends today?’

- (21) *lorā-bor-loi bar dukh hai*
 lorā-bor-loi bar dukh hai
 boy-DCM-DAT very sorrow be
 ‘I feel so sorry for our young soldiers.’

The use of *bor* as an empathy marker is, however, not restricted to humans alone; its use can be extended also to non-humans:

- (22) *dukh-bor hiyāte thalo!*
 dukh-bor hiyāte thalo
 sorrow-DCM heart-LOC locked
 ‘I have locked my sorrows in my heart!’
- (23) *mā, mā, xiāl-bor kiya imān beyā?*
 mā, mā, xiāl-bor kiya imān beyā?
 mother, mother, fox-DCM why so bad
 ‘Mom, Mom, why are the foxes so cunning?’

In (22), the hearer understands that the speaker is talking about sorrows not in the abstract, but all the real sorrows that his heart has experienced, and hence not simply *dukh*, but *dukh-bor*. In (23), a young curious child groups all the foxes he has come to know from the stories he has read for himself, or heard from his mother: why are *they* so cunning? These are not *any* foxes, an abstract class, but a situationally anchored definite group of foxes. Compare (23) with *mā, mā xiāl kiya iman beyā?* ‘Mom, mom, why are foxes so cunning?’

2.1.3. *hat, lok, xakal, brinda, gana*

As noted, while *bor* and *bilāk* are used with to refer to humans, animals, and inanimate objects, the other collective markers are used to refer to only humans.

Note that a qualified *you* is formed with these definite collective markers as in (24)-(26) below. This shows that they are, like *bor*, are empathy markers: the group is defined by the speaker relative to himself.

Of the five, we will discuss here only *hat* and *lok*.

2.1.3.1. *hat*

In the examples below *hat* is used with animals to address a group of people in contempt; with pseudo anger; or with comradely feelings. The NPs thus form a qualified 2nd person, that is, a qualified *you*.

- (24) *garu-hat!*
 cow-DCM
 ‘You, fools!’
- (25) *bāndar-hat!*
 monkey-DCM
 ‘You, - my little naughty children!’
- (26) *bhāi-hat!*
 brother-DCM
 ‘My dear brothers!’

In the following examples *hat* is used to “pluralize” a proper name, i.e. to form a group of humans centering the 3rd person, i.e. him/her. In this case, *hat* functions rather like *et al*, the abbreviation of the Latin *at alii*:

(27) *erik-hat gol*
 erik-hat gol
 erik-DCM leave.PERF
 ‘Erik and the other people with him have left.’

(28) *māhat āhil*
 mā-hat āh-il
 mother-DCM arrive-PERF
 ‘Mother and the other people with her have arrived.’

Thus *hat* also “pluralizes” the 2nd and 3rd person personal pronouns:

2nd person

(29) *toi-hat*
 you (SG)-DCM
 ‘You’ (PL)

3rd person

(30) *xi-hat*
 he-DCM
 ‘They’

2.1.3.2. *lok*

This is used to pluralize personal pronouns. For the 1st person Assamese has two independent words, *moi* ‘I’, and *āmi* ‘we’. In the case of the 2nd and the 3rd persons, the plurals are derived from their singular counterparts by using *lok*. Note, however, that this applies only to the honorifics: the non-honorific 2nd and 3rd personal singulars go with *hat*:

2nd person (Honorific)

(31) *tumi-lok* ‘you (SG)-DCM’ = *tomālok* ‘you’ (PL)

(32) *āpuni-lok* ‘you (SG)-DCM’ = *aponāok* ‘you’ (PL)

3rd person (Honorific)

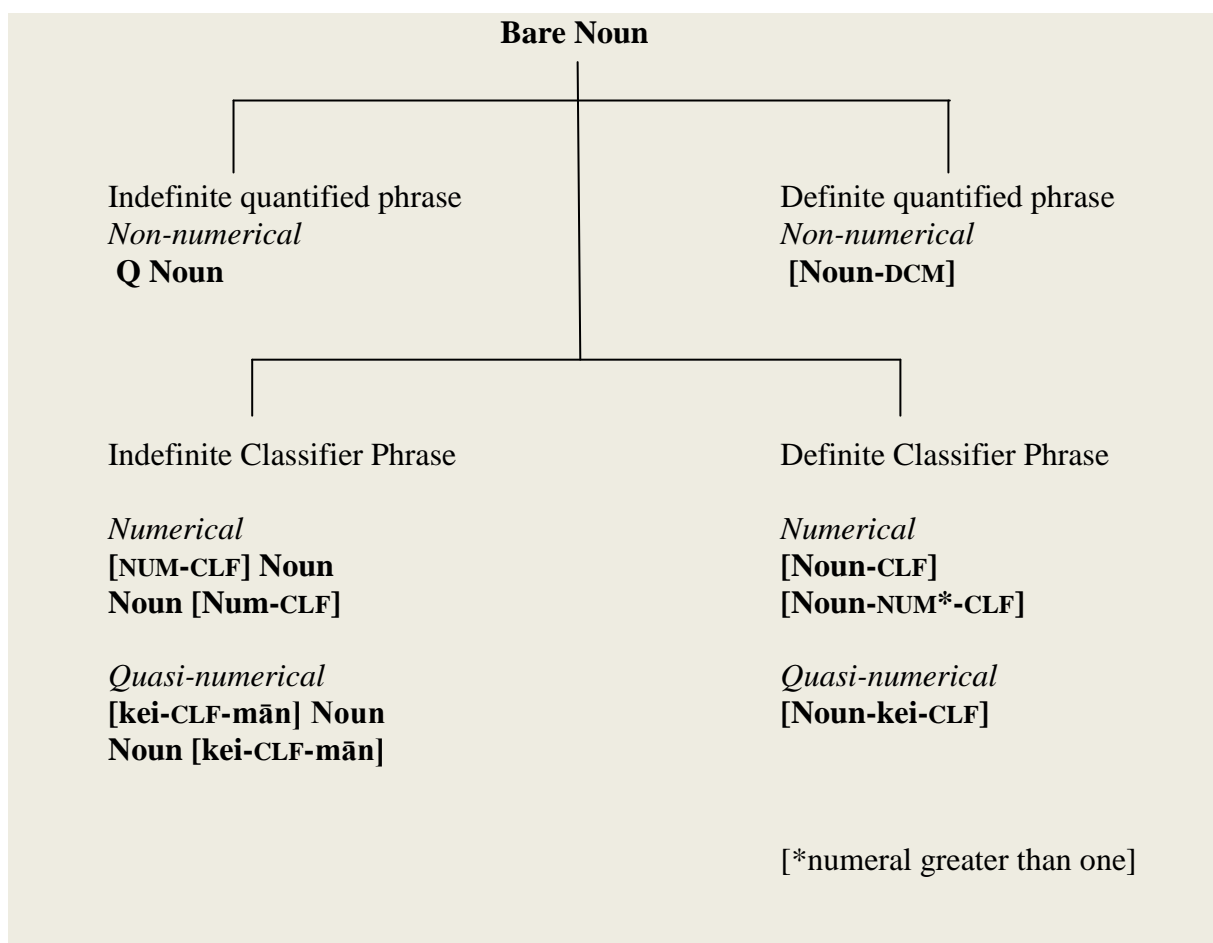
(33) *teo-lok* ‘he-DCM’ = *teolok* ‘they’

