The papers for this volume were initially presented at the sixth and seventh meetings of the North East Indian Linguistics Society, held in Guwahati, India, in 2011 and 2012. As with previous conferences, these meetings were held at the Don Bosco Institute in Guwahati, Assam, and hosted in collaboration with Gauhati University. The present collection of papers are testament to the ongoing interest in North East India and continued success and growth in the community of North East Indian linguists. As in previous volumes, all the papers here were reviewed by leading international specialists in the relevant subfields. This volume, in particular, highlights the recent research of many scholars from the region. Out of eleven contributions, eight are from North East Indian scholars themselves. This book therefore brightly shines light on the work being done by North East Indian linguists on the languages of their own region. The remaining contributions are authored by international scholars from Australia, Singapore, Germany/USA, and Nepal.
## Contents

*About the Contributors*

Foreword
Shobhana Chelliah

*A Note from the Editors*

### Phonology

1. Acoustic-statistical and perceptual investigations of Karbi tones: ..............................13
   *A peculiar case of incomplete neutralization of F0*
   *Linda Konnerth and Amos Teo*

2. Syllable Restrictions in Chungli (Ao) ........................................................................39
   *T. Temsunungsang*

### Morphosyntax

3. Coordination in Manipuri ...........................................................................................63
   *S. Indrakumar Singh and Ch. Yashawanta Singh*

4. Passive Constructions in Assam Sadri .................................................................77
   *Lucky Dey*

5. Negation in Rongmei Naga .....................................................................................95
   *Debajit Deb and Kh. Dhiren Singha*

6. Differential marking of cases in Asamiya ...............................................................107
   *Runima Chowdhary*

7. Non-nominative subjects in Assam Sadri ...............................................................127
   *Lucky Dey*

### Historical Linguistics

8. Proto-Khasian: An emerging reconstruction .........................................................149
   *Paul Sidwell*

### Language Contact

9. Contact-induced changes in Baram .........................................................................167
   *Dubi Nanda Dhakal*
10. Phonological changes in the Hindi lexicon: a case of Meghalaya Hindi ..................193
   Maansi Sharma

11. Spirantization in Tibeto-Burman-Bengali bilinguals in south of Assam ..................213
   Kakoli Dey
About the Contributors

Editors

Gwendolyn HYSLOP (gwendolyn.hyslop@anu.edu.au) is a Research Fellow in Linguistics at the Australian National University’s College of Asia and Pacific. She received her PhD in Linguistics from the University of Oregon in 2011, producing a grammar of the Tibeto-Burman language Kurtöp. She specializes in languages of Bhutan and the eastern Himalayan region.

Linda KONNERTH (lkonnert@uoregon.edu) is a Postdoctoral Research Associate in Linguistics at the University of Oregon, USA, where she also completed her PhD dissertation ‘A grammar of Karbi’. She has recently started working on the Northwestern Kuki-Chin languages of Chandel District in Manipur and is writing a grammatical description of Monsang. Her publications include topics in historical-comparative linguistics, syntax, and information structure in Tibeto-Burman.

Stephen MOREY (s.morey@latrobe.edu.au) is an Australian Research Council Future Fellow at the Centre for Research on Language Diversity, La Trobe University. He is the author of two books on tribal languages in Assam, from both Tai-Kadai and Tibeto-Burman families. He is the co-chair of the North East Indian Linguistics Society and has also written on the Aboriginal languages of Victoria, Australia.

Priyankoo SARMAH (priyankoo@iitg.ernet.in) is an Assistant Professor in Linguistics at the Department of Humanities and Social Sciences, Indian Institute of Technology Guwahati. He specialises in the phonetics and phonology of the languages spoken in North East India, such as Assamese, Bodo, Dimasa, Mizo and Tiwa. He has published articles on tones and vowels of North East Indian languages.

Authors

Runima CHOWDHARY (runimachowdhary@yahoo.com) is Associate Professor in Linguistics at Gauhati University in Assam, India. She has authored a monograph on Assamese verbs and a number of research articles on various aspects of Assamese linguistics. She is a life member of the Linguistic Society of Assam and the Linguistic Society of India.

Debajit DEB (debojit.rm@gmail.com) works at Assam University Silchar. He has published many papers in various journals of linguistics and attended/presented many national and international seminars, conferences, symposiums and workshops.

Kakoli DEY (kakolidey6@gmail.com) has a Doctorate from University of Delhi, Delhi. Her specialization is in language variation and sound change. Her research interest lies in sociolinguistics, phonetics and phonology, morphosyntax, language acquisition, communication and technical writing, and she has worked on languages of northeastern India and language contact. Currently, she is a faculty in a private engineering college, Delhi-NCR teaching Professional Corporate Communication and Technical Writing.
Lucky DEY (lucky@tezu.ernet.in) worked on a variety of Sadri, an Indo-Aryan language spoken in Assam, for her doctoral thesis. She completed her PhD at the Department of English and Foreign Languages at Tezpur University with Prof. Madhumita Borbora. She is presently working as a Project Scientist on a project titled "Digital Language Preservation" sponsored by DeitY in the Department of English and Foreign Languages at Tezpur University. For the project, she is studying two endangered languages spoken in Arunachal Pradesh, India.

Dubi Nanda DHAKAL (dubidhakal@yahoo.com) is an Associate Professor at Central Department of Linguistics, Tribhuvan University, Kathmandu, Nepal. His interests include language documentation, language contact, language typology and field linguistics. He has worked and published on lesser-described Indo-Aryan and Tibeto-Burman languages of Nepal.

Linda KONNERTH (lkonnert@uoregon.edu) is a Postdoctoral Research Associate in Linguistics at the University of Oregon, USA, where she also completed her PhD dissertation 'A grammar of Karbi'. She has recently started working on the Northwestern Kuki-Chin languages of Chandel District in Manipur and is writing a grammatical description of Monsang. Her publications include topics in historical-comparative linguistics, syntax, and information structure in Tibeto-Burman.

Maansi SHARMA (maansi16@gmail.com) is a field linguist and completed her doctorate on the grammar of contact Hindi. It was the first grammar ever written on contact Hindi. She is interested in language contact and its impact on various languages.

Paul SIDWELL (paulsidwell@gmail.com) received his PhD in linguistics from Melbourne University in 1999. He is based at the Australian National University (Canberra), where since 2012 he has held an Australian Research Council Future Fellowship. Sidwell is a leading specialist in Austroasiatic linguistics, especially the Katuic and Bahnaric branches; he also has interests in forensic linguistics and phonetics. He also is the editor for the Journal of the Southeast Asian Linguistics Society and the Mon-Khmer Studies Journal.

Ch. Yashawanta SINGH (chungkhamyash@gmail.com) is a Professor in the Department of Linguistics, Manipur University, and specializes in Tibeto-Burman studies. He has written grammars of Manipuri, Tarao and Koireng. His research publications include topics in descriptive linguistics, morphology, phonology, syntax, sociolinguistics, pragmatics, and the structure of Tibeto-Burman.

S. Indrakumar SINGH (iksagol@gmail.com) has a PhD in Linguistics, having researched the topic "Manipuri Clause Structure" at Manipur University, Canchipur, Imphal. He is currently working as a linguist in the project entitled "Indian Languages Corpora Initiative ILCI". His publications include topics in syntax in the Tibeto-Burman area.

Kh. Dhiren SINGHA (dhirensingha@rediffmail.com) is Associate Professor in Linguistics at Assam University, Silchar, Assam. He is the author of three books: An Introduction to Dimasa Phonology, Dimasa Word Book: A Classified Vocabulary and Ahni Grao: My Language (A Dimasa English Bilingual Primer). His areas of interest are language typology, morphosyntax and Tibeto-Burman linguistics.
T. TEMSUUNSANG (ttjamir@gmail.com) is Assistant Professor at the English and Foreign Languages University in Shillong, India. He received his PhD in Linguistics from English and Foreign Languages University in Hyderabad in 2010. He specializes in phonology (especially tone), morphology, Tibeto-Burman languages, and language acquisition.

Amos TEO (amosbteo@gmail.com) is a current PhD graduate student in the Linguistics Department at the University of Oregon. He completed an MA in Linguistics at the University of Melbourne. He has been working on Tibeto-Burman languages of Northeast India since 2007, when he started to work on Sumi, a language of Nagaland. His main research interests include prosodic typology, as well as tone production and perception.
Foreword

Shobhana L Chelliah
University of North Texas

It is my distinct pleasure to write a foreword for this volume of selected papers presented at the 6th International Conference of the North East Indian Linguistics Society (NEILS) held on the banks of the majestic Brahmaputra at the Don Bosco Institute in Guwahati, Assam. Congratulations to the conference organizers, Stephen Morey, Mark Post, Jyotiprakash Tamuli and to the volume editors (Gwendolyn Hyslop, Linda Konnerth, Stephen Morey, Priyankoo Sarmah) for their continued dedication to encouraging quality linguistic training, research, and dissemination in Northeast India.

This volume, along with others in the series, has the broad impact of encouraging local and international linguistic scholarship from Arunachal Pradesh, to Assam, Meghalaya, Manipur, Mizoram, Nagaland, and Tripura as well as to bordering states. The articles are theory-neutral, example-rich descriptions based on original field data. They can readily be used as examples by junior scholars on how best to document a language in the Northeast Indian context. The mentoring that takes place at the NEILS conferences extends seamlessly into mentoring on research dissemination to an international audience, an audience that would otherwise most likely not benefit from this local scholarship. The fact that many of the junior scholars participating at NEILS conferences are community members documenting their heritage language makes the articles even more valuable both from cultural and scientific perspectives.

NEILS 6 includes articles on a range of topics from nuanced descriptions of language to contact situations, morphosyntax (passive, differential case marking, coordination), phonetics and phonology of tone, and phonological reconstruction. It is my hope that many more NEILS meetings are held and many more such volumes are produced to continue to enrich our understanding of human language and cognition.
A Note from the Editors

The papers for this volume were initially presented at the sixth and seventh meetings of the North East Indian Linguistics Society, held in Guwahati, India, in 2011 and 2012. As with previous conferences, these meetings were held at the Don Bosco Institute in Guwahati, Assam, and hosted in collaboration with Gauhati University. The present collection of papers are testament to the ongoing interest in North East India and continued success and growth in the community of North East Indian linguists. As in previous volumes, all the papers here were reviewed by leading international specialists in the relevant subfields.

This volume, in particular, highlights the recent research of many scholars from the region. Out of eleven contributions, eight are from North East Indian scholars themselves. This book therefore brightly shines the light on the work being done by North East Indian linguists on the languages of their own region. The remaining ones are authored by international scholars from Australia, Singapore, Germany/USA, and Nepal.

As with previous volumes, the papers reflect the diversity of the North East through their diverse array of topics and approaches. Our section on phonology presents an acoustic analysis of tone, on the one hand, and an OT-inspired analysis of syllable restrictions, on the other. The findings of both—in addition to presenting data on the relatively undescribed languages of Karbi and Chungli Ao—offer real contributions to our theories of phonology. Konnerth and Teo find that pitch is almost completely neutralized in the realization of tone in some Karbi contexts, leading to an interesting mismatch between tone perception and production. Temsunungsang shows that while a relationship between rimes and codas is found in Chungli (as is to be expected), there is also evidence for an important relationship between the onset and rime.

Five papers deal with morphosyntax, with two paper devoted to Tibeto-Burman languages and three papers dealing with Indic varieties. As for the studies in Tibeto-Burman morphosyntax, Singh and Singh’s contribution looks at the syntax and semantics of syntactic and asyndetic coordination in Manipuri. Deb and Singha present an analysis of a range of data on negation in the little-researched Rongmei Naga language. Turning to studies in the morphosyntax of Indic languages, Lucky Dey has two papers that present analyses from the little-studied Assam Sadri. Her first paper in this volume looks at passivization while the second paper presents an analysis of non-nominative subjects in the language. The contribution from Runima Chowdhary is in a similar vein, discussing differential case marking in Asamiya (Assamese).

Paul Sidwell’s paper brings into fruition the result of years of labour in data collection and analysis, presenting for the first time a reconstruction of Proto-Khasian. Though widely spoken in Meghalaya and surrounding regions of Assam, the Khasian languages have managed to remain relatively unknown to the outside world and seem to resist being shaped by the typological features that otherwise influence the neighbouring Indo-Aryan and Tibeto-Burman languages. This contribution is the first look at the historical development of the Khasian languages.

The final section of this volume addresses language contact between large Indo-Aryan languages and smaller languages from the Tibeto-Burman and Austroasiatic language families. Dhakal’s paper examines Nepali borrowing into the highly endangered Baram language. Dhakal shows that much of the Baram lexicon and a considerable amount of the grammar have already been replaced by Nepali. Parallel situations of language endangerment and language replacement threaten many languages in North East India. Sharma’s paper
examines the phonological changes in Hindi as spoken in Meghalaya among native Khasi speakers. Finally, Kakoli Dey offers a sociolinguistic study of the variable of spirantization in Tibeto-Burman/Bengali bilinguals.

With the move to Asia-Pacific Open Access publications, this book marks the beginning of a new and welcomed era in the publication of *North East Indian Linguistics*. We are pleased that, for the first time, NEILS papers will be available for free download and hope with this move the topics covered in these pages will be more easily accessible to our colleagues in India as well as those people about whose languages we write.

Lastly, we wish to thank everyone who helped bring this book into fruition, including the authors, peer reviewers, and the local team in Canberra, especially Paul Sidwell and Jack Plane.

**Gwendolyn Hyslop**  
*Canberra, Australia*

**Linda Konnerth**  
*Liwachangning, Chandel, Manipur, India*

**Stephen Morey**  
*Melbourne, Australia*

**Priyankoo Sarmah**  
*Guwahati, India*
Phonology
1. Acoustic-statistical and perceptual investigations of Karbi tones

1. Acoustic-statistical and perceptual investigations of Karbi tones: A peculiar case of incomplete neutralization of $F_0$1

Linda Konnerth
University of Oregon

Amos Teo
Australian National University

Abstract
This article presents two studies that investigate the phonetics and phonology of the Karbi (Tibeto-Burman; Assam) lexical tone system: an acoustic-statistical study and a follow-up perception study. Although currently called a tone system, the two studies show how calling it such is not uncontroversial. A previous account of the Karbi tone system proposes three tonal categories (low, mid, high) for open or sonorant-final syllables, whereas checked syllables (ending in /p,t,k/) may only be low or high; additionally, while low and high tones are realized as differences in pitch height ($F_0$), the mid tone also exhibits characteristic glottalization word-finally (Grüßner 1978). In fact, a preliminary acoustic investigation by Konnerth (2010) suggests that the pitch height of the mid tone is non-unique and especially overlaps with the high tone in the case of bare monosyllabic stems. It is proposed that the word-final glottalization is the more important phonemic cue. The current acoustic-statistical study expands the scope of investigation by looking not only at bare monosyllabic stems but also the same stems featuring one of two suffixes: a low tone suffix - pó and a high tone suffix - jí. By obtaining expanded data sets from both a male and a female speaker, we were able to include a preliminary statistical analysis of $F_0$ contours, which confirms the pitch height overlap between mid and high tones for bare stems, - pó suffixed stems, and - jí suffixed stems in the case of the female speaker. The male speaker shows an interestingly different pattern with sufficient differentiation between mid and high tones in all cases, which other native speakers easily replicate when contrasting minimal pairs. To investigate the phonemic basis of the mid vs. high tone contrastive production, we followed up with a perception study, for which seven native speakers engaged in an identification task of tonal minimal pair items, listening to the stimuli employed in the original acoustic study from the male and the female speaker. Our results suggest that the pitch height differentiation between the mid and high tones as produced by the male speaker is not phonemic but may in fact indicate a covert contrast.

Citation

Volume Editors
Gwendolyn Hyslop, Linda Konnerth, Stephen Morey, Priyankoo Sarmah

Copyright © 2014, the author(s), release under Creative Commons Attribution license

1. Introduction

Karbi is a Tibeto-Burman language of Northeast India, spoken by about half a million people mostly in the Karbi Anglong district and neighboring areas of Assam. It is a largely agglutinating and to some degree isolating, verb-final language. This article presents two studies in Karbi phonetics and phonology to investigate the tone system: an acoustic study and a follow-up perception study.

The Karbi tone system (i.e., the suprasegmental phenomena described here, which perhaps are not uncontroversially referred to as a tone system) was first described by Grüßner (1978) as a lexical tone system with three categories: low, mid, and high. All three categories occur in open or sonorant-final syllables, and can be exemplified by a few minimal triplets.

---

1 We would like to thank all Karbi native speaking participants of the two studies: Punyalata Engtipi, Sashikala Hansepi, Sampri Kro, Maloti Rongpharpi, Liladhon Teron, Rani Teronpi, and especially Sikari Tisso, who also helped facilitate the study. In addition, we thank participants of NEILS7 and an anonymous reviewer for their comments and suggestions. The first author acknowledges support from the National Science Foundation under grant BCS-0951749 ‘A descriptive grammar of Karbi’.
Checked syllables (ending in /p, t, k/) may only carry low or high tone according to Grüßner's (1978) analysis.\(^2\)

As already noted by Grüßner (1978), all three tones are realized as differences in respective F\(_0\) levels or pitch height, while the mid tone also exhibits characteristic glottalization word-finally. However, the pitch height of the mid tone might be so variable that the word-final glottalization could become the actual phonetic cue for this tonal category in word-final contexts. This is the finding of a preliminary acoustic investigation by Konnerth (2010), which suggests a strong overlap in pitch height between the mid tone and especially the high tone in the case of bare monosyllabic stems. We still continue to call this tone the mid tone following Grüßner (1978) as it is currently not clear what a more appropriate alternative label would be.

Karbi is written but exact rules for orthography have not been agreed upon at this point. There are currently no conventions in place to represent tone in the orthography. There have been proposals to include some tonal information in the spelling system, but they have never been comprehensive. For example, one proposal was to write a homorganic stop after a syllable with a nasal coda if that syllable is mid tone, presumably to represent the glottalization (e.g., write “langk” for mid tone lāng ‘water’). However, this proposal did not include the representation of the low or the high tone, and not even the consistent representation of the mid tone across all possible rhymes. Native Karbi speakers generally find it very difficult to compare and identify the tonal category of morphemes/words (particularly beyond monosyllables) (see Konnerth 2014). Furthermore, there are only a few minimal triplets for tone, and typically not all items in these triplets belong to the same word class. This suggests that lexical tone has a low functional load in the language, as has been noted for other languages of the area, including the Tibeto-Burman language Boro (Scott DeLancey, p.c.).

The current acoustic-statistical study expands the scope of investigation by looking not only at bare monosyllabic stems but also the same stems featuring one of two suffixes: a low tone suffix -pò and a high tone suffix -jí.\(^3\) The majority of stems we investigate are verbal stems, but some nominal stems are included as well (see below, Table 2); the suffixes are irrealis markers (see discussion below). By adding suffixed forms, we investigate the phonetic realization of the mid tone on stems, where glottalization does not occur due to the moved word boundary. Compared to previous studies, we obtained expanded data sets from both a male and a female speaker, which made it possible to include a statistical analysis of F\(_0\) values. Our results confirm the pitch height overlap between mid and high tones for bare stems, -pò suffixed stems, and -jí suffixed stems in the case of the female speaker. However, the male speaker shows an interestingly different pattern with sufficient differentiation between mid and high tones in all cases. This pattern of contrasting the mid and the high tone in fact appears to be replicated, according to our auditory observations, by other native speakers, including our female speaker, when contrasting minimal pairs. To investigate the phonemic basis of the mid vs. high tone contrastive production, we followed up with a perception study, for which seven native speakers engaged in an identification task of tonal minimal pair items, listening to the stimuli employed in the original acoustic study from the male and the female speaker. Our results suggest that the pitch height differentiation between the mid and high tones as produced by the male speaker is not phonemic but may in fact indicate a covert contrast.

This article is structured as follows. Section §2 provides information on Karbi phonology and the relevant literature as well as formulates the research questions we are interested in.

---

\(^2\) Note, however, that Konnerth (2014) now also finds a category of mid tone checked syllables. This category is not further discussed in this article.

\(^3\) Low tone is indicated by /ˈ/, high tone by /ˈ/, and mid tone by /ə/.
1. Acoustic-statistical and perceptual investigations of Karbi tones

Section §3 discusses the methodology, results, and a brief discussion of the results of our acoustic-statistical study. Section §4 deals with the follow-up perception study, and also describes methodology, results, and includes a brief discussion. Section §5 discusses the implications that the results from both studies described in §3 and §4 bear on our analysis of the Karbi tone system. We draw a conclusion and discuss future directions in §6.

2. Language background

Karbi has 3 syllable types: open, ending in a monophthong or a diphthong; closed, with a sonorant in coda position; or checked, with a stop in coda position (see Table 1). According to Grüßner (1978), checked syllables can only take a low or high tone. There are 17 consonants that are permissible in onset position, but only seven that are permissible in coda position.

Table 1: Karbi syllable types

<table>
<thead>
<tr>
<th>Syllable type</th>
<th>Schematic</th>
<th>Tone possibilities (Grüßner 1978)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open</td>
<td>(C)(CRL)V(V)³</td>
<td>L-M-H</td>
</tr>
<tr>
<td>Sonorant-final</td>
<td>(C)(CRL)VN</td>
<td>L-H</td>
</tr>
<tr>
<td>Checked</td>
<td>(C)(CRL)VS</td>
<td>L-H</td>
</tr>
</tbody>
</table>

N: /m,n/ or /r/; S: /p,t,k/; C RL: /r,l/

Previous literature on the Karbi tone system is limited to two sources: Grüßner (1978) and Konnerth (2010). The comprehensive analysis of the Karbi tone system by Grüßner (1978) was part of his descriptive grammar of the phonology and the morphology of the language. At the time, Grüßner did an excellent job of describing Karbi phonology while having to rely only on auditory observations. Grüßner did the groundwork of describing the Karbi phoneme inventory, syllable structure, and tone system while providing minimal pairs and triplets, as well as prosodic features; he gave an account of the phonological processes at work in borrowings from languages such as Khasi, Assamese, and English; and he worked out interactions between tones and tone changes as morphophonemic processes in the language (Chapters 2-4). Grüßner (1978: 20-1) also addressed the phonetic realization of the three tones. We can summarize his account as follows:

- only pitch height, not pitch contour, is phonemic
- L tone is produced at the normal pitch height of individual speakers' pitch level, and exhibits a slightly falling contour in open syllables
- H tone exhibits a slightly rising contour in open syllables
- for the M tone, the phonetic realization depends on other prosodic and morphological factors:
  a. on a stressed syllable either before a pause (word-finally) or before a vowel; or on an unstressed syllable preceding another M tone syllable: mid level tone with a glottal stop following the vowel or sonorant coda (mid tone does not occur in checked syllables)
  b. within a word, on a stressed syllable, preceding a L or H tone syllable: mid level tone without glottal stop, somewhat shortened vowel on open syllables and significantly shortened vowel on closed syllables
  c. within a word, on a stressed syllable, preceding another M tone syllable: becomes H tone and vowel gets shortened, with following M tone getting realized as a L tone

Footnote: ³ Only the following onset consonant clusters are allowed. /pr,pl,pʰr,pʰl,tʰr,tʰl,t,kr,kl,kʰr/
We see that the M tone is the complex and irregular tone category, as it sometimes induces glottalization; seemingly sometimes shortens its vowel (likely connected to the glottalization); may change to a H tone; and is generally sensitive to stress and syllable position, unlike the L and H tones. In addition and contrary to Grüßner's description, our auditory observations suggest that there does not actually appear to exist a unique mid tone pitch level located between the low and high tone level. It still needs to be noted that while the mid tone may not have a clear pitch correlate, it is in fact part of the tone system together with the low and the high tone, because it participates in morphophonemic tone changes. For example, low tone monosyllabic roots regularly become mid tone following the reflexive/reciprocal prefix che-, and mid tone monosyllabic roots regularly become high tone following that same prefix, such that low tone làng ‘see’ becomes mid in che-làng ‘RR-see’, and mid tone ēn ‘take’ becomes high tone che-ēn ‘RR-take’.

The auditory observations suggesting that there is no unique mid tone pitch level led to a preliminary phonetic investigation, where a sample of monosyllabic stems representing all three tones was recorded (Konnerth 2010). The results showed that the pitch range of the mid tone overlapped with the pitch range of the high tone and, to a lesser degree, with the pitch range of the low tone. There are several caveats, however, that need to be noted for the 2010 preliminary study. First, the stimuli were recorded as three repetitions in isolation and once in a carrier phrase, and only the ones in isolation were measured. Tone contours over items in isolation may be affected (and distorted) by other prosodic factors such as phrase-final list intonation, which makes this methodology problematic. Without the context of a phrase, it was also difficult to judge if differences in pitch height between lexical items (or lack thereof) were due to the consultant producing lexically contrastive pitch or simply speaking in a different key across the utterance. Second, the 2010 study was limited to stimuli of only monosyllabic stems. Finally, all recordings were done with a single male speaker.

Keeping the caveats of the 2010 study in mind, for the current study we recorded four tokens of our stimuli in the pattern of one repetition in isolation and three in a carrier phrase. In addition to re-recording the same male speaker from the 2010 study, we also recorded the same stimuli from a female speaker living in the same town. The current acoustic study then aims at exploring two major research questions. Our first question is: By eliminating the caveats from the 2010 study, can we still replicate the result that M tone items have F0 values that overlap especially with H tone items on monosyllabic bare stems? In other words, is pitch height, measured as F0, a reliable cue to identify the M tone? Since the glottalization, which is characteristic of word-final M tone, seems to disappear word-medially, our second main research question is: What happens to the M tone F0 if suffixes are added, which move the word boundary? In order to get at this question, we recorded stems not only in their bare forms but also with one of two suffixes: a low tone suffix -pò and a high tone suffix -jì.

As §3 will show, the male speaker's productions differentiate pitch levels for M and H tones, while the female speaker's productions do not. This leads us to the perception study in §4 with the larger research question in mind: Which aspects of the Karbi tone system - and here specifically regarding the differential pitch realization of M and H tones - are truly phonological and which are perhaps only phonetic?

---

5 Note that the 2010 study did not find notable differences in pitch height between tokens produced in isolation and tokens produced in a carrier phrase.

6 Note that interestingly, the male speaker's production of tone in the 2010 study is quite different from his production in the current study. In the 2010 study, his productions were more similar to the female speaker's productions presented below.
3. Acoustic-statistical study

The acoustic study was conducted with the following research questions in mind:

- Do Karbi speakers distinguish the M and H tones by producing a difference in pitch height (in addition to producing word-final glottalization in conjunction with the M tone)?
- How are the M and H tones distinguished when they do not occur in word-final position, i.e. when the lexical stem receives a suffix, thus removing the salient cue of glottalization?

After discussing our methodology in §3.1, we consider these questions by first looking at the results of the acoustic study of pitch on monosyllabic bare stems in §3.2, followed by a brief discussion of our findings in §3.3. We then look at the results of the acoustic study of pitch on stems with suffixes in §3.4. This is followed by a discussion of these results in §3.5.

3.1. Methodology

For this study, one male and one female speaker were recorded in Diphu, the capital of the Karbi Anglong District. The female speaker was 36 years old and the male speaker 52 years old. The recordings were conducted with a Shure SM10A head-mounted, unidirectional, dynamic microphone and a Zoom H4N digital audio recorder.

Speakers were asked to first produce the lexical item once in isolation and then repeat it three times in the carrier phrase shown in (1). The carrier phrase consists of the first person exclusive pronoun with the honorific suffix, followed by the lexical item under investigation in a quotative construction that consists of postposed pù ‘say’ with the non-final suffix -si, and finally the main verb pù ‘say’ with the nominalizer kV-, which indicates the imperfective.

(1)  Nè-li ______ pù-sì’  ki-pù.
     1EXCL-HON ______ say-NF  NMLZ-say
     ‘I am saying ______.’

A carrier phrase was deemed necessary since we were looking for differences in pitch height between the M and H tones. The phrase provided important auditory information that allowed us to judge if the speaker was modifying pitch on only the investigated item or across the whole phrase – while the occurrence of the former might correspond to a phonological tone category, the occurrence of the latter would indicate only a change in key across the entire utterance. This particular phrase was chosen because it allowed us to insert both verbs and nouns (with additional suffixes, due to the quotative construction). By placing the item in phrase-medial position, the effects of post-lexical intonation at the phrase boundaries, especially list intonation, were also reduced.8

The investigated lexical items consist of bare monosyllabic stems of 59 verbs and three nouns (see Table 2), as well as the same stems with the suffixes: -pò ‘irrealis1’ and -jí ‘irrealis2’. These suffixes are two of the four most frequently used endings of final verbs in

---

7 Although Karbi has a (contrastive) focus marker =si ‘FOC’, the -si in pu-si is not this focus marker but rather the non-final marker -si. The grammaticalization of the stem ‘say’ with a non-final suffix (often developing purposive or complementizer functions) is commonly found in Tibeto-Burman, e.g., in Galo (Post 2007: 635) or in Baram (Kansakar et al. 2011: 162).

8 It is acknowledged that no carrier phrase is a prosodically-neutral context and there is a chance that speakers may place either narrow focus on the investigated item or broad focus across the verbal phrase. An anonymous reviewer suggested that there may have been differences in which the speakers interpreted the information structure of the carrier phrase, given that the female speaker appears to deaccent the /pusi kipu/ following the lexical item under investigation to a greater degree than the male speaker. At present, it is unclear to us how this could explain the differences in the productions of the mid versus the high tone that are described later in this article, but it represents an important direction for future research.
Karbi, the other two being a realis and a negative suffix, while bare stems also figure as final verbs. They may attach to noun stems as well (which then function as nominal predicates), since suffixes in Karbi arguably are not specified for the part of speech of the stem (Konnerth 2014).

Table 2: List of stimuli recorded as bare stems, -jí suffixed stems, and -pò suffixed stems for acoustic-statistical study

<table>
<thead>
<tr>
<th>Item</th>
<th>Gloss</th>
<th>Item</th>
<th>Gloss</th>
<th>Item</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>àp</td>
<td>'shoot'</td>
<td>jâng</td>
<td>'contain'</td>
<td>sâng</td>
<td>'spread'</td>
</tr>
<tr>
<td>áp</td>
<td>'be able'</td>
<td>jâng</td>
<td>'fall'</td>
<td>sâng</td>
<td>'raw rice (n.)'</td>
</tr>
<tr>
<td>bân</td>
<td>'be old'</td>
<td>jùn</td>
<td>'drink'</td>
<td>sâng</td>
<td>'rest'</td>
</tr>
<tr>
<td>bân</td>
<td>'persist'</td>
<td>klô</td>
<td>'fall'</td>
<td>sêk</td>
<td>'attach'</td>
</tr>
<tr>
<td>bî</td>
<td>'be small'</td>
<td>kôm</td>
<td>'surround'</td>
<td>sêk</td>
<td>'burn'</td>
</tr>
<tr>
<td>bî</td>
<td>'keep'</td>
<td>kóm</td>
<td>'be less'</td>
<td>tô</td>
<td>'spread in sun'</td>
</tr>
<tr>
<td>bôk</td>
<td>'put on back'</td>
<td>lâng</td>
<td>'see'</td>
<td>thâk</td>
<td>'weave; answer'</td>
</tr>
<tr>
<td>bôk</td>
<td>'put on back'</td>
<td>lâng</td>
<td>'water (n.)'</td>
<td>thâk</td>
<td>'divide'</td>
</tr>
<tr>
<td>bû</td>
<td>'carry on back'</td>
<td>lâng</td>
<td>'water (n.)'</td>
<td>thân</td>
<td>'tell'</td>
</tr>
<tr>
<td>bû</td>
<td>'overflow'</td>
<td>lêt</td>
<td>'make mistake'</td>
<td>thâng</td>
<td>'pound'</td>
</tr>
<tr>
<td>châm</td>
<td>'be wet'</td>
<td>lêm</td>
<td>'play with toys'</td>
<td>thât</td>
<td>'slaughter'</td>
</tr>
<tr>
<td>chôm</td>
<td>'eat'</td>
<td>lîm</td>
<td>'spread a mat'</td>
<td>thì</td>
<td>'die'</td>
</tr>
<tr>
<td>dâr</td>
<td>'break'</td>
<td>mân</td>
<td>'have burning sensation'</td>
<td>thî</td>
<td>'be short'</td>
</tr>
<tr>
<td>dâm</td>
<td>'oust'</td>
<td>mân</td>
<td>'become'</td>
<td>thî</td>
<td>'be short'</td>
</tr>
<tr>
<td>dô</td>
<td>'stay'</td>
<td>mê</td>
<td>'be good'</td>
<td>thî</td>
<td>'snatch'</td>
</tr>
<tr>
<td>è</td>
<td>'plant'</td>
<td>nîm</td>
<td>'soak'</td>
<td>thû</td>
<td>'be tasty'</td>
</tr>
<tr>
<td>èn</td>
<td>'take'</td>
<td>nông</td>
<td>'plough'</td>
<td>thû</td>
<td>'be rotten'</td>
</tr>
<tr>
<td>hûm</td>
<td>'pick up'</td>
<td>nông</td>
<td>'fame (n.)'</td>
<td>thû</td>
<td>'cut'</td>
</tr>
<tr>
<td>hûm</td>
<td>'starve'</td>
<td>pî</td>
<td>'give'</td>
<td>tôn</td>
<td>'point'</td>
</tr>
<tr>
<td>i</td>
<td>'lie down'</td>
<td>rõ</td>
<td>'hit'</td>
<td>tông</td>
<td>'hurry'</td>
</tr>
<tr>
<td>i</td>
<td>'put on'</td>
<td>rõ</td>
<td>'praise'</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The items were recorded by tone category, i.e., all high tone roots at once, all mid tone roots at once, and all low tone roots at once. This did not follow the order of minimal sets used for presentational purposes in Table 2.

The software program ‘EMU Labeler’ (version 2.3.0) (Cassidy and Harrington 2001) was used in conjunction with the statistical program ‘R’ (version 2.13.1) (Harrington 2010) to analyze the recordings and extract F0 values. The Pitch and Formant Tool in ‘EMU Labeler’ was used to generate pitch traces. Two interval tiers: ‘Word’ and ‘Segment’; and one point tier ‘Tone’ were created to do the annotations in the program. The assignment of tone label followed tonal categories previously established with language consultants, which were by and large the same tone category assignments found in Grüßner’s work.

Using the statistical program ‘R’, F0 values were extracted across the entire tone-bearing vowel segment at 10% intervals. Only the pitch across the vowel portion of each token was extracted – the start and end of the tone-bearing vowel segment was defined by the start and end of clear formant structure, and prior to the setting in of any syllable-final glottalization.

9 Predicate post-stem morphological slots including the negative, an aspect, and a modal (realis versus irrealis) slots are not word class specific and can occur on verb stems as well as noun stems (although there is clearly a strong tendency for them to occur on verb stems), see Konnerth (2014).
1. Acoustic-statistical and perceptual investigations of Karbi tones

This was to allow us to group the open and sonorant-final syllables together for the acoustic analysis.10

A complementary acoustic study looked at the production of pitch relative to the carrier phrase, measured in semi-tones. This was done in order to determine whether differences in pitch height occurred only on the investigated item or across the entire carrier phrase. The formula used for this method compares the pitch of the investigated item with the pitch on the -li suffix in the carrier phrase, and follows the formula used by Mazaudon and Michaud (2006) in their study of Tamang tones:

\[ F_{REL} = 12 \times \frac{\log(F_{TARGET} / F_{FRAME})}{\log(2)} \]

where

- \( F_{REL} \) is the pitch of the investigated item relative to the carrier phrase, measured in semi-tones;
- \( F_{TARGET} \) is the pitch of the investigated item, measured in Hz; and
- \( F_{FRAME} \) is the pitch on the -li suffix in the carrier phrase, measured in Hz.

However, the semi-tone analysis and the absolute Hz values turned out to yield almost identical results. Given space constraints, we only present results from the semi-tone analysis for the male speaker’s productions of bare stems and -ji suffixed stems.

Finally, the software program Praat (version 5.3.52) (Boersma and Weenink 2013) was used to generate spectrograms for the purpose of identifying word-final glottalization associated with the M tone.

3.2. Results - bare stems

For both the female speaker and the male speaker, the production of the M tone is accompanied by word-final glottalization, as can be seen in Figure 1. Such word-final glottalization is not present in the production of the L tone or the H tone.

---

10 Consequently, comparisons in duration between mid and high tone were not made for this study, since vowel length in sonorant-final syllables was typically shorter than in open syllables. Future work will need to compare tone duration for each syllable type separately.
The results of the acoustic study of the female speaker’s production of tones are presented in Figure 2.

![Figure 2: Time-normalized pitch trajectory for L, M, and H tones on vowel portion of monosyllabic bare stems (female speaker)](image)

As can be seen, prior to the onset of word-final glottalization, the M and H tones do not differ significantly in terms of pitch.

Table 3: Mean F₀ realizations of tone at vowel mid-point of monosyllabic bare stems (female speaker)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Mean (Hz)</th>
<th>Standard deviation</th>
<th>No. of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>262.09</td>
<td>21.84</td>
<td>51</td>
</tr>
<tr>
<td>M</td>
<td>262.70</td>
<td>25.67</td>
<td>75</td>
</tr>
<tr>
<td>L</td>
<td>205.16</td>
<td>14.70</td>
<td>32</td>
</tr>
</tbody>
</table>

Table 3 gives the mean F₀ value of each tone at the vowel mid-point. A one-way ANOVA shows that tone category does have a significant effect on pitch, F(2,155), p < 0.001. However, post-hoc Tukey’s HSD tests show that only L tone pitch values are significantly lower than those of M and H tones. The phonetic realization of M and H tones is not significantly different in pitch at the vowel mid-point.

For the male speaker, M tones in word-final position are also accompanied by word-final glottalization, as shown in Figure 3.
1. Acoustic-statistical and perceptual investigations of Karbi tones

Figure 3: Acoustic waveform and spectrogram of bī ‘be small’ showing glottalization in word-final position (male speaker)

However, unlike the female speaker, the male speaker produced all three tones at significantly different pitch heights, as seen in Figure 4.

Figure 4: Time-normalized pitch trajectory for L, M and H tones on monosyllabic bare stems (male speaker)

Table 4 gives the mean F0 value of each tone at the vowel mid-point. A one-way ANOVA shows that tone category does have a significant effect on pitch, F(2,161), p < 0.001. A post-hoc Tukey’s HSD test shows that the phonetic realizations of all three tones are significantly different in pitch from each other at the vowel mid-point.  

---

11 A one-way ANOVA also shows that tone category has a significant effect on pitch at 10% of the vowel segment: F(2,161), p < 0.001; and at 90% of the vowel segment: F(2,161), p < 0.001.
Table 4: Mean $F_0$ realizations of tone at vowel mid-point of monosyllabic bare stems (male speaker)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Mean (Hz)</th>
<th>Standard deviation</th>
<th>No. of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>171.19</td>
<td>16.46</td>
<td>57</td>
</tr>
<tr>
<td>M</td>
<td>148.05</td>
<td>16.19</td>
<td>78</td>
</tr>
<tr>
<td>L</td>
<td>124.04</td>
<td>9.96</td>
<td>29</td>
</tr>
</tbody>
</table>

Figure 5 shows the male speaker’s production of pitch relative to the carrier phrase, measured in semi-tones. Here we can see that the male speaker did produce the H tone items at a higher pitch than the M tone items, even relative to the carrier phrase – he was not producing the entire utterance at a high pitch, or in a higher key.

![Figure 5: Time-normalized pitch trajectory, relative to carrier phrase, for L, M and H tones on monosyllabic bare stems (male speaker)](image)

Table 5 gives the mean pitch of each tone at the vowel mid-point, relative to the carrier phrase. A one-way ANOVA shows that tone category does have a significant effect on pitch, $F(2,161), p < 0.001$. A post-hoc Tukey’s HSD test shows that the phonetic realizations of all three tones are significantly different in pitch from each other at the vowel mid-point.\(^{12}\)

Table 5: Mean $F_0$ realizations of tone at vowel mid-point of monosyllabic bare stems (male speaker)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Mean (semi-tones)</th>
<th>Standard deviation</th>
<th>No. of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>6.37</td>
<td>1.04</td>
<td>57</td>
</tr>
<tr>
<td>M</td>
<td>4.87</td>
<td>1.49</td>
<td>78</td>
</tr>
<tr>
<td>L</td>
<td>1.65</td>
<td>1.54</td>
<td>29</td>
</tr>
</tbody>
</table>

3.3. Discussion – bare stems

For both speakers, word-final glottalization is generally produced in conjunction with the M tone. This confirms Grüßner's (1978) original analysis that the M tone occurs with glottal stops in word-final position.

\(^{12}\) A one-way ANOVA also shows that tone category has a significant effect on pitch at 10% of the vowel segment: $F(2,161)$, $p < 0.001$; and at 90% of the vowel segment: $F(2,161)$, $p < 0.001$. 
However, let us return to the crucial part of our first research question: Do Karbi speakers also distinguish the M and H tones by producing a difference in pitch height? It can be clearly seen that the female speaker does not distinguish M and H tone in terms of pitch height, which supports the findings of Konnerth’s preliminary (2010) study, as well as our own previous auditory impressions that M and H tone are often not distinguished by a difference in pitch height.

In contrast, the male speaker does produce a difference in pitch height to distinguish M and H tone, both in terms of absolute F0, and also relative to the carrier phrase. This is surprising because it contradicts findings from previous recordings of the same speaker (used by Konnerth (2010)), where he did not distinguish the two by pitch height.

Moreover, if we compare the spectrograms of the male and female speakers’ productions of bī ‘be small’ in a carrier phrase, the relatively darker vowel formants for bī produced by the female speaker (compared to the rest of the phrase) suggest that she is accenting the investigated word more than the male speaker. This makes the finding that she does not distinguish M and H tone by pitch height even more surprising, given that from a typological perspective, accented syllables are more likely to be the sites of tonal contrasts, e.g. in Scandinavian languages (Riad 2006).

After seeing these results, we asked the female speaker to concentrate on the minimal triplet thī ‘snatch’, thī ‘be short’, and thī ‘die’. We recorded her say the triplet once in isolation and then three times in a carrier phrase (following the same procedure outlined above), starting with the high tone item, then the mid tone item, and then the low tone, all in immediate succession. The result was surprising. Figure 6 below was created by extracting the second production of each item in its carrier phrase and combining them into a single graph. The pitch trajectories clearly show that even the female speaker, when asked to exaggerate the differences between items of a minimal triplet, makes very salient pitch distinctions. Just looking at the vowel midpoint in these particular productions, the high tone is at 461 Hz, the mid tone at 311 Hz, and the low tone at 185 Hz.

---

13 Thanks to an anonymous reviewer for drawing our attention to this.
14 Note that we originally recorded the stimuli by tone category and not by minimal set.
15 There are other phonetic peculiarities about the pitch trajectories in Figure 6, including the sharp fall in pitch of the high tone item. What we are concerned with here, however, are the exaggerated pitch height differences.
Figure 6: Exaggerated differences in minimal triplet showing pitch differentiations (female speaker)

It therefore appears that our initial results still remain inconclusive. However, findings from our study of pitch on stems with suffixes will serve to clarify some of our ongoing questions.

3.4. Results – stems with suffixes

As mentioned earlier, stems with suffixes were chosen because the glottalization associated with M tones only occurs in word-final position. Indeed, both speakers produced stem M tones without any glottalization if the stem was followed by a suffix.

In our presentation of the results, we first deal with the suffix -pò ‘irrealis1’ and start with the female speaker’s production of tones. Figure 7 shows each averaged pitch trajectory across the vowel segment on the stem, while Figure 8 shows the average pitch trajectory across the suffix -pò. Tables 6 and 7 give the mean pitch value at the mid-point of the vowel of the stem and of the suffix respectively.
1. Acoustic-statistical and perceptual investigations of Karbi tones

Table 6: Mean $F_0$ realizations of tone at vowel mid-point of stem before -pò suffix (female speaker)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Mean (Hz)</th>
<th>Standard deviation</th>
<th>No. of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>256.68</td>
<td>34.40</td>
<td>49</td>
</tr>
<tr>
<td>M</td>
<td>265.99</td>
<td>42.06</td>
<td>72</td>
</tr>
<tr>
<td>L</td>
<td>204.19</td>
<td>18.63</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 7: Mean $F_0$ realizations of tone at vowel mid-point of -pò suffix (female speaker)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Mean (Hz)</th>
<th>Standard deviation</th>
<th>No. of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>after H</td>
<td>243.74</td>
<td>17.15</td>
<td>49</td>
</tr>
<tr>
<td>after M</td>
<td>233.86</td>
<td>22.68</td>
<td>72</td>
</tr>
<tr>
<td>after L</td>
<td>184.73</td>
<td>14.12</td>
<td>30</td>
</tr>
</tbody>
</table>

A one-way ANOVA shows that tone category does have a significant effect on pitch on the stem, $F(2,148)$, $p < 0.001$. However, post-hoc Tukey’s HSD tests show that only the pitch values of L tones are significantly lower than those of M and H tones. The phonetic realization of M and H tones is not significantly different in pitch at the vowel mid-point.

A one-way ANOVA shows that tone category of the preceding stem does have a significant effect on pitch on the suffix -pò, $F(2,148)$, $p < 0.001$. However, post-hoc Tukey’s HSD tests show that the pitch at the midpoint of the suffix is only significantly lower after L tones. The pitch on the suffix after M tones is not significantly different from the pitch on the suffix after H tones.

Now we turn to the surprisingly different results of the male speaker’s production of tones. Figure 9 shows each averaged pitch trajectory across the vowel segment on the stem, while Figure 10 shows the average pitch trajectory across the suffix -pò. Tables 8 and 9 give the mean pitch value at the mid-point of the vowel of the stem and of the suffix respectively.
Figure 9: Time-normalized pitch trajectory for L, M and H tones on vowel segment of stem before -pò suffix (male speaker)

Figure 10: Time-normalized pitch trajectory for L, M and H tones on vowel segment of -pò suffix (male speaker)

Table 8: Mean F₀ realizations of tone at vowel mid-point stem before -pò suffix (male speaker)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Mean (Hz)</th>
<th>Standard deviation</th>
<th>No. of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>170.63</td>
<td>18.56</td>
<td>48</td>
</tr>
<tr>
<td>M</td>
<td>147.90</td>
<td>18.71</td>
<td>71</td>
</tr>
<tr>
<td>L</td>
<td>126.89</td>
<td>9.72</td>
<td>25</td>
</tr>
</tbody>
</table>

Table 9: Mean F₀ realizations of tone at vowel mid-point of -pò suffix (male speaker)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Mean (Hz)</th>
<th>Standard deviation</th>
<th>No. of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>after H</td>
<td>148.82</td>
<td>17.00</td>
<td>48</td>
</tr>
<tr>
<td>after M</td>
<td>136.77</td>
<td>18.61</td>
<td>71</td>
</tr>
<tr>
<td>after L</td>
<td>113.09</td>
<td>10.75</td>
<td>25</td>
</tr>
</tbody>
</table>

A one-way ANOVA shows that tone category does have a significant effect on pitch, F(2,141), p < 0.001. A post-hoc Tukey’s HSD test shows that the phonetic realizations of all three tones are significantly different in pitch from each other in the case of the male speaker.

A one-way ANOVA shows that tone category of the preceding tone does have a significant effect on pitch on the suffix -pò, F(2,141), p < 0.001. However, post-hoc Tukey’s HSD tests show that the pitch at the midpoint of the suffix is only significantly lower after L tones. The pitch on the suffix after M tones is not significantly different from the pitch on the suffix after H tones.

Besides -pò, the other suffix that was examined is the suffix -jì ‘irrealis2’. We again start with the results of the female speaker’s production of tones. Figure 11 shows each averaged pitch trajectory across the vowel segment on the stem, while Figure 12 shows the average pitch trajectory across the suffix -jì. Tables 10 and 11 give the mean pitch value at the mid-point of the vowel of the stem and of the suffix respectively.
1. Acoustic-statistical and perceptual investigations of Karbi tones

Figure 11: Time-normalized pitch trajectory for L, M and H tones on vowel segment on stem before -jì suffix (female speaker)

Figure 12: Time-normalized pitch trajectory for L, M and H tones on vowel segment of -jì suffix (female speaker)

Table 10: Mean F₀ realizations of tone at vowel mid-point of stem before -jì (female speaker)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Mean (Hz)</th>
<th>Standard deviation</th>
<th>No. of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>235.44</td>
<td>31.18</td>
<td>48</td>
</tr>
<tr>
<td>M</td>
<td>226.75</td>
<td>20.47</td>
<td>73</td>
</tr>
<tr>
<td>L</td>
<td>189.42</td>
<td>8.66</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 11: Mean F₀ realizations of tone at vowel mid-point of -jì suffix (female speaker)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Mean (Hz)</th>
<th>Standard deviation</th>
<th>No. of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>after H</td>
<td>265.78</td>
<td>25.77</td>
<td>48</td>
</tr>
<tr>
<td>after M</td>
<td>279.27</td>
<td>30.03</td>
<td>73</td>
</tr>
<tr>
<td>after L</td>
<td>268.60</td>
<td>21.63</td>
<td>30</td>
</tr>
</tbody>
</table>

A one-way ANOVA shows that tone category does have a significant effect on pitch, F(2,148), p < 0.001. However, post-hoc Tukey’s HSD tests show only L tone pitch values are significantly lower than those of M and H tones. The F₀ realizations of M and H tones are not significantly different at the vowel mid-point.

A one-way ANOVA shows that the tone category of the preceding syllable does not have a significant effect on pitch on the suffix -jì, F(2,148), p = 0.02003.

We now turn to the results of the male speaker’s production of tones, which are again surprisingly different. Figure 13 shows each averaged pitch trajectory across the vowel segment on the stem, while Figure 14 shows the average pitch trajectory across the suffix -jì. Tables 12 and 13 give the mean pitch value at the mid-point of the vowel of the stem and of the suffix respectively.
A one-way ANOVA shows that tone category does have a significant effect on the pitch of the stem, \( F(2,149), p < 0.001 \). A post-hoc Tukey’s HSD test shows that the F\(_0\) realizations of all three tones are significantly different from each other.

In contrast to the female speaker, a one-way ANOVA shows that tone category of the preceding tone also has a significant effect on pitch on the suffix -\( ji\), \( F(2,149), p < 0.001 \). However, post-hoc Tukey’s HSD tests show that the pitch at the midpoint of the suffix is only significantly higher after H tones. The pitch on the suffix after M tones is not significantly different from the pitch on the suffix after L tones.

The semi-tone analysis of the male speaker’s production of pitch in -\( ji\) suffixed items yielded the following results. The tone of the stem is presented in Figure 15, while the tone of the suffix (still relative to the preceding syllable -\( li\) in the carrier phrase) is presented in Figure 16. The mean pitch values relative to the carrier phrase at the midpoint of the vowel are presented in Tables 14 and 15: for the stem before the suffix and for the suffix, respectively.
1. Acoustic-statistical and perceptual investigations of Karbi tones

Figure 15: Time-normalized pitch trajectory, relative to carrier phrase, for L, M and H tones on vowel segment of stem before -jí suffix (male speaker)

Figure 16: Time-normalized pitch trajectory, relative to carrier phrase, for L, M and H tones on vowel segment of -jí suffix (male speaker)

Table 14: Mean pitch, relative to carrier phrase, at vowel mid-point of stem before -jí suffix (male speaker)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Mean (Hz)</th>
<th>Standard deviation</th>
<th>No. of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>5.53</td>
<td>1.38</td>
<td>51</td>
</tr>
<tr>
<td>M</td>
<td>3.22</td>
<td>1.58</td>
<td>75</td>
</tr>
<tr>
<td>L</td>
<td>0.25</td>
<td>0.91</td>
<td>26</td>
</tr>
</tbody>
</table>

Table 15: Mean pitch, relative to carrier phrase, at vowel mid-point of -jí suffix (male speaker)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Mean (Hz)</th>
<th>Standard deviation</th>
<th>No. of tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>after H</td>
<td>6.39</td>
<td>1.12</td>
<td>51</td>
</tr>
<tr>
<td>after M</td>
<td>5.40</td>
<td>1.40</td>
<td>75</td>
</tr>
<tr>
<td>after L</td>
<td>5.41</td>
<td>1.38</td>
<td>26</td>
</tr>
</tbody>
</table>
A one-way ANOVA shows that tone category does have a significant effect on pitch on the stem, relative to the carrier phrase, $F(2, 149), p < 0.001$. A post-hoc Tukey’s HSD test shows that all three tones are significantly different from each other. A one-way ANOVA further shows that the tone category of the preceding tone also has a significant effect on pitch on the suffix -jì, relative to the carrier phrase, $F(2, 149), p < 0.001$, with the tone on -jì significantly higher in pitch after H tones, compared to both L and M tones. The pitch on the suffix (relative to the carrier phrase) after M tones is not significantly different from the pitch on the suffix after L tones.

### 3.5. Discussion – stems with suffixes

Looking first at the pitch on the stem before the two suffixes, the results for the female speaker remain consistent with those of her production of bare stems: M and H are not distinguished by a difference in pitch height. The male speaker, in contrast, consistently produced H tones at a higher pitch than M tones on the stem.

If we consider the pitch over the suffix -pò, we find that both speakers produced -pò at a lower pitch after L tones, compared to after M and H tones. However, there was found to be no contrast in pitch height on -pò after M and H tones for both speakers.

On the other hand, when we look at the suffix -jì, we find that the female speaker produced -jì at the same pitch height after all tones. More importantly, the male speaker produced -jì after H tones at a significantly higher pitch than after both L and M tones. This suggests that for these H tone stems, the male speaker was producing the whole word at a higher pitch, even relative to the carrier phrase. Given that cross-linguistically, the perception of tone relies on relative (not absolute) differences in pitch, this raises the question of whether the phonetic difference in pitch height that the male speaker has been producing is actually phonologically contrastive.

These results also suggest that the female speaker was simply neutralizing the contrast between M and H stem tones in non-word-final position. However, other acoustic factors that we had not investigated in this study, such as length and loudness, may also be important phonetic cues in tone identification. It is for these reasons that it was decided to conduct a tone perception experiment using these recordings.

### 4. Perception study

The surprising result of the acoustic study that shows the female speaker consistently merging the pitch height of M and H tones in the cases of a) bare stems, and b) -pò or -jì suffixed stems, while the male speaker almost equally consistently produced the H tone at a higher pitch level led us to conduct a follow-up perception study. The perception study seeks to answer the following questions:

- Can listeners correctly identify M and H tone stems in the case of a) bare stems (given the presence of M tone glottalization), compared to the case of b) -pò or -jì suffixed stems (given the lack of M tone glottalization)?
- Does the pitch difference between M and H tones in the productions of the male speaker help listeners identify the intended tone category?
1. Acoustic-statistical and perceptual investigations of Karbi tones

4.1. Methodology

For the perception study we selected six native Karbi speakers to listen to a set of either the female speaker's or the male speaker's stimuli; both sets contained the same stimuli in the same order. The male speaker also acted as a listener (with the listener code L1) and was the only listener who listened to both sets. The female speaker did not participate in the perception study. Among the remaining five listeners (i.e., L2-L6) there were four listeners aged between 47 to 53 years and one younger listener aged 29. Two different listeners listened to the female speaker (FS), three different listeners listened to the male speaker (MS), and the male speaker listened to both his own and the female speaker's productions.

We asked listeners to put on Sony Professional MDR-7506 headphones and helped adjust the volume so they could comfortably listen to the stimuli. We then had them listen to a total of 152 stimuli, which consisted of a randomly ordered set of 76 x 2 different stimuli: 20 bare stem types, 20 -ji suffixed stems, 20 -pò suffixed stems, as well as 16 stems with two other suffixes, which are excluded from the current study. We used the same 20 stems across the three categories of bare stems, -ji suffixed stems, and -pò suffixed stems. The 20 stems consist of ten minimal pairs (see Table 16). The order was once manually randomized so that the same stimulus would not occur twice in a row. All listeners - whether listening to the female speaker's or the male speaker's productions - listened to the stimuli in the same order.

Table 16: Tone minimal pairs used as stimuli as either bare stem, -ji suffixed stem, or -pò suffixed stem for perception study

<table>
<thead>
<tr>
<th>Lexeme</th>
<th>Tone</th>
<th>Gloss</th>
<th>Lexeme</th>
<th>Tone</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 bī</td>
<td>M</td>
<td>'be small'</td>
<td>11 bī</td>
<td>H</td>
<td>'keep'</td>
</tr>
<tr>
<td>2 bū</td>
<td>M</td>
<td>'carry child'</td>
<td>12 bū</td>
<td>H</td>
<td>'overflow'</td>
</tr>
<tr>
<td>3 jāŋ</td>
<td>M</td>
<td>'contain'</td>
<td>13 jāŋ</td>
<td>H</td>
<td>'fall'</td>
</tr>
<tr>
<td>4 kōm</td>
<td>M</td>
<td>'surround'</td>
<td>14 kōm</td>
<td>H</td>
<td>'fall'</td>
</tr>
<tr>
<td>5 mān</td>
<td>M</td>
<td>'feel burning sensation'</td>
<td>15 mān</td>
<td>H</td>
<td>'become'</td>
</tr>
<tr>
<td>6 nōng</td>
<td>M</td>
<td>'cultivate'</td>
<td>16 nōng</td>
<td>H</td>
<td>'fame (noun)'</td>
</tr>
<tr>
<td>7 rō</td>
<td>M</td>
<td>'hit'</td>
<td>17 rō</td>
<td>H</td>
<td>'praise'</td>
</tr>
<tr>
<td>8 sāng</td>
<td>M</td>
<td>'raw rice (noun)'</td>
<td>18 sāng</td>
<td>H</td>
<td>'take rest'</td>
</tr>
<tr>
<td>9 thī</td>
<td>M</td>
<td>'be short'</td>
<td>19 thī</td>
<td>H</td>
<td>'snatch'</td>
</tr>
<tr>
<td>10 thū</td>
<td>M</td>
<td>'be rotten'</td>
<td>20 thū</td>
<td>H</td>
<td>'cut'</td>
</tr>
</tbody>
</table>

Listeners listened to the stimuli as produced by the speakers, i.e., as one isolated repetition and three repetitions in carrier phrases. Usually listening to the stimulus (i.e., four repetitions) once was enough for the listener to decide which lexeme they thought it was, although we gave them the option to listen to the stimulus as many times as they wanted. We recorded the answer on a sheet of paper where all 152 stimuli were represented with the written Karbi (without any indication of tone as is the typical way of writing Karbi) together with the possible meanings (as English glosses) depending on the tone.

