Vol. 10, Issue II 2025

ISSN: 2426-2149

SEL

Society for Endangered and lesser known Languages

वाक् मंथन

सोसाइटी फॉर इंडेन्जर्ड एंड लेसर नोन लैंग्वेज़ेज की शोध पत्रिका

वाक् मंथन
ISSN: 2426-2149

An Overview of the Consonant Sounds in Poula

Vibeituonuo Mere
vibeituonuomerer@gmail.com

Abstract

Poula belongs to the Tibeto-Burman language of Kuki-Chin sub branch, it is one of the three varieties of languages spoken by the Chakhesang tribe of Nagaland, with the majority of speakers residing in parts of Manipur. This research study is based on the variety of Poula spoken in Phek district of Nagaland. The other two languages spoken by the tribe are Chokri and Khezha. As such, linguistically, this language shares a contiguous border with Manipur. This paper is an attempt to phonologically unlock the consonantal sounds in the language and map their distribution. This study presents 29 consonant phonemes in Poula. The distinctions are made according to the place and manner of articulation. Place of articulation such as; bilabial, labio-dental, alveolar, post-alveolar, retroflex, palatal, velar and glottal. Manner of articulation such as; plosives, fricatives, affricates, nasals, lateral and approximant. Out of the 29 consonant phonemes, eight are plosives, four are nasals, nine fricative sounds, six affricate sounds and two approximants sounds. The present study will also bring into light the phonemic contrast based on the minimal pairs and its distribution in the environment. Interestingly, Poula permits both consonant clusters and consonant sequence and it is both open syllabic and closed syllabic allowing alveolar approximant in the coda position.

Keywords: Poula, Tibeto-Burman, Linguistics, Phonology, Consonants

1. Introduction

The Chakhesangs are a Naga tribe who primarily inhabit the Phek district of Nagaland along with Pochury tribe who consider the district as their tribal headquarters. The name Chakhesang itself reflects three acronyms; *'Cha'* which stands for Chokri, *'Khe'* for Khezha and *'Sang'* for Sangtam. Phek district is made up of three areas and two ranges, of which the present study uses the Razeba range as a point of departure for studying the Poula language. The range is a small one consisting of only three villages and a town who identify ethnically as Poumai. They are Zhavame (Zhamei), Zelome, Tsüpfüme (Chobama), and the town of Razeba with approximately only 6000-10,000 Poula speakers in Nagaland itself. The Poumais are classified as a sub-group of Chakhesang who in turn are further classified as a sub-group of the larger Tenyimia community. Poula language behaves differently from the slightly bigger languages of Chakhesangs- Khezha and Chokri in all aspects, phonologically, morphologically and syntactically.

2. Conceptual Framework

Till date, no linguistics work is found in this variety of Poula. It was only in 2021, a small booklet titled *'Poula Primer Dictionary'* was published by the Poula literature committee which is the first written literature in the language. Besides this, there is absolutely no literature available in this variety of Poula spoken by the Chakhesangs. A few linguistics works are available in the variety spoken in Manipur; Veikho (2014), Veikho & Khyriem (2015) and Veikho & Sarmah (2018), which provides a preliminary phonological description of Poula, Consonants and Vowels. Veikho (2021) builds upon previous works by expanding the findings through the recording of additional data and broadening the analysis.

The Poula Literature Board Committee was formed in the year 2013 with the primary intention to develop its orthography and provide reading materials in the language. One of their main focus lies on the development of Bible and hymnals in their own language. At present, the literature board is working on developing school textbooks for the primary sections.

3. Methodology

This study was based on qualitative research method which includes both primary and secondary sources for data collection and analysis. Primary data were collected from native speakers of the community varying in gender, age and occupation by applying methods such as observation method, structured and unstructured interviews, and collection of community and personal narratives. All the data incorporated in this paper are first-hand information collected by the researcher. Secondary sources were in the form of books, journals, articles, internet sources, official documents and related literature written on the tribe to obtain information on the language and the people.

4. Consonants

Consonants are nearly always movements at the beginning or end of a vowel (Ladefoged, 2005). All sounds during the production of which we hear friction are consonants, but not all consonants are produced with friction (Balasubramanian, 2000) Consonants are classified along three dimensions: voicing, place of articulation and manner of articulation. In the production of consonant sounds, the air from the lungs escapes through the oral passage with friction and the nasal passage is completely blocked.

This study presents 29 consonant phonemes in Poula. The distinctions are made according to the place and manner of articulation. Place of articulation such as; bilabial, labio-dental, alveolar, post-alveolar, palatal, velar and glottal. Manner of articulation such as; plosives, fricatives, affricates, nasals, lateral and approximant. Poula has eight plosives having three-way contrast; bilabial, dental-alveolar and velar, /p, b/ /t, d/ /k/ /p^h, t^h, k^h/. The voiced velar plosive[1] /g/ is absent in Poula thus, there is no corresponding voiced sound to /k/.

[1] The voiced velar plosive occurs only in borrowed words, eg: /gari/ 'vehicle' (source- Hindi)

There are five nasals in Poula, and the sounds are distinguished based on the following places of articulation; bilabial nasal /m/, alveolar nasal /n/, palatal nasal /ɲ/, velar nasal /ŋ/ and aspirated voiceless velar nasal /ŋ̊ʰ/.

Fricatives are also distinguished based on the four places of articulation; labio-dental /f, v/; alveolar /s, z/; post-alveolar /ʃ, ʒ/; glottal /h/ and the voiceless retroflex fricative /ʂ/.

Affricates in Poula are alveolar affricates /ʦ, ʣ/, post-alveolar affricates /ʧ, ʤ/ and the voiceless aspirated counterpart /ʧʰ/.

Approximants in Poula are of different types; voiced alveolar approximant /ɹ/. Voiced alveolar lateral approximant /l/ and voiced bilabial approximant /w/.

Table 1. Poula Consonant Inventory

	Bilabial	Labiodental	Alveolar	Post-alveolar	Retroflex	Palatal	Velar	Glottal
Voiced (+) Unvoiced (-)	-v +v	-v +v	-v +v	-v +v		-v +v	-v +v	-v +v
Plosive	p b		t d				k	
Aspirated plosive	pʰ		tʰ				kʰ	
Fricative		f v	s z	ʃ ʒ	ʂ			h
Affricate Aspirated affricate			ʦ ʣ	ʧ ʤ ʧʰ				
Nasal	m		n			ɲ	ŋ̊ʰ ŋ	
Lateral Approximant			l					
Approximant	w		ɹ					

4.1 Stops/Plosives

In a stop, the airflow through the mouth is momentarily closed off. This can be done by the two lips, forming [p] or [b]; by the tongue tip touching the alveolar ridge, forming [t] or [d]; by the tongue body touching the palate, forming [k] or [g]; and in other ways (Hayes, 2009)

Poula has eight phonemic stops /p/, /pʰ/, /b/, /t/, /tʰ/, /d/, /k/ and /kʰ/. Phoneme aspirations normally occurs with voiceless stops in labial, alveolar and velar places of articulation. The voiceless aspirated stops /pʰ/, /tʰ/ and /kʰ/ occurs in word-initial and word-medial positions. Voiced stops do not have its aspirated counter parts and it also occurs word initially and word medially.

/p/: [p] is a voiceless bilabial plosive which is realized as [p], it occurs in word-initial and word-medial positions.

Initial	**Medial**
/paomodu/ 'fault'	/asadʒəpa/ 'happiness'
/paomusi/ 'criticism'	/mepu/ 'groom'
/pekʰo/ 'sound'	/səpapitsume/ 'publisher'

/pʰ/: [pʰ] is a voiceless aspirated bilabial plosive, it is realized as [pʰ]. It occurs in word-initial and word-medial positions.

Initial	**Medial**
/pʰutsu/ 'pride	/apʰu/ 'spade'
/pʰusi/ 'apple'	/kobipa/ 'cauliflower'
/pʰimoja/ 'sock'	/kʰɪopʰao/ 'graveyard'

/b/: [b] is a voiced bilabial plosive, it is realized as [b]. It occurs in word-initial and word-medial positions.

Initial	**Medial**
/balə/ 'stay'	/risutsubu/ 'field'
/bamotsu/ 'fold (arm)'	/tʰubuvə/ 'grain sack'
/bako/ 'hour'	/trinaba/ 'butterfly'

/t/: [t] is a voiceless alveolar plosive, it is realized as [t]. It occurs in word-initial and word-medial positions.

Initial	**Medial**
/titʰɪə/ 'air'	/natə/ 'little'
/tihə/ 'cloud'	/ŋatitu/ 'mole'
/tizə/ 'dark'	/kʰatu/ 'bread'

/tʰ/: [tʰ] is a voiceless aspirated alveolar plosive, it is realized as [tʰ]. It occurs in word-initial and word-medial positions.

Initial	**Medial**
/tʰaotə/ 'younger sister'	/məlitʰu/ 'heart'
/tʰubu/ 'rice'	/tatʰu/ 'pickle'
/tʰpatsu/ 'print'	/datʰa/ 'butcher'

/d/: [d] is a voiced alveolar plosive, it is realized as [d]. It occurs in word-initial and word-medial positions.

Initial	**Medial**
/daoruki/ 'hospital'	/bedu/ 'cheek'
/duraʃi/ 'grape'	/mode/ 'greens'
/daotsu/ 'cutting'	/ɲiedu/ 'orange'

/k/: [k] is a voiceless velar plosive, it is realized as [k]. It occurs in word-initial and word-medial positions.

Initial	**Medial**
/ki/ 'house'	/zaoki/ 'aeroplane'
/kihupʰetsume/ 'sweeper'	/ʤəpikinitsume/ 'sailor'
/ko/ 'reason'	/ʤəpiki/ 'boat'

4.2 Fricatives

In a fricative, a tight constriction is made, so that air passing through the constriction flows turbulently, making a hissing noise [f], [v] etc. In the production of sibilant fricatives like [s], [z]. [ʃ] and [ʒ], the mechanism of production is more complex: a stream of air is directed at the upper teeth, creating noisy turbulent flow (ibid). Fricative sounds in Poula are relatively rich. There are eight fricative phonemes, out of which five are voiceless fricatives /f/, /s/, /ʃ/, /h/ and /ʂ/; three voiced fricatives /v/, /z/ and /ʒ/. All these fricatives occur word initially and word medially. Contrast of fricative sounds in Poula are given below:

/f/: [f] is a voiceless labiodental fricative, it is realized as [f]. It occurs in word-initial and word-medial positions.

Initial	**Medial**
/fomoʃi/ 'essence'	/səfəro/ 'body'
/fokutsu/ 'to hide'	/səfətsu/ 'breath'
/fəpə/ 'parents'	/batʰɹuafə/ 'bat'

/v/: [v] is a voiceless labiodental fricative, it is realized as [v]. It occurs in word-initial and word-medial positions.

Initial	**Medial**
/vi/ 'bull'	/levə/ 'comb'
/ve/ 'breakfast'	/ʤovə/ 'arm-pit'
/vitsu/ 'spine'	/kʰavu/ 'curry'

/s/: [s] is a voiceless alveolar sibilant fricative, it is realized as [s]. It occurs in word-initial and word-medial positions.

Initial	**Medial**
/sali/ 'frying pan'	/rasəʤə/ 'fruit juice'
/su/ 'pain'	/pisu/ 'headache'
/saləɹe/ 'adoption'	/asame/ 'friend'

/z/: [z] is a voiced alveolar sibilant fricative, it is realized as [z]. It occurs in word-initial and word-medial positions.

Initial	**Medial**
/zəpi/ 'to wrap'	/zezi/ 'surname'
/zi/ 'name'	/mozao/ 'swallow'

/ʃ/: [ʃ] is a voiceless postalveolar sibilant fricative, it is realized as [ʃ]. It occurs in word-initial and word-medial positions.

Initial	**Medial**
/ʃi/ 'bad'	/muʃa/ 'greed'
/ʃo/ 'to produce sweat'	/ɹaʃi/ 'ghost'
/ʃa/ 'pest found in rice'	/səumoʃi/ 'spoil'

/ʒ/: [ʒ] is a voiced postalveolar sibilant fricative, it is realized as [ʒ]. It occurs in word-initial and word-medial positions.

Initial	**Medial**
/ʒi/ 'to sleep'	/baʒo/ 'palm'
/ʒoʧo/ 'similar'	/ʧaʒa/ 'salary'

/ʂ/: [ʂ] is a voiceless retroflex fricative, it is realized a [ʂ]. It occurs in word-initial and word-medial positions.

Initial	**Medial**
/ʂu/ 'perspiration'	/nameʂə/ 'pregnant'
/ʂə/ 'plan'	/meʂə/ 'population'

/h/: [h] is a voiceless glottal fricative, it is realized as [h]. It occurs in word-initial and word-medial positions.

Initial	Medial
/hutsu/ 'to blow'	/monaha/ 'birthday'
/huk^{h}aki/ 'jail'	/ɹuha/ 'reputation'
/hat^{h}ə/ 'to fast'	/ʒoha/ 'busy'

4.3 Affricates

An affricate is a stop followed by a fricative, made at the same location in the mouth in rapid succession so that the result has the typical duration of a single speech sound. (ibid). There are five affricate sounds in Poula. /ʦ/ is a voiceless alveolar affricate; /ʣ/ is a voiced alveolar affricate; /ʧ/ is a post-alveolar affricate; /ʤ/ is a voiced post-alveolar affricate and /ʧ^{h}/ is a voiceless aspirated post-alveolar fricative. /ʧ^{h}/ do not have its voiced aspirated counterpart. All the affricate sounds occur in word-initial and word-medial positions.

/ʦ/: [ʦ] is a voiceless alveolar affricate, it is realized as [ʦ]. It occurs in word-initial and word-medial positions.

Initial	Medial
/ʦa/ 'to tease'	/vutsu/ 'to bury'
/ʦisa/ 'to obey'	/patsu/ 'to bloom'
/ʦamutsi/ 'to harvest'	/zitsu/ 'to divide'

/ʣ/: [ʣ] is a voiced alveolar affricate, it is realized as [ʣ]. It occurs in word-initial and word-medial positions.

Initial	Medial
/ʣəp^{h}i/ 'sorrow'	/muʣə/ 'urinate'
/ʣədəu/ 'stable'	/hiʣə/ 'tear'
/ʣəap^{h}ɪa/ 'sprinkle'	/ʃuʣə/ 'abroad'

/ʧ/: [ʧ] is a voiceless post-alveolar affricate, it is realized as [ʧ]. It occurs in word-initial and word-medial positions.

Initial	Medial
/ʧaʧəu/ 'swear'	/zeʧi/ 'strict'
/ʧifə/ 'rumor'	/tuʧome/ 'runner'

/ʧ^{h}/: [ʧ^{h}] is an aspirated voiceless post-alveolar affricate, it is realized as [ʧ^{h}]. It occurs in word-initial and word-medial positions.