3. Conclusion

What we hope to have shown in the current paper is that Assamese offers two basic ways for scalar quantification. One of these is numerical quantification done by using a classifier, (which classifies a countable noun in terms of the shape (or shape related properties) of the referent) along with the numeral. Depending on the order of the numeral, the classifier, and the noun, such a quantified phrase (which we have termed a classifier phrase) is either indefinite or definite. We have also noted that Assamese offers ways to blur the precision as a means of politeness. On the other hand, Assamese has two sets quantifiers, which are either inherently indefinite or inherently definite. While a numeral is an individuating quantifier, these non-numerical scalar quantifiers are mostly non-individuating so that they quantify both the countable and the mass noun. The non-individuating character of these quantifiers can be correlated with the absence of a grammatical mass-count distinction in Assamese. In Figure 2

below, we present a schematic representation of the ways Assamese offers for scalar quantification.

Figure 2: Scalar quantification of the bare noun in Assamese



Abbreviations

- 2 2nd person
- 3 3rd person
- CLF classifier
- DAT dative
- DCM definite collective marker
- dmnsl dimensional
- FUT future tense
- IMP imperative
- INTER interrogative
- LOC locative
- NOM nominative
- NUM numeral
- PERF perfective
- PL plural
- PROG progressive
- SG singular

References

- Allan, Keith. (1980). "Nouns and Countability". *Language* 53: 285–311.
- Borah, Gautam K. (2006). "Macbeth, the Birnam Wood, and the meaning of [s]". *CIEFL Bulletin (New series)*, 16 (2). 52–68.
- _____. (2012). "Classifiers in Assamese: Their Grammar and Meaning Chains." In Gwendolyn Hyslop, Stephen Morey and Mark W. Post, Eds. *North East Indian Linguistics Vol. 4*. New Delhi, Cambridge University Press, India: 292–314.
- _____. (2014). "The fox or Mr Fox? (On particularization of the bare noun)". *Journal of Modern Languages* 24: 62–70.
- _____. (forthcoming). "Are classifiers redundant? Individuation and meaning manipulation by classifiers".
- Crystal, David. (1995). *The Cambridge Encyclopaedia of the English Language*. Cambridge, Cambridge University Press.
- Ikegami, Yoshihiko. (1993). "What does it mean for a language to have no singular-plural distinction? Noun-verb homology and its typological implication." In Richard A. Geiger and Brygida Rudzka-Ostyn, Eds. *Conceptualizations and Mental Processing in Language*. Berlin, Mouton de Gruyter: 800–814.
- Lakoff, George. (1987). *Women, Fire and Dangerous Things: What Categories Reveal About the Mind*. Chicago, Chicago University Press.
- _____. (1988). "Cognitive Semantics". In Umberto Eco, Macro Santambrogio and Patrizia Violi, Eds. *Meaning and Mental Representations*. Bloomington and Indianapolis, Indiana University Press: 119–154.
- Langacker, Ronald. (1987). *Foundations of Cognitive Grammar (Vol. 1)*. Stanford, Stanford University Press.
- Talmy, Leonard. (1988). "The Relation of Grammar to Cognition". In Brygida Rudzka-Ostyn, Ed. *Topics in Cognitive Linguistics*. Amsterdam, John Benjamins: 165–205.

Changing forms and their usage in Assamese among Assamese youths¹

Pushpa Renu Bhattacharyya
Tezpur University

Abstract

This paper is based on a sample survey carried out on the changing patterns of speech of native speakers of Assamese youths due to the influence of English and Hindi. It also seeks to show the impact of multilingualism, its socio-cultural implications, and its effect on the mastery of their mother tongue and on the linguistic cognitive aspect of the individual concerned. The existence of multilingualism tied to various social factors is a common scenario of an urban society of Assam. The changing patterns in Assamese which have been observed in Standard Assamese are found to be the results of the immense support by youths for the thriving of multilingual situations along with the influence of globalization and western culture. The survey was conducted at Tezpur University. In this study 100 (50 males and 50 females) students were participants, all of them were multilingual speakers. A questionnaire was used to collect information. The findings reveal that English is the most preferred language especially for reading books, magazines, writings, e-mailing and texting messages and watching movies. Hindi is the most preferred listening medium for songs. Assamese is used mostly for speaking at home. The regular patterns of Assamese grammar is moving towards a potential change due to the impacts of these factors. Each and every participant feels proud as an Assamese though s/he has not acquired 100% fluency in his/her mother tongue.

Citation	Bhattacharyya, Pushpa Renu. 2018. <i>Changing forms and their usage in Assamese among Assamese youths</i> . North East Indian Linguistics (NEIL) 8, 208-227. Canberra, Australian National University: Asia-Pacific Linguistics Open Access. ISBN: 978-1-922185-41-9.
Volume Editors	Linda Konnerth, Stephen Morey, Amos Teo
Copyright	© 2018, the author(s), release under Creative Commons Attribution license
URL	http://hdl.handle.net/1885/155280

1. Introduction

Language changes over time and this process of change is natural. The processes of analogy and borrowing, code-switching and code-mixing are common situations in multilingual societies. Different languages like English and Hindi used in different media such as written, spoken, listening and watching have impacted on Assamese youths in the age group 18-30. The present study is an exploration of the sociolinguistic dimension of the use of Assamese in different social variables so as to be able to throw light on their changing patterns of usage in Assamese. With this end in view, a sample survey was carried out on 50 males and 50 females.

For the collection of information from the respondents, a pre-tested questionnaire (enclosed in appendix) was distributed and the collected information was analysed statistically and the results presented by bar diagrams below in section 5.