The 'answer sheet' also included the option for listeners to say they were unsure what the intended meaning was. Figure 17 provides a sample displaying thīji with three gloss options as there are three segmentally identical stems 'thī' forming a tone minimal triplet. The "Unsure"
option was only chosen four times out of a total of 7 response sets times 152 responses, although we encouraged listeners to make use of that option if they indeed were unsure.

| 1 | thiji | O Unsure. | O snatch | O short | O die |

Figure 17: Sample item "thiji" from stimulus response sheet

Those listeners who were fluent readers of English simply marked the answers themselves, whereas those who weren't were asked to tell us the intended meaning in Karbi by adding uniquely identifying affixes or perform uniquely identifying gestures (for example, the meanings for 'snatch' and 'short' from Figure 17 could easily be identified by gesturing).

4.2. Results

Figure 18 shows the error rates of listeners when they misidentified the intended meaning of the female speaker's or the male speaker's productions, respectively. They indicate what percentage of the overall number of stimuli was misidentified. Generally we assume that if the error rate is at about 50% or higher, the listener was not able to identify the stimulus, and was simply making a guess.

Looking at the individual bars, what is most striking is the large amount of individual differences between listeners. There are only two generalizations we want to suggest. First, the white bars representing error rates in identifying the intended meaning of bare stem productions do not reach the 50% mark once, although L5 and L6 listening to the male speaker's (MS) stimuli show surprisingly high error rates. That is, bare stems have consistently lower error rates than suffixed stems, which is true for every single speaker. Second, the gray and the black bars representing error rates in identifying the -jì and -pò suffixed stems are very high, i.e., mostly at around 50% or higher.
1. Acoustic-statistical and perceptual investigations of Karbi tones

Figure 18: Error rates in responses by listeners L1, L2, and L3 to the female speaker's (FS) stimuli on the left and to the male speaker's (MS) stimuli by listeners L1, L4, L5, and L6 on the right (including the male speaker himself as L1).

Figure 19 offers error rate averages across those who listened to the female speaker's (FS) productions and those who listened to the male speaker's (MS) productions, respectively. The error rate average of those listening to the MS excludes the error rates of the MS listening to his own productions, which are instead listed separately next to it.

Figure 19: Error rates in responses by all listeners to female speaker's (FS) productions and male speaker's (MS) productions, excluding the error rates in responses by MS listening to his own productions.
We see that for bare stems, the MS listeners have clearly higher error rates, while in the case of the suffixed stems, both FS and MS listeners have error rates above 50%. Furthermore, we may point out again that bare stems have a consistently lower error rate than suffixed stem, for both FS and MS listeners. Also, even while listening to his own productions, the MS shows a considerable percentage of misidentifications.

4.3. Discussion

We may now go back to the two questions we set out to answer with this perception study. The first question asked whether listeners could distinguish between M and H tones on stems depending on whether they were bare or suffixed stems. Recall that in bare stems, the M tone features a characteristic glottalization, which disappears when suffixes are added. The perception study showed that although listeners could identify the difference between M and H tones on bare stems to some degree (i.e., 100% for Listener 1, 80% for Listener 2, and 60% for Listener 3), there was still some considerable error involved. Compared to the suffixed stems, however, the error rates for bare stems were much lower. We propose that the distinctive glottalization - as the putatively major phonetic cue in bare stems - does help listeners, since error rates go up when the glottalization disappears due to added suffixes. With such suffixed stems, the contrast appears to be essentially neutralized, with listeners simply guessing what the lexeme is, in the absence of a semantic or pragmatic context.

The second question asked whether the differences in pitch height between the M and H tones that the male speaker consistently produced would help listeners identify the intended lexeme. In other words, the second question asked whether those listeners who listened to the male speaker's productions would have lower error rates than those listeners who listened to the female speaker's productions. The answer is clearly no: the male speaker's differentiation of pitch height between M and H tones appeared to confuse listeners even more (and even he has still considerable error rates listening to his own productions).

The question, of course, is why the male speaker produced a statistically significant pitch height difference between the M and the H tones at all. Here we see a clear mismatch between tone production and tone perception, i.e., the speaker was articulating a difference in pitch that he himself was not using as a perceptual cue, at least for the suffixed forms (where his error rates are at 20% and 50%, see Figure 19). We assume that at least part of the answer here is that the male speaker was over-articulating the difference between the M and the H tone as a result of his awareness of the tone minimal sets. Although the order in which we recorded the stimuli was by tone category and not by minimal set, the male speaker was heavily involved in finding stimuli for this study and almost certainly was very aware of each of the minimal set and the respective M or H counterpart. Moreover, recall that the female speaker also distinguished M and H tones by pitch height when asked to contrast items of a minimal set, as shown in Figure 6 above.

5. Conclusion

The first study in this article discussed the F₀ realization of the three tone categories L, M, and H of Karbi. We gathered data from one male and one female speaker, which included productions of bare stems and stems either with the suffix -ji or the suffix -pò. We found that the female speaker consistently merged the pitch of M and H stem tones (across both bare and suffixed stems) while the male speaker almost equally consistently produced the H stem tone at a higher
1. Acoustic-statistical and perceptual investigations of Karbi tones

F₀ level. However, prior to the recording session with the female speaker and even afterwards, she also produced the H stem tone at a higher pitch level when asked to contrast M and H tones.

In order to investigate whether the male speaker's differentiation between M and H tones, i.e., the higher pitch on the H tone was phonologically contrastive, we conducted a perception study in which listeners were asked to identify the intended lexeme. As we expected, the M vs. H tone contrast was fairly recoverable in the case of bare stems, though not always perfectly. This was most probably due to the presence of glottalization in word-final M tones. However, listeners consistently misidentified the lexemes in suffixed stems, which interestingly held for the male speaker's stimuli as much as for the female speaker's stimuli despite the male speaker's F₀ differentiation of M and H tones. Furthermore, the male speaker misidentified his own productions. This suggests that the male speaker's production of higher pitch on H tones is not phonemic. As a result, we may say that M and H tones are not differentiated by F₀ in Karbi.

However, this result cannot be conclusive: Why does Grüßner (1978) describe three pitch levels? Why do native speakers seem to generally produce three pitch levels when contrasting minimal triplets, and produce H tones at a higher pitch level than M tones (according to auditory observations)? Why did the male speaker produce three pitch levels? Despite the results of our two studies, there seems to exist some level of representation at which M and H tones are in fact differentiated by F₀. That level of representation was (presumably somewhat consciously) accessed by the male speaker in our data and is accessed by other native speakers when they contrast minimal pairs. Ongoing research into the F₀ realization of M vs. H tones in stems followed by one of two suffixes other than -jí and -pò suggests that in those cases that same level of representation surfaces and M and H tones are distinguished by pitch. As a result of these considerations, we may want to say that M and H tones are differentiated by pitch height but this contrast is neutralized in bare stems and those suffixed by -pò and those suffixed by -jí.

However, even this is too simplistic a conclusion. The contrast is not entirely neutralized, since the male speaker did produce H tones at a significantly higher pitch height than M tones, and the female speaker likewise produced H tones at a higher pitch height than M tones when contrasting minimal pairs. Therefore, we suggest to consider these data a case of incomplete neutralization (see Labov et al. (1972: Chapter 6); Yu (2007, 2011: §4), among others), although a somewhat unusual case. Incomplete neutralization typically refers to instances of "small but consistent phonetic differences in supposedly neutralized environments" (Yu 2007: 188). For the data presented here, the phonetic difference is the M vs. H tone F₀ difference, and the supposedly neutralized environments are -jí suffixed stems, and -pò suffixed stems, since M and H stem tones are misidentified by listeners in those cases. However, the data presented here represent an unusual case, since what we find in our studies is that only one of our two speakers produces this consistent phonetic difference. Furthermore, he does not produce a "small" pitch height difference between M and H tones, but the F₀ difference between M and H tones at the vowel midpoint is in fact as large or slightly larger than the F₀ difference between L and M tones.

What is certainly even more puzzling are the large individual differences both among speakers and among listeners (keeping in mind, of course, the overall small number of participants for this study). This fact calls for a sociolinguistic study to investigate potentially ongoing sound change, and to record a larger number of speakers.¹⁶

¹⁶ Considering the age difference between the female and the male speaker, a sociolinguistic investigation of the tone system would be worthwhile, as pointed out by an anonymous reviewer. Note that the male speaker, who produced the three-way pitch contrast was 52 years old, and the female speaker, who did not, was younger, at 36 years old. Perhaps the three-way pitch contrast is original and currently being lost.
Finally, recall that one characteristic of the Karbi tone system that is also found in other languages of the area is the overall low functional load of tones. In Karbi, this is evidenced by native speakers’ difficulty to assign tone categories to words especially of more than one syllable, as well as the fact that there are barely any tone minimal triplets, even less so tone minimal triplets within the same word class (cf. Konnerth 2014). While we believe these factors play an important role in accounting for the present data, they certainly call for further research into this peculiarly unstable looking tone system.
1. Acoustic-statistical and perceptual investigations of Karbi tones

References


2. Syllable Restrictions in Chungli

T. Temsunungsang
The English and Foreign Languages University Shillong

Abstract
In this paper, the syllable structure of Chungli, a Tibeto-Burman language spoken in Nagaland, India is examined by providing a detailed description of the restrictions that are in force. What is attested in Chungli is markedly different from Coda Condition in that the nucleus and the coda consonant are in a harmonic relation, making it a syllable internal restriction as attested in some Chinese dialects (Duanmu 2003/2007) and Vietnamese dialects (Pham 2006). On the other hand, there also seems to be a relationship between the onset and the nucleus. This paper claims that rather than the Coda Condition, a condition on the harmonic relation within the rime (Rime harmony) is best suited for Chungli. It is concluded that Rime harmony, licensing of onsets by the nucleus and OCP play an important role in determining the Syllable restrictions in Chungli.

Citation

1. Introduction

Since Kahn (1976), the syllable has come a long way in being established as an important prosodic unit in phonological theory. The evidence for the syllable as a prosodic unit comes from various phonological processes and phonotactic restrictions (Ito 1986; prosodic licensing, locality and directionality. However, cf. Blevins (2002) for a different perspective.). In Chungli, the evidence for postulating the syllable as the lowest unit in the hierarchy comes from phonotactic restrictions, segment harmony, licensing requirements and certain OCP related constraints within the syllable. Considering the numerous arguments for the syllable, it would not be surprising if the syllable is considered to be an important unit in forming prosodic words.

One major observation on syllables in the Chungli dialect of Ao is the restriction on coda consonants. There have been numerous studies on the type of consonants that can occupy the syllable final position (coda) in languages, with restrictions on place features playing a central role.

A number of constraints have been postulated in phonological theory over the years, dealing with such coda restrictions. The most basic is the syllabic well-formedness principle No-Coda which disallows codas in a syllable.

(1) No-Coda  Codas are disallowed

In addition, different forms of the Coda-Condition (Ito 1986, McCarthy and Prince 1993a,b; Ito and Mester 1994), states that codas are not allowed unless linked to a following onset.

---

1 This paper is a reworked version of chapter two of my thesis and was presented at NEILS6. The chapter was written in the framework of Optimality Theory but in this paper, the theoretic formalisations have been omitted. Thanks to K.G. Vijayakrishnan, three external examiners (anonymous) and a NEILS reviewer who made insightful comments and suggestions. Also my gratitude to the participants at NEILS6 for their valuable contributions.
(2) **Coda-Cond**  
Codas are not allowed unless linked to a following onset.  
A coda consonant is a nasal homorganic to following stop or affricate.  
A coda consonant can have only Coronal place or place shared with another consonant. (Prince & Smolensky 1993)

In the above conditions, a following onset is central to most analyses. Such examples are also noted in Chungli where the following consonant influences the preceding coda. However, such processes are not very productive in the language.

(3) \textit{sun} + \textit{apak} \Rightarrow \textit{sumpak} \Rightarrow \textit{summak}  
money flat ‘currency’

In (3), the nasal assimilates to a bilabial which further leads to the complete assimilation of the following stop to a bilabial. In compounding, the truncation of the initial vowel (or even a prefix) is a common phenomenon.

On the contrary, what is attested in Chungli is markedly different in that the vowel nucleus and the coda consonant are in a harmonic relation, making it a syllable internal restriction as attested in some Chinese dialects (Duanmu 2003/2007) and Vietnamese dialects (Pham 2006). On the other hand, there seems to be a relationship between the onset and the nucleus, i.e., only a small subset of onsets is allowed when the nucleus is a high back spread vowel /\textit{ɯ}/.

In this paper, we examine the facts of syllable structure in Chungli, providing a detailed description of the restrictions that are in force, which we believe are a result of Rime harmony, Onset licensing and Obligatory Contour Principle (OCP).

2. **Inventory and syllable structure**

The phonemic inventory for Chungli is shown in Table 1 where the allophonic variants are in square brackets (Temsunungsang 2007, 2010).

<table>
<thead>
<tr>
<th>Consonants</th>
<th>Vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>p t k ? m n nɡ</td>
<td>i uʊ u ʃ z ʃ (h)ə</td>
</tr>
<tr>
<td>s [ʃ] z (h)²</td>
<td>[ə]</td>
</tr>
<tr>
<td>ʃ [ts] ʃ [ts]</td>
<td>a</td>
</tr>
<tr>
<td>r l w j</td>
<td></td>
</tr>
</tbody>
</table>

² While /h/ is not part of the standard inventory, it is attested in many of the Chungli varieties.
2. Syllable Restrictions in Chungli

Table 2: Onset and coda consonants

<table>
<thead>
<tr>
<th>C₁ consonants</th>
<th>C₂ consonants</th>
</tr>
</thead>
<tbody>
<tr>
<td>p, t, k</td>
<td>p, t, k, ŋ</td>
</tr>
<tr>
<td>m, n, ŋ</td>
<td>m, n, ŋ</td>
</tr>
<tr>
<td>s, z</td>
<td></td>
</tr>
<tr>
<td>ŋ</td>
<td></td>
</tr>
<tr>
<td>l, r, w, j</td>
<td>r</td>
</tr>
</tbody>
</table>

A maximal syllable template would consist of \((C₁)V₁T₁(V₂T₂)\) where the nucleus with a tone \(T₁\) is obligatory. \(C₁\) allows all the consonants except the glottal stop while \(C₂\) allows only a set of consonants which include stops, nasals and a flap as in Table 2.

Chungli shows a three way tonal contrast and bear a tone on every syllable, either lexically or by default. The three contrastive tones attested are High, Mid and Low. In citation forms, the realisation of the three tones depends on the nature of the syllable and hence we make a distinction between checked (where the coda is a stop) and smooth syllables (where the coda is a sonorant or is null). While all the three tones can be realised on smooth syllables, only the High and Low tones are attested in checked syllables. A three way contrast of tones is hence observed only in smooth syllables.

Table 3: Smooth syllables

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[áʧə́m]</td>
<td>‘fear’</td>
</tr>
<tr>
<td>[áʧə̄m]</td>
<td>‘drink’</td>
</tr>
<tr>
<td>[áʧə̀m]</td>
<td>‘decrease’</td>
</tr>
</tbody>
</table>

A single syllable can also have a maximum of two tones. However, this is treated as the concatenation of two level tones rather than as a contour tone since they do not show any contrast. The tones are realised on the vowels or the coda if it is a sonorant.

Table 4: Concatenation of tones

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>áâp</td>
<td>‘rot’</td>
</tr>
<tr>
<td>míŋ</td>
<td>‘accompany’</td>
</tr>
<tr>
<td>pʊ̜r</td>
<td>‘snake’</td>
</tr>
</tbody>
</table>

3. Monosyllabic words

We begin by looking at monosyllabic words in order to arrive at certain generalisations with regard to the distribution of segments within the syllable. We examine the three possible syllable structures \(C₁V\), \(VC₂\) and \(C₁VC₂\) where \(V\) can be \(V:\). The question of whether the nucleus is \(V\), \(V:\) or a diphthong has no significant bearing on our analysis and hence these have been clubbed together. It must be noted that Chungli has very few monosyllabic words and hence many of the unattested forms may perhaps be accidental gaps. By analysing the above mentioned syllable
structures, we attempt to look at the relation, if any, between the onset-nucleus, nucleus-coda and the onset-coda.

3.1. Open monosyllabic words C₁V

In this syllable pattern, any consonant apart from /ʔ, ɲ, z, w/ can occur in C₁ with any of the vowels V=/a, i, u, u/.  

<table>
<thead>
<tr>
<th>Table 5: C₁V words</th>
</tr>
</thead>
<tbody>
<tr>
<td>mì</td>
</tr>
<tr>
<td>tú</td>
</tr>
<tr>
<td>lā</td>
</tr>
<tr>
<td>nū</td>
</tr>
</tbody>
</table>

The unattested segments /ņ, z, w/, apart from the glottal stop, are however possible C₁ consonants in disyllables as will be seen below in § 4.

Another constraint on C₁V monosyllabic words is the ban on words of the form [*Cə]. We observe that while the high back spread vowel /ɯ/ occurs with the consonants [n, s, r, ts] as seen in Table 6, it does not occur with the other consonants.

<table>
<thead>
<tr>
<th>Table 6: Occurence of /ɯ/</th>
</tr>
</thead>
<tbody>
<tr>
<td>nūu</td>
</tr>
<tr>
<td>stū</td>
</tr>
<tr>
<td>ruū</td>
</tr>
<tr>
<td>tsūū</td>
</tr>
</tbody>
</table>

The reason is that with other consonants, /u/ is reduced to a schwa as in (4).

(4)  tuū.kā =⇒  [tū.kā] ‘hand’

Hence, while the /u/ reduction is attested in disyllabic forms, there are no instances of such reduction in monosyllabic words resulting in the form [*Cə].

3.2. Onsetless monosyllabic words VC₂

In VC₂, a sequence of high back spread /u/ and any of the consonants (*uC) is not attested while /a/ seems to be least marked in that it allows all the possible consonants except /t/.

<table>
<thead>
<tr>
<th>Table 7: VC₂ words</th>
</tr>
</thead>
<tbody>
<tr>
<td>āp</td>
</tr>
<tr>
<td>ān</td>
</tr>
<tr>
<td>ām</td>
</tr>
<tr>
<td>ār</td>
</tr>
<tr>
<td>át</td>
</tr>
</tbody>
</table>
/u/ does not allow the labial and flap consonants (*u+labial, *u+flap) while /i/ allows only /t, p, n, m/ along with the glottal stop.

### Table 8: Occurrence of /i/

<table>
<thead>
<tr>
<th>Syllable</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>āīt</td>
<td>‘come in’</td>
</tr>
<tr>
<td>āip</td>
<td>‘slash’</td>
</tr>
<tr>
<td>āṃ (jān)</td>
<td>‘iron’</td>
</tr>
<tr>
<td>āiʔ</td>
<td>‘thorn’</td>
</tr>
<tr>
<td>ām (jām)</td>
<td>‘village’</td>
</tr>
</tbody>
</table>

The absence of *u+labial could be treated as a case of OCP but as will be seen in § 3.3 and § 5, we shall consider this to be a violation of rime harmony.

### 3.3. Closed monosyllabic words C₁VC₂

In this syllable structure, a regular occurrence of similar segments within the rime is attested. Hence, while /i/ and [ə] occurs with /m, t, n, p, r/, the back vowels /a, u/ and /u/ occur with the consonants /k, ŋ, ʔ, r/. The consonant /r/ appears to occur with all vowels. However, we do not go into the details of this behaviour. The exceptions are hun ‘gold’, san ‘sheep’, lun ‘chutney’ and tiʔ ‘sister’ which do not follow the above pattern.

### Table 9: Onset/Coda Distribution

<table>
<thead>
<tr>
<th>C₁</th>
<th>V</th>
<th>C₂</th>
<th>C₁VC₂</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>t, k, m, s, [ʧ], l, j</td>
<td>a</td>
<td>/k, ŋ, ʔ, r/</td>
<td>[t aŋ]</td>
<td>‘rice’</td>
</tr>
<tr>
<td>t, k, m, s, [ʧ], r</td>
<td>i</td>
<td>/m, t, n, p, r/</td>
<td>mim</td>
<td>‘love’</td>
</tr>
<tr>
<td>p, n, s, [ʧ], l, r, j</td>
<td>u</td>
<td>/k, ŋ, ʔ, r/</td>
<td>nuk</td>
<td>‘knife’</td>
</tr>
<tr>
<td>k, n, ŋ, s, [ʧ], r, l</td>
<td>[ə]</td>
<td>/m, t, n, p, r/</td>
<td>[kən]</td>
<td>‘song’</td>
</tr>
<tr>
<td>s, [ts]</td>
<td>[u]</td>
<td>/k, ŋ, ʔ, r/</td>
<td>[tsuk]</td>
<td>‘grain’</td>
</tr>
</tbody>
</table>

As far as onsets are concerned in C₁VC₂, we do not find any uniformity in their distribution with regard to the nucleus V. If we combine and compare C₁ segments in both C₁VC₂ and C₁V monosyllabic words, only the consonants /z, w/ are unattested anywhere in Chungli monosyllabic words which may be considered to be accidental gaps.

Some of the observations made above needs to be examined in greater detail. This can be done by looking at disyllabic words which we take up in the next section.
4. Disyllabic words

As was observed in the monosyllabic words, similar phonotactic restrictions are observed in both the syllables of disyllabic words. Under a general observation, for disyllabic words:

- All consonants can occur in the onset position except the glottal stop.
- /i/ and [ə] occur with /m, t, n, p, r/, while the back vowels /a, u/ and /u/ occur with /k, η, ? , r/ which is similar to monosyllabic words.
- The stop series, nasals /m, η/, lateral and glides do not occur as onsets when the nucleus is a high back spread vowel /u/. This is in line with the restricted occurrence of /u/ observed in monosyllabic words.
- The distribution of /z/ is very limited in that it occurs in the onset only with the high back spread vowel /u/.

Apart from these, there does not seem to be other patterns of interest in disyllabic words. Hence, we do not pursue this any further. Our focus will be on the possible coda consonants and their distribution with regard to the nucleus.

4.1. Rime patterns in disyllabic words

In the coda position, eight consonants are licensed. The stop series /p, t, k, ?/, the nasal series /m, n, η/ and the flap /r/. Below, we examine the rime sequences of the first and second syllable in disyllables and monosyllables. Both the first and second syllable will be analysed together since there is no significant difference between the two syllables as far as co-occurrence restrictions are concerned.

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>t</th>
<th>k</th>
<th>?</th>
<th>m</th>
<th>n</th>
<th>η</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>i</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>u</td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>e</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>u</td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>

It can be seen from Table 10 that there are gaps in the distribution. We argue, in § 5, that these gaps are a result of restrictions between the coda and the nucleus and not accidental gaps. Below, we lay out the restrictions that are in place with respect to the nucleus and the coda.

1. If the vowel is /u/ or /a/, the possible consonants in the coda are /η, k, ?/ or null.
2. The vowel /i/ behaves differently from the rest of the vowels. While it occurs frequently with the glottal stop or in an open syllable, it also occurs with the other consonants: nasals /m, n, η/ (mûtîn ‘devil’), stops (except /k/) and flap.
3. The high back spread vowel /u/ occurs with all coda consonants. However, it undergoes reduction to a schwa when followed by /p, t, m, n/. Hence [u] occurs with /k, η, ?/ while [ə] occurs with /p, t, m, n/, again satisfying feature harmony within the Rime.
2. Syllable Restrictions in Chungli

4. In addition, the onset is always [ts], /z/, /s/, /r/ or /n/ when the vowel is /ɯ/. It is interesting to see that the onset and the coda have an asymmetric distribution on the lines of [± coronal] (with the exception of /t/).

It is apparent that the above points are not very different from those made for monosyllabic words at the end of § 3.3 and only reinforce the earlier observations. Below, we shall examine a few of the above observations vis-a-vis,

- Occurrence of /i/ and [ə] with /p, m, t, n/ (§ 5)
- Occurrence of /u/, [ɯ] and /a/ with /k, η, ʔ/ (§ 5)
- The choice of a small subset of onsets by the high back spread vowel /ɯ/ (§ 6.1)
- The OCP constraint on labials and glides (§ 8)
- The status of the glottal stop (§ 9)
- Some exceptions in monosyllabic words (§ 10)

5. The Rime sequence

In this section, we examine the distribution of the coda consonants /p, m, t, n, k, η, ʔ/ in both mono and disyllabic words. The coda consonants can be classified into four places of articulation on the surface: labial, coronal, velar and glottal. The coda /r/ seems to be the only consonant which occurs with all vowels. For example, [puɾ] ‘snake’, [marla] ‘mosquito’, [puɾ] ‘forest rat’, [tir] ‘head of an organisation’. The occurrence of /r/ with both [ə] and [ɯ] is however, never contrastive.

Following the Articulator Model of the feature tree, Chungli appears to make a distinction of consonants in terms of the feature [± back], where /p, m, t, n/ are considered [-back] and /k, η/ are considered [+back]. If this distinction of [± back] holds for consonants, there appears to be a harmonic relation within the rime as attested in Tables 9 and 11. Hence, while /i/ and [ə] (the assumption is that [ə] is [-back] in Chungli) occur with /p, m, t, n/, the back vowels /a/, /u/ and

---

3 This occurrence may be a result of /i/ not being specified for [Dorsal] and hence, the [±back] requirement (explained in this section) becomes redundant. However, this is not pursued further in this paper.
[ɯ] occur with /k, ŋ/. Looking at the examples above, it is evident that the nucleus and coda must match in some feature ([±back] here) to satisfy Rime harmony or the coda must be empty or must be filled by a default place for /Ɂ/. The glottal stop will be considered to be unspecified for place in §9.

Note that this condition on Rime harmony applies to all words. However, for onsetless monosyllabic words (in §3.2) a few words in Table 7 do not follow this harmonic pattern. It should be noted that the exceptions to Rime harmony (e.g., ām ‘biscuit’, ān ‘hen’) attested in monosyllabic words are never found in disyllabic forms. Hence, in comparing the analysis of monosyllabic and disyllabic words, we find that while syllables in disyllabic words follow the feature sharing between nucleus and coda in a strict manner (for /a, u, u/), monosyllables do not. In other words, a number of exceptions are found in monosyllabic words with respect to rime harmony. Such words are considered to be contracted words as shown in §10 and hence do not follow Rime harmony.

A direct consequence of Rime harmony is the allophonic variation between [ɯ] and [ə]. We assume /ɯ/ to be the phoneme, while [ɯ] and [ə] are the allophonic variants. Hence, the underlying forms /atsuk/ ‘roast’ and /amun/ ‘ripe’ surface as [atsuk] and [amun] respectively. The assumption here is that schwa is [-back] in Chungli. Phonetically, [ə] in Chungli is lower and more fronted than either [ɯ] or [u]. While vowel reduction in many languages occur as a consequence of being in unstressed positions, in the case of Chungli, it results due to the following [-back] consonant and hence, the postulation that the schwa loses its feature.

6. The onset-rime relation

Having looked at the nucleus-coda sequences, we now turn to the onset-rime sequences which attest certain restrictions. There are 22 possible rime sequences including V and VC which should ideally combine with the 13 possible onset consonants, resulting in 286 possibilities as onset. As is evident from Tables 12 and 13, this is not true.

The onset-rime combinations are shown in Tables 12 and 13 where the allophones [ɯ], [ə] and [ts], [ʧ] are shown in separate rows and columns for more clarity. While the possible sequences are ticked, the shaded boxes imply that they can be explained as violations of Rime harmony or, as will be seen in the later sections, OCP and Onset licensing. The blank boxes may be treated as accidental gaps.

In Table 12, most of the unattested onset-rime combinations are explainable with only a few accidental gaps. The only onset which seems to have a number of gaps is /z/ which occurs only with /ɯu/. However, we do not have any evidence to explain this restricted occurrence.

---

4 Whether the schwa as [-back] can be grouped together with the other [-back] vowel /i/ needs further examination which we do not take up in this paper.
### Table 12: The onset-rime sequence in Chungli (with [+back] vowels)

<table>
<thead>
<tr>
<th></th>
<th>ak</th>
<th>aŋ</th>
<th>uk</th>
<th>uŋ</th>
<th>ukŋ</th>
<th>aʔ</th>
<th>uʔ</th>
<th>uʔ</th>
<th>a</th>
<th>u</th>
<th>uŋ</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>t</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>k</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>m</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>n</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ŋ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ts</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ʃ</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>s/f</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>z</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>l</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>r</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>w</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>j</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

The rime sequences with /i/ appear to have the maximum number of gaps in distribution Table 13. In fact, apart from /i/ and /iʔ/ which allows almost all onsets, other possible onset-rime sequences involving /i/ are negligible, occurring only with the stops, nasal /m/ and the fricative /s/. In most cases, the occurrence of /i/ with /p, t, m, n/ is a result of word contraction. On the other hand, the occurrence of /s/ with /i/ is a case of allophonic variation, changing to [ʃ].

Hence, /i/ may be considered to occur only in open syllables and with glottal stop codas. Other gaps can be explained by phonotactic restrictions, OCP and allophonic distribution while some must be considered to be accidental gaps.
6.1. Onset Licensing

In Table 12, a consistent co-occurrence is seen with respect to \([ɯ]\) and the coronal consonants \([s, z, r, n, t, ts]\), exemplified in Table 14.

### Table 14: Co-occurrence of \([ɯ]\) and coronal consonants

<table>
<thead>
<tr>
<th></th>
<th>áṣùk</th>
<th>‘draw’</th>
<th>áżùŋ</th>
<th>‘read’</th>
</tr>
</thead>
<tbody>
<tr>
<td>rʊ</td>
<td>‘ice’</td>
<td>ántùŋ</td>
<td>‘sky’</td>
<td></td>
</tr>
<tr>
<td>tʊzʊ</td>
<td>‘vein’</td>
<td>[tsʊk]</td>
<td>‘grain’</td>
<td></td>
</tr>
</tbody>
</table>

The example \(tʊzʊ\) ‘vein’ is perhaps a result of vowel copy since it is not common for /ɪ/ and /ʊ/ to co-occur. However, for the rest of the coronals, the co-occurrence with /ʊ/ is widely attested. In addition to the coronal feature of the above sounds, they are all [-distributed]. This distinction is necessary considering the absence of /ʧ/ from the set of coronal sounds which is [+distributed]. Hence, the group of sounds which can occur with /ʊ/ must be [+coronal, -distributed].

The suggestion here is that, there is a relation of dependence between the nucleus and the coda on one hand; and between the onset and the rime on the other. While we have treated the
nucleus-coda relation as one of harmony, the same cannot be said of the relation in the onset-rime; there is nothing harmonic between the sound segments [ɯ] and [ts] or [ə] and [ʃ]. Hence, it can only be a case of licensing.

Table 15: Occurrence of coronals and vowels

| ánàk | ‘scratch’ |
| ãsū | ‘cook’ |
| àʧī | ‘see’ |
| [ásɔp] | ‘off’ |

However, the question arises as to which element licenses what; whether it is the nucleus which license onsets or vice-versa. We conclude that it is the nucleus which licenses the onset based on two observations.

Firstly, it must be noted that the above mentioned coronals can also occur with other vowels but [ɯ] cannot occur with other consonants Table 15.

Secondly, consider the allophonic distribution of [ʃ] and [ts]. While [ts] occurs with [ɯ], [c] occurs elsewhere.

(5) [á.ʦùŋ] ‘fill’
    [á.ʧəp] ‘cry’

The implication here is that the nucleus has some property that allows it to choose the onset element (even though both are coronals).

(6) a.ʧɯp => a.ʧəp ‘cry’
    a.ʧɯŋ => a.tsɯŋ ‘fill’

Hence, from the above, it is apparent that there is some condition which allows /ɯ/ to license only coronals (with the feature [-distributed]) in the onset position.

In standard theory, the generally held view was that only codas need licensing. However, this is far from true as many languages need onset licensing just like the codas. In Chungli, the glottal stop is never licensed in the onset position. In addition to this ban on glottal stops in the onset position, we observe that only coronals (barring [+distributed]) are licensed in the onset by the vowel /ɯ/. This phenomenon must be captured by a constraint on licensing rather than a co-occurrence or harmonic constraint, given that there is nothing in common between these sounds.

While we have seen a case of the onset not licensing the glottal stop and the nucleus (/ɯ/) licensing specific onsets, the question remains as to whether onset licensing precedes or follows rime harmony. Our hypothesis is that Rime harmony takes place before licensing of the onset. The data in (6) suggests that this is what happens exactly. Note that /ɯ/ is reduced to a schwa as a result of the following front consonant. This results in the licensing of [ʃ] and not [ts] in the onset. If onset licensing precedes Rime harmony, we would expect, [aʧuŋ] => *[atsuŋ] => *[atsəp] which is unattested.

The story so far is that while the nucleus licenses the coda, leading to a harmonic relation in the rime, the resultant rime/nucleus licenses certain onset consonants.
7. Vowel reduction

So far, what we have seen is the case of /ɯ/ reduction to schwa when the following consonant (coda) is [-back] /p, t, m, n/. However, this is not the only environment where vowel reduction takes place. If the onset is [-coronal], vowel reduction is again attested since sequences of [-coronal] and /ɯ/ are banned in the language.

(7) \( pɯ.lũm \Rightarrow pə̄.lũm \) ‘tire’

In the above data, the first schwa is a case of vowel reduction due to *[-coronal][±back, + high, -round], a ban on the sequence [-coronal][ɯ], while the second schwa is reduction due to the following front consonant /m/.

Interestingly, as mentioned earlier in § 3.1, the sequence *Cə, where C=[± coronal], is not attested word finally. If C=[+coronal], its absence can be attributed to the constraint just mentioned above. However, what disallows *Cə when C=[-coronal], a sequence well attested word internally? We assume a morphological well-formedness requirement in word final position or right edge of the word, (*əɷ) disallowing reduced vowels word finally.

While it is difficult to find data to exemplify this constraint, considering that *Cə is totally banned, the lone instance of alternation in the 1st person possessive pronoun /kɯ/ illustrates the working of this requirement. The sequence [kɯ] must be banned by virtue of the onset being [-coronal] and the nucleus /ɯ/. However, what we observe in native speakers are two possible surface forms, [kə] and [kɯ], based on its domain of occurrence. If it occurs as a word on its own, [kɯ] surfaces in order to satisfy the well-formedness condition *əɷ. If the occurrence is within a larger word, [kə] surfaces as a result of /ɯ/ reduction, since the schwa is no longer at the word final position and consequently satisfies the coronal constraint.

(8) \([kɯ]_o[mɪjuŋ]_o \sim [kə-mɪjuŋ]_o\) ‘mine’

Note that these two forms are in free variation and the criteria that distinguishes the two seem to be a slight pause between [kɯ] and [mɪjuŋ]; and a slight prominence on [kɯ]_o.

8. OCP: Onset glides and labials

It is assumed in phonological theory that sound segments have some feature organization, allowing processes to target specific tiers/nodes/planes. One of the oft cited definitions is that of McCarthy (1986).

(9) OCP At the melodic level, adjacent identical elements are prohibited.

The OCP, first discussed by Leben (1973) and further taken up by Goldsmith (1976), was initially based on the nature of tonal representations and how tones were represented. It basically disallowed similar tones to co-occur in sequence. Two important notions need to be addressed in the context of OCP: adjacency and local domain. The local domain could be considered as language specific, attested in a stem sized domain for Tashlihiyt Berber (Alderete 1997 and references therein) where a stem may have, at most, one labial consonant /b, f, m/. If more than one is attested, the labial delabializes to a coronal.
2. Syllable Restrictions in Chungli

(10)  m-xazar ‘scowl’    n-fara ‘disentangle’
      (where m- is the reflexive prefix)

The domain could also be a morpheme as in Japanese (Lyman’s law, Ito and Mester 1986, 1998) where multiple obstruent voicing is not allowed. E.g.: *buda

In the Semitic languages, the domain of OCP application is the root (McCarthy 1981, 1986) and hence forms like sasam are not attested where the root is ssm⁵. This is attributed to an OCP constraint within Arabic roots where consonants with identical place features do not co-occur.

With regard to adjacency, we follow Rose’s (2000) stipulation for guttural OCP in Tigre:

(11)  Two consonants in sequence are adjacent irrespective of intervening vowels.

Without committing to the nature of representation, we look at OCP constraints below. We first look at the labial OCP restrictions.

Table 16: OCP restrictions in Chungli

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>pap</td>
<td>pəp</td>
<td>pəp</td>
<td>pip</td>
<td>wəp</td>
<td></td>
</tr>
<tr>
<td>pam</td>
<td>pəm</td>
<td>pəm</td>
<td>pim</td>
<td>wəm</td>
<td></td>
</tr>
<tr>
<td>map</td>
<td>məp</td>
<td>məp</td>
<td>mip</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>mam</td>
<td>məm</td>
<td>mum</td>
<td>mim</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Apart from the two forms in bold, none of the forms in Table 16 are attested in Chungli. While column I and III can be explained as violations of Rime Harmony, column II, IV and V must be permissible. This seems to be a clear case of OCP where labials are not allowed in adjacent positions.

The glides in onsets are also restricted in their distribution. While /w/ never occurs with /u/, /j/ does not occur with /i/⁶. (also see Table 8)

(12)  (a)  *wu
      (b)  *ji

The two conditions required to account for the OCP restrictions are:

  OCP-Labial: Prohibit adjacent labials
  OCP-Glide: The onset and nucleus cannot have the same specification for [±round, + high]

While OCP-Labial accounts for the forms in Table 16, OCP-Glide accounts for the restriction involving glides as in (12). But what allows the forms in bold in Table 16? They are in fact disyllabic forms undergoing which undergo contraction to form monosyllabic words. Hence, the OCP holds in the non-contracted forms and underlying monosyllables. A more detailed examination of exceptions is taken up in § 10.

⁵ *sm being a possible root form for sasam is rejected since spreading is rightward in Arabic. The root sm results in samam.

⁶ One occurrence is that of ji ‘wine’ which also has an alternate form ii.
9. The glottal stop

As has been noted in § 5, the glottal stop occurs with all vowels in the syllable. In many of the North East Indian languages like Mizo (Fanai 1992), Garo (Burling 1992), and many of the Naga languages, the distribution of the glottal stop is limited to the coda and, more often, to the word final position.

In Chungli, its distribution is confined to the word final position in non-derived words and the glottal stop deletes when morpheme concatenation takes place. In Waromung Mongsen, Coupe (2003 : 27) analyses the glottal stop as “a non-predictable word prosody that simultaneously functions as a contrastive unit at the segmental tier when it is realised word-finally; it is otherwise subject to a deletion rule in the environment of a word internal syllable boundary”. We list out three of his observations on Waromung Mongsen.

Firstly, its realisation is not predictable and can contrast with other phonemes as well as zero.

Secondly, he mentions an alternative analysis by Okell (1969) for Burmese, treating the glottal stop as an exponent of a particular tone. He claims that this cannot be true for Mongsen as the glottal stop occurs with all three tones. Examples in Table 18 are from Coupe (2003 : 24).

| Table 17: Contrast of the glottal stop |
| sàk | ‘scratch’ | [sɔp] | ‘off’ |
| sā | ‘say’ | sū | ‘break’ |
| sāʔ | ‘measure’ | stūʔ | ‘rot’ |

| Table 18: Occurrence of glottal stop with all tones |
| /jāʔ/ | ‘beat.PAST’ | /jāʔ/ | ‘hear.PAST’ |
| /ā-ṁʔ/ | ‘NPF-spear’ | /ā-(miʔ) | ‘NPF-person’ |
| /tāʔ/ | ‘intestine’ | /tāʔ/ | ‘height’ |

Thirdly, the glottal stop, in contrast to oral stops, undergoes deletion when followed by a suffix (Coupe 2003 : 25).

| (14) | /ŋūʔ/ ‘bite’ + /-āʔ/ | IMPER => | /ŋūʔ/ |
| /jāʔ/ ‘hear’ + /-ūkū/ | ANT => | /jāukū/ |

The suggestion here is that the glottal stop has a dual role; its primary function being boundary marking and its contrastive function being “a corollary of the word final realisation”.

Yip (1994) argues that the glottal [c.g] and nasalization are features of the morpheme and does not belong to a segment alone in Chaoyang where the morpheme is mostly co-extensive with the syllable. Hence, syllables of the structure (C)(G)V(G)(ʔ) are treated as open syllables while (C)(G)V(C) are closed syllables where the possible coda consonants are [m, ŋ, p, k]. The argument is that codas consist only of place features and surface as voiceless glottalised stops if the syllable is [c.g], or as nasals if the syllable is not [c.g].
2. Syllable Restrictions in Chungli

In Garo, Burling (1992) notes that the glottal stop behaves like a tone rather than a consonant. The arguments put forward by Burling are summarized as given in (15) by Duanmu (1994).

(15) (a) [ʔ] occurs only in the coda but not in the onset.
(b) [CVʔ] syllables are similar to [CV] instead of [CVC].
(c) [ʔ] is the only coda that is transparent to vowel assimilation.
(d) [ʔ] is subject to deletion, but other C codas are not.
(e) [ʔ] is not restricted to the coda position, but may occur on other segments.

Contrary to Burling, Duanmu argues, with special emphasis on 15(b), (c) and (e), that the glottal stop in Garo is segmental.

9.1. Feature or segment?

The question remains as to whether the glottal stop should be treated as a segment or some suprasegmental feature in Chungli. We will argue that the glottal stop functions based on its dual property: it behaves as a segment as well as a supralaryngeal feature which we discuss below.

Its property as a stop segment comes based on its behaviour in not allowing mid tones in checked syllables. Additionally, just like the back consonants, it occurs with [ɯ] but never [ə]. Hence, we will postulate that the glottal stop must be treated as a stop segment.

In addition to its status as a segment, the glottal stop has a second property, evidenced on preceding vowels, i.e., the vowel is always accompanied by irregular vibration patterns similar to creakiness. Note that creaky voice is not phonemic and arises when a glottal is present. This property is clearly attested in the deletion of the glottal stop word internally in morpheme concatenation.

(16) (a) cáʔ-úkū ‘eat-ASP’
(b) cá-úkū ‘call-ASP’

In (16)(a), the glottal stop deletes but what remains is the irregular vibrating pattern on the vowel [ə], suggesting that though the segment is deleted, its feature [c.g] remains on the vowel. This irregular vibrating pattern is not attested in (16)(b) where no glottal stop is present. This clearly shows that the glottal stop can also behave as a supralaryngeal feature. While its function is not clear, we hypothesise that this is a result of satisfying non-deletion of consonant segments in Chungli which is never attested elsewhere. In other words, while the glottal stop is deleted, its feature is retained on the vowel.

So what allows the deletion of a glottal stop word internally, unlike the other stop consonants? Firstly, in (16), resyllabification results in the glottal stop being in the onset position. Since the glottal stop is not allowed as onsets, it is deleted. Secondly, we assume that the glottal stop is peculiar from other stops as it has no specification for place underlyingly and gets its specification to satisfy a requirement which demands that all codas be specified for place.

(17) CodaPlace Codas must be specified for place

---

7 The glottal stop can occur on nasals or a lateral in a CVʔN structure. The glottal and N is simultaneous rather than in sequence. It can also occur as a glottalized vowel (Burling 1992).
The reason for the choice of a glottal stop to fill this placeless position perhaps comes from the fact that the glottal stop is the least marked of the consonants and hence can occur with all vowels irrespective of [±back]. Such an analysis, where the glottal stop is the least marked in the place markedness hierarchy, is proposed by Lombardi (2001a,b, 2002), with the assumption that the glottal has a pharangeal place (McCarthy 1994).

\[(18) \quad *\text{Dor}, *\text{Lab} \gg *\text{Cor} \gg *\text{Phar}\]

Due to its status as the least marked, it is a common epenthetic consonant, preferred over another less marked segment; the coronal.

In our analysis, we assume that due to its least marked status, it is the segment that is inserted in placeless codas. Hence,

\[(19) \quad \text{Underlying} \quad \text{Surface} \]
\[
\begin{array}{lll}
\text{aʧu}C & \Rightarrow & \text{atsu}? \\
\text{aʧi}C & \Rightarrow & \text{aʧi}? \\
\end{array}
\Rightarrow
\begin{array}{ll}
\text{atsu}? & \Rightarrow \text{atsu}? \\
\text{aʧi}? & \Rightarrow \text{aʧi}? \\
\end{array}
\quad \text{‘pull’ ‘eat’}
\]

Additional evidence about the peculiarity of the glottal stop comes from other kinds of deletion attested cross dialectically. Some of the Chungli varieties attest a corresponding glottal stop for the velar /k/.

Table 19: Cross dialectal deletion

<table>
<thead>
<tr>
<th>Std Chungli</th>
<th>Mongsen</th>
<th>Ungma variety</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>ámàk</td>
<td>màk</td>
<td>ámàʔ</td>
<td>‘attack’</td>
</tr>
<tr>
<td>ákùk</td>
<td>kùk</td>
<td>ákóʔ</td>
<td>‘win’</td>
</tr>
<tr>
<td>ámāŋ</td>
<td>māŋ</td>
<td>ámāʔ</td>
<td>‘believe’</td>
</tr>
<tr>
<td>áʧtúp</td>
<td>tʃúp</td>
<td>áʧtúp</td>
<td>‘cry’</td>
</tr>
</tbody>
</table>

In the Ungma variety of Chungli, while most velar stops of standard Chungli have a corresponding glottal stop, the velar nasal has no corresponding form except for the nasalization on the vowel. Other stops remain unchanged. This is of particular interest in two ways. Firstly, it points to the fact that the glottal stop may in part function as some sort of closure to mark word boundaries as suggested by Coupe (2003 : 26). Note that while the glottal takes the coda position, the nasal having no such function is realised on the vowel resulting in an open syllable.

Secondly, this is a case of loss of place in one variety, where the loss leads to insertion of the glottal stop; equivalent to a placeless coda being filled by the glottal stop as in the standard dialect. Similar cases of cross dialectal variation are also attested in Mongsen.

\[(20) \quad átīk \quad átīʔ \quad \text{‘release’}\]

This deletion (cross dialectically) is not attested for the labial and alveolar stops to our knowledge. Hence, the treatment of the glottal stop as a segment as well as a suprasegmental feature seems to be a favourable one considering the evidence presented. If the glottal stop is realised word finally, it surfaces to function as a closure to word boundary. If it occurs word
2. Syllable Restrictions in Chungli

internally, this function no longer holds and hence can be deleted. However, due to a strong requirement for non-deletion of consonants, although the segment is deleted, its feature remains thus satisfying the non-deletion of the glottal stop. Deletion of other stops is not attested, perhaps for the reason that its feature cannot be realised on the preceding vowel. If attested, it would amount to postulating alveolarized and labialized vowels.

10. Return to Exceptions

In the course of this paper, we have come across a number o
exceptions with regard to Rime harmony and OCP most notably in monosyllables. The question is whether these should be treated as exceptions or whether there can be a principled explanation for such occurrences. Given that most words, barring three, are all monosyllables, can there be such a case whereby the constraints are inactive in monosyllables? This would seem too far-fetched given the high number of monosyllables attested in Chungli.

On a closer look at the monosyllabic exceptions, it is apparent that these are reduced or contracted forms and hence need not be treated as exceptions. For instance, forms like mim and miip were considered to violate OCP-Labial and were treated as exceptions.

(21) (a) [míjɔ́m] => míń ‘love’
(b) [míjɔ́p] => miıp ‘‘turn’’

In both the cases above, the deletion of the glide results in monosyllabic forms. In (21)(b), there is an additional lengthening of the vowel in order to accommodate the tone which cannot be deleted.

To account for the contracted monosyllabic words, we claim a distinction between derived and underived forms where only underived forms adhere to the OCP constraints and do not affect derived forms. Hence, mim and miip being contracted forms will not violate OCP-Labial. This line of argument will be followed in dealing with the other exceptions that we have seen so far.

The other set of problematic words are the nouns below which do not follow the regular phonotactic restrictions.

Table 20: Exceptions

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>hùn</td>
<td>‘gold’</td>
<td>sàn</td>
</tr>
<tr>
<td>ám</td>
<td>‘banana leaf’</td>
<td>ān</td>
</tr>
<tr>
<td>āút</td>
<td>‘camel’</td>
<td>tām</td>
</tr>
<tr>
<td>nāriŋ</td>
<td>‘orange’</td>
<td></td>
</tr>
</tbody>
</table>

The regular phonotactic restriction would require back vowels to occur with [+back] codas, which is not the case above. We will show that four are reduced forms, while three are borrowed words and hence are derived forms where phonotactic restrictions do not apply.

(22) āhm => [ām] => ām ‘banana leaf’
āhən => [ān] => ān ‘hen’
These assumptions come from cross dialectal data where Mangmetong Mongsen clearly attests a fricative as the onset of the second syllable. Hence, at one point of the derivation, these words do not violate phonotactic restrictions since the harmony between the nucleus and the coda is respected.

On the other hand, two other words hun ‘gold’ and uut are borrowed words from Hindi/Assamese.

(23) (a) \( \text{sun}a \Rightarrow \text{hû}nå \Rightarrow \text{hû}n ‘gold’ \)
(b) \( \text{u}t \Rightarrow \text{û}t ‘camel’ \)

The s > h is a common phenomenon in Chungli since most dialects attest /h/ instead of /s/. While deletion of final vowel is attested in (23)(a), forcing the tone onto the sonorant, cases like (23)(b) are common where, in a borrowed monosyllabic word, the vowel is always lengthened to accommodate two tones and phonotactic restrictions are seldom active in such loanwords.

(24) \( \text{ká}t ‘cut’ \)
\( \text{ká}p ‘cup’ \)

For såñ, based on cross dialectal variation in Chungli varieties, we postulate the base to be såjñ or såm which was probably incorporated into the Mongsen dialect through the translated bible. The last example nàrìñ ‘orange’ is a Mongsen word which has been borrowed into Chungli. However, two cases múcìñ ‘devil’ and ìmsù ‘taboo’ appear to be true exceptions.

One common characteristic of these forms is that they all violate certain restrictions. While múñ and múñp violate OCP-Labial, forms such as såñ violate phonotactic restrictions. However, based on our analysis, what brings them together as belonging to one class of words is that they are all reduced forms. Such reduced forms are in a sense derived words. We thus make a generalization that phonotactic and OCP restrictions do not hold in derived words. Loanwords are derived in the sense that tones need to be minimally binary on the surface.

We had postulated that phonotactic restrictions are observed in all cases, derived and underived. Considering that all underived words, suffixes and prefixes adhere to the phonotactic restrictions, it is not possible for us to check whether this is true in the case of derived words formed by affixation. Thus, the only evidence for this comes from reduction of word structure and loanwords.

11. Conclusion

In this paper, we started by examining how restrictions in the coda have been dealt with by invoking the Coda Condition in many languages. We claim that rather than the Coda Condition, a condition on the harmonic relation within the rime is best suited for Chungli. We examine the six points observed in Chungli words:

---

8 In fact, the form hûnå is also seen in some of the other Chungli varieties.
9 A reviewer suggests an internal stratification of the lexicon as attested in Japanese (Ito and Mester 1999). While this is possible, space constraints do not allow us to pursue this line of analysis in this paper.
2. Syllable Restrictions in Chungli

- Occurrence of /i/ and [ə] with /p, m, t, n/ (§ 5)
- Occurrence of /u/, [u] and /a/ with /k, ɲ, Ɂ/ (§ 5)
- The choice of a small subset of onsets by the high back spread vowel /u/ (§ 6.1)
- The OCP constraint on labials and glides (§ 8)
- The status of the glottal stop (§ 9)
- Some exceptions in monosyllabic words (§ 10)

We conclude that Rime harmony, licensing of onsets by the nucleus and OCP plays an important role in determining the Syllable restrictions in Chungli. On the other hand, we claim that the glottal stop is a result of default insertion in a placeless position to fulfil CodaPlace and argue that the glottal stop can be treated as a segment as well as a feature based on the phenomenon where a glottalized vowel remains even on deletion of the glottal segment.

We also re-examined the exceptions which were pointed out in the course of the paper. We argue that apart from two cases mucin ‘devil’ and amsur ‘taboo’, all the so-called exceptions can be explained as cases of contraction or word reduction. Hence, we conclude that since such forms are derived, phonotactic and OCP restrictions do not hold in such words including loanwords.
References


2. Syllable Restrictions in Chungli


_____. (2010). *Aspects of the Prosodic Phonology of Ao: An Inter-dialectal study*. PhD Dissertation. The English and Foreign Languages University, Hyderabad.

Morphosyntax
3. Coordination in Manipuri

S. Indrakumar Singh and Ch. Yashawanta Singh
Department of Linguistics, Manipur University, Imphal.

Abstract
The present paper describes the syntax and semantics of coordination in Manipuri (Meiteilon). Regarding the presence and distribution of coordinators that identify coordinated elements, Manipuri has syndetic (including mono- and bisyndetic) and asyndetic coordination constructions. Furthermore, coordinators are discussed: conjunctive coordination ('and'), disjunctive coordination ('or') and adversative coordination ('but'). Positions of coordinators in both monosyndetic and bisyndetic coordinations are discussed. Lastly, semantic types of coordination are discussed.

Citation

Copyright © 2014, the author(s), release under Creative Commons Attribution license

1. Introduction

Manipuri is a verb-final language and it shares a number of characteristic features of SOV languages as proposed by Greenberg (1963): adpositions are all postpositional rather than prepositional, and the genitive precedes the governing noun. Specifically, Manipuri has extensive verb morphology, extensive suffixation with more limited prefixation.

Coordinate constructions are claimed to be a universal category in languages of the world (Zhang 2009). In Manipuri, lexical coordinators are used in monosyndetic coordination, while suffixal coordinators are used in bisyndetic coordination (Bhat & Ningomba 1997). Chelliah (1997:196ff.) discusses clausal coordination in Manipuri, and specifically illustrates that finite clauses can be related to each other through the use of coordinators which are based on əә 'that' and one of the oblique case markers. Each coordinator of the form determiner-case marker has a specialized function.

The outline of this paper is as follows: §2 discusses syndesis in this language. The next section §3 is devoted to the semantic types of coordination. §4 analyses the position of coordinators, which shows the conjunction strategies in both monosyndetic and bisyndetic coordination. The next section is devoted to the discussion of category-sensitivity. §6 of this paper discusses semantic distinctions. Finally, §7 summarizes the findings.

This paper offers a descriptive account of coordination in Manipuri. Data incorporated in this paper are collected by the authors (Manipuri native speakers) themselves from reliable and authoritative books, published or unpublished works, local newspapers, news broadcasted over television and radio, and various conferences held in the university or other academic places. On the basis of these data an attempt has been made to analyse Manipuri coordination.
2. Syndesis

If one or more coordinators occur in a coordinating construction, it is called syndetic. There are two types of syndetic constructions: monosyndetic and bisyndetic constructions (Haspelmath 2004).

2.1. Monosyndetic Coordination

Monosyndetic coordination involves only a single coordinator if two coordinands are present. In Manipuri, monosyndetic coordination occurs with lexical coordinators such as əməsəŋ/əmədi ‘and’, ədugu ‘and’, ədudıg ‘then’, əduhu ‘but’, əntrəgo ‘or’, ədumonəmək ‘however’. In addition, Manipuri allows multiple coordinands. Examples are given below. The single coordinator əməsəŋ/əmədi ‘and’ conjoins two conjuncts: magi jum ‘his house’ and magi inkəol ‘his land’ in (1) and lairik ‘book’ and kolom ‘pen’ in (2).

(1)  magi jum əmədi inkəol jonkʰre
    ma-gi jum əmədi inkəol jon-kʰ-re
    he-GEN house and land sell-DEF-PERF
    ‘He has sold his house and land.’

(2) punsi-də lairik əməsəŋ kolom cəŋ-i
    punsi-də lairik əməsəŋ kolom cəŋ-i
    life-DET book and pen need-ASP
    ‘Life needs book and pen (education).’

The coordinators əmədi and əməsəŋ ‘and’ are synonymous, so either one can be used without a difference in meaning. In (1), əmədi ‘and’ can be replaced by əməsəŋ showing no difference in meaning. Similarly in (2), əməsəŋ ‘and’ is replaceable by əmədi.

In Manipuri, monosyndetic constructions with multiple coordinands are possible to occur such that all but the last coordinator are omitted as in (3) and (4). This construction also exists in other Tibeto-Burman languages, such as Hakha Lai (Peterson and VanBik 2004).

(3) tombə, ram, məni, caobə əmədi mohon cətkʰre
    tombə, ram, məni, caobə əmədi mohon cət-kʰ-re
    tomba, ram, mani, chaoba and mohon go-DEF-PERF
    ‘Tomba, Ram, Mani, Chaoba and Mohon have gone’.

(4) əikʰoiəə məi, isiŋ, nuŋsi əməsəŋ jum dərkər oi
    əi-kʰoi-ŋə məi isiŋ nuŋsi əməsəŋ jum dərkər oi
    I-PL.-ERG fire water air and house need be
    ‘We need fire, water, air and house’.
2.2. Bisyndetic Coordination

Bisyndetic coordination involves as many coordinators as conjuncts. This type of coordination occurs with suffixed coordinators (SCs), such as -gə~kə, -ne, -nə~nə and -su. In (5), the two conjoining conjuncts are coordinated with the suffixed coordinator -gə.

(5) ənəŋə jahirə lake
ənən-gə jahir-gə lək-e
Anand-SC Jahir-SC come-ASP
‘Anand and Jahir have come.’

Similarly, -ne can be used, as in (6). Unlike monosyndetic coordination with multiple coordinands, it is not possible to delete all but the last suffixed coordinator, as shown in (7).

(6) pʰuritne kʰongrəone kʰongupne lairikne laire
pʰurit-ne kʰongrao-ne kʰongup-ne lairik-ne lai-re
shirt-SC trouser-sc book-sc buy-PERF
‘Shirt, trouser, shoes and book have been bought.’

(7) *pʰurit kʰongrao kʰongup lairikne laire
pʰurit kʰongrao kʰongup lairik-ne lai-re
shirt trouser shoe book-sc buy-PERF

2.3. Asyndetic Coordination

When coordinators are absent in a coordinating construction, it is said to be asyndetic coordination, i.e., the conjoining coordinands are concatenated with no grammatical morpheme between them. The concatenated structure can be seen in (8) and (9) below.

(8) magi jum iŋkʰol lai
ma-ɡi jum iŋkʰol lai-i
he-GEN house land have-ASP
‘He has his house and land.’