Initial	Medial
/ʧ^{h}i/ 'shiver'	/monaʧ^{h}i/ 'rebirth'
/ʧ^{h}əna/ 'third'	/motsəʧ^{h}i/ 'rethink'

/ʤ/: [ʤ] is a voiced post-alveolar affricate, it is realized as [ʤ]. It occurs in word-initial and word-medial positions**.**

Initial	Medial
/ʤoli/ ‘to tickle’	/buʤə/ ‘trunk of a wood’
/ʤosu/ ‘sake	/ʃiʤo/ ‘tomorrow’

4.4 Nasals

In a nasal consonant, the velum is lowered, allowing air to escape through the nose. Most nasal consonants have a complete blockage within the mouth at the same time. The place of articulation for nasals are mostly the same as those for stops (ibid). Poula has four nasals /m/, /n/, /ŋ/ and /ɲ/. /m/ which is realized as [m] is a voiced bilabial nasal which occurs in word-initial and word-medial positions. /n/ which is realized as [n] is a voiced alveolar nasal which occurs in word-initial and word-medial positions. /ŋ/ which is realized as [ŋ] is a voiced velar nasal which occur in word-initial and word-medial positions. /ɲ/ which is realized as [ɲ] is a voiced palatal nasal which occur in word-medial and word-final positions.

/m/: [m] is a voiced bilabial nasal, it is realized as [m]. It occurs in word-initial and word-medial positions.

Initial	Medial
/mote/ ‘salt’/	/kʰamotsu/ ‘spoon’
/moʤə/ ‘urine’	/same/ ‘blanket’
/moha/ ‘yawn’	/sabamuni/ ‘cloth’

/n/: [n] is a voiced alveolar nasal, it is realized as [n]. It occurs in word-initial and word-medial positions.

Initial	Medial
/ni/ ‘press’	/tsini/ ‘sugar’
/nitsu/ ‘laughter’	/belani/ ‘tomato’
/nituɹə/ ‘smile’	/musuna / ‘public

/ŋ/: [ŋ] is a voiced velar nasal, it is realized as [ŋ]. It occurs in word-initial and word-medial positions.

Initial	Medial
/ŋu/ ‘village’	/paomuŋao/ ‘recommendation’
/ŋumetʰɹo/ ‘villagers’	/remuŋu/ ‘fifty’
/ŋaobame/ ‘enemy’	/kemeŋu/ ‘fifteen’

/ɲ/: [ɲ] is a voiced palatal nasal, it is realized as [ɲ]. It occurs in word-initial and word-medial positions.

Initial	Medial
/ɲali/ ‘earthen-pot’	/haɲi/ ‘rich’
/ɲa/ ‘soil’	/duɲa/ ‘coriander’

4.5 Approximants

Approximants are consonants in which the constriction is fairly wide, so that air passes through without creating turbulence or trilling. In lateral approximants, the air passes around the sides of the tongue (ibid). Poula has three approximants, /w/, and /ɹ/ are the central approximants and /l/ is the liquid approximant. The occurrence of the bilabial approximant /w/ is very limited and it is found in word initial only. The only word with this sound is the response word /we/. /l/ is a voiced alveolar lateral approximant and /ɹ/ is a voiced alveolar approximant both occurring in word-initial and word-medial positions.

/ɹ/: [ɹ] is a voiced alveolar lateral approximant, it is realized as [ɹ]. It occurs in word-initial, medial and final positions. /

Initial	**Medial**	**Final**
/ɹə/ 'stitch'	/ɹaɹa/ 'rash'	/kʰəɹ/ 'sepulchre'
/ɹi/ 'horse'	/bekʰaɹao/ 'jaw'	/məɹ/ 'mouth'

/l/: [l] is a voiced alveolar lateral approximant, it is realized as [l]. It occurs in word-initial and word-medial positions.

Initial	**Medial**
/laɹəvə/ 'book'	/hilatə/ 'eyeball'
/lao/ 'diarrhea'	/məlithoə/ 'heart'

4.6 Phonemic Distribution of Consonants

Table 2.2 illustrates the distribution of the consonant phonemes in the word-initial, medial and final positions.

Table 2.2 Distribution of Consonant Phonemes at the Syllable Level

Consonant	**Initial**	**Medial**	**Final**
/p/	+	+	-
/pʰ/	+	+	-
/b/	+	+	-
/t/	+	+	-
/tʰ/	+	+	-
/d/	+	+	-
/dʰ/	+	-	-
/k	+	+	-
/kʰ/	+	+	-
/f/	+	+	-

/v/	+	+	-
/s/	+	+	-
/z/	+	+	-
/ʃ/	+	+	-
/ʒ/	+	+	-
/ʂ/	+	+	-
/h/	+	+	-
/ʦ/	+	+	-
/ʣ/	+	+	-
/ʧ/	+	+	-
/ʧʰ/	+	+	-
/ʤ/	+	+	-
/m/	+	+	-
/n/	+	+	-
/ŋ/	+	+	-
/ŋ̊ʰ/	+	+	-
/ɲ/	+	+	-
/l/	+	+	-
/w/	+	-	-
/ɹ/	+	+	+

4.7 Phonemic Contrast of Consonants

A contrastive pair is a pair of word that differ only in one phoneme. Such phonemes are environmentally contrastive as their difference lies within the specific environment. By substituting one sound with another sound in the same environment, the semantic meaning of the word completely changes. Given below are some examples of consonant contrastive/ minimal pairs in Poula.

4.7.1 Contrast of Stops/Plosives

a. Bilabial plosives /p, pʰ, b/

/pe/ ‘grandmother’ /pʰe/ ‘to wash (hair)’ /be/ ‘to leave (the work)’

/pu/ 'forehead' /pʰu/ 'to bind' /bu/ 'gun'

b. Alveolar plosives /t, tʰ, d/

/ta/ 'go' /tʰa/ 'fast' /da/ 'to beat an object with a stick'

/tu/ 'reach' /tʰu/ 'thousand /du/ 'big'

c. Velar plosives /k, kʰ/

/ka/ 'broad' /kʰa/ 'a type of leaf'

/ki/ 'house' /kʰi/ 'not straight/ curve'

4.7.2 Contrast of Fricatives

a. Labiodental fricatives /f, v/

/fo/ 'to pick' /vo/ 'to trap or surround'

/fo/ 'wrestling' /vo/ 'pig'

b. Alveolar fricative /s, z/

/su/ 'deer' /zu/ 'to look'

/si/ 'dog' /zi/ 'to sleep'

c. Postalveolar affricates /ʃ, ʒ/

/ʃo/ 'ladder' /ʒo/ 'wages'

/ʃa/ 'to announce' /ʒa/ 'one's share'

d. Glottal fricative /h/ and retroflex fricative /ʂ/

/ha/ 'fresh' /ʂa/ 'to slice (meat)'

/ha/ 'veins' /ʂa/ 'to bless'

/ha/ 'buffllo' /ʂa/ 'blessing'

4.7.3 Contrast of Affricates

a. Alveolar Affricate /ʦ, ʣ/

/ʦə/ 'thought' /ʣə/ 'to cook'

/ʦə/ 'a type of traditional instrument used by female' /ʣə/ 'to cut a bamboo'

/ʦə/ 'support' /ʣə/ 'water'

4.7.4 Contrast of Nasals

Near-minimal segmental pairs for Poula nasals /m/, /n/, /ŋ/ and /ɲ/ are given below:

a. Bilabial nasal /m/ and alveolar nasal /n/

/ma/ 'swell' /na/ 'low'

/ma/ ‘fault’ /na/ ‘child’

/ma/ ‘pumpkin’ /na/ ‘sacrifice (to ghost)’

b. Velar Nasal /ŋ/ vs Palatal Nasal /ɲ/

/ŋa/ ‘mole’ /ɲa/ ‘to apply (powder)’

/ŋe/ ‘thatch’ /ɲa/ ‘to dirty oneself in mud’

/ŋe/ ‘wither (flower)’ /ɲa/ ‘soil’

4.7.5 Contrast of Approximants

The minimal pairs for voiced alveolar lateral approximant /l/ and the voiced alveolar approximant /ɹ/ are illustrated below:

a. /l/ vs /ɹ/

/la/ ‘language’ /ɹa/ ‘a traditional basket placed above the hearth’

/la/ ‘to stand’ /ɹa/ ‘stripes/design’

/la/ ‘to pass’ /ɹa/ ‘dog’s bark’

4.8 Consonant Clusters

By consonant clusters we mean a sequence of two or more consonants at the beginning or end of a syllable. In other words, a sequence of two consonants will have to form part of the same syllable if it has to be considered a consonant cluster (Balasubramanian, 2000). The consonant clusters in Poula are onset cluster and it exclusively occur with stops and approximants. Given below are the formation of onset clusters between stops and the alveolar approximant [ɹ].

4.8.1 Voiceless Bilabial Plosive [p] + Alveolar Approximant [ɹ]

The voiceless bilabial plosive form clusters with the voiced alveolar approximant as shown below:

Initial	**Medial**
/pro.le/ ‘too much’	/ki.pri/ ‘garden’
/pri/ ‘needle’	/ki.pɹao/ ‘tent’
/pɹe/ ‘cup’	/kʰao.pɹi.tao/ ‘terror’

4.8.2 Voiceless Bilabial aspirated plosive [pʰ] + Alveolar Approximant [ɹ]

The voiceless bilabial aspirated plosive form clusters with the voiced alveolar approximant as shown below:

Initial	**Medial**
/pʰɹa/ ‘talk’	/me.pʰɹi.me/ ‘spectator’
/pʰɹa/ ‘generous’	/la.ɹə.pʰɹə.me/ ‘student’

4.8.3 Voiceless Alveolar Plosive [t] + Alveolar Approximant [ɹ]

The voiceless alveolar plosive form clusters with the voiced alveolar approximant as shown below:

Initial	Medial
/tro.to/ ‘maize’	/mu.tɹu/ ‘worship’
/tru.si/ ‘peach’	/pe.tɹo/ ‘talkative’
/tro/ ‘game’	/pʰu.tɹəu/ ‘tight’

4.8.4 Voiceless Alveolar Aspirated Plosive [tʰ] + Alveolar Approximant [ɹ]

The voiceless alveolar aspirated plosive form clusters with the voiced alveolar approximant as shown below:

Initial	Medial
/tʰɹe/ ‘sour’	/itʰɹume/ ‘we’
/tʰɹo.lu/ ‘victory’	/li.tʰɹi/ ‘only’
/tʰɹu/ ‘shed’	/a.tʰɹau/ ‘strength’

Consonant clusters in Poula also occurs in onomatopoeic words. Given below are some examples of sounds produced by animals and inanimate objects.

bɹe~bɹe ‘of cows’

fɹi~fɹi ‘of whistling’

In Poula, consonant clusters are permitted, but with certain restrictions. The language allows initial and medial consonant clusters, with a minimum requirement of two consonants. Vowel-initial clusters, where the clusters are preceded by a vowel, are also observed. However, the number of consonants in a cluster does not exceed two, and the syllable structure is analyzed as CCV, CCVV, CVCCV, VCCV, or VCCVCV.

While clusters involving plosives with the alveolar approximant are common in Poula, clusters with nasals and approximants, as well as fricatives and alveolar approximants, are limited and primarily found in onomatopoeic word constructions. Additionally, Tibeto-Burman consonant clusters are only found in root-initial position (Benedict, 1972).

Table 3 Consonant Clusters in Monosyllabic words

Onset Cluster	Gloss
pɹe (CCV)	‘cup’
pʰɹa (CCV)	‘talk’
tɹo (CCV)	‘game’
tʰɹu (CCV)	‘shed’

4.9 Consonant Sequence

The occurrence of consonant sequences is very rare in Poula and it is found only in a limited number of words. Consonants which occur in sequence across the syllable boundaries and have no intervening vowel are termed as consonant sequence. The following data shows consonant sequence between an alveolar approximant followed by a consonant cluster; an aspirated alveolar plosive and an alveolar approximant (a) and, an alveolar approximant and an alveolar fricative (b).

a) Approximant + Aspirated stop

/kaɪ.tʰɹi/ 'how'

b) Approximant + Fricative

/məɪ.su.ʃi/ 'chilly'

c) Approximant + Fricative

/mə.həɪ.həɪ/ 'brown'

5. Conclusion

Poula exhibits 29 consonant phonemes, the distinctions are made according to the place and manner of articulation. The presence of the voiceless retroflex fricative [ʂ] in Poula, marks an interesting phonetic difference. This sound, which is not found in any of the Angami-Pochuri languages, adds to the unique phonetic inventory of Poula. The voiceless retroflex fricative sound is produced by constricting the airflow with the tongue tip curled back and raised toward the hard palate, creating a turbulent noise. In Poula, this sound may play a distinctive role in the language's phonological system, potentially contrasting with other sounds to differentiate meaning.

Unlike other Angami-Pochuri languages, Poula has an aspirated voiceless velar nasal [ŋ̊ʰ], but no voiceless nasal at any other place of articulation Veikho (2021). Dantsuji (1986); Bhaskara Rao & Ladefoged (1991) stated that many languages in South East Asia have voiceless nasal consonants that contrast with their voiced counterparts. In Angami, three voiceless nasals (bilabial, alveolar, and palatal) have been identified in Khonoma Angami (Bhaskara Rao & Ladefoged, 1991). However, there are no reported occurrences of a voiceless velar nasal in any Angami-Pochuri languages. Therefore, the presence of the voiceless velar nasal in Poula, is an intriguing linguistic feature that distinguishes Poula from its Angami-Pochuri counterparts.

Poula permits both consonant sequence and consonant cluster and the number of consonants does not exceed more than two. This study does not claim to be a complete work on consonants. Further research may be conducted using phonological tools.

References

Balasubramanian, T. (2000). *A textbook of English phonetics for Indian students*. India: Macmillan Publisher. Print.

Benedict K, P. (1972). *Sino-Tibetan: A Conspectus. California*. Berkeley: Cambridge University Press. Print.

Bhaskar Rao, P. & Ladefoged, P. (1991), 'Two types of voiceless nasals', *Journal of the International Phonetic Association* 21(2), 80–88.

Dant Suji, M. (1986), 'Some acoustic observations on the distinction of place of articulation for voiceless nasals in Burmese', *Studia Phonologica* 20(1–11).

Hayes, Bruce. (2009). *Introductory phonology*. Oxford: Blackwell Publisher. Print.

Ladefoged, Peter. (2005), *Vowels and consonants*. Oxford: Blackwell Publishing. Print.

Veikho, S.L. (2021), *Grammar of Poumai Naga (Poula).* Boston: Brill. Print.