Assamese is an Indo-European language, recognized by the constitution of India as one of the official languages of India. It is spoken in the Brahmaputra valley of Assam. It may be mentioned that Bangla (Bengali) an Indo-European language recognized by the constitution of India as one of the official languages of India which is spoken in the Barak valley of Assam. According to census report of 2011, the population of Assam is 3,11,69,272, the total

¹ My sincere gratitude is due to Dr. Runima Chowdhary for her support and suggestions in writing this paper. I would like to express my special thanks of gratitude to all participants without whom the present study could not have been completed. I am also thankful to the reviewers for their various kind suggestions to be incorporated in the paper.

number of native speakers of Assamese being 1,31,68,484. The languages spoken in Assam belong to different language families, viz., Indo-European (of which Assamese is one example), Sino-Tibetan, Austroasiatic, Tai-Kadai, and Dravidian family.

According to Weinreich (1953) factors like usefulness, opportunity for communication, age and order of acquisition, degree of emotional involvement, social functions, literary and cultural values play an important role in the pattern of communication and choice of language of an individual. The result is noticed as an effect on the level of command in different languages. Urban contexts are highly multilingual, where the existence and use of more than one language in more than one context is very common; as such multilingualism has always been a feature among the elites. In the rural areas, where Assamese is the native language (in the many tribal areas it may not apply) Assamese is spoken in most social interactions. A typical trilingual scenario commonly found in Assam is Assamese as the mother tongue spoken at home, Hindi as the national language and English as the official language and as the medium of instructions in educational institutions.

The implementation of trilingual language formula of the Government of India and the influence of these languages in their childhood remains throughout their entire life, seen in their choice of preferences in different contexts. As a result the usage of multilingualism causes some impact on changing patterns in Assamese which is discussed below in Section 5. Although this study is confined to these three languages some youths interviewed spoke even more than three languages depending on their demographic and economic requirements. It must be mentioned that all the participants of the survey were native speakers of Assamese, none of the participants was a native speaker of a tribal or other language.

It is worth noting here that the Upper Assam dialect, which is recognized as the standard Assamese, and which is taken to represent the best instructional practices and patterns for the most formal use, does not reflect the general pattern of Assamese among youths as is found at this point of time and attempts have been made further to find out how and why different languages (among Assamese, Hindi and English) are chosen by participants of both the genders in similar contexts and also to find out how this effect on their fluency in mother tongue which have been discussed in the following structure in the paper-

- Objectives
- Methodology
- Data Analysis
- Influence of English and Hindi on Assamese Language

2. Objectives

The following are the objectives of the study-

- To find out what languages are preferred in different domains and the percentage of the three languages used by youths.
- To identify various factors contributing to change in patterns including urbanization, national language policies, contact with other languages and culture, the context of upbringing (rural/urban), parents' educational level, participant's attitude towards the language etc.
- To find out the percentage of participants' level of self-declared fluency in their mother tongue.
- To find out what forms of Assamese language are used while translating the words and sentences from English or Hindi and how those forms are influenced by contact of Hindi and English languages.

3. Methodology of the survey

The study was conducted by adopting sample survey methodology. The sampling design adopted was a purposive sampling. The sampling units were the students of Tezpur University from undergraduate to Ph.D. level. Although the students were the sampling units no complete sampling frame was formed. The students were met personally in hostels, canteens, classrooms, streets. They were asked whether they were native speaker of Assamese or not. Whenever someone confirmed that s/he was a native speaker of Assamese s/he was selected for the sample. The present study was conducted in Tezpur University, in Assam, India. The students (50 male and 50 female) selected for the study were native speakers of Assamese in the age grouping of 18-30, representing the youths from various parts of Assam. The participants were classified into different age-groups viz. 18-22, 23-27 and 28-32. In the age-group 18-22 there were 49 participants, in the age group 23-27 there were 37 participants and in the age group 28-32 there were 11 participants. As the survey was confined to student participants only, the range of education level of the participants was from undergraduate to Ph.D. level. Participants of the survey belonged to both rural and urban areas. The necessary information was collected from them on the spot at a single setting itself. As the study is descriptive in approach, all the successive stages are oriented accordingly.

As mentioned above, the study was confined only to those speakers whose mother tongue is Assamese. Towards this end, keeping in view the objectives of the study, a structured questionnaire to get uniformity, comparability and generalization had been developed which was tested among a small sample of ten students where five were males and five were females to see whether the questions were appropriate or some modifications were necessary. Whenever it was found some modifications were necessary and wherever it was necessary it was accordingly modified and the questionnaire was finally structured, e.g., the question number (iv), under the Part VI in the questionnaire covers questions in Hindi (written in Roman script). This is because the second person pronouns in Hindi exhibit a three-term distinction for social status of referents which is not found in English, but in Assamese. This questionnaire was given to the participants chosen for sampling for the collection of necessary information. The questionnaire of the survey is designed to explore the usage of changed of morpho-syntactic forms in standard Assamese language among youths. After a brief introduction about the purpose of the study and instruction to the questionnaire, each selected student was given the questionnaire to fill up instantly within 30 minutes. The questionnaire was designed with both open and closed questions to elicit information relating to the language use from the participants in the survey. The questionnaire consisted of six (6) parts I, II, III, IV, V and VI. The questionnaire was aimed at providing an opportunity to the participants to express their views in their own words, as their choices were not fixed. The questionnaire was designed to know the preference in choice of languages, argumentation, interactional and conversational behaviour etc. between the responses and attitudes of the participants.

Although the questionnaire was in English the medium of the study, the interaction between the participants and the researcher was in Assamese to create a normal, relaxed, more friendly and not strictly formal situation. The interactions covered more detail than the responses received from the direct question of the questionnaire. The questionnaire was designed to check the correlation in choice of patterns between the responses and attitudes of the participants. To explore whether the responses shown in the questionnaire tally with conversations and practical situations, some follow-up interviews were conducted to confirm the results immediately after the questionnaire was completed and before a considerable span of time had taken place. Thus, data were collected using this interview method by the researcher personally. The collected data were tabulated and statistically analyzed for

achieving the objectives. Many a times a lot of participants helped the researcher to approach other participants. During interviews, a number of participants shared their life-experiences and feelings, apart from the issues covered in the questionnaire. The questionnaire covered following six parts-

Part-I

1. Participants made choices about language usage according to the different domains and an attempt was made to find out the percentage of different languages used by the selected youths in various domains.
2. Opinion polls were conducted to elicit responses from participants about the percentage of youths who preferred Assamese as the most preferred language in all domains.