(9) tʰəwai həkcaŋ jaunə hotnəu
[tʰəwai həkcaŋ] jau-na hotnə-u
heart body with-ADV try-COMD
‘Try with heart and soul.’

The concatenated structure, as in (8) and (9), is used for familiar combinations which are most often experienced as a single conceptual unit. The single conceptual unit is iconically encoded with a more compact grammatical construction with no intervening coordination. In Manipuri, asyndetic coordination is limited to two coordinands and to nouns only.
The following sentence (10) is also acceptable without the overt coordinator.

(10) əigi ima ipa lakte
əi-ɡi i-ma i-pa lak-te-i
I-GEN 1PP-mother 1PP-father come-NEG-ASP
‘My parents did not come.’

In Manipuri, nouns in a coordinate construction, such as ima ‘my mother’ and ipa ‘my father’, are treated as a familiar and conscious combination, i.e., ima-ipa ‘parents’ and hence, a concatenate structure occurs and speakers of this language more economically use this structure (ima-ipa) than the coordinate structure (ima əməsun/əmədi ipa). Likewise, this kind of concatenation is also found in Karen, a Sino-Tibetan language (Carol & Benson 2004).

Similarly, kʰongraø ‘trouser’ and pʰurit ‘shirt’ in (11) are a familiar combination and a conceptual unit, and hence they can occur without the coordinator as kʰongraø pʰurit. Here the combination might be translated as ‘garment’ in English.

(11) məhak kʰongraø pʰurit setkʰi
məhak kʰongraø pʰurit set-kʰi
he trouser shirt wear-PERF
‘He has worn clothes.’

3. Semantic types

Coordinators are generally classified as conjunctive (‘and’), disjunctive (‘or’) or adversative (‘but’). In Manipuri, conjunctive includes coordinators such as əməsun/əmədi ‘and’, ədudəgi ‘after that’, ədugə ‘and’, as discussed further in §3.1. Disjunctive coordinators are discussed in §3.2 and include nattrəgo ‘or’, nattrədəbu ‘if it is not so’, nattrədədi ‘except/or’. Lastly, adversative coordinators, discussed in §3.3, include ədubu ‘but’, ədumoinəmək ‘however’, and təuigumbəsuŋ ‘but’.

3.1. Conjunctive Coordination

Conjunctive coordination in Manipuri occurs with lexical coordinators such as əməsun/əmədi ‘and’, ədudəgi ‘after that’, ədugə ‘and’. Examples are given below.

(12) ram əməsun sita cətkʰore
ram əməsun sita cət-kʰ-o-re
Ram and Sita have gone.’

(13) məhak cak cakʰore ədugə cətkʰore
məhak cak ca-kʰ-o-re ədugə cət-kʰ-o-re
he rice eat-DEF-PERF and go-DEF-PERF
‘He has eaten rice and has gone away.’
3. Coordination in Manipuri

The coordinator aməsuŋ ‘and’ in (12) shows simple conjoining of the two conjuncts Ram and Sita. In (13), verb phrase (VP) conjuncts caḵʰəre ‘has eaten’ and caolkʰəre ‘has gone’ are conjoined by ədugə ‘and’ indicating the additive with the implication that the second VP occurs after the first.

3.2. Disjunctive Coordination

Disjunctive coordination in Manipuri occurs with lexical coordinators such as nattrəga ‘or’, nattradəbu ‘if it is not so’, nattrəbədi ‘except/or’. Examples are (14) and (15).

(14) tombə nattrəga caoba soi-də-na lakkəni
    tombə nattrəga caoba soi-də-na lakkəni
    ‘Either Tomba or Chaoba will come without fail.’

(15) kʰamba nattrəbədi kəna aməttəna kao pʰəba nəm-loi
    kʰamba nattrəbədi kəna aməttə-na kao pʰə-bə nəm-loi
    khamba except who one-NEG-ADV cow catch-NZR be.able-FUT
    ‘No one except Khamba can catch the cow.’

In sentence (14), there is an alternative between the two coordinands Tomba or Chaoba to perform the action of coming. However, the disjunctive coordinator nattrəbədi in (15) describes that no one except Khamba can catch the cow.

In VP coordination, the final VP coordiand which follows the coordinator nattrəbədi is always negated. But the coordinator nattrəga occurs with both positive and negative elements. More examples can be examined as follows.

(16) mahak lakkəni nattrəga lak-loi pau pirəkəni
    mahak lak-kəni nattrəga lak-loi pau pi-rək-kəni
    he come-FUT or come-NEG news give-INCT-FUT
    ‘The news of either his coming or his not coming will be conveyed.’

(17) əcuməŋŋu nattrəbədi pʰəroi
    ə-cum-bə ŋəŋ-u nattrəbədi pʰə-roi
    ATT-be right-NZR speak-COMD or be good-NEG
    ‘Speak the truth or it won’t be good!’

Example (16) shows the possibility of the occurrence of nattrəga in a negative environment, while (17) shows that the final VP is negated and such coordination is very frequent in use. Speakers of this language do not use this construction in cases such as (18). In this construction, the coordinator nattrəbədi ‘or’ cannot occur in the positive environment of conjoining the clauses -əi ca təkkəni ‘I will drink tea’ and əi kəpʰi təkkəni ‘I will drink coffee’.
(18)* əi ca təkkəni nattrəbədi kopʰi təkkəni
    əi ca təkkəni nattrəbədi kopʰi təkkəni
I tea drink-FUT or coffee drink-FUT

Coordinators which carry the meaning of offering a choice between two or more conjuncts in Manipuri are nattrəgə ‘or/either’, or nattrəbədi ‘except/or’. They conjoin two conjuncts expressing an alternative.

(19) kʰəmba nattrəgə no̱ban tʰao
    kʰəmba nattrəgə no̱ban tʰa-o
    khamba or nongban send-COMD
‘Send Khamba or Nongban!’

(20) tuminə pʰəmmə nattrəbədi pʰurəkkəni
    tumin-ə pʰəmm-u nattrəbədi pʰu-rək-kəni
    be.quiet-ADV sit-COMD lest beat-INCT-FUT
‘Sit quiet otherwise I shall beat you!’

The presence of the coordinator nattrəgə ‘or/either’ in (19) gives a choice between the referents of the two conjuncts kʰəmba and no̱ban to be sent. However, in (20), the presence of the coordinator nattrəbədi ‘lest/otherwise’ keeps a condition that is, ‘sitting quiet’ or ‘beating’.

Sentences (19) and (20) above can be transformed into bisyndetic and asyndetic structures, respectively, as in the following examples.

(21) kʰəmboɾə no̱ban-ə əma tʰao
    kʰəmboɾə no̱ban-ə əma tʰa-o
    khamba-SC nongban-SC one send-COMD
‘Send either Khamba or Nongban!’

(22) tuminə pʰəmmə pʰurəkkəni
    tumin-ə pʰəmm-u pʰu-rək-kəni
    be.quiet-ADV sit-COMD beat-INCT-FUT
‘Sit quiet otherwise I shall beat you!’

In (21), when the coordinator nattrəgə ‘or’ is withdrawn, suffixal coordinators are attached to the conjuncts. Such a construction is very frequently used in this language. For instance, the suffixal coordinators -lə--rə is added to the final and non-final conjuncts, i.e. kʰəmbo and no̱ban respectively with a numeral əma ‘one’ which occurs after the last conjunct.

3.3. Adversative Coordination

In adversative clause coordination, lexical coordinators such as ədubu ‘but’, ədumoinəmək ‘however’, and təuigungəbəmən ‘but’ are employed. In adversative coordination, one of the two clauses is negated and/or is understood as an opposition to the other one. The coordinators ədubu

68
‘but’ and *taugumbəsun* ‘but’ carry the adversative meaning and they do not occur in NP coordination. The most frequently used adversative coordinator is adubu ‘but’.

In (23) and (24), the final clauses (*kəptu *kaide* ‘(the cup) did not break’ and *(məhak) *parikʰ*ə *ŋəmdə* ‘(he) did not pass the exam’) are negated. They are coordinated by adubu ‘but’ with the clauses *kəptu *tare* ‘the cup fell down’ and *məhak *lairik *kənnə *pai* ‘he worked hard’, respectively.

(23) *kəptu *tare *adubu *kaide*

\[
\text{kəp-tu ta-re adubu kai-de-i} \\
\text{cup-DEM fall-PERF but break-NEG-ASP} \\
\text{‘The cup fell down but it did not break.’}
\]

(24) *məhak *lairik *kənnə *pai *adubu *parikʰ*ə *ŋəmdə*

\[
\text{məhak lairik kan-no pai adubu parikʰə ŋəm-de-i} \\
\text{he book be.hard-ADV read but exam pass-NEG-ASP} \\
\text{‘He worked hard but did not pass the exam.’}
\]

Sentences (25) and (26) reveal that *taugumbəsun* ‘but’ also functions as an adversative coordinator. It is found in written language but is not very frequent in spoken language. In (25), though the first clause describes the height of Rani, the second clause posits the beauty of Rani.

(25) *rani *nemmi *taugumbəsun* *məsək *pʰəjəi*

\[
\text{rani nem-i təugumbəsun mə-sək pʰə-jə-i} \\
\text{rani be.short-ASP but 3PP-feature be.beautiful-REFL-ASP} \\
\text{‘Rani is short but beautiful.’}
\]

In sentence (26), the conjoining of the two clauses by the coordinator *taugumbəsun* ‘but’ shows that the non-final clause mentions the news heard however, the final clause rejects the news.

(26) *paedu *tare *taugumbəsun* *mədu *əcumə *nətte*

\[
\text{paudu ta-re təugumbəsun mədu ə-cum-bə} \\
\text{news-DEM hear-PERF but it ATT-be.right-NZR} \\
\text{nət-te-i} \\
\text{not-NEG-ASP} \\
\text{‘I heard the news but it is not true.’}
\]

In (27), the coordinator *ədumoinəmək* ‘however’ conjoins the two clauses expressing that the reason of Tomba’s coming is understood as an opposition to the first clause.
(27) tombə lakə̃m-mi adumoinəmək mərə̃m-di hairə̃nde
   tombə lak-lə̃m-i adumoinəmək mərə̃m-di
   tombə come-COMPL-ASP however reason-DEM
   hai-rə̃m-de-i
   say-COMPL-NEG-ASP
   ‘Tomba came here however he did not say why he came.’

4. Position of coordinator(s)

In Manipuri, monosyndetic and bisyndetic coordination constructions prevail in the following structures:

   Monosyndetic: [A] [Co] [B]
   Bisyndetic: [A-Co] [B-Co]

In these structures, A and B stand for two coordinating conjuncts and Co stands for the coordinator. In monosyndetic coordination, the structure is just a symmetrical tripartite structure where the coordinator arises from lexical categories. In bisyndetic coordination, coordinators are all suffixal and they attach to the conjuncts. Examples of the monosyndetic and the bisyndetic construction respectively are given in (28) and (29).

(28) [tombə] [əməsən] [caubə] cət-kə-re
    tombə əməsən caubə cət-kə-re
    Tomba and Chaoba go-DEF-PERF
    ‘Tomba and Chaoba have gone.’

(29) [tombə-ne] [caubə-ne] cət-kə-re
    tombə-ne caubə-ne cət-kə-re
    Tomba-SC Chaoba-SC go-DEF-PERF
    ‘Tomba and Chaoba have gone.’

One exception is offered in (30). Here, the suffixal coordinator -ne is used along with the lexical coordinator adugo ‘and’.

(30) tonne ramne manine adugo kumame lakkəni
    tom-ne ram-ne mani-ne adugo kumar-ne lak-kəni
    tom-SC ram-SC mani-SC and kumar-SC come-FUT
    ‘Tom, Ram, Mani and Kumar will be coming’.

5. Category-sensitivity

The conjunctive coordinators əməsən/əmədi ‘and’, adugo ‘and’ and the disjunctive coordinators nəttrə-gə ‘or’, nəttrə-bədi ‘except/or’, nəttrə-dəbu ‘if it is not so’, can link different
categories such as noun phrases (NPs), verb phrases (VPs), adjective phrases (Aps), postpositional phrases (PPs), and clauses. The adversative coordinators ədubu ‘but’, touigungmosan ‘but’, adumoinamak ‘however’, are mostly confined to clauses and VPs.

The conjunctive coordinators conjoin NPs, like VPs, APs, PPs and clauses. In (31), əmədi conjoins NPs; in (32), ədugə conjoins VPs; in (33), əmədi conjoins APs; in (34), əməsan conjoins PPs; and in (35), ədugə conjoins clauses.

(31)  məhak əmədi əi cətkəni
məhak əmədi əi cətkəni
he and I go-FUT
‘He and I will go there.’

(32)  məkʰoi cək cəkʰre ədugə cətkʰre
məkʰoi cək cəkʰre ədugə cətkʰre
3PP-PL rice eat-DEF-PERF and go-DEF-PERF
‘They have taken rice and gone away.’

(33)  əcumbo əmədi əsenəba wa ənənə
əcumbo əmədi əsenəba wa ənənə
ATT-be true-NZR and ATT-be real-NZR word speak-COMD
‘Speak the true and real word.’

(34)  nəŋgi əməsan ən ənənə cətlə
nəŋgi əməsan ən ənənə cətlə
you-GEN front and back look-ADV go-COMD
‘Go looking your front and back.’

(35)  məhakŋə isai sakʰi ədugə əkʰənə nok-i
məhakŋə isai sakʰi ədugə əkʰənə nok-i
he-ERG song sing-ASP and I-PL-ERG laugh-ASP
‘He sings a song and we laugh at it.’

Like conjunctive coordinators, disjunctive coordinators also can conjoin NPs, VPs, APs, PPs and clauses as in the examples below. Specifically, natrəgo ‘or’ coordinates NPs in (36), VPs in (37), APs in (38), PPs in (39), and clauses in (40).

(36)  məhak natrəgo əi cətkəni
məhak natrəgo əi cətkəni
he or I go-FUT
‘He or I will go there.’

(37)  məkʰoi cək cəkʰre natrəgo cade kʰənde
məkʰoi cək cəkʰre natrəgo cade kʰənde
3PP-PL rice eat-DEF-PERF or eat-DEF-ASP know-PERF-ASP
‘I don’t know whether they have taken rice or not.’
(38) əcumbo nattrəgo əsenə bo wa ɲaŋ-u
   ə-cum-bo  nattrəgo  ə-seŋ-bə  wa  ɲaŋ-u
ATT-be.true-NZR  or  ATT-be.real-NZR  word  speak-COMD
‘Speak the true or real word.’

(39) nəŋgi maŋ nattrəgo tuŋ jeŋ-u
   nəŋ-gi  maŋ  nattrəgo  tuŋ  jeŋ-u
you-GEN  front  or  back  look-COMD
‘Look at your front or back.’

(40) isoi sakkəni nattrəgo jəgoi sa-gəni
   isoi sək-kəni nattrəgo jəgoi sa-gəni
song  sing-FUT  or  dance  play-FUT
‘(We) will sing or dance.’

Unlike conjunctive and disjunctive coordinators, adversative coordinators conjoin only clauses and VPs as in (41) and (42), respectively.

(41) mahakəna isai səki adubu ai-kəoi-na noki
   mahak-na  isai  sək-i  adubu  ai-kəoi-na  noki
he-ERG  song  sing-ASP  but  I-PL-ERG  laugh-ASP
‘He sings a song but we laugh at it.’

(42) məkəoi jaurəkkhre adubu cəŋləktri
   mə-kəoi  jau-rəkkh-re  adubu  cəŋ-lək-tə-ri
3PP-PL  arrive-INCP-DEF-PERF  but  enter-INCT-NEG-PROG
‘They have arrived but did not enter here.’

6. Semantic distinctions

The conjunctive coordinators əmsənəŋ əmədəi ‘and’ and ədugə ‘and’ are used independently of the meaning of the conjuncts. Manipuri does not employ different conjunctive constructions dependent on the animacy of the conjuncts. NPs with human and non-human referents are conjoined with no difference e.g. ram əmsənəŋ sita ‘Ram and Sita’, ɲəŋ əmədəi əi ‘You and I’. No distinction is there regarding conjunctive coordination between proper names and common nouns, e.g. imphəl əmsənəŋ delhi ‘Imphal and Delhi’, hui əmsənəŋ səgol ‘dog and horse’.

Manipuri distinguishes between tight and loose coordination. Tight coordinators are used with items that can be thought of as couples or pairs which are considered closely associated in the real world. Loose coordinators are used with items which are less closely associated (Haspelmath 2004). For instance, ɲəŋ nəŋəsit ‘rain and wind’, ima ipa ‘my mother and my father’, hui həudəŋ ‘dog and cat’ are some examples of tight coordination. Here the coordinative conjuncts are simply juxtaposed and they are often treated as coordinative compounds which are conjoined with no overt coordinator. However, ima əmsənəŋ inəu ‘my mother and my brother’, u əmsənəŋ ləi ‘tree and flower’ are the examples of loose coordination where the conjuncts are conjoined with coordinators.
Manipuri does not distinguish clause coordination between same-subject and different-subject clause combinations. Consider (43) and (44).

(43)  
\[ \text{mə} \text{kə} \text{lak} \text{lə} \text{mle} \text{ə} \text{dugə} \text{cətə} \text{kə} \text{re} \]  
\[ \text{mə-kə} \text{lak-ləm-le} \text{ə} \text{dugə} \text{cət-kə} \text{re} \]  
\[ 3\text{PP-PL} \text{come INCPT PERF} \text{and go DEF PERF} \]  
‘They have come here and then left.’

(44)  
\[ \text{ramnə pəməmmi ədugə sitanə leppi} \]  
\[ \text{ram-nə pəm-i ədugə sita-nə lep-i} \]  
\[ \text{Ram-ERG sit-ASP and Sita-ERG stand-ASP} \]  
‘Ram sits and Sita stands.’

While (43) is a same-subject clause combination and (44) is a different-subject clause combination, both instances use the same coordinator ədugə ‘and’.

7. Conclusion

From the above analysis, it can be concluded that there is a difference between morphological and syntactic marking of coordination. Syntactic marking of coordination is found frequently in written language but very rarely in spoken language; while morphological marking of coordination is found frequently in spoken form but very rarely in written form. Very interestingly, another level of marking, namely asyndetic coordination, is unmarked. It is limited to two noun conjuncts and is not the same as echo formation.

Regarding the lexical categories of the conjuncts, the coordinator ədugə ‘and’ is found to conjoin only VP conjuncts, while the coordinator əmsəm/əmsədi ‘and’ conjoins VP as well as NP conjuncts.

The use of the disjunctive coordinator nattrəbədi ‘except/or’ always occurs in a negative context. In alternative coordination, when the coordinator nattrəga ‘or/either’ is withdrawn from a coordinated structure, the suffixal coordinator -ɾə~lə is added to the final as well as non-final conjuncts along with a numeral əmə ‘one’, which occurs after the last conjunct.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1PP</td>
<td>First Person Pronoun</td>
</tr>
<tr>
<td>2PP</td>
<td>Second Person Pronoun</td>
</tr>
<tr>
<td>3PP</td>
<td>Third Person Pronoun</td>
</tr>
<tr>
<td>ACC</td>
<td>Accusative</td>
</tr>
<tr>
<td>ADV</td>
<td>Adverb</td>
</tr>
<tr>
<td>ASP</td>
<td>Aspect</td>
</tr>
<tr>
<td>ASS</td>
<td>Associative</td>
</tr>
<tr>
<td>ATT</td>
<td>Attributive</td>
</tr>
<tr>
<td>COMD</td>
<td>Command</td>
</tr>
<tr>
<td>COMPL</td>
<td>Complete</td>
</tr>
<tr>
<td>CONJ</td>
<td>Conjunction</td>
</tr>
<tr>
<td>COP</td>
<td>Copulative</td>
</tr>
<tr>
<td>DEF</td>
<td>Definitive</td>
</tr>
<tr>
<td>DEM</td>
<td>Demonstrative</td>
</tr>
<tr>
<td>DES</td>
<td>Desire</td>
</tr>
<tr>
<td>ERG</td>
<td>Ergative</td>
</tr>
<tr>
<td>FUT</td>
<td>Future</td>
</tr>
<tr>
<td>GEN</td>
<td>Genitive</td>
</tr>
<tr>
<td>INCT</td>
<td>Inceptive</td>
</tr>
<tr>
<td>LOC</td>
<td>Locative</td>
</tr>
<tr>
<td>NEG</td>
<td>Negative</td>
</tr>
<tr>
<td>NZR</td>
<td>Nominalizer</td>
</tr>
<tr>
<td>PERF</td>
<td>Perfect</td>
</tr>
<tr>
<td>PL</td>
<td>Plural</td>
</tr>
<tr>
<td>PROG</td>
<td>Progressive</td>
</tr>
<tr>
<td>SC</td>
<td>Suffixal Coordinator</td>
</tr>
</tbody>
</table>
3. Coordination in Manipuri

References

4. Passive Constructions in Assam Sadri

Lucky Dey
Dept of English and Foreign Languages, Tezpur University

Abstract
Sadri was originally the mother tongue of the Sadans, an Aryan group in Chota Nagpur Plateau in central India (Navarangi 1965:5). It evolved as a link language of the heterogeneous non-Aryan groups of the region known as the Adivasis. The present paper gives a descriptive account of the passive constructions in Sadri spoken in Assam (henceforth AS). In AS, the periphrastic passive consists of the perfective participle marker -al or the non-finite -ɛk suffixed to the main verb, followed by the auxiliary verb hɔ- ‘happen’ or dja ‘go’, which carries the 3rd person singular agreement marker by default. There is a difference in the syntax and semantics of the two dependent markers -al and -ɛk. The former carries a causative implication and also of a defocused agent, while the latter does not.

Citation

1. Introduction

Languages typically have different constructions to foreground the patient or the action in a clause. This is done in order to meet specific pragmatic and situational requirements to suppress the agent. Passive constructions are one such means. Siewierska (2005) defines a typical passive construction as fulfilling the following five criteria. First, it contrasts with another construction, the active; second, the subject of the active corresponds to a non-obligatory oblique phrase of the passive or is not overtly expressed; third, the subject of the passive, if there is one, corresponds to the direct object of the active; fourth, the construction is pragmatically restricted relative to the active. Finally, the construction displays some special morphological marking of the verb.

Languages with basic passives may employ either strict morphological or periphrastic passive constructions. Strictly morphological passives do not use auxiliary verbs, while “periphrastic passives,” in the line of Keenan (1985), “are usually marked by the use of auxiliary verbs. These verbs are of four kinds: (i) verbs of being or becoming, (ii) verbs of reception, (iii) verbs of motion and (iv) verbs of experiencing”.

Assam Sadri, an Indic language spoken in Assam (henceforth AS), has two periphrastic passive constructions that involve auxiliary verbs: one uses the verb of motion dja ‘go’, and the other uses the verb of being or becoming hɔ- ‘happen’. In AS passive constructions, the topicalised argument, which corresponds to the object of the active clause, does not act as a

---

1 I would like to thank my informants Mr. Rajen Rohidas, Mr. Pradeep Hemrom, Luke Horo and others who are native speakers of Sadri living in Assam. I am grateful to my PhD supervisor Prof. M. Barbora for her guidance and support. I extend my sincere gratitude to Dr. Gautam Borah for his valuable suggestions and insight.
subject, as it retains the semantics and case morphology of the object of the corresponding active construction.

The paper primarily gives a descriptive account of passive constructions in AS, which includes a detailed discussion of both the basic and non-basic passive types. Secondarily, the paper brings forth a comparative discussion of passive constructions in Nagpuriya Sadri (NS) variety and AS, in order to assess the influence of the dominant regional language Assamese on AS. AS data were collected from Sadri speakers of Lakhimpur and Sonitpur districts in Assam. NS data have been collected from native speakers of this variety during a field trip to Ranchi, Jharkhand. Some of the NS examples are also cited from Goswami (1976).

1.1. The Sadri language and its typological features

Sadri or Sadani originated as the mother tongue of the Sadans, an Aryan group amongst the Non-Aryans in Chota Nagpur Plateau (Navarangi 1965:5). It evolved as a link language of the Adivasis mainly living in and around Chota Nagpur Plateau, which spreads over present day Bihar, Jharkhand, Chhattisgarh and parts of West Bengal and Orissa. About 97 or so heterogeneous communities like the Munda, Kharia, Ho, Oraons and others come under the umbrella term Adivasis. These communities, mostly belonging to three language families, namely Austro-Asiatic, Indo-Aryan and Dravidian, used Sadri as their link language for inter- and intra-community communication. Starting out as a pidgin, Sadri gradually evolved to a creole, primarily due to inter-community marriages between the various linguistic groups. The offspring of these bilingual parents adopted this link language as their mother tongue. According to the 2001 Census Report, there are 2,044,776 Sadri speakers in India.

When in the 19th century, the British tea planters brought these adivasi people to Assam as labourers, Sadri, as their link language, came along with them to Assam. Over the last two hundred years, Sadri has been under tremendous influence of the dominant regional languages: Assamese in the Brahmaputra Valley, and Bangla in the Barrack Valley. Sadri as spoken in the Brahmaputra Valley in Assam has been labelled Assam Sadri to distinguish it from the Sadri spoken in the Barrack Valley (Chota Nagpur Plateau), better known as Nagpuria Sadri (henceforth NS). AS differs from NS at the phonological, morphological and lexical level.

Sadri essentially belongs to the Indo-Aryan branch of the Indo-European language family, even though it has been shaped by creolization. It has many typically Indic features such as verbs being inflected for Tense, Aspect, and Mood and the postpositions being free morphemes. Besides, Sadri is a Nominative-Accusative language: S and A arguments are in unmarked nominative case, different from the O argument, which takes the accusative case. Consider (1) and (2).
4. Passive Constructions in Assam Sadri

(1) \([\text{birsi}]_A [\text{etowa ke}]_O \text{ pit-l-ak}\)

birsi: NOM etowa ACC beat-PFV-3SG
‘Birsi has beaten Etowa.’

(2) \([\text{birsi}]_S \text{ ekhon nach-ət} h-ɛ\)

birsi: NOM now dance-PROG AUX.PRES-3SG
‘Birsi is now dancing.’

In (1), \text{birsi} ‘Birsi’ is the A argument, and \text{etowa} ‘Etowa’ is the O argument of the transitive verb \text{pit} ‘beat’. In (2), the NP \text{birsi} ‘Birsi’ is the S argument of the intransitive predicate \text{nach} ‘dance’. In (1) the NP functioning as the O argument is marked differently from the S and A in (1) and (2), respectively.

According to Masica (1991:340), the subject position in Indo-Aryan languages can be occupied by an agent and also by non-agents, which normally do not trigger verbal agreement. The language under study being an Indo-Aryan language has a similar phenomenon. In AS, besides the usual nominative case, subject NPs also receive non-nominaive case markers, namely, the dative, genitive, and locative. Although these subjects occur at the sentence-initial position, they normally do not trigger verbal agreement and bear the role of an experiencer, possessor and goal (Dey, 2014). Indo-Aryan languages often exhibit differential object marking related to animacy, definiteness and specificity of the O argument (Blake 2001, Butt 1993, Mohanan 1994). Likewise in AS, O arguments can either be marked by accusative case or can remain unmarked. The specific DOM criterion for AS is that if the object is \([+\text{animate}]\), it takes accusative case \(kɛ\); and if it is \([-\text{animate}]\), it takes nominative case (or is not marked for case). The object case will depend upon the ‘animacy hierarchy’ where the object higher in animacy takes the accusative case. The definite \([+\text{animate}]\) object NPs take accusative and the definite \([-\text{animate}]\) object NPs remain unmarked for case. Consider example (3) and (4).

(3) \(\text{etowa birsi ke pit-l-ak}\)

etowa: NOM birsi ACC beat-PFV-3SG
‘Etowa beat Birsi.’

(4) \(\varnothing \text{ gilas-thɔ bhaŋ-l-ak}\)

3SG: NOM glass-CL break-PFV-3SG
‘He broke the glass.’

In (3), the object \text{birsi}, being a \([+\text{animate}]\) NP, takes the accusative case marker \(kɛ\), while in (4), \text{gilas-thɔ} ‘the glass’\(^2\) takes the unmarked case by virtue of being the \([-\text{animate}]\) NP.

2. Morphosyntactic elements of passivisation in AS

In AS, periphrastic passives consist of a dependent-marked main verb which is followed by an auxiliary verb. Dependent marking of the verb is achieved either by the perfective participle marker \(-al\) or the non-finite suffix \(-ɛk\), as discussed in §2.1. There are two auxiliary verbs

\(^2\) \text{-thu} and \text{-ta} are so-called classifiers in AS that are suffixed to nominals in order to indicate definiteness or specificity.
involved in passive constructions: \( hɔ \) ‘happen/become’ (i.e., an auxiliary verb of being or becoming), and \( dża \) ‘go’ (i.e., a verb of motion), as discussed in §2.2.

2.1. Dependent marking of the verb: -al and -ek markers

The non-finite suffix -ek is used as a nominaliser and to mark complement clauses. Consider (5) and (6).

(5) \( ɛnta \ kha-ek-\t^{a} \ kharap \ ah-ɛ \) (Nominaliser)
    too.much eat-NF-CL bad COP.PRES-3SG
    ‘Too much eating is bad’.

(6) \( ʊ \ ghar \ dża-ek \ khōd-æt \ h-ɛ \) (Infinitive)
    3SG:NOM home go-NF want-PROG AUX.PRES-3SG
    ‘He wants to go home.’

In (5), the suffix -ek on kha ‘eat’ functions as a nominaliser and allows for the classifier -\( t^{a} \)\(^3\) to be added, resulting in an NP kha\( k^{t}a \) ‘the eating’. In (6), -ek is suffixed to dża ‘go’ to form a complement clause of the main verb khōd\( ð \) ‘want’. As a cover term for this functional range, -ek has been glossed as ‘non-finite (NF).’

When -ek ‘NF’ is part of the passive construction, it occurs on the main verb, which is then followed by the auxiliary hɔ ‘happen,’ as in (5). This passive construction emphasizes or foregrounds the action. However, the non-finite -ek marker needs to be causativised as in (7) in order to imply an external agent. Instead of the sequence of -aa and -ek, the -al marker can be used with the same effect, as in (8).

(7) (a) \( kam \ kar-ek \ ho-l-ak \)
    work do-NF happen-PFV-3SG
    ‘Work has been done.’

    (b) \( kam \ kar-aa-ek \ ho-l-ak \)
    work do-CAUS-NF happen-PFV-3SG
    Lit: ‘Work has been done (by someone)’.

    (c) \( kam-thu \ chōra-man-k\_s \ e \ kar-aa-ek \ ho-l-ak \)
    work-CL boy-PL-GEN INS do-CAUS-NF happen-PFV-3SG
    Lit: ‘The work has been done by the boys’.

(8) \( kam \ kar-al \ ho-l-ak \)
    work do-PFV.PTCP happen-PFV-3SG
    Lit: ‘Work has been done (by someone).’

(7)(a) is an impersonal passive construction that lacks the notion of a referential or external agent and merely foregrounds the notion of ‘action was done’ (this is further discussed in §4).

\(^3\) See footnote 2.
On the other hand, in (7)(b), the causative marker -aa implies that the work was accomplished with someone’s help or by someone. Example (7) shows the possibility of an added agent chōra-man-ker se ‘by the boys’ with the verb taking the causative marker -aa followed by the -ek marker. Consequently, the -al marker can co-occur with a causer marked by instrumental se as in (9)(b), whereas, the -ek marker lacking the causative implication cannot co-occur with a causer as in (9)(a).

(9) (a) *tē-ker se kam kar-ek hɔ-l-ak
3SG-GEN INS work do-NF happen-PFV-3SG
‘The work was done by him.’

(b) tē-ke se kam kar-al hɔ-l-ak
3SG-GEN INS work do-PFV.PTCP happen-PFV-3SG
‘The work was done by him.’

2.2. Passive auxiliaries in AS: hɔ-l-ak vs ge-l-ak

Periphrastic passives in Sadri are marked by the use of an auxiliary verb of motion, ḍga ‘go’, or one of being or becoming, hɔ ‘be’. This is also seen in other Indic languages like Assamese and Bangla where both the verb of motion and that of being/becoming have grammaticalized as auxiliaries in passive constructions (i.e., hol and gol in Assamese, and holo and gelo in Bangla). Regarding verbs of ‘being or becoming’, Keenan and Dryer (2007:337) remark that such passive auxiliaries “commonly exhibit a certain ambiguity as to whether they are interpreted dynamically or purely statively”. They illustrate this with the English example of ‘the vase was broken,’ which is ambiguous in the sense that it can either specify ‘a state of the vase’ or ‘an activity performed upon the vase’. In the first case, the broken state of the vase might not have been caused by an external agent, and in case of the latter, it might have been. Conversely, the passive auxiliary derived from the verb of motion ‘go’ in Indo-Aryan languages is less ambiguous while interpreting dynamicity. For instance, in the Hindi sentence murgi mari gayee glossed as ‘chicken killed went’, meaning ‘the chicken was killed’ (Keenan and Dryer 2007:11), the dynamic interpretation can be drawn from the fact that an activity is performed upon the patient ‘chicken’ by some external agent. This construction is unlikely to have a stative implication of a ‘dead chicken’.

Both the auxiliaries hɔ-l-ak and ge-l-ak, though inherently suggesting motion or a process, also emphasize a ‘resultant state’ in their passive reading. However, these two auxiliary verbs have some semantic differences. In passive constructions, the verb of motion ge-l-ak may be interpreted with an abilitative sense or a non-volitional reading. The non-volitional reading of the verb ge-l-ak is used in a more restricted context than the former reading as it occurs only with perception verbs. Examples (10) and (11) illustrate the non-volitional/abilitative and volitional readings of the passive auxiliaries ge-l-ak and hɔ-l-ak, respectively.

(10) cor-thɔr ke dekh-al ge-l-ak (non-volitional)
thief-CL ACC see-PFV.PTCP go-PFV-3SG
‘The thief was seen/noticed.’ (or: ‘The thief could be noticed.’)
In (10) the use of ge-l-ak after dekh-al indicates non-volition whereas, in (11) the use of ho-l-ak after the verb dekh-al suggests volition. Regarding passives in Indo-Aryan languages, Pandharipande (1981:16-7, 163-4) states that “all Indo Aryan languages as well as Kannada (Dravidian) require verbs to express a volitional act if they are to be passivised”. In other words, those constructions where a volitional act is expressed with action verbs like ‘eat’, ‘kill’, ‘do’ and so on, can have passive counterparts. On the other hand, constructions which do not express a volitional act, for instance copula constructions, cannot be passivized. In AS however, not only constructions implying volitional acts can be passivised. It is important to note that because of the presence of the two passive auxiliary verbs ho-l-ak and ge-l-ak, a distinction can be drawn between volitional and non-volitional passives. Consequently, the language allows perception verbs like dekh ‘see’, which imply a non-volitional act, to be passivised as in (10).

The passive auxiliary ge-l-ak, besides giving a non-volitional implication with perception verbs, also has an additional modal connotation indicating ability. Unlike ho-l-ak, which suggests ‘state of the action done’, it suggests the manner of how the action is done, semantically implying ‘overcoming restraints’. When the manner of ‘how the action is done’ is suggested, the action done can be conceived as involving some degree of dynamicity.

With action verbs like kar ‘do’ in (12), ge-l-ak conveys an abilitative passive sense, while ho-l-ak implies the ‘state of the action done’. The nature of the action verb kar-al ‘do-PTCP’ implies both volition and causation. But with ho-l-ak in (12)(b), it can be interpreted as ‘was done’ and with ge-l-ak in (12)(a), it can be interpreted as ‘could be done’. A similar interpretation can be drawn from (13).

In (13), the sentence with auxiliary ho-l-ak can be interpreted as ‘the snake was killed’ and with auxiliary ge-l-ak, the sentence can be interpreted as ‘the snake could be killed’. The same non-modal versus modal interpretation occurs with the auxiliary verbs hol and gol in passives of Assamese.
4. Passive Constructions in Assam Sadri

(14) kam-tu kor-a ho-l/go-l
     work-CL do-NF COP-PFV/go-PFV
     ‘The work was done/could be done.’

2.3. Case markings in passive constructions

In AS passives, the core arguments undergo a change in syntactic role. In the passive of a transitive clause, the O argument becomes the S argument in passivisation and the A goes in a peripheral position. The A argument in the active sentence receives the unmarked nominative case and in passivisation it takes the instrumental se or dwara. The O argument in the active sentence is differentially marked depending upon animacy and definiteness criteria, as has been discussed in §1.1. In passivisation, the O argument retains the accusative case marking of the active form. This is illustrated in (15), and further discussed in §3.1.

(15) [birsi ke]$_S$ pit-al ge-l-ak [(etowa-ker se)]$_{OBL}$
    birsi ACC beat-PFV-PTCP go-PFV-3SG etowa-GEN INS
    ‘Birsi was beaten (by Etowa).’

3. Passivisation of Different Verb types

Traditionally, passives of transitive verbs are considered basic. This might be because in Western Indo-European languages like English, intransitive verbs cannot be passivized. In Indo-Aryan languages, however, passivisation of both transitive and intransitive verbs is possible (Masica 1991:357-358). In this connection, Shibatani (1998:96) argues that ‘the notion of action is expressed by both transitive and intransitive verbs and accordingly the action-based understanding of voice would predict that the active-passive opposition obtains with both intransitive verbs and transitive verbs as long as they denote actions’.

In AS both transitive and intransitive verbs that denote actions can be passivized. In the following subsections, passivisation of transitive, ditransitive and intransitive verbs are discussed.

3.1. Transitive passives in AS

In AS passivisation, the O argument of the original transitive verb is normally topicalised and occurs in sentence-initial position. In passive constructions, the verb does not agree with the topicalised argument (or the passive S argument) in person and number like the corresponding active construction. Instead, the verbs carry the default 3rd person singular agreement marker -ak. Keenan and Dryer (2007:336) consider such agreement markers that do not actually agree with an argument “semantically empty”. Nevertheless, Masica (1991:361) also points out that the passive subjects have verb agreement in some languages like Hindi, Sindhi, Kashmiri, Gujarati and Marathi, primarily. Again, in most Indo-Aryan languages, the O argument has DOM features, and in passivisation, the passive subject retains the DOM marking of the corresponding active sentence (Masica 1991:361). There are different opinions regarding the notion of passive subject in Indic languages. Indo-Aryan languages do not require the subject to be overtly expressed in active clauses. Moreover, the promotion of the object to subject position in passive constructions is often termed as “incomplete” in these languages (Masica 1991). The reason
behind calling the promotion “incomplete” is that the object of the corresponding active sentence is not always fully promoted to subject. The fact that the object retains its DOM case features even after passivisation suggests the lack of a subject NP in the AS passive constructions. In AS, however, preliminary evidence suggests that the passive S argument fulfills the syntactic criterion of being an antecedent of a reflexive or a pronominal, which points towards its subjecthood. Nevertheless, case marking and verb agreement show the lack of subjecthood of the passive S argument. Consider the following examples.

(16) \[\text{etowa} \quad \text{birsi} \quad \text{ke} \quad \text{pit}-\text{ak} \]
    etowa:NOM birsi ACC beat-PFV-3SG
    ‘Etowa beat Birsi.’

(17) \[\text{birsi} \quad \text{ke} \quad \text{pit-al} \quad \text{ge}-\text{ak} \quad (\text{etowa-ker se}) \]
    birsi ACC beat-PFV.PTCP go-PFV-3SG etowa-GEN INS
    ‘Birsi was beaten (by Etowa).’

(18) \[\text{birsi} \quad \text{ke} \quad \text{pit-ek} \quad \text{ho-lak} \]
    birsi ACC beat-NF happen-PFV-3SG
    Lit: ‘Birsi was beaten.’
    ‘The action of beating of Birsi was done.’

In (17) \text{birsi ke} is the passive S argument which has the object case marker \text{ke} and is the argument affected by the action of ‘beating’. The demoted passive agent \text{etowa} receives the instrumental case \text{se} preceded by the genitive base \text{-ker}. Overt mention of the agent is optional in the passive construction. Again, in (18), we can have a passive construction with the verb carrying that non-finite \text{-ek}, though it is less frequently used. Here the emphasis is more on the ‘resultant state’. Thus the auxiliary \text{ho-lak} is used.

Again, an inanimate O argument remains unmarked, as in (4), repeated here as (19)(a).

(19) (a) \[\text{gilas-tho} \quad \text{bhaŋ-lak} \]
    3SG:NOM glass-CL break-PFV-3SG
    ‘He broke the glass.’

(b) \[\text{gilas-tho} \quad \text{bhaŋ-al/*ek} \quad \text{ho-lak} \]
    glass-CL break-PFV.PTCP/*-NF happen-PFV-3SG
    ‘The glass was broken (by him).’

In (19)(b), \text{gilas-tho} ‘the glass’ is the passive S argument. Here, the use of \text{-ek} yields an incorrect construction with the auxiliary \text{ho-lak}. Even with the auxiliary \text{ge-lak}, the construction would be unacceptable because of the inherent nature of the verb \text{bhaŋ} ‘break,’ which needs a causer and a volitional implication.\(^5\)

---

4 Hook (1979:120) remarks that overt accusative marking of the object can be retained after passivisation in Hindi as in \text{mujh ko fauran pehchaan liyaa jaaegaa} ‘I will be recognized immediately’. But the accusative on the NP \text{mujh ko ‘me’} suggests that the object may not have been promoted to the grammatical subject position.

5 A separate construction is used for a non-volitional interpretation. This is the conjunctive participle construction. It consists of morpho-phonemic change in the vowel sound of the verb: \text{bhaŋ} ‘break’ becomes \text{bhaŋy} ‘break.CP’ in the expression \text{bhaŋy gel-}
3.2. Ditransitive passives in AS

While the word order in AS is flexible, in case of active sentences, the basic word order is SOV. In active ditransitive constructions, both objects can either precede or follow each other depending upon which object needs to be focused. In AS the case marker ke occurs with the O argument and with R. This ke marking is however differentially marked (DOM) on the O argument depending upon animacy and definiteness factors. So the O argument receives both unmarked and accusative case ke. The T argument always receives the unmarked irrespective of animacy factors. For instance, the T argument chawa-thɔ ‘the child’ is in unmarked case despite being an animate NP in the sentence moi ɛ-ke chawa-thɔ delo ‘I gave him the child’. Thus, in AS it is seen that animate O arguments align with R arguments (corresponding to the primary-secondary object alignment) and inanimate O arguments align with T arguments (corresponding to the direct-indirect object alignment). This is illustrated in the following example.

(20) (a) birsi siṭa ke citthi bhedʃ-1-ak
birs:i NOM sita DAT letter send-PFV-3SG
‘Birsi sent a letter to Sita.’

(b) siṭa ke citthi bhedʃ-al hɔ-l-ak
Sita DAT letter send-PFV.PTCP happen-PFV-3SG
‘The letter was sent to Sita.’

(c) citthi-thɔ siṭa ke bhedʃ-al hɔ-l-ak
letter-CL sita DAT send-PFV.PTCP happen-PFV-3SG
‘THE LETTER was sent to Sita.

The R argument ‘Sita’ is marked by the dative case ke and the T argument is citthi ‘letter’ in unmarked case. In the passive construction in (20)(b), siṭa is in sentence-initial position with the verb carrying the passive morphology, while in (20)(c) the T argument citthi ‘letter’ is in the initial position. In both these passive constructions, the verb bhedʃ ‘send’ takes the perfective -al followed by ge-l-ak. The verb bhedʃ ‘send’ followed by hɔ-l-ak will imply ‘state of the letter being sent’. The passive auxiliary ge-l-ak here would suggest that the letter could be sent to ‘Sita’, in an abilitative sense. In both the cases if the ditransitive verb bhedʃ carries the -ek, it must be followed by hɔ-l-ak but not with *ge-l-ak.

(21) (a) *siṭa ke citthi bhedʃ-ek ge-l-ak
(b) *citthi-thɔ siṭa ke bhedʃ-ek ge-l-ak

ak: ‘was broken (accidentally)’. The conjunctive participle form is also used in perfective aspect. The verb taking such change of vowel sound, henceforth, will be glossed as ‘verb.CP’.
As has been mentioned in §2.1, the use of -ɛk as passive morphology focuses merely on the ‘action done’ without any explicit notion of causation. In other words, a stative implication is expressed by the non-finite -ɛk. Again, an ablative sense is expressed by the passive auxiliary ge-l-ak, which in turn emphasises the dynamicity of the action. Consequently, the properties of dynamicity and the manner of ‘how the action is done’ suggested by ge-l-ak, are incompatible with the stative implication expressed by the non-finite -ɛk. This explains the ungrammaticality of examples (21).

3.3. Intransitive passives in AS

Passives on intransitives, in the lines of Keenan and Dryer (2007:332) are “passives that lack the property of proto-typical passives in which the subject corresponds to an object in a corresponding active clause”. Nevertheless, these passives do entail the existence of an agent and, in AS, employ the same morphology as passives of transitives and ditransitives. According to Masica (1991:356), “Periphrastic or suffixal, in the New Indo Aryan (NIA) passives are notable in their applicability to intransitives as well as transitives”. There is, however, an exception to this generalization as Shibatani (1985:834) observes: “passives of non-agentive intransitive are not permitted”. This leads us to draw a distinction between two types of intransitives, namely, unergatives and unaccusatives (Perlmutter and Postal 1983). The former is the agentive and the latter is the non-agentive intransitive construction. In AS, unergatives can be passivized as in (22), since agentivity is associated with the unergative argument. Nevertheless, unaccusatives in AS can also be passivized with the passive morphology -al, since -al has a causative implication.

(22) (a) chöri-thɔₚ kand-l-ak
girl-CL:NOM cry-PFV-3SG
‘The girl cried.’

(b) chöri-thɔₚ ke kand-ɛk ho-l-ak
girl-CL ACC cry-NF happen-PFV-3SG
‘The girl had to cry/* crying was done by the girl.’

(c) chöri-thɔₚ ke kand-al ho-l-ak
girl-CL ACC cry-PFV.PTCP happen-PFV-3SG
‘The girl was made to cry.’

Conversely, in case of the unergative construction with the verb kand ‘cry’ (22), it is observed that the suffix -ɛk does not convey a passive reading (an asterisk* is put before the English translation). Instead, the result is a deontic interpretation where kand-ɛk ho-l-ak meaning ‘had to cry’ is suggestive of obligation or compulsion. Thus, in (22)(b) the verb will require the causative marker -aa, preceding the -ɛk, in order to be a possible passive construction in the language. On the other hand, the use of the perfective participle marker -al gives a causative implication to the passivized unergatives in (22)(c) and hence is a possible construction.

In the case of unaccusatives, non-finite -ɛk does not provide the expected passive reading. This is because unaccusative constructions are by nature non-agentive. Even if the causative -aa
is suffixed to the verb of an unaccusative construction, this does not provide an expected passive reading. Conversely, the use of the perfective participle marker -al with causative implication gives a possible passive reading of the unaccusative construction. This is illustrated in (23).

(23)  
(a)  
| darḍa-thès | kʰoilo | ge-l-ak |
| door-CL | open.CP | go-PFV-3SG |
‘The door opened.’

(b)  
| darḍa-thès | kʰo-il-al | hɔ-l-ak |
| door-CL | open-PFV.PTCP | happen-PFV-3SG |
‘The door was opened.’

(c)  
| darḍa-thès | kʰo-il-ɛk | hɔ-l-ak |
| door-CL | open-NF | happen-PFV-3SG |
‘The time of opening the door has happened/*the door opening is done.’

The unaccusative construction (23)(a) is passivized in (23)(b) with the passive morphology -al followed by the auxiliary verb suffixed by the perfective marker and the default 3rd personal singular agreement marker -ak. In (23)(b), the verb takes the perfective participle -al with a causative implication. On the other hand, if the verb takes the non-finite -ɛk, it cannot be passivised because there is no causative implication. The construction is still acceptable but has a perfective implication, as in (23)(c): ‘the time of opening the door has happened’.

Unlike AS, in Assamese, the use of non-finite -a with unaccusative passives gives a possible construction, as in (24) and (25).

(24)  
| dɔrza-kʰɔn | kʰo-il-a | hɔ-l |
| door-CL | open-NF | COP-PFV |
(Assamese)  
Lit: ‘The door has been opened.’

(25)  
| sinema-kʰɔn | su-wa | hɔ-l |
| movie-CL | watch-NF | COP-PFV |
Lit: ‘the movie has been watched/seen.’

In (24) and (25), the non-finite marker -a suffixed to the verb kʰo-il ‘open’ and su ‘see’ gives a causative implication.

4. -ɛk vs. –al as impersonal vs. personal passives in AS

The -al and -ɛk distinction in AS discussed earlier can be seen in terms of personal and impersonal passivisation. The latter is distinguished from the former since it cannot take a referential agent. The presence or absence of a referential agent classifies passives into personal and impersonal. Shibatani (1998) defines impersonal passives as ‘constructions lacking a referential subject instigating an action thereby contrasting with active constructions whose subject expresses an agent instigating an action’. In other words, impersonal passives are generally passives of intransitives that lack a subject NP. However, both personal and impersonal passives share a fundamental grammatical function. At the same time in most languages both
personal and impersonal passives use the same syntactic and morphological means (Keenan and Dryer 2007:344). In AS, impersonal passives also take the same periphrastic passive form as the personal passives. Here also, the action is foregrounded without any reference to the agent. Moreover, the sentence lacks a subject NP. In AS, the non-finite -ek form followed by the auxiliary hɔ- ‘happen’ is used to imply impersonal passives.

(26)  
\[
\begin{align*}
\text{kail} & \quad \text{rait} & \quad \text{deri} & \quad \text{se} & \quad \text{nind-aa-ekhɔ-l-ak} \\
yesterday & \quad \text{night} & \quad \text{late} & \quad \text{INS} & \quad \text{sleep-CAUS-NF} & \quad \text{happen-PFV-3SG}
\end{align*}
\]
Lit: ‘It was slept very late last night.’

The passive construction of the intransitive verb nind-aa-ek ‘sleep-CAUS-NF’ in (26) gives an impersonal implication, that is, without any reference to the agent or the passive subject.

Though Shibatani (1998:106) mentions that “impersonal passives are closely, but by no means exclusively associated with intransitive”, he also agrees that occasionally, impersonal passives with transitive objects are also found “where the direct object retains its grammatical object function”. In the light of the above statement, it can be concluded that both intransitive and transitive verbs can have impersonal passive counterparts with the use of non-finite -ek. Example (27) can be considered an impersonal passive of a transitive verb.

(27)  
\[
\begin{align*}
\text{ti̲an} & \quad \text{kat-ek} & \quad \text{hɔ-l-ak} \\
\text{vegetable} & \quad \text{cut-NF} & \quad \text{happen-PFV-3SG}
\end{align*}
\]
Lit: ‘vegetable cutting was done.’

In (27) the unmarked object ti̲an ‘vegetable’ of the transitive verb kat ‘cut’ remains in the object position and retains its unmarked case feature in the passive construction. In this passive construction, the ‘action done,’ that of ‘vegetable cutting,’ is foregrounded without referring to any agent or any causative implication. Similar constructions are found in Assamese, see (24) and (25) in § 2.2.1.

It is to be noted that in both personal and impersonal passives, the non-referential agent ‘someone’ appears to be manifested in the 3rd person singular agreement marker suffixed to the verb. This implies that the passives in AS are possibly grammaticalized from constructions having an indefinite agent which reflects as the 3rd person singular agreement on the passive verb. A similar instance of a non-referential agent has been observed by Givón (2006) in Kimbundu. Givón (2006) classifies a ‘left dislocation passive clause type’ in Kimbundu where the impersonal ‘they’ has a de-topicalised, non referring agent. This is illustrated in (28), cited from Givón (2006:340).

(28)  
\[
\begin{align*}
\text{Nzu, a-mu-mono} & \quad \text{(imperfective/passive)} \\
\text{john, they-him-saw} & \\
\text{‘John, he was seen.’}
\end{align*}
\]

In this example, the patient ‘John’ is foregrounded and the non-referring agent denoted by the pronominal ‘they’ is de-topicalised or de-focused.
5. Causative-Passive interaction in AS

As has already been mentioned in §2, Shibatani (1985:831) has termed passivisation to be an ‘agent centered’ construction, as the fundamental pragmatic function of this construction is to defocus the agent. It rather brings into focus the action done and the argument that is affected by the action. However, passives cannot normally defocus the notion of causation; this is the reason why in many languages (especially the Indo-Aryan languages) the passive and causative interact morphologically. In AS, except for the derived transitive verbs, where the causative -aa is obligatory, and in case of passives of indirect causatives usually marked by the -vaa marker, passivisation is not dependent on the causative morphology or the causative -aa/-vaa marker like other Indo Aryan languages. For instance, in Assamese, the passive morphology occurs after the causative (Barbora 1994), as is illustrated in (29).

(29) mari-r hotɔwai hari-k pit-ɔwa hɔl
    mary-GEN INS hari-ACC beat-CAUS PASS
    ‘Hari was made to be beaten by Mary.’

In (29), the passive form follows the causative -ɔwa. Similar is the case in Hindi.

(30) mary sɛ hari-ko pit-vaa-yaa gaya
    mary INS hari-ACC beat-CAUS-PFV PASS
    ‘Hari was made to be beaten by Mary.’

However, unlike Hindi and Assamese, in AS in addition to the non-finite plus causative marking option, there also is the -al option that already includes the causative.

6. Passive morphology in NS and AS

The perfective participle -al is used as the passive morphology in both AS and NS, and has a causative implication in both. Apart from the perfective participle -al, in NS passes the perfective marker -e/-ae is used in passives with a causative implication. There is no equivalent of this NS marker in AS.

(31) ɛtowa sɛ birsi ke citthi bhedɔ-ae ge-l-ak (NS)
    etowa INS birsi ACC letter send-PFV go-PFV-3SG
    Lit: ‘Birsi was sent a letter by Etowa.’

(32) ɛtowa sɛ gilas-thɔ bhaŋ-al ge-l-ak(NS)
    etowa INS glass-CL break-PFV.PTCP go-PFV-3SG
    Lit: ‘The glass was broken by Etowa.’

The NS equivalent (in function and distribution) of AS non-finite -eɔk is not used in passivisation. The Assamese equivalent (in function and distribution) of AS non-finite -eɔk is -a, which is also used in passivisation as shown in (33).
Thus, one can assume that the presence of this marker in the AS variety could be because of its contact with Assamese, the regional language of Assam.

The perfective participle -al in passivisation has a causative connotation. This is unlike Assamese, Bangla and Hindi passivised causative constructions where a separate causative marker precedes the passive marker or morphology. But in NS, there are instances where causative-passive interaction can be seen in certain cases. According to Goswami (1976:68), in NS, apart from the perfective -ɛ marker and participial marker -al, passivisation can take the progressive aspect -t/ -at finally followed by the causative -aa. These are two rare instances with such morphology; however, since he has not provided the Hindi translation, it is difficult to get the context in which it is used.

In (34)(b) the subject takes the oblique instrumental case se and the verb is inflected for the direct causative marker -aa followed by the progressive marker t finally followed by the auxiliary he. The causative in (34)(b) is marked by vowel length. Similarly, in (35)(b) the passive meaning has been derived by instrumental se with the subject and the causative -aa suffixed to the verb in progressive aspect. The use of causative marker -aa is evident in the passive construction in (35)(a) and (b) cited from Goswami (1976:55).

The difference between AS and NS varieties can be drawn from the above NS data. The NS example (36) is an instance of an imperfective passive sentence, which has not been found in AS. Again, AS mostly has periphrastic passive constructions while NS has instances like (36), which is not a periphrastic passive construction.
7. Conclusion

From the above discussion it can be concluded that in AS, passive constructions are mainly used to emphasize or foreground the patient or the action. It is noteworthy that though it is an Indic language, it has its unique characteristics like passivisation of unergatives and unaccusatives and the perfective participial -al bearing both passive and causative implications. Although the passive auxiliaries ḍa ‘go’ and ḥo- inherently suggest motion or a process, they also emphasize a ‘resultant state’ in their passive reading. Nevertheless, they differ semantically in their usage; with action verbs ge-l-ak conveys an ablitative passive sense, while hɔ-l-ak implies a ‘state of the action done’. Again, with perception verbs, ge-l-ak suggests non-volitional action whereas the hɔ-l-ak suggests volition. Further, it is observed that the non-finite -εk occurring with transitive verbs is mostly followed by the passive auxiliary hɔ-l-ak. Since it lacks the causative implication, the marker -εk is mostly used in impersonal passivisation. Besides, the brief comparative analysis of NS and AS passives suggests some interesting changes in the latter variety, which are a possible result of contact with the Assamese language.
Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>Accusative</td>
</tr>
<tr>
<td>AS</td>
<td>Assam Sadri</td>
</tr>
<tr>
<td>AUX</td>
<td>Auxiliary</td>
</tr>
<tr>
<td>CL</td>
<td>Classifier</td>
</tr>
<tr>
<td>COP</td>
<td>Copula</td>
</tr>
<tr>
<td>CAUS</td>
<td>Causative</td>
</tr>
<tr>
<td>EXIST</td>
<td>Existential</td>
</tr>
<tr>
<td>GEN</td>
<td>Genitive</td>
</tr>
<tr>
<td>IMPF</td>
<td>Imperfective</td>
</tr>
<tr>
<td>INS</td>
<td>Instrumental</td>
</tr>
<tr>
<td>NOM</td>
<td>Nominative</td>
</tr>
<tr>
<td>NP</td>
<td>Noun Phrase</td>
</tr>
<tr>
<td>NS</td>
<td>Nagpuriya Sadri</td>
</tr>
<tr>
<td>NASS</td>
<td>Passive</td>
</tr>
<tr>
<td>PASS</td>
<td>Past Tense</td>
</tr>
<tr>
<td>PFV</td>
<td>Perfective</td>
</tr>
<tr>
<td>PFV.PTCP</td>
<td>Perfective Participial</td>
</tr>
<tr>
<td>PL</td>
<td>Plural</td>
</tr>
<tr>
<td>PRES</td>
<td>Present</td>
</tr>
<tr>
<td>PROG</td>
<td>Progressive</td>
</tr>
<tr>
<td>SG</td>
<td>Singular</td>
</tr>
</tbody>
</table>
4. Passive Constructions in Assam Sadri

References


5. Negation in Rongmei Naga

Debajit Deb & Kh. Dhiren Singha
Assam University Silchar

Abstract
Rongmei is one of the undocumented Tibeto-Burman languages of Northeast India. It has about 61,197 speakers (Census of India, 2001), and is mainly spoken in the states of Assam, Manipur and Nagaland. The goal of the present study is to describe some of the morphosyntactic aspects of negation in Rongmei Naga as spoken in Barak Valley. Negation in Rongmei is mainly expressed by means of suffixation. However, in addition to the main negative suffix -mak and the prohibitive suffix -ri, a negative interjection ha is also used. Negation in Rongmei can be formed at the clausal level but not at the constituent level. The same negator -mak is shown to be used to negate different types of declarative clauses, including verbal predicates, negative indefinite, and normalized clauses.