Veikho, S.L. & Khyriem, B. (2015), Poula phonetics and phonology; an initial overview, *in* L. Konnerth, S. Morey, P. Sarmah & A. Teo, eds, 'North East Indian Linguistics', Vol. 7, 47–62, Canberra Australian National University: Asia-Pacific Linguistics, Open Access.

Classifiers in Uchoi

Niloy Chakraborty
nilaychk@gmail.com
&
Rajkumari Monimala Sinha
monimalarajkumari@gmail.com

Abstract

The Uchoi language is spoken in the north-eastern state of Tripura, specifically in the South Tripura and Gomati districts. The two primary types of classifiers used in Uchoi's numeral classifier system are sortal and mensural. Morphologically, classifiers in Uchoi follows the constituent order of N [CL-NUM], and these classifiers serve as free morphemes or affixes. Classifiers that differentiate between animate and inanimate nouns are further classified within the sortal category according to their animacy. For sortal classifier *kʰroʔ-* is used by Uchoi for human nouns. When it comes to non-human or inanimate nouns, Uchoi uses sortal classifier *ma-*. Mensural classifiers in Uchoi are measure items that fall into three semantic subtypes: Divisive, Quantitative, and Collective measure terms. Conversely, divisive measure items denote parts or pieces of distinct masses or things. Uchoi makes a distinction between extended (long and flat) and non-extended (round) shapes when it comes to shape classification. In this paper, a thorough yet understandable summary of the classifiers and measure terms in Uchoi is given. The intricate system of sortal and mensural classifiers is examined, providing insight into its nuanced classification and application.

Key Words—Uchoi, Sortal, Mensural, Classifier, Semantic.

1. Introduction

Uchoi is a language, which falls under the Tibeto-Burman language family. According to Thomas H. Lewin's 1869 the nomenclature of the tribe is known as '*Ulsoi*', however, the other alternate names are '*Uchai*' and '*Uchoi*'. The Uchoi are a small ethnic group that primarily inhabits in the southern regions of Tripura. As of the 2011 census, the total population of the tribe is 2447. The Chittagong Hill Tract in the Khagrachori Hill District is home to fewer speakers of the tribe. However, they are always identified as 'Tippera', whereas their neighbouring tribes call them 'Mrung', which is also the 'Mranma' (Arakanese) term for Reang. However, in terms of ethnicity, Uchoi is more similar to 'Bru' than to Tippera; according to Debnath's description (2010, p. 31), their tradition unites Bru and Uchoi as brothers, and they too experienced early Tippera dominance. Even though there are significant discrepancies across clans in the language, these are mostly due to the speaker's accent, and the language is still mutually intelligible. The '*Jawlai*' variety is frequently used in literary contexts[1].
The *Jawlai* variety of the Uchoi language, which is spoken in the districts of Gomati and South Tripura, is the subject of this study. The study intends to offer a thorough examination of the language's structure

[1]There is an active organization, cum NGO '*Uchoi Youth Association*' who have been started working recently for the welfare of the community. In this regard, the first Conference was held in the year 2022. Apart from this particular organisation, there are other organizations also like *Uchoi Welfare Society* (UAWS) &*Tripura Uchoi Youth Association* (TUYA), jointly they have been working for the development of the community as well as for the language also.

and usage, with an emphasis on the numerical system, by looking at one particular variant. This study, in our opinion, will contribute to our understanding of the Uchoi language and its unique features.

2. Methodology and Analysis

2.1 Research Design:

This study looks on the features of classifiers in Uchoi using a linguistic fieldwork methodology. Primary data from twelve native speakers covering a range of genders, ages, and educational backgrounds makes up the corpus for study. To guarantee a thorough representation of the Uchoi-speaking people, the research was carried out in a number of places, particularly in the villages of Depacherra and Ratanpur in the Gomati and South Tripura districts, respectively.

2.2 Data Collection:

During the fieldwork, several methods were used to collect primary data. These consist of:

Word lists—are collections of numerals and associated words that provide a thorough documentation of the numeral system.

Sentence lists—are collections of words and sentences that contain numerals, allowing for a thorough examination of how they are used in context.

Narratives—oral traditions, traditional tales, and spoken narratives to examine the historical and cultural significance of Uchoi numbers.

Surveys and Interviews—comprehensive questionnaires and interviews with language consultants to extract information about their usage and comprehension of numbers.

2.3 Sample Population:

Twelve language consultants from the community participated in the study. In order to provide a more comprehensive picture of the number system usage across various populations, the sample was specifically selected to guarantee variety in terms of gender, age, and educational attainment.

2.4 Data Elicitation:

Audio recorder (Zoom H4N Pro), questionnaires, and interviews were used to elicit data in order to guarantee accurate and trustworthy spoken data records. Numerous data elicitation techniques were used to enable a thorough collection of numeral expressions in diverse communicative contexts.

2.5 Data Analysis:

The collected data was rigorously analysed, using recognized linguistic analysis methods. The following steps were part of the analysis:

Transcription—the original structure and linguistic characteristics of the spoken material were preserved when it was transferred from the audio recordings into a standardized format.

Translation—to make analysis and understanding easier, Uchoi transcriptions were translated into English.

Thematic Coding—number features, usage patterns, and cultural implications were categorized using thematic coding applied to the translated and transcribed data.

3. Classifiers in Uchoi

According to Aikhenvald (2000:17), Uchoi numeral classifier system offers an alternate approach to classifying nouns. These are used as attributive noun phrases (NPs) and are not part of the noun. They can be found in numeral NPs or statements about quantities. Classifiers for numbers can be either independent words or affixes that are commonly attached to numerals or quantifiers. They use the noun's inherent characteristics to describe it. Morphologically classifiers serve as separate lexemes or affixes, in Uchoi. According to Aikhenvald (2000:104), Greenberg (1972) identified four potential constituent orders in the construction of numeral classifiers:

i. [NUM-CL]-N
ii. N[NUM-CL]
iii. [CL-NUM]-N
iv. N[CL-NUM]

Uchoi falls under category no (iv) i.e. N [CL-NUM].

Sortal and Mensural are the two main classificatory categories into which classifiers in Uchoi fall. Sortal classifiers are categories of classifiers that individuate referents according to the type of item they indicate or the way individuals connect to them (Lyons 1977:163), whereas mensural classifiers individuate referents according to quantity (Lyons 1977:163).

3.1 Sortal Classifiers

In Uchoi, sortal classifiers are divided into a number of categories, including animate and inanimate. They are divided into three categories under animate nouns: humans, animals, and groups of humans or animals. Under inanimate nouns they are categorized as plants, shapes and dimensions, consistency, function, specific, etc. Uchoi uses a strong classificatory system in which classifiers are used to group nouns into different classes. Human, non-human, group of humans and non-humans are the subcategories of animate nouns, which are the two categories (animate and inanimate nouns) into which sortal classifiers have been separated. Inanimate nouns are further subdivided into one-dimensional, two-dimensional, and three-dimensional which are covered below—

3.1.1 Animate Nouns [+human +animate]

Humans

The classifier *kʰroʔ-* is used to quantify or categorize individuals in the context of human nouns, as seen in the following examples—

1. *čala* *kʰroʔ-ha*
   man CL:HUM-one

   'One Man'

2. *čala* *kʰroʔ-noi*
   man CL:HUM-two

   'Two Men'

Examples (1) and (2) demonstrate how Uchoi uses classifier to categorize human nouns; while indicating the quantity of "one man," the classifier *kʰroʔ-* is used.

Non-humans

When discussing birds and animals, or non-humans, Uchoi uses a specific classifier *ma-* as shown in the examples (3) to (5):

3. *takʰoŋ* *ma-ha*
   duck CL-one

   'One Duck'

4. *pʰuŋ* *ma-ha*
   goat CL-one

   'One Goat'

5. *siŋyoh* *ma-ha*
   rat CL-one

   'One Rat'

Nouns denoting the living entities especially non-humans as shown in the examples (3) to (5), incorporate with the classifier *ma-*.

Group of humans and non-humans

Both non-human and human groupings are indicated by the classifier *doba*. By indicating the combined presence of elements in a single unit, this classifier adds a crucial organizational feature as shown in the examples no (6), (7) and (8).

6. *gabaŋma* *broh* *doba* *kutʰuŋ* *laiha*
   many people CL:group gather PST

   'Many groups of people gathered'

7. *gabaŋma* *broirao* *doba* *kutʰuŋ* *laiha*
   many women CL:group gather PST

   'Many groups of women gathered'

8. *doba* *kainoi* *soi* *bai* *pʰiyaʔ-ha*
   CL:group two dog it destroy-PST

   'It was destroyed by two groups of dogs'

The word *doba* is used in conjunction with 'broh' (people) and 'broirao' (women) in phrases (6) and (7) to indicate that there are several groups of people who have gathered together. In (8), it illustrates the idea of animal (dog) groupings by using the classifier. It is clear from this formation that multiple groups of dogs were responsible for the destruction. The classifier *doba* is essential in defining the collective

character of groups in these situations, regardless of whether they are made up of human or non-human entities.

3.1.2 Inanimate Nouns (-human, -animate)

Plants, natural items, forms, and so on are examples of inanimate nouns. Following Denny and Creider (1976), Croft (1994:149) highlighted a typical distinction that is important for categorizing inanimate objects, specially forms. He separated shapes into two categories: non-extended, which comprises round (three-dimensional) shapes, and extended, which includes long (one-dimensional) and flat (two-dimensional) shapes.

Dimensionality	Lexeme	Root	Gloss	Classification
1-D	*ɯpʰaŋ*	*-pʰaŋ*	Tree	Tree-like objects
	utʰuŋ	*-tʰuŋ*	Rope	Rope-like objects
2-D	*ɯblai*	*-lai*	Leaf	Leaf-like objects
	kaipʰe	*kai-*	Flat	Flat-like objects
3-D	*utʰai*	*-tʰai*	Fruit	Fruit-like objects
	učloi	*-čloi*	Seed	Seed-like objects

Table 1 shows the Uchoi classifier database, which includes each classifier's lexeme, root, forms, gloss, and semantic properties. One can attest that Uchoi's classifiers are distinct from lexical nouns.

One-dimensional category (long)

A one-dimensional classifier in Uchoi is one that assigns height and length and may have orientation features like vertical and horizontal, as well as consistency features like rigidity or flexibility. Uchoi describes one-dimensional objects that resemble stems or ropes using *pʰaŋ* or *tʰuŋ* classifiers. In Uchoi, the noun *ɯpʰaŋ*, which signifies 'tree', has been reduced to the stem-like classifier *-pʰaŋ*. It is frequently used to characterize hard, long, cylindrical objects that resemble sticks and are upright, like trees, as shown in example (9) below.

9. *ɯpʰaŋ* *pʰaŋ-tʰaŋ*
   tree CL-three

   'Three trees'

But interestingly, for the branches or twigs of a tree or plant Uchoi employs a separate classifier *pra-* with the numerals. For instance—

10. *ndei* *pra-tʰaŋ*
    branch CL-three

    'Three branches'

11. *borpʰaŋ* *ɯpʰaŋ* *ndei* *pra-ba*
banyan tree branch CL-five

'Five branches of banyan tree'

12. *ɯpʰaŋ* *ɯpra* *pra-ha*
tree twigs CL-one

'Five twigs of tree'

In Uchoi, the noun *utʰuŋ*, which means 'rope' has been reduced to the rope-like classifier *-tʰuŋ*. As seen in the examples (13) and (14) below, it is used to categorize objects that resemble ropes and are long, flexible, thin, thread-like, etc.

13. *utʰuŋ* *tʰuŋ-ha*
rope CL-one

'One Rope'

14. *katuŋ* *tʰuŋ-noi*
thread CL-two

'Two threads'

Apart from this *utʰuŋ* which means rope, Uchoi has also an alternative term for the same concept i.e. *dukʰoi*[2]

In the One-dimensional category, Uchoi has a separate classifier *koŋ-* to denote objects like tooth, hands, pen etc. The examples are given below:

15. *buwa* *koŋ-ha*
tooth CL- one

'One tooth'

16. *ŋkruŋ* *koŋ-noi*
hand CL- two

'Two hands'

17. *kolon* *koŋ-broi*
pen CL- four

'Four pens'

[2]But this particular term is not used in the concerned variety of Uchoi. In Uchoi, there are total twelve clans as namely, Paingk'toma, Owauchu, Jawlai, Tuimoi Yaphow, Kaisni, Chongpraing, Paitokra, Meska (the same clan also found in Reang community), Dousoi, Skang, Towma YaphowTowma Yak'cho. For this classifier study we completely rely on the data of Jawlai.

Two-dimensional category (flat)

'Two-dimensional category' describes classifiers that show length and width and can contain either vertical or horizontal objects. Uchoi describes two-dimensional objects using *kai-* and *lai-*. In Uchoi, the noun *ɯblai*, which signifies 'leaf', has been reduced to the two-dimensional classifier *lai-*. As seen below, *lai-* is used to describe flat, two-dimensional objects.