Part-II

It focused on the percentage of participants' level of self-declared fluency in Assamese.

Part-III

This section attempted to find out the various factors that have contributed to the fluency or lack of fluency among youths in Assamese.

Part-IV

This section sought to verify whether the youth aims to show off when speaking to his/her friends in Assamese.

Part-V

This section focused on the participants' own judgement in introducing himself/herself as an Assamese.

Part-VI

By asking students to translate some sentences into Assamese, this part sought to focus on the influence of English and Hindi on Assamese Language in usage.

4. Data Analysis

4.1. Part I

Part 1 was designed to elicit answers to questions concerning actual use of languages viz., Assamese, Hindi and English in different situations. Several functions were listed including questions in language used, e.g., what languages do they prefer in the different domains while – (a) (i) Reading novels (ii) Stories (iii) Poems (iv) Books (v) Magazines (vi) Journals and (vii) Newspapers (b) (i) Sending text messages (ii) E-mails (c) (i) Listening music and radio and (ii) Watching movies and the programs telecast by television (d) (i) While conversing with parents (ii) Parents' friends (iii) Participant's friends and (iv) in social gatherings.

4.1.1. Findings

- (a) The results of Part1 show that the categories for which the highest proportion of participants prefer Assamese, is that of reading novels and poems (25%). Hindi was chosen by 24% (where the score is highest) of the participants to read poems and

English was listed as the most preferred language in reading novels, short stories, poems, books, magazines, journals and newspapers.

- (b) The texting language (Assamese words written in Roman script with some English numerals) was chosen as the most preferred language by 98% (where the score is highest) of the participant in sending text messages. 98% of the participants use English in sending e-mails. Most of the participants admit that they do not know how to type in Assamese script. A few participants admitted that although they know how to type in Assamese script, they feel more comfortable typing in Roman script.

The results of (a) and (b) or percentage of youths preferring different languages in different domains has been shown by bar diagram in Figure 1.

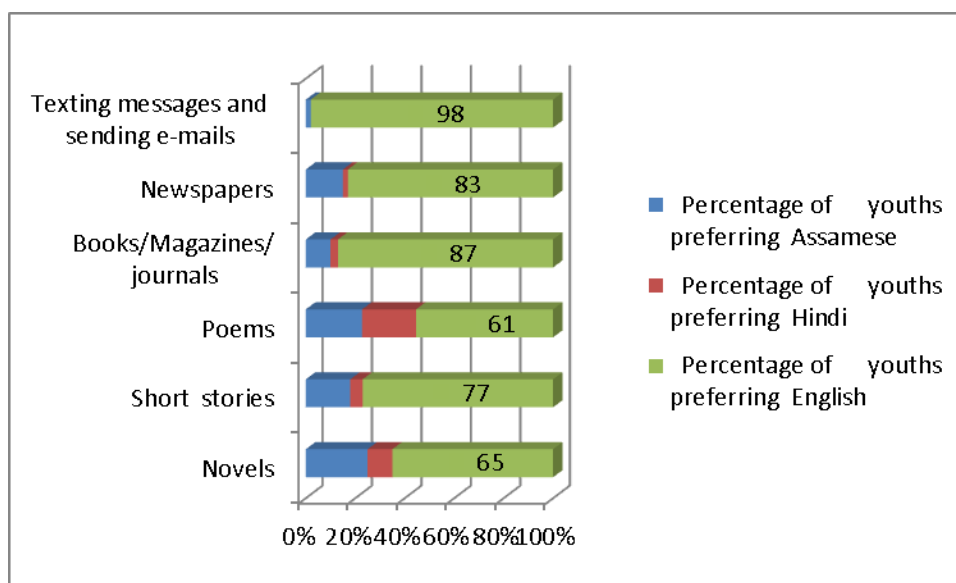


Figure 1 – Percentage of youths preferring different languages in different domains

- (c) Hindi was chosen as the most preferred language by 70% (where the score is highest) of the participants in listening music and radio. The participants opined that they do not have sufficient choices when listening to Assamese songs, because there are very few good singers in Assamese. But in Hindi they have different choices with more good singers and better performances. On the other hand 70% of youths preferred English for watching movies and the programs on television. It must be mentioned that programmes are telecast by television in all the three languages as such that the spectators have the freedom to select any language. Only 10% of the participants prefer Assamese for listening to songs and watching movies or programs telecast by T.V. channels. The results of (c) or percentage of youths preferring different languages in different domains has been shown by bar diagram in Figure2.

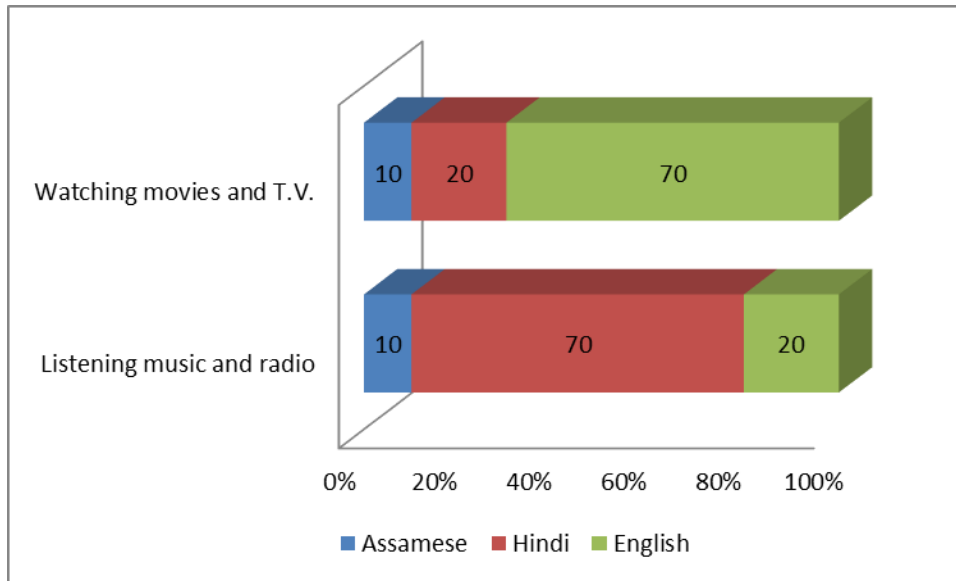


Figure 2 – Languages preferred listening to music and radio and watching movies

There were only two youths who preferred Assamese as the most preferred language in all domains.