Citation

Volume Editors
Gwendolyn Hyslop, Linda Konnerth, Stephen Morey, Priyankoo Sarmah

Copyright
© 2014, the author(s), release under Creative Commons Attribution license

1. Introduction

Rongmei2 is one of the undocumented tribal languages of Northeast India with about 61,197 speakers (Census of India 2001). It is mainly spoken in the three states of Assam, Manipur and Nagaland in Northeast India. The majority of the Rongmei population is found in the Tamenglong district of Manipur. However, a good number of Rongmei speakers are also found in two districts of the Barak Valley of Assam viz., Cachar and Hailakandi. There are thirty-six Rongmei villages in the Barak Valley, of which thirty-five are in Cachar and only one village is found in the Hailakandi district. A large number of Rongmei speakers live in the Dimapur, Kohima and Jaluke provinces of Nagaland as well.

Rongmei belongs to the Zeme group of Tibeto-Burman languages (Burling 2003). Linguistically, Rongmei has close affinities with Liangmai, Paomei, and Zeme. Some speakers of these languages recognize their unity by calling themselves Zeliang or Zeliangrong, names constructed from the first syllables of Zeme, Liangmai and Rongmei. Typologically, Rongmei is a tonal, agglutinating and verb-final language, with most roots being monosyllables. The tones of the language form a four-way tone contrast, i.e., level, rising, falling and rising-falling (Deb 2010), but tone is not indicated in the present study. In

---

1The present paper is the revised version of the one presented under the same title at seventh NEILS, 2012, in Guwahati. We are thankful to anonymous reviewers for their valuable comments and suggestions. Special thanks go to Linda Konnerth, Stephen Morey and Mark Post for their academic support and encouragement in preparation of this paper. We are thankful to Rongmei informants namely Rev. Neihlalung, Mr. Abhijan Rongmei, Haron Kamei, Lungai Kamai, Khangai Rongmei and Alex Kamei for their assistance to collect data for the present study. We are also thankful to Mr. Samir Debbarma, research scholar, Department of Linguistics, Assam University, Silchar for providing us relevant materials in preparation of this paper.

2The term Rongmei likely derives from the word Ruang ‘south’ and mei ‘people’ which literally means ‘the people from the south’. Rongmei Naga is formerly known as Kabui or Kapwi. It is worth mentioning here that Nruanghmei is the alternate name of Rongmei (Burling 2003).
Rongmei, relative clauses precede their head noun whereas numerals and genitive modifiers follow their head noun.

Like many Tibeto-Burman languages of Northeast India, Rongmei is under-documented. Previous linguistic work on Rongmei is limited to two sources: Grierson (1903) and Shreedhar (1979). In the Linguistic Survey of India (LSI), Grierson offers a grammatical sketch of Rongmei in the name of Kabui along with some text data. Shreedhar (1979) describes the phonemes and their allophonic variations of Rongmei Naga as spoken in Nagaland.

The goal of this paper is to describe morphosyntactic aspects of negation in Rongmei as spoken in the Barak Valley of Assam. The present study specifically investigates three different negators and their usage and functions in the language.

2. Rongmei Naga morphosyntax

2.1. Verbs

Verb roots in Rongmei are free, i.e., they are capable of standing independently on their own in larger constructions. Most of the verbs in the language are monosyllabic while disyllabic verbs are also attested in the language. In Rongmei, verbs are neither inflected for agreement with arguments, nor are they morphologically marked for transitivity. However, intransitive verbs can be transitivized by the causative prefix ti- as in (4) and (5). As an agglutinating language, Rongmei is abundant with verbal affixes that express various verbal categories, such as tense and aspect, as in (1) and (2), intensification in (3), causative in (4) and (5), prohibitive in (6), and hortative in (7) and (8):

(1)  
\[
\text{\textit{kamai} \, \text{kəl}^{t} \text{iam-kəu} \, \text{tat-tə-e}}
\]
\begin{align*}
\text{she} & \quad \text{market-LOC} \quad \text{go-PST-DECL} \\
\text{‘She went to the market.’}
\end{align*}

(2)  
\[
\text{\textit{kamai} \, \text{kəl}^{t} \text{iam-kəu} \, \text{tat-mak-bəm-e}}
\]
\begin{align*}
\text{she} & \quad \text{market-LOC} \quad \text{go-NEG-PROG-DECL} \\
\text{‘She is not going to the market.’}
\end{align*}

(3)  
\[
\text{\textit{ai} \, \text{kəu} \, \text{tu-zat-mag-e}}
\]
\begin{align*}
\text{I} & \quad \text{fish} \quad \text{eat-INTS-NEG-DECL} \\
\text{‘I do not eat fish at all.’}
\end{align*}

(4)  
\[
\text{\textit{kamai-rui} \, \text{naŋ-ta} \, \text{ti-nui-dai?}}
\]
\begin{align*}
\text{he-NOM} & \quad \text{you-ACC} \quad \text{CAUS-laugh-IQ} \\
\text{‘Does he make you smile?’}
\end{align*}

(5)  
\[
\text{\textit{naŋ-rui} \, \text{takuan-ta} \, \text{ti-pak-bom-e}}
\]
\begin{align*}
\text{you-NOM} & \quad \text{horse-ACC} \quad \text{CAUS-run-PROG-DECL} \\
\text{‘You are making the horse run.’}
\end{align*}

(6)  
\[
\text{\textit{tat-ri-o}}
\]
\begin{align*}
\text{go-PROH-IMP} \\
\text{‘Don’t go (command)!’}
\end{align*}
5. Negation in Rongmei Naga

(7) *kamai tat-kani-ri-o*

he go-HORT-PROH-IMP
‘Let him not go!’

(8) *kaniu lam-kani-ri-ṭ’o*

they dance-HORT-PROH-IMP
‘Let them not dance!’

Table 1 shows a simplified summary of the position classes of verb construction in Rongmei Naga.

<table>
<thead>
<tr>
<th>PFX</th>
<th>ROOT</th>
<th>SFX 1</th>
<th>SFX 2</th>
<th>SFX 3</th>
<th>SFX 4</th>
<th>SFX 5</th>
<th>SFX 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>ti</td>
<td>0</td>
<td>-zat</td>
<td>-kani</td>
<td>-mak</td>
<td>-tan</td>
<td>-i</td>
<td>-e</td>
</tr>
<tr>
<td>‘CAUS’</td>
<td>‘INTS’</td>
<td>‘HORT’</td>
<td>‘NEG’</td>
<td>‘HAB’</td>
<td>‘PST’</td>
<td>‘DECL’</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>‘č’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>‘bo’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>‘dai’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>‘IQ’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>‘o’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>‘IMP’</td>
</tr>
</tbody>
</table>

2.2. Word classes

Rongmei distinguishes two major word classes: nouns and verbs. Adjectives fall under the category of verbs, particularly stative verbs. They share morphosyntactic features of verbs, i.e., when used predicatively, the adjectivals take verbal affixes; consider (9) and (10).

(9) *ṭ’iŋnuì kaŋ-ṭ’o-e*

leaf be.dry-PST-DECL
‘The leaf is dry.’

(10) *o-pu  ṭ’iŋnuì ti-kaŋ-ṭ’o-e*

1-father leaf CAUS-be.dry-PST-DECL
‘My father made the leaf dry.’

Conversely, when used attributively, the nominalizer -mai is attached to the adjectival, as illustrated in (11).

(11) *ṭ’iŋnuì kaŋ-mai*

leaf dry-NMLZ
‘Dry leaf.’

2.3. Speech acts

All languages have grammatical constructions that indicate whether a clause is an assertion (declaratives), a request (interrogatives), or a command (imperatives). In Rongmei, different
suffixes are employed to form the various types of clauses. The distinction between declarative, interrogative and imperative clauses is signaled by the morphological distinction between the clause final suffixes -e (for declaratives); -dai, -bo, and -čo (for interrogatives); and -o or -t’o (for imperatives).

3. Negation in Rongmei Naga

3.1. Negative construction in Rongmei

Negation in Rongmei is mainly expressed by a suffix. It can be formed at the clausal level but not at the constituent level. In addition to the declarative negative suffix, there is a prohibitive marker as well as a negative particle used for negative interjections. Unlike English lexical negative indefinites, e.g., nobody or nothing, the Rongmei negative indefinite construction is formed by negating the verb rather than the pronominal. Finally, the nominalised clause is negated by the general negator as discussed in §7.

3.2. Types of negators

The regular verbal negator in Rongmei is -mak (pronounced -mag if preceding a vowel). Two other negative elements are the prohibitive -ri and the particle ha. The general negator -mak and prohibitive -ri cannot occur simultaneously. The negative particle ha is used as an interjection that typically occurs clause-initially.

In Rongmei, -mak is the most common negator which is used to negate different types of clauses, including basic verbal clauses, negative indefinite, nominalised clauses, and so on. Interestingly, the negator -mak (-mag) is also found in other Tibeto-Burman languages namely Liangmai (Kuki-Naga; Moita 2013), Koireng (Kuki-Chin; Singh 2010) and Kom (Kuki-Chin; Singh 1996). This must be related to the PTB negative *ma.

4. Negation in declarative clauses

The declarative clause is marked by the clause final suffix -e. To negate a declarative clause, the negator -mak is employed to negate the predicate. Past tense is marked by the morpheme -t’ə as illustrated in (12)(b) and (13)(b), whereas present tense is unmarked as illustrated in (12)(a) and (13)(a). However, this tense distinction is neutralized in the negated form of Rongmei clauses as shown in sentences (12)(c) and (13)(c).

(12)  (a)    tunə-hai  nui-e
         girl-DET    laugh-DECL
         ‘The girl laughs.’

         (b)    tunə-hai  nui-t’ə-e
         girl-DET    laugh-PST-DECL
         ‘The girl laughed.’

         (c)    tunə-hai  nui-mag-e
         girl-DET    laugh-NEG-DECL
         ‘The girl doesn’t/didn’t laugh.’
5. Negation in Rongmei Naga

(13) (a) kamai-rui latrik hi-e
he-NOM book write-DECL
‘He writes a book.’

(b) kamai-rui latrik hi-tə-e
he-NOM book write-PST-DECL
‘He wrote a book.’

(c) kamai-rui latrik hi-mag-e
he-NOM book write-NEG-DECL
‘He doesn’t/ didn’t write a book.’

5. Negation in interrogative clause

In Rongmei, question words such as tʰau ‘who’ and tabui ‘what’ are used to form the content questions, along with the clause final suffix -bo as in (14).

(14) tʰau tə-tʰai tat-mak-puni-bo?
who 1-SOC go-NEG-FUT-IQ
‘Who will not go with me?’

The clause-final suffix -čo, is used for polar questions, as illustrated in (15).

(15) naŋ latrik čam-mak-čo?
you book study-NEG-IQ
‘Don’t you study?’

Finally, the suffix -dai is used for tag questions, as in (16).

(16) kamai latrik tʰai-mak-mai-čo? kiu-dai?
he book know-NEG-NMLZ-IQ be real-IQ
‘He isn’t a learned man, is he?’

In all types of interrogatives, the general negator -mak is used to negate the interrogative clause followed by the question suffixes.

6. Prohibitive and negative hortative

Prohibitive is a type of clausal negation that cross-linguistically often differs from standard negation (Miestamo 2005, 2007). Likewise, Rongmei has a special form of the prohibitive that does not occur in non-imperative clauses (declaratives and interrogatives). A separate prohibitive marker is one of the typological features of most Tibeto-Burman languages of Northeast India such as Ao (Gowda 1975), Boro (Basumatary 2005), Dimasa (Singha 2004), Galo (Post 2009), Garo (Burling 1961), Kokborok (Jacquesson 2007), Lotha (Acharya 1983), Manipuri (Bhat & Singh 1997, Chelliah 1997 and Singh 2000), Paite (Singh 2006), among others. Rongmei has the prohibitive marker -ri whose use correlates with second person and there is no first or third person prohibitive in the language.

Prohibitive -ri is followed by the polite imperative marker -tʰo or command imperative marker -o as shown in (17) and (18).
The use of imperative markers -o or -t’o is obligatory with the prohibitive -ri, as is evident from the ungrammaticality of the following sentences (19)(b) and (20)(b).

(19) (a)  
\textit{dui} laguàŋ-’t’o  
water bring-IMP  
‘Bring water (polite)!’

(b)  
*\textit{dui} laguàŋ  
water bring  
‘Bring water!’

(20) (a)  
\textit{dui} laguàŋ-ri-’t’o  
water bring-NEG-IMP  
‘Don’t bring water (polite)!’

(b)  
*\textit{dui} laguàŋ-ri  
water bring-NEG  
‘Don’t bring water (polite)!’

In Rongmei, the hortative construction is formed by the marker -\textit{kani} which is attached to the verb root followed by the imperative markers -o or -t’o as in (21)(a) and (22)(a). Furthermore, the hortative construction in Rongmei expresses the proposal from the speaker to allow somebody to perform the action. However, it is opposite to express the negative hortative where the speaker proposes somebody not to perform the action. Structurally, hortative negative in Rongmei is formed with the prohibitive marker -ri which always co-occurs with the imperative markers -o or -t’o, and the hortative marker -\textit{kani} is attached to the verb root. In other words, the negator -ri is not directly attached to the verb root, instead it attaches to the hortative marker -\textit{kani}. Subsequently, the order of negator in negative hortative construction in Rongmei is S + V + HORT + NEG + IMP as revealed in the following sentences (21)(b), and (22)(b).

(21) (a)  
\textit{kamai} tat-\textit{kani}-o  
he go-HORT-IMP  
‘Let him go!’

(b)  
\textit{kamai} tat-\textit{kani-ri}-o  
he go-HORT-NEG-IMP  
‘Let him not go!’
5. Negation in Rongmei Naga

(22) (a) әniu tali-kani-ә
we play-HORT-IMP
‘Let us play!’

(b) әniu tali-kani-ри-ә
we play-HORT-NEG-IMP
‘Let us not play!’

7. Negation in nominalised clause

In Rongmei, nominalised or relative clauses are negated via the general negator -mak, which is attached to the verb root followed by the nominalizer -mai: [VERB + NEG + NMLZ]. Negation in nominalised clauses in Rongmei is illustrated in (23), (24) and (25).

(23) tunә zan tu-mak-mai-hai ә-tuaŋ čampan-e
girl meat eat-NEG-NMLZ-DET 1-GEN friend-DECL
‘The girl who does not eat meat is my friend.’

(24) ai zian-mak-mai tunә-hai έei-έә-e
I love-NEG-NMLZ girl-DET dead-PST-DECL
‘The girl whom I do not love is dead.’

(25) bеŋ-mak-mai si-hai ә-tuaŋ-e
white-NEG-NMLZ dog-DET 1-GEN-DECL
‘The dog which is not white is mine.’

8. Negative indefinite pronouns

Iwasaki and Ingkaphirom (2005) point out that most of the languages of East and Southeast Asia typically lack indefinite pronouns. Post (forthcoming) also makes a similar statement that negative indefinite pronouns like ‘no-one’, ‘nobody’ and ‘nothing’ in English are completely alien to Tani languages. Similarly, Rongmei doesn’t have negative indefinite pronouns. Nevertheless, different strategies are used to express the negative indefinite pronouns in the language, i.e. the negative indefinite pronoun is formed by negating the verb rather than the pronominal forms like ‘someone’ or ‘something’ as exemplified in (26) and (27).

(26) әŋkәatsini ә-kai guaŋ-mag-e
someone 1-house come-NEG-DECL
‘No one came to my house.’

(27) john қәatsini sa-mag-e
John something say-NEG-DECL
‘John says nothing.’ (Literally John does not say anything.)

9. Negative interjection ә

In Rongmei, there is a particle ә which functions as negative interjection. It frequently occurs as an answer to a question or to contradict a statement perceived to be incorrect. The
negative particle *ha* occupies the left-most position in a clause in Rongmei as illustrated in the following sentences (28)(b) and (29)(b):

(28) (a)  
\[-nąŋ \ kət'iam-\textit{təŋ} \ tat-bom-dai?\]  
you market-LOC go-PROG-IQ  
‘Are you going to the market?’

(b)  
\[-ha! \ ai \ kət'iam-\textit{təŋ} \ ta-bom-mag-e\]  
No, I market-LOC go-PROG-NEG-DECL  
‘No, I am not going to market.’

(29) (a)  
\[-ai \ niŋ-\textit{təŋ} \ naŋ \ tiŋku \ ručərok-na-\textit{təŋ} \ e\]  
I think-PST-DECL you age sixteen-only-PST-DECL  
‘I thought that you are just sixteen.’

(b)  
\[-ha! \ ai \ tiŋkum \ rupadai-e\]  
No, I age twenty-four-DECL  
‘No, I am twenty four.’

10. Negative strengthening

Negative statements in Rongmei are strengthened by the suffix *-zət* which is directly attached to the verb root followed by the negative marker *-mak*, then followed by the declarative suffix. In other words, *-zət* occurs between the verb root and the negative marker as schematised as follows: [\textit{verb} + \textit{INTS} + \textit{NEG} + \textit{DECL}]. It occurs only in a negative environment expressing intensification as is evident from the ungrammaticality of (30)(b).

(30) (a)  
\[-ai \ k^b-a \ tu-zət-mag-e\]  
I fish eat-INTS-NEG-DECL  
‘I do not eat fish at all.’

(b)  
\[-*ai \ k^b-a \ tu-zət-\textit{təŋ} \ e\]  
I fish eat-INTS-PST-DECL  
‘I ate fish at all.’

11. Conclusion

The present paper has explored some of the morphosyntactic aspects of negation in Rongmei Naga as spoken in the Barak Valley of Assam. The paper has shown that the negator *-mak* negates declarative clauses: verbal, negative indefinite, nominalised clause etc. In addition, there is a prohibitive suffix *-ri*. The negative suffixes *-mak* and *-ri* usually occur in the post-verbal position followed by clause final suffixes. The negative and the prohibitive *-mak* and *-ri* cannot co-occur. The negative indefinite construction is formed by suffixing *-mak* to the verb; indefinite pronouns remain the same in both affirmative and negative constructions. The present paper has also shown that the negative particle *ha* is merely employed to form the negative interjections clause initially. Finally, the negative statement is strengthened by the intensifier *-zət* in Rongmei negative construction.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC</td>
<td>Accusative</td>
</tr>
<tr>
<td>CAUS</td>
<td>Causative</td>
</tr>
<tr>
<td>DET</td>
<td>Determiner</td>
</tr>
<tr>
<td>DECL</td>
<td>Declarative</td>
</tr>
<tr>
<td>FUT</td>
<td>Future</td>
</tr>
<tr>
<td>GEN</td>
<td>Genitive</td>
</tr>
<tr>
<td>HAB</td>
<td>Habitual</td>
</tr>
<tr>
<td>HORT</td>
<td>Hortative</td>
</tr>
<tr>
<td>IMP</td>
<td>Imperative marker</td>
</tr>
<tr>
<td>INTS</td>
<td>Intensifier</td>
</tr>
<tr>
<td>IQ</td>
<td>Interrogative question</td>
</tr>
<tr>
<td>LOC</td>
<td>Locative</td>
</tr>
<tr>
<td>LSI</td>
<td>Linguistic Survey of India</td>
</tr>
<tr>
<td>NEG</td>
<td>Negative</td>
</tr>
<tr>
<td>NOM</td>
<td>Nominative</td>
</tr>
<tr>
<td>NMLZ</td>
<td>Nominalizer</td>
</tr>
<tr>
<td>PROG</td>
<td>Progressive</td>
</tr>
<tr>
<td>PROH</td>
<td>Prohibitive</td>
</tr>
<tr>
<td>PST</td>
<td>Past</td>
</tr>
<tr>
<td>PRF</td>
<td>Perfective</td>
</tr>
<tr>
<td>PFX</td>
<td>Prefix</td>
</tr>
<tr>
<td>SOC</td>
<td>Sociative</td>
</tr>
<tr>
<td>SFX</td>
<td>Suffix</td>
</tr>
<tr>
<td>1</td>
<td>First person pronominal prefix</td>
</tr>
</tbody>
</table>
References


6. Differential marking of cases in Asamiya

Runima Chowdhary
Gauhati University

Abstract
The paper attempts at a descriptive account of the intriguing system of differential case marking in Asamiya manifested by asymmetric and syncretic forms of both core and non-core cases. The study shows that the language syntactically follows nominative-accusative case system with a mixed surface ergativity in that absolutive is used as the null counterpart of both the overt core cases, nominative and accusative. The canonical marking on A, which is almost consistently governed by transitivity, is additionally conditioned phonologically, while that on S, it is determined by the degree of agentivity as encoded by the given predicate. The marking on P, on the other hand, is conditioned by an animacy hierarchy in addition to pragmatic factors. Both the canonical and non-canonical non-core cases are highly correlatable to the semantics of predicates, with a closed set being capable of assigning case lexically.

1. Introduction
Case is usually defined as ‘a system of marking dependent nouns for the type of relationship they bear to their heads’ (Blake 1994:1). Case marking is one of the three basic strategies that a language utilizes to indicate grammatical relations, the other being agreement and constituent order. Cross-linguistic typological studies have shown that morphological cases besides being correlatable to grammatical relations highly correlate with given semantic as well as pragmatic roles (Comrie 1981:117-130). The cases used to encode nuclear syntactic roles like typical complements of intransitive and transitive verbs are referred to as core cases, as contrasted with the rest, designated non-core (peripheral) cases bearing peripheral relations (Blake 1994:34, 119).

Following the lines of assumptions in Comrie (1981:104-105) and Croft (1990:10) the automated (i.e., institutionalized) syntactic roles expressed by variation in case-marking are as in Table 1.

<table>
<thead>
<tr>
<th>I</th>
<th>A</th>
<th>Subject of transitive/ditransitive clause</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>S</td>
<td>Subject of intransitive clause</td>
</tr>
<tr>
<td>III</td>
<td>P</td>
<td>Direct object of transitive clause</td>
</tr>
<tr>
<td>IV</td>
<td>G</td>
<td>Indirect object of ditransitive clause</td>
</tr>
<tr>
<td>V</td>
<td>T</td>
<td>Direct object of ditransitive clause</td>
</tr>
</tbody>
</table>

The paper attempts at a descriptive account of the assignment of various morphological cases in Asamiya. §2 is an overview of the problematic issues of identifying cases in Asamiya.

---

1 I would like to express my sincere gratitude to Prof. Scott DeLancey for his comments and suggestions on the first draft of the paper.
(popularly known as Assamese\(^2\)). §3 describes the canonical and non-canonical\(^3\) marking of core cases and the motivating factors underlying these. §4 deals with the marking of non-core cases used for encoding various semantic functions followed by concluding remarks.

2. An Overview

Asamiya is characterized by SOV as the unmarked constituent order. The subject of a sentence canonically agrees with the verb in terms of person, with the second person being further distinguished in terms of degrees of honorificity indicative of social status. The language exhibits an intriguing system of differential case marking manifested by asymmetric\(^4\) and syncretic forms making it problematic to ascertain the exact number of cases with their correlatable functions. Consequently, the few works dwelling on the subject have conflicting views making it all the more difficult even to draw a holistic overview on the topic. Following Sanskrit tradition seven cases are generally identified in the language with the postposition \(\text{pɔra} \) ‘from’ being accorded the status of ablative case (Barua 1984:89-98; Kakati 2007:302). While Kakati (2007:302-3) and Goswami (1982:277; 2008:152) have postulated null/absolute cases as counterparts of nominative and accusative for the absence of overt forms for subjects and objects respectively, others have identified the overt and null cases for subjects as ergative and absolutive\(^5\) cases respectively (Amritavalli and Sarma 2002). The syncretism of the instrumental case with that of the subject as shown in Goswami (1982:266) has been accounted for by the shared feature of agentivity in Amritavalli (ibid) leading her to postulate an ergative case. According to it, the language has ‘a mixed nominative-ergative case system with an ergative, an absolutive and an accusative case, but without a nominative’. It is important to note that non-nominative (or non-ergative) subjects have been completely overlooked till date in any works on the language. This study aims at examining all such issues and attempts to look for ways to account for the conflicting views in this regard.

The view taken here is that the language has only inflectional cases expressed by suffixation of markers on nominals and that, postpositions, as traditionally assumed, are not case markers. The case in Asamiya is used to indicate the syntactic relationship between a NP and a verb at the sentence level, and between two NPs or between a NP and a postposition at the phrase level. All nominals (with or without dependents) may be suffixed with morphological cases. However, all case marked forms are not nominal in function, but may be correlatable to various adverbial functions. Asamiya is classified as nominative-accusative type on the basis of the behavioral properties of accusativity: (i) S aligns with A, not with P; (ii) verbs agree with S and A, not with P. The language exhibits a mixed surface ergativity in that absolutive is used as the null counterpart of both the core cases nominative and accusative. It is worth noting that the core cases have non-canonical variants. The non-core cases include dative, genitive, instrumental, locative, accusative and oblique\(^6\). All case inflections are phonologically conditioned. It is important to note that the marking of cases is insensitive to grammatical categories of tense, aspect, modality as well as polarity.

\(^2\) Asamiya, pronounced /\ɔhɔmija/ in Standard Asamiya, stands for the Assamese language (as well as the people of Assam), where the phoneme /\ɔ/ as used in the word may have several dialectal or ideolectal variants, viz., /h/, /x/, /k\ɔ/ and /s/.

\(^3\) Canonical case marking refers to default (prototypical) case as contrasted to non-canonical case (Andrews 2001; Aikhenvald 2001).

\(^4\) Case asymmetry refers to application of non-identical sets of case marking on arguments sharing common semantic and functional characteristics in appropriate syntactic positions (Iggeson 2005).

\(^5\) Morphologically unmarked case is also referred to by some as nominative (Blake 1994:122).

\(^6\) The term ‘oblique’, originally used to refer to all non-nominative cases and in some studies to all non-core or adverbal cases (i.e., cases marked on nominalized verbs) (Blake 1994:203), is used in this study as a nomenclature for descriptive convenience to distinguish the case marking on the nominal complement of a postposition head from that of the syncretic genitive case marking on the nominal complement of a noun head in Asamiya.
3. Marking of core cases

This section deals with the case marking of subjects of transitive as well as intransitive sentences and objects of (mono-) transitive, and examines the factors determining their differential case markings.

3.1. Canonical marking of subjects

The A or the subject of a transitive verb in Asamiya is canonically marked by the nominative case realised as -e.7

(1) tomalok-e prɔb∧ndɔ-tɔ.Ø porh-is-a-ne
   2PL-NOM article-CL.ABS read-IPFV-2QP
   ‘Have you read the article?’

Here the consonant-ending personal pronoun in the subject position of the mono-transitive clause semantically functioning as an initiator/controller described by the verb takes the nominative case. The same holds true in case of a ditransitive verb as shown in (2).

(2) hihx-e hɔmpadɔk-ɔlɔi e-kɔn siti.Ø lik2-is-il-(e)8
   3PL-NOM editor-DAT one-CL letter.ABS write-IPFV-PST-(3)
   ‘They wrote a letter to the editor.’

(3) tɔɔ.Ø hɔmpadɔk-ɔlɔi e-kɔn siti.Ø lik2-is-il-(e)
   3SG.ABS editor-DAT one-CL letter.ABS write-IPFV-PST-(3)
   ‘He wrote a letter to the editor.’

However, a vowel-ending personal pronoun as A (3)(31)/(52) is canonically in absolutive case (Ø) following the phonological conditioning rule even in co-occurrence with the same predicate. From the example above, it can be inferred that the absolutive case on A is a variant of the nominative case (see footnote 5).

One of the features shared by all the three (1-3) is transitivity. As Observed by Hopper (Hopper et al 1980:251), ‘Transitivity involves a number of components, only one of which is the presence of an object of the verb.’ Another factor is the semantic specification of agentivity. Agentivity associated with the degree of control of the subject over the situation (Comrie 1978:356) in terms of various semantic features like force, movement, volition and wilfulness or purposefulness. The transitivity of the examples above with the agentive role of the subjects is explicitly clear.

It is worth noting here that the overt case marking on subjects co-occurring with transitive verbs remains intact even in a context of omission of the object. Some other frequently used transitive verbs (simple and conjunct) are hik ‘learn’, hun ‘listen’, mar ‘beat/kill’, pa ‘get

6. Differential marking of cases in Asamiya

7 The variants of the case marker -e, as shown in Amritavalli and Sarma (2002) are not phonological, but orthographic variants only.

8 The 3rd person marker -e either on a transitive or an intransitive verb, suffixed with the aspect plus past tense marker -is-il may be optional, e.g., parh-is-il-(e) ‘read/studied’, kʰa-is-il-(e) ‘ate’, di-is-il-(e) ‘gave’, ah-is-il-(e) ‘came’, por-is-il-(e) ‘fell down’.

However, the person marker is obligatory on a transitive verb suffixed only with past tense marker -il, e.g., parh-il-e ‘has read/studied’, kʰa-il-e ‘has eaten’, di-il-e, as contrasted to that of an intransitive, in which it is obligatory in some as in hu-il-e ‘slept’, kand-il-e ‘cried’ or may be optional in some as in ah-il-e ‘arrived’, por-il-e ‘fell down’. There are also some verbs which do not require the overt 3rd person marker in this form as in go-l/e *go-l-e (as in 10) ‘left’, ho-l/e *ho-l-e ‘became’.

109
Asamiya exhibits two types of causatives - direct and indirect, with the latter being differentiated morphologically as well as syntactically by the actual/potential presence of a fourth argument. Only the direct causatives are exemplified in this study.

The difference of case marking on A as in (1)(2)(3)(5) and S as in (4)(6)(7)(8) can be accounted for on the basis of the capacity of the verb to take objects. Semantically, the A is an initiator/controller, while the S of this subset is a patient/affected or theme which undergoes change of states encoded by its predicate. Some other intransitive verbs of this kind are mor ‘die’, hera ‘get lost’, p’ut ‘blown up’, nam ‘get down’, u’r ‘rise/get up’, bas ‘live/survive’,
Differential marking of cases in Asamiya

6. Differential marking of cases in Asamiya


That the result of a volitional act can be under control of an initiator is exemplified by (9) with a nominative case marked subject in contrast to (8) having a subject in absolutive in the patient role.

(9) lora-tɔ-e por-i di-l-e
    boy-CL-NOM fall-NF GIVE-PST-3
    ‘The boy (deliberately/purposefully) fell down.’

Here the vector or lexical aspectual auxiliary di ‘GIVE’ (Chowdhary 2008:225) used with the main verb lends the predicate an interpretation of volitionality of the subject which has control over the situation, thus triggering the suffixation of the overt case marker. However, the alternation of the vector by another like za- ‘GO’ (realized suppletively as go-) encodes loss of control by the subject. Subsequently, the subject remains unmarked.

(10) lora-tɔ.Ø por-i go-l
    boy-CL.ABS fall-NF GO-PST.3
    ‘The boy (accidently) fell down.’

The examples (9) and (10) are indicative of the distinction between the vectors in terms of the degree of agentivity inherent in them. The subject of a copular predicate in the language is unmarked for case, as it lacks an object as shown by (11).

(11) tekʰet.Ø hikkʰɔk as-il
    3SG.ABS teacher be-PST.3
    ‘He was a teacher.’

(12) ʰogɔban.Ø as-e
    God. ABS exist-3
    ‘There is God.’ (lit. ‘God exists’)

(13) bagisa-t sari-zɔpa am-ɡɔs.Ø as-il
    garden-LOC four-CL mango-tree.ABS exist-PST.3
    ‘There were four mango trees in the garden.’

Examples (14) and (15) exemplify S of another subset of intransitive verbs which is non-canonically marked for nominative case -e, as distinguished from the subset encountered earlier in (6)-(8).

(14) sowali-zoni-e ʰal nas-e
    girl-CL-NOM good dance-3
    ‘The girl dances well’.

---

10 Copular sentences are a type of intransitives, designated ‘complex-intransitives’ as opposed to ‘ordinary’ intransitives (Huddleston and Pullum 2002:53).

11 The other copular verb used in Asamiya besides as ‘be’ (11) is hɔ ‘be/become’ (Chowdhary 2011:199).
The distinctive feature of case marking on the subject of intransitive verb of this kind can be accounted for by its property of transitivity which presupposes the presence of an actual or underlying argument in the role of P. This is substantiated by (16) and (17) with the potential cognate objects nas ‘dance’ and kand ‘cry’ respectively fulfilling that requirement. Additionally, unlike the verbs ah ‘arrive’, za ‘go’, pɔr ‘fall’ (6)(7)(8) which describes some telic situations signaling end points, nas ‘dance’, kand ‘cry’ (with or without potential objects) imply atelic situations which have no such terminal points, but can be extended or stretched further indefinitely (Comrie 1976:44). The subjects are undoubtedly agentive in that they are in full control over the situation.

Some other intransitive verbs of this subset having nominative case marked subjects are dour ‘run’, zuz ‘fight’, hah ‘laugh’, kah ‘cough’, buz ‘understand’, zopia ‘jump’, hator ‘swim’, ga ‘sing’. Compare the following pair of sentences with analogous propositional content. The S in (18) contains a simple verb, while (19) has its conjunct counterpart shown in brackets. As evidenced, both the sentences have nominative case-marked S. The same also holds true for those conjunct verbs lacking simple counterparts, as exemplified by (20) and (21).

12 Creissels (2008) discusses similar cases of alignment variations motivating split intransitivity in languages.

13 Bihu is a popular folk dance of Assam.
6. Differential marking of cases in Asamiya

The conjunct verbs *agat pa* ‘get hurt’ and *pitɔn* *k̊a* ‘get spanking/beating’ composed of the agitative light verbs *pa* ‘find/get’ and *k̊a* ‘eat’ in conjunction with the nominal cores as in (20) and (21) carry transitivity property in that it presupposes the presence of objects. This requirement is satisfied by the nominal cores they have at their disposal as dummy objects, viz., *agat* ‘injury’ and *pitɔn* ‘beating/spanking’. Similar conjunct verbs requiring case-marked *S* are *duk̊a* *pa* ‘become sad’, *har pa* ‘get awake’ *boe k̊a* ‘get afraid’, *k̊a* *ɔz karo* ‘walk’ etc.

As contrasted with an agitative light verb, a patiitive light verb in a conjunct verb projects the subject as the patient/theme, which remains invariantly unmarked for case. Some patientive light verbs acting as integral constituents of the conjunct verbs are *ho* ‘become/be’ and *por* ‘fall’ as in *hec ho* ‘get exhausted’ (22) and *dora por* ‘get caught’ (23).

(22) *kagɔz.*[O] *[hec ho]-is-e*  
*paper.ABS [end be-IPFV-3]*  
‘The stock of papers has run out.’

(23) *sor-to.*[O] *[dɔr-a por]-il-(e)*  
*thief-CL.ABS [catch-NOML fall-PST.3]*  
‘The thief got caught.’

The language exhibits asymmetric case marking on the subject of matrix clause containing subjectless nonfinite embedded clause with either transitive or intransitive verb, as in (24) and (25) respectively.

(24) *nitin-e/[O] *[bɑt k̊a-i] kɔlez-loi go-l*[^14]  
*Nitin-NOM/ABS [rice eat-NF college-DAT go-PST.3]*  
‘Nitin has left for college after having his meal.’

(25) *nitin-e/[O] *[gɔr-loi za-boloi] bisar-is-e*  
*Nitin-NOM/ABS [home-DAT want-IPFV-3]*  
‘Nitin wants to go home.’

The subject of the matrix clause in (24) being *S* may have absolutive case, but may optionally take nominative case by being induced by the adjacent subject of the embedded transitive clause. The opposite holds for (25), where the subject of the matrix clause being *A* may be invariantly marked or may remain unmarked following the unmarked subject of the embedded intransitive clause. This statement can be reinforced by the following pair of sentences (26) and (27).

(26) *manuhandjɔn*[O] *[hɔt’at utk-i go-i] agɔr sok-i t boh-il-(e)*  
*man-CL.ABS [suddenly rise-NF GO-NF] front chair-LOC sit-PST.3*  
‘Getting up suddenly the man sat down on the front chair.’

(27) *proti bɔsɔr-e porisɛm mi so rave-bor*[O] *[bat-ɔt kotɔs-εi]*  
*each year-LOC migrant bird-PL.ABS way-LOC nowhere-EMP*  
*Neg-stop-NF here-DAT fly-ADVL come-3*  
‘Each year the migrant birds come here flying without stopping anywhere on the way.’

[^14]: *go-l* is suppletive form of the verb *za* ‘go’.
Besides the matrix verbs ‘sit’ and ‘come’ here, the verbs of the embedded clauses ‘go’ and ‘stop’ too being inherently patientive are incapable of assigning overt case on the subjects of the matrix clauses.

### 3.2. Non-canonical marking of subjects

Now moving on to the non-canonical marking of subjects, both A and S can be marked genitive case. However, unlike the canonically marked subject, non-canonical marked one does not trigger agreement with the verb.

The genitive case on subject is used to encode possessional relation, physiological state/sensation, psychological experience or inner feeling/emotion. The restricted set of simple verbs taking genitive subjects includes lag ‘feel’, as ‘possess’/‘have’ and ho ‘happen’/‘become’. Some of the conjunct verbs of this subset are mon za ‘intend/wish’, laz lag ‘be ashamed of’, boc lag ‘be afraid of’, hol lag ‘like’, bok lag ‘be hungry’, hop ho ‘be zealous’, bissah ho ‘believe’, hahi ut ‘feel like laughing’, kosto ho ‘undergo hardship’, hagol lag ‘get tired’, amoni lag ‘get bored’, tsponi dhr ‘be sleepy’ etc.

The subject status of the genitive case-marked argument is evidenced from the fact that like its corresponding nominative case-marked argument (28), it occurs in the sentence-initial position (29). Syntactically, the argument in the former is A, as contrasted to S in the latter. Another piece of evidence of its subject status is provided by the coordination of two sentences (30) and (31) with two differently marked subjects as (32), with one of its subjects being omitted under identity.

(28) deuta-k-e lora-tɔ-k [kʰɔŋ] kor]-is-e
father-33-NOM boy-CL-ACC anger do-IPFV-3
‘(His) father is rebuking his son.’

(29) deuta-k-ɔr lora-tɔ-r oprɔɔt [kʰɔŋ] ut]-is-e
father-33-GEN boy-CL-OBL at anger rise-IPFV-3
‘(His) father is getting angry at his son.’

(30) mʊ-r [mur ɡ̈ur-a]-is-il
1SG-GEN head spin-CAUS-IPFV.3
‘I was suffering from giddiness.’

(31) mœ.Ø rasta-t por-i go-is-il-ɔ
1SG.ABS street-LOC fall-NF GO-IPFV-PST-1
‘I fell down on the street.’

(32) mœ. Ø [mur ɡ̈ur]-a-i rasta-t por-i go-is-il-ɔ
1SG.ABS head spin-CAUS-NF street-LOC fall-NF GO-IPFV-PST-1
‘Suffering from giddiness I fell down on the street.’

As substantiated by (30), irrespective of the category of the subject the verb of this type is in default 3rd person form. The following complex sentence (33) exemplifies control of the genitive case marked subject of the matrix clause over the reflexive of the copular embedded clause as an evidence of its subject-hood.
6. Differential marking of cases in Asamiya

(33)  
\[\text{mɔ-r} \quad [\text{ijat} \quad \text{niz-ɔk} \quad \text{asohua} \quad \text{zen}] \quad \text{lag-e}\]  
1SG-GEN  here  REF-ACC  stranger  like  feel-3  
‘I feel as if I am a stranger here.’

Compare the pair of intransitive sentences with an analogous proposition.

(34)  
\[\text{soali-zɔn-i-e} \quad \text{hah-is-e}\]  
girl-CL-F-NOM  laugh-IPFV-3  
‘The girl is laughing.’

(35)  
\[\text{soali-zɔn-i-r} \quad \text{[hahi} \quad \text{ut}^{b}\text{-is-e}}\]  
girl-CL-F-GEN  laughter  rise-IPFV-3  
‘The girl feels like laughing.’

While the nominative case-marked S (34) is agentive, having control over the situation as in (36), the genitive case marked S (35) is not, as evidenced from (37).

(36)  
\[\text{soali-zɔn-i-e} \quad \text{[hahi-boloi} \quad \text{sesta} \quad \text{kor-is-e}}\]  
girl-CL-F-NOM  laugh-NF  try  do-IPFV-3  
‘The girl is trying to laugh.’

(37)  
\[*\text{soali-zɔn-i-r} \quad \text{[hahi} \quad \text{ut}^{b}\text{-iboloi} \quad \text{sesta} \quad \text{kor-is-e}}\]  
girl-CL-F-GEN  laughter  rise-NF  try  do-IPFV-3  

The following compound (38) and simple (39) sentences exemplify the use of the possessive verb as ‘have’/’possess’ invariably taking genitive case-marked subjects.

(38)  
\[\text{apona-r} \quad \text{doń-bɔl.Ø} \quad \text{as-e} \quad \text{zodiɔ} \quad \text{zɔnɔ-bɔl. Ø}\]  
2|SG-GEN  money-power.ABS  possess-3  although  man-power.ABS  
\text{n-a(s)-e}\]  
NEG-be-3  
‘Although you have got money-power, you lack manpower.’

(39)  
\[\text{mɔ-r} \quad \text{du-kɔm} \quad \text{gari.Ø} \quad \text{as-il}\]  
1SG-GEN  two-CL  car  possess-PST.3  
‘I had two cars.’

The obligatory genitive marking on subject in co-occurrence with the possessive verb as ‘possess’ as in (38) and (39) can be attributed to the ‘lexicality’\(^{15}\) of the verb. Dative is another non-canonical marking on subject encoding the semantic specification of goal (directional or spatial), benefactive or recipient. The marking of this case on A in co-occurrence only with the desiderative verb lag-‘want/ seek/ need/ desire’ is also attributable to the lexicality of the verb (40).

\(^{15}\) ‘Lexicality’ refers to the capacity of a particular verb to specify a specific case on some arguments (Davies 1988, cited in Andrews 2001).
The subject-hood property of the dative case-marked argument is substantiated by the omission of the coreferential subject of the second clause in the compound sentence (40). Another piece of evidence of its subject-hood is provided by (41), where the desiderative verb takes a subjectless nonfinite clause with a transitive verb as its complement.

(41)  
\[ mʊ-k [ k^b-a-boloi kiba.Ø] lag-e \]

1SG-DAT eat-NF something. ABS want-3

‘I want something to eat’.

The language exhibits differential case marking on the subject of an embedded complement clause, whereby the nominative case may be optionally alternated with the accusative case (42).

(42)  
\[ robin-e [nitin-e/ɔk sobi-loi kapɔr kin-a] dek^b-is-il-(e) \]

Robin-NOM Nitin-NOM/ACC Sobi- DAT dress buy-NF see-IPFV-PST-(3)

‘Robin saw Nitin buy a dress for Sobi.’

The alternation of nominative case by accusative is used as a functional device to distinguish the subject of the embedded clause from the subjacent similar case marked subject of the matrix clause. Significantly, the accusative case marking on subject semantically identifies a referent to be in contrastive focus in comparison to that of a nominative case marked one in a corresponding sentence.

That syntactically the accusative case-marked argument in (42) is not the object of the matrix clause, but the subject of the embedded clause, can be evidenced from its control over reflexivization (43).

(43)  
\[ robin-e [nitin-ɔk niz-oloi kapɔr kin-a] dek^b-is-il-(e) \]

Robin-NOM Nitin-ACC REF-DAT dress buy-NF see-IPFV-PST-(3)

‘Robin saw Nitin buy a dress for himself (i.e. Nitin).

3.3. Marking of direct objects

The accusative case is semantically correlatable to patient, theme, undergoer or affected. The canonical accusative case marking on P has two variants, one of which is realized as -k or -ɔk and the other is null/ absolutive. The overt marker is phonologically conditioned, with a vowel-ending argument taking -k and a consonant ending argument taking -ɔk. Dative may be used as a non-canonical variant of accusative case (as in 53). The language exhibits a hierarchy of animacy in marking of accusative case. Additionally, pragmatics plays a role in the selection of a variant. Let us consider the following examples.

(44)  
\[ mœ.Ø teø-k dek^b-is-ɔ \]

1SG.ABS 3SG-ACC see-IPFV-1

‘I have seen him.’
6. Differential marking of cases in Asamiya

As seen, the personal pronoun and the proper name of a person as in (44) and (45) respectively, being identifiable and hence, definite, are obligatorily marked for overt accusative case. However, a human common noun as in (46) may have split case marking in that it may be optionally marked or unmarked for accusative case. The motivating factor is pragmatic, which depends on whether the referent is considered by the speaker as something specific/salient in contrast to general, or topic worthy or not.

(46) tevlok-e manuh-zon-ɔk/Ø dekʰ-is-e
3pPL-NOM man-CL-ACC/ABS see-IPFV-3
‘They have seen the man.’

On the other hand, a non-human referent as in (47) is invariantly unmarked for accusative case. But a proper name used to refer to a non-human like ‘cow’, the overt case is obligatory as in (48).

(47) tevlok-e goru-zoni-*k/Ø dekʰ-is-e
3pPL-NOM cow-CL-ACC/ABS see-IPFV-3
‘They have seen the cow.’

(48) tevlok-e kazoli-k dekʰ-is-e
3pPL-NOM Kajoli-ACC see-IPFV-3
‘They have seen Kajoli.’

(49) tevlok-e tazmohol-*ok/Ø dekʰ-is-e
3pPL-NOM Tajmahal-ACC/ABS see-IPFV-3
‘They have seen the Tajmahal.’

(50) tevlok-e sobi-kʰon-*ok/Ø dekʰ-is-e
3pPL-NOM picture-CL-ACC/ABS see-IPFV-3
‘They have seen the picture.’

An inanimate referent of any sub-class of noun is invariantly unmarked, irrespective of whether it is a proper name or common noun, as in (49) and (50). The data above exhibiting case marking on direct objects reveal that a human common noun is more animate than a non-human or an inanimate common noun, while a pronoun or an animate proper noun is more animate than a human common noun, as represented by Table 2.

However, an inanimate referent can be marked for accusative case when it is intended to be presented as a topic worthy element, as exemplified by the following slogan against tobacco used in a billboard. The accusative case canonically used to represent P may be alternated with dative in the same syntactic environment to encode change of its semantic role from patient (52) to that of source (53).
Table 2: Animacy hierarchy

<table>
<thead>
<tr>
<th>Pronoun/Animate proper noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human common noun</td>
</tr>
<tr>
<td>Non-human /Inanimate common noun</td>
</tr>
</tbody>
</table>

(51)  
\( \text{zibən-ok} \quad \text{adər-øk} \quad \text{dəpat-øk} \quad \text{nə-hə-e} \)  
Life-ACC \quad Welcome-IMP\(_1\) \quad Tobacco-ACC \quad Neg-be-3  
‘Welcome life, not tobacco’.

(52)  
\( \text{hi} \quad \text{deutak-øk} \quad [\text{həe} \quad \text{kər}]\)-e  
3SG father-ACC \quad Fear \quad Do-3  
‘He is afraid of his father.’ (Lit. ‘He fears his father.’)

(53)  
\( \text{hi} \quad \text{deutak-əloi} \quad [\text{həe} \quad \text{kər}]\)-e  
3SG father-DAT \quad Fear \quad Do-3  
‘He is afraid of his father.’

3.4. Marking of core cases revisited

As inferred from the discussion above, the canonical marking of nominative case on A is syntactically determined by the transitivity inherent in the predicate, while that on S is governed by the intrinsic property of a select set of predicates. The data presented above show that there are two kinds of predicates of S, one is P-like or patientive predicate, the subject of which is canonically in absolute (4)(6)(7)(8) and the other is A-like or agentive predicate with a nominative case-marked subject. The agentive predicates of S fall into three categories—(i) which co-occurs with an agentive vector as in (9); (ii) which is potentially capable of having a cognate object as in (14) and (16) and (15) and (17); and (iii) which has an agentive light verb as its constituent as in (18)(19)(20) and (21). Additionally, the semantic property of telicity encoded by its predicate has an important role to play in assigning case to the subject\(^{17}\). As for the non-canonical marking of A by genitive or dative case and S by dative, it is found to be conditioned by lexical and/or semantic specification of a subset of predicates.

The asymmetrical accusative case marking on P can be accounted for by animacy hierarchy, correlatable to nominals. On the highest scale of hierarchy lie the pronouns and proper names of human referents that are invariantly marked, on the lowest scale lie the common nouns encoding inanimate referents that are invariantly in absolutive. In between lie the common nouns, encoding human or inanimate referents, the optional marking of which is determined by speakers pragmatically.

Notwithstanding the consistent morphological case marking on A (excepting vowel-ending pronouns) in contrast to that on S and P, which may be marked or unmarked, both A and S are syntactically treated alike, as substantiated by the following compound sentences with S and A, one being omitted under identity.

\(^{16}\) In some linguistic literature the intransitive verbs with case-marked subjects are designated unergative, while those with unmarked subjects are termed unaccusative (Amritavalli and Sarma 2002).

\(^{17}\) Georgian and Nepali are the only languages reported so far to have split intransitivity conditioned by telicity (Creissels 2008:149).
6. Differential marking of cases in Asamiya

(54) daktør. Ø ah-il aru rogi-k [porikk*ja kor]-il-e
  doctor. ABS come-PST.3 and patient-ACC examination do-PST-3
  ‘The doctor came and examined the patient.’

4. Marking of non-core cases

The accusative case -(o)k or its unmarked variant used for marking core case on P (direct object of transitive verb) is also used for marking non-core case on T (direct object of ditransitive verb). The dative is another non-core case which is marked on G (indirect object of a ditransitive verb), encoding semantic role of recipient or benefactive or directional/temporal goal. The dative case is realized as -(o)loi or -(o)k, the latter being syncretic with the accusative case. The indirect object precedes the direct object in the unmarked order. The following exemplifies the marking of accusative case on T and dative case on G.

(55) ma-k-e mastør-øk lora-το-k/Ø hop-il-e
  mother-3SG-NOM teacher-DAT boy-CL-ACC/ABS entrust-PST-3
  ‘(His) mother entrusted (the custody of) the boy to the teacher.’

Following the animacy hierarchy the human common noun lora-το ‘the boy’ in the role of T as in (55) may or may not be overtly marked for accusative case. But a proper noun (56) used as T is marked overtly for the same.

(56) ma-k-e mastør-øk pona-k hop-il-e
  mother-3SG-NOM teacher-DAT Pona-ACC entrust-PST-3
  ‘(His) mother entrusted (the custody of) Pona to the teacher.’

Any problem arising out of interchange of positions of G and T marked with syncretic cases can be resolved pragmatically. The sentence (57) exemplifies the dative case -loï marked on a proper noun ‘Sobi’ in the role of G, with the inanimate referent ‘letter’ in the role of T being invariably unmarked. The dative case encoding directional and temporal goal/destination respectively is also used to mark adjuncts of verbs as in (58) and (59).

(57) nitin-e sobi-loï sit h i. Ø likh-is-il-(e)
  Nitin-NOM Sobi-DAT letter.ABS write-IPFV-PST-(3)
  ‘Nitin wrote letters to Sobi.’

(58) ami. Ø kailoi dibrugær-oïloï za-m
  1PL.ABS tomorrow Dibrugarh-DAT go-FUT.1
  ‘We will go to Dibrugarh tomorrow.’

(59) tumi. Ø pas-bøza-loï bat sa-bɔ na-lag-e
  2SG.ABS five-sound-DAT way see-NF NEG-need-3
  ‘You don’t need to wait till five o’clock.’

Locative is another non-core case which is marked on a referent used as an adverbial complement or adjunct encoding spatial/temporal location or source. It is realized as -t or -ɔt as exemplified by (60) and (61).

(60) robin-ør gør.Ø guwahati-t
  Robin-GEN house. ABS Guwahati-LOC
  ‘Robin’s home is in Guwahati’.
The locative case may be optionally alternated with another variant -e in the same environment to encode directional location. The locative case marked referent in (63) expresses source or cause. When a natural phenomenon like wind, storm, rain, lightning or flood is considered as an agentive force instead of a source, the locative case (64) may be alternated with nominative (65).

(62) himalœ porbot. Ø barot-er uttor-ot/e
Himalaya mountain.ABS India-GEN north-LOC
‘The Himalaya mountain is in the north of India.’

(63) kot²-o-t kot² a. Ø barh-e
gossip-LOC gossip.ABS increase-3
‘Gossip begets gossip’, (lit. ‘Talk increases from talk.’)

(64) botah-ot gos-zo pa. Ø bag-il
wind-LOC tree-CL.ABS break-PST.3
‘The tree got broken due to the wind.’ (lit. The tree broke from the wind.’)

(65) botah-e gos-zo pa. Ø baŋ-il-e
wind-NOM tree-CL.ABS break-PST-3
‘The wind broke the tree.’

The genitive as a non-core case is used to encode possessorial relationship between a dependent nominal and the head noun in a NP (66) or between a subject and the predicative complement in a copular clause (67). The phonologically conditioned instrumental case is realized as -(e)re as in (68) and (69).

(66) mœ -r kitap
1SG-GEN book
‘My book.’

(67) kitap-k² on. Ø mœ -r
book-CL. ABS 1SG-GEN
‘The book is mine’.

(68) am-tœ. Ø ei-k² on kotari-re kat-a
mango-CL.ABS this-CL knife-INS cut-IMP₂
‘Cut the mango with this knife’.

(69) sobi-bœr. Ø hat-ere nu-su-b-i
picture-PL.ABS hand-INS NEG-touch-FUT-2₃
‘Don’t touch the pictures with (your) hands.’

It is worth noting here that the instrumental case -(e)re may be alternated with its reduced form -e, syncretised with the nominative marker as in hat-ere > hat-e, used in
compound adjectives, e.g., hat-e bɔa (kapɔr) ‘hand-woven (cloth)’, hat-e kɔta (huta) ‘hand-spun (yarn)’\(^{11}\). The study considers the -r/ɔr- ending form, distinguished from its syncretized genitive case, as oblique case on the basis of their syntactic and semantic distinctions. Syntactically, as contrasted with the genitive case used to encode the syntactic relationship between the dependent noun and the head noun in a NP as in (66), the oblique case is invariantly imposed by the head postposition to its dependent noun to encode the dependent-head relationship in a PP. Semantically, unlike that of a genitive case which is used to encode possessional or ownership relationships or analogous extended interpretations between two referents in a NP, the oblique case encodes an orientational relationship between a referent within a PP with another across the phrase\(^{19}\).

(70) hi. Ø dilli-r pɔra ah-il
3SG.ABS Delhi-OBL from come-PST.3
‘He returned from Delhi.’

(71) tɔma-r proti ama-r ast\(^b\)a. Ø as-e
2SG-OBL for 2PL-GEN faith. ABS be-3
‘We have faith in you’ (lit. ‘Our trust is in you’).

The oblique case-marked forms here have co-occurred with postpositions pɔra ‘from’ (70), proti ‘for’ (71). The nominative case marked subject of an active (72) is relegated to an oblique case marked complement of the postposition head dara ‘by’ in its passive counterpart (73).

(72) porikk\(^b\)ok-e uṭṭor-bohi-homuh. Ø punorikk\(^b\)on kor-il-e
examiner-NOM answer-script-PL.ACC re-examination do-PST-3
‘The examiner re-examined the answer scripts.’

(73) porikk\(^b\)ok-ɔr dara uṭṭor-bohi-homuh. Ø punorikk\(^b\)on kor-a ho-l
examiner-OBL by answer-script-PL.ACC re-examination do-NF pass-PST.3
‘The answer scripts were re-examined by the examiner.’

5. Conclusion

The differential marking of cases in Asamiya resultant from their asymmetric and syncretic morphological manifestations is too complex a topic with each issue worth investigating in details. An attempt has been made in this study to throw light on such issues relating to cases from various perspectives.

Table 3: Marking of cases

<table>
<thead>
<tr>
<th>Syntactic role</th>
<th>Canonical case</th>
<th>Non-canonical case</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>NOM, ABS</td>
<td>GEN, DAT, ACC, LOC</td>
</tr>
<tr>
<td>S</td>
<td>ABS</td>
<td>NOM, GEN</td>
</tr>
<tr>
<td>P</td>
<td>ACC, ABS</td>
<td>DAT</td>
</tr>
</tbody>
</table>

\(^{18}\) The claim (Amritavalli and Sarma 2002) that the use of -e as instrumental case is suggestive of its link to agentivity is untenable, as evidenced from the unacceptability of the given case marked nominal as the agentive subject, e.g., ‘hat-e kapɔr bɔa-ɛ ‘The hand weaves clothes!’ or ‘hat-e huta kɔt-ɛ ‘The hand spins yarn!’

\(^{19}\) Following the assumptions in Huddleston and Pullum (2002:648) we assume the orientational relationship to be a relation between a ‘landmark’ (the reference point) and a ‘tranjector’ (the item whose location is specified).
The study reveals that grammatical relations play a less significant role in marking of cases than semantic and/or pragmatic roles. As observed, all the core arguments may be marked or unmarked for cases. The marking on A, which is almost consistently governed by transitivity, is additionally conditioned phonologically, while that on S, it is determined by the degree of agentivity encoded by the given predicates and on P, by the animacy hierarchy in addition to pragmatic factors. The study unravels that the selection of the non-canonical core cases is based on the semantic parameter of the predicates, with a closed set being capable of assigning case lexically. Both the canonical and non-canonical non-core cases are highly correlatable to the semantics of predicates. The case marking on core arguments, distinguished as canonical and non-canonical may be summed up schematically as in Table 3.