18. *ɯblai* *lai-noi*
    leaf CL-two

    'Two leaves'

19. *mada* *ɯblai* *lai-noi*
    basil leaf CL-two

    'Two basil leaves'

The Uchoi word *kepʰe / kaipʰe* have been reduced to the flat-like 'two dimensional' classifier *kai-*. In Uchoi, *kai-*refers to flat, two-dimensional items, as shown in the following examples (20) to (23):

20. *lama* *kai-ha*
    road CL-one

    'One road'

21. *ɟaŋtʰai* *kai-ha*
    mat CL-one

    'One mat'

22. *ŋkruŋ* *kai-doh pʰen*
    hand CL-six fans

    'Six hand fans'

23. *kʰutai* *kai-tʰaŋ*
    shirt CL- three

    'Three shirts'

Three-dimensional category (round)

The category 'three-dimensional' encompasses classifiers for three dimensions, such as cubes, round or spherical objects, and large, heavy objects. Uchoi describes three-dimensional objects using *ko-*, *tʰai-, le-*, *kʰo-*, *toi-*, *kʰoʔ-*.
The term *boko* which signifies 'circular' has been reduced to the circular-like classifier *ko-*. As seen in the examples (25) and (26), *ko-* is used to describe three-dimensional items that resemble sphere, such as round, oval, tiny objects, and grains. For seed like-structure they also use an alternative word *ɯčloi*, which is again reduced to *čloi* as shown in the example (24):

24. *ɯčloi* *čloi-broi*
    seed CL-four
    'Four seeds'

25. *ɯkoh* *ko-ha*
eye CL-one

'One eye'

26. *bokol* *ko-tʰaŋ*
grain CL-three

'Three grains'

In Uchoi, the noun *ɯtʰai*, which means fruit, has been reduced to the fruit-like classifier *-tʰai*. *tʰai-* is used to describe three-dimensional items that resemble fruits; it is not limited to fruits alone; it can also be used to describe vegetables, as demonstrated in (27) and (28):

27. *tʰaiču* *tʰai-noi*
mango CL-two

'Two mangoes'

28. *tʰailih* *tʰai-ba*
banana CL-five

'Five bananas'

Classifier *le-* is used to denote the coin-like structure. The use of the classifier *le-* is very restricted to only coins or paisa (while collecting data, we are not able to find any suitable counterpart for the word paisa in Uchoi, so we are using the borrowed form, which they use in day-to-day communication) as described in the following example (29):

29. *poiča* *le-ha*
poisa CL-one

'One coin of paisa'

Classifier *-kʰo* that identifies hole-like items has been derived from the noun *ŋəkʰo* as shown in the following examples (26) and (27) respectively:

30. *hakʰo* *kʰo-noi*
earth-hole CL-two

'Two holes in earth'

31. *ɯpʰaŋ* *ŋəkʰo* *kʰo-tʰaŋ*
tree hole CL-three

'Three tree holes'

The classifier *toi-* is used to describe the egg-like oval shape structure, which has been reduced from the noun *nutoi*, as shown in the following example (32):

32. *nutoi* *toi-noi*
egg CL-two
'Two eggs'

In Uchoi, the language has a separate classifier *kʰoʔ-* to classify rupee as illustrated in the following example (33) and (34):

33. *raŋ* *kʰoʔ-ba*
rupee CL-five

'Five rupees'

34. *raŋ* *kʰoʔ-ha*
rupee CL-one

'One rupee'

Except these classifiers, Uchoi uses a different classifier i.e. *ba-* to denote flowers, which also come under the objects of three-dimension, as we can see in the following examples (35) and (36):

35. *kʰu* *ba-ha*
flower CL-one

'One flower'

36. *kʰu* *mali* *ba-ha*
flower jasmine CL-one

'One jasmine flower'

3.2. Mensural Classifiers

According to Lyons (1977:463), Mensural classifiers are those that individuate in terms of amount. They are employed to measure countable nouns and mass units. The quantity or measure of an entity and its physical characteristics are the two elements that influence the choice of a mensural classifier, as stated by Aikhenvald (2000:115). According to Lyons' definition, mensural classifiers in Uchoi are classified as mensural terms, which can be further classified into three subtypes: measure terms that are Quantitative, Collective, and Divisive.

3.2.1. Quantitative measure terms:

Quantitative indicator terms describe numerical standards used to assess persons or groups. Uchoi employs some measuring terms to measure rice, husked rice, water, tea, wine, wood, and other items in Uchoi. Uchoi uses the following items to indicate how individuals or groups are being measured:

37. *kuŋdoh* *mačoŋ-noi*
husk CL-two

'Two fistful amounts of husk'

38. *maituh* *tuh-noi*
rice CL-two

'Two pots of rice'

39. *arouh* *tuh-broi*
rice-beer CL-four

'Four pots of rice beer'

3.2.2. Collective measure items:

Collective measure words, such as group, bunch, pair, etc., indicate a collection or undifferentiated mass of individuals and convey a specific sense of a well-defined distinct object or entity. The collective measure terms in Uchoi are indicated by the following:

kʰo- refers to a mouthful of any solid things like a mouthful of rice, betel-nut etc.

40. *mai* *kʰo-ha*
    rice CL-one

    'One mouthful of rice'

41. *mai* *kʰo-noi*
    rice CL-two

    'Two mouthfuls of rice'

42. *kuwaiŋ* *kʰo-broi*
    betel-nut CL-four

    'Four mouthfuls of betel nut'

Apart from this *kʰo-* classifier, Uchoi uses *pʰau-* classifier to denote bundle of wood or bamboo-like structure. The classifier has been derived from the noun *wapʰau*, which means bamboo. The illustration has been given in the following examples: (43) and (44)

43. *hostuŋ* *pʰau-ha*
    fire-wood CL-one

    'One bundle of fire wood'

44. *wapʰau* *pʰau-noi*
    bamboo CL-two

    'Two bundles of bamboo'

The classifier ɟro- is used to signify a pair of things, in a collective way for instance—a pair of books, pair of shirts etc. as detailed in the following examples:

45. *boi* *ɟro-ha*
    book CL-one

    'One pair of books'

46. *kʰutai* *ɟro-noi*
    shirt CL-two

    'Two pairs of shirts'

3.2.3. Divisive measure items:

In Uchoi, the term "divisive measure" refers to a part or a piece of any specific object, mass of objects, etc. The following is a discussion of several examples of divisive measure terms:

The classifier *du-* usually refers to clod (mud) related terms as shown in examples (47) and (48):

47. *hadu* *du-ha*
    clod CL-one

    'One clod of mud'

48. *hadu* *du-noi*
    clod CL-two

    'Two clods of mud'

The classifier *tʰo-* has been derived from the noun *ntʰoh* which means node. This particular *tʰo-* classifier is being used to refer to nodes of bamboo, sugar cane, wood etc. as discussed in the examples below:

49. *wa* *nətʰoh* *tʰo-ha*
    bamboo node CL-one

    'One node of bamboo'

50. *kru* *nətʰoh* *tʰo-noi*
    sugar cane node CL-two

    'Two nodes of sugar cane'

The same classifier *tʰo-* is also used to signify drops of any liquid, such as drops of milk, drops of oil etc. as illustrated in the following examples:

51. *abutʰoi* *tʰo-noi*
    milk CL-two

    'Two drops of milk'

52. *tʰaou* *tʰo-broi*
    oil CL-four

    'Four drops of oil'

The classifier *pʰoŋ-* is used to denote pieces of flesh, such as pieces of meat, pork, fish etc., which has been discussed in details in the following examples:

53. *ba-haŋ* *pʰoŋ-tʰaŋ*
    chicken-flesh CL-three 'Three pieces of meat'

54. *wa-haŋ* *pʰoŋ-noi*
    pork-flesh CL-two

    'Two pieces of pork'

55. *aha* *pʰoŋ-broi*
    fish CL-four
    'Four pieces of fish'

4. Conclusion

Uchoi illustrates several classifiers[3] systems that divide nouns into distinct groups using numeral classifiers. Both sortal and mensural classifiers are essential linguistic tools for classification and quantification. The sortal classifiers discriminate between animate and inanimate creatures, as well as between various dimensional shapes, by classifying nouns according to animacy and dimensionality. Interestingly, Uchoi uses classifiers such as *kʰro?* to quantify people among humans and *ma* to indicate a non-human entity such as birds, animals etc. As it represents groups of both human and non-human things, the classifier *doba* is significant because it provides an organized way to indicate the presence of a collective entity. The classifiers classify a variety of properties, including plants, forms, and dimensions, within the domain of inanimate nouns. Classifiers like—*pʰaŋ, tʰuŋ, lai, kai, tʰai, čloi* are used to neatly categorize shapes that are one-dimensional, two-dimensional, and three-dimensional. Furthermore, mensural classifiers serve as measure words for divisive, collective, and quantitative measures. From grains and fruits to groups and parts, these measure phrases quantify and characterize a wide range of items, substances, and entities. To conclude, the Uchoi classifier system is a prime example of linguistic creativity, providing a framework for categorizing and measuring a wide range of objects. This system demonstrates the linguistic ability of its speakers to organize and communicate complicated ideas in addition to reflecting the language's ingrained cultural understanding.

References

Allan, Keith. 1977. *Classifiers*. Language 53. 2: 285-311.
Aikhenvald, A. Y. 2000. *Classifiers: A Typological of noun categorization devices*. Oxford: Oxford University Press.
Bradley, D. 1997. Tibeto-Burman languages and classification. In David Bradley (ed.), Papers in Southeast Asian Linguistics No. 14. Tibeto-Burman Languages of the Himalayas. Canberra: Pacific Linguistics.
Comrie, Bernard. 1981. *Language Universals and Linguistic Typology: Syntax and Morphology* Chicago: University of Chicago Press.
Debnath, R. 2010. *Exploring Highlanders of Tripura and Chittagong Hill Tracts*. New Delhi: Akansha Publishing House.
Dryer, M. 2008. *Word Order in Tibeto-Burman Languages: Linguistics of Tibeto-Burman Area* 31 (1)
Greenberg, J. 2013. 10. Numeral Classifiers and Substantive Number: Problems in the Genesis of a Linguistic Type. In F. Keifer and P. Van Sterkenburg (eds.), *Eight Decades of General Linguistics* (pp 305-333). Netherland: Brill.
Higginbotham, J. 1994. *Mass and Count Quantifiers*. Linguistics and Philosophy, 17: 447-480.
Lyons, J. 1977. *Semantics*. Volume I. Cambridge: Cambridge University Press.
Lewin, T. H. 1869. *The Hill Tracts of Chittagong and the Dwellers Therein.* Calcutta: Bengal Publishing Company.

[3]While analysing the data, one of the most common phenomenon found in Uchoi is that majority of the classifiers are derived from nouns, so we can assume that all the classifiers might have come from the nouns, and through the passage of time it has lost or the young generation speakers might have forgotten the root form of the classifiers; though the researchers are not claiming this statement confidently, need to look into more data to comment with such certainty.

Gender in Simte

H. Kapginlian
lianhangluah2017@gmail.com
&
Dr Saralin A. Lyngdoh
saralyngdoh@gmail.com
North Eastern Hill University, Shillong

Abstract

Simte is one of the Kuki-Chin languages spoken mostly in Churachandpur and Pherzawl districts of Manipur. The language falls under the North Kuki-Chin sub-groups of the Kuki-Chin language family (Lewis. et.al, 2013). This paper attempts to describe the gender markers in Simte. The gender markers for [+human] are '*pa*' for masculine and '*nu*' for feminine gender. Whereas the gender markers for [-human, + animal] are '*tal*' for masculine and '*nu*' for feminine. The neuter gender can be identified by their own lexical term such as *'naupaŋ'* for a child. It is observed that the morphemes '*pu*' and '*pi*' have multiple functions. First, '*pu*' serves as a gender marker for both the maternal grandfather and paternal grandfather while '*pi*' functions as a gender marker for both the maternal grandmother and paternal grandmother. Second, they are used as honorific terms showing respect to elders. Third, they are used to categorized male and female genders in authority or persons highly regarded in the society.

Keywords: Simte, Kuki-Chin, North east India, gender markers, neuter gender.

1. Introduction

Simte is one of the Kuki-Chin language family inhabiting in the North-Eastern part of India. They are predominantly concentrated in Churachandpur and Pherzawl districts of Manipur (Hangluah, 2019). According to Singh and Suantak (2011), scores of Simte population are scattered in Sikkim, Arunachal Pradesh, Nagaland, Mizoram, Meghalaya, Tripura and Assam. According to 2011 Census, this tribe constitutes a total population of 6,728.

Language endangerment is evident in this case, as intergenerational transmission is largely restricted to specific social domains, such as familial and religious (e.g., church) settings. Although the language is incorporated into the educational system as a medium of instruction up to Class X under the Board of Secondary Education, Manipur, its visibility remains limited due to the absence of promotion through print media. This lack of broader institutional and media support raises concerns about the language's long-term vitality and sustainability.

Mention may also be made that the tribe is also recognized as one of the Scheduled Tribes by the Government of India in 1956. It is officially recorded in the Government of India Tribe Recognition Order of 1956 vide S.R.O, 24774 Gazette of India, Part II, Section III, No. 316A, New Delhi, 29th October, 1956 (Ching, 2012).

Chelliah (2015) agrees with Haokip (2011, pp.60) argument on multiple languages prevalent in Manipur. She further states that Manipur hosts more than 29 languages, and bilingualism and code-switching are prominent especially within the Kuki-Chin languages. It also results in the convergence of languages with fewer speakers being subsumed by structurally and similar languages which have political and economic prestige. Due to code-switching and code-mixing, Haokip (2011) is also of the opinion that some of the young Gangte speakers are shifting to Thadou and young Simte speakers shifting towards Paite.

2. Methodology

The present study has been carried out by conducting a field work in Pamjal Village and Joutung village of Pherzawl district of Manipur. More than 10 informants, aged between 45-55 were interviewed, and the data were also cross-examined with different age groups of the same community. Journals, books and e-materials were also accessed for the secondary data.

3. Literature Review

According to Corbett (1991:1), the term 'gender' has its origins in the Latin word 'genus' and was passed down through Old French as '*gendre*'. The original interpretation of the term is rendered as 'kind' or 'sort'. The term 'gender' is employed to refer to both a collection of nouns and the entire category. Thus, a particular language encompasses three genders: masculine, feminine, and neuter. Lyons (1968:283) argues that the three genders of Greek and Latin, namely masculine, feminine, and neuter, were the primary noun categories acknowledged in their respective grammars. The classical Indo-European languages classify gender into three categories: masculine, feminine, and neutral.

Grierson (1904) describes the gender markers of some of the Kuki-Chin languages such as Thadou, Paite and Lushai (Mizo). These languages have distinct markers for both the masculine gender and feminine gender. They employ '*pa'* for masculine human gender and '*nu'* for feminine human gender. However, there is distinction in the gender markers for non-human among these Kuki-Chin groups. Thadou, Vaiphei and Gangte exhibit '*cal'* for masculine and *'pi'* for feminine gender. Languages such as Zou, Paite, Simte and Teddim Chin exhibit '*tal'* gender marker for non-human masculine gender while '*pi'* is employed for feminine gender. (Haokip, 2009; Haokip, 2014; Singh, 2009; Cing, 2017). The similar gender distinction is attested by Kuki-Chin languages such as Hmar (Bapuia& Baruah, 1996), Zou, Gangte, Vaiphei, Simte (Haokip, 2009) and Teddim Chin (Cing, 2017). In the existing literature of Simte languages, gender markers were discussed in parts by Naorem Brojen Singh in his unpublished thesis '*Simte Grammar'* in 2009. The gender markers are

also discussed in Connie Champeon book '*Simte Writers' Handbook'* in 2019. The above studies were carried out by etic researchers who highlight the morphemes for gender distinction. The present study is emic research who not only highlights the morphemes for the gender distinction but also attempts to describe the functions of these gender markers.

4. Gender in Simte

Like other Kuki-Chin languages, Simte does not exhibit any grammatical gender but gender is determined based on meaning (cf. Haokip, 2009). In Simte, the gender markers for human beings and animals are different. Essentially, the lexical meaning of '*pa*' means 'father' and '*nu*' means 'mother'. Generally, both these morphemes can be used as an identification of gender thus serve as a gender marker.

4.1 Gender markers for animate [+human]

The gender markers for human beings are '*-pa'* and '*-nu*'. '*-pa'* is the masculine marker while '*-nu'* is the feminine marker. Both of these markers are suffixed to the noun. The gender markers for human beings are illustrated in the following examples.