- (d) Assamese was chosen as the most preferred language, while conversing with parents, parents' friends, participant's friends and in social gatherings. In the household domains Assamese is mostly preferred. But the respondents can converse in Hindi and English also, depending upon the context of discourse even if they belong to same cultural background. Code mixing and code switching are also common phenomena among those youths which is not being discussed in the present study.

4.2. Part II

The findings of part II which is based on the present survey reveal the percentage of participants' level of self-declared fluency in their mother tongue. It is seen that a number of participants (80%) reported that they had 80% fluency, while the rest of 18% reported that they believed they did not. The remaining 2% did not respond. When it was asked if they could read and write equally well in their mother tongue, 80% reported that they could read well, while 17% reported that they found difficulties in reading. Most of those participants who had their education in English medium schools admitted that in writing Assamese they made lots of spelling mistakes. That means that, even though they have their own individual mother tongue, their first language is English and in most cases they are using English rather than Assamese.

4.3. Part III

In part III an attempt was made to identify the various factors that have contributed to the fluency or lack of fluency in Assamese, by asking participants to identify the factors.

4.3.1. Findings

The proficient participants in Assamese attribute their fluency to these common factors-

- The context of upbringing in a rural area. Some of the participants consider their staying in a rural area helps them to get fluency in their mother tongue. A number of participants believe that their visits to rural area during holidays help them to acquire and maintain fluency.
- It is the result of encouragement from their parents who insist on them learning Assamese at home and especially as a medium of instruction at the beginning of school.
- They are also in a constant contact with their grandparents.

The urban dwellers who are fluent in their mother tongue attributed fluency to -

- Constant communication with family members, relatives, grandparents who are proficient in Assamese.
- Talking with their friends and neighbors.
- Assamese was their medium of instruction at school. It must be mentioned that Assamese is not a compulsory medium of instruction in Assam.
- Continuous encouragement received from their parents and relatives to read and write in their mother tongue.

According to the non-fluent participants, following are the common reasons for the lack of fluency in their mother tongue-

- The limited use of Assamese at school, either as a second language or not at all as a medium of instruction.
- At home, where the parents belong to different ethnic groups² or working parents who leave the child with the domestic help (who belong to Tribal, Adivasi, Bengali or Bihari community etc.), whose native language is different from that of the child.
- Hindi is most frequently used among friends when most of them belong to either different cultural backgrounds or to the same cultural background. For example, Hindi is used most frequently as a medium of communications in those schools while playing where the children belong to various ethnic groups and the school does not belong to a vernacular medium.
- English is the main language of most available printed materials and medium of instruction for higher studies and also serves as a lingua franca.
- Influence of national language policies like the Three Language Formula, that is -(i) The mother tongue or the regional language (ii) The official language of the union i.e. Hindi or the associate official language i.e. English (iii) Modern Indian Language or a foreign language, not covered under (i) and (ii) above and other than that used as the medium of instruction.
- Many parents send their children to English medium schools from the kindergarten stage and such parents encourage their children to speak English at home at the beginning of their school going period, so that they can make improvement in their academic performance.

4.4. Part IV

This question helped to verify whether the youth aims at showing off, particularly in English and Hindi, when speaking to his/her friend in mother tongue in different age groups.

² It refers to those parents who are equally capable to speak in Assamese like native speakers of Assamese. They have used Assamese as their first language since their childhood and as their medium of instruction at school.

4.4.1. Findings

In response to the above question most of the participants answered positively. Figure 3 is representing the response of the participants in showing off when speaking.

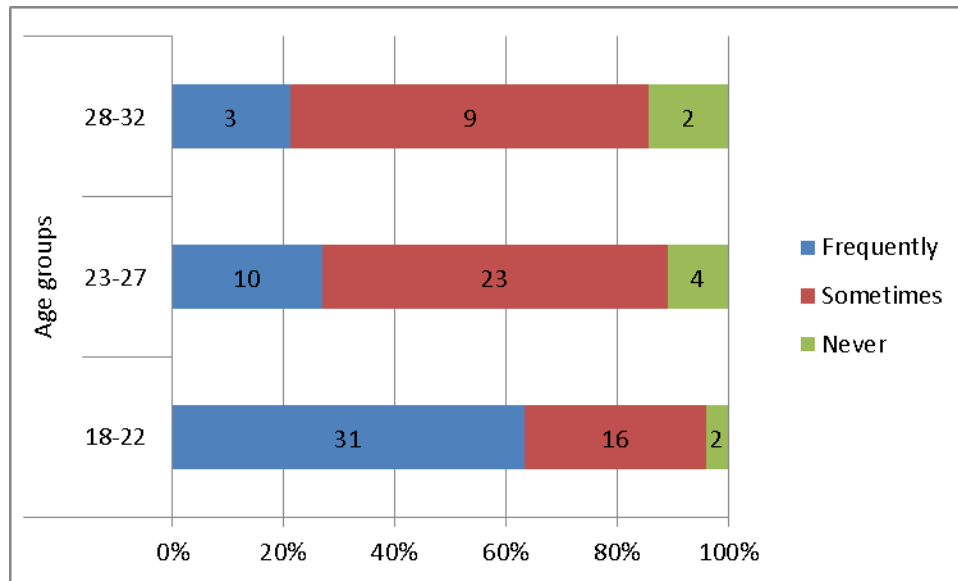


Figure 3 – The number of youths of the sample showing off by using English or Hindi when speaking.

4.5. Part V

This part focused on the participants' own judgement in introducing himself / herself as an Assamese.

4.5.1. Findings

Based on the result of self-assessment each and every participant feels proud as an Assamese, though s/he has not acquired 100% fluency in his/her mother tongue.

5. Influence of English and Hindi on Assamese Language

Since the researcher herself collected data, any inconsistency e.g., differences of the responses to the same forms of questions by the same participant, were cross-checked at the time of interview itself. Responses to some open questions were analyzed separately. The result of the survey found that, depending on the strategy and context of communication, the influence of English and Hindi on Assamese is resulting in a changing pattern in the Assamese language. It is not only the words or expressions that are borrowed in greater numbers, but even the structure of regular grammar in Assamese language is undergoing considerable modification. Some very common and frequently used examples of changing forms and their usage in Assamese are given below. These examples are found in the translated works of the youths during the survey (Part VI), and can be regarded as representing the norm one can find in the speech of youths at present.

(i) Negation in Standard Assamese is expressed by the process of prefixation. But, because of the influence of Hindi on Assamese, a similar pattern of an independent negative form is being developed in the language and it is found in the collected data from the survey.