On the basis of the analysis of the data on core cases, both canonical and non-canonical, it can be inferred that underlying the differential morphological case marking on core arguments, Asamiya syntactically follows a Nominative-Accusative case system.
6. Differential marking of cases in Asamiya

**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS</td>
<td>Absolutive</td>
</tr>
<tr>
<td>ACC</td>
<td>Accusative</td>
</tr>
<tr>
<td>ADVL</td>
<td>Adverbial Marker</td>
</tr>
<tr>
<td>CAUS</td>
<td>Causative</td>
</tr>
<tr>
<td>CL</td>
<td>Classifier</td>
</tr>
<tr>
<td>DAT</td>
<td>Dative</td>
</tr>
<tr>
<td>EMP</td>
<td>Emphatic</td>
</tr>
<tr>
<td>F</td>
<td>Feminine</td>
</tr>
<tr>
<td>FUT</td>
<td>Future</td>
</tr>
<tr>
<td>GEN</td>
<td>Genitive</td>
</tr>
<tr>
<td>IMP1</td>
<td>Imperative Honorific</td>
</tr>
<tr>
<td>IMP2</td>
<td>Imperative Familiar</td>
</tr>
<tr>
<td>INS</td>
<td>Instrumental</td>
</tr>
<tr>
<td>IPFV</td>
<td>Imperfective</td>
</tr>
<tr>
<td>LOC</td>
<td>Locative</td>
</tr>
<tr>
<td>NEG</td>
<td>Negative</td>
</tr>
<tr>
<td>NF</td>
<td>Nonfinite</td>
</tr>
<tr>
<td>NOM</td>
<td>Nominative</td>
</tr>
<tr>
<td>NP</td>
<td>Noun Phrase</td>
</tr>
<tr>
<td>NOML</td>
<td>Nominal Marker</td>
</tr>
<tr>
<td>OBL</td>
<td>Oblique</td>
</tr>
<tr>
<td>PL</td>
<td>Plural</td>
</tr>
<tr>
<td>PST</td>
<td>Past</td>
</tr>
<tr>
<td>QP</td>
<td>Question particle</td>
</tr>
<tr>
<td>SG</td>
<td>Singular</td>
</tr>
<tr>
<td>1</td>
<td>1st person</td>
</tr>
<tr>
<td>3</td>
<td>3rd person</td>
</tr>
<tr>
<td>1&lt;sub&gt;1&lt;/sub&gt;</td>
<td>1st person honorific</td>
</tr>
<tr>
<td>2&lt;sub&gt;2&lt;/sub&gt;</td>
<td>2nd person familiar</td>
</tr>
<tr>
<td>2&lt;sub&gt;3&lt;/sub&gt;</td>
<td>2nd person ordinary</td>
</tr>
<tr>
<td>3</td>
<td>3rd person</td>
</tr>
<tr>
<td>3&lt;sub&gt;2&lt;/sub&gt;</td>
<td>3rd person familiar</td>
</tr>
<tr>
<td>3&lt;sub&gt;3&lt;/sub&gt;</td>
<td>3rd person ordinary</td>
</tr>
<tr>
<td>Ø</td>
<td>Null case</td>
</tr>
</tbody>
</table>
References


7. Non-nominative Subjects in Assam Sadri

Lucky Dey
Dept of English and Foreign Languages, Tezpur University

Abstract

Syntactic categories such as ‘subjects’ often do not map clearly onto semantic roles such as ‘agents’ even though a prototypical subject may be an agent and a prototypical object may be a patient. Assam Sadri is unique in the context of other Indo-Aryan languages in that it has evolved as a ‘link language’ and is not a direct descendent of Sanskrit in the strict sense. Nonetheless, like other Indo-Aryan languages, Assam Sadri (AS) marks subjects with a variety of cases, often reflecting local case roles. AS is also a typical Indo-Aryan language in that it has SOV word order, uses postpositions, and marks verbal categories through suffixation. This chapter presents an overview of AS non-nominative subjects. Dative, genitive and locative subjects are discussed in contrast to nominative subjects.

Citation

Volume Editors
Gwendolyn Hyslop, Linda Konnerth, Stephen Morey, Priyankoo Sarmah

Copyright © 2014, the author(s), release under Creative Commons Attribution license

1. Introduction

In clause structure, the subject is higher than the other arguments both syntactically and semantically. We say it is ‘higher’ because the ‘subject’ mostly occurs at clause initial position and also because it is notionally ‘what is being talked about’ (Palmer 1994). In this way, subjects are considered higher than objects and obliques. The hierarchy of the grammatical relations as proposed by Comrie (1976) is shown in example (1).

(1) Subject>Direct object>indirect object>oblique

Semantically, the role associated with the subject is usually that of an actor or agent. However, subject position can also be occupied by non-agents. Thus, the semantic roles of the argument selected for the subject are also organized in a hierarchy. Fillmore (1967) proposed the thematic hierarchy as shown in (2).

(2) Agent<experiencer<instrument<object<source<goal<location< time

The argument with the semantic role of agent will be higher in the hierarchy than the other oblique arguments and occupies the subject position. In this connection Dowty (1991) distinguishes between Proto-agent and Proto-patient properties of arguments in a clause structure. He puts forth the Argument Selection Principle in the following lines.

In predicates with grammatical subject and object, the argument for which the predicate entails the greatest number of Proto-agent properties will be lexicalized as the subject of the predicate; the argument having the greatest number of Proto-patient entailments will be lexicalized as the direct object. (Dowty 1991:576)

---

1 I would like to thank my informants Mr. Rajen Rohidas, Mr. Pradeep Hemrom Luke Horo and others. I am grateful to my supervisor Prof. M. Barbora for her guidance and support. I also extend my sincere thanks to Asst. Prof. Gatuum Borah, Dept of EFL, Tezpur University, for his kind suggestions.
Dowty (1991) proposes the following entailments the agent Proto-Role:

(3)  
(a) Volitional Involvement in the event or state: John is being polite to Bill/ is ignoring Mary (cf. Dowty 1979:164-66)

(b) Sentience (and/or perception): John knows/believes/ is disappointed at the statement.

(c) Causing an event or change of state in another participant: His loneliness causes his unhappiness.

(d) Movement (relative to the position of another participant): The bullet overtook the arrow.

(e) Exist independently of the event named by the verb: John needs a car.

Among the above mentioned entailments, Dowty (1991:578) states that ‘at least one Proto-agent entailment, in absence of any of the P-patient entailment is enough to qualify an argument for subjecthood, and conversely with Proto-patient entailments for objects.’

As far as the subject position in Indic languages are concerned, we have seen that have both agentive and non-agentive semantics. According to Masica (1991:340), the subject position, in Indo-Aryan languages, can be occupied by an agent and also by other non-agents, which normally do not have verbal agreement\(^2\). The language under study, being an Indo-Aryan language, has a similar phenomenon. The present paper discusses subjects in Assam Sadri in general and the non-nominative subjects in particular. This includes a detailed discussion of the grammatical marking and semantic relationship of the predicate with these arguments. In addition, it attempts to show that the grammar of argument marking of the non-nominative subjects is not arbitrary but it is organized in a hierarchy.

2. **Assam Sadri and its typological features**

Sadani/Sadri, originally the mother tongue of the Sadans\(^3\), evolved as a *link language* of the heterogeneous Adivasis like, Munda, Kharia, Oroan and so on, mainly living in and around Chota Nagpur Plateau\(^4\) (P.S. Navarangi 1965:5). These communities, mostly belonging to three language families namely, Austro- Asiatic, Indo-Aryan and Dravidian used Sadri as their *link language* for inter and intra community communication. From pidgin, Sadri gradually evolved as a creole, primarily due to inter community marriages between the various linguistic groups amongst them.

When the British tea planters brought these adivasis to Assam, as labourers, in the 19\(^{th}\) century, Sadri as link language came with along with them. Over the period of two hundred years, Sadri came under tremendous influence of the dominant regional languages: Assamese and Bangla. Sadri as spoken in Assam has been labeled as *Assam Sadri* (henceforth AS) to distinguish it from the Sadri spoken in Chota Nagpur Plateau better known as *Nagpuria Sadri*.

\(^2\) Keenan (1976:316) states that verb agreement ‘fails to be a necessary condition basic-subjecthood since in many languages verbs agree with no NP… in a very few case verbs may agree with object… particularly in Hindi’.

\(^3\) According to Navarangi (1965), Sadans were an Aryan group amongst the non-Aryans in Chota Nagpur Plateau.

\(^4\) Chota Nagpur Plateau spreads over present day Bihar, Jharkhand, Chattisgarh and parts of West Bengal and Orissa.
The typological description of AS\(^5\) brings forth its Indo-Aryan features. It has Subject Object Verb word order and verbs are inflected for Tense Aspect and Mood. Postpositions are free markers.

3. Assam Sadri and its typological features

In AS, the subject of both the transitive (A) and intransitive (S) verb takes the nominative case, as in, (4) and (5). The nominative case remains unmarked (Ø) in the language.

(4) birsi ɛtowa-pit l-ak
birsi.NOM etowa-ACC beat-PAST-3SG
‘Birsi beat Etowa.’

(5) chõra-tho kand-q h-e
boy-CL.NOM cry-PROG AUX-PRES.3SG
‘The boy is crying.’

In (4) birsi ‘Birsi’ is the subject that takes the nominative case. In (5) chõra-tho ‘boy-CL’ is the subject of the intransitive verb taking nominative case. In both the sentences the subject bears the grammatical role of ‘agent’.

In AS, nominative subject usually receives the agent thematic role. The dynamic predicates like pit ‘beat’ and similar action verbs like mar ‘hit’, bol ‘tell’ and so on attribute proto-agent properties like volitionality, causation, and independent existence to the subject. Verbs of perception like dekh can have both the volitional interpretation of ‘look’ and the non-volitional interpretation of ‘see’. The former involves a sense of sentience and thus may assign the agent role to the subject. The latter, however, does not; may assign the thematic role of experiencer to the nominative subject case.

(6) 3SG.NOM chõri-tho ke dekh-l-ð/ djan bødgh ke
3SG.NOM girl-CL ACC see-PAST-1SG/ deliberately
see-PAST-1SG
‘He saw the girl/ He looked at the girl deliberately’

The volitional interpretation of dekh is facilitated by use of the adverb ‘deliberately’. The same is true for psychological (psych) verbs such as ‘fear’. In case of stative verb like ‘know’ and the copular verb, the subject normally takes nominative case, whereas, the thematic role is not that of proto-agent. In such constructions we see that the two-place predicates have a subject in the nominative case and the object in the nominative or absolutive case. In these cases the subject is the one ‘who is being talked about’ and the object ‘what is being predicated or talked about with relation to the subject’. Focusing mainly on such subjects, it is observed that they are different from proto-agents and experiencers of the action named by the verb. Thus, slightly different to the way the terms used elsewhere in linguistics, such subjects are referred to here as theme and their objects as rheme.

---

\(^5\) Language data have been mainly collected from Sadri speakers of Lakhimpur and Sonitpur districts in Assam.
The copula verbs like *hek* ‘COP.PRES’ are treated separately from typical action verbs, as they do not entail the proto-agent properties of volition and causation. They mostly refer to notions of ‘being’ and ‘existence’ and belong to a group of stative verbs. The thematic role of such nominative subjects can be that of themes, and the thematic role of the objects can be that of rhemes.

Thus, in AS, the thematic roles of nominative subject can be arranged in a hierarchy where the agent is higher than the experiencer followed by the themes Figure 1.

![Figure 1: Thematic hierarchy of Nominative subject in AS](image)

The hierarchy of the thematic roles shows the degree of control or agentivity of the nominative subjects as continuum between dynamic and stative verbs, where the dynamic verbs outranks perception verbs followed by psych verbs, and finally stative verbs, in agentivity.

4. Non-nominative subjects

In AS, besides the usual nominative case, the subject NP may also receive various non-nominative case markers, such as dative, genitive, and locative. These subjects, though they occur at the sentence initial position, normally do not have the verbal agreement and in this way differ from canonical subjects in AS. They bear the role of an experiencer, possessor and goal.

In §3, it is seen that agentivity is mostly associated with dynamic verbs and stative verbs ensure no agentivity. Unlike the Proto-agent which implies control over the action, non-nominatives do not exert control and are expected to occur with stative verbs.

The following sub-sections discuss the non-nominative subjects with various types of predicates and the semantic differences between those NPs which take nominative case and those which take non-nominative case.

4.1. Dative subject

In AS, subjects can take the dative case marked by *ke*, which is assigned to the subject by stative verbs whereas nominative case is assigned by the dynamic counterparts of these verbs. This difference is expanded upon below.

4.1.1. Stative versus dynamic

The difference between arguments with nominative and those with dative case can be illustrated with the use of dynamic and stative predicates respectively:
7. Non-nominative subjects in Assam Sadri

(8) (a)  
\[
\begin{array}{llllll}
\text{mɔi} & \text{phəl} & \text{bhal} & \text{pa-ʃ} & \text{h-ð} \\
\text{1.NOM} & \text{flowers.NOM} & \text{good} & \text{get-PROG} & \text{AUX.PRES-1SG}
\end{array}
\]
‘I like flowers.’

(b)  
\[
\begin{array}{llllll}
\text{biriş} & \text{ke} & \text{phəl} & \text{bhal} & \text{lag-ɛ-la} \\
\text{1.NOM} & \text{DAT} & \text{flowers.NOM} & \text{good} & \text{feel-1SG-IMPERF}
\end{array}
\]
‘Birsi likes flowers.’

The subject of the dynamic verb \( \text{bhal pa-ʃ h-ð} \) ‘liking’ (lit. ‘good get-PROG’) in (8)(a) is in nominative case. While, in (9)(b) the verb \( \text{bhal lag-ɛ-la} \) ‘like’ (lit. ‘good feel-3SG-IMPERF’) indicates a state and requires the subject NP to be in the dative case. The former subject is an agent and the latter is an experiencer.

In AS, predicates like the stative verb \( \text{lag} \), expressing ‘want’ or ‘need’, can have a dative subject. The dynamic counterpart of the verb \( \text{khodʃ} \) ‘want/search’ takes the nominative subject. This is exemplified in (9)(a)-(b).

(9) (a)  
\[
\begin{array}{llllll}
\text{mɔi} & \text{dher} & \text{pɔisa} & \text{khodʃ-ət} & \text{h-ð} \\
\text{1.NOM} & \text{much} & \text{money} & \text{want-PROG} & \text{AUX.PRES-1SG}
\end{array}
\]
‘I am searching for/wanting lots of money.’

(b)  
\[
\begin{array}{llllll}
\text{mɔ-ke} & \text{dher} & \text{pɔisa} & \text{lag-ɛ} \\
\text{1-DAT} & \text{much} & \text{money} & \text{want.PRES-3SG}
\end{array}
\]
‘I want lots of money.’

In (9)(a), the dynamic verb \( \text{khodʃ} \) ‘want’ is followed by the progressive marker \(-ət\) and the auxiliary in the present tense \( \text{h-ð} \) gets 1st person singular agreement marker \(-o\). The verb \( \text{lag} \) ‘want’ in (9)(b) takes the default 3sg agreement marker \(-ɛ\) and assigns dative case to the subject \( \text{mɔ-ke} \) ‘1-DAT’. Note in both (8) and (9) the verb which conditions nominative marking on the subject is also marked for progressive aspect.

4.1.2. Volitional vs non-volitional

In AS, there are certain dynamic verbs that denote non-volitional actions on the part of the dative subject. The volitional counterparts of these dynamic verbs take nominative subjects.

4.1.2.1 Verbs of perception

Experiencers are also marked by nominative in case of perception verbs such as \( \text{dekh} \) ‘see’ and \( \text{sun} \) ‘hear’.
In (10)(b) the use of the light verb de ‘give’ with the perception verb sön ‘hear’ requires a dative subject while when sön ‘hear’ is not followed by the light verb de ‘give’ the subject takes nominative case. The difference in subject case marking can be explained with the notion of control or agentivity. The perception verb sön ‘listen’ in (10)(a) can be a voluntary action where the subject has some control over the action and therefore takes nominative case. However, the subject of the compound verb sön-ai de ‘hear give’ in (10)(b) is definitely an experiencer, as it refers to an involuntary action where the subject does not have any control over the action. Experiencers are involuntary recipients of the action of the verb and hence are non-agentive (Cole 1983).

The use of light verbs plays an important role in the unified semantics of the compound verb structure in Indo-Aryan languages. Paul (2003:4) in this connection identifies twelve light verbs or V2 in Bangla, an Indo-Aryan language which ‘add subtle nuances to the overall meaning of compound verb sequence’. However, the focus of the present paper is on the use of light verbs in compound verbs and conjunct verbs sequence that affects the argument marking.

4.1.2.2. N+V Conjunct

Predicates which require a dative subject can form an N+V conjunct with verbs like, aa ‘come’ preceded by a noun. The N+V conjunct here, mainly, indicates non-volitional action.

(11) (a) mə- sən-l-ə  yaad  kar-ət
     1.NOM sën-PAST-1SG  remember  do-PROG
     h-ə
     AUX.PRES-1SG
     ‘I am remembering him.’

(b) mə-ke  sən-ai  de-l-ək
     1-DAT  sën-PERF  give-PAST-3SG
     ‘I heard.’

In (11)(a) the light verb kar ‘do’ has a nominative subject whereas, in (11)(b) aa ‘come’ has a dative subject, when used in the N+V conjunct. The conjunct yaad kar ‘remember do’ refers to a volitional act and yaad aa ‘remember come’ implies non-volitional act.

4.1.2.3. Psych verbs

In AS, the dative subject can occur with the use of certain psych verbs such ‘to feel’, ‘to like’, to be frightened’ or verbs expressing other mental or physical states like ‘hunger’, ‘fever’, ‘sleep’ and so on. These verbs appear in progressive aspect, essentially having a stative reading. Consider (12).
7. Non-nominative subjects in Assam Sadri

(12) mɔ-ke bɔkhar lag-ɑ̃ h-ɛ
1-DAT fever feel-PROG AUX.PRES-3SG
‘I have fever.’
‘Lit: I am feeling fever’

In (12) the verb lag ‘feel’ functions as a psych verb expressing a physical state, like bɔkhar lag ‘feel fever’, assigning the subject the thematic role of an experiencer.

4.1.3. Syntactic properties of dative subjects

An argument, in order to be the subject in a clause structure, needs to fulfill certain syntactic properties. Keenan (1976) states certain properties of subjecthood like reflexivisation, deletion of co-referential NP, controlled participial tests and so on. The following subsections show the syntactic properties of dative subjects in the light of these three tests of subjecthood.

4.1.3.1 Reflexivisation

According to Keenan (1976:315), ‘basic subjects can control reflexive pronouns’ i.e., only subjects can be the antecedent of reflexives. The subjects and the controlled reflexives are co-indexed. The dative subjects in AS undergo reflexivisation as can be seen (13)(a)-(c).

(13) (a) mɔ-kei niʤe dher poisa lag-ɛ
1-DAT self a lot money want.PRES-3SG
‘I myself want lots of money.’

(b) mɔ-kei niʤe bɔkhar lag-ɑ̃ h-ɛ
1-DAT self fever feel-PROG AUX.PRES-3SG
‘I myself have come down with a fever.’

(c) mɔ-kei niʤe sɔn-ai de-l-ak
1-DAT self hear-PERF give-PAST-3SG
‘I myself heard.’

In (13)(a)-(c) the subjects act as antecedent of reflexive niʤe ‘self’. The reflexive niʤe ‘self’ is a free word and normally occur immediately after the subject. The co-indexation indicates the subject controls the reflexives.

4.1.3.2. Deletion of co-referential NP

As per Keenan (1976:315) ‘basic subjects are the possible controller of co-referential deletion and pronominalisations’. That is, only subjects can control the co-referential NP of a non-finite clause. The co-referential NP can thus be deleted. In AS, dative subjects undergo the deletion of co-referential NP test.
In (14)(a)-(b) the sentences consist of two clauses – the main clause which is finite and the embedded clause which is non-finite (shown in square brackets). The dative subjects control the non-finite clause and the subjects (indicated by a gap and a subscript) of the non-finite clauses are deleted by virtue of being the co-referential NP. The dative subjects and the co-referential NP positions have the same indexation.

4.1.3.3. Controlled participial clause

As per Keenan (1976) subjects can control participial clause. In AS, dative subjects control participial clauses, as is illustrated in (15)(a)-(b).

(15) (a) 
\[
\text{baris me bhid̄y ke, mɔ-ke, bɔkhar ho-l-ak}
\]
\[
\text{rain LOC wet CP 1SG-DAT fever happen-PAST-3SG}
\]
‘Having been drenched in the rain, I became ill.’

(b) 
\[
\text{ekhon bɔkhar aw-ek ni lag-ɛ}
\]
\[
\text{1-DAT now fever come-NF NEG feel-PRES-3SG}
\]
‘I should not fall sick now.’

In (15)(a)-(b) we see that the dative subjects control participial clause (shown in square brackets). The subject position of the participial clause and the subject mɔ-ke ‘1SG-DAT’ are co-indexed so as to illustrate the control. The syntactic properties of the arguments with dative case in the above discussion show how they can be considered as syntactic subjects. However, in order to see how much agentivity they imply, it is important to observe the nature of the predicates that assign the dative subjects that is, whether they are pure dynamics or exhibit certain degrees of dynamicity. The following section (4.1.4) discusses the aspectual properties of the predicates that assign dative subject case.

4.1.4. Aspectual and proto-properties of the dative subject

In order to ascertain the thematic role of the dative subjects, I have conducted the tests of aspectual properties of verb suggested by Van Valin and LaPolla (1997:95). The aspectual properties will determine the durative nature of of the each type of verbs that assign dative subject case and entail proto-properties to the subject argument. As per the test, the use of in-phrase denotes the telic nature of the verb and the use of for-phrase denotes atelic or durative nature of the verb.
7. Non-nominative subjects in Assam Sadri

4.1.4.1. The use of the phrase *ek ghanta se/ me ‘in/ for an hour’

The phrase *ek ghanta se ‘for an hour’ does not indicate when the action began and ended and has the possibility of going on later. The phrase *ek ghanta me ‘in an hour’, on the other hand, denotes a specific time (Van Valin and LaPolla, 1997). This test reveals whether a verb can have a telic or terminal end point or not. In AS, neither *ek ghanta se ‘for an hour’ nor *ek ghanta me ‘in an hour’ can occur with dative subjects assigned by stative verbs, as in (16).

(16) *mɔ-keit ek ghanta se/ me phol bhal lag-e-la
    1-DAT    for /in an hour   flowers       good   feel-1SG-IMPERF
    ‘I like flowers for /in an hour.’

Dative subjects with predicates expressing physical states can occur with the phrase *ek ghanta se, indicating the durative nature or the possibility that the action can be ‘going on for a certain stretch of time’, or is atelic. But they cannot occur with *ek ghanta me ‘in an hour’, as such verbs cannot denote terminal endpoint or telicity, where the action has a beginning and an end as in (17).

(17) mɔ-keit ek ghanta se/#me bɔkhar lag-ət h-e
    1-DAT    for /in an hour   fever         feel-PROG AUX.PRES-3SG
    ‘I am feeling feverish for/*in an hour’.

4.1.5. Proto-properties of dative subject

In AS, dative subjects occur with perception verbs like ‘hear’ and ‘see’, followed by the auxiliary de ‘give’, which indicates that sentence is not involved on the part of the subject. From the discussion, it appears that we cannot have two distinct verbal categories of stative and dynamic that assign dative and nominative case, respectively. Rather, the verbal categories are dynamic non-volitional action, mental and physical states that assign dative case to the subject. So we can conclude that verbs that assign dative case to their subjects belong to a position between the two extreme categories of dynamic and stative. The nominative subject occurs with dynamic volitional action and has proto-agent properties. Unlike nominative subjects, dative subjects with a thematic role of experiencer do not have control over the action. But their involvement in the action or the experience, even though non-volitional, can be shown to exist on a continuum between dynamic non-volitional action followed by mental state and finally physical state Figure 2.

![Thematic hierarchy of dative subject in AS](image)

Figure 2: Thematic hierarchy of dative subject in AS

4.2. Genitive Subjects

In AS, subjects takes a genitive case marker (-r or ker)\(^6\) with predicates like *ahe ‘have’, implying possession or ‘to possess’. It can be either concrete or abstract possession.

\(^6\)The genitive marker is -r in case of 1st and 2nd person singular pronominal as in mɔ-r and to-r. –ker is used with plural pronouns and 3rd person singular pronouns.
4.2.1. Stative vs dynamic

The genitive case indicated by the -r marker can also have non-possessive implications. These genitive subjects with non-possessive implications mostly occur with stative verbs expressing mental states of ‘anger’ and ‘desire’. The dynamic counterparts of these verbs take the nominative case. This is illustrated in the following sub sections.

4.2.2. Verbs expressing desire or intension and psych verbs

The genitive subject in AS can occur with verbs expressing mental states like desire, liking or intention. Again, psych verbs like ‘to be angry’ denote a mental state when they occur with the copula ah- in its existential function (Dey and Barbora 2012). The subject of such verbs takes the genitive case as in (19)(b)\(^8\). The dynamic counterpart of this verb takes the progressive aspect of the auxiliary verb kar ‘do’ and takes nominative subjects as in (19)(a). However, the mental state of ‘anger’ can appear with the auxiliary ʊth ‘rise’ in its progressive aspect, where again the subject is in genitive case as in (19)(c).

\[ (19) \]
\[ (a) \quad mɔi \quad ʊ-ker \quad ʊpre \quad raag \quad kar-ət \quad h-ə \]
\[\text{1.NOM 3SG-GEN LOC angry do-PROG AUX.PRES-1SG} \]
\[ ‘I am getting angry with him.’ \]

\[ (b) \quad mɔ-r \quad ʊ-ker \quad ʊpre \quad raag \quad ah-ə \]
\[\text{1-GEN 3SG-GEN LOC angry EXIST.PRES-3SG} \]
\[ ‘I have anger on him.’ \]

\[ (c) \quad mɔ-r \quad ʊ-ker \quad ʊpre \quad raag \quad ʊth-ət \quad h-ə \]
\[\text{1-GEN 3SG-GEN LOC anger rise-PROG AUX.PRES-3SG} \]
\[ ‘My anger is rising in me for him.’ \]

4.2.3. Verbs expressing physical ailment

Verbs expressing ‘physical ailment’ can have a genitive subject as in (20). The physical ailment usually refers to the physical state of the subject argument.

\[ (20) \]
\[ mɔi \quad ʊ-ker \quad ʊpre \quad raag \quad ʊth \quad h-ə \]
\[\text{1.NOM 3SG-GEN LOC anger rise AUX.PRES} \]
\[ ‘He has many/much friends.’ \]

Subjects take the genitive case marker -ker with the verb ahe ‘have.PRES’, indicating concrete possession, as with saŋti ‘friend’ and abstract possession, as with sahəs ‘courage’ in (18). The genitive case ker is free morpheme when it occurs with the nominals but is affixed with pronominals.

---

\(^7\) The verb ahe ‘have.PRES’ is homophonous with the attributive copula ah in the language. However, their functions vary with regard to their morphological features. The attributive copula agrees to the subject in person and number, whereas the possessive verb ahe does not.

\(^8\) Here the copula ah- in its existential function has been glossed as EXIST.
7. Non-nominative subjects in Assam Sadri

(20) (a) $m\-r$ bimar lag-$\epsilon$
I-GEN ailment COP-PRES-3SG
‘I have an ailment.’

(b) $m\-r$ bimar ho-$i$ h-$\epsilon$
I-GEN ailment happen-PERF AUX-PRES-3SG
‘I have an ailment.’

(c) $m\-r$ bimar ho-$t$ h-$\epsilon$
I-GEN ailment happen-PROG AUX-PRES-3SG
‘I have an ailment.’

In (20) the subjects take the genitive case marker $-r$ and have the semantic role of experiencer, undergoing physical ailments such as ‘fever’ and ‘cold’. The subject takes genitive case with the equative copula $lag$ in (20)(a) indicating ‘state’, with perfect aspect of verb ‘happen’ in (20)(b) and with progressive aspect of ‘happen’ in (20)(c) denoting ‘event’.

4.2.4. Volitional vs non-volitional

Genitive subjects in AS occur normally with non-volitional actions whereas the volitional counterparts of these actions take nominative case. This contrast is illustrated with the use of the light verb $kar$ ‘do’, denoting ‘activity’ versus a verb like $h\-\alpha$ ‘become’, denoting ‘achievement’ or ‘event’.

(21) (a) moi der kar-$l\-\delta$
1.NOM late do-PAST-1SG
‘I came late.’

(b) $m\-r$ der $h\-\alpha$-$l$-$ak$
1.GEN late happen-PAST-3SG
‘I am late.’

(c) $m\-r$ der $h\-\alpha$-$t$ h-$\epsilon$
1.GEN late happen-PROG AUX-PRES-3SG
‘I am getting late.’

In (21)(b) the perfective aspect of $h\-\alpha$ ‘happen’ implies achievement or event, while in (21c) the progressive aspect gives a habitual reading. In (21)(a), the auxiliary verb $kar$ ‘do’ assigns nominative subjects, in the thematic role of agent. The action of $der$ $kar$ ‘late do’ refers to a volitional acts, where the subject has control over the action. On the other hand, in (21)(b)-(c) the actions indicated by the light verb $h\-\alpha$ ‘happen’ are non-volitional acts, indicating the subjects’ lack of control over the action. The predicate $h\-\alpha$ ‘happen’ takes a genitive subject when in the semantic role of beneficiary.

As far as the syntactic properties of genitive subjects are concerned it is seen that in AS, the reflexive $nij\-\acute{e}$ ‘self’ can have a genitive subject as its antecedent. They are therefore co-indexed. The reflexives occur after the subject and take the genitive case marker $-r$ as in (22) and (23).
(22) ʊ-ker, nɪdɛ-r, bahøŋ sahøs ah-ɛ  
3SG-GEN self-GEN much courage have.PRES.3SG  
‘He himself has enough courage.’

(23) mɔ-r, nɪdɛ-r, bimar hɔ-i h-ɛ  
1-GEN self-GEN illness happen-PERF AUX.PRES-3SG  
‘I myself have got an illness.’

Genitive subjects with stative verbs do not undergo co-referential NP deletion.

(24) *mɔr, [_. sahøs ah-ek] lag-ɛ  
I-GEN courage have-NF COP.PRES-3SG

(25) mɔ-r, [_. parh-ek] mon ah-ek  
I-GEN study-NF mind EXIST-NF  
‘I want to study.’

In (24), the genitive subject does not control the non-finite clause. But, in case of a predicate indicating a deliberate mental action such as ‘desire’ or ‘liking’, we see that they control the deletion of co-referential NP as in (25).

Again, the genitive subject with the predicate dāɪva ahe ‘responsibility have.PRES’ in example (26) does not control the participial clause. While, in (27)-(28), the genitive subjects of the predicate bimar hɔlak and mon ahe, having the thematic role of experiencer, control the participial clause.

(26) [*i, ba moir ke] ʊ-ker, dher dāɪva ahe  
father die CP 3SG-GEN much responsibility have.PRES

(27) [i, sahoɾ dɔai ke] mɔ-r, parh-ek mon ah-ɛ  
city go CP 1SG-GEN study-NF mind EXIST.PRES-3SG  
‘After going to the city, I would like to study.’

(28) [i, baris me bhidʒ ke] mɔ-r, bimar hɔ-l-ak  
rain POSP wet CP 1SG-GEN illness become-PAST-3SG  
‘Having been drenched in the rain, I became ill.’

4.2.5. Aspectual Properties of the verbs that take Genitive Subject

In AS, predicates that take genitive subjects cannot be divided into two distinct categories of stative and dynamic. Rather, the predicates can occur as state and also in progressive aspect of the state. In order to reveal the aspectual properties of these predicates, I have conducted the test of the use of the phrase ek ghanta se/ me ‘for /in an hour’. The use of the former will indicate the atelic or durative nature of the verb property of the verb, while the latter phrase will reveal the telic nature of the verb.

4.2.5.1. The use of the phrase ek ghanta se/ me ‘for /in an hour’

In AS, the predicate bimar lage 'ailment feel' indicates a state and therefore does not occur with the phrase ek ghanta se/ me, as shown in (29)(a). However, the predicate bimar ho
7. Non-nominative subjects in Assam Sadri

‘illness happen’ can occur with the phrase *ek ghanta se*, suggesting that the predicate is durative (29)(b). However, the use of the phrase *ek ghanta me* ‘in an hour’ is ungrammatical.

(29) (a) *mɔ-r*  
I-GEN  
for /in an hour  
*bimar*  
ailment  
*lag-e*  
COP.PRES-3SG  

(b)  
mɔ-r  
I-GEN  
for /in an hour  
*bimar*  
ailment  
*ho-t*  
AUX-PRES-3SG  

‘I am having ailment for /*in an hour.’

4.2.6. Proto-Properties of Genitive Subjects

Semantically, the genitive subject occurs with non-volitional predicates in AS. The genitive subject can be both possessor and experiencer. As far the independent existence of the genitive subject is concerned, the distinction can be made between concrete and abstract possession. In concrete possession the possessor exists independent of the possession whereas, in inalienable possessions like sahas ‘courage’, the subject co-exists with the possession; they cannot be separated. It is seen that the predicates that denotes more stativity do not have the proto-agent properties like volition, causation, movement and independent existence, whereas the predicates that can appear in progressive aspect of the state, can have some proto-agent properties, such as causing a change of mental state and sentience. The involvement or participation of genitive subjects in the action can be seen as a continuum between agent and patient roles as in Figure 3.

![Thematic hierarchy of Genitive Subjects in AS](image)

From the analysis of the genitive subjects, it comes out that the predicates indicating dynamic mental states ensure more involvement of the subject followed by predicates denoting physical states, events and finally possession.

4.3. Locative Subjects

Locative subjects indicate the abstract location of property and are associated with the notion of space. These subjects mostly have the semantic role of goal/possessor but also utilize the genitive construction. In such constructions we find the locative postposition *opre* or *me* occurring with the genitive-marked subject NP.

(30)  
mɔ-r  
1-GEN  
*opre*  
LOC  
*dher*  
many  
daigya  
responsibilities  
*ah-e*  
EXIST.PRES-3SG  

‘I have many responsibilities’

(31)  
a-kær  
3-GEN  
*me*  
LOC  
*bahuqt*  
much  
sakji  
strength  
*ah-e*  
EXIST.PRES-3SG  

‘He is very strong.’
In (30) and (31), we find that the subject NPs take compound post-positions, that is, the genitive -r and the locative postposition uprɛ and meɛ as case markers.

4.3.1. Syntactic Properties of Locative Subjects

In AS, locative subjects control reflexivisation and allow for deletion of co-referential NPs. But they do not control arguments in participial clauses. Consider (32)-(34).

(32) ə -ker i meɛ i nidge i bahy sakti ah -ɛ
3SG-GEN LOC self much strength EXIST.PRES.3SG
‘He himself is very strong.’

In (32) the locative subject is the antecedent that controls the reflexive nidge ‘self’.

(33) mɔ -r meɛ i dher [ˌi_ kam kar-ɛk] samta
I-GEN LOC much work do-NF capability
ah -ɛ
EXIST.PRES. 3SG
‘I have the capability to do lots of work.’

In (33) the locative subject controls deletion of co-referential NP.

(34) * [ˌi_ likh-a porh-a koiɾ ke] ə-ker meɛ i
write-NZ read-NZ do CP 3SG-GEN LOC
bes bɔddhi ho-ɪ-ak
much intelligence happen-PAST-3SG

In (34) the locative subject does not allow for control the argument in the participial clause. Instead of the locative subject a genitive subject will be appropriate here.

4.3.2. Aspectual Properties Locative predicates

The analysis of locative subjects in section 4.3.1. shows that they display two of the syntactic properties of subjecthood. In the following sub section, the aspectual properties of the locative predicates will determine whether locative subject asserts any control over the action or not.

4.3.2.1. The use of the phrase ek ghanta se/ me ‘in/ for an hour’

The predicates that take locative subjects mostly co-occur with the copula ahe in its existential function. Hence, these constructions mainly refer to location, abstract or concrete.
7. Non-nominative subjects in Assam Sadri

(35) (a) mɔ-r opre ek ghanta se/*me dher daïṭva
ah-e
EXIST.PRES-3SG
‘I have had many responsibilities for one hour’

(b) *o -ker me ek ghanta se/me bahuṭ sakṭi
ah-e
EXIST.PRES-3SG

In (35)(a)-(b) both the predicates refer to location of abstract entities like daïṭva ‘responsibility’ and sakṭi ‘strength’ but we see that the former is compatible with the ek ghante se phrase whereas the latter is incompatible with both ek ghante se and me. The reason behind such a selection could be that daïṭva is a relatively less permanent attribute than sakṭi and is accordingly less stative.

Locative subjects have syntactic subject properties. But semantically, the predicates do not denote proto-agent properties like volitionality, sentience, causation, or independent existence as locative subjects refer to the location of inherent property possession.

5. Thematic hierarchy of the non-nominative Subjects in AS

Discussing argument selection, Dowty (1991:578) states ‘not only do strong agents outrank strong patients, but both instruments and experiencers outrank any relatively patient-like argument for subjecthood.’ An agent is volition + causation + sentience + movement, or in some usage just volition + causation, or just volition (Dowty 1979, 1991). As per the properties of proto–agents proposed by Dowty (1991), in AS, nominative subjects can be considered proto-agents as they have properties like volition, sentience (deliberate perception) and causing change of state. They have independent existence in the sense that they cannot be created or destroyed by the event named by the verb. Nominative subjects can have both proto-agent and non proto-agent roles depending upon the nature of the verb. For instance, the experiencer nominative subject with the perception verb ‘see’ has less control over the action compared to that of a proto-agent nominative subject.

In AS, experiencer subjects are not proto-agents but they do not belong to proto-patients either as their involvement in the experience can be considered different from that of the proto-patients who are affected arguments. A proto-patient as per Dowty (1991) is the one that undergoes change of state, is causally affected, has no independent existence and so on. As has been mentioned in 3(c) in section 1., one of the proto-agent properties is to cause an event or change of state in another participant. In case of the experiencer subjects it is seen that they undergo a change of state like ‘feeling sleepy/hungry/feverish’ rather than causing any change to another participant. In such case, they are close to be claimed as patients, but they lack the proto-patient property of ‘being causally affected’. Non-nominative experiencer subjects are not volitional actors, nor are they causally affected by the action named by the verb. Thus, they occur in between the proto-agent and proto-patient roles in the thematic hierarchy.

In AS, these subjects display a degree of control along a continuum between proto-agent and proto-patient as illustrated in Table 1.
Table 1: The degree of control of the Non-nominative Subjects in AS

<table>
<thead>
<tr>
<th>The degree of control</th>
<th>Proto Agent</th>
<th>Proto-Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative</td>
<td>Dative</td>
<td>Genitive</td>
</tr>
<tr>
<td>Implication of the Predicates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volition</td>
<td>Non-volitional</td>
<td>Desire/intension</td>
</tr>
<tr>
<td>Deliberation</td>
<td>Obligation, Necessity</td>
<td>Non-volitional</td>
</tr>
<tr>
<td>Intension</td>
<td>Psche verbs</td>
<td>Perception verbs</td>
</tr>
<tr>
<td></td>
<td>Physical states</td>
<td>Mental states</td>
</tr>
<tr>
<td>Predicate Types</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;pit&quot; 'beat'</td>
<td>&quot;lag&quot; 'want/need'</td>
<td>&quot;mon ahe&quot; 'desire'</td>
</tr>
<tr>
<td>&quot;kand&quot; 'cry'</td>
<td>&quot;lag&quot; 'feel'</td>
<td>&quot;bhāk lag&quot; 'hungry'</td>
</tr>
<tr>
<td>&quot;d̄jan&quot; 'know'</td>
<td>&quot;bēkhar lag&quot; 'fear'</td>
<td>&quot;bimarhōk&quot; 'illness happen'</td>
</tr>
<tr>
<td>&quot;sac&quot; 'think'</td>
<td>&quot;nind lag&quot; 'sleepy'</td>
<td>&quot;rag uṭh&quot; 'get anger'</td>
</tr>
<tr>
<td>&quot;dekh&quot; 'look'</td>
<td>&quot;dikhai de&quot; 'appear'</td>
<td>&quot;bhēt ḍa&quot; 'meet happen'</td>
</tr>
<tr>
<td>&quot;son&quot; 'listen'</td>
<td>&quot;son-ai de&quot; 'heār'</td>
<td></td>
</tr>
<tr>
<td>&quot;yad kar&quot; 'remember do'</td>
<td>&quot;yad aa&quot; 'remember come'</td>
<td></td>
</tr>
<tr>
<td>&quot;bhēt kar&quot; 'meet do'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;bhāl pa&quot; 'like'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;bāt kar&quot; 'talk do'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;der kar&quot; 'late do'</td>
<td>&quot;bhāl lag&quot; 'like'</td>
<td>&quot;dēr ḍa&quot; 'late happen'</td>
</tr>
</tbody>
</table>

Table 1 shows the different predicate types and their implications for case of subject. The predicates that imply volition, deliberation and intension, as shown in the list of verbs, take a nominative subject with the thematic role of proto-agent. Whereas the dative and genitive subjects come next to the nominative subjects as per degree of control over the action. The dative subjects that occur with this list of predicates can be either a beneficiary or experiencer. Genitive subjects can be possessors and also experiencers depending upon the verb types. The locative subjects are not associated with the notion of control or involvement. Rather, they take the role of goal and hence occur next to the patient role. The proto-patient role is normally the object argument, a topic beyond the scope of this chapter. Hence, in Table 1, the cell indicating the object case below the proto-patient has been shown in a different shade.

Table 1 gives the Case hierarchy of the argument in the subject position, where nominative outranks the non-nominative. The mapping of case and thematic role in AS can be schematized as given in Table 2.
Table 2: The mapping of case and thematic hierarchy in AS

<table>
<thead>
<tr>
<th>Thematic Hierarchy</th>
<th>Case Hierarchy</th>
<th>Beneficiary</th>
<th>Experiencer</th>
<th>Theme</th>
<th>Possessor</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative</td>
<td>A, E, T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dative</td>
<td>E, B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genitive</td>
<td>E, P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Locative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G</td>
</tr>
</tbody>
</table>

In Table 2 the agent outranks the non-agentive thematic roles such as beneficiary and experiencer, theme, and possessor. Again, the case hierarchy shows that the nominative is higher than the non-nominatives. The thematic roles assumed by each Case features in the subject position is given in initials, A=agent, B=Beneficiary, E=Experiencer, T=Theme, P=Possessor and G=Goal.

6. Conclusion

In keeping with the findings that have been reported for other Indo-Aryan languages such as Bangla (Dasgupta 2004) and many others and more recently by Verbeke (2013), the analysis of subjects in AS brings forth the notion of agentivity and non-agentivity of these arguments and their corresponding morphological representation. In AS, nominative subjects can have the thematic roles of agent, experiencer and that of theme. Dative subjects can assume the role of experiencer and beneficiary. Genitive subjects, again, can have the thematic role of experiencer and possessor. In AS there is no one to one correspondence between case marking and thematic roles. Nevertheless, if we consider the role of the predicates in assigning the thematic roles to these case marked arguments, we can definitely discern a pattern. The predicates can be arranged in a hierarchy where the dynamic outranks the non-dynamic. The analysis of the AS data brings forth the observation that there is no definite demarcation between dynamic and stative, instead they are broken down into further categories showing the degree of control as continuum between proto-agent and proto-patient. Locative subjects indicate abstract location, having the semantic role goal and thus occurs after the patient role in the thematic hierarchy. From the above analysis it can be concluded that the grammar of argument marking of the subjects in the language is not arbitrary but it is organized in a hierarchy.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS</td>
<td>Assamese Sadri</td>
</tr>
<tr>
<td>AUX</td>
<td>Auxiliary</td>
</tr>
<tr>
<td>CL</td>
<td>Classifier</td>
</tr>
<tr>
<td>COP</td>
<td>Copula</td>
</tr>
<tr>
<td>CP</td>
<td>Conjunctive Particle</td>
</tr>
<tr>
<td>GEN</td>
<td>Genitive</td>
</tr>
<tr>
<td>IMPERF</td>
<td>Imperfective</td>
</tr>
<tr>
<td>NO</td>
<td>Nominative</td>
</tr>
<tr>
<td>NP</td>
<td>Noun Phrase</td>
</tr>
<tr>
<td>NF</td>
<td>Non Finite</td>
</tr>
<tr>
<td>NZ</td>
<td>Nominalizer</td>
</tr>
<tr>
<td>PAST</td>
<td>Past Tense</td>
</tr>
<tr>
<td>POSP</td>
<td>Postposition</td>
</tr>
<tr>
<td>PRES</td>
<td>Present</td>
</tr>
<tr>
<td>PROG</td>
<td>Progressive</td>
</tr>
<tr>
<td>PERF</td>
<td>Perfective</td>
</tr>
<tr>
<td>PL</td>
<td>Plural</td>
</tr>
<tr>
<td>SG</td>
<td>Singular</td>
</tr>
</tbody>
</table>
References


Historical Linguistics
8. Proto-Khasian: An emerging reconstruction

Paul Sidwell
Australian National University

Abstract
The paper summarizes an emerging phonological reconstruction of pKhasian being developed by the author. The basic framework combines etymological and typological considerations with insights achieved by earlier scholarship (especially Handricourt 1965 and Shorto 2006), modelling the phonological history of the Khasian languages and the reconstruction of pKhasian lexicon. The reconstruction finds that Khasi and Pnar are most similar phonologically and lexically, Lyngngam is more removed lexically, and War has significantly restructuring with vowel shifts and stop devoicing. An appendix is included that provides a pKhasian lexicon based on the Swadesh 100 list, with reflexes in the criterion languages.

Citation

Volume Editors
Gwendolyn Hyslop, Linda Konnerth, Stephen Morey, Priyankoo Sarmah

Copyright
© 2014, the author(s), release under Creative Commons Attribution license

1. Introduction

The Khasian\(^2\) languages are spoken by a substantial population in Meghalaya State of North East India, and to a lesser extent in neighbouring areas of Banglesh (see Brightbill et al. 2007). As Austroasiatic branches go Khasian is one of the smaller ones, being comprised of (apparently) four distinct languages (and their dialects and some mixed varieties, see Daladier 2011 for some commentary) although the group itself is extremely asymmetrical, dominated socially by Standard Khasi (henceforth simply Khasi, for which the 2001 census population count is 843,000) while several other members of the group and transitional dialects can be quite small (for example, Lyngngam has apparently only a few thousand of speakers).

While vernacular literary and prescriptive works on Khasi are extensive and well developed, comparative-historical analyses of the branch have been lacking in the literature. In particular, since the work of Grierson (1903) and Schmidt (1904) the usual practice of scholars has been to cite Khasi forms as representative of the branch and pay scant attention to other Khasian languages, labeling them ‘dialects’. Still, Standard Khasi is important for comparative purposes; being very well documented (at least prescriptively) its rich inventory of consonant clusters and vestiges of morphology (typical of the Khasian branch) have usefully informed reconstruction.

Recently, more extensive sources of lexical and other data for Khasian languages have become available (e.g. Nagaraja 1996, 2004, Choudhary 2004, Brightbill et al. 2007, Bareh 2010, Ring 2012, Baker 2013 and others) and these have permitted the progressive compilation of a comparative Khasian lexicon, and an emerging pKhasian reconstruction. Details of the emerging reconstruction were discussed in Sidwell (2011), and in early 2012 a

---

\(^1\) This paper began as the write-up for this author’s presentation “Proto-Khasian and Khasi-Palaungic” at the NEILS meeting in Guahati on February 4th 2012, but with the passing of time since then the details of the emerging reconstruction have evolved, so it was decided to focus more on the historical phonology and less on the classification question. I would especially like to thank Stephen Morey and Gwen Hyslop for encouraging me to follow through and actually submit the written draft, and to mention Adele Gregory for initially proofing the text. Finally I would like to acknowledge the support of the Australian Research Council for the Future Fellowship which has afforded me the time and resources to continue my work on Austroasiatic languages.

\(^2\) In the past I have also used the term Khasic for this branch, and more recently I have been advocating the more neutral geographical term Meghalayan, although Khasian is used consistently in this paper.
working version of this author’s comparative lexicon and reconstruction was posted online at sealang.net/monkhmer.

The main finding of this work can summarized as follows:

- The Khasian languages form a phonologically and lexically well-defined branch of Austroasiatic;
- Khasian is only moderately diverse, and relatively young compared to other AA branches (statistical methods suggest between 1,500 and 2,000 years);
- The consonant system is marked by a “Germanic shift” and loss of implosive versus plain contrast among voiced stops;\(^3\)
- The vocalism shows a trend towards loss of quantity contrast, continuing an apparently group wide trend;
- pKhasian shows lexical and structural features that suggest a special relationship with Palaungic.

These findings contribute to our understanding of Austroasiatic prehistory, and in particular to ideas concerning the migration and diversification of peoples in NE India in recent prehistory. It is also potentially significant that, being isolated geographically from SEAsia, Khasian has not been under the same areal pressures to restructure phonologically as other AA groups, so the reconstruction may reveal important clues about pAustroasiatic morphology and/or phonology.

It is also important that the recent computational phylogenetic of Nagaraja et al. (2013, see tree diagram at Figure 1) has independently confirmed the internal structure of the Khasian group, consistent with the indications from historical phonology discussed in this paper. The family tree indicated a relative young and compact Austroasiatic branch, which has been diversifying internally for perhaps around 1500 years.

![Figure 1: Khasian family tree from Nagaraja et al. (2013), with Palaung included to root the tree](image)

Although the family tree is found to have a strongly nested configuration, this actually has relatively little impact on the phonological reconstruction. The most divergent within the lineage are the War dialects, which have undergone significant phonological changes and lexical replacements after the break-up of pKhasian. But these changes have the effect of reducing the value of War as a witness to the reconstruction, except in marginal cases. More important is the evidence of the best known languages, Khasi and Pnar, which can be readily compared directly to wider Austroasiatic comparisons, demonstrating that their forms are relatively close to pKhasian, with few changes to model.

\(^3\) The so-called Germanic shift in this case is the general change in terms of greater positive VOT in oral stops. The effect is to shift plain voiceless stops to aspirates, e.g. \(p > p^h\) etc. In the case of Khasian it does not apply to the voiced stops. The term ‘Germanic Shift’ comes from the famous Indo-European sound-law discovered by Grimm which reflects spirantization of stops in Germanic.
2. Proto-Khasian Phonology

The pKhasian sound system is reconstructed with reference to four languages, Khasi, Pnar, War and Lyngngam. Descriptive materials are available for all four languages, although the quality is certainly mixed, and it can be problematic to reconcile the various notational conventions, analytical frameworks, and motivations of their authors. Fortunately we have now reached a point at which the phonetics and/or phonology of the four principle languages have been described in modern structural linguistic frameworks, the most important of these studies for the present work being the following:

- Khasi: the descriptive grammar of Rabel (1961) and older dictionaries such as Singh (1906) that mark more phonological distinctions than contemporary orthography;
- Pnar: the phonetics/phonology of Ring (2012) and recent theses such as Choudhary (2004), Bareh (2010);
- War: several dialects are known, e.g. the Amwi dialect is described in the grammar of Weidert (1975), and the Lamin dialect in the manuscript dictionary of Gashnga (in preparation);
- Lyngngam: the phonetics/phonology is described by Baker (2013) and there is also published data by Nagaraja (1996, 2004), Nagaraja et al. (2013);
- There is also a useful comparative lexicon for Khasi, Pnar, War and Lyngngam provided by Daladier (2011), plus the present author is in possession of various manuscript comparative Khasian materials prepared by Harry Shorto in the 1970’s (apparently with input from Weidert) but these are not directly referenced in this paper.  

The above sources provide sufficient detail and analysis to support a preliminary comparative phonology, which is outlined in the following sub-sections.

2.1. Word/syllable structure

The first stage of the comparative reconstruction pursued here is the modeling of word and syllables structures. Within Khasian languages there are very productive word-formational strategies - especially compounding - and within Khasi words of up to five syllables are attested (see Nagaraja 1985). In this study we are effectively restricting our treatment to words formed with one lexical root that may also be prefixed and/or infixed. Such forms can have initial clusters of two or three consonants, and can include obstruents and sonorants, with relatively few restrictions. Sonorant consonants and/or epenthetic vowels between stop sequences (written y in Khasi orthography) can create phonetic unstressed initial syllables, but syllabification in such cases is treated as secondary/superficial. Baker (2013) and Ring (2012) do report some examples of contrastive unstressed vowels in Lyngngam and Pnar respectively, but so far I have found no examples of cognates which would indicate reconstruction of phonemic unstressed vowels at any time-depth, although we must acknowledge the possibility going forward.

Baker (2013:65) describes Lyngngam as having a maximal word structure CVCCVC; Ring (2012:155) characterises the maximal Pnar monosyllable as CCVVC while his lexicon includes maximal word shape CVCCVVC; Weidert (1975:21) gives a maximal War-Amwi syllable as CCVVC, with longer words formed with additional initial syllables with structures

---

4 PDF copies of all the relevant Shorto materials can be obtained directly on request by email from Sidwell.
CV, CVC and CS (were S is a sonorant); Rabel (1961:15-29) characterizes Khasi words as consisting of at least a major syllable with maximal structure CCVVC, plus additional possible minor syllables of CC type (where the second element is a nasal, rhotic or liquid). Some of Rabel’s poly-syllabic forms have more than one lexical root, so this needs to be taken into account in considering her analysis.

The above structures can be reconciled into a common template which we can take as indicative of pKhasian. Keeping in mind that we are restricting our consideration to words with only one lexical root, we can distinguish two classes:

- monosyllables \(*C_i(C_m)VC_f\) which may or may not have initial clusters, and in which zero is a possible final, and
- iambic disyllables \(*(C_p(r/n))C_iVC_f\) in which the initial syllable has a sonorant peak or phonetically a short vowel epenthesis, corresponding to the maximal CVCCVC structure in the sources mentioned above.

The distinction between mono- and disyllables is potentially ambiguous; Khasian languages permit onset clusters that can include falling sonority, and from a phonetic view-point we might want to characterize such cases as disyllabic (or “sesqui-syllabic”, although I disprefer the latter term).

Vowel length is apparently contrastive in Khasi, but not in Pnar, Lyngngam or War, although quantity differences do exist in the other languages. Clear vowel quantity contrasts are apparently limited to Khasi, yet are not available in all syllable types. According to Nagaraja (1985), there is a structural symmetry in Khasi whereby main syllables with long vowels have short codas and main syllables with short vowels have long codas. In contemporary Khasi orthography short final stops are written \(b, d, id, j\) while long stops are written \(p, t, it, k\), and there is no orthographic indication of quantity with final sonorants. The early dictionaries are more helpful in this regard as they frequently indicated long vowels with various diacritics.

All the languages are reported to have diphthongal vowels possible in stressed position, and the descriptions of these are diverse and complicated by a historical tendency to misanalyse orthographic vowel sequences and glide transitions to adjacent palatals as indicative of diphthongs. Without going into detail in this short paper, suffice it to say that the comparative reconstruction indicates a single \(*ia\) diphthong historically (with good AA etymologies) while other diphthongs in Khasian languages are evidently secondary.

Consonant clusters in the languages are common, and have been the subject of discussion, including consideration of whether they violate some of Greenberg’s (1965) phonological universals (e.g. Henderson 1976; 1989-90). In summary, initial stop sequences show a tendency to dissimilate voicing and place of articulation, but this still allows for a tremendous diversity of clusters. Such sequences are typically written orthographically with an intervening letter \(y\), reflecting a tendency to vocalise the transition between stops with a short central vowel. Comparing Khasi to Greenberg’s universals, Henderson remarks:

---

5 E.g. Pnar /o/ is typically longer than /ɔ/ according to Ring (2013). Baker (2013) finds that some Lyngngam vowel phonemes are characteristically short. Such patterns are interpreted as indicating that there is a historical trend of loss of contrastive length which is still in progress, leaving numerous etic traces that should become clearer as more extensive phonetic documentation is accumulated.

6 These orthographic finals are typically taken as indicating voiced codas. Voicing may be involved, but there does not appear to have been any published investigations of the phonetics of Khasi codas. Nagaraja’s suggestion, which is impressionistic rather than based on instrumental analysis, appears to anticipate by several decades the idea of an AA “bimoraic constraint” of Anderson & Zide (2002).
Underlying this generalization would seem to be the idea that there is a tendency to have the voiced consonants of a cluster closest to the vowel, i.e. adjacent to the voiced nucleus of the syllable. This may indeed be a tendency, but it is not a universal. Khasi has initial clusters $bt$, $bth$, $bs$, $bsh$ [bj], $dk$, $dkh$, $dp$. In fact, there seems in Khasi to be a deliberate dissimilation of voicing in initial clusters, especially when the cluster is of two stops. Thus, alongside $bt$, $dk$- and $dp$-, we can have $tb$, $kd$- and $pd$, but $pt$- and $tk$- are reserved for a few onomatopoeic and what one might call ‘expressive’ words only. The general word stock prefers voice dissimilation. (Henderson 1989-90:62)

Traces of nasal and rhotic infixation - apparently now unproductive - are frequent. Indications of infixes can often be seen in geminate sonorants, which apparently have arisen from the assimilated of infixal $*-n/-*r$- to following segments (e.g. Khasi $pillac$ ‘set free’ shows an assimilated nasal infix, while the Pnar cognate, $pnlac$ ‘to cause’, retains an unassimilated reflex).

2.2. Proto Consonants

Overall the consonant correspondences across the branch are quite trivial with most involving no change, so in this short presentation we will focus on those correspondences where significant sound changes have occurred. In the consonantism these are as follows:

- War dialects show devoicing of stops, while the contrast is stable elsewhere;
- Prevocalic palatals show different patterns of lenition according to language.