1(a). puan sɔp ***-pa*** *huŋ tuŋ - ta*

cloth wash M COP arrive PFV

'The washer man has arrived'

1(b). puan sɔp ***-nu*** *hung tuŋ -ta*

cloth wash F COP arrive PFV

'The washer woman has arrived'

2(a) inveŋ ***-pa*** *toʔ bazar ka -hoʔ -uʔ*

neighbour M and bazaar 1SG go PL

'I and my male neighbour go to the market'

2(b). inveŋ ***-nu*** *toʔ bazar ka -hoʔ -uʔ*

neighbour F and bazaar 1SG go PL

'I and my female neighbour go to the market'

4.2 Gender markers for animate [-human]

For gender reference of animals, '*-tal'* is suffixed to the noun to indicate the masculine gender and '*-pi'* is suffixed to the noun to indicate the feminine gender. The argument is exemplified from 3(a)-4(b).

3(a). ui ***-tal*** *-in mi* k^{h}*at a- pet*

dog M ERG person one 3SG bite

'The male dog (stud) bite one person'

3(b). ui ***-pi*** *-in nou a -nei -ta*

dog F ERG baby 3SG own PFV

'The female dog (bitch) has delivered a baby'

4(a). aʔ ***-tal*** k^{h}*uaŋ -in ka-* t^{h}*ou zou -ta*

chicken M crook ADV 1SG awake finish PFV

'I was already awake when the rooster crowed'

4(b). aʔ ***-pi*** *-in baŋ*c^{h}*i ciaʔ tu:i a- tha diai*

chicken F ERG when time egg 3SG lay Q

'When will the hen lay eggs?'

a. Functions of the gender markers *'-pu'* and *'-pi'*

The gender markers-*pu* and *-pi are used in kinship terms such as grandfather and grandmother* respectively. Besides, they occur before the referent nouns. This is illustrated in examples 5(a) to 5(d).

5 (a.) ka ***-pu*** *kum 80 ahi*

1SG grandfather year 80 DECL

'My grandfather is 80 years old'

5 (b) ka ***-pi*** *a- upat vaŋin a- dam thei hi*

1SG grand-mother 3SG old although 3SG well possible COP

'Although my grand-mother is of age, she is still healthy'

5(c) ka ***-pu*** *lian -in a- huŋ panpiʔ*

1SG grand-father lian ERG 3SG COP help

'My grand-father lian helped me'

5(d) ***pi*** *luni a dam -lou*

Grand-mother luni 3SG well NEG

'Grand-mother Luni is sick'

Both these gender markers function as honorific terms signifying respect to the elders. They are suffixed to the referent noun as exemplified in 6(a) and 6 (b).

6(a.) *hausa* ***-pu*** *-in* *kʰosuŋ* *suʔ* *-siaŋ* *ding* *a-* *ʧi*

chief- *M-* *ERG village* *CAUS* *clean* *COP* *3SG say*

'The male chief instructs the villagers to clean the village'.

6(b)*hausa* ***-pu*** *siʔzou in* *hausa* ***-pi*** *-in* *khua* *a-* *enkol*

chief -M *die* *SEQ* *chief - F - ERG village* *3SG* *look after*

'After the male chief died, his wife is taking charge of the village'.

Interestingly, the gender markers '*-pu'* and '*-pi'* are also used to refer to persons who have authority or person with high social status. Structurally, the two gender markers are suffixed to the referent NP. Examples 7(a) and 7 (b) are illustrative.

7(a) *lamkai* ***-pu*** *-in* *nasem -te* *toʔguan* *a-* *pia*

leader *M* *ERG* *worker PL* *job* *3SG* *give*

'The male leader is giving a job to their workers'.

7(b) *lamkai* ***-pi*** *-in* *nupi* *teŋ* *nataŋ* *a-* *kuan* *diŋin* *a* *-ʧial*

leader *F* *ERG* *women* *all* *social work* *3SG* *go* *PURP* *3SG* *invite*

'The woman leader invites all the womenfolk to participate in the social work'

4.4 Neuter or common gender in Simte

Simte also exhibits a neuter or common gender, where the gender is indicated by the presence of the lexical noun. This The neuter gender is exemplified in (8-10)

(8).***naupaŋ*** *a-* *ʧiaʔ* *-na* *diŋ* *lampi* *diʔ* *a* *-ʧil* *a-* *hoiʔ*

child *3SG* *go* *NMZ* *UNR ASP* *way* *right* *3SG* *train* *3SG good*

'It is good to train up the child in the right path'

(9).***gari*** *-in* *kel* *a* *-phu* *-lum*

vehicle ERG *goat* *3SG* *hit* *sleep*

'A vehicle knock down/ kill a goat'.

(10) ***vapal*** *in* *aʔ* *ka* *kʰoi* *-ut*

pigeon *house* *LOC* *1SG* *rear* *want*

'I want to rear pigeons at home'.

There are certain animate (human) beings where the gender markers are not attached to it. But it is innately understood by the native speaker that the lexical noun itself carries the gender specifics. It is explained in the following column.

Table 1: Gender specifics intrinsic in the lexical items

ni	*Father's sister Or, grandfather sister*	*gaŋ*	*Father sister's husband or grand-father sister's husband*
zi	*Wife*	*pasal*	*Husband*
nuŋa:ʔ	*Female youth/ girl*	*taŋval*	*Male youth/ boy*
meigoŋ	*Widow*	*pagoŋ*	*Widower*
pasal	*Male*	*numei*	*Female*

In a rare case, the gender of a person is highlighted twice by the presence of gender specifics intrinsic in a lexical noun and the morphological gender suffixes. In examples 11-12 the gender specific lexical nouns occur along with the gender suffixes. This is because the focus or attention of the discourse is on the referent subject NP. Whereas in examples 13 and 14, the bone of contention is that someone is narrating an incident without necessarily focusing on a particular person or thing. Hence, the absence of the focus provides room for the omission of the gender suffixes.

(11) ***pasal*** *naupaŋ* ***-pa*** *komaʔ na leʔkhathon ka- pia*

male child M to 2SG letter 1SG give

'I give your letter to the boy child'

*(12).***numei** *naupaŋ* ***-nu*** *komaʔ na leʔkhathon ka- pia*

female child F to 2SG letter 1SG give

'I give your letter to the girl child'.

(13).ka- nu dam -lou ʧu ka nitoʔ ka gaŋ -in a huŋveʔ -uʔ

1SG mother well NEG DEF 1SG aunt and 1SG uncle ERG 3SG COP visit PL

'My uncle and aunt came to visit my sick mother'

(14).nuŋa:ʔ -te -n misi __in aʔ_____siŋpi a-___hom_-uʔ

____girl PL ERG dead house LOC tea 3SG serve PL

'The girls were serving tea at the funeral house'

4.5 Word formation processes of gendered nouns in Simte

The word formation processes for gendered nouns in Simte primarily involve affixation, compounding, and occasionally reduplication, though gender marking is not as extensive as in Indo-European languages. Here's a breakdown of the key processes:

4.5.1 Affixation (Prefixation/Suffixation)

Simte uses gender-specific affixes to distinguish male and female referents, particularly in kinship terms, occupations, or natural gender distinctions.

(15) (a) ***pa-*** *t^haŋ*

M thang

'Uncle Thang'

(15) (b) ***nu-*** *lun*

F lun

'Aunt Lun'

(15) (c) annhuan ***-pa***

cook M

'Male Cook'

(15) (d) innp^{h}iat ***-nu***

sweeper F

'Female sweeper'

4.5.2 Compounding

Gendered nouns in Simte are formed by combining a neutral root with a gender-specific morpheme or word. It is illustrated in the following examples.

(16) *naupaŋ* + ***pa***

child M

'Male Child'

(17) houtu + ***nu***

teacher F

'Female teacher

4.5.3 Reduplication

Reduplication in Kuki-Chin languages is more frequent for plurality, intensification, or ideophones rather than gender marking. It is somehow found in Hmar, but not in Simte. In Hmar, the reduplicated form *tu-tu* refers to "small male child!" while *nu-nu* refers to "small female child," or an "affectionate child". Simte does not exhibit such construction to identify the gendered nouns.

5. Conclusion

In Simte, the lexical items are extensively used to accentuate the distinction of the genders. The lexical items such as '*pa'* and '*nu'* have multiple functions. It serves as a gender marker for both the masculine and feminine human class of gender. It is also observed that the lexical items such as '*pu'* and '*pi'* are employed to address the person of high authority or used as honorific titles. These gender markers are suffixed to the referent NP. In matter concerning the common or neuter gender, the referent lexical noun carries the gender specifics. To emphasize the referent NP, the gender specific lexical nouns are followed by the morphemes '*pa'* and '*nu'*.

Abbreviations

1SG	First Person Singular
2SG	Second Person Singular
3SG	Third Person Singular
ADV	Adverb(ial)
COP	Copula
DECL	Declarative
DEF	Definite marker
ERG	Ergative
F	Feminine
LOC	Locative
M	Masculine
NEG	Negation
NMZ	Nominalizer
PFV	Perfective
PL	Plural
PURP	Purposive Marker
Q	Question Marker
SEQ	Sequential
UNR.	ASP Unrealized Aspect

References

Champeon, C. (2019). *Simte Writers Handbook.* U.S.A: Bibles International.

Chelliah, S. L. (2015). Is Manipur a linguistic Area*? Journal of South Asian Language and Linguistics* 2(1), 87-109

Ching, G. (2012) *History and Culture of Simte*. [Ph.D. Thesis, Department of History]. Imphal: Manipur University.

Cing, Z.N. (2017). *A descriptive grammar of Teddim Chin.* [PhD. Thesis, Department of Linguistics]. Shillong: North Eastern Hill University.

Corbett, G. G. (1991). *Gender.* Cambridge: Cambridge University Press.

Grierson, G.A. (1904). *Linguistic Survey of India. Vol.3, Part III.* Delhi: Motilal Banarasi das.

Hangluah, T.K.L. (2019). The Simte. In STC (Ed.), *60 Diamond Jubilee Simte Tribe Council Souvenir* (pp.16-17). Churachandpur: STC Publication.

Haokip, P. (2009). Noun Morphology in Kuki-Chin languages*. Language in India*.ISSN 1930-2940. Vol. 9: 11 November 2009, 109-129.

Haokip, P. (2011). The Language of Manipur: A Case Study of The Kuki-Chin Languages. *Linguistics of the Tibeto-Burman Area.* Volume 34.1 — April 2011, 85–118.

Haokip, M. K. (2014). *Grammar of Thadou-Kuki: A Descriptive Study*. [Ph. D Unpublished thesis, Centre for Linguistics] New Delhi: Jawaharlal Nehru University.

Lewis, M.P, Simons, G.F. & Fennig, C.D. (eds.). (2013). *Ethnologue: Languages of*

the World, Nineteenth edition. Dallas, Texas: SIL International. Online version:

http://www.ethnologue.com

Lyons, J. (1968). *Introduction to theoretical linguistics.* Cambridge: Cambridge University Press.

Singh, N. B. (2009). *Simte Grammar*. [Ph.D. Thesis, Department of Linguistics]. Imphal: Manipur University.

Singh, S.K., & Suantak, K.S.K. (2011). Epilogue. In Singh, S.K &Suantak, K.S.K (Eds.), *Linguistics Ecology: Manipur,* (pp.285-292). Guwahati: EBH Publishers (India).

Nasal coda loss in Northern Naga: Revising W. T. French's *–Vŋ rimes[1]

Kellen Parker van Dam
kellenparker@gmail.com
Chair for Multilingual Computational Linguistics, Universität Passau, Germany
Department of Languages & Linguistics, La Trobe University, Australia

Abstract
The earliest reconstruction of Patkaian (sometimes known in the literature "Northern Naga") was produced by W. T. French in 1983. While this was a monumental work and of major significance to the field of Tibeto-Burman linguistics, much of the analyses were based on limited data available to the author at the time. As a result, a number of the reconstructed rimes do not hold up today in light of a considerably greater amount of dialectal data to pull from. This paper investigates two recently reconstructed rimes in the language group, previously labeled *#–aŋ*^{I} and *#–aŋ*^{II} in **van Dam, Kellen Parker** (2019). Additional rimes for which French (1983) reconstructed nasal codas are also investigated. In addressing these rimes, an explanation is given for some notable exceptions. In particular the concept HORSE, which defies any clear placement in a single rime across the modern languages. Borrowing from a close relative is proposed as the source of this discrepancy, and phonetic motivations for the modern-day reflexes are investigated.

Keywords Patkaian, historical linguistics, reconstruction, phonology.

1. Introduction

Two distinct *#–aŋ* rimes are reconstructable in Tangsa-Nocte (Glottolog: kony1249), the most internally diverse branch of Patkaian, formerly known as Northern Naga, (**van Dam, Kellen Parker**. in press). These times previously been referred to as *#–aŋ*^{I} and *#–aŋ*^{II} in **van Dam, Kellen Parker** (2019). In the majority of language varieties within this subgroup, reflexes of the two rimes developed distinctly from each other, with one retaining the velar coda and resulting in a shape similar to /-aŋ/ and the other undergoing coda loss, typically resulting in an open CV syllable with either an /a/ or /o/.

Complicating reconstruction of the phonological shape of the rime in the proto language is an inconsistency regarding which of the two rimes occur today from one language to another. Specifically, in some variety's reflexes of *#–aŋ*^{I} retain the coda while the other does not, and in other varieties the opposite is seen, where *#–aŋ*^{II} is the one to regain the coda. In still other varieties, the rimes are indistinguishable due to having merged entirely, with both pronounced /-aŋ / today. While there are no attested cases in which a variety will have retained the nasal coda on *#–aŋ*^{II} but not *#–aŋ*^{I}, the exact motivations for the sound change resulting in the loss of nasal coda on *#–aŋ*^{II} requires explanation. What's more, despite the overall regularity of reflexes, a number of etyma create some complication in the reconstruction, most notably that for HORSE *#kVm.raŋ*, suggesting a close borrowing rather than common inheritance.

Previously, this had been assumed to be a case of two distinct rimes rather than an incomplete but otherwise regular sound change for a number of reasons. First, should it be the case that there was a single **–aŋ* rime, phonological motivations for the split would need to be accounted for, but no conditioning factors have been identified. For the most part, reflexes of the two rimes are consistently distinct across Northern Naga, with only a minority of varieties merging them. An explanation for the way in which the two sides of the split would

[1] Parts of this paper were originally presented at the 52nd International Conference on Sino-Tibetan Languages and Linguistics at the University of Sydney, Australia.

be required which could explain the direction of sound change as well as why it did not occur consistently across varieties.