Interestingly, the following example is mentioned in a book on Assamese grammar which is used as a text book in higher education, Saikia Bora (2006:145).

- (1) *toma-r na rup as-e na gun as-e*
 2₂SG-GEN not beauty be-3 not quality be-3
 ‘Neither have you had beauty nor quality.’

But the standard form is, as in (2)

- (2) *toma-r rup- υ n-ae gun- υ n-ae*
 2₂SG-GEN beauty-EMPH NEG-be3 quality-EMPH NEG-be3
 ‘Neither you have beauty nor quality.’

(ii) The changing trends that have been found from the survey are that the youth speakers of the Assamese language who are brought up under the influence of Western culture, become linguistically unable to correlate the socio-cultural specific expressions due to language contact. The questionnaire included questions to look at whether the respondent could differentiate among the tripartite distinctions of 2nd person pronouns (i.e., the honorific person 2 form *apuni* is shown with subscript 2₁, the familiar form *tumi* is shown with subscript 2₂ and the ordinary form *t ω e* is shown with subscript 2₃) ‘you’. The result shows that they tend to use the ordinary form (*t ω e*) as referring terms in place of the honorific (*apuni*) or familiar form (*tumi*). It must be mentioned that the personal pronouns in Assamese have inherent features of honorificity or status super imposed on the person category. The fused category of person and honorificity in the language is morpho-syntactic as it requires the verb to be in concordial relation with the subject belonging to relevant persons accordingly.

(iii) In Assamese the various degrees of honorificity are marked by the suffixation of classifiers. However, none of the participants was able to encode distinction while choosing between the non-honorific and diminutive (*-t ω*) and (*-z ω ni*) and honorific (*-z ω n*, *-g ω raki*) classifiers. For example the youths used-

- (3) *sar-t ω*
 sir-CL(SG:MASC:NON-HON)
 ‘Sir.’ (Lit. ‘The sir.’)
- (4) *baideu-z ω ni*
 madam- CL(SG:FEM:NON-HON)
 ‘Madam.’ (Lit. ‘The madam.’)

instead of-

- (5) *sar-z ω n*
 sir-CL(SG:MASC:HON)
 ‘Sir.’ (Lit. ‘The sir.’)
- (6) *baideu-g ω raki*
 madam-CL(SG:FEM:HON)
 ‘Madam’ (Lit. ‘The madam.’)

(iv) Assamese exhibits a set of five-term pronominal enclitics with phonologically conditioned allomorphs attached to a kinship term, encoding a concordial relationship between the optional possessor modifier and the possessed head. Table 1 shows two

phonologically conditioned allomorphs, with the vowel-ending forms taking /-k/, /-ra/ and /-r/ and consonant-ending forms taking /-ek/, /-era/ and /-er/ respectively.

Table 1 - Pronominal enclitics (1)

Person	Enclitic	Word form	
1	- ϕ	<i>deuta-ϕ</i> <i>ħok^hi-ϕ</i>	'my father', 'my friend'
2 ₁	-(e)k	<i>deuta-k</i> <i>ħok^hi-ek</i>	'your father', 'your friend'
2 ₂	-(e)ra	<i>deuta-ra</i> <i>ħok^hi-era</i>	'your father', 'your friend'
2 ₃	-(e)r	<i>deuta-r</i> <i>ħok^hi-er</i>	'your father', 'your friend'
3	-(e)k	<i>deuta-k</i> <i>ħok^hi-ek</i>	'his/her father', 'his/her friend'

As seen from the Table 1, excepting the first person, there are overt enclitics for all the person categories - 2₁, 2₂, 2₃ and 3. The kinship terms thus formed can occur in all the syntactic positions with different case markings. In case of the first and second person categories the overt presence of the possessor modifier is optional. But for the 3rd person, which can have a wide range of referents, the presence of the pronominal possessor is obligatory, e.g.,

(7) (*toma-r*) *deuta-ra-e* *toma-k* *mat-is-e*
2₂SG-GEN father-2₂-NOM 2₂SG-ACC call-IPFV-3
'Your father is calling you.'

(8) (*tɔ-r*) *deuta-r-ək* *kɔ^ha-tɔ* *ko-b-i*
2₃ SG-GEN father-2₃-ACC matter-CL tell-FUT-2₃
'Tell your father about the matter.'

However, as the kinship terms marked by the 2₁ and 3 pronominal enclitics are syncretized, the possessor modifiers are obligatory to resolve ambiguities. In order to avoid ambiguity in such cases, the possessor needs to be expressed overtly. e.g.,

(9) *apona-r* *deuta-k-ək* *bat-ɔt* *lɔg* *pa-is-il-ɔ*
2₁SG-GEN father-2₁-ACC way-LOC meet get-IPFV-PST-1
'I met your father on the way.'

(10) *ta-r* *deuta-k-ɔloi* *sit^hi* *lik^h-is-ɔ*
3₁SG-GEN father-3₃-DAT letter write-IPFV-1
'I am writing letter to his father.'

While the participants are asked to translate the word forms of Table 1 from Hindi to Assamese it is noticed that the use of these enclitics are completely avoided by the youths. The youths used the possessor modifier as obligatory to resolve ambiguities as shown in Table 2.

Table 2 -Pronominal enclitics (2)

Person	Word form
1	<i>mor deuta</i> 'my father'
2 ₁	<i>apɔnar deuta</i> 'your father'
2 ₂	<i>tomar deuta</i> 'your father'
2 ₃	<i>tɔr deuta</i> 'your father'
3	<i>iar /tair deuta</i> 'his/her father'

(v) The dropping of genitive, dative, instrumental, and locative case markers and adverbial marker in the speech habit is a very common pattern among youths.

In standard Assamese, the genitive case marker *-ɔr* would be present in the following example. But the data from the survey shows that participants dropped the genitive case marker *-ɔr* while translating the sentence from English to Assamese. The dropping of the genitive case marker *-ɔr* (*pet bih* instead of *pet-ɔr bih*) is the result of an influence of Hindi.

Hindi

- (11) *pet dɔrd kɔmɔr dɔrd me je dɔwa accha he*
 stomach pain waist pain LOC this medicine good be
 'This medicine is useful for stomach and waist pains.'

Assamese

- (12) *pet bih kɔkal bih-ɔt ei dɔrɔb ɣal*
 stomach pain waist pain-LOC this medicine good
 'This medicine is useful for stomach and waist pains.'
 (Lit. 'This medicine is useful for stomach, waist pains.')