The four languages all show aspirated stops at $C_i$ corresponding to plain voiceless stops in many other AA tongues, indicating a sound change characterised by Haudricourt (1965) as a Germanic shift (note that the reflex of $^*k^h$ is /h/ in War). A similar change has occurred in some individual languages/sub-branches elsewhere in AA, but only in Khasian has it affected a whole branch, and is therefore reconstructed to the proto-level. Additionally, there was a merger of pAA implosives and plain voiceless stops yielding a single voiced series. A moderate number of etyma retain initial voiceless stops without the Germanic shift, for reasons that are unclear; also there are apparent lexical innovations at pKhasian that include voiced, voiceless and aspirated $C_i$ stops. Changes among the other consonants are minimal; other than a tendency to confuse prevocalic /j/ and /ŋ/ the sonorants are stable. Consequently, we propose the following pKhasian consonants according to their positional constraints at Table 1.
Table 1: pKhasian consonants

<table>
<thead>
<tr>
<th>Ci Consonants</th>
<th>Ct Consonants</th>
</tr>
</thead>
<tbody>
<tr>
<td>*/ pʰ tʰ kʰ</td>
<td>*/ p t c k h /</td>
</tr>
<tr>
<td>p t c k ?</td>
<td>m n p n n</td>
</tr>
<tr>
<td>b d j</td>
<td>w r j h /</td>
</tr>
<tr>
<td>m n p n n</td>
<td>w r j h /</td>
</tr>
<tr>
<td>w r j h /</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cp Consonants</th>
<th>Cm Consonants</th>
</tr>
</thead>
<tbody>
<tr>
<td>*/ pʰ tʰ kʰ</td>
<td>*/ w r l j h /</td>
</tr>
<tr>
<td>p t c k</td>
<td></td>
</tr>
<tr>
<td>b d j</td>
<td></td>
</tr>
<tr>
<td>r s h /</td>
<td></td>
</tr>
</tbody>
</table>

Henderson (1976) makes an important point about the distribution of velar stops in Khasi. Despite an otherwise strong tendency for voicing dissimilation in clusters of obstruents, there are none-the-less sequences such as *kti ‘hand’, *kpa ‘father’ but none with /g/. This correlates with the lack of a /g/ generally in Standard Khasi, and it is also absent in Pnar and War varieties. A voiced velar /g/ is reported for Lyngngam, but a direct reflex of this segment is not found in corresponding etyma in other Khasian languages. It appears to occur as a result unpacking of nasals, and as prefix of recent origin, see Table 2.

The Lyngngam forms with initial voiced velar are isolated, perhaps they originate in the k-3rd person marker with voicing dissimilation yielding a /g/ (e.g. this is clearly the case in regard to the ‘water’ etymon, see below). Other examples suggest either unpacking of the velar nasal, or perhaps even simply mistranscription at the data collection stage. These considerations allow us to suggest that there was no *g in proto-Khasian, consistent with a general devoicing of AA voiced stops, and the fact that no velar implosive is indicated for pAA.

Table 2: Lyngngam voiced velar stop

<table>
<thead>
<tr>
<th>Gloss</th>
<th>Lyngngam</th>
<th>S.Khasi</th>
<th>Pnar</th>
<th>War-J.</th>
<th>War-Lamin</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘to hear’</td>
<td>sŋgu</td>
<td>sŋew</td>
<td>sŋaw</td>
<td>snao</td>
<td>sãã?</td>
</tr>
<tr>
<td>‘to stop’</td>
<td>sygeʔ</td>
<td>syeʔ</td>
<td>--</td>
<td>--</td>
<td>soyeʔ</td>
</tr>
<tr>
<td>‘water’</td>
<td>gum</td>
<td>?am</td>
<td>?um</td>
<td>?um</td>
<td>?am</td>
</tr>
<tr>
<td>‘that’</td>
<td>gteʔ</td>
<td>taj</td>
<td>taj</td>
<td>taj (tun)</td>
<td></td>
</tr>
</tbody>
</table>

The general claim of Haudricourt (1965) is that within Khasian there was a restructuring of oral stops (occupying the C₁ position), such that Austroasiatic voiceless stops became aspirated, and implosives became plain voiced stops. Haudricourt could not find unambiguous examples indicating the fate of Austroasiatic plain voiced stops, but these were relatively infrequent in proto-Austroasiatic, so this is not so surprising. More problematic, however, we do find some apparent violations of the Germanic shift, such as indicated in Table 3.

---

7 War-J the Jaintiapur dialectes spoken in Bangladesh, from Brightbill et al. (2007).
8. Proto-Khasian: An emerging reconstruction

Table 3: Khasian stops that violate the Germanic shift rule

<table>
<thead>
<tr>
<th>Gloss</th>
<th>Lyngngam</th>
<th>Khasi</th>
<th>Pnar</th>
<th>War-Lamin</th>
<th>Shorto PMK</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘mushroom’</td>
<td>--</td>
<td>tit</td>
<td>tit</td>
<td>tet</td>
<td>1903 *ptis ‘fungus’</td>
</tr>
<tr>
<td>‘to blow’</td>
<td>--</td>
<td>put</td>
<td>put</td>
<td>pet⁸</td>
<td>1023 *pwt ‘to blow’</td>
</tr>
<tr>
<td>‘to bite’</td>
<td>kap</td>
<td>kap</td>
<td>--</td>
<td>--</td>
<td>1231 *kap ‘to bite’</td>
</tr>
</tbody>
</table>

Examples such as these show that it is difficult to argue that any phonological conditioning might be involved - an apparent violation of the neogrammarian principle in which a sound change has failed to apply to a portion of the vocabulary. If we did not have external comparisons to inform our analysis, we might suggest that these represent vestiges of the Austroasiatic plain voiced series, but in this case we have compelling evidence that they reflect voiceless stops that we would otherwise expect to see reflected as aspirates. Given the modest extent of the phenomenon, and the indicative nature of the external comparisons, my approach is to straightforwardly treat these as plain voiceless stops in proto-Khasian. There are weak indications that consonant clusters may have conditioned these exceptions to the Germanic shift, but this is not discussed further in this paper.

Taken together, the changes in labial, apical and velar stops from pre- to pKhasian can be summarized as follows:

pAA > pKhasian

*ɓ/*b-, *ɗ/*d-, *b-, *d-

*ph, *th, *kh- (and infrequently *p-, *t-, *k-, see Table 3)

Among the palatals and *s there is a different pattern of correspondence; the voiced palatal stop is not devoiced, and the voiceless stop is not generally aspirated like *p, *t, *k, but instead shows conditioned lenition and merger with /ʃ/. The Khasi orthographic /ʃ/ mostly corresponds to a palatal stop/affricate in the other languages, with a proportion of forms with /ʃ/ in War that may be loans or conditioned allophones, it is not quite clear.

Table 4: pKhasian palatal consonants and /s/

<table>
<thead>
<tr>
<th>pAA</th>
<th>pKhasian</th>
<th>Lyng.</th>
<th>Khasi (orth.)</th>
<th>Pnar</th>
<th>War</th>
<th>pKhasian</th>
</tr>
</thead>
<tbody>
<tr>
<td>*ʃ</td>
<td>*ʃ</td>
<td>j</td>
<td>j</td>
<td>j</td>
<td>c</td>
<td>*ja: ‘rice, jαw ‘sour’</td>
</tr>
<tr>
<td>*c/C_</td>
<td>*s</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>*ksaw ‘dog’, *pʂa: ‘nephew’</td>
</tr>
<tr>
<td>*c elsew.</td>
<td>*c</td>
<td>c</td>
<td>sh</td>
<td>b̥j</td>
<td>/ɕ\textsuperscript{h}-tj</td>
<td>*c\textgamma ‘to sit’, *c\textalpha:ŋ ‘bone’</td>
</tr>
<tr>
<td>*s</td>
<td>*s</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>*sim ‘bird’, *s\textalpha:\textalpha m ‘blood’</td>
</tr>
</tbody>
</table>

In addition to the Germanic shift mentioned above, there are several striking phonological innovations among the C\textalpha segments (main syllable codas); all Khasian languages reflect the changes listed in Table 5.

---

⁸ Weidert records War-Amwi forms *pʰ\texttilde t ‘sound of blowing your nose’ and *pʰ\texttilde h\texttilde t ‘wind; to blow’ which suggest a regular aspirated reflex of this phoneme.
Table 5: pKhasian C₁ segmental changes

<table>
<thead>
<tr>
<th>Change</th>
<th>pAA (Shorto#)</th>
<th>pKhasian</th>
<th>Khasi examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>*£-l &gt; -n</td>
<td>*diel (#1745a) &gt; *de:n ‘tracks’</td>
<td>dien /dien/</td>
<td></td>
</tr>
<tr>
<td>*-l &gt; zero elsewhere</td>
<td>*priel (#1791) &gt; *pʰria ‘hail’</td>
<td>phria /pʰria/</td>
<td></td>
</tr>
<tr>
<td></td>
<td>*kial (#1710) &gt; *kʰia ‘cucumber’</td>
<td>khia /kʰia/</td>
<td></td>
</tr>
<tr>
<td>*-s &gt; t</td>
<td>*nas (#1912) &gt; *nːt ‘heart/liver’</td>
<td>nūd /nuːt/</td>
<td></td>
</tr>
<tr>
<td></td>
<td>*ʔas (#1871) &gt; *ʔat ‘to swell’</td>
<td>at /ʔat/</td>
<td></td>
</tr>
<tr>
<td></td>
<td>*kpa:s (#1915) &gt; *kmpʰa:t ‘cotton’</td>
<td>kymphād /kmpʰaːt/</td>
<td></td>
</tr>
<tr>
<td>*-h &gt; t</td>
<td>*kiah (#1967) &gt; *kʰiat ‘deer’</td>
<td>kʰiat /kʰiat/</td>
<td></td>
</tr>
<tr>
<td></td>
<td>*cuh (#1984) &gt; *prsut ‘bellows’</td>
<td>pyrsut /prsut/</td>
<td></td>
</tr>
<tr>
<td>*-ʔ &gt; zero</td>
<td>*caʔ (#8) &gt; *bsa ‘to feed’</td>
<td>bsa /bsa/</td>
<td></td>
</tr>
<tr>
<td></td>
<td>*kaʔ (#16) &gt; *kʰa ‘fish’</td>
<td>dohka /doʔ kʰaː/</td>
<td></td>
</tr>
<tr>
<td>*-k &gt; -ʔ</td>
<td>*lʔok (#277a) &gt; *lʔaʔ ‘cloud’</td>
<td>lyoh /lʔaʔ/</td>
<td></td>
</tr>
<tr>
<td></td>
<td>*tiak (#305) &gt; *tʰiaʔ ‘sleep’</td>
<td>thiaʔ /tʰiaʔ/</td>
<td></td>
</tr>
</tbody>
</table>

The changes among main syllable codas include the apparently remarkable shift of -h > -t, presumably through an intermediate merger with -s; the number of examples are few, but the environment seems to involve a high vowel, which is a plausible conditioning factor. The other changes are more typologically common, but still have distinctive Khasian features, in particular:

- The chain shift *-ʔ > zero, *-k > -ʔ preserves the etymological velar-glottal distinction,
- The split reflexes of pre-Khasian *-l are apparently conditioned, but the conditioning is unusual, as typologically it is more common to have either complete loss of final lateral or complete nasalisation (the latter being the typical outcome in Tai languages).

2.3. Proto Vowels

A thorough reconstruction of the vocalism of proto-Khasian is not possible with the data presently at my disposal, but the outlines of the proto-system seem clear enough, and have good typological justification. The vowel inventories reported in the sources for the four criterion languages are as follows:

---

9 The formulation of this rule is tentative due to shortage of examples, it may be that other conditioning factors are involved.
8. Proto-Khasian: An emerging reconstruction

Table 6: Khasian vowel phoneme inventories

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>/ i, i: i u</td>
<td>/ i, i: i α, u</td>
</tr>
<tr>
<td>e, o o</td>
<td>e, o o</td>
</tr>
<tr>
<td>a, a:</td>
<td>e a o</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Khasi Rabel (1961):</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ i, i: i u, u:</td>
</tr>
<tr>
<td>e, o: a, a: e, o:</td>
</tr>
<tr>
<td>ie, ia uo</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>/ i i u</td>
<td>/ i u</td>
</tr>
<tr>
<td>e o</td>
<td>e o</td>
</tr>
<tr>
<td>e e</td>
<td>e e</td>
</tr>
<tr>
<td>ia a</td>
<td>a</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>War-Amwi Weidert (1975)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ i u</td>
</tr>
<tr>
<td>e o o</td>
</tr>
<tr>
<td>e a o</td>
</tr>
<tr>
<td>ia ua</td>
</tr>
</tbody>
</table>

The vowel inventories given at Table 6, although showing superficial differences, can be readily reconciled. In respect of Lyngngam, Nagaraja lists nine members, including a short-long contrast for two pairs while Baker has 11 members including two pairs with tense-lax contrast. Having worked with both scholars, this writer is in a position to interpret the differences. It is evident that Nagaraja has failed to distinguish /e, e/ and /o, o/, treating them as simply /e/ and /o/. Also, Nagaraja has treated the distinction between the high front vowels as short-long, while Baker treated them as tense-lax, and Nagaraja misses the same tense/lax in respect of the high back vowels.\(^\text{10}\)

The Pnar vowels of Bareh and Ring are very similar; the high central vowel is lacking in Ring’s table because it is restricted to the minor-syllables, so this is not an issue. Additionally, Ring does not list diphthongs among his vowel phonemes, instead he treats them separately, providing a list of eight. All of these involve [i] but only one - /ia/ - is not conditioned by an adjacent palatal or laminal, so the net effect is that Bareh and Ring have the same inventory.

The War-Amwi vowel inventory is straightforward, the only remark I make here is that there has evidently been significant restructuring among War vowels, so although the inventory is similar to that of the other languages, the corresponding values in cognates can be very different (see discussion below).

The Khasi vowel system is more interesting, as there are systematic vowel length differences, and these frequently correspond to length values in cognate vocabulary elsewhere.

---

\(^{10}\) While tense/lax captures the present phonetics well, the distinction may well continue in part the older AA long-short quantity contrast.
in Austroasiatic (established by the etymologies in Shorto 2006). Rabel’s diphthongs ie, ou often correspond to /e/, /o/ elsewhere in Khasian, and her short-long high vowel correspond to the equivalent tense-lax pairs in Lyngngam. Also, there are indications that the Khasi short-long pairs /e, e:/ and /o, o:/ correspond to the height distinction /e, e/ and /o, o/ in the other languages (for example, Ring (2012) shows that Pnar /o/ is consistently long while /ɔ/ is short in similar environments).

Consequently, having regard to the above considerations, we reconstruct the inventory of pKhasian vowels as laid out at Table 7.

<table>
<thead>
<tr>
<th>Table 7: pKhasian vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>*/ i u i: u: e ə o e: ə: o: e a ə iə a: ə: /</td>
</tr>
</tbody>
</table>

The main discrepancies between the proto-inventory and contemporary Khasi relate to the mid and high central vowels. The short proto-vowel *ə is reconstructed to account for a modest number of correspondences of front and back vowels (such as *ksəw ‘dog’ < Khasi ksew, Lyngngam ksu). In much the same way, the long proto-vowel *əː is lacking direct reflexes as a central vowel, but is reconstructed to account for correspondences between Khasi ie and central back reflexes elsewhere (such as *-səːm ‘breath’ < Khasi msiem ‘soul, spirit’, Lyngngam hənsəm ‘breathe, inhale’).

Many Khasian vowel correspondences are trivial, while others are more complex. Among the latter the syllables with final glides show interesting patterning, especially War which dissimilates some sequences to produce new diphthongs. These are illustrated at Table 8.

<table>
<thead>
<tr>
<th>Table 8: Khasian correspondences with glide finals</th>
</tr>
</thead>
<tbody>
<tr>
<td>proto-rhymes</td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>*aːj</td>
</tr>
<tr>
<td>*aːw</td>
</tr>
<tr>
<td>*əw</td>
</tr>
<tr>
<td>*əj</td>
</tr>
<tr>
<td>*uː</td>
</tr>
<tr>
<td>*iː</td>
</tr>
</tbody>
</table>

More dramatically, the War dialects (Amwi, Lamin, also Mnar11 and others) are marked by a raising of *a, *aː, often to [ə] and [i], and *ə followed by glides have also raised, dissimilating the place of articulation of the following glide. See Table 9.

11 A Mnar lexicon is included in the materials of Harry Shorto. Although spoken in the north of the Khasian area well away from the War in the south, it is clearly a War dialect.
8. Proto-Khasian: An emerging reconstruction

Table 9: Khasian correspondences with glide finals

<table>
<thead>
<tr>
<th>pKhasian</th>
<th>Lyngngam</th>
<th>Khasi (orth.)</th>
<th>Pnar</th>
<th>War-Mnar</th>
<th>War-Amwi</th>
</tr>
</thead>
<tbody>
<tr>
<td>*kʰla: ‘tiger’</td>
<td>kʰla</td>
<td>kʰla</td>
<td>kʰla</td>
<td>kʰli</td>
<td>kʰli</td>
</tr>
<tr>
<td>*kʰa: ‘fish’</td>
<td>akʰa</td>
<td>ʔkʰa</td>
<td>dəkʰa</td>
<td>-</td>
<td>ʔi</td>
</tr>
<tr>
<td>*snaːm ‘blood’</td>
<td>snam</td>
<td>snam</td>
<td>snam</td>
<td>cʰim</td>
<td>rna</td>
</tr>
<tr>
<td>*ʔaːr ‘two’</td>
<td>ʔaːr</td>
<td>ʔaːr</td>
<td>ʔaːr</td>
<td>iːr</td>
<td>ʔiːr</td>
</tr>
<tr>
<td>*ksəw ‘dog’</td>
<td>ksi</td>
<td>ksew</td>
<td>ksaw</td>
<td>ksou</td>
<td>ksia</td>
</tr>
<tr>
<td>*hnəj ‘moon’</td>
<td>hni</td>
<td>hni</td>
<td>hni</td>
<td>pni</td>
<td>pnu</td>
</tr>
</tbody>
</table>

3. Final remarks

The above discussion gives a summary of the emerging phonological reconstruction of pKhasian being pursued by this researcher. The basic framework, combining etymological and typological consideration, and insights achieved by earlier scholarship, establishes a basis for modelling the phonological history of the Khasian languages and reconstruction of pKhasian lexicon. To illustrate this potential, a pKhasian lexicon based on the Swadesh 100 list, with reflexes in the criterion languages, is given at Appendix 1. The reconstruction also provides independent support for the family classification offered by Nagaraja et al. (2013) (see Figure 1) derived by computational phylogenetic methods.

The comparative reconstruction finds that Khasi and Pnar are most similar phonologically and lexically, and presently provide the most useful witnesses for reconstructing pKhasian. Lyngngam is less well known, but the indications are that while it is more removed lexically, it resembles Khasi and Pnar phonologically. War is significantly restructured by extensive vowel restructuring and stop devoicing, which presently limits its usefulness for reconstruction. Beyond this Sidwell (2011) has presented evidence and arguments that Khasian sub-groups with Palaungic within Austroasiatic, although that relation must be quite old and we have no particular indications of when or how the pre-Khasian speakers arrived in Meghalaya. It is expected that as the pKhasian reconstruction is further developed the issue of inter-branch relations, origins and migrations will be further clarified.
References


## Appendix 1

Swadesh list of pKhasian and reflexes
Non-cognates bracketed.
Data sourced from Nagaraja et al 2013.

<table>
<thead>
<tr>
<th></th>
<th>pKhasian</th>
<th>Khasi</th>
<th>Pnar</th>
<th>Lyngngam</th>
<th>War-Lamin</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>all</td>
<td>*-rɔʔ</td>
<td>rɔʔ</td>
<td>warɔʔ</td>
<td>prok</td>
<td>barɔʔ</td>
</tr>
<tr>
<td>2.</td>
<td>ashes</td>
<td>*tɔʔj</td>
<td>dpei</td>
<td>tpai</td>
<td>apaw</td>
<td>tvu</td>
</tr>
<tr>
<td>3.</td>
<td>bark (of tree)</td>
<td>*sneʔ</td>
<td>snep</td>
<td>snieʔ</td>
<td>snieʔ</td>
<td>sniaʔ</td>
</tr>
<tr>
<td>4.</td>
<td>belly</td>
<td>*kpsʔ</td>
<td>kpsʔ</td>
<td>kpsʔ</td>
<td>lawhaʔ</td>
<td>-</td>
</tr>
<tr>
<td>5.</td>
<td>big</td>
<td>*heʔ</td>
<td>heʔ</td>
<td>heʔ</td>
<td>(kimba)</td>
<td>(mia)</td>
</tr>
<tr>
<td>6.</td>
<td>bird</td>
<td>*(kisim)</td>
<td>sim</td>
<td>sim</td>
<td>sim</td>
<td>ksem</td>
</tr>
<tr>
<td>7.</td>
<td>to bite</td>
<td>*kap, *kʰat</td>
<td>(dait)</td>
<td>(dait)</td>
<td>kap, kinnap</td>
<td>hit</td>
</tr>
<tr>
<td>8.</td>
<td>black</td>
<td>*jəŋ</td>
<td>jəŋ</td>
<td>jəŋ</td>
<td>inŋəŋ</td>
<td>(prin)</td>
</tr>
<tr>
<td>9.</td>
<td>blood</td>
<td>*snaːm</td>
<td>snam</td>
<td>snam</td>
<td>snam</td>
<td>rnaʔ</td>
</tr>
<tr>
<td>10.</td>
<td>bone</td>
<td>*cʔəŋ</td>
<td>fəŋ</td>
<td>cəŋ</td>
<td>cəŋ</td>
<td>nʔəŋ</td>
</tr>
<tr>
<td>11.</td>
<td>breast</td>
<td>*bun, *-buː</td>
<td>jimbun</td>
<td>imbu</td>
<td>(eŋkriŋ)</td>
<td>buŋ</td>
</tr>
<tr>
<td>12.</td>
<td>to burn (vt.)</td>
<td>*təŋ</td>
<td>thəŋ</td>
<td>thəŋ</td>
<td>intʔəŋ, təŋəŋ</td>
<td>təŋ</td>
</tr>
<tr>
<td>13.</td>
<td>claw/nail</td>
<td>*trsim</td>
<td>trsim</td>
<td>trsim</td>
<td>torsim</td>
<td>snem</td>
</tr>
<tr>
<td>14.</td>
<td>cloud</td>
<td>*lʔɔʔ, *lŋəm</td>
<td>lʔɔʔ</td>
<td>lʔɔʔ</td>
<td>lʔoʔ</td>
<td>lŋəm</td>
</tr>
<tr>
<td>15.</td>
<td>cold</td>
<td>*kjam</td>
<td>kjam</td>
<td>(kʰiət)</td>
<td>kjam</td>
<td>ktjam</td>
</tr>
<tr>
<td>16.</td>
<td>to come/arrive</td>
<td>*ləˑr, *wan</td>
<td>wan</td>
<td>wan</td>
<td>ləˑr, linnar</td>
<td>va, van</td>
</tr>
<tr>
<td>17.</td>
<td>die (of a person)</td>
<td>*jap</td>
<td>jap</td>
<td>jap</td>
<td>jəp, j̥ənnap</td>
<td>jip</td>
</tr>
<tr>
<td>18.</td>
<td>dog</td>
<td>*ksəw</td>
<td>kəsw</td>
<td>kəsw</td>
<td>kəsw</td>
<td>kisia</td>
</tr>
<tr>
<td>19.</td>
<td>to drink (water)</td>
<td>*dɨʔ</td>
<td>dɨʔ</td>
<td>dɨʔ</td>
<td>deit, dinniet</td>
<td>deʔ</td>
</tr>
<tr>
<td>20.</td>
<td>dry (adj./stat.)</td>
<td>*raw, *r(ə)kʰəŋ</td>
<td>kʰiəŋ</td>
<td>raw</td>
<td>rianʔkʰoŋ</td>
<td>riəŋ</td>
</tr>
<tr>
<td>21.</td>
<td>ear</td>
<td>*cəkər</td>
<td>fəkər</td>
<td>cəkər</td>
<td>ləkər</td>
<td>(təran)</td>
</tr>
<tr>
<td>22.</td>
<td>earth/soil</td>
<td>*-teʔ</td>
<td>(kʰiədəw)</td>
<td>(kʰiədəw)</td>
<td>(kmən)</td>
<td>pərθa</td>
</tr>
<tr>
<td>23.</td>
<td>to eat</td>
<td>*baːm, *baːŋ</td>
<td>baːm</td>
<td>baːm</td>
<td>baːŋ</td>
<td>buə</td>
</tr>
<tr>
<td>24.</td>
<td>egg</td>
<td>*pnləŋ</td>
<td>pləŋ</td>
<td>pləŋ</td>
<td>piliŋ</td>
<td>(hunsʔi)</td>
</tr>
<tr>
<td>25.</td>
<td>eye</td>
<td>*-mat</td>
<td>khmat</td>
<td>khmat</td>
<td>khmat</td>
<td>mat</td>
</tr>
<tr>
<td>26.</td>
<td>fat/grease/oil</td>
<td>*lʔɔt</td>
<td>səqai dred</td>
<td>səqai (immir)</td>
<td>ləʔɔt</td>
<td></td>
</tr>
<tr>
<td>27.</td>
<td>feather</td>
<td>*-nəʔ</td>
<td>sner</td>
<td>ɭəwənər</td>
<td>sniʔ</td>
<td>ɭəbəniar</td>
</tr>
<tr>
<td>28.</td>
<td>fire</td>
<td>*sməŋ, *diŋ</td>
<td>diŋ</td>
<td>diŋ</td>
<td>adon</td>
<td>ʃmən</td>
</tr>
<tr>
<td>29.</td>
<td>fish (n.)</td>
<td>*kʔaː</td>
<td>doʔk’a</td>
<td>dak’a</td>
<td>k’a</td>
<td>hi</td>
</tr>
<tr>
<td>30.</td>
<td>fly (v.)</td>
<td>*heːr</td>
<td>her</td>
<td>phnher</td>
<td>(kindej)</td>
<td>piar</td>
</tr>
<tr>
<td>31.</td>
<td>foot</td>
<td>*kʃət</td>
<td>kʃat</td>
<td>kʃat</td>
<td>kʃat</td>
<td>(nía)</td>
</tr>
<tr>
<td>32.</td>
<td>full (as a basket)</td>
<td>*dəp</td>
<td>dap</td>
<td>dap</td>
<td>dap</td>
<td>dap</td>
</tr>
<tr>
<td>33.</td>
<td>give</td>
<td>*ʔaːj</td>
<td>aj</td>
<td>e</td>
<td>ʔənnaj</td>
<td>ʔaː</td>
</tr>
<tr>
<td>34.</td>
<td>good</td>
<td>*miət</td>
<td>(bʰa)</td>
<td>(bʰa)</td>
<td>(mirr⁴aŋ)</td>
<td>miət</td>
</tr>
</tbody>
</table>
35. green *ŋỳngam, *ŋa: *ŋỳngam, ŋa: *ŋỳngam (sǐnjìŋ) $ñỳngam
36. hair (of head) *su?, *spu? ʃu? spu? su? sn?
37. hand *ktì: kti kti kie jì: ta:
38. to hear/listen *sìpaw sjèw, snàp sìwà sìugu sà?
39. heart *nɔt, *kɔŋ- sìnan m kloŋ snam kloŋ snam ?ànòt kloŋ run子弟
40. horn *rèq rèq rèq òrèq rèq
41. I *nà: nà nà nà nà nà nominative form
42. to kill *pìnàp pìnàp pìnàp pìnàp pìnàp
43. knee *su: (kəʔ) kəʔ kəʔ kəʔ kəʔ Palaung karəm
44. know *ti: p tip tip (he?kon) tə?
45. leaf *sìla: sìla sìla sìla sìli
46. to lie (down) *drəm dem dem dem - -
47. liver *n̥o.t do/n̥u.d no.d no.d kətɨm Palaung kartɨm
48. long *r̥ŋəŋ r̥ŋəŋ r̥ŋəŋ jìr̥ŋəŋ kərŋəŋ
49. louse (head) *ksi: ksi ksi (silliet) səa
50. man/husband *cɔŋrəŋ ʃrəŋ cəŋrəŋ kəŋkɔŋrəŋ tərma Palaung ime
51. many *bʊ:n bən bən bən fìbəu War initial is voiced but vowel change is regular
52. meat/flesh *dɔʔ, *mìm doʔ dɔʔ mɛm mən doʔ
53. moon *bənaj bna:j bna:j bni pənəu
54. mountain/hill *lɔ:m, *peŋ/l jəm jəm dom pənəj
55. mouth *kətən (ʃɪntur) ktein (gap) (tkəŋ)
56. name *prtu:e (krəŋəŋ) prtud (krəŋəŋ) (tvəŋəŋ) Lyng. borrowed < Khasi
57. neck *rəndəŋ rəndəŋ (krəŋəŋ) rəndəŋ rəndəŋ War borrowed < Pnar
58. new *təmə:j təməj təməj təməj təmə: Lyng. borrowed < Khasi
59. night *məːt məːt (səŋpəu) məːt ləmaʔ?
60. nose *kəmut kəmut kəmut leumət (mərkəŋəŋ)
61. not *ʔəm, *təʔ im m inji təʔ təʔ
62. one *-wɨ: wəe wəi wəwə mi
63. person/human *briə: bɾiəw bɾiə bɾu tˈŋpɾəw
64. rain *slap slap slap slap sələː
65. red *səa:w saw so ʃəntsəw səa
66. road, path *ləntə ləntə (sərək) (tvər) (sərək) lyng
67. root (of a tree) *trəj, *təvət trəj təvəj təvəj təvəj
68. round (object) *pəlnən plənən plənən plənən plənən
69. sand *ŋəʔap ʃəʔap tʃə.esp cəʔap (sərvəu)
70. see *prəc, *ʔəji:j ʃəʔi:j pət (məʔpəʔ) (məʔ)
71. sit *kəŋ, *skəw fəŋ cəŋ cəŋ fəŋ fəŋ
72. skin *snəʔ, *snəʔ ʃnəʔ ʃnəʔ ʃnəʔ ʃnəʔ ʃnəʔ
73. sleep *təjəʔ təjəʔ təjəʔ (təŋəŋ) təjəʔ
təjəʔ
74. small *diʔ, *rət rət (kəʔpə) doh-diʔ (səbiət)
75. smoke (n.) *tədmə tədəm tədəm (inəʔak) tədəm War < Khasi/Pnar
76. to speak, say *ʔəŋ əŋ əŋ ənəŋ əŋəŋ
77. to stand *səŋ ʃəŋ ʃəŋ ʃəŋ ʃəŋ ʃəŋ
78. to talk *kələr kələr kələr kələr kələr kələr
79. to swim *jənəp, *jənəp jəŋ jəŋ jəŋ jəŋ (rəŋəŋ)
80. to take *tədəŋ tədəŋ tədəŋ tədəŋ tədəŋ
81. this (prox.) *ni: -ni -tə -tej katai ga-tej (keʔu tom)
82. stone *sməwə məw məw məw ʃma
83. sun *səŋəŋ ʃəŋ ʃəŋ ʃəŋ ʃəŋ ʃəŋ
84. to talk *tədəŋ tədəŋ tədəŋ tədəŋ tədəŋ
85. that (dist.) *təj -tə -tej katai ga-tej (keʔu tom)
86. to speak, say *ʔəŋ əŋ əŋ ənəŋ əŋəŋ
87. to stand *səŋ ʃəŋ ʃəŋ ʃəŋ ʃəŋ ʃəŋ
88. to talk *kələr kələr kələr kələr kələr kələr
89. to swim *jənəp, *jənəp jəŋ jəŋ jəŋ jəŋ (rəŋəŋ)
90. to take *tədəŋ tədəŋ tədəŋ tədəŋ tədəŋ
91. this (prox.) *ni: -ni -tə -tej katai ga-tej (keʔu tom)
92. to stand *səŋ ʃəŋ ʃəŋ ʃəŋ ʃəŋ ʃəŋ
93. to talk *kələr kələr kələr kələr kələr kələr
94. to swim *jənəp, *jənəp jəŋ jəŋ jəŋ jəŋ (rəŋəŋ)
95. to take *tədəŋ tədəŋ tədəŋ tədəŋ tədəŋ
96. this (prox.) *ni: -ni -tə -tej katai ga-tej (keʔu tom)
8. Proto-Khasian: An emerging reconstruction

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>87.</td>
<td>thou/you</td>
<td>*me:</td>
<td>me</td>
<td>me</td>
</tr>
<tr>
<td>88.</td>
<td>tongue</td>
<td>*tnleːc</td>
<td>tʰillieʃ</td>
<td>tʰleiʔ</td>
</tr>
<tr>
<td>89.</td>
<td>tooth</td>
<td>*lmep</td>
<td>(bniaʔ)</td>
<td>ɬameʃn</td>
</tr>
<tr>
<td>90.</td>
<td>tree</td>
<td>*deːŋ</td>
<td>diŋŋ</td>
<td>deŋ</td>
</tr>
<tr>
<td>91.</td>
<td>two</td>
<td>*ʔaːr</td>
<td>aːr</td>
<td>ar</td>
</tr>
<tr>
<td>92.</td>
<td>to walk, go</td>
<td>*diʔ</td>
<td>(jaːid)</td>
<td>(lai)</td>
</tr>
<tr>
<td>93.</td>
<td>warm/hot</td>
<td>*sʔaːc, *cit</td>
<td>sʔaːid</td>
<td>cit</td>
</tr>
<tr>
<td>94.</td>
<td>water</td>
<td>*ʔum</td>
<td>um</td>
<td>um</td>
</tr>
<tr>
<td>95.</td>
<td>we (excl.)</td>
<td>*ʔiː</td>
<td>(ŋi)</td>
<td>i</td>
</tr>
<tr>
<td>96.</td>
<td>what?</td>
<td>-</td>
<td>-eʃ</td>
<td>(ilɛʔ)</td>
</tr>
<tr>
<td>97.</td>
<td>white</td>
<td>*leːʔ</td>
<td>lieʔ</td>
<td>leiʔ</td>
</tr>
<tr>
<td>98.</td>
<td>who?</td>
<td>-</td>
<td>(-no)</td>
<td>u/kə ji</td>
</tr>
<tr>
<td>99.</td>
<td>woman/wife</td>
<td>*kntʰaj</td>
<td>kntʰej</td>
<td>kntʰai</td>
</tr>
<tr>
<td>100.</td>
<td>yellow</td>
<td>*stem</td>
<td>stem</td>
<td>stəm</td>
</tr>
</tbody>
</table>
Language Contact
Abstract
The greater Himalayan region is home to many Tibeto-Burman as well as a number of Indo-Aryan languages. This article focuses on the area inside Nepal that borders Northeast India. Its aim is to investigate the impacts of Nepali (Indo-Aryan) on lexical and morphosyntactic features of the Baram language (Tibeto-Burman). Although the article is largely focused on morphosyntactic borrowing into Baram grammar, it also incorporates a study of loanwords in Baram due to contact with Nepali, which is the lingua franca of the region. In a corpus of a total of 33 hours of transcribed speech of a wide variety of genres, native Baram vocabulary is now limited to a few hundred words. Nepali grammatical morphemes such as the plural suffix and numeral classifiers in nouns have been borrowed. In addition, the evidential construction represents a fusion of the original Baram and an innovative Nepali construction. Generally, speakers use both native as well as borrowed constructions. Other cases of grammatical borrowing include a copula, the comparative construction, and coordinating conjunctions.

Citation

Volume Editors
Gwendolyn Hyslop, Linda Konnerth, Stephen Morey, Priyankoo Sarmah

Copyright © 2014, the author(s), release under Creative Commons Attribution license

1. Introduction

Nepal is home to languages from four major language families: Indo-Aryan, Tibeto-Burman, Austro-Asiatic, and Dravidian (as well as Kusunda, a language isolate). These four major language families are also represented in Northeast India, where furthermore Tai-Kadai languages are spoken. A common modern contact situation is the same in Northeast India and Nepal: it involves a dominant Indo-Aryan (IA) language (such as Assamese in Northeast India) that impacts non-dominant Tibeto-Burman (TB) languages (e.g., Rabha in Northeast India, see Joseph 2007). In fact, Assamese serves as a lingua franca in Northeast India (Moral 1997:43), and has thus had an impact on non-dominant TB languages. In this type of contact situation, we find a number of changes induced in TB languages by the dominant Indo-Aryan language. This article discusses an instance of this scenario in Nepal, where the dominant Nepali language (IA) has had an enormous impact on Baram, a TB language spoken in western Nepal. The contact of Nepali speaking people with TB languages has resulted in different contact situations in the Nepalese Himalayas (Noonan 2003: 65, Noonan 2006).

Baram (ISO ‘brd’) is a severely endangered language spoken mainly in one village named Dandagaun (literally, “hill village”) of the Takukot Village Development Committee (VDC) in the Gorkha district of western Nepal (van Driem 2007). In addition to Dandagaun, Mailung of the Takukot VDC is also a village where speakers with very low proficiency use this language. The Baram people are engaged in subsistence farming and are Hindus by religion. According to CBS (2002), there are 342 Baram people, and a more recent census (CBS 2011)

I thank the Linguistic and Ethnographic Documentation of the Baram Language Project, hosted by Central Department of Linguistics, Tribhuvan University Nepal and funded by Hans Rausing Endangered Languages Documentation Grant (MDP0158). I acknowledge thanks to Tej Ratna Kansakar, Yogendra Prasad Yadava, Krishna Prasad Chalise, Balaram Prasain and Krishna Poudel for their contribution while building the corpus for this project, and the members of the Baram community for the data they provided. Most of the Baram examples are drawn from this Baram corpus, which consists of 33 hours of transcribed speech of a wide variety of genres.
reports that only 155 people speak the language; however, the number of speakers may actually be even lower than this figure.

Bradley’s (1997) classification of TB languages includes Baram in the Eastern subbranch of West Himalayish, which itself belongs to the Bodic branch of TB, as presented in Figure 1.

![Figure 1: Genetic classification of Baram (Bradley1997)](image)

Both Bradley (1997) and a genetic classification by Noonan (2006:14) suggest that Baram and Thangmi are closely related languages. Thangmi is now spoken in proximity to complex pronominalized Kiranti languages (Turin 1998:477) to the east of the Kathmandu Valley, whereas Baram is spoken to its west, where it is in close contact with Nepali and non-pronominalized TB languages. While Thangmi has a much higher number of native speakers than Baram, an approximate total of 23,151 (CBS 2011), it also has to be considered threatened (Eppele et al. 2012). Thangmi has retained a verb agreement system (Turin 1998:477) similar to that of Kiranti languages (Turin 1998: 477), whereas Baram has not despite the fact that it was also classified as a complex pronominalized language by Grierson (1909). Turin therefore notes, “while the Baram system of verbagreement is all but decayed, the verbal morphology of Thangmi is complex and reminiscent of the Kiranti model”(1998:477). The obvious question of whether Baram has lost this system of complex pronominalization due to contact with Nepali (and perhaps other non-pronominalized TB languages), and over what period of time it would have lost it, is yet to be investigated.

We find some previous literature related to Baram lexicon and grammar. Hunter (1978[1868]) and Hodgson (1992[1880]) provide some lexical items. Similarly, Grierson (1909) discusses a few grammatical features of Baram in addition to offering some lexical items. Recent research includes Kansakar et al. (2011a), a discussion of the sociolinguistic situation of Baram. In addition, Kansakar et al. (2011b, c) are a dictionary and a grammar of the language respectively. They do not, however, discuss the contact-induced changes in Baram. Since the villages where Baram is spoken are in geographical proximity, there are no significant dialectal differences but only minor variations.

Evidence as to when the Baram people moved to their present location is not available. Neither is the history of the Baram villages recorded. However, there is a belief that they migrated to their present location from the east (Baramu et al. 2011:8). There is a legend which claims that “Barams are the progenitors of Sunuwars living in the eastern Nepal”

2 Exact information about the sources of the data in Hunter (1978[1868]) and Grierson (1909) is not mentioned, but it appears safe to assume that there was not much dialect variation of Baram in these villages.
9. Contact-induced changes in Baram

(Thapa 1996:10). Their migration to the western part of the Kathmandu Valley led to contact with the IA language Nepali. The contact between Baram and Nepali thus led to a number of contact-induced changes and borrowings in Baram, where borrowing is taken as “the incorporation of foreign features into a group’s native language by speakers of that language” (Haugen 1950; Thomason and Kaufman 1988:37). It is significant to note that Baram and Nepali share typological similarities in noun affixation and verb morphology (§3). We see borrowed morphology mainly if the position of affixes is identical in both Baram and Nepali.

The cases of grammatical convergence in Baram and Nepali that are presented in this article all underlie the same direction of borrowing: from Nepali to Baram. This is expected given the prolonged socioeconomic dominance of Nepali over Baram.

This article is structured as follows. §2 deals with the sociolinguistic setting of the Baram language. §3 briefly presents the typological features of Baram and Nepali. We present the analysis of loan words in Baram in §4. Contact-induced morphology is presented in §5, and Nepali influence in Baram syntax is discussed in §6. Finally, the paper is summarized in §7.

2. Sociolinguistic setting

The use of Nepali is gaining ground not only as a lingua franca but also as a mother tongue among the Baram people (Kansakar et al. 2011a). Since the Baram constitute a small ethnic group, they often marry speakers of other languages, which is one of the reasons for the interruption of intergenerational language transmission. The Baram are, in fact, the “rememberers” of the Baram language based on a typology of language endangerment by Grinevald (2003), because they rarely speak it in their daily lives.

All Baram speak Nepali, and there is not a single monolingual Baram speaker. The young generation has altogether stopped speaking Baram. In interviews conducted as part of a sociolinguistic study (Kansakar et al. 2011a), some speakers who are now above 55 years responded that Baram is the first language they learnt at home. The use of Baram is characterized by the following aspects: (a) since Nepali is used predominantly, the use of Baram has shrunk in most domains. In fact, it is difficult to meet a fluent Baram speaker; (b) the Baram speaking area is surrounded by other ethnic groups (such as Brahmins, Chhetris, and others), which increases the interaction of Baram speakers with non-Baram speakers.

Nowadays, Baram is hardly used in the everyday life of the Baram people. Speakers report that the language was used when they were young. Mina Baram (one of the fluent speakers) comments that she is happy to talk to her sisters in her mother tongue at regular intervals of several months (see Kansakar et al. 2011a for details). Prolonged bilingualism of the Baram also speaking Nepali has thus given rise to contact-induced language change.

3. Baram and Nepali: Some typological features

Despite Baram being Tibeto-Burman and Nepali being Indo-Aryan, there are some apparent typological similarities. Synchronically, the languages are similar in the number of consonants and vowels they possess. Both are verb-final languages and modifiers precede the head nouns in these languages. Baram, on the one hand, is mainly prefixing in verb morphology but allows suffixing in some cases of verb morphology and nominal morphology. Nepali, on the other hand, employs more suffixes compared to prefixes. Owing to these typological similarities, we can assume that the transfer of Nepali suffixes into Baram is facilitated (Aikhenvald and Dixon 2001:18). Thus, morphological borrowing is often found where the position of the morphemes is the same as in Nepali, the source language.
4. Lexicon

There is evidence that the native Baram lexicon has shrunk. Hunter (1978 [1868]) is the earliest wordlist of Baram vocabulary. The Baram dictionary by Kansakar et al. (2011b) suggests that there are only 1,022 native Baram words among a total of 3,652 words, i.e., less than one third. The other two thirds constitute borrowings mostly from Nepali, English, or other languages via Nepali.

It is evident that even basic vocabulary and function words have been borrowed. Although a detailed analysis of loanwords from Nepali into Baram is still lacking, the reduction of the Baram lexicon is seen when we compare the word list in Hunter (1978 [1868]) and the vocabulary Barams use these days. From a total of 176 words given in Hunter’s list, only 30 words were borrowings. After 143 years (from 1868 until 2011), the picture has changed. In Kansakar et al. (2011b), that same word list offered by Hunter was reelicited. It is found that the number of Nepali borrowings has doubled to 60 words when contrasted with Hunter’s list (1978 [1868]). A large number of Nepali loanwords have substituted native Baram words. A brief survey in this section shows how the modern lexical inventory of Baram contrasts with data in the previous literature. Note that for some lexical items, both native words and loanwords are used in the Baram corpus that was originally collected for the dictionary and the grammar by Kansakar et al. (2011b, c).

4.1. Nouns

Nepali nouns are freely borrowed into Baram. Added to this, we also find a lot of examples in the corpus where native nouns are used along with borrowed synonyms. Nouns can be derived in Baram via the nominalizer ki-, as can be seen in (1)-(3). In some cases, however, we find examples where the derivation is also influenced by the Nepali construction. For example, kjaku in (1) is a noun derived from aku ‘steal’. Baram speakers either use this native derived noun, or the Nepali loan cor ‘thief’.

(1)  
naŋ kjaku kilik.
   n-target   kil-ca   kil-liŋ
   you NMLZ-steal PST-become
   'You became thief.'

Similar to kjaku in (1), speakers also use either simply the derived noun kica ‘food’ as in (2), or add the general noun kuro, ‘thing(<NPL)>3 as in (3).

(2)  
kica gidumŋiŋa.
   ki-ca gi-dum ŋiŋa
   NMLZ-eat PRF-find COP.NPST
   '(I) havefoundfood.'

(3)  
kica kuro
   ki-ca kuro
   NMLZ-eat thing(<NPL)
   'food'

3Morphemes borrowed to Baram from Nepali are indicated by ‘(<NPL)’ throughout this article.
9. Contact-induced changes in Baram

This is a case where the native morphological process of nominalization via ki- is gradually losing ground. The construction in (3) can be considered a form of calquing: in colloquial Nepali, the identical participant nominalization construction including a head noun is used, as in khane kura ‘food (lit., things to eat)’. Even though kica by itself means ‘food’, Barams often make use of the Nepali structure as in (3). Nominalization that is only marked by ki- is very rarely used although examples like (1) and (2) do occur in the corpus.

The discussion leads to the following conclusion. Firstly, Baram borrows nouns from Indic sources, such as cor ‘thief (<NPL)’. This includes borrowings recorded by Hunter (1978 [1868]) as well as ones that have been borrowed since (see Appendix 1). Secondly, as a result of Nepali influence, the native way of deriving nouns is very rarely used, and instead a construction calqued from Nepali is used.

4.2. Verbs

Verbs are often considered difficult to borrow (Curnow 2001:415). When verbs are borrowed into Baram, they generally take the nativizing suffix -di, e.g., lekh-di ‘write (<NPL)-NTVZ’ or pod-di ‘read(<NPL)-NTVZ’, where lekh and pod are borrowed from Indic. In some rare cases, Baram verb roots of Indic origin, such as rop ‘plant(<NPL)’, do not take the nativizing suffix -di. Some verb stems undergo slight morphophonemic changes when followed by the nativizing suffix. For example, Nepali pəәhirənu ‘to put on’ becomes pəәirdido ‘to put on’ in Baram. Similarly, cjapnu ‘to press’ becomes cjepdido ‘to press’, khijaunu ‘to row a boat’ becomes khoidido ‘to row a boat’, lagnu ‘to feel’ becomes logdido ‘to feel’.

As is the case with nouns, some native Baram verbs that are recorded by Hunter (1978 [1868]) have nowadays been replaced with verbs from Indic. Such verbs include aprito ‘kill’, moto ‘strike’, and cisojo ‘tell’.

4.3. Adjectives

In Baram, some adjectives are simple whereas others are either derived from verbs or borrowed from Nepali. There is a list of native adjectives recorded in Hunter (1978 [1868]) but many of those are replaced by loan Nepali adjectives in modern times (see the Appendix).

Native adjectives in Baram are mainly derived from verbs with the adjectivizer ki-/gi-. For example, Hunter (1978 [1868]) provides kicho ‘fat’, and macho ‘thin’. However, the modern dictionary entries for ‘fat’ and ‘thin’ in Kansakar et al. (2011b) are moto and patlo respectively. Both of them are borrowed from Nepali. Instead of abo ‘white’, speakers now use giabo. The word giabo ‘fat’ is the same form with the prefix gi-. Compared to the earlier list, colour adjectives in particular are no longer used.

4.4. Other word classes

Native postpositions recorded by Hunter (1978 [1868]) are: hai ‘in’, thaci ‘on’, to ‘to’, nanj, trka ‘within’, andambu ‘without’ (now meaning ‘outside’). In the modern Baram corpus, however, most of the postpositions are borrowed, e.g., lagi ‘for’, war ‘this side’, pari ‘across’, etc.

---

4In the single instance of ban-dis’make(<NPL)-NTVZ’, the root ban-takes the form -dis of the nativizing suffix. The suffix -dis is thus lexically conditioned and not productive.

5 Hunter (1986 [1886) also mentions the nativizing affix -di in Baram budziu ‘understand’ (Appendix).

6 These forms suggest the existence of a separate verb stem cho ‘become fat’, considering that the negative marker in Baram is either ma~ ~ ma- or a-. This hypothesized verb stem, however, is not recorded anywhere as such. Note that the root for ‘eat’ is ca.
Conjunctions have largely been borrowed. The following conjunctions listed in Hunter (1978 [1868]) are no longer used today: wo:ŋ ‘and’, dzu:n ‘as’, k e’or’, u:chi ‘so’, he, khakṣa ‘thus’. There is not a single native coordinating conjunction in Baram anymore.

Native numerals up to ‘three’ are still used today, i.e., d e’one’, ni s’two’, som ‘three’. Grierson (1909:405) reports the use of native numerals bi ‘four’ and bọya ‘five’. Otherwise, Nepali numerals are used nowadays. In addition, a number of discourse particles such as rọ, tọni, ki appear to have been borrowed from Nepali into Baram. By contrast, pronouns seem to be less affected by contact. Most of the pronouns mentioned in the earlier lists are retained nowadays.

The borrowing of postpositions has also been reported for language contact in Northeast India. For example, the Rabha ablative postposition para ‘from’ is a loan from Assamese (Joseph 2007:363), see (4).

(4) nokgiribe panini para thé mànna cisona
  nokgiri-be pani-ni para thé màn-na cison-a
  master-DEF tree-ABL (GEN-POSTPOS) fruit get-INF expect-PRES
  'The master looks forward to get fruit(s) from (the) tree.' (Joseph 2007:363)

4.5. Summary of contact-induced changes in the Baram lexicon

This discussion is summarized in Table 1. The table shows contact-induced changes that have occurred within 143 years, as it compares the native vs. borrowed vocabulary as recorded in Hunter (1978[1868]) and Kansakar et al. (2011b). It includes a total of 176 vocabulary items, divided into word classes.

Table 1: Baram native vs. borrowed lexicon across word-classes

<table>
<thead>
<tr>
<th>Word classes</th>
<th>Number of lexical items</th>
<th>Hunter (1978[1868])</th>
<th>Kansakar et al. (2011b)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Native</td>
<td>Borrowed</td>
</tr>
<tr>
<td>Numerals</td>
<td>15</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Adverbs and particles</td>
<td>37</td>
<td>35</td>
<td>2</td>
</tr>
<tr>
<td>Adjectives</td>
<td>28</td>
<td>22</td>
<td>6</td>
</tr>
<tr>
<td>Nouns</td>
<td>63</td>
<td>56</td>
<td>7</td>
</tr>
<tr>
<td>Verbs</td>
<td>16</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>Pronouns</td>
<td>17</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>176</strong></td>
<td><strong>146</strong></td>
<td><strong>30</strong></td>
</tr>
</tbody>
</table>

Table 1 presents some interesting findings. Numerals from ‘one’ to ‘five’ were used when Hunter recorded the data but two of the five numerals are no longer used. Adverbs and particles have undergone drastic changes. Only two borrowed adverbs and particles were listed in Hunter. By contrast, borrowed adverbs and particles have increased to 25 at present. This is the category in which the highest percentage of Nepali loans is found. This is the most vulnerable word class in Baram in terms of being replaced by borrowings. Among 63 nouns mentioned in Hunter, only six were borrowed. The borrowed nouns have increased to 16. In the sample we are limited to, verbs are the class least affected by borrowing with an increase from one to two. No changes have occurred in Baram pronouns despite the change of inflectional morphology. Nevertheless it has to be noted that Table 1 only offers a preliminary
9. Contact-induced changes in Baram

perspective on the increase in borrowings, based on a convenience sample of 176 lexical items that are available from Hunter’s data.

Despite being a new and under-researched field of linguistic investigation, there are some studies of contact-induced changes in the lexica of languages of Northeast India. For example, Satyanath and Laskar (2008) note that lexical items of various genetic affiliation are found in the Bishnupriya lexicon. Furthermore, Devi (2004) points out that in Manipuri, almost half of the loanwords are from Indic. As shown above, Rabha has borrowed the ablative postposition from Assamese, among other items (Joseph 2007). Lexical borrowing from IA languages into TB languages is a shared linguistic phenomenon in many TB languages of Northeast India. Discussing the situation of loanwords in Galo varieties, Post mentions that “in certain Assamese contact areas in the foothills, loanword use is often extreme, to the point where native grammar has come to largely provide a partial shell for an almost wholly borrowed lexicon” (2007:64). Lexical borrowings from Hindi and Nepali are also reported in Kurtöp, a TB language spoken in Bhutan (Hyslop 2011:251). Thus, lexical borrowings from Indic into Tibeto-Burman are very common in the extended Northeast Indian region.

5. Morphology

Some cases of contact-induced morphological change are also evidenced in Baram. This happens more easily when there is a typological fit between the affixes of the languages under study (Thomason and Kaufman 1988:75). This is the case with Baram and Nepali, as in both languages all nominal inflections are suffixes.

5.1. Plural suffix

The Baram plural suffix -həru is a borrowing from Nepali. Interestingly, the plural suffix attaches to the second person pronoun in Baram, consider nayʼ ’2(SG)ʼ and nay-ru ’2-PLʼ.7 Both Hunter (1978 [1868]) and Grierson (1909) assumed that the native plural suffix is -du in the word hu-du ’they’. This suffix can not be obtained in elicitation, nor does it occur in the corpus. Instead, the plural suffix -həru is used. Consider (5).

(5) ghãs rə sin-həru nïlik.
ghãs rə sin-həru ni-lik
grass(<<NPL) and(<NPL) wood-PL(<NPL) NPST-become

‘It can be (used as) fodder and firewoods.’

Although the Nepali plural suffix occurs most frequently, the suffix -baŋ is also occasionally used as a plural marker in the Baram corpus.

(6) ubaŋgo palo kja.
ubaŋ-go palo ki-jâ
he-PL-GEN time(<NPL) PST-go

‘Their turn passed (lit. went).’

The suffix -baŋ is used as a plural marker in many contexts in the Baram corpus, in addition to its lexical meaning ‘man’. As the plural marker, it appears with a number of pronouns, such as tibay ’they’, ibaŋ ’these’, ubaŋ ’they’. In addition, it also appears in nouns,

---

7 Note that the form nay’2PL’ is also used interchangeably.
such as *mama-baŋ* 'girl-PL', *papa-baŋ* 'man-PL', *caca-baŋ* 'child-PL'. While *baŋ* is a lexical item of Baram origin, the suffix *-həәru* is borrowed from Nepali.

### 5.2. Numeral classifiers

The Baram classifier system is limited to making a human/non-human distinction although it only occurs with a few numerals. The numeral classifiers follow the nouns directly in Baram. There is a form *dheŋ*, which may be a fusion of *de* 'one' and a non-human classifier *-ey*, but it is difficult to tell whether *dheŋ* can be analyzed into these two components: This hypothesized classifier *-ey* does not occur with any other numeral except ‘one’ in the corpus. Two pieces of evidence suggest that *-ey* might be a numeral classifier in Baram. First, the form of the numeral ‘one’ is *de* (Hunter (1978 [1868]); Grierson 1909; Kansakar et al. (2011a)). Second, when the word *dheŋ* occurs in the corpus, it is not followed by other classifiers.

Similarly, the native human classifier is *-wa*, but its only occurrence in the corpus is in the forms *dze-wa*‘one-CLF’ (notice the morphophonemic change from *de* to *dze*), *nis-wa* ‘two-CLF’ (as in (7)), and *pəәccis-wa* ‘twenty.five(<NPL)-CLF’. The use of these native classifiers is thus very limited.

The head noun is *bal* ‘man’ in (7) and the native classifier *-wa* is attached to the numeral *nis* ‘two’, whereas the noun in (8) refers to the non-human noun ‘goat’, using the classifier *-ey*.

(7) *niswa bal likiŋ kja.*

    *nis-wa* bal lik-iŋ ki-ŋa  
    two-CLF man become-SEQ PST-go  
    'Gathering (lit. becoming) two men, (they) went.'

(8) *dheŋ michja*

    *dheŋ* michja  
    one.CLF goat  
    'one goat' (Kansakar et al. 2011c:56)

In modern Baram, both native and borrowed classifiers exist side by side. In the corpus, borrowed classifiers *-ta* and *dzəәna* in (9) and (10) are predominant. These numeral classifiers are borrowed from Nepali. The semantics of the classifier is the same: the distinction between human and non-human. Examples follow.

(9) *ŋa-go som-ŋa nam ki-ŋa.*

    *ŋa-go* som-ŋa nam ki-ŋa  
    I-GEN three-CLF(<NPL) house PST-COP  
    'I had three houses.'

(10) *nisdzəәnalakho.*

    *nis dzəәna bal akho*  
    two CLF(<NPL) man call  
    'Call two men!'

Aikhenvald and Dixon (2001:17) claim that there is a ready borrowability of classifiers if the position of the classifier in relation to its head noun is the same in source and recipient

---

8 It is not clear why *de* becomes *dze* when the classifier *-wa* is added.
language. This is the case with Baram and Nepali, where classifiers follow their head nouns in both languages. Similarly, the classifier system in Chhantyal (TB) has also been borrowed from Nepali (Noonan 2003:75).

The borrowing of classifiers from a dominant IA language has also been reported in Rabha. Joseph (2007:402) mentions, “Many classifiers that are used with native numerals have an Assamese origin, but they must be considered to be naturalised and incorporated into the Rabha system.” There are Assamese classifiers borrowed and attached to Rabha numerals, or the Assamese classifiers may be attached to Assamese numerals. In the Rabha example (11), the native numeral -sa ‘one’ attaches to -sak ‘CLF(<Asm)’ and follows the head noun kai ‘person’.

In the Rabha example (12), we see that the number sari ‘four’ is borrowed from Assamese along with the classifier jon ‘CLF’. When both the numeral and the numeral classifier are borrowed, their ordering is also changed as in (12), reflecting the order in the Assamese enumeration construction. The case of Rabha thus shows that typological convergence in ordering the head noun vis-à-vis the numeral plus the classifier is not always a prerequisite for borrowability.

(11) kaisaksa
    kai sak-sa
    person CLF(<Asm)-one ‘one person’ (Rabha; Joseph 2007:391)

(12) sarijonkai
    sari-jon kai
    four(<Asm)-CLF(<Asm) person ‘four persons’ (Rabha; Joseph 2007:402)

5.3. Genitive marker

The genitive marker in Baram, as reported by Hunter (1978 [1868]) and Grierson (1909), is -ku. In contrast to this, Baram speakers use the genitive marker -ko/-go at present. Consider the genitive forms of pronouns as reported in Hunter, Grierson, and Kansakar et al., as shown in Table 2. While in the earlier two sources, viz. Hunter and Grierson, the genitive marker is -ku, the modern allomorphs are -go/-ko, as reported in Kansakar et al. (2011c).