The issue of reconstructing *#–aŋ* rimes in Patkaian was first pointed out in **van Dam, Kellen Parker** (2019) but without a suggestion of the possible mechanisms resulting in the current state of the two rimes. Resolving this uncertainty has implications on further historical, typological and phylogenetic research in the region and particularly with members of the purported Sal language group within Tibeto-Burman, as it may help better establish the place of the language group within the Tibeto-Burman genealogy. There is significance as well when taking a more zoomed-in approach to the languages including basic documentation of what is an otherwise largely under-documented group of languages, as the etyma which fall under these two rimes are incredibly common within the basic vocabulary. Such basic concepts covered by these rimes include SKY, BONE, RED, YEAR and ALIVE, as well as the corresponding fourth-tone *#–ak*^{I} rime for concepts such as BLACK and BITE.[2] By resolving historical changes, low-level classification and cognate identification can be more accurately done in the future.

While a reconstruction has previously been published for Northern Naga, being that of French (1983), the provided rime reconstructions have a degree of inconsistency in terms of the modern reflexes which is all the more apparent today as a much greater amount of linguistic data has become available. To give an example of some of the mismatches, ELEPHANT, YEAR and RAPTOR are reconstructed for French's proto–Northern Naga (hereafter PNN) as **Ç-gla:ŋ* **pa:ŋ* and **la:ŋ* respectively, all reconstructed with long vowels and all falling under the *#–aŋ*^{II} rime. LIGHTWEIGHT *#kʲaŋ*^{II} is reconstructed by French as **gyaŋ* with short <a>, while many other frequent *#–aŋ*^{II} stems are not found in French. This is the case across the family, in large part due to the quality of data available to W. T. French at the start of the 1980s.

An additional issue relating to data availability is not just that there were fewer language varieties which had reliable data sets was the fact that much of the data available to French simply did not cover as large a portion of the lexicon of these languages as is available today. Thus, a number of stems which would be informative for a reconstruction in 1983 were simply unattested. For example, INTOXICATED *#paŋ*^{II} and SIT *#ŋaŋ*^{I} are both absent in French, as they were in Needham (1897).

As with *#–aŋ*^{II} rimes, corresponding *#–aŋ*^{I} stems are also reconstructed for PNN again with the resulting proto-language proving less reliable in demonstrating regular sound changes than may have actually been the case. Such examples include alive **Ç-raŋ*, CORPSE **maŋ* and DREAM **maŋ* compared to MACHETE **gləŋ* and SKY **rəŋ*. A pattern already begins to emerge in these cases, being a three-way split combining vowel quality and vowel length, but one which does not consistently map to modern data where more detailed phonemic analyses do not suggest phonemic vowel length.

Additionally, a handful of stems do not easily reconstruct for one rime or the other. Most significant of these is HORSE *#kVm.raŋ*. In some varieties the second syllable patterns as though a reflex of *#–aŋ*^{I}, and in others of *#–aŋ*^{II}. However, despite some significant examples, the majority of stems do consistently pattern with one rime or the other across modern varieties.

With all of this in mind, this paper presents a resolution for these issues, providing an explanation for the paths of development of the two rimes for the language group as well as an explanation for the irregularity of HORSE in modern Patkaian varieties. This is done in part through comparison to sound change patterns elsewhere in Northern Naga, as well as through comparison to a very similar sound change that has been previously described for the Sūzhōu dialect of Wú. By considering the different reflexes in Northern Naga in relation to attested sound changes in Northern Wú where the historical phonology is much better established, a more accurate transcription of the two rimes can be determined. This results in the reconstructions of **–ɐŋ* for what

[2] CV/CVN syllables will occur with tones 1-3, and typically have a corresponding 4th tone form with a homo-organic stop coda not unlike the A-D tone system of Tai or the four tones of Middle Chinese (van Dam 2018).

was previously described as *#–aŋ*^{I}, and **–aŋ* for *#–aŋ*^{II}, bringing the pair more in line with two of French's many reconstructions – **–əŋ* and **–aŋ* respectively – while correcting inconsistencies due to the lower quality of data available at that time. Data are presented in support of the conclusion from a range of Northern Naga varieties including those belonging to Tangsa-Nocte, Wancho, Konyak, Phom, Chang, Lainong, Makyam, Chang, Pounyiu and Khiamniungan.

While French (1983) provides a much wider range of possible reconstructed rimes for the series of stems which fall under *#–aŋ*^{I} and *#–aŋ*^{II}, this is also the case for other rimes ending with velar codas, as well as a number of non-velar codas are also reconstructed for proto-Northern Naga (PNN) which are less easily justified in light of greater data availability, especially when compared to cognates in the wider Tibeto-Burman area.

This is especially the case for stems for which a final liquid coda is reconstructed, such as COUNT which French reconstructs as **ʔ-we*, breaking from Matisoff's PTB **wyal* and which I reconstruct for Northern Naga as having final **–al* which has undergone a shift to /j/ in many varieties resulting in the <e> vowel posited by French, but which is also widely retained in many varieties not known to French in the early 1980s. In other cases, what I have reconstructed as final **–l* is reconstructed by French as **–r* or **–n* for many stems which have PTB reconstructions with **–l*. Again this is understandable given the data available at the time. One reason for this is that French was operating with a much smaller data set, covering a few at-the-time well-documented language varieties, but also he was limited to varieties which did not so easily retain **–l*. Specifically, French relied on data from Joglei (Yogli), and Ha'wa Nocte (Namsangia) for Tangsa-Nocte varieties, along with Muishaung (Moshang) which we will come back to in a moment, as well as a single variety of Wancho, Konyak, Phom and Chang each. However, a number of varieties – notably Ngaimong, Muklom, *and Muishaung* – prove quite conservative in retaining final **–l* codas. But with the Muishaung data being largely tied to that of Needham (1897) which featured data from a speaker for whom coda /l/ was less stable **van Dam, Kellen Parker (**Forthcoming 2025) and instead often given as /r/, French would have missed much of the evidence for a retained /l/ coda in Tangsa-Nocte. Additionally, the doculect of Muishaung was of a time when the now nearly-complete process of diphthongisation was only part way complete, resulting in greater variation of **–Vŋ* rime pronunciation than was the case either 100 years earlier or 100 years later.

As for the other varieties mentioned above, French had no access to these data, and so the liquid coda not showing up as prominently in his reconstructions is understandable. This is the case as well for sun, which French reconstructs as **cər* but which I reconstruct with the **–al* rime, corresponding to the same *#sal* stem after which Burling (1983) named the language family. It is less feasible for the higher-order **sal* reconstruction for sun to be correct — which I believe to be the case — and then Northern Naga to have undergone a shift to final **–r*, only for coda /l/ to still be widely attested in the group. The potential outsize influence of Needham's 1897 transcription of a language at the midpoint of a major sound change on French's reconstruction will be discussed further below.

Across the Sal languages, Northern Naga is the most internally diverse today, both in terms of number of distinct varieties and how much they differ from one to another. Tangsa-Nocte has somewhere around 80 different varieties (Morey, 2017) with a number of internal subgroups differing in terms of phonology, morphology, lexicon, and syntax. Muishaung and Joglei, the varieties included in French's reconstruction, are very closely related and both part of the Rangpang/Pangwa group. They share an incredibly high degree of lexical similarity, with only a handful of significant phonological differences. Not included in French's data were highly divergent varieties such as those in the Champhang group, Kon-Nyinshau, Chuyo-Gaqkat, Aasen-Yaqsawa, other Noctean varieties such as Hakhun, Muklom or Phontai (Phong), all three of which now have well documented descriptive grammars published (as Boro 2017, Mulder 2020, and Dutta 2023 respectively). In addition to the wide range of Tangsa-Nocte varieties which were inaccessible to French, other Northern Naga varieties such as Tanhai Konnyak (Jacques and M. T. Konyak 2010), Chen (H. Konyak and Mulder 2022), Karyaw, Khiamniungan, Lainong, Makyam, Lao, Nahen, Makyamic, and Gongvangpounyiu were also unaddressed.

French's proposed correspondences were well motivated and often well described given the data available at the time. However, with access to a much larger set of data for a more representative sample of Patkaian languages, a clearer picture can be painted regarding what correspondences exist, and which cases are likely exceptions to these patterns due to sound changes which have not diffused across the lexicon. Furthermore, with the data currently available for those varieties which French did analyse, gaps in the lexicon have been able to be filled in, allowing greater confidence in reconstructing rimes with more consistency.

An additional difficulty faced by French was a likelihood of inconsistency in transcription of the data. Sometimes this inconsistency was due to ongoing sound changes, in many cases regarding velar nasal codas, discussed in greater detail below.

In settling these uncertainties, this paper contributes to ongoing efforts at reconstruction and typology in North East India and Myanmar, while also providing a first look at a new reconstruction of Northern Naga (Patkaian), revising that done published in French (1983).

2. Conventions used in this paper

In order to distinguish the revised reconstruction from proto-Northern Naga as reconstructed by French (1983), the branch will hereafter be referred to as Patkaian, and the new reconstruction as proto-Patkaian (PPk). Further discussion on the name can be found in Parker, K. (in press), with the main motivation being that there is no link between Northern Naga and the other so-labelled 'Naga' languages spoken further south which would otherwise separate Northern Naga from the rest of Sal. Post and Burling (2017:225) offer the alternative label "Northern Sal" for the same reasons of recognising a break between it and the Naga languages.

Unless marked otherwise, reconstructions that follow are my own. Those of French will be marked with PNN and – where deemed beneficial to the reader – the revised reconstructions with PPk . Tone information is included when useful for illustrating different otherwise homophonous rimes, but has been left out of tables in order to have more room for additional segmental word forms. When present, subscript numerals indicate tone categories, while superscript numerals indicate pitch contours following the five-level system established in Chao (1930).Reconstructed etyma given with an asterisk are ones for which the correspondences are clear. The reconstruction should be taken at least in part as symbolic, and not necessarily a proposed recreation of the pronunciation of the proto-language. Forms which are not fully reconstructed but point to likely cognate classes are marked with a hash (#), as is the case with *#–aŋ*^{I} and *#–aŋ*^{II}, which have been transcribed as such elsewhere, but which should be more realistically transcribed **–ɐŋ* and **–aŋ* respectively. Concepts are given in small caps and glosses in single quotes.

3. Data sources

In addition to recent contributions to the literature, considerable as-yet unpublished data has been collected across Northern Naga varieties. Much of this has been based upon the *Culturally Appropriate Lexicostatistical Model for Southeast Asia* word list, containing 235 concepts (Matisoff 1978) as well as another list of 1,467 concepts of undetermined provenance. Much of this data was collected by the author or others operating in the region around the same time, and remains largely unpublished. A few historical sources are also to be found. One of the earliest – and perhaps most significant in terms of illustrating sound changes – is the Muishaung (Moshang) word list collected by Needham (1897) already mentioned above, much of which was reprinted in Grierson 1930). Much of French's reconstruction thus indirectly relies on data from Needham, as interpreted by Grierson and later Shafer (1953). These are also the primary sources of Tangsa-Nocte data for Marrison (1967), and it is not until later when researchers such as Dutta (1969) and Walker (1948) elicited newer data. In many cases data were provided by speakers directly.

Due to the lack of additional documentation happening for much of the early 20th century, Needham remains one of the most valuable sources on Tangsa-Nocte as spoken at the time both for the extensive account of

Muishaung as well as additional data for Shecyü in the text. By comparing between sources, we are able to see a number of rimes have remained quite stable throughout the past century and a half. Table 1 shows the *–*an* rime across three time periods. Many of the early word lists and descriptions were sparse, resulting in many holes in the table. However, it is clear even with these gaps that there has been little if any change in *–*an* across recent generations.

Table 1: *–*an* in Muishaung Tangsa

gloss	Needham	Marrison	Das Gupta	modern
IRON	yân	yan	–	ʒan$_2$
PLATE		–	–	pan$_2$
COW	mân	–	man	man$_2$
BRACELET	sân	san	–	san$_2$
BREAD	–	–	–	βan$_2$

4. Possible errors in French's sources: Examples from Needham (1897)

Despite the value of Needham's contribution, there are still many inconsistencies. Some of those are themselves valuable as they provide insight into changes that must have been going on at that time. Others, however, are likely to be errors in transcription. Muishaung consultants have on multiple occasions told me that this is the case, believing an Assamese interpreter to be the one to blame. From Table 1 Needham's <ân> may seem to reliably reflect *–*an*, but many exceptions are also to be found. Table 2 shows a number of such cases, with <ân> clearly corresponding to *#–aŋ*^{II}*. This is significant as *#–aŋ*^{II} lost the nasal coda and went through a stage of nasalisation on the vowel before this too was lost, and Needham used underlined <n̲> as the typographic convention for nasal vowels.

While modern Muishaung shows much greater uniformity than what is described in Needham's account, this apparent inconsistency is also of value, however. According to Needham's own descriptions of the orthography employed in his text, <â> likely corresponds to /ɑ/ and <o> to /ɒ/. This /ɑ/ value for <â> provides the first clue for the likely phonetic values of the vowels in our competing *#–aŋ* forms, discussed in more detail in Section 6 below. Note the similarity of transcriptions from Marrison (1967) with those of Needham, the underline being lost entirely by the time of Marrison.

Table 2: Shift from **aŋ*^{II} to /ɔ/ in Muishaung

gloss	Needham 1897	Marrison 1967	Das Gupta 1980	modern
ELEPHANT	jân	jan	co:	tɕɔ$_1$
RED	shân	shan	–	ʃɔ$_2$
LIGHT WEIGHT	chân	chan	–	tɕɔ$_1$
LOOK	lân	lan	–	lɔ$_3$
BONE	rân̲	–	–	rɔ$_1$
BAMBOO	wân̲	wan	–	βɔ$_2$
OLD. MAN	–	likang	–	li$_2$kɔ$_2$
OLD. MAN	vî-kân̲ wâ	vikanwa	–	–
BEAR	cha'bâ	chaba	–	tɕəp$_4$bɔ$_2$

It's also entirely possible that Needham would have interpreted nasalisation as segmental in some cases. The data provided in the original 1897 publication shows considerable variety in a small number of pages.

Another significant point is that of nasal/non-nasal alternation. Tangsa-Nocte varieties have a four-way tonal distinction, likely derived from phonemic phonation distinctions. Three of the four tones are found on open syllables, including those with nasal or liquid codas, and the fourth tone is found only on closed syllables (van Dam 2018). This pattern is similar to that of the traditional four-tone system of Middle Chinese, with *ping, shang* and *qu* versus checked *ru*, or Tai *A B* and *C* versus checked *D* tone.