In standard Assamese, the dative case marker *-ɔloi* would be present in the following example. But the data from the survey shows that participants dropped the dative case marker *-ɔloi* while translating the sentence from English to Assamese. In both English and Hindi there is no use of case marker. The dropping of the dative case marker *-ɔloi* (*gɔr* instead of *gɔr-ɔloi*) appears to be the result of an influence of English and Hindi. The dropping of the dative case marker *-ɔloi* has become a normal and common practise among the youths e.g.,

English

- (13) *I am going home today.*

Hindi

- (14) *me-ϕ aaz gɔr za rah-a hu*
 I today home go be-PROG. 1
 'I am going home today.'

Assamese

- (15) *azi mæ-φ gər za-m*
 today 1SG-NOM home go-FUT1
 ‘I am going home today.’

(vi) In standard Assamese the sociative case marker *-hoite* would be present in the following example. But the data from survey shows that participants dropped the sociative case marker *-hoite* and used instrumental case marker *-re*, while translating the sentence from English to Assamese. The differences between the sociative (*-dwara, -hoite*) and the instrumental (*-re, -di*) case markers is being dropped, and both are merged in use of Assamese by the youths.

- (16) *mæ-φ tɛv-re alɔsɔna kor-i nɔ-zɔ-a-tɔ*
 1SG-NOM 3₂SG-INST discuss do-NF NEG-go-NF-EMPH

tʰik kor-is-ɔ
 decide do-IPFV-1
 ‘Discussing with him/her I have decided not to go.’

In standard Assamese the locative case marker *-ɔt* would be present in the following example whether the response is in a single word or in a complete sentence. While translating the same sentence from English to Assamese most of the participants dropped the locative case marker *-ɔt*. Following is an example where the locative case marker *-ɔt* (*zɔrhat* instead of *zɔrhat-ɔt*) is dropped.

- (17) Q: *tɔma-r gər kot*
 2₂SG-GEN home where
 ‘Where is your home?’

- (18) Ans: *zɔrhat*
 Jorhat
 ‘Jorhat.’

(vii) In Assamese, adverbs of manner are derived by the suffixation of an adverbial marker *-koi* to some adjectives. In standard Assamese the adverbial marker *-koi* would be present in the following example. While translating the same sentence from English to Assamese most of the participants dropped the adverbial marker *-koi* (*pɔn* instead of *pɔn-koi*).

- (19) *bejam kor-a-r ħɔmɔj-ɔt mɛrudɔndɔ pɔn rakh-ɔk*
 exercise do-NOM-GEN time-LOC spine straight keep-3
 ‘While doing exercises keep your spine straight.’

(viii) In the following example the use of noun+emphatic marker (*biħɔj-e*) in standard Assamese is substituted by a calque (*ɔpɔr-ɔt*) of the English preposition *on*.

- (20) Q: *tɔma-r gɔweħɔna-r biħɔj ki*
 2₂ SG-GEN research-GEN topic what
 ‘What is the topic of your research?’

- (21) Ans: *mɔe-φ* *ɓupen* *hazɔrika-r* *upɔr-ɔt* *gɔwehɔna* *kor-*
 1SG-NOM Bhupen Hazarika-GEN on-LOC research do-

i *as-ɔ*
 NF be-1
 ‘I am pursuing research on Bhupen Hazarika.’

(xi) Following is a common example of use of finite verbs in both relative and correlative constructions while translating the sentence into Assamese from English and Hindi.

Hindi

- (22) *tum* *ne* *zɔ* *bat* *ɓɔta-ja tʰa* *vo* *sɔch* *nɔhi* *tʰa*
 2₂SG ERG which matter say-NF PST that true not PST
 ‘The matter which you have told is not true.’

Assamese

- (23) *tumi* *zi-tɔ* *kɔtʰa* *ko-is-ila* *hei-tɔ* *hɔsa* *n-as-il*
 2₂SG which-CL matter say-IPFV-PST that-CL true NEG-be-PST
 ‘The matter which you have told is not true.’

The older section of the speakers preferred the following pattern of the relative clause which is antecedentless and the verb is marked by non-finite form -

- (24) *tumi* *kɔw-a* *kɔtʰa-tɔ* *hɔsa* *n-as-il*
 2₂SG say-NF matter-CL true NEG-be-PST
 ‘The matter, you have told is not true.’

(x) In standard Assamese a zero copular sentence is formed by juxtaposition of copula subject and copula complement in a sequential order, the verb is absent. The addition of copula verb in such syntactic constructions is also an influence of English and Hindi among youths, e.g.,

Hindi

- (25) *vo* *ek* *sikʰjɔk* *he*
 3₂SG one teacher be
 ‘He is a teacher.’

Assamese

- (26) *tɛɔ* *ɛ-zɔn* *hikʰjɔk* *hɔ-e*
 3₂SG one- (CL:SG:MASC:HON) teacher be-3
 ‘He is a teacher.’

6. Summary and conclusion

The effect of the socio-cultural implications on the mastery of languages and on the cognitive aspect of the youths, have immense fascination with western customs and ideas. It has become fashionable for Assamese youths to have a conversation in English as international language regarding a common topic. Though the government tried to purge Assam from the influence of English by replacing semantically equivalent words, yet there are some English words which have no semantic equivalents in Assamese. Even though some semantically equivalent Assamese words corresponding to English words have been developed, these words are not used by the common speakers rather they prefer to use the English words in their day to day life, e.g., ‘computer’ which is an English word is used by all common speakers instead of using the corresponding Assamese equivalent *ganonik z̄ontr̄o*. Same is the case with ‘mobile phone’. Speakers of Assamese commonly use the English word instead of using Assamese equivalent *bramj̄ob̄ah̄*.

The shifting of language depends on the amount of pressure or attraction that is received from the dominant language. As multilingual speakers, the Assamese youths have always juggled their language to fit in the various contexts. Those youthful members of the Assamese speech community value novelty and by adopting the global language i.e. English, they were taking the western culture, life style and ways of thinking which has also become a social need in the competitive world of the young generation. Part 1 of Section 4.1 reveals that a significantly higher number of youths choose English as the most preferred language, while reading novels, short stories, poems, books/ magazines/ journals, and newspapers. English is also mostly preferred for texting messages and sending e-mails. Assamese is preferred mostly for reading novels and poems. The survey also reveals that while translating the sentences of the questionnaire into Assamese, the spoken form of Assamese of the youths is converging towards English and Hindi structures, as exemplified in Section 5 and this is different from the standard Assamese. The youths used to write the same way as they speak.