Table 2: Pronouns and their genitive forms

<table>
<thead>
<tr>
<th>Pronouns</th>
<th>Hunter (1978 [1868])</th>
<th>Grierson (1909)</th>
<th>Kansakar et al. (2011c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st singular</td>
<td>na</td>
<td>-ku</td>
<td>-go</td>
</tr>
<tr>
<td>1st plural</td>
<td>ni</td>
<td></td>
<td>nuŋru-ko,’2PL-GEN’</td>
</tr>
<tr>
<td>2nd singular</td>
<td>naŋ</td>
<td></td>
<td>usko,’3SG-GEN’</td>
</tr>
<tr>
<td>2nd plural</td>
<td>nuŋ</td>
<td>unku,’3PL-GEN’</td>
<td>ubaŋ-го’3PL-GEN’</td>
</tr>
<tr>
<td>3rd singular</td>
<td>u</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd plural</td>
<td>hudu</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>potu’3PL-GEN’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Since there is no regular sound change /u/ > /o/ in Baram, the question is, why did -ku turn into -go/-ko? Let us first ignore the voicing alternation in /k~g/, and try to explain the vowel change. One possible explanation is that this vowel change has happened due to contact with Nepali. In Nepali, the genitive marker is -ro/-ko, where -ro is used with 1st and 2nd person pronouns, and -ko is used with 3rd person pronouns and common nouns. Because the genitive
marker follows a now grammaticized plural marker and is part of a suffixal string, it is not subject to voicing assimilation rules in the way it would be if it were attached directly to the pronoun. One possible explanation is that Baram borrowed voiceless -ko from Nepali.

What strengthens this hypothesis is the form of the 2nd person plural pronoun nuŋru, seen in Table 2: if we compare this form to the form recorded previously, nuŋ, the question is where -ru comes from. Again, we can look to Nepali, which has a plural suffix -həəru that may be abbreviated as -ru. This plural suffix has been borrowed into Baram as well, as discussed above. Thus, the -ko genitive forms are indeed likely to have developed due to contact with Nepali.

Now, let us consider the voicing alternation in /k~g/. Note that this is due to allomorphy, whereby the voiced [g] occurs after voiced segments and voiceless [k] occurs after voiceless segments. For example, kəəpu ‘head’ ends in a vowel and therefore the genitive marker will be -go but when a noun ends in a voiceless sound it is -ko, such as cəap-ko ‘Chap tree-GEN’.

This exact same allomorphy in fact also exists in Baram in a suffix homophonous with the genitive suffix: the infinitive marker -go/-ko. The obvious explanation for the voicing allomorphy in the genitive suffix is therefore that it has developed in analogy with the infinitive suffix. The exception of the 2nd person plural genitive form nuŋruko may be explained due to both -ru and -ko being borrowed from Nepali, which may have prevented the native Baram allomorphy rule to apply.

Nevertheless, it is not inconceivable that this could have occurred internally. Sporadic vowel lowering could have changed -ku to -ko.

5.4 Evidential

Moving to verb morphology, the use of the borrowed suffix -chəə or -cəə deserves some examination. Baram has a distinct way of expressing inferred evidence, in which the suffix -a/-o is added as exemplified in (13).

(13) cəapa i dar lako.
    cəapa i         dar         lak-o
    old man-ERG     wood(<NPL)    cut-EVID

'The old man felled the wood. /(The old man seems to have felled the wood.)'

Example (13) is different from the simple past tense construction, which would be ki-lak ‘PST-cut’. The construction in (13) instead expresses an evidential distinction, showing that the information is inferred from the results. Under the influence of Nepali, in addition to the native morpheme -a/-o, the borrowed marker of inferred evidence -chəə is used, as in (14), where both markers occur in sequence (see Curnow 2001:431). Peterson (2000:19) maintains that the suffix -chəə marks inferred evidence in Nepali.

(14) baləi dalinhoəru cunochə.
    bal-i         dalin-haru         cun-o-chə
    man-ERG        purlin(<NPL)-PL (<NPL) put-EVID-EVID(<NPL)

'The men fixed (lit. put) the purlin (unexpectedly).'

The men seem to have fixed (lit. put) the purlin.

The borrowing -chəə is found in both elicited examples as well as in the corpus. The evidential marking may thus be expressed in two ways: either using the suffix -a/-o only, or

---

9We also find some counter examples, such as ibi-ko ‘self-GEN’, dzhəədə-ko ‘quarrel (<NPL)-GEN’.

176
9. Contact-induced changes in Baram

else using -a/-o followed by -chə. The single marking via -a/-o is more common than double marking via both -a/-o and -chə, but both express the identical meaning interchangeably. Note that in Nepali, besides occurring in an evidential construction, chə is a copula used in locative clauses as well to express perfect and progressive aspects (Acharya 1991). In Baram, only its use in the evidential construction has been borrowed. In Nepali, chə is inflected for person, number, gender and honorificity (Acharya 1991), but Baram only borrowed -chə as a suffix.

6. Syntax

Baram has borrowed some syntactic constructions from Nepali. In certain contexts, native morphology or syntax is employed, whereas in other contexts, borrowed morphology or syntactic mechanisms are employed. Whether the native or borrowed morphosyntactic construction is used is decided by speakers, contexts, and/or discourse topics.

6.1. Comparative marker

The comparative morpheme has been borrowed from Nepali. The comparative marker in Nepali, as in a number of Indic languages, consists of a non-finite marked verb of saying, i.e., bhon-da 'say-SIM(<NPL)'. In Baram, the verbs of saying are da-go 'say-INF' and kham-go 'say-INF', but the comparative is bhonda, borrowed from Nepali. Note that bhon 'say' never occurs in Baram independently as a verb of saying.

(15) uca bhonda maca nekhla.
    uca bhanda maca ni-khala
child than(<NPL) mother NPST-speak
'The mother speaks more than the child.'

As in Baram, other TB languages of Nepal such as Gurung, Tamang, and Chhantyal have borrowed a comparative morpheme from Nepali (Noonan 2006).

6.2. Desiderative

The native desiderative construction is encoded by the suffix -se~si followed by a copula. These two forms occur in free variation although -se is more common. In addition, a construction borrowed from Nepali is also sometimes used. Nepali does not mark the desiderative morphologically, and the native Baram desiderative construction exists along with the borrowed construction from Nepali. This is thus a case of constructional borrowing. The more common, native way of marking desideratives is given in (16) and (17), in the past and non-past tense respectively.

(16) ṇa māŋ jase kiŋa.
    ṇa māŋ ja-se kiŋa
    I also go-DESID PST-COP
'I also wanted to go.'

(17) ṇa hapse ƞiŋa.
    ṇa hap-se ƞi-ŋa
    I weep-DESID NPST-COP
'I want to weep.'
The desiderative mood is marked by the suffix -se followed by the auxiliary ṇiŋa (for non-past) or kiŋa (for past tense) in the Baram native construction as illustrated in (16 - 17). The desiderative suffix attaches to the verb and the copula (auxiliary) forms the desiderative construction in Baram. By contrast, Nepali has a different construction to convey the desire of the speaker, as shown in (18). There is no morphologically desiderative marked verb in Nepali.

(18) məәlai ghəә dzanəә məәn lagchəә.
    məә-lai ghəә dza-nəә məәn lag-chəә
'I want to go home.' (Nepali)

The desiderative in Nepali involves a dative subject, and the verb lag 'feel', which is preceded by məәn 'soul'. The expression məәn lag 'soul feel' always appears in this kind of construction. In (18), the finite clause is preceded by another clause dza-nəә 'go-INF'.

The Baram construction in (19) is analogous to the Nepali construction in (18). In a number of cases, Baram verbs marked with the desiderative suffix are additionally followed by the borrowed word lag 'feel'. This structure is partially analogous to (18). It occurs frequently, not only in direct elicitation but also in corpus data. The desiderative marker –se along with the copula following it expresses the desiderative construction in Baram. However, the verb lag 'feel' has replaced the copula in some sentences in the Baram corpus. Example (19) shows that it is more like the Nepali desiderative construction in (18) than the native construction given in (16) and (17).

(19) mamabangəoi ucwa kispu bela niuwa casi ɳilətgdi.
    mama-baŋ-goï ucwa ki-tsəpu bela niuwa ca-si
    woman-PL-ERG child NMLZ-carry time citron eat-DESID
    ɳi-lag-di
    NPST-feel(<NPL)-NTVZ
'The pregnant woman likes to eat the citron.'

6.3. Copula ho

Baram has two copulas. First, the native existential/locative copula is ṇiŋa with its suppletive past form kiŋa/kã. It is used in possessive clauses (20), and in predicative adjective and predicative locative clauses (21).

(20) ɳago niswa ucməoi ɳiŋa.
    ŋa-go nis-wa ucmai ɳiŋa
    I-GEN two-GEN daughter COP.NPST
'I have two daughters.'

(21) məori pərakkai ɳiŋa.
    məori pərak-goï ɳiŋa
    bee(<NPL) cliff-LOC COP.NPST
'The bees are in the cliff.'

We see that the copula ɳiŋa is used in different copular clauses in (20) and (21). In equational contexts, a typical construction is simple juxtaposition, as in (22).
9. Contact-induced changes in Baram

(22) sugo kisen.
    su-go             ki-sen
    who-GEN           NMLZ-match

    'Whose is good?' (Kansakar et al. 2011c:144)

The second copula is ho, which has been borrowed from Nepali and is used in
identificational clauses (predicative nominals), just like in Nepali. It seems that the native
strategy instead was simple juxtaposition as in (22).

An example of the borrowed copular identificational clause is given in (23). Nepali marks
number, person and gender on the copula (Acharya 1991), but Baram does not because it does
not have an elaborate person agreement paradigm.

(23) jo dar cin ho.
    jo             dar             cin             ho
    this(<NPL)     dar.wood(<NPL)    EMPH  COP.NPST(<NPL)

    'This is a Dar tree.'

The copula does not agree with a plural subject as in (24) in Baram. Notice that the copula
ho would inflect as hun in Nepali to agree with a plural subject but this does not take place in
Baram due to its lack of an agreement system.

(24) ubaŋru car bhai ho.
    ubaŋ-ru          car             bhai             ho
    they-PL(<NPL)    four(<NPL)     brother(<NPL)    COP.NPST(<NPL)

    'They are four brothers.'

6.4. Relative-correlative

The native correlative construction of Baram as reported by Hodgson (1992 [1880]:168) is
relative (hetu) - correlative (hotu). These relative-correlative pronouns do not occur in the
corpus, however. Instead, the Nepali relative-correlative construction is used, which involves
dzun (relative) and u/tjo (correlative). In (25), the common noun kjaŋ ‘bread’ is preceded by
the relativizer dzun. This noun is coreferential with the correlative tei ‘that’, which acts as a
pro-form.

(25) dzun kjaŋ kisen njña bal-ai tei kithja
    dzun            kjaŋ           ki-sen         njña            bal-ai     tei
    REL(<NPL)       bread         NMLZ-match    COP.NPST       man-ERG   COREL(<NPL)

    ki-thaŋ
    PST-bring

    'The good bread was what the man brought.'

As a comparison, Joseph (2007:486) argues that the relative-correlative construction is
borrowed into Rabha from IA sources. He maintains that this borrowed element is also
frequently used in Rabha. Example (26) illustrates this.
The relative pronoun *je* in (26) is a borrowed element from IA sources. Joseph further notes that the three closely related languages Bodo, Garo and Rabha share the same relative pronoun (2007:570). Structural borrowing of relative-correlative constructions from IA to TB languages is also reported in other languages of Northeast India, such as Mongsen Ao (Coupe 2007), as well as in Nepal (Noonan 2003:75).

### 6.5. Directionals

Another case of contact-induced change is the use of a biclausal construction to mark directionals in purposive clauses.

The native Baram directional purpose construction involves two directional prefixes, *he*-andative (movement away from the speaker) as in (27), and *hjuŋ*- venitive (movement towards the speaker) as in (28). These directionals also yield a purposive reading. This construction occurs frequently in the corpus.

(27) *dzhuŋ keilak.*
   dzhuŋ  ki-he-lak
   bamboo  PST-ANDA-cut
   '(He) went to cut the bamboo.'

(28) *bal kjoŋkham.*
   bal     ki-hjun-kham
   man    PST-VEN-say
   'The man came to say (it).'

In sharp contrast to (27) and (28), directional coding with purposive clauses is also expressed in another way in Baram. This other construction seems to be a borrowing from Nepali. First of all, let's consider a Nepali sentence in (29).

(29) *sita-ko logne bhat kha-nə a-jo*  
   Sita-GEN husband rice eat-INF come-PST.3SG.M.NH  
   'Sita's husband came home to eat rice.' (Nepali)

In a Nepali purposive clause, the infinitive -*nə* is followed by the main verb, such as *kha-nə ajo* 'came to to eat' in (29). As can be seen in this example, the purposive clause precedes the main clause. The infinitive is also used to mark the purposive clause in Nepali. Structurally, the Nepali example (29) differs from (27) and (28) in that it does not have a morphological way of expressing the directional. Examples (30) and (31) from Baram are similar to the construction in Nepali.

(30) *mamaca ahugo kitai.*
   mamaca ahu-go    ki-tai
   girl ask-INF  PST-come
   'The girl came to ask.'
9. Contact-induced changes in Baram

(31) məhə cago kiṯi məhuri.
    məhə     ca-go     ki-tai     məhuri
    honey(<NPL) eat-INF PST-come bee(<NPL)
'A bee came to eat honey.'

While relatively younger speakers use the borrowed construction shown in (30) and (31), the elderly speakers in the corpus use the native construction as shown in (27) and (28). These two strategies appear to be semantically and pragmatically identical. Although the periphrastic construction could have arisen internally, the fact that it is analogous to Nepali suggests that it was borrowed.

6.6. Coordinating conjunctions

The native coordinating conjunction to ‘and’ reported in Hunter (1978 [1868]) is no longer used. Instead, the coordinating conjunction rə is used, which is a loan from Nepali. Examples such as (32) abound in the corpus.

(32) ŋə-i nun-gəi niŋla rə hadzago aju kipi.
    ŋə-i     nun-gəi     niŋla     rə     hadza-go     aju     ki-pi
    I-ERG     you-DAT     millet     and     (<NPL)     paddy-GEN     seed     PST-give
'I gave you the seed of millet and rice.'

Similarly, Tamang has borrowed the conjunctions ‘and’ and ‘or’ from Nepali (Noonan 2006), and the borrowing of coordinators is also found in languages of Northeast India. For example, Rabha has borrowed the coordinator aro ‘and’ from Assamese (Joseph 2007:469). Joseph claims that aro “is a firmly established coordinating device in Rabha at present.” Likewise, Coupe (2007:127) suggests that the coordinating conjunction ãlə ‘and’ in Mongsen Ao might have been borrowed from either Assamese or Nagamese.

6.7. Sentence conjunction with əni

The native clause combining devices in Baram are the conjunctions tigəŋ and tjaŋ 'then/and then.' The native markers occur most frequently, but there are also instances where the Nepali conjunction əni 'then/and then' is used instead, as in (33).

(33) niru git kituk kə əni balban həkən kja.
    ni-ru     git     ki-tuk     kə     əni     bal-baŋ     hāku-iŋ
    we-PL     song(<NPL)     PST-do     COP.PST     then     (<NPL)     man-PL     lose-SEQ
    ki-ja
    PST-go
'We used to sing a song. Then, they (now) have forgotten (lit. lost) their songs.'

6.8. Disjunctive coordination

The disjunctive coordinators nə...nə 'neither... nor' are common in a number of IA languages, such as Hindi (Kachru 2006:241), Maithili (Yadav 1997 [1996]:241), and Nepali (Pokharel 2056 vs:187). This construction is also found in Baram. While the element -n or -n- is a negative element in IA languages (Masica 1991:390), the negative markers in Baram are a- or
As the Nepali disjunctive coordination construction was borrowed, it was changed into *na...na* 'neither nor' in Baram, as shown in (34).

(34) **na nisi na nidziwa.**

na ni-si na ni-dziwa
neither(<NPL) NPST-die nor(<NPL) NPST-live

'It is neither dead nor alive.'

Despite the vowel change, the disjunctive coordinators *na...na* clearly represent borrowings from Nepali. In a similar fashion, Subbarao et al. (2012) mention that the conditional subordinator *judi* ‘if’ is borrowed from Assamese to a variety of Rabha spoken mainly in lower Assam.

### 7. Conclusion

The constructions discussed in this paper are summarized in Table 3.

<table>
<thead>
<tr>
<th>Areas of change</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lexicon</strong></td>
<td>Reduction/replacement of native vocabulary</td>
</tr>
<tr>
<td><strong>Morphology</strong></td>
<td>Borrowing of plural suffix</td>
</tr>
<tr>
<td></td>
<td>Borrowing of numeral classifiers</td>
</tr>
<tr>
<td></td>
<td>Double-marking in evidential construction</td>
</tr>
<tr>
<td><strong>Syntax</strong></td>
<td>Borrowed comparative marker</td>
</tr>
<tr>
<td></td>
<td>Borrowed copula in predicative nominal</td>
</tr>
<tr>
<td></td>
<td>Relative-correlative</td>
</tr>
<tr>
<td></td>
<td>Coordinator</td>
</tr>
<tr>
<td></td>
<td>Sentence coordinator</td>
</tr>
<tr>
<td></td>
<td>Disjunctive coordinator</td>
</tr>
</tbody>
</table>

Note that two changes discussed in this paper are not listed in Table 3, i.e., the genitive (§5.3) and the biclausal directional purpose construction (§6.5), because they might either be the result of language contact or internal change.

This paper has presented a number of constructions in Baram that are argued to have developed as a result of intense contact with Nepali. While some changes are more clearly the result of language contact, others are still open to further analysis, but have been presented here for comparison and discussion. In this scenario of Baram contact with Nepali, speakers use a range of native and borrowed structures because the Baram language is not used on a day-to-day basis any more. In some cases, Baram speakers make use of native morphology or syntax whereas the borrowed morphology or syntactic mechanism is employed in other cases. The changes in Baram can be categorized into two types: (a) Some borrowed constructions have replaced native constructions, such as the plural suffix *-həru* that is borrowed from Nepali, the comparative marker *bhanda*, the coordinator *rə*, and the relative-correlative construction. The native forms are no longer found in these domains. (b) Other types of changes can be seen in which the native and new structures from another language (Nepali) exist side by side, i.e., clause chaining with *əni*, the introduction of the copula *ho* in predicative nominal construction and a new mechanism of encoding directional and desiderative clauses. We find a number of contact-induced changes in the TB languages from...
9. Contact-induced changes in Baram

Nepali in Nepal. Similar contact-induced changes are found in a number of Tibeto-Burman languages spoken in Northeast India, such as in Rabha.
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANDA</td>
<td>Andative</td>
</tr>
<tr>
<td>AUX</td>
<td>Auxiliary</td>
</tr>
<tr>
<td>ASM</td>
<td>Assamese</td>
</tr>
<tr>
<td>CLF</td>
<td>Classifier</td>
</tr>
<tr>
<td>COP</td>
<td>Copula</td>
</tr>
<tr>
<td>CORR</td>
<td>Correlative</td>
</tr>
<tr>
<td>DAT</td>
<td>Dative</td>
</tr>
<tr>
<td>DEF</td>
<td>Definite</td>
</tr>
<tr>
<td>EMPH</td>
<td>Emphasis</td>
</tr>
<tr>
<td>EVID</td>
<td>Evidential</td>
</tr>
<tr>
<td>FUT</td>
<td>Future</td>
</tr>
<tr>
<td>GEN</td>
<td>Genitive</td>
</tr>
<tr>
<td>NEG</td>
<td>Negative</td>
</tr>
<tr>
<td>NMLZ</td>
<td>Nominalizer</td>
</tr>
<tr>
<td>NOM</td>
<td>Nominative</td>
</tr>
<tr>
<td>NPST</td>
<td>Non-Past</td>
</tr>
<tr>
<td>NTVZ</td>
<td>Nativizer</td>
</tr>
<tr>
<td>PL</td>
<td>Plural</td>
</tr>
<tr>
<td>POSTPOS</td>
<td>Postposition</td>
</tr>
<tr>
<td>PRES</td>
<td>Present</td>
</tr>
<tr>
<td>PRF</td>
<td>Perfect</td>
</tr>
<tr>
<td>PRT</td>
<td>Particle</td>
</tr>
<tr>
<td>PST</td>
<td>Past</td>
</tr>
<tr>
<td>PURP</td>
<td>Purposive</td>
</tr>
<tr>
<td>REL</td>
<td>Relative</td>
</tr>
<tr>
<td>SEQ</td>
<td>Sequential Converb</td>
</tr>
<tr>
<td>SIM</td>
<td>Simultaneous Converb</td>
</tr>
<tr>
<td>TB</td>
<td>Tibeto-Burman</td>
</tr>
<tr>
<td>VEN</td>
<td>Venitive</td>
</tr>
</tbody>
</table>
References


### Appendix: Comparative word list

The dash (-) in the column “Kansakar et al. (2011b)” indicates the gap or missing words in Baram nowadays. The genitive forms of pronouns (in 23-26) show that the stems of pronouns have retained but the inflectional morpheme has undergone change.

<table>
<thead>
<tr>
<th>Item #</th>
<th>English gloss</th>
<th>Hunter (1978 [1868])</th>
<th>Kansakar et al. (2011b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>one</td>
<td>de</td>
<td>de</td>
</tr>
<tr>
<td>2</td>
<td>two</td>
<td>ni</td>
<td>nis</td>
</tr>
<tr>
<td>3</td>
<td>three</td>
<td>swom</td>
<td>som</td>
</tr>
<tr>
<td>4</td>
<td>four</td>
<td>bi</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>five</td>
<td>bəŋa</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>I</td>
<td>ŋa</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>you (thou)</td>
<td>naŋ</td>
<td>naŋ</td>
</tr>
<tr>
<td>18</td>
<td>he (she, it)</td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>19</td>
<td>we</td>
<td>ni</td>
<td>ni</td>
</tr>
<tr>
<td>20</td>
<td>ye</td>
<td>nung</td>
<td>nung, naŋru</td>
</tr>
<tr>
<td>21</td>
<td>that</td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>22</td>
<td>this</td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>23</td>
<td>their</td>
<td>un-ku</td>
<td>usko</td>
</tr>
<tr>
<td>24</td>
<td>your</td>
<td>nung-ku</td>
<td>ŋago</td>
</tr>
<tr>
<td>25</td>
<td>our</td>
<td>ni-ku</td>
<td>ŋi-ɡo</td>
</tr>
<tr>
<td>26</td>
<td>his</td>
<td>u-ku</td>
<td>usko</td>
</tr>
<tr>
<td>27</td>
<td>thine</td>
<td>naŋ-ku</td>
<td>naŋko</td>
</tr>
<tr>
<td>28</td>
<td>mine</td>
<td>ŋaku</td>
<td>ŋago</td>
</tr>
<tr>
<td>29</td>
<td>who</td>
<td>ʰai</td>
<td>su</td>
</tr>
<tr>
<td>30</td>
<td>what</td>
<td>ʰai</td>
<td>hai</td>
</tr>
<tr>
<td>31</td>
<td>anybody</td>
<td>suŋ</td>
<td>sumŋ</td>
</tr>
<tr>
<td>32</td>
<td>anything</td>
<td>ʰŋ</td>
<td>ʰai-ŋo</td>
</tr>
<tr>
<td>33</td>
<td>above</td>
<td>hu:khəi</td>
<td>ʰkuŋ</td>
</tr>
<tr>
<td>34</td>
<td>and</td>
<td>wo:ŋ</td>
<td>-</td>
</tr>
<tr>
<td>35</td>
<td>as</td>
<td>dzuːn</td>
<td>-</td>
</tr>
<tr>
<td>36</td>
<td>below</td>
<td>hu-ṃəi</td>
<td>ʰuməi</td>
</tr>
<tr>
<td>37</td>
<td>between</td>
<td>ə-səl</td>
<td>-</td>
</tr>
<tr>
<td>39</td>
<td>far</td>
<td>kaloh</td>
<td>kalok</td>
</tr>
<tr>
<td>40</td>
<td>from</td>
<td>dzaŋ, ɡaŋ</td>
<td>gaŋ</td>
</tr>
<tr>
<td>41</td>
<td>here</td>
<td>hidi</td>
<td>hidi</td>
</tr>
<tr>
<td>42</td>
<td>how</td>
<td>ʰeː-tu</td>
<td>ʰkhaloŋ</td>
</tr>
<tr>
<td>43</td>
<td>how much</td>
<td>kuwo</td>
<td>kwa</td>
</tr>
<tr>
<td>44</td>
<td>in</td>
<td>tha-ci</td>
<td>-</td>
</tr>
<tr>
<td>45</td>
<td>little</td>
<td>soːnbi</td>
<td>ami</td>
</tr>
<tr>
<td>46</td>
<td>much</td>
<td>bu:d-he</td>
<td>-</td>
</tr>
<tr>
<td>47</td>
<td>near</td>
<td>kəŋjak</td>
<td>kanek</td>
</tr>
<tr>
<td>48</td>
<td>no</td>
<td>məmi, əlik</td>
<td>mami</td>
</tr>
<tr>
<td>49</td>
<td>not</td>
<td>man</td>
<td>mahə</td>
</tr>
<tr>
<td>50</td>
<td>now</td>
<td>thachi</td>
<td>-</td>
</tr>
<tr>
<td>51</td>
<td>of</td>
<td>kuː</td>
<td>ko</td>
</tr>
<tr>
<td>52</td>
<td>on</td>
<td>thachi, gai</td>
<td>ɡəi</td>
</tr>
<tr>
<td>53</td>
<td>or</td>
<td>ke</td>
<td>ki</td>
</tr>
<tr>
<td>-----</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>54</td>
<td>outside</td>
<td>ōmbu</td>
<td>ōmbu</td>
</tr>
<tr>
<td>55</td>
<td>so</td>
<td>u:ci</td>
<td>-</td>
</tr>
<tr>
<td>56</td>
<td>then</td>
<td>wa-lhe</td>
<td>-</td>
</tr>
<tr>
<td>57</td>
<td>there</td>
<td>hu:di</td>
<td>hudi</td>
</tr>
<tr>
<td>58</td>
<td>thus</td>
<td>he, khaksa</td>
<td>-</td>
</tr>
<tr>
<td>59</td>
<td>to</td>
<td>tu:</td>
<td>-</td>
</tr>
<tr>
<td>60</td>
<td>today</td>
<td>ti-jə</td>
<td>t̪əja</td>
</tr>
<tr>
<td>61</td>
<td>tomorrow</td>
<td>wogai</td>
<td>hogaï</td>
</tr>
<tr>
<td>62</td>
<td>when</td>
<td>kəilhe</td>
<td>kəile</td>
</tr>
<tr>
<td>63</td>
<td>where</td>
<td>kunși</td>
<td>kuni</td>
</tr>
<tr>
<td>65</td>
<td>with</td>
<td>chon</td>
<td>-</td>
</tr>
<tr>
<td>66</td>
<td>within</td>
<td>trko, nəŋ</td>
<td>-</td>
</tr>
<tr>
<td>67</td>
<td>without</td>
<td>ōm-bu</td>
<td>-</td>
</tr>
<tr>
<td>68</td>
<td>yes</td>
<td>mo, lik</td>
<td>-</td>
</tr>
<tr>
<td>69</td>
<td>yesterday</td>
<td>mi-ljə</td>
<td>minja</td>
</tr>
<tr>
<td>70</td>
<td>air</td>
<td>əsi</td>
<td>asi</td>
</tr>
<tr>
<td>71</td>
<td>ant</td>
<td>əŋəp</td>
<td>anap</td>
</tr>
<tr>
<td>72</td>
<td>arrow</td>
<td>para</td>
<td>-</td>
</tr>
<tr>
<td>73</td>
<td>bird</td>
<td>dzja-liŋ</td>
<td>dzjaliŋ</td>
</tr>
<tr>
<td>74</td>
<td>blood</td>
<td>chi:wi:</td>
<td>cihui</td>
</tr>
<tr>
<td>75</td>
<td>boat</td>
<td>du:ŋa</td>
<td>-</td>
</tr>
<tr>
<td>76</td>
<td>bone</td>
<td>wot</td>
<td>hət</td>
</tr>
<tr>
<td>77</td>
<td>buffalo</td>
<td>bhəisə</td>
<td>-</td>
</tr>
<tr>
<td>78</td>
<td>cat</td>
<td>məndzi</td>
<td>məndzi</td>
</tr>
<tr>
<td>79</td>
<td>cow</td>
<td>sjə</td>
<td>sjə</td>
</tr>
<tr>
<td>80</td>
<td>crow</td>
<td>kəŋkəŋ</td>
<td>kəŋkəŋ</td>
</tr>
<tr>
<td>81</td>
<td>day</td>
<td>dina</td>
<td>-</td>
</tr>
<tr>
<td>82</td>
<td>dog</td>
<td>əkja</td>
<td>akja</td>
</tr>
<tr>
<td>83</td>
<td>ear</td>
<td>kana</td>
<td>kuna</td>
</tr>
<tr>
<td>84</td>
<td>earth</td>
<td>nasa</td>
<td>nəsa</td>
</tr>
<tr>
<td>85</td>
<td>egg</td>
<td>hom</td>
<td>homa</td>
</tr>
<tr>
<td>87</td>
<td>eye</td>
<td>mik</td>
<td>mik</td>
</tr>
<tr>
<td>88</td>
<td>father</td>
<td>bəbəi</td>
<td>abəi</td>
</tr>
<tr>
<td>89</td>
<td>fire</td>
<td>mai</td>
<td>mui</td>
</tr>
<tr>
<td>90</td>
<td>fish</td>
<td>nəŋa</td>
<td>nəŋa</td>
</tr>
<tr>
<td>91</td>
<td>flower</td>
<td>əwəi</td>
<td>-</td>
</tr>
<tr>
<td>92</td>
<td>foot</td>
<td>u:ndzik</td>
<td>undzik</td>
</tr>
<tr>
<td>93</td>
<td>goat</td>
<td>mi:chhə, michjə</td>
<td>michjə</td>
</tr>
<tr>
<td>94</td>
<td>hair</td>
<td>sjam</td>
<td>sjam</td>
</tr>
<tr>
<td>95</td>
<td>hand</td>
<td>bhit</td>
<td>bit</td>
</tr>
<tr>
<td>96</td>
<td>head</td>
<td>kapa</td>
<td>kəpu</td>
</tr>
<tr>
<td>97</td>
<td>hog</td>
<td>pəksjə</td>
<td>pəksjə</td>
</tr>
<tr>
<td>98</td>
<td>horn</td>
<td>u:ŋə, u:ŋu:</td>
<td>uŋu</td>
</tr>
<tr>
<td>100</td>
<td>house</td>
<td>nəm</td>
<td>nam</td>
</tr>
<tr>
<td>101</td>
<td>hunger</td>
<td>ujəŋkehe</td>
<td>ujəŋke</td>
</tr>
<tr>
<td>102</td>
<td>iron</td>
<td>phəiam</td>
<td>-</td>
</tr>
<tr>
<td>103</td>
<td>leaf</td>
<td>sou</td>
<td>su</td>
</tr>
</tbody>
</table>
9. Contact-induced changes in Baram

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Baram</th>
<th>Baram</th>
</tr>
</thead>
<tbody>
<tr>
<td>105</td>
<td>man</td>
<td>bāl, bār</td>
<td>bal</td>
</tr>
<tr>
<td>106</td>
<td>monkey</td>
<td>pāju:k</td>
<td>pājuk</td>
</tr>
<tr>
<td>107</td>
<td>moon</td>
<td>chalwawi, chalwaw</td>
<td>chaunii</td>
</tr>
<tr>
<td>108</td>
<td>mosquito</td>
<td>amin</td>
<td>-</td>
</tr>
<tr>
<td>109</td>
<td>mother</td>
<td>amai</td>
<td>amēi</td>
</tr>
<tr>
<td>110</td>
<td>mountain</td>
<td>danda</td>
<td>-</td>
</tr>
<tr>
<td>111</td>
<td>mouth</td>
<td>anam</td>
<td>anam</td>
</tr>
<tr>
<td>112</td>
<td>oil</td>
<td>asa</td>
<td>asa</td>
</tr>
<tr>
<td>113</td>
<td>name</td>
<td>min</td>
<td>min</td>
</tr>
<tr>
<td>115</td>
<td>plantain</td>
<td>upsje</td>
<td>-</td>
</tr>
<tr>
<td>116</td>
<td>river</td>
<td>gu:du:l</td>
<td>gudul</td>
</tr>
<tr>
<td>117</td>
<td>road</td>
<td>umma</td>
<td>ēmma</td>
</tr>
<tr>
<td>118</td>
<td>salt</td>
<td>cha</td>
<td>cha</td>
</tr>
<tr>
<td>121</td>
<td>snake</td>
<td>paigu</td>
<td>paigu</td>
</tr>
<tr>
<td>123</td>
<td>stone</td>
<td>ku:ṇba</td>
<td>kumba</td>
</tr>
<tr>
<td>124</td>
<td>sun</td>
<td>uni:</td>
<td>uni</td>
</tr>
<tr>
<td>125</td>
<td>thirst</td>
<td>awaphoŋ</td>
<td>ciga</td>
</tr>
<tr>
<td>126</td>
<td>tiger</td>
<td>bu:man</td>
<td>buman</td>
</tr>
<tr>
<td>127</td>
<td>tooth</td>
<td>su:a, swa</td>
<td>swa</td>
</tr>
<tr>
<td>128</td>
<td>tree</td>
<td>simma:</td>
<td>senma</td>
</tr>
<tr>
<td>129</td>
<td>village</td>
<td>handu:ŋ</td>
<td>handuŋ</td>
</tr>
<tr>
<td>130</td>
<td>water</td>
<td>awa</td>
<td>awa</td>
</tr>
<tr>
<td>131</td>
<td>weariness</td>
<td>kitukhwia</td>
<td>kitkhui</td>
</tr>
<tr>
<td>132</td>
<td>yam</td>
<td>jak</td>
<td>koi</td>
</tr>
<tr>
<td>133</td>
<td>bad</td>
<td>mādo</td>
<td>māsen</td>
</tr>
<tr>
<td>134</td>
<td>bitter</td>
<td>kjakhiai</td>
<td>kekhja</td>
</tr>
<tr>
<td>135</td>
<td>black</td>
<td>chilin</td>
<td>cilin</td>
</tr>
<tr>
<td>136</td>
<td>cold</td>
<td>chi:so</td>
<td>-</td>
</tr>
<tr>
<td>137</td>
<td>crooked</td>
<td>ba:yo</td>
<td>-</td>
</tr>
<tr>
<td>138</td>
<td>fat</td>
<td>kicho</td>
<td>ucho</td>
</tr>
<tr>
<td>139</td>
<td>flat</td>
<td>nimbule</td>
<td>-</td>
</tr>
<tr>
<td>140</td>
<td>good</td>
<td>gado</td>
<td>gōsā</td>
</tr>
<tr>
<td>142</td>
<td>green</td>
<td>siksik</td>
<td>-</td>
</tr>
<tr>
<td>143</td>
<td>handsome</td>
<td>kusjen</td>
<td>kisen</td>
</tr>
<tr>
<td>144</td>
<td>hot</td>
<td>u:du:m</td>
<td>keuhon</td>
</tr>
<tr>
<td>145</td>
<td>long</td>
<td>ku:bho, alhok</td>
<td>alok</td>
</tr>
<tr>
<td>146</td>
<td>raw (green)</td>
<td>pon</td>
<td>paggu</td>
</tr>
<tr>
<td>147</td>
<td>red</td>
<td>phajō</td>
<td>phaja</td>
</tr>
<tr>
<td>148</td>
<td>ripe</td>
<td>kimin</td>
<td>kimin</td>
</tr>
<tr>
<td>149</td>
<td>round</td>
<td>düllo</td>
<td>-</td>
</tr>
<tr>
<td>150</td>
<td>short (of a man)</td>
<td>anjak</td>
<td>-</td>
</tr>
<tr>
<td>151</td>
<td>short (of a thing)</td>
<td>anjak</td>
<td>-</td>
</tr>
<tr>
<td>152</td>
<td>small</td>
<td>ami</td>
<td>ami</td>
</tr>
<tr>
<td>153</td>
<td>sour</td>
<td>kjaso</td>
<td>keuwak</td>
</tr>
<tr>
<td>154</td>
<td>square</td>
<td>carpajō</td>
<td>-</td>
</tr>
<tr>
<td>156</td>
<td>sweet</td>
<td>kjosja</td>
<td>kjosja</td>
</tr>
<tr>
<td>157</td>
<td>tall</td>
<td>alhok</td>
<td>alok</td>
</tr>
<tr>
<td>158</td>
<td>thin</td>
<td>macho</td>
<td>-</td>
</tr>
</tbody>
</table>
There are some words that are treated as native Baram words in Hunter (1978 [1868]), but in fact they are of Indic origin. These words include *budziu* ‘understand’, *carptja* ‘square’, *dando* ‘mountain’, *kəihe* ‘when’, *duja* ‘boat’, *bhəisa* ‘buffalo’, *dina* ‘day’. The numbering is discontinuous in the table because some words which were not mentioned in Hunter (1978 [1868]) are not included in the table.
10. Phonological changes in the Hindi lexicon: a case of Meghalaya Hindi

Maansi Sharma
Jawaharlal Nehru University

Abstract
Language contact is a multidimensional phenomenon which is studied in bilingualism, language acquisition, typology, language processing etc. It can affect any aspect of a language. The typological study of any contact variety unravels interesting insights for linguistic researches. In this paper we discuss the contact-induced phonological changes in the Hindi lexicon of Contact Hindi (CH) in Meghalaya. CH is the link language in Meghalaya and is used at all the public places. Though, Hindi is the primary lexifier for CH but it also borrows words and various grammatical features from diverse pool of several languages such as Eastern Indo Aryan languages (EIA), languages of Meghalaya, English etc. But this paper seeks an enquiry into the phonological changes only of the Hindi lexicon in CH. The study also seeks an enquiry into the sources and processes that lead to these phonological changes. The results would be seen in light of Matras’s (2009) paradigm on types of processes leading to phonological changes. Further, a brief discussion on type of bilingualism and language attitudes of CH speakers is done in the conclusion.

Citation

Volume Editors
Gwendolyn Hyslop, Linda Konnerth, Stephen Morey, Priyankoo Sarmah

Copyright © 2014, the author(s), release under Creative Commons Attribution license

1. Introduction

A language can resemble another language for reasons such as geographic proximity, genetic relation and contact. Contact-induced changes occur when communication takes place among the speakers of different languages in a single geographical area. As a contact-induced phenomenon, the languages of such speakers can mix, change and undergo reanalysis. There are primarily three types of contact languages; pidgins, creoles and mixed languages. There are several examples in literature of these, such as, Michif, Media lingua, Tok pisin, plantation creoles etc. The development of contact variety has been aptly explained by (Siegel 2008: 1–3) that

“At the stage of initial contact, speakers first develop their own ways of communicating. It is generally done by using words and phrases they have learned from other languages (most often from the lexifier) which they believe that others might be familiar with. The combination of these individualized ways of communicating is called a ‘jargon’ or ‘pre-pidgin’. After a sustained contact, and constant use of pre-pidgin as a lingua franca, certain communicative conventions may develop, resulting in new language—a pidgin.”

Thus, in this paper we discuss an aspect of such a pidgin, called Meghalaya Hindi (MH). Before we move ahead, it is important to distinguish between Contact Hindi (CH), Meghalaya Hindi (MH) and Standard Hindi (SH). “Hindi the official language, based on Western Hindi,

---

1 I am extremely indebted to Kharsanu and Mellinda who are the primary informants for this paper. I further thank Prof. Awadesh Mishra who gave invaluable insights on the phonological changes in CH. I thank Radha Mohan Meena sir, Prof. V.S Shukla without whom my field trip would not have completed. I am extremely thankful to my supervisor Prof Anvita Abbi for all the support. I thank my mother Mrs Shobha Sharma who stood by me throughout the field trip. Lastly, I thank Shashank who has always been there as a critique, friend and a firm support.

2 This can be used as a term of reference but it does not represent entire state of Meghalaya. There can be many more varieties of Hindi in Meghalaya.
is the standardized form of a language that is also referred to as Modern Standard Hindi, to
distinguish it from the colloquial and regional varieties of the language” (Kachru 2006: 1–2). Contact Hindi is a variety of Hindi that is used as a language of wider communication and convenience among the people for whom Hindi is not the first language. Thus, CH is composed of inputs from first language of the people using it and SH. Hence, CH is deviant from SH. Further, MH is one of such CH spoken in Meghalaya. In this paper, the phonological changes in the Hindi lexicon of MH will be discussed. The aim of the paper is not to prove whether or not MH is deviant from SH. But the aim of the paper is to ascertain the types, causes and sources of phonological deviation. This has been studied by comparing the sound system of Hindi lexicon in MH and Hindi itself. The paper begins with a brief discussion on; Hindi as a link language, varieties of Hindi, CH in North East India and genesis of MH. This is followed by a brief theoretical discussion on language contact and phonology. Towards the end we discuss the data and results followed by a brief conclusion.

1.1. Hindi as a link language

Hindi was chosen as the second official language of India together with English, at the time of
constitution framing, as stated in Article 343, the constitution of India 1949 and Official
languages act 1963 (Basu, D.D 2012: 407–411). At present, it is the official language of 9
states and 1 union territory. These places are Bihar, Jharkhand, Uttar Pradesh, Uttarakhand,
Madhaya Pradesh, Rajasthan, Chattisgarh, Himachal Pradesh, Haryana and Delhi. It is
important to note here that Indian constitution provides the opportunity to state legislatures (Article 345) and the President (Article 347) to recognise languages other than Hindi for intra-state official transactions. In a state, the speakers of a substantially spoken language have a right to demand their language as an official language within state. Thus, other states have chosen English and a regional language/s as their official language. Hindi being an official language of several states has been the integral part of governance, administration and daily public life across the country. But its feature pool varies with the change in geographical area, languages in contact and social environment of the area it is situated in. (Das 1997) broadly classifies Hindi speakers into following six groups:

- Speakers of Standard Hindi (SH) who use it as their mother tongue.
- Speakers of Hindi whose primary language is certain dialect of Hindi.
- Speakers of Hindi whose primary language is another major Indo-Aryan
  language such as, Punjabi, Gujarati, Bangla etc.
- Speakers of Hindi whose mother tongue is a Dravidian language such as Tamil,
  Telegu, Malyalam etc.
- Native speakers of Urdu.
- Speakers of Hindi whose primary language is English i.e., a variety of Standard
  English or English in India.

Thus, these distinct speakers develop varieties of Hindi such as, Bambaiya Hindi, Kalkattai Hindi, Dakkhini Hindi etc. These varieties have developed into distinct forms. Thus, some of these are restricted pidgins whereas others may be expanded pidgins and

---

3 “It is the codified variety which is used for official purposes and has several registers. Hence, it is not a point but a range which defines Standard Hindi” (Abbi & Sharma: forthcoming).
4 Contact Hindi of Bombay
5 Contact Hindi of Calcutta
6 An auxiliary language which is used only for intergroup communication and has small vocabulary and grammatical morphology is called restricted pidgin (Siegel 2008: 3).
nearing towards becoming creoles. “Hindi is a fast growing link language and has reached to all corners of India. Thus, the classification proposed by (Das 1997) needs to be elaborated. It should also include the speakers whose primary language is an Austro-Asiatic language or a Tibeto-Burman language or languages of Andaman. Extensive bi/multilingualism prevailing in the country contributes in shaping different varieties of Hindi. Each of these varieties is marked by distinct structure and has emerged in varying circumstances pertaining to historical, political, and geographical factors” (Sharma 2013: 5).

In the Hindi states, this bilingualism is constituted of varieties of Hindi (Census 2001). For instance, Magahi and Bhojpuri in Bihar, Braj and Awadhi in Uttar Pradesh, Marwari and Mewati in Rajasthan, Kumaoni and Garhwali in Uttarkhand, Chattisgarhi in Chattisgarh. It is important to note that though all these varieties have their own unique literary traditions yet these are considered varieties of Hindi. In the Hindi belt, CH develops as a combination of regional languages and SH. It was observed in Jharkhand that “the young generation is happy to forget their mother tongues. The elders prefer to teach their children Hindi than their indigenous languages”8 (Abbi 1997:306).

In the non-Hindi belt, the emergence of CH is needs-based. Such CH is a source of communication with non-Hindi speakers as well as to the Hindi speakers in that place. It is clear that there are several sources of input for CH such as, schools, television, newspapers etc. But, the speaker does not receive SH as an input. Be it from the restricted formal teaching at school and offices, language learning from television9, partial multi-acquisition10 of various tongues or the language of trade and work. The speaker does not receive SH from any source. In fact, they receive the simplified version of Hindi produced by the native Hindi speakers and the regional Hindi from other speakers such as the Eastern Indo Aryan (EIA) speakers in NE. At this point, it is significant to note that the effort is towards knowing Hindi that can be used for daily communication. Neither the native speakers nor the learners aim towards SH. Hence, the result is the emergence of a variety which is not close to the standard but has acceptability and prestige among the users. It is this Hindi that we are going to discuss in the following pages. We have taken Meghalaya state of the Northeast (NE) India. We shall call this Hindi as Meghalaya Hindi (MH).

1.2. Contact Hindi in NE

Not much has been written about the inception and development of Hindi in NE India. There are no records of the forms and pattern of early CH in NE. Earlier Assamese was the link language in NE. But CH is gradually replacing it as a link language. It is primarily because of the increased government role in the governance, administration and education of the area. It is most likely that in the coming years CH will completely replace Assamese in the role of link language. The development of Hindi language in the NE region is undertaken through Central Hindi Directorate (CHD) and Kendriya Hindi Sansthan (KHS), Agra. The objective of the Directorate is propagation and development of Hindi as a link language. The objective of the Sansthan is to improve the standards of teaching of Hindi at various levels.11 The Central

---

7 When the pidgin becomes the everyday lingua franca in a multilingual community and is used in government and religion, it expands lexically and grammatically. Such pidgin is called an expanded pidgin (Siegel 2008: 3).
8 Exception is Santhali
9 The Hindi used in television is primarily Bombay Hindi and not SH.
10 The children whose parents are from different tribes or community often learn several languages at home. Apart from distinct languages of parents they also learn CH which is used as link language between parents. In addition to this, such children learn dominant regional language and official language outside home.
Hindi Directorate was established in 1960 by the Presidential order in pursuance of the special directive of the Constitution envisaged in Article 351, which reads as under:

"It shall be the duty of the Union to promote the spread of the Hindi language, to develop it so that it may serve as a medium of expression for all the elements of the composite culture of India and to secure its enrichment by assimilating without interfering with its genius, the forms, style and expressions used in Hindustani and in the other languages of India specified in the Eighth Schedule, and by drawing, wherever necessary or desirable, for its vocabulary, primarily on Sanskrit and secondarily on other languages."\(^{12}\)

The development of CH in NE and its growth is closely linked to the development of the hill frontiers as full-fledged states and hence the penetration of administrative policy in form of induction of soldiers, teachers, policemen, guides, officers etc (Modi 2005: 30). The contributing factors behind the spread of Hindi as a link language are several. These are:

a. Hindi speaking states cover vast geographical areas that make them gigantic participants as entrepreneurs and consumers in the trade and economic practices. Thus, for economic benefits the non-Hindi speakers have to learn Hindi.

b. In addition to this, affirmative Government policies such as, Article 351 of the constitution makes Union Government through the Ministry of Education and Youth Services responsible for the promotion and development of Hindi.

c. There are several autonomous institutes in AP and Meghalaya that are working for the promotion of Hindi.

d. Media and television are majorly responsible for the spread of Hindi in NE.

The CH of NE develops in typologically rich environment that comprises participants from several language families. It is also correct that contact languages often develop interesting typological features as a contact-induced change. For instance, “Bombay Hindi has an instance of such contact-induced change where the question particle has switched from sentence-initial to sentence final position under Marathi influence” (Thomason & Kaufman 1998: 98). Thus, it would be interesting to note the features of MH that develop in the environment, which comprises Eastern Indo Aryan languages, English, Hindi and Austro-Asiatic languages.

1.3. Genesis of MH

Meghalaya is a multi-ethnic state in character and has people from various parts of the country and neighbouring countries such as Assam, Bihar, West Bengal, Uttar Pradesh, Bangladesh, Nepal etc. These outsiders/non indigenous population bring along their languages in Meghalaya. Thus, several languages can be found in Meghalaya. The primary languages spoken there are Khasi, Jaintia, Garo, Nepalese, Bangla and English. Out of all these languages, English is the official language. “Khasi and Garo have recently been declared as the associate official languages under Meghalaya State Language Act 2005\(^{13}\)”.

Further, the Indo-Aryan and the EIA language speakers bring Hindi and regional Hindis into Meghalaya. The consolidation of regional Hindis is used as a means of communication among non indigenous population in Meghalaya. The primary sources through which various forms of Hindi flourish in Meghalaya are through migrants, traders, media, schools and

---


government workers. The non-indigenous language speakers occupy several significant positions in Meghalaya such as administrative officers, teachers, traders, vendors and migrants. Thus, Hindi is the language of governance and service providers. Hence, MH, a variety of Hindi is used as a link language. It is clear that IA and EIA language speakers are the initial perpetrators of MH. For inter-communication these speakers produce a consolidation of their individual regional Hindis. This regional Hindi is the primary input for the indigenous speakers who further innovate it through semantic extension and readjustments. (In the non-Hindi belt, this is a typical pattern for the development of CHs.) As aforementioned in §1.1, the speaker does not receive SH from any source. Hence, MH resultanty deviates from SH. We can demonstrate the development of MH in Figure 1.

![Figure 1: Genesis of Meghalaya Hindi](image)

It is clear in Figure 1 that IA speakers establish the base of MH through their regional Hindis as the primary input. Thus, in Figure 1 the regional Hindis are denoted by subscript numerals. For instance, (A1)H1) can be read as Assamese Hindi which is a type of Hindi that is recognizably Hindi because of Hindi lexicon but has Assamese grammatical features. Further, the consolidation of these regional Hindis serves as the primary input for indigenous language speakers of Meghalaya. Initially, these speakers only use simplified structures from the received input. But gradually, with the increase in sources of learning, constant communication and increased usage, they innovate the simple structures with the inputs from their own indigenous languages. Hence, MH is formed with forms and features of both indigenous and non indigenous languages.

---

14 Khasi, Jaintia and Garo are the primary indigenous languages of Meghalaya. The first two are Austro-asiatic languages and Garo is a Tibeto-Burman language. Thus, both the speakers have different varieties of CH. But in this study we have focused only on the most common features of MH used by all ethnic groups.
1.4. Phonology and language contact

Phonological change can be an addition, deletion or modification of sounds. “The phonic interference arises when a bilingual indentifies a phoneme of the secondary system and, in reproducing it, subjects it to the phonetic rules of the primary language” (Weinreich 1953: 14).

There are instances of entire system convergence where the complete phonological system is borrowed. The clear instance can be found in South Asian English, the phoneme /t/ (in English an aspirated dento-apical stop) is usually pronounced as an unaspirated retroflex [ʈ] (Matras 2009: 223). (Matras 2009: 225) has proposed four processes that demonstrate the different types of contact-induced phonological changes. These types can be demonstrated in Table 1.

### Table 1: Types of processes leading to contact-induced phonological change

<table>
<thead>
<tr>
<th>Type</th>
<th>Process</th>
<th>Description</th>
<th>Speakers/bilingualism</th>
<th>Language attitudes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Phonological adaptation of word-forms</td>
<td>Replicated word-forms are adjusted to match the sound patterns of the recipient language</td>
<td>Semi-bilinguals or monolinguals</td>
<td>Strong loyalty towards, and stability of the recipient language; superficial contact</td>
</tr>
<tr>
<td>B</td>
<td>Borrowing of phonological features along with word-forms</td>
<td>Borrowed and inserted word-forms maintain (fully or partly) the original sound patterns of the donor language (‘authentication’)</td>
<td>Fairly widespread bilingualism</td>
<td>Flexibility in the use of the recipient Language, prestigious bilingualism</td>
</tr>
<tr>
<td>C</td>
<td>Convergence of systems during second-language acquisition</td>
<td>Word-forms of the target language are systematically adjusted to match the sound patterns of the native language</td>
<td>Emerging bilingualism; stable minority bilingualism; emergence of ethnolect or language shift</td>
<td>Strong group identity coupled with a need (pressure) to acquire the target language</td>
</tr>
<tr>
<td>D</td>
<td>Convergence of systems in stable, intensive bilingualism</td>
<td>Sound patterns of the native language are adjusted to match those of the second language</td>
<td>Intensive and widespread bilingualism</td>
<td>Second language is ‘prestige’ language</td>
</tr>
</tbody>
</table>

In each type of phonological change, the speakers, language attitudes and the description varies. In the present paper, the phonology of MH will be examined through these various types of contact-induced processes. This discussion will be done while concluding the study.

2. Phonological outcomes: data and results

The phonological changes in contact situation are a result of several linguistic and social variables such as age, gender, education, occupation, bi/multilingual profile etc. It is correct to state that “phonological replication or ‘interference’ or ‘transfer’ or ‘borrowing’ may affect
any level of sound structure: the articulation of individual phones or phonemes within words, length and gemination, stress and tone, prosody and intonation” (Matras 2009: 222). In the case of MH, Hindi is the lexifier and other indigenous and non indigenous languages in Meghalaya are other source languages. Thus, the phonological inventory is a mix of all these languages.

MH has borrowed lexicon and grammatical categories from several languages such as Hindi, EIA languages and Khasi (Sharma 2012). Thus, in the process it has also borrowed several sounds or adapted words forms as per their sound system. The Hindi lexicon borrowed in MH also undergoes several changes, such as where Hindi has a short schwa Assamese/Bangla has low rounded vowel. Thus, MH speakers who are bilinguals in Assamese/Bangla replace schwa /ə/ by a low rounded vowel. This low round vowel has made way along with Assamese lexical items and now occurs in MH. Further, there is an addition of a lamino-dental stop phoneme in MH inventory (as illustrated in R.MH.14) replacing retroflex stops of Hindi. MH does not have the front higher-low vowel /ɛ/. In fact, it uses the higher-mid vowel /e/ of Hindi in its place as seen in Hindi lexicon of MH. The absence of higher-low vowel in MH may be because Khasi also does not have the higher-low vowel. Thus, MH has omitted, substituted and adapted several sounds that it has borrowed from Hindi. Before moving ahead, it is important to demonstrate the sound system of Hindi, MH and Khasi. These can be exhibited in Table 2, Table 3, Table 4, Table 5 and Table 6.

Table 2: Vowel Chart of Hindi

<table>
<thead>
<tr>
<th></th>
<th>Front rounded</th>
<th>Central unrounded</th>
<th>Back +/- round</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i</td>
<td></td>
<td>u</td>
</tr>
<tr>
<td>Low high</td>
<td>i</td>
<td></td>
<td>o</td>
</tr>
<tr>
<td>Higher mid</td>
<td>e</td>
<td></td>
<td>o</td>
</tr>
<tr>
<td>Mean mid</td>
<td>e</td>
<td></td>
<td>e</td>
</tr>
<tr>
<td>Higher low</td>
<td>e</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>a</td>
<td></td>
<td>a</td>
</tr>
</tbody>
</table>

Table 3: Vowel chart of Meghalaya Hindi

<table>
<thead>
<tr>
<th></th>
<th>Front rounded</th>
<th>Central unrounded</th>
<th>Back +/- round</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i</td>
<td></td>
<td>u</td>
</tr>
<tr>
<td>Low high</td>
<td>i</td>
<td></td>
<td>o</td>
</tr>
<tr>
<td>Higher mid</td>
<td>e</td>
<td></td>
<td>o</td>
</tr>
<tr>
<td>Mean mid</td>
<td>e</td>
<td></td>
<td>o</td>
</tr>
<tr>
<td>Higher low</td>
<td>e</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>a</td>
<td></td>
<td>a</td>
</tr>
</tbody>
</table>

The vowel chart of MH is different from Hindi. MH does not have a higher low vowel /ɛ/. In Hindi lexicon of MH it is replaced by higher mid vowel /e/.

Further, the consonant inventory of MH also deviates greatly from Hindi. It lacks retroflexes, aspirated affricates, palatal nasals etc. Unlike Hindi, MH consonant chart has an additional place of articulation i.e. lamino-dentals replacing retroflexes of Hindi. The consonant chart of Hindi, MH and Khasi are demonstrated in Table 4, Table 5 and Table 6 respectively.
### Table 4: Consonant Chart of Hindi

<table>
<thead>
<tr>
<th></th>
<th>Bilabial vl vd</th>
<th>Labiodental vl vd</th>
<th>Dental vl vd</th>
<th>Alveolar vl vd</th>
<th>Retroflex vl vd</th>
<th>Palatal vl vd</th>
<th>Velar vl vd</th>
<th>Glottal vl vd</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stop</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-asp</td>
<td>p b</td>
<td></td>
<td>t d</td>
<td></td>
<td>l d</td>
<td>k g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+asp</td>
<td>pʰ bʰ</td>
<td></td>
<td>tʰ dʰ</td>
<td></td>
<td>lʰ dʰ</td>
<td>kʰ gʰ</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fricative</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Affricate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nasal</strong></td>
<td>m</td>
<td></td>
<td>n</td>
<td></td>
<td>n</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Flap or trill</strong></td>
<td></td>
<td></td>
<td>r</td>
<td></td>
<td>r</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Approximant</strong></td>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 5: Consonant chart of Meghalaya Hindi

<table>
<thead>
<tr>
<th></th>
<th>Bilabial vl vd</th>
<th>Labiodental vl vd</th>
<th>Dental vl vd</th>
<th>Alveolar vl vd</th>
<th>Lamino-dentals vl vd</th>
<th>Palatal vl vd</th>
<th>Velar vl vd</th>
<th>Glottal vl vd</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stop</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-asp</td>
<td>p b</td>
<td></td>
<td>t d</td>
<td></td>
<td>t d</td>
<td>k g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+asp</td>
<td>pʰ</td>
<td></td>
<td>tʰ</td>
<td></td>
<td>tʰ</td>
<td>kʰ</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fricative</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Affricate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nasal</strong></td>
<td>m</td>
<td></td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Flaps or trill</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Approximant</strong></td>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

200
Table 6: Consonant chart of Khasi

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Retroflex</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>vl v d</td>
<td>vl v d</td>
<td>vl v d</td>
<td>vl v d</td>
<td>vl v d</td>
<td>vl v d</td>
<td>vl v d</td>
<td>vl v d</td>
</tr>
<tr>
<td>Stop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-asp</td>
<td>p b</td>
<td></td>
<td>t d</td>
<td></td>
<td>k</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+asp</td>
<td>pʰ bʰ</td>
<td></td>
<td>tʰ dʰ</td>
<td></td>
<td>kʰ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td></td>
<td></td>
<td>s</td>
<td></td>
<td>h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>jʰ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>j</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td></td>
<td>n</td>
<td></td>
<td>p</td>
<td>η</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+asp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flaps or trill</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laterals</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the present paper, the phonological rules are represented as ‘R.MH.x,’ where ‘R’ stands for rule, MH stands for Meghalaya Hindi and ‘x’ is occupied by a roman number. For instance, R.MH.3 can be read as ‘third rule of Meghalaya Hindi.’ Further, the social parameters reflect on the variation in sound system. Thus, in echo word formation the replacing vowel and consonant in echo word varies according to the age, education and exposure to Hindi as illustrated in R.MH.1 and R.MH.2. There are other cases where there is a phonemic variation because of difference in tribes. The Khasi and Garo language speaker speak distinctly as illustrated in R.MH.8. All this can be demonstrated through following phonological rules where we begin with the phonology of echo word formation and then other phonological rules found in Hindi lexicon of MH.