These 'rime families' are common throughout Tangsa-Nocte. Most often these will involve an alternation between a nasal coda on one hand and a non-nasal stop coda with the same place of articulation. One example of such an alternation would be between MACHETE **ʒˡaŋ*^{I} and HAND **ʒˡak*^{I}. With incredibly few exceptions, rime families undergo the same sound changes within a given variety. Thus, while those two concepts differ on the segmental level, we may also say they differ simply by tone. While perhaps a step too far for most readers, a single **ʒˡaƗ* form could be proposed. We see this same pattern across the lexicon when looking at verbal stem suppletion correspondences across dialects (Morey 2017:364f), but also in splits which have occurred between two tonal variants for a given lexeme. As an example, the verb HEAR/LISTEN corresponding to French's PNN **Ç-ta:t* has two possible etyma in Northern Naga varieties, namely **tat* and **talˀ*. In most Rangpang Tangsa varieties it is the latter, while among Heimi varieties it tends to be the former. Alternations also occur for other phonemes, notably liquids /l/ and /r/ where a fourth tone equivalent coda becomes /t/. Fully open CV syllables shift to a glottal stop, and final bilabials likewise retain place of articulation when undergoing denasalisation.

Thus, one finds *#–aŋ*^{I} finals corresponding to *#ak*^{I} final words such as BLACK *#njak* and BITE **kak*, among others. No non-nasal equivalents of *#–aŋ*^{II} are to be found. All vocalic changes which occurred on *#–aŋ*^{I} also occurred in the same way on *#–ak*^{I} finals.

Despite potential concerns with some of the earliest data, Needham (1897) provides considerable assistance in terms of dating sound changes. This will be discussed in Section 6 below. First, further explanation of the confusion regarding Patkaian *#–aŋ* rimes is required.

5. Competing *#–aŋ* rimes in Patkaian

Table 3 shows examples of *#–aŋ*^{I} and *#–aŋ*^{II} rimes from none Tangsa-Nocte varieties and Wancho. In order, these are Cholim, Muklom, Haidley, Joglei, Khalak, Chamchang (Kimsing), Lochhang, Lungkhi, Muishaung, Tutsa, Wancho and Phom. PNN indicates French's Proto-Northern Naga reconstruction, with the exception of sit **ŋaŋ*, which French did not reconstruct. Instead, the proto-Tibeto-Burman (PTB) form is given (Matisoff 2003). PPk refers to the revised reconstructions, but corresponding to the same proto-language as referred to in French as PNN.

Dream, sky and sit are in the *#–aŋ*^{I} rime while lightweight, raptor and bone have *#–aŋ*^{II}, as does elephant, although in Gaqlun and some other varieties elephant is showing a reflex of *#–aŋ*^{I}.

Table 3: Examples of **–aŋ*^{I} & **–aŋ*^{II} rimes

	PNN	PPk	Ch	Ml	Hl	Jg	Kl	Km	Lo	Lk	Ms	Tu	UW	Pm
DREAM	**maŋ*	*#maŋ*^{I}	maŋ	mʌŋ	mo	məŋ	maŋ	ma	ma	məŋ	mɐuŋ	maŋ	məŋ	maŋ
SKY	**C-raŋ*	*#graŋ*^{I}	raŋ	ɹʌŋ	ro	rəŋ	raŋ	ra	ra	rəŋ	rɐuŋ	raŋ	gəŋ	ɣaŋ
SIT	**ŋaŋ*	*#naŋ*^{I}	ŋaŋ	-	ŋəŋ	ŋaŋ	ŋa	ŋa	ŋəŋ	ŋɐuŋ3	-	-	-	
LIGHT	**gyaŋ*	*kˈaŋ*^{II}	tɕa	ʧaŋ	tʃa	tɕʰaŋ	tʃa	ʒɯ	cʰəu	tɕʰa	tɕɔ	ʧaŋ	-	ʃaŋ
HAWK	**la:ŋ*	*laŋ*^{II}	la	laŋ	la	laŋ	la	lɤ	ləu	la	lɔ	laŋ	lã	laŋ
BONE	**ra:ŋ*	*graŋ*^{II}	ra	ɹaŋ	ra	raŋ	ra	rɤ	rəu	ra	rɔ	raŋ	gã	ɣaŋ
ELEPH.	—	*glaŋ*^{II}	tɕʰa	-	tɕʰaŋ	tʃʰa	tɕɤ	cʰəu	tɕʰa	tɕɔ	-	-	-	

6. Resolving the uncertainty

As mentioned in Section 4, Needham's description of his orthographic conventions gives us an idea of where to begin in explaining the split and apparent flip-flop, wherein some varieties have *#–aŋ^{I}* as closer to /aŋ/, and others instead *#–aŋ^{II}* as more similar to /aŋ/ today. As Needham states, <â> is long as in 'father', and <a> short as 'comp*a*ny' or the u in 'but'. Thus, we can assume <â> to be close to /ɑ/, and <a> as something more mid-central.

Additionally, as in Table 3, a number of languages show a clear difference in vowel quality, where the *#–aŋ^{I}* form is higher as is the case with Muklom /ʌ/ vs /a/, Joglei and Wancho /ə/ vs /a/, and Muishaung /ɐu/ vs /ɔ/. If a difference of vowel quality is what separates *#–aŋ^{I}* and *#–aŋ^{II}*, then an explanation can be given with evidence from a very distant relative in the Sino-Tibetan family: The variety of Wú Chinese spoken in Sūzhōu in eastern China.

As the historical prestige variety within Northern Wú, Sūzhōu dialect has retained a number of older features lost elsewhere in the Yangzte River Delta, and those changes which have occurred have been the subject of much attention over many years. One such change is the loss of velar coda **ŋ*, but only on certain lexemes.

In the case of Sūzhōu dialect, the loss of **ŋ* only occurred when preceded by a low vowel, while in cases where the preceding vowel was not low, it was instead fronted and nasalised upon coda deletion (Zhu 2019). These nasalised vowels still exist throughout Wú dialects today. This intermediate step of a CṼ syllable likely occurred for Northern Naga as well among those varieties which today lack a segmental coda among these rimes. In many Patkaian, however, nasalisation of vowels is not generally phonemic, and so the nasalisation on the vowel was likely lost quite quickly in such cases. There is evidence of phonemic vowel nasalisation a century ago in Needham (1897), and Wancho retains it through to today, but it is otherwise uncommon in the group. Retention in Wancho lends some support to this explanation, however, as *#–aŋ^{II}* rime is today realised as /ɑ̃/.

With the case of Wú dialects such as Sūzhōu as a reference, *#–aŋ^{II}* is thus reconstructed as **–aŋ* and *#–aŋ^{I}* as **–ɐŋ*. The apparent flip-flop found in varieties such as Lungri where /–aŋ/ corresponds to **–ɐŋ* can therefore be explained through coda deletion on **–aŋ* followed by loss of nasalisation on the vowel, followed by a lowering of the vowel in **–ɐŋ*. Unlike Sūzhōu Wú, the vowels in these cases did not typically undergo fronting, although this did occur with Jiingu (Dunghi) and the Hahcheng-Ngaimong subgroup as shown in Table 8.

Additionally, **–aŋ* (*#–aŋ^{II}*) has lost the final velar nasal in the majority of Tangsa-Nocte varieties and much of Patkaian / Northern Naga as whole. Some of the few exceptions are Chodok, Hawi, Muklom, Tutsa, Phom, the Tikhak-Yongkuk group and the Ngaimong-Hahcheng group. Conversely, the rime **–ɐŋ* (*#–aŋ^{I}*) has retained the nasal coda in the majority of not just of Tangsa-Nocte but Patkaian more generally. Lochhang, Lama, Mungre, Shecyü and Bote are some of the rare exceptions.

Needham's account provides another useful indicator of the timing of one notable sound change in Muishaung, namely the diphthongisation or vowel fracture of a number of rimes. In the case of Table 4, this is the **–ɐŋ* (formerly *#–aŋ^{I}*) rime which has become /auŋ/ in Muishaung. Muishaung shows a number of similar phonological innovations, many of them likely to be recent. In this case in particular, we can see precisely how recent due to the apparently inconsistencies in how **–ɐŋ* words were transcribed in Needham.

Table 4: Shift from **aŋ^{I}* to /ɐuŋ/ in Muishaung

gloss	Needham 1897	Marrison 1967	Das Gupta 1980	modern
MACHETE	yaung	yaung	ja[u]ng	ʒɐuŋ$_1$
GO UPRIVER	–	–	wang	βɐuŋ$_2$
FOREHEAD	khâng	khang	–	kʰɐuŋ$_2$

BEGIN	–	–	phaung	$p^hau\eta_3$
MANY	pâng	pang	–	$pɐu\eta_3$
PLACE	–	–	baung	$bɐu\eta_1$
WING	rong	rong	–	$rɐu\eta_2$
SIT	nong	nong	ngaung	$ŋɐu\eta_3$
SKY	–	rong	rang	$rɐu\eta_2$
FUT.	khâung	–	–	$k^hɐu\eta_2$
FUT.TR	–	–	shaung	$ʃɐu\eta_2$
FUT.INT	–	–	–	$t^hɐu\eta_2$
SAND	phân̲	phan	–	$p^hɐu\eta_2$
2PL	m'nang	–	mnang	$m_1nɐu\eta_1$

Das Gupta (1980), a much more recent source, shows at least some evidence of this change not having yet been complete at the time of writing. MACHETE is transcribed in two different ways in the text, once with the diphthong and once without, and both SKY and 2SG are given without diphthongisation. Needham (1897), echoed in Marrison (1967), also gives a different vowel for WING and SIT, however French (1983) reconstructs the former as **Ç-rəŋ* and the reconstructed PTB form of the latter also shows a lower vowel than /o/, so we may disregard Neeham's <o> here as either a mistake or something characteristic of his consultants but which did not survive into the present day, where all **–ɐŋ* finals in Muishaung are /ɐuŋ/ and all **–ək* (formerly *#–ak*^{I}) finals are /ɐuŋ/. In the present day the sound change is complete, and no phonemic /ɐŋ/ or /ɑŋ/ finals occur in Muishaung.

6.1 Timing of coda loss in Muishaung

Based on Needham's writing it can also be determined that in the late 1800s the loss of the velar coda on **–aŋ* (*#–aŋ*^{II}) was also not yet complete, as a number of these terms are given with a less-clear <an̲> with an underlined <n̲>. This is not explicitly explained in the text, although an oblique <*n*> is described as a "nasal sound like the French n in 'enfant"'. We may assume this is another typographical error and the underlined <n> is intended to mark nasalisation of the preceding vowel. We can be even more confident of this interpretation based on the sound changes described for Sūzhōu Wú discussed in Section 6 – wherein the loss of a coda resulted in a nasalised vowel which has since been lost – as well as the presence of a nasalised vowel in Wancho and other varieties which have phonemic vowel nasalisation, for example bone **graŋ* as /gã$_2$/ in the Upper Wancho dialect of Kamhua Noknu, or RAPTOR **laŋ* as /lã$_1$/ in the same variety. Burling (1998) in writing about a different variety of Wancho did not include vowel nasalisation, giving RAPTOR as /o^{44}la^{44}/. Whether this is indicative of a loss of nasalisation in that case or simply an oversight in transcription is unclear. This is the case as well for YEAR **grəŋ.ɯ^{ɓ}aŋ* which Burling gives as /zaŋ^{44}pwa^{44}/ but which in Kamhua Noknu is /gəŋ$_1$põ$_3$/, the shift of the final vowel /a/ to nasalised /o/ being a result of influence from the labialised onset.

Table 5: **aŋ*^{II} rime

	PPk	Ch	Jg	Kl	Km	Lo	Lk	Ms	Mr	Ng	Pt	Rk	Sc
ELEPHANT	**tɕ^{h}aŋ*	tɕ^{h}a	tɕ^{h}aŋ	tʃ^{h}a	tɕɤ	c^{h}əu	tɕ^{h}a	tɕɔ	tɕ^{h}ɤ	tɕeŋ		tɕ^{h}a	tɕɤ
YEAR	**βaŋ*	βa	paŋ	va	βɤ	vo	va	βɔ	ve			wa	vɤ
BE.BORN	**tɕ^{h}aŋ*	tɕ^{h}a		tʃa		c^{h}əu	t^{h}a	tɕɔ	tɕ^{h}ɤ	tɕeŋ	tʃaŋ		tɕɤ
LIGHTWEIGHT	**tɕ^{h}aŋ*	tɕa	tɕ^{h}aŋ	tʃa	ʒɯ	c^{h}əu	tɕ^{h}a	tɕɔ	tɕ^{h}ɤ	tɕeŋ	tʃaŋ	tɕ^{h}a	tɕɤ
RED	**ʃaŋ*	xa	ʃaŋ	sa	ʃɯ	səu	ʃa	ʃɔ	ʃɤ	seŋ	saŋ	aʃa	ʃɤ
RAPTOR	**laŋ*	la	laŋ	la	lɤ	ləu	la	lɔ	laŋ			la	lɤ
LOOK	**laŋ*	la	laŋ		lɤ	ləu	–	lɔ	le	leŋ		la	lɤ

BONE	ra1	raŋ1	ra	rɤ1	rəu1	ra1	rɔ1	əre1	reŋ1	raŋ1	ra1	rɤ1
HORNBILL		-raŋ2				-ra2	-rɔ2	-joŋ2		-raŋ	-ra2	-jɔŋ2

The inconsistencies between PNN and PPk are also not limited just to the *#–aŋ* rimes. Another rime having a velar nasal coda which is widespread throughout the basic lexicon has been tentatively reconstructed as *#–juŋ*. As with the *#–aŋ* rimes, French's reconstructions show a considerable range in shape. Examples of such terms include FLLOWER and BREAST (likely a case of polysemous stems), both PNN **puːŋ*, NOSE **C/V-guːŋ*, BACK **kʰeːŋ*, WHITE **luaŋ* and FLAME **Ç-loːŋ*. These are another case where having considerably more data provides insight into a fairly consistent rime, despite a number of nuclei reconstructed for PNN. Not included in French (1983) but widely attested is the stem for MOUTH, *#njʊŋ*. As with examples above, # here indicates that the correspondences are clear, but the underlying phonological values are still being formalised. These reconstructions may be considered accurate as correspondences to the proto-language but not as pronunciations. This is distinct from another rime, reconstructed for PPk as *#–ʊŋ*, which includes etyma such as STONE *#lʊŋ* (PNN *Ç-luŋ*), INSECT *#ʒʊŋ* (PNN **gluŋ)* and HEART *#mʊŋ* (PNN **moŋ*), although in this case there is perhaps something to be said for influence of the /w/ glide as reconstructed for PTB). These too show impressive consistency across Tangsa-Nocte, Khiamniungan, Lainong, Makyam, Konyak and Wancho. Table 6 shows reflexes of these two rimes for a sample of Patkaian languages.