During the interview besides the questions asked in the questionnaire, some other related matters also came up in the course of discussion. The respondents opined that the writing style and quality of providing information is better in English than the other two languages. Most of the participants feel comfortable in using English.

The reported trends in language preference by the participants in the study seem to point to a process of language change. Most of the youths, especially those raised in urban areas, have grown up in families where, even in interpersonal domains such as communication with members of the family, frequent use is made of code-mixing or code-switching for communicating needs in the context of discourse. There are some families where the parents encourage the children to speak in English in a bid to improve their academic performance. As a result they have not developed enough proficiency in their mother tongue during their childhood which is the best period of acquiring language.

While Assamese remains the most widely used language in Assam. The process of language change described above in Section 5 is leading to complaints about the falling standards of Assamese. So, since the English words are used very often in formal contexts in which the use of English is prevalent, they easily or automatically come to mind and are quickly adopted in the process of speech production. Thus it also gives the connotation of sharing the same identity and hence a deeper understanding of the situation. The continual use of English and Hindi expressions while speaking Assamese, have had several linguistic consequences on the speakers' proficiency in Assamese. It reflects the linguistic outcomes of language contact which has also become a part of the broader spectrum of communication. As a result, the regular pattern of Assamese grammar is moving towards a potential change.

Abbreviations

1	1 st person	GEN	Genitive case
2 ₁	2 nd person honorific	HON	Honorific
2 ₂	2 nd person familiar	INST	Instrumental
2 ₃	2 nd person ordinary form	IPFV	Infinite verb
3	3 rd person	LOC	Locative case
ACC	Accusative case	MASC	Masculine
ANS	Answer	NEG	Negative
CL	Classifier	NF	Non-finite
EMPH	Emphatic marker	NOM	Nominative
ERG	Ergative case	PST	Past
FEM	Feminine	Q	Question
FUT	Future	SG	Singular

References:

- Census of India (2011). Retrieved from:
<http://www.languageinindia.com/feb2014/censusofindia2011data.pdf>, accessed 12/10/2015.
- Coulmas, Florian. (2005). *Sociolinguistics Discourse Strategies*. 3rd printing, Cambridge University Press, Cambridge.
- Duttabarua, Phanindra Narayan. (2006). *Prayug Bhasa Bigyanar Rooprekha*. Banalata, Guwahati
- Fromkin, Victoria and Robert Rodman. (1974). *An Introduction to Language*. Harcourt Brace Jovanovich College Publishers.
- Government of India, Ministry of Human Resource Development. Retrieved from <http://mhrd.gov.in/language>, accessed 09/10/2015.
- Gumperz, John J. (1982). *Discourse Strategies*. Cambridge University Press Cambridge.
- Peyravbe, Alain & Chaofen Sual. Eds. (1999). *Studies on Chinese Historical Syntax and Morphology*. Ecole Des Hautes Etudes En Sciences Sociales, Centre de Recherches Linguistiques Sur l'Asie Orientale Paris.
- Rabha, T. K. (1984). *Assamese as a second language*. (with modern methods and techniques) Lal Bahadur Shastri National Academy of Administration.
- Saikia Bora, Lilabati. (2006). *Asamiya Bhasar Roptatta*, Banalata, Panbazar, Guwahati.
- Weinreich, Uriel. (1953). *Languages in contact, findings and problems*. New York, Linguistic Circle of New York

Appendix

Questionnaire and interview format for the survey on changing forms and their usage in Assamese among youths.

Namaskar, I am from Tezpur University and I have been working on changing forms and their usage in Assamese among Assamese youths. Towards this I am also carrying a survey, especially on the usage of these forms and the factors behind the use. This may take about 30-35 minutes. Whatever information I shall gather will be used for research purpose only. I thank you all for your help and co-operation.

1. Participant Information

Name:

Gender:

Mother tongue:

Age:

Education:

Medium of instruction at school:

Do you live in

- a. Nuclear family
- b. Joint / Extended Family
- c. Adopted Family

Residential Area

- a. Urban
- b. Rural

What are the other communities in the area?

- a. Language
- b. Tribe / Caste

Place of Interview:

Date:

2. Part I

Following is a list of domains that are used in one's daily social life. The question is about frequency of choosing the language. Numbers are given against each language. Choose the appropriate language and placed it in the box in the right side.

Languages	Number
Assamese	1
Hindi	2
English	3

- (a) (i) Reading novels
- (ii) Stories
- (iii) Poems

- (iv) Books
- (v) Magazines
- (vi) Journals
- (vii) Newspapers.
- (b) (i) Sending text messages (ii) E-mails
- (c) (i) Listening music and radio
- (ii) Watching movies and the programs telecast by television.

- (d) (i) While conversing with parents
- (ii) Parents' friends
- (iii) Participant's friends
- (iv) in social gatherings

Part-II Answer the following-

(i) What is your level of self-declared fluency in your mother tongue in percentage?

(ii) Are you able to read and write equally well in your mother tongue?

Part III What are the factors responsible for the fluency or possible loss of fluency in mother tongue?

Part IV Do you try to show off by using English and Hindi in front of your friends when speaking in your mother tongue? Mark tick in one of the following box by choosing one as the appropriate answer to the above question.

(i) Frequently

(ii) Sometimes

(iii) Never

Part V What is your self-judgement in introducing yourself as an Assamese?

Part VI Following is a conversation among A, B and C which contains some slots with hidden words, Fill in the blanks with appropriate forms of personal pronouns to make complete the sentences and translate the sentences into Assamese.

A: I am feeling hungry. But I do not know how to cook.

B: O.K. I am going to make it. What would _____ like to have?

C: _____ is not at home and _____ went to the market. I would like to have my breakfast with _____

(iii) **Translate the following sentences into Assamese.**

- (a) He does not know how to sing and dance.
- (b) Sir is kind.
- (c) Madam is beautiful.

(iv) **Translate the following Hindi phrases into Assamese.**

- (a) *mere pitaji*
- (b) *apke pitaji*
- (c) *tumhare pitaji*
- (d) *tere pitaji*
- (e) *inke/unke pitaji*

(v) **Translate the following sentences into Assamese.**

- (a) This medicine is useful for stomach and waist pains.
- (b) I am going home today.

(c) Discussing with him/her I have decided not to go.

(d) Q: Where is your home?

Ans: Jorhat.

(e) While doing exercises keep your spine straight.

(f) Q: What is the topic of your research?

Ans: I am pursuing research on Bhupen Hazarika.

(g) The matter which you have told is not true.

(h) He is a teacher