3. Phonological rules

3.1 Partial reduplication: rule of replacive morpheme

In echo word formation, MH follows two rules of replacing the initial sound and one rule of vowel alternation. Due to insufficient data, no generalization can be derived on the vowel alteration. Thus, we will only demonstrate the rule of initial sound replacement in this section.

In the rule of replaciveness, a +back, +high, +rounded vowel /u/ replaces the first syllable of a lexeme. This rule is found among the –young MH speakers and the ones who have no formal training of Hindi. This can be represented in R.MH.1 and exemplified in (1) as follows:

---

16 By formal training we refer to Hindi language learning at school.
3.1.1. First Rule of Meghalaya Hindi (R.MH.1)

\[
\begin{align*}
\text{MH Single lexeme} & \quad \text{MH Partial reduplication} \\
\{ \text{young} \} & \quad \{ \text{formal training of H} \} \\
\[\text{ru} \, \text{ŧ}=\text{young} & \quad [\text{ru} \, \text{ŧ} \, \text{u} \, \text{ŧ}]=\text{bread and other edible things} \]
\end{align*}
\]

Thus, example 1 illustrates the echo word formation where the lexical noun \text{ru} ‘bread’ becomes \text{ru} \, \text{ŧ} \, \text{u} \, \text{ŧ} ‘bread and other edible things.’

In another rule of replaciveness in echo word formation, any initial consonant is replaced by a +voice, +labiodentals, +approximant consonant /v/. This rule is found among the speakers who are +young and have formal training of Hindi. This rule can be represented in R.MH.2 and exemplified in (2) as follows:

3.1.2. R.MH.2

\[
\begin{align*}
\text{MH Single lexeme} & \quad \text{MH Partial reduplication} \\
\{ \text{young} \} & \quad \{ \text{formal education of H} \} \\
\[\text{ru} & \quad [\text{ru} \, \text{v} \, \text{u} \, \text{ŧ}]=\text{bread and other edible things} \]
\end{align*}
\]

3.2 Word final consonant cluster simplification

Khasi does not allow consonant clusters, word finally\textsuperscript{17}. This rule is borrowed in MH also and is seen in Hindi lexicon. The cluster simplification is done by either deletion of one consonant as shown in R.MH.3 and exemplified in (3) or by vowel insertion as shown in R.MH.4 and exemplified in (4). It can be demonstrated as follows:

3.2.1. R.MH.3

\[
\begin{align*}
\text{MH Single lexeme} & \quad \text{MH Partial reduplication} \\
\{ \text{young} \} & \\
\[\text{dork} \, \text{ŧ} \, \text{u} \, \text{ast} & \quad \text{dork} \, \text{ŧ} \, \text{a} \, \text{ŧ} \quad \text{‘request’} \]
\end{align*}
\]

3.2.2. R.MH.4

\[
\begin{align*}
\text{MH Single lexeme} & \quad \text{MH Partial reduplication} \\
\{ \text{young} \} & \\
\[\text{dord} & \quad \text{dord} \, \text{ŧ} \quad \text{‘pain’} \]
\end{align*}
\]

\textsuperscript{17} Source: Khasi language structure accessed on http://www.lisindia.net/Khasi/Khasi.html on 7/7/2013.
3.3 Phonological changes in the Hindi lexicon

In language contact the speakers replace the sounds of the target language with the sounds of their own language. (Matras 2009: 226) points out that “the speakers re-interpret the sounds in a loanword or target language by matching them to sound patterns of the recipient or native language. This procedure of approximation enables the production of foreign word-forms within a familiar phonological framework.” Similar approximation is found in Indian English also as in substitution of English labiodentals fricative /f/ by SH voiceless bilabial aspirate /pʰ/. There are several such substitutions in MH also. These are:

3.3.1. Devoicing

The MH speakers who receive formal training in Hindi and use it regularly do not devoice consonants. The devoicing is an obligatory phenomenon among people who lack formal training in Hindi or who do not use it regularly. This rule can be represented in R.MH.5 and exemplified in (5), (6) and (7) as follows:

3.3.2. R.MH.5

\[
\begin{align*}
\text{[+voice]} & \rightarrow \text{[−voice]} \quad / \quad \{−\text{formal training in H} \} \quad \text{such as,} \\
H & \quad MH & \quad \text{Gloss} \\
(5) & \quad \text{nasib} & \quad \text{nasip} & \quad \text{‘luck’} \\
(6) & \quad bʰaša & \quad pəsa & \quad \text{‘language’} \\
(7) & \quad subəh & \quad šo̱pe & \quad \text{‘morning’}
\end{align*}
\]

It can be seen in examples (5), (6) & (7) that voiced sounds are devoiced at final, initial and medial positions respectively. It is also important to note that Khasi does not have voiced velars (refer table 6) thus the Khasi speakers of MH who lack formal training in Hindi devoices /g/ and /gʰ/ to /k/ and /kʰ/ respectively. Though, now with increased sources of Hindi learning, voiced velars are evident in MH. The rule of devoicing of velars can be demonstrated through instances from Hindi lexicon in MH in examples (8) and (9) as follows:

3.3.3. R.MH.6

\[
\begin{align*}
\text{[+voice velar]} & \rightarrow \text{[−voice velar]} \quad / \quad \{−\text{formal training in H} \} \quad \text{such as,} \\
H & \quad MH & \quad \text{Gloss} \\
(8) & \quad gaʃɔr & \quad kafɔr & \quad \text{‘carrot’} \\
(9) & \quad gʰɔr & \quad kʰɔr & \quad \text{‘home’}
\end{align*}
\]

3.3.4. Fricativization

Khasi does not have palatal stops and affricates (refer table 6). Thus, the Hindi palatal affricates /c/ and /cʰ/ are realized as fricatives /s/ or /ʃ/ by MH speakers because fricatives are the nearest substitute. The lack of affricates in Khasi triggers fricativization in MH. This rule is also superimposed by Assamese influence on MH. Thus, fricativization can be found in the
MH speakers who are either bilinguals in Khasi or Assamese. The rule can be represented in R.MH.7 and exemplified in (10) and (11) as follows:

### 3.3.5. R.MH.7

$$\begin{align*}
\{+\text{affricate}\} & \rightarrow \{+\text{fricative}\} \big/ \{+\text{Khasi bilingual/ +Assamese bilingual}\} \\
\text{such as,}
\end{align*}$$

<table>
<thead>
<tr>
<th>H</th>
<th>MH</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>əččʰa</td>
<td>əšša</td>
<td>‘good’</td>
</tr>
<tr>
<td>čay</td>
<td>ša</td>
<td>‘tea’</td>
</tr>
</tbody>
</table>

### 3.3.6. Spirantization

A sound change where a stop changes to a fricative is called spirantization. The Garo speakers of MH typically show this sound change. Garos replace bilabial stop /p/ of Hindi with a labiodental fricative /f/ while using MH. “This p~f alternation is commonly found in NE languages where IA words are borrowed” (Dey, K: personal communication). The phonological rule for such substitution can be represented in R.MH.8 and exemplified in (12), (13) and (14) below:

### 3.3.7. R.MH.8

$$\begin{align*}
\{+\text{bilabial} \\
\{+\text{stop}\} & \rightarrow \{+\text{labiodentals} \}
\end{align*}$$

$$\begin{align*}
\{+\text{fricative}\} \big/ \{+\text{Garo speaker}\}
\end{align*}$$

<table>
<thead>
<tr>
<th>H</th>
<th>MH</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>kəpɾa</td>
<td>kəfɾa</td>
<td>‘cloth’</td>
</tr>
<tr>
<td>pura</td>
<td>fura</td>
<td>‘whole’</td>
</tr>
<tr>
<td>nap</td>
<td>naf</td>
<td>‘measurement’</td>
</tr>
</tbody>
</table>

It was interesting to find that a single speaker can display both instances of absence and presence of this rule. Such presence of parallel structures is common in contact situations. “The languages pass through the stage of parallel structures which are in free variations (and thus create optional use)” (Abbi 2000: 53). Thus, clearly MH has a parallel use of absence and presence of the spirantization rule.

### 3.3.8. Lateralization

The intervocalic /ɾ/ or which occurs between a vowel and a consonant, changes to /l/ in the speech of –young and less educated speakers. This can be represented in R.MH.9 and exemplified in (15) as follows:
10. Phonological changes in the Hindi lexicon

3.3.9. R.MH.9

\[
\begin{cases}
+\text{trill} \\ +\text{lateral}
\end{cases} / \begin{cases}
-\text{young} \\ +\text{less educated}
\end{cases}
\]
such as,

<table>
<thead>
<tr>
<th>H</th>
<th>MH</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(kərof)</td>
<td>(klur)</td>
<td>‘crore’</td>
</tr>
</tbody>
</table>

3.3.10. Trillation

In MH, the –young and +less educated speakers replace lateral \(/l/\) by trill \(/r/\). This can be represented in R.MH.10 and exemplified in (16) and (17) as follows:

3.3.11. R.MH.10

\[
\begin{cases}
+\text{lateral} \\ +\text{trill}
\end{cases} / \begin{cases}
-\text{young} \\ +\text{less educated}
\end{cases}
\]
such as,

<table>
<thead>
<tr>
<th>H</th>
<th>MH</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(ləsən)</td>
<td>(rəsən)</td>
<td>‘garlic’</td>
</tr>
<tr>
<td>(rəməl)</td>
<td>(rəmar)</td>
<td>‘handkerchief’</td>
</tr>
</tbody>
</table>

3.3.12. Labialization

In MH, the labio-dental approximant of Hindi is substituted by a bilabial stop among the speakers who are less educated. In addition to this, the contact with EIA languages superimposes this substitution because unlike Hindi, EIA languages like Assamese, Nepalese and Bangla do not have a labio-dental approximant \(/\nu/\). Thus, MH replaces the Hindi sound by a nearest substitute. This can be represented in R.MH.11 and exemplified in (18), (19) and (20) as follows:

3.3.13. R.MH.11

\[
\begin{cases}
+\text{labio-dental} \\ +\text{approximant}
\end{cases} / \begin{cases}
+\text{bilingual EIA language}
\end{cases}
\]
such as in,

<table>
<thead>
<tr>
<th>H</th>
<th>MH</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(jəvəb)</td>
<td>(jəbəb)</td>
<td>‘reply’</td>
</tr>
<tr>
<td>(həuə)</td>
<td>(həba)</td>
<td>‘wind’</td>
</tr>
<tr>
<td>(nəv)</td>
<td>(nəb)</td>
<td>‘boat’</td>
</tr>
</tbody>
</table>

4. Vowel raising

In MH, the Hindi higher-low vowel \(/e/\) is raised to higher-mid vowel \(/e/\) because Khasi does not have a higher-low vowel. Such vowel raising is found in CH of Arunachal Pradesh also. The rule of vowel raising can be represented in R.MH.12 and exemplified in (21) and (22) as follows:
4.1. R.MH.12

\[
\begin{align*}
(+\text{low mid}) & \rightarrow (+\text{higher mid}) / \{+\text{Khasi speaker}\} \\
H & \quad \text{MH} & \quad \text{Gloss} \\
(21) & \quad \text{bet}^h & \quad \text{bet} & \quad \text{sit}' \\
(22) & \quad \text{per} & \quad \text{per} & \quad \text{foot/leg}'
\end{align*}
\]

5. Rounding

The MH speakers who lack formal training in Hindi and are bilinguals in Assamese/Bangla replace Hindi schwa /ə/ by low rounded vowel of EIA languages.\(^\text{18}\) This can be represented in R.MH.13 and exemplified in (23) as follows:

5.1. R.MH.13

\[
\begin{align*}
(+\text{central}) & \rightarrow (+\text{back}) / \{–\text{formal training in H} \\
–\text{round} & \quad +\text{round} / \{+\text{bilingual Assamese/Bangla}\} \\
\text{such as,} & \\
H & \quad \text{MH} & \quad \text{Gloss} \\
(23) & \quad \text{məʈəɾ} & \quad \text{məɾ} & \quad \text{peas}'
\end{align*}
\]

6. Phoneme addition

Fronto/Lamino-dentals\(^\text{19}\) sounds in Khasi were first discovered by (Abbi 1979: 227–229, 1987: 101) by doing the deep feature study of the Khasi sounds. She noted that these sounds are phonetic variants of alveolar, retroflex and dental stops and have been ignored in the phonetic literature. She added that “these sounds occur word finally after a glide. In order to explain these sounds she introduced a new phonetic term then ‘contact at lower teeth’” (Abbi 1979: 227–229). Henderson in 1967 termed these sounds as ‘puzzling’. But now we see one of these variants as a stable phoneme in MH. This can be considered as an innovation in MH as compared to SH sound system. Thus, retroflex stops of Hindi are realized as lamino-dental stops among all MH speakers whose first language is Khasi. Thus, it is a phonological interference from Khasi phonemic inventory. But the phonotactic constraints of Khasi that these sounds occur word finally after a glide does not apply to MH. The lamino dental stops can be represented as /t/ and /d/. The addition of lamino-dentals can be represented in R.MH.14 and exemplified in (24), (25) and (26) as follows:

---

\(^\text{18}\) Eastern Indic languages have low rounded vowels whereas Western Indic have schwa, short /a/. Thus, Hindi has a schwa and Assamese/Bangla has low rounded vowel. As an effect of Assamese/Bangla MH speakers who are bilinguals in Assamese/Bangla has low rounded vowel.

\(^\text{19}\) In its production the blade of the tongue is pressed against the lower teeth, and the body of the tongue is raised high towards the pre-palatal region.
10. Phonological changes in the Hindi lexicon

6.1. R.MH.14

\[
\begin{array}{c}
\{+\text{retroflex} \} \\
\{+\text{stops} \}
\end{array} \rightarrow \begin{array}{c}
\{+\text{lamino-dental} \} \\
\{+\text{stop} \}
\end{array} / \{+\text{khasi speaker} \}
\]

<table>
<thead>
<tr>
<th>H</th>
<th>MH</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(24) ṭokri</td>
<td>tukri</td>
<td>‘basket’</td>
</tr>
<tr>
<td>(25) ṭuṭ</td>
<td>tūṭ</td>
<td>‘break’</td>
</tr>
<tr>
<td>(26) ḍibiya</td>
<td>ḍibya</td>
<td>‘small box’</td>
</tr>
</tbody>
</table>

Thus, it is correct to state that MH does not always simply borrow from source languages. In fact, it omits, substitutes and innovates sounds. This has been found through studying the MH lexicon borrowed from Hindi in above examples. These phonological processes are contact-induced from Khasi and other source languages. Both social and linguistic factors, such as, education, bi/multilingualism profile, age etc, determine the phonological sketch of MH. Apart from the aforementioned phonological outcomes, several other changes were also found in MH. But due to lack of data, we cannot generate generalizations for these. For instance, it is observed that the people who work on lower jobs often centralize the vowels of loanwords from English in MH. Thus, ṭeksi is produced as ṭaksi ‘taxi.’ Further, sometimes a vowel occurring in the first syllable after a single consonant changes to high, central, unrounded vowel [Ɨ], such as, ḡʰ̃ənta becomes ƙạnta ‘hour.’

7. Conclusion

The conclusion is that various social and linguistic variables are determinant for the contact-induced phonological changes in MH. These variables are education, age, type of exposure to Hindi (formal training or not) and linguistic profile. There is certainly emerging bilingualism; stable minority bilingualism, and emergence of separate contact variety called MH. But unlike Matras’s assessment there is no language shift from Khasi to MH among the bilinguals. Khasi speakers have strong language loyalty. Though, there is certainly a need (pressure) to acquire MH because it is the variety of language of administration and governance. The phonological inventory of MH is not just conditioned by Khasi but also by other participant languages such as regional Hindis and EIA languages (Assamese, Bangla, Nepalese etc.) Thus, not all the Hindi sounds make way into MH phonemic inventory via Hindi lexicon. In fact, these sounds undergo substitution, omission, addition and innovation which ultimately form the phonemic inventory of MH that is distinct to Hindi phonemic inventory.

These contact-induced phonological changes of MH are nearly symmetrical to Matras’s paradigm. The types of changes of MH can be explained from type A-C (Table1). To begin with, the semi bilinguals phonologically adapt the word forms according to the sound pattern of their language. For instance, the Khasi speakers do not allow the final consonant clusters in Hindi lexicon because Khasi also does not allow final consonant clusters. In other instances, the phonological features are borrowed along with forms. For instance, MH has low rounded vowel of Assamese rather than Hindi schwa. In this case, Khasi speakers have borrowed the low rounded vowel along with Assamese word forms. Lastly, there are cases of system convergence also. For instance, MH speakers replace Hindi retroflex by Khasi lamino-dentals because Khasi does not have retroflexes.

Lastly, intensive and widespread bilingualism will soon appear in Meghalaya because the language attitudes are slowly changing. The speakers are gradually realizing that Hindi is a
prestigious language. Thus, such change in language attitude will be coupled with a change in type of bilingualism and its phonological correlates. With such a change, the Hindi sounds might also make way into the sound system of Khasi. In fact, this can be a potential research area for further studies in language contact.
10. Phonological changes in the Hindi lexicon

Abbreviations

AH    Arunachalee Hindi
CH    Contact Hindi
EIA    Eastern Indo Aryan
H    Hindi
MH    Meghalaya Hindi
NE    North East
R.MH    Rule of Meghalaya Hindi
SH    Standard Hindi
References

11. Spirantization in Tibeto-Burman-Bengali bilinguals in south of Assam

Kakoli Dey
University of Delhi, Delhi

Abstract
Silchar Bengali, unlike other Indo-Aryan languages or its westerly Bengali dialects, shows a tendency for spirantization of plosives. Specifically, the voiceless bilabial plosive [p] and the aspirated bilabial plosive [ph] are variably realized as the labio-dental fricative [f]. This study seeks to address how the local bilingual Tibeto-Burman groups (such as Bodos, Meitheis, Bishnupriyas, etc.) respond to the variable spirantization, found in the Bengali monolingual population in Silchar. The research questions addressed in the study are as follows. (1) To what extent do the various Tibeto-Burman-Bengali bilinguals participate in the process of spirantization? (2) To what extent do the various groups form a part of a single multilingual speech community? Is spirantization a similar sociolinguistic variable across all of the various groups? (3) What role do linguistic factors as well as social and psychological factors play for the variable of spirantization, given the multilingual context?

It is found that bilingual Tibeto-Burman Bengali speakers use spirantization more than bilingual Hindustani Bengali speakers. Among the various Tibeto-Burman groups, the Bishnupriya show a high percentage of spirantization, which is unsurprising given their high level of integration into the Bengali community. However, the Meithei also show a particularly high degree of spirantization. While the various Tibeto-Burman groups differ in terms of the rate of spirantization, the structural constraints on spirantization are shared.

Citation

Volume Editors
Gwendolyn Hyslop, Linda Konnerth, Stephen Morey, Priyankoo Sarmah

Copyright © 2014, the author(s), release under Creative Commons Attribution license

1. Introduction

Assam is located in the north eastern region of India and is home to many diverse languages and cultures. Besides the numerous Tibeto-Burman languages spoken in this part of the region, the two Indo-Aryan languages, Assamese and Bengali have a remarkable presence. Bengali has a long history in this region and adjoining East Bengal going back to pre-colonial period. It is the economic prosperity of the region of Bengal deltas which further promoted Bengali to become an important language in the entire region. Bengali further spread to the neighbouring regions such as Assam, Tripura, Manipur, Meghalaya and elsewhere, owing to the British policies in the region (Satyanath and Laskar 2008).

The paper is an extension of my doctoral work. This paper deals with the issue of the participation of the various local indigenous ethno-linguistic groups (Tibeto-Burman/Bengali bilingual) in the local linguistic changes that are essentially introduced and led by the dominant ethno-linguistic group (Bengali) in a multilingual-multietnic society of Silchar, Assam.

Silchar is located in the southern part of Assam (also known as Barak Valley) under the Cachar district. It shares an international boundary with Bangladesh and has several Tibeto-Burman speaking groups which may be considered as indigenous to the region such as Meithei, Bishnupriya, Dimasa/Kachari, Bodo and Naga. According to Census of India, 2001 (Assam), the total Schedule Tribe population is of about 3,308,570 out of which 18,631 Tibeto-Burman (TB) speakers are found in Cachar district and of about 1,869 TB speakers are in Silchar (Municipality).
Bengali constitutes the dominant language in the region. However, it should be mentioned that Silchar Bengali Vernacular (SBV) is a new dialect that emerged as a direct consequence of the movement of population triggered by the Tea Plantations which started coming up in the early 19th century in this part of the country (Griffith 1967, Satyanath 1998, Dey 2010). It also serves as the lingua franca across the Barak Valley among different linguistic groups (Census of India 1991). Hence, the non-Bengali groups are bilingual in their community language and Bengali (Wise 1883, Grierson 1903).

In Dey (2010), I focused on variation and change in Silchar Bengali. I proposed that Silchar Bengali, unlike other Indo-Aryan languages and westerly dialects of Bengali, shows a tendency for spirantization of plosives which was essentially introduced by the immigrants coming from eastern Bengali speaking regions of erstwhile Bengal (now Bangladesh).

Building on the findings of Dey (2010), the proposed study attempts to approach Bengali in Silchar in a multilingual framework. It seeks to address ‘how the local bilingual Tibeto-Burman groups (such as Meiteis, Bishnupriyas, Bodos, Nagas etc) respond to variation and change of spirantization’, which are essentially the characteristics of the Bengali monolingual population in Silchar, Assam, in a multilingual framework.

The population of Silchar consist of three groups: Bengali, Hindi and Tibeto-Burman (see Figure 1) of which Bengali constitutes the dominant linguistic group. The Bengalis are essentially the immigrants from the adjoining provinces of erstwhile Bengal and maintained ‘Bengali identity’. Secondly, there are those who came from Hindi dialect areas (Bihar, Jharkhand, Chota Nagpur) and maintained a non-Bengali identity, namely ‘Hindustani’\(^1\). The majority of the speakers of this speech community are bilingual in their mother tongue and also in Bengali (Dey 2010). The third group are the local indigenous Tibeto-Burman speakers representing different linguistic speech community and are mostly bilinguals in Bengali and in their home language, who also might be presumably immigrants (Grierson 1903, 1904)\(^2\).

---

\(^1\) The non-Bengali includes not only the so called ‘Hindustani’ but also the ‘Marwaris’ and similar groups whose ancestors arrived as traders and not necessarily as tea-labour. Unlike, the rest of the Tea Gardens in Assam, the language spoken by the Tea Garden labourers is primarily ‘Hindustani’ in this part of Assam.

\(^2\) It is to be noted that various Tibeto-Burman languages are also spoken over large tracts of adjoining East-Bengal (present day Bangladesh) alongside Bengali dialects (Wise 1883, Grierson 1903, Dey 2010).

---

Figure 1: Major speech communities in Silchar

It is to be noted that most of the non-Bengali speakers are bilingual in SBV and their Mother tongue. The first language is learnt at home whereas Bengali is learnt naturally i.e. outside the home domain which is used in other domains and with members of other ethnic groups. Bengali is used widely and spoken everywhere in Silchar and serves as a lingua franca. With this trend, it is now seen that the non-Bengali speakers are now assimilating into the dominant local culture linguistically (Dey 2010).
11. Spirantization in Tibeto-Burman-Bengali bilinguals

Keeping in line with Dey (2010), where I have looked at the Bengali and Hindustani speakers in the participation of spirantization, data from five speakers representing three Tibeto-Burman groups was drawn from Silchar town (Table 1).

Table 1: Speakers of Tibeto-Burman languages

<table>
<thead>
<tr>
<th>Language</th>
<th>No of Speakers</th>
<th>Gender</th>
<th>Quantum of data (natural speech)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kabui Naga</td>
<td>3</td>
<td>Female</td>
<td>1 Hour 30 mins</td>
</tr>
<tr>
<td>Meithei</td>
<td>1</td>
<td>Male</td>
<td>1 Hour 25 mins</td>
</tr>
<tr>
<td>Bishnupriya</td>
<td>1</td>
<td>Male</td>
<td>45 mins</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5 Speakers</strong></td>
<td></td>
<td><strong>3 Hours 40 mins</strong></td>
</tr>
</tbody>
</table>

Silchar on the whole is semi-urban in character. While some of its areas can be considered as relatively urban, other areas can be considered as relatively rural. The data was drawn from the same neighbourhoods within Silchar i.e. from where I carried out my doctoral research work (Dey 2010), keeping in mind the history of these residential settlements, socio-economic, linguistic and the ethnic composition of the areas. The major guiding principles for selection of the speakers I considered were those who are born and brought up in Silchar or those who have spend a considerable number of years in Silchar without any significant breaks.

The study is based primarily on conversational data drawn through sociolinguistic interviews. The five different Tibeto-Burman speakers with different language backgrounds were interviewed in local Bengali vernacular. This paper explores the nature and behavior of the TB speakers with that of the Bengali and Hindustani speakers in the use of spirantization in a multilingual framework.

2. Spirantization of labial plosives

In SBV, it is the voiceless obstruents [[p ph], [k kh], [c]] that undergo spirantization and are variably realized as [f, x and s] respectively.

The study focuses on voiceless bilabial plosive /p/ and the aspirated bilabial plosive /ph/ which are variably realized as labio-dental fricative [f] as in Figure 2.

---

3 This work is an extension of my doctoral research. The aim of this study is to find out the role of the forces of contact and diffusion in bringing spirantization in SBV and to what extent the TB speakers participate in the process of spirantization. Though they are limited instances of spirantization found attested in TB, it tries to examine whether they are the result of diffusion due to the spread of Bengali across speakers of TB languages or vice versa? Since this is my ongoing research, I would add more speakers eventually from different age groups and social background.

4 Kabui Naga also known as Rongmei or Rongmei Naga are found in various parts of north eastern states. I prefer to use ‘Kabui Naga’ because this is what I found the speakers refer them as, locally. The neighborhood where this study was carried out is situated within Silchar. It is interesting to find that the older speakers refer to ‘Kabui Naga’ more often than the younger generations.
Figure 2: Spirantization of bilabial plosives

More examples on variable realization of /p/ and /ph/, i.e., the same word is realized as a plosive or with a spirant as in Table 2.

Table 2: Examples of labial plosives

<table>
<thead>
<tr>
<th>Variables</th>
<th>Examples</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>[p]:</td>
<td>pore ~ fore</td>
<td>‘after’</td>
</tr>
<tr>
<td></td>
<td>biporith ~ biforith</td>
<td>‘opposite’</td>
</tr>
<tr>
<td></td>
<td>kapor ~ kafor</td>
<td>‘cloth’</td>
</tr>
<tr>
<td></td>
<td>phola ~ pula ~ fula</td>
<td>‘swell’</td>
</tr>
<tr>
<td></td>
<td>phutani ~ putani ~ futani</td>
<td>‘proud’</td>
</tr>
<tr>
<td></td>
<td>maph ~ map ~ maf</td>
<td>‘forgive’</td>
</tr>
<tr>
<td>[ph]</td>
<td>phola ~ fula</td>
<td>‘swell’</td>
</tr>
<tr>
<td></td>
<td>phutani ~ futani</td>
<td>‘proud’</td>
</tr>
<tr>
<td></td>
<td>phurti ~ furti</td>
<td>‘enjoyment’</td>
</tr>
<tr>
<td></td>
<td>phand ~ fand</td>
<td>‘trap’</td>
</tr>
<tr>
<td></td>
<td>phen ~ fen</td>
<td>‘again’</td>
</tr>
</tbody>
</table>

Spirantization\(^6\) as used in the present study and as is generally understood refers to the transformation of a plosive into its corresponding fricative. Spirantization can affect both voiceless and voiced plosives and the instances of the two are both well attested across languages.

As far as spirantization is concerned, it does not have a well attested long and strong history in Indo-Aryan (IA). In Assamese /s/ is frequently realized as /h/, (Kakati, 1941). Aspirated stops like /jh/ and /ch/ are often realized as affricates in Marathi (Kelkar 1958, Ghatage 1965). Outside Indo-Aryan (p > h) is attested in the history of Kannada (Narasimhai 1941, Gai 1946, Schiffman 1983). However spirantization is well attested in numerous languages spoken outside India (Dey 2010).

The sociolinguistic studies of speech communities argue that linguistic knowledge is a property of a community and not simple of an individual. Therefore a language cannot be understood in isolation and without taking society as the primary unit of investigation.

---

5 In comparison with standard Western Bengali, vowels in SBV, in general, are relatively shorter in duration. As many tokens of /u/ in SBV are derived from /o/ (for e.g. sutto < choto ‘small’), there is an opposition across tokens of [u] in terms of duration. Majority of the tokens of [u] are relatively shorter, thereby blurring the difference between a regular [u] and a checked variant represented as [ù] (for further explanation see Dey, 2010).’

6 A detailed discussion on why spirantization cannot be termed as lenition here is discussed in my Doctoral thesis: ‘A sociolinguistic study: Silchar Bengali’ (2010). Spirantization as a lenition is seen as a process of weakening i.e. a phonological process involving change from a stronger to a weaker one. According to Shiraishi (2006:6) spirantization is a process of weakening (lenition) in many languages, and weakening typically targets prosodically weak positions.
11. Spirantization in Tibeto-Burman-Bengali bilinguals

Though the grammar is shaped by linguistic factors, such a grammar cannot be understood without taking into account the social differentiation and the role of various ethno-linguistic groups present in a speech community. Numerous studies have further shown that the different groups present in a society stand in a different relationship to a given linguistic variable (Roberts 1997, Guy 2011, Stanford 2008, Takemura 2011).

2.1. Spirantization across Tibeto-Burman languages

A quick survey of the Tibeto-Burman languages spoken in Assam based on Grierson (1903 and 1909) suggest that spirantization (labial plosives) though not a very productive process, is nevertheless attested as shown in Table 3.

Table 3: Spirantization of labials across the Tibeto-Burman languages

<table>
<thead>
<tr>
<th>Tibeto-Burman Dialects</th>
<th>Spoken in Assam</th>
<th>Examples</th>
<th>Spirants</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plain Kachari (Bodo)</td>
<td>Darrang Dt.</td>
<td>a-fa ‘my father’</td>
<td>Yes</td>
<td>Grierson (1903: 18-19)</td>
</tr>
<tr>
<td>Mech (Bodo)</td>
<td>Goalpara Dt.</td>
<td>a-fa ‘my father’</td>
<td>Yes</td>
<td>Grierson (1903: 38-40)</td>
</tr>
<tr>
<td>Lalung (Bodo)</td>
<td>Nowgaon Dt.</td>
<td>fā ‘father’</td>
<td>Yes</td>
<td>Grierson (1903: 53-55)</td>
</tr>
<tr>
<td>Dimasa Kachari (Bodo)</td>
<td>North Cachar Dt.</td>
<td>bufa ‘father’</td>
<td>Yes</td>
<td>Grierson (1903: 60-61)</td>
</tr>
<tr>
<td>Hojai (Bodo)</td>
<td>Nowgaon Dt.</td>
<td>pafa ‘father’</td>
<td>Yes</td>
<td>Grierson (1903: 65-67)</td>
</tr>
<tr>
<td>Tipura (Bodo)</td>
<td>State Hill Tipperah</td>
<td>bu-fa-no ‘his-father-to’</td>
<td>Yes</td>
<td>Grierson (1903: 114-116)</td>
</tr>
<tr>
<td>Aka / Hrusso (Siamese-Chinese)</td>
<td>North hills of Assam valley (Darrang)</td>
<td>phum ~ pfumu ‘five’</td>
<td>Yes</td>
<td>Grierson (1909: 582-583)</td>
</tr>
<tr>
<td>Kabui Naga</td>
<td>State Manipur</td>
<td>a-po ‘my father’, paibam ‘place’, fai ‘cloth’</td>
<td>Not Known</td>
<td>Grierson (1903(94): 421-422)</td>
</tr>
<tr>
<td>Bishnupuriya/Mayang</td>
<td>Dist Sylhet</td>
<td>bapoke ‘father’</td>
<td>No</td>
<td>Grierson (1903: 426-428)</td>
</tr>
<tr>
<td>Bishnupuriya/Mayang</td>
<td>State Manipur</td>
<td>pa-ba ‘father’</td>
<td>Yes</td>
<td>Grierson (1903: 432-434)</td>
</tr>
<tr>
<td>Meithhei</td>
<td>State Manipur</td>
<td>bapoke ‘father’</td>
<td>No</td>
<td>Grierson (1904(94): 421-422)</td>
</tr>
</tbody>
</table>

One finds that lexical items of Indo-Aryan origin show variability in spirantization as in \( \text{pop} \sim \text{faf} \) ‘sin’ (see Table 3). Moreover, among the words of Tibeto-Burman origin also one finds variation between [ph] and [f] as in a frequently occurring word ‘father’, ‘place’ and ‘field’. As Silchar is home to various Tibeto-Burman groups, this paper intends to look at how the various Tibeto-Burman linguistic groups present in Silchar (Cachar) respond to

7 The reason why I choose Grierson's data and not recent data from Robert Burling or Shobhana Chelliah is because of the nature of the data i.e. it is one story which has been narrated in different languages which makes sense for a study like this. I have also drawn comparison between (1) apparent time Vs real time (those who came around 1900 and those who were born in subsequent decades); (2) Grierson's Data Vs the oldest speakers born at that time when Grierson carried out his work. Secondly, my focus was to trace the process of change (plosive to spirants) from historical data as well as from real speech data across generations. Moreover, Grierson's study on the languages of this region was carried out in the later part of the century.

8 The etymology of the words cited in the column such as fai ‘cloth’; faithop ‘shoes’ are not known whether the underlying sound is ‘p’ or ‘f’.
spirantization, and whether they share same linguistic constraints with SBV. In the present study my data represents from Kabui-Naga, Bishnupuriya and Meithei speakers in Silchar.

The phonetic system of Kabui-Naga, Bishnupuriya and Meithei (as compiled from Grierson’s texts) is given in Table 4. A quick look at the sound inventory of the three languages suggests that these languages are not particularly rich in spirants.

**Table 4: The consonant pattern in the Tibeto-Burman Languages**

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Dental</th>
<th>(Post) Alveolar</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosives Unasp</td>
<td>p</td>
<td>b</td>
<td>t</td>
<td>d</td>
<td>ch</td>
<td>j</td>
</tr>
<tr>
<td>Asp</td>
<td>ph (bh)</td>
<td>th (dh)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasals</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td>η</td>
</tr>
<tr>
<td>Fricatives Vls</td>
<td>(f)</td>
<td>s/sh</td>
<td></td>
<td></td>
<td></td>
<td>h</td>
</tr>
<tr>
<td>Vd</td>
<td>v</td>
<td>(z)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricates Unasp</td>
<td>ts</td>
<td></td>
<td></td>
<td></td>
<td>tʃ'</td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>tʃh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laterals</td>
<td>l</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on the earliest specimens of Tibeto-Burman languages from Grierson’s *Linguistic survey of India*, there are some interesting sound patterns. Consonants such as /bʰ/, /dʰ/, /g/ in Naga-Bodo sub group occur in borrowed words (as shown in brackets in Table ). Consonants such as /f/ and /z/ (in brackets see Table ) are unknown. The sounds /l/ and /r/ are interchangeable in Kabui (Naga-Bodo sub group) i.e. in compound words /l/ > [r] after a vowel such as lan ‘weath’ ka-ran ‘his wealth’ but it is otherwise [l] as in pahut-lana ‘having run’. This phenomenon is also seen in Meithei language (Grierson 1903: 416). In Meithei, there is a distinction between soft and hard consonants and are often interchangeable such as /b/ ~ /p/; /d/ ~ /t/; /g/ ~ /k/ ; /l/ ~ /n/ and /ch/ ~ /ʃ/.

2.2. Overall distribution of the labial plosives in Indo Aryan (Bengali speakers) and Tibeto-Burman Bengali bilinguals

The over all distribution of the two variables: /p/, /ph/ and their variants in Indo-Aryan (IA) speakers and Tibeto-Burman (TB) speakers are shown in Table 5.
11. Spirantization in Tibeto-Burman-Bengali bilinguals

Table 5: Overall distribution of /p/ and /ph/ in IA Bengali speakers and TB Bengali bilinguals

<table>
<thead>
<tr>
<th>No</th>
<th>Variables</th>
<th>[p]</th>
<th>[ph]</th>
<th>[f]</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>IA(^9)</td>
<td>/p/</td>
<td>%</td>
<td>--</td>
<td>38.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>439</td>
<td>61.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1) *(^{10})</td>
<td>705</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1144</td>
</tr>
<tr>
<td>2.</td>
<td>IA</td>
<td>/ph/</td>
<td>%</td>
<td>--</td>
<td>13.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>--</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>61</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(p+ph)</td>
<td>37.1 %</td>
<td>[447]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>62.9 %</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1205</td>
</tr>
<tr>
<td>1.</td>
<td>TB</td>
<td>/p/</td>
<td>%</td>
<td>--</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>271</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>67</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>338</td>
</tr>
<tr>
<td>2.</td>
<td>TB</td>
<td>/ph/</td>
<td>%</td>
<td>--</td>
<td>07.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>N</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(p+ph)</td>
<td>77.5%</td>
<td>[272]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>351</td>
</tr>
</tbody>
</table>

In Dey 2010, it is seen that among the Indo-Aryan speakers, out of a total 1205 tokens there are more instances of spirants (62.9%) than plosives (37.1%). However, among the Tibeto-Burman speakers, out of a total of 351 tokens, there are more instances of plosives (77.5%) as opposed to spirants (22.3%). It is interesting to note that, the aspirated bilabials show much higher degree of spirantization compared to the unaspirated plosives in both the two linguistic groups. Despite this difference in the frequency of spirantization, the two /p/ and /ph/ behave alike with respect to the linguistic constraints. In addition, the total numbers of tokens of /ph/ are very small. Considering the two behave alike, /p/ and /ph/ were combined and treated as a single labial variable /p/ (Dey 2010).

3. Phonological constraints

Considering the nature of the spirantization in SBV, the findings of the study suggest that spirantization is conditioned by phonological constraints (Dey 2010). In the case of labial plosives, the clusters (consonant sequences), the individual preceding phonological segments, the following vowels and lexical factors constitute the strongest linguistic constraints on spirantization\(^{11}\) (as shown in Table 6). Table 6 shows the summary of all the linguistic constraints which were favoured and which were not favoured that were subjected to multivariate analysis. However, the strongest constraint on spirantization is gemination as it categorically blocks spirantization.

---

\(^{9}\) Here Bengali speakers and Hindustani/Hindi speakers have been clubbed together as IA.

\(^{10}\) * not considered for analysis

\(^{11}\) The linguistic and the socio factors were combined while running the multivariate program. **Multivariate analysis** - a generic term for any statistical technique used to analyze data for more than one variable. The subject of multivariate analysis deals with the statistical analysis of the data collected on more than one (response) variable. These variables may be correlated with each other, and their statistical dependence is often taken into account when analyzing such data.
Table 6: Linguistic constraints on spirantization in SBV

<table>
<thead>
<tr>
<th>No</th>
<th>Constraints which favour</th>
<th>Constraints which disfavour</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Clusters</td>
<td>Geminates</td>
</tr>
<tr>
<td>2</td>
<td>Preceding segments</td>
<td>Word position</td>
</tr>
<tr>
<td>3</td>
<td>Following vowels</td>
<td>Syllabic structure of a word</td>
</tr>
<tr>
<td>4</td>
<td>Lexical constraints</td>
<td>Structure of a syllable</td>
</tr>
<tr>
<td>5</td>
<td>--</td>
<td>Weight of a syllable</td>
</tr>
<tr>
<td>6</td>
<td>--</td>
<td>Compound words and phrases</td>
</tr>
<tr>
<td>7</td>
<td>--</td>
<td>Syllable boundary in a word</td>
</tr>
<tr>
<td>8</td>
<td>--</td>
<td>Following segment</td>
</tr>
<tr>
<td>9</td>
<td>--</td>
<td>Preceding vowel</td>
</tr>
</tbody>
</table>

Even though, in both IA Bengali speakers and TB Bengali speakers, the rate of spirantization seems to behave differently as TB Bengali speakers have lesser rate of spirantization, the linguistic constraints are same.

The effect of clusters favours maximally at the syllable boundary and block spirantization within a syllable in IA Bengali speakers. However, in TB Bengali speakers it categorically blocks spirantization within a syllable (as shown in Table 7 and in Figure 3). Examples of clusters (TB Bengali speakers) are such as:

(1) Clusters within syllable boundary
   
   (a) \textit{pro.cur} \\
       ‘plenty’

   (b) \textit{aro} \textit{praе duи-arai} \textit{gonta rasta} \\
       and almost two-half hour road \\
       ‘Again almost two and a half hours journey’

(2) Clusters across syllable/word boundary

   (a) \textit{he-r khelar gol.fо shuru} \\
       he-GEN play story start \\
       ‘His sports related story started’

   (b) \textit{ami ar far.si.nа} \\
       I again can-Pst.T.Neg \\
       ‘I could not (do) it anymore’

   (c) \textit{pac-jon pac-jon kor-iya gesi} \\
       five-Cl five-Cl do-Prog. went \\
       ‘Five each (we) went’
Table 7: Spirantization of labials in clusters and single consonants in IA and TB SBV speakers

<table>
<thead>
<tr>
<th>Clusters</th>
<th>IA</th>
<th>TB</th>
</tr>
</thead>
<tbody>
<tr>
<td>[f %]</td>
<td>[N]</td>
<td></td>
</tr>
<tr>
<td>Within Syllable boundary</td>
<td>10.2</td>
<td>10/98</td>
</tr>
<tr>
<td>Across Syllable/word boundary</td>
<td>58</td>
<td>69/119</td>
</tr>
<tr>
<td>No Clusters</td>
<td>69</td>
<td>679/988</td>
</tr>
<tr>
<td>Total</td>
<td>63</td>
<td>758/1205</td>
</tr>
</tbody>
</table>

Figure 3: The effect of clusters on spirantization in IA and TB SBV speakers

Though the clusters in SBV are formed generally when labials are either preceded or followed by a liquid, sibilant or a nasal, there is an overlapping between the clusters and the effect of the preceding segment. Firstly, the clusters allow us to see what happens when a labial plosive is preceded or followed by another segment. Secondly, the effect of the preceding segment allows us to determine specific effects of the individual segments. The TB Bengali speakers behave in a similar pattern with that of the IA Bengali speakers with regard to the preceding phonological segment as shown in Table 8 and Figure 4. The preceding liquids have the most favourable effect on spirantization compared to all other consonants. Among the liquids the lateral liquids favour spirantization much more than the non-lateral liquids.

Among the preceding segments, the nasals i.e., a bilabial nasal [m-] blocks spirantization categorically within a word as in *shom.pro.dae* ‘creed’, *fam.pa.bari* ‘place name’. However, preceding nasals categorically favour spirantization at word boundaries.

For instance:

(3)  *taim faile jais bashat*  
    time (if)-get go-T home-L  
    ‘If (you) have time then visit my house’

This is because the process of assimilation is weakened between a nasal and an oral plosive across word boundary which in turn facilitates spirantization to operate.
Table 8: Preceding segments and clusters in IA and TB SBV speakers

<table>
<thead>
<tr>
<th>Preceding phonological segments</th>
<th>Clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IA [f %]</td>
</tr>
<tr>
<td>Lateral liquid</td>
<td>80</td>
</tr>
<tr>
<td>Non-lateral liquid</td>
<td>68.8</td>
</tr>
<tr>
<td>Sibilant</td>
<td>50</td>
</tr>
<tr>
<td>Nasal</td>
<td>14.3</td>
</tr>
<tr>
<td>Plosive</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 4: Effect of preceding consonants on spirantization in IA and TB SBV speakers

In addition, there are also constraints which add local character (language specific) to it. As far as the effect of the following segments is concerned, the quality of the following vowels turned out to have a significant effect on spirantization and not the features of the following consonant. That is the findings in IA Bengali speakers do not show any effect on the following consonants, but shows a considerable effect of spirantization on the following vowels (examples in Table 9).

Table 9: Examples of following vowels

<table>
<thead>
<tr>
<th>[a]</th>
<th>[ɔ]</th>
<th>[o]</th>
<th>[u]</th>
<th>[i]</th>
<th>[e]</th>
<th>[e]</th>
</tr>
</thead>
<tbody>
<tr>
<td>fas ‘five’</td>
<td>khafɔr ‘cloth’</td>
<td>fousaia ‘having reached’</td>
<td>fiuja ‘festival’</td>
<td>fison ‘back’</td>
<td>pɛchi ‘(I) got’</td>
<td>pesha ‘job’</td>
</tr>
<tr>
<td>farus ‘to be (hon.)’</td>
<td>ʃɔrika ‘exam’</td>
<td>for.to ‘read-T’</td>
<td>fush.ta-fish ‘post-office’</td>
<td>safiya ‘having caught’</td>
<td>phɛn ‘again’</td>
<td>inisfɛktɔr ‘inspector’</td>
</tr>
<tr>
<td>fara ‘colony’</td>
<td>ʃɔre ‘after’</td>
<td>phon ‘phone’</td>
<td>fishkuni ‘pond’</td>
<td>firot ‘return’</td>
<td>fɛl ‘fail’</td>
<td>fɛnt ‘pant’</td>
</tr>
<tr>
<td>fɑn ‘betal-leaf’</td>
<td>ʃɔr ‘to wear’</td>
<td>pondit ‘priest’</td>
<td>ʃɔrti ‘enjoyment’</td>
<td>fit-o ‘back Loc’</td>
<td>afeł ‘apple’</td>
<td>pet ‘stomach’</td>
</tr>
</tbody>
</table>

In Table 10 and in Figure 5, we find that generally the following back vowels show maximum effect on spirantization and it is the mid low back vowel [ɔ] which favours spirantization maximally among the IA Bengali speakers. In SBV some vowels show tense-lax distinction, and it is possible that the vowels which are relatively more checked (and being

---

12 The table 9 shows examples of unaspirated/unaspirated plosives and spirants which are extracted form the speaker’s data.
shorter and represent unstressed syllables) favour spirantization. In the case of TB Bengali speakers, we find a similar pattern, that is, the following vowels which are relatively more checked favour spirantization. The following mid low back vowel [ɔ] has a higher rate of spirantization as opposed to high back vowel [u] among TB Bengali speakers. However, the mid vowels [e] and [o] show less effect on spirantization on both IA and TB Bengali speakers.

Table 10: Following vowels and spirantization in IA and TB SBV speakers

<table>
<thead>
<tr>
<th>-following vowels</th>
<th>IA</th>
<th>TB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[f %]</td>
<td>[N]</td>
</tr>
<tr>
<td>Mid low back vowel [ɔ]</td>
<td>91</td>
<td>269/295</td>
</tr>
<tr>
<td>High back vowel [u]</td>
<td>76</td>
<td>94/124</td>
</tr>
<tr>
<td>Low back vowel [a]</td>
<td>66</td>
<td>28/425</td>
</tr>
<tr>
<td>High front vowel [i]</td>
<td>57</td>
<td>25/44</td>
</tr>
<tr>
<td>Mid low front vowel [ɛ]</td>
<td>47</td>
<td>9/19</td>
</tr>
<tr>
<td>Mid high back vowel [o]</td>
<td>28</td>
<td>29/105</td>
</tr>
<tr>
<td>Mid high front vowel [e]</td>
<td>20</td>
<td>3/15</td>
</tr>
</tbody>
</table>

The effect of lexical constraints also turned out to be significant in SBV. SBV lexicon comes essentially from three sources. One, a great majority of its lexical items is of Indo-Aryan origin. Secondly, a small number of lexical items come from English. These are of relatively recent origin and were added only in the recent past. Thirdly, a significant proportion of its lexicon might be of Tibeto-Burman origin though unspecified\(^\text{13}\) in terms of individual languages. These must have entered into the various Bengali dialects over a period of long time due to their coexistence with Tibeto-Burman languages in Bengal Basin, and it is possible that many of these lexical items were brought by the numerous incoming Bengali dialects of the immigrants. Lexical items of Indo-Aryan and Tibeto-Burman origin favour spirantization maximally than the words of English origin as in Table 11 and in Figure 6.

\(^\text{13}\) Unspecified: I could not find Indo-Aryan etymologies for many of the words in Turner’s Dictionary. It does not mean that many of these words are not attested in Indo-Aryan languages. Additionally, what makes the task of etymology hunting more difficult is that the changed phonological shapes of these words. For example, ‘\textit{khuf}’ meaning ‘small space’ may or may not be related to ‘\textit{gupha}’ meaning ‘cave/tunnel’. Other words such as \textit{khofol} ‘papaya’; \textit{fatak} ‘place name’; \textit{tifa} ‘to press’; \textit{futla} ‘a kind of bag’. Many words that are attested in Indo-Aryan might have come from non-IA sources.
Table 11: Lexical constraints on spirantization in IA and TB SBV speakers

<table>
<thead>
<tr>
<th>Clusters</th>
<th>IA [f%]</th>
<th>IB [f%]</th>
<th>IA [N]</th>
<th>IB [N]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indo-Aryan</td>
<td>72</td>
<td>24</td>
<td>334/464</td>
<td>21/86</td>
</tr>
<tr>
<td>Unspecified (TB?)</td>
<td>65</td>
<td>26</td>
<td>374/572</td>
<td>46/175</td>
</tr>
<tr>
<td>English</td>
<td>30</td>
<td>13</td>
<td>50/169</td>
<td>12/90</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>758/1205</td>
<td>79/351</td>
</tr>
</tbody>
</table>

Figure 6: The effect of lexical constraints on spirantization in IA and TB SBV speakers

4. Sociolinguistic constraints

Social factors are necessary to understand the social differentiation of linguistic variants in the community. Social factors are typically far less independent than linguistic factors and tend to show a lot of interaction. Existing studies such as Satyanath (1991, 2001) have demonstrated the usefulness of generation over age groups in similar studies of contact situations.

As mentioned above Silchar is embedded in a Tibeto-Burman speaking region, the overall rate of spirantization of Bengali, Hindustani and the indigenous Tibeto-Burman speakers are shown in Table 12. The Bengali speakers and the Tibeto-Burman speakers have higher rate of spirantization as that of the Hindustani speakers.

Table 12: Linguistic ethnicity and spirantization

<table>
<thead>
<tr>
<th>Language descent</th>
<th>[f%]</th>
<th>[N]</th>
<th>[p%]</th>
<th>[N]</th>
<th>Total [N]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bengali</td>
<td>69</td>
<td>739</td>
<td>31</td>
<td>330</td>
<td>1069</td>
</tr>
<tr>
<td>Hindustani</td>
<td>14</td>
<td>19</td>
<td>86</td>
<td>117</td>
<td>136</td>
</tr>
<tr>
<td>Tibeto-Burman</td>
<td>23</td>
<td>79</td>
<td>78</td>
<td>272</td>
<td>351</td>
</tr>
</tbody>
</table>

The Bengalis are far ahead in the use of spirantization than the Non-Bengali speakers. It is interesting to note that among the non-Bengali speakers it is the Tibeto-Burman speakers who have more of spirantization than the Hindustani speakers as shown in Figure 7. Moreover, even several generations after the two groups (Bengali and the Hindustanis) arrived on Tea plantation in Assam, the two groups do not behave alike with respect to spirantization.
11. Spirantization in Tibeto-Burman-Bengali bilinguals

In my dissertation (2010), it is shown that gender plays an important role in social differentiation of spirantization in SBV speech community. On the whole women use more spirants compared to men (see Table 13 and Figure 8). Interestingly, among Tibeto-Burman speakers it is the males who are ahead in the use of spirants then women.

Table 13: Linguistic ethnicity, Gender and spirantization

<table>
<thead>
<tr>
<th>Gender</th>
<th>Language descent</th>
<th>Bengali</th>
<th>Hindustani</th>
<th>Tibeto-Burman</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[f %]</td>
<td>[N]</td>
<td>[f %]</td>
<td>[N]</td>
</tr>
<tr>
<td>Males</td>
<td>59</td>
<td>249/424</td>
<td>12</td>
<td>12/98</td>
</tr>
<tr>
<td>Females</td>
<td>76</td>
<td>490/645</td>
<td>18</td>
<td>7/38</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>739/1069</td>
<td>19</td>
<td>136</td>
</tr>
</tbody>
</table>

There appears to be a gradual decline in the use of spirantization with time. However, the second and the third generations do not show progressive decline in the use of spirantization. In the fourth generation we found that there is a reversal of patterns, i.e. men (both Bengali and Hindustani) show more of spirantization. In the case of Tibeto-Burman speakers it is the third generation and males who favour more of spirantization (see Table 14 and Table 15).
Table 14: Speakers of Bengali, Hindustani and Tibeto-Burman speakers across generations

<table>
<thead>
<tr>
<th>Language Descent</th>
<th>Immigrant [%]</th>
<th>Married [%]</th>
<th>2nd gen [%]</th>
<th>3rd gen [%]</th>
<th>4th gen [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bengali</td>
<td>97 (31/32)</td>
<td>85 (50/59)</td>
<td>68 (300/438)</td>
<td>73 (274/374)</td>
<td>51 (84/166)</td>
</tr>
<tr>
<td>Hindustani</td>
<td>-</td>
<td>-</td>
<td>10 (8/80)</td>
<td>-</td>
<td>21 (11/56)</td>
</tr>
<tr>
<td>Tibeto-Burman</td>
<td>8 (3/36)</td>
<td>-</td>
<td>7 (1/15)</td>
<td>25 (75/300)</td>
<td>-</td>
</tr>
</tbody>
</table>

Though social class by itself did not turn out to be a significant factor, cross tabulation of gender, generation and class revealed that though in the second generation women use more of spirants than men, there were variability in the use of spirantization in the third generation.

Table 15: Overall spirantization across generation, language descent and social class

<table>
<thead>
<tr>
<th>Generation</th>
<th>Language Descent</th>
<th>Bengali Speakers</th>
<th>Hindustani Speakers</th>
<th>Tibeto-Burman Speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Immigrants</td>
<td>-</td>
<td>Nearly categorical in (o) speaker</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td>2nd Generation</td>
<td>Lower spirantization</td>
<td>High spirantization</td>
<td>Negligible (2.4%)</td>
<td>Less</td>
</tr>
<tr>
<td>3rd Generation</td>
<td>High spirantization among lower class</td>
<td>High spirantization among lower class</td>
<td>No data</td>
<td>No data</td>
</tr>
<tr>
<td>4th Generation</td>
<td>Rise in spirantization</td>
<td>Decline in the use of spirantization</td>
<td>More use of spirants</td>
<td>No data</td>
</tr>
</tbody>
</table>

* (o) stands for the oldest immigrant and (r) stands for the recent immigrants
* UMC stands for Upper Middle Class

In the third generation, the lower class men and women14 have more of spirants than the higher classes as in Table 15.

5. Conclusions

What do we make out of this?

1. Overall, it is the IA speakers who use spirantization more than the TB speakers. However, it appears that contrary to what I thought in Dey (2010), it is the Tibeto-Burman Bengali speakers who are following closely on the spirantization than the Hindustani Bengali speakers. The Hindustani speaking groups are lagging behind the Tibeto-Burman speakers.

2. It is possible that much of the higher rate among the Tibeto-Burman is accounted for Bishnupriya alone. Bishnupriyas are definitely better integrated into the Bengali linguistic culture. But the higher spirantization among the Meithei suggest that this is not so (see Table 16).

14 The social class was constructed on the basis of the socio-economic index and each speaker was coded on a scale for factors such as education, occupation, housing and possessions.
11. Spirantization in Tibeto-Burman-Bengali bilinguals

Table 16: Overall distributions of spirants among the Tibeto-Burman speakers

<table>
<thead>
<tr>
<th>Tibeto-Burman speakers</th>
<th>Gender</th>
<th>Age</th>
<th>Generation</th>
<th>Class</th>
<th>$[f]$ %</th>
<th>Total</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kabui-Naga Female</td>
<td>65</td>
<td>immigrants</td>
<td>LC</td>
<td>-</td>
<td>0/9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kabui-Naga Female</td>
<td>60</td>
<td>immigrants</td>
<td>LMC</td>
<td>11.1%</td>
<td>3/27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kabui-Naga Female</td>
<td>30</td>
<td>2nd Gen</td>
<td>LMC</td>
<td>06.7%</td>
<td>1/15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meithei Male</td>
<td>27</td>
<td>3rd Gen</td>
<td>UMC</td>
<td>22.9%</td>
<td>37/166</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bishnupriya Male</td>
<td>28</td>
<td>3rd Gen</td>
<td>UMC</td>
<td>28.4%</td>
<td>38/134</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22.5%</td>
<td>79/351</td>
<td></td>
</tr>
</tbody>
</table>

Note: Gen= Generation, LC= Lower Class, LMC=Lower Middle Class, UMC= Upper Middle Class

3. Though it appears that spirantization was introduced by the Bengali immigrants, the other groups in Silchar are also participating in the local changes.
4. The rate of spirantization varies across different groups. As we add more data from both in terms of number of speakers and languages, the results would become clear.
5. Most interesting findings are that even though the rate of spirantization may vary across groups, what is more important is that the structural constraints on spirantization are shared. However, it needs to be tested whether there are significant differences across generations and linguistic groups.


References:


11. Spirantization in Tibeto-Burman-Bengali bilinguals