In order from left to right those are Wakching Konyak, Upper Wancho of (Losu & Morey 2003), Phom, Ngaimong-Shangwan, Mungre, Shangthi, Raqnu, Pinku, Muishaung, Joglei and Yaqsawa.

Table 6: **–ʊŋ* and **–jʊŋ* rimes in PPk

	WK	**Ln**	**UW**	**Pm**	**Ng**	**Mr**	**St**	**Rn**	**Pi**	**Ms**	**Jg**	**Ys**
–ʊŋ	oŋ	uaŋ	oŋ	oŋ	uŋ	oŋ	uŋ	uŋ	oŋ	uŋ	oŋ	uŋ
–jʊŋ	eŋ	an	õ	oŋ	oŋ	uŋ	y	o	o	ɨ	oŋ	ɤ

While semantic shifts have resulted in many of these etyma no longer in use among some varieties, and poor data availability for some Patkaian varieties means there will be some holes regardless, these two rimes prove quite regular across Northern Naga.

6.2 Irregular correspondences on HORSE

As for the **–aŋ* and **–ɐŋ* rimes, we find considerable regularity across Patkaian although a few notable exceptions exist. The most significant of these is for HORSE, PTB **k-m-raŋ* (Matisoff 2003), which in Tangsa-Nocte varieties does not consistently pattern with one of the two rimes. Instead, some varieties have reflexes of **raŋ* while in others it appears to derive from **rɐŋ*. It is here suggested that this word had previously been lost in Tangsa-Nocte to be replaced later by a loan from a close relative, most likely Singpho or another Jinghpawic variety with which Tangsa-Nocte speakers were in close contact. Evidence of such a replacement can be seen in Muishaung, for which the word today is /kum$_2$rɔ$_2$/, with a reflex of **raŋ* on the second syllable. Despite considerable phonological innovation in Muishaung compared to other Rangpang Tangsa varieties, this is often in the form of splits rather than mergers. Thus, much can be gleaned from Muishaung in terms of rime categories. Much of Tangsa-Nocte otherwise shows reflexes of **rɐŋ*, even those varieties which are closest to Muishaung. It is therefore possible that a handful of varieties — notably Muishaung, Ringkhu and Hakhun — saw a shift of horse to the **raŋ* stem during some intermediate stage while vowel nasalisation was still present. Marrison (1967) gives <gimrang>, while Das Gupta (1980) transcribes <kumro>. With an onset voicing merger having occurred in Tangsa-Nocte, the <g> and <k> onsets transcribed here should be taken as equivalent. Marrison's transcription supports interpretation as a reflex of **raŋ* in this case, since we would expect to see something more like <raung> for **rɐŋ* at this stage, as it is clear from Needham (1897) that this

shift was well underway nearly 80 years prior. This is further supported by Das Gupta's account (1980). Meanwhile, while alternation between [i~ɨ~u] is common in Tangsa-Nocte, there is no clear indication that this should have been /i/ at this stage, and consulting older Muishaung speakers who were alive at the time of Marrison do not support an older /kim/ pronunciation of the first syllable.

Table 7 shows the distribution of HORSE across varieties of Tangsa Nocte. Note that in Mungre and Shecyü, the second syllable in horse does not appear to be a reflex of either of the two stems, suggesting the recentness of borrowing (Parker, K. 2020).

Table 7: Second morpheme of HORSE across Tangsa-Nocte varieties

variety	*–ɐŋ	HORSE σ2	*–aŋ
Bote	-aŋ	raŋ	-a
Champang	-əŋ	rəŋ	-ə
Cholim	-aŋ	raŋ	-a
Gaqlun	-aŋ	raŋ	-a
Hahcheng	-aŋ	raŋ	-eŋ
Hakhun	-ɤ	ra	-a
Jiingi	-əŋ	rəŋ	-ɛ
Joglei	-əŋ	raŋ	-aŋ
Khâlak	-aŋ	raŋ	-a
Longri	-aŋ	raŋ	-a
Maitai	-aŋ	raŋ	-a
Muishaung	-auŋ	rɔ	-ɔ
Mungre	-ɔ	raŋ	-ɛ
Ngaimong	-aŋ	raŋ	-eŋ
Rinkhu	-aŋ	ra	-a
Shecyü	ɑ	raŋ	ɨ
Tikhak	-aŋ	raŋ	-aŋ

Another exceptional stem is ELEPHANT, which, like HORSE, is not consistent in terms of which rime it patterns with. It has been suggested that this is a loan from Southwestern Tai (van Dam 2018; Morey & van Dam 2019) along with a number of other stems found in Tangsa-Nocte relating to wet rice cultivation.

Table 6 shows the reflexes of the two rimes discussed here as they appear across Patkaian. Columns represent regular correspondences for Tutsa, Chuyo, Haqchum, Jiingi, Hahcheng, Shecyü, Muishaung, Pinku, Phom, Wancho and Lainong in that order.

Table 8: Reflexes of *–ɐŋ and *–ɑŋ in some Patkaian varieties

		Tu	Cy	Hc	Ji	Hh	Sc	Ms	Pi	Pm	UW	Ln
*–ɐŋ	#–aŋ^{I}	aŋ	ɑŋ	ɑŋ	əŋ	ɑŋ	ɑ	ɐuŋ	ɑŋ	æŋ	aŋ	ɑi
*–aŋ	#–aŋ^{II}	aŋ	ɑŋ	ɑ	e	iŋ~eŋ	ɨ	ɔ	ɑ	æŋ	ã	o

7. Conclusion

From the above analysis a few conclusions can be made. First, with considerable consistency of reflexes, two distinct rimes can be reconstructed for stems reconstructed as *–*aːŋ*, *–*aŋ* and *–*ɐŋ* in French (1983). These are reconstructed here as *–*ɐŋ* and *–*aŋ*. Very loosely, French's *–*aŋ* and *–*ɐŋ* tend to correspond to PPk *–*ɐŋ*, while French's PNN *–*aːŋ* roughly aligns with PPk *–*aŋ*. However, many exceptions are to be found, almost certainly due to a paucity of data available for the PNN reconstruction.

Second, similar to sound changes seen with the Sūzhōu dialect of Northern Wú, in many Patkaian varieties, the *–*aŋ* rime saw a loss of the velar nasal coda, resulting in an intermediate stage involving nasalisation of the vowel. This nasalisation was then lost in nearly all varieties outside of Wancho which has retained phonemic vowel nasalisation. Meanwhile, *–*ɐŋ* rimes generally retained the coda across Patkaian varieties with only a few exceptions. In many cases the vowel was lowered resulting in an /aŋ/ reflex. In the case of Muishaung, the centrality of the vowel instead underwent vowel fracture, a process that was still in its early stages at the time of Needham (1897). A parallel change occurred to the PPk *–*ɐk* rime, which consistently patterns with *–*ɐŋ* across much of Patkaian.

Additionally, the lack of a reconstructable *–*ak* rime may be explained as the result of an early merger with *–*ɐk* during the development of Patkaian tone systems, where in nearly all varieties but Khiamniungan CVC syllables formed their own 'checked' tone category. For many of the Patkaian tone systems, duration is a secondary feature, and such checked syllables are almost always significantly shorter in duration. Mulder (2020) provides one possible counter-example in Muklom, with a length distinction on checked tones described for /puk/ for stomach and /puːk/ for ELEPHANT. For this to be conclusive, more investigation is needed. Gogoi (2025 p.c.) has suggested a similar length distinction on checked syllables for Tutsa, although voicing may play a role in this and it too requires further investigation.

The work undertaken by W. T. French in his 1983 reconstruction of Northern Naga is still of considerable importance. It was no small task, especially given the overall lack of data for what is otherwise an incredibly diverse and vibrant branch within Tibeto-Burman. its significance cannot be understated, not just in terms of showing relatedness of stems across the languages included in the study, but also in solidifying the name of the family for decades to come, with its only shortcoming being that it was a product of a time where data was much more difficult to come by, as would be the case for the following two and a half decades after its publication. Only in the past decade has the situation improved substantially in terms of how much data access there is, in large part due to the increasing number of community member scholars contributing to the collection and publication of linguistic works, boding well for an ever-improving understanding of the linguistic and cultural history of the Patkai mountains.

Abbreviations

PPk	Proto-Patkaian
PNN	Proto–Northern Naga of French (1983)
σ	syllable within a lexeme

Language abbreviations

Ch	Cholim	Mr	Mungre
Cy	Chuyo	Ng	Ngaimong
Hh	Hahcheng	Pm	Phom
Hl	Haidle	Pt	Phong
Hc	Haqchum	Pi	Pinku

Ji	Jiingi	Rn	Raqnu
Jg	Joglei	Rk	Ringku
Kl	Khalak	St	Shangthi
Km	Kimsing	Sc	Shecyü
Ln	Lainong	TK	Tanhai Konyak
Lo	Lochhang	Tu	Tutsa
LW	Lower Wancho	UW	Upper Wancho
Lk	Lunkhi	WK	Wakching Konyak
Ms	Muishaung	Ys	Yaqsaq
Ml	Muklom		

References

Bano, Atiqua (2008). "A descriptive study of Phom language". PhD thesis. North Eastern Hill University, Shillong.

Boro, Krishna (2017). "A grammar of Hakhun Tangsa". PhD thesis. University of Oregon.

Burling, Robbins (1983). "The Sal Languages". In: *Linguistics of the Tibeto-Burman Area* 7.2.

Burling, Robbins and L. Amon Phom (1998). "Phom phonology and word list". In: *Linguistics of the Tibeto-Burman Area* 21.2, pp. 13–71.

Burling, Robbins and Mankai Wangsu (1998). "Wancho phonology and word list". In: *Linguistics of the Tibeto-Burman Area* 21.2, pp. 43–71.

Chao, Yuen Ren (1930). "ə Sistim əv Toun-letəz". In: *Le Maître Phonétique* 30, pp. 24–27.

Dam, Kellen Parker van (forthcoming). "Patkaian (Northern Naga)". In: *Oxford Handbook of the Tibeto-Burman Languages.* Oxford University Press

— (2025). "'A Classified Account of J. F. Needham's A Collection of A Few Môshâng Naga Words'". In: *Journal of Asian and African Studies*.

— (2020) "Borrowing and re-borrowing in Southeast Asia: Complicating factors in historical reconstruction". *Workshop on Tracing contact in closely related languages*. Zürich

— (2019). "Horsing around: The Complicated State of Tangsa-Nocte *aŋ Rhymes". 52nd International Conference on Sino-Tibetan Languages and Linguistics. University of Sydney, Sydney, Australia.

— (2018). "The Tone System of Tangsa-Nocte and Related Northern Naga Varieties". PhD thesis. Melbourne, Australia: La Trobe University.

Dam, Kellen Parker van and Syed Iftiqar Rahman (2022). "Tangsa-Nocte as a Continuum: A diagnostic feature list for classification of varieties". In: *Himalayan Linguistics: Languages and Peoples of the Eastern Himalayan Region* 21.2.

Das Gupta, Kamalesh (1980). *The Tangsa Language: A Synopsis*.

Dutta, Niharika (2023). "A descriptive grammar of Phong". PhD thesis. Gauhati University.

Dutta, Parul (1969). *The Tangsas of the Namchik and Tirap Valleys*. North East Frontier Agency.

French, Walter Thomas (1983). *Northern Naga: A Tibeto-Burman Mesolanguage*. City University of New York.

Gogoi, Sayanjyoti (2025). personal correspondence.

Grierson, George A (1903). *Linguistic Survey of India*. Calcutta: Office of the Superintendent of Government Printing.

Jacques, Guillaume and Methna Tanhai Konyak (2010). "A Preliminary Study of Tanhai Konyak Historical Phonology". In: *La linguistique* 46.2, pp. 41–57.

Konyak, Hoipo and Mijke Mulder (2022). "A Brief Outline of Chen Phonology". In: *Proceedings of the Payap University Research Symposium*. Chiang Mai: Payap University.

Losu, Banwang and Stephen Morey (2023). "The Wancho language of Kamhua Noknu village". In: *Linguistics of the Tibeto-Burman Area* 46.2, pp. 201–234.

Marrison, Geoffrey Edward (1967). "The classification of the Naga languages of north-east India." PhD thesis. School of Oriental and African Studies (University of London).

Matisoff, James A (1978). A culturally appropriate lexicostatistical model for Southeast Asia languages.

— (2003). *Handbook of Proto-Tibeto-Burman: System and philosophy of Sino-Tibetan reconstruction*. Berkeley, Los Angeles: University of California Press.

Morey, Stephen (2017a). "Tangsa". In: *Sino-Tibetan Languages*. Ed. by Graham Thurgood and Randy LaPolla. London and New York: Routledge.

— (2017b). "The sociolinguistic context of the Tangsa languages". In: *Sociohistorical Linguistics in Southeast Asia*. Brill, pp. 169–187.

Morey, Stephen and **Van Dam, Kellen Parker** (2019). "Material culture and Agriculture in Tangsa (Naga) languages – evidence from Tai borrowings and traditional song texts". 24th International Conference on Historical Linguistics. Australian National University, Canberra, Australia.

Mulder, Mijke (2020). "A descriptive grammar of Muklom Tangsa". PhD thesis. La Trobe.

Needham, Jack F (1897). *A Collection of A Few Môshâng Naga Words*. Shillong: Assam Secretariat Printing Office.

Post, Mark W and Robbins Burling (2017). "The Tibeto-Burman languages of Northeast India". In: *The Sino-Tibetan Languages*, pp. 213–242.

Shafer, Robert (1953). *Classification of the northernmost Naga languages*. Vol. 39, pp. 225–264.

Walker, G.E.D. (1948). *Outline grammar of Yogli Naga*. Shillong.

Wayesha, Ahsi James (2010). "A phonological description of Leinong Naga". MA thesis. Chiang Mai: Payap University.

Zhu, Yuhong (2019). "Diachrony of Coda **ŋ* in Suzhou, Wu Chinese: Regular Change and Differing Readings". In: Proceedings of the 30th North American Conference on Chinese Linguistics (NACCL-30). ed. by Yuhong Zhu et al. Vol. 2. Columbus, pp. 444–461.

Language Vitality and Endangerment: A Case Study of Uchoi

Niloy Chakraborty
Research Scholar
Department of Linguistics and Tribal Languages
Tripura University (A Central University)
nilaychk@gmail.com

Abstract

Uchoi is a Tibeto-Burman language of Bodo-Garo group. The Language is largely spoken in the north-eastern state Tripura, especially in South Tripura and Gomati Districts. According to the 2011 census report of India, the tribe has only 2447 native speakers. The language is in close contact with the Kokborok language (the dominant language among the tribes in Tripura). For that reason, there is a language shift towards Kokborok. If the process continues, it will become no one's speech. Here in this paper, trying to examine the endangerment status based on UNESCO's framework. The data has been collected in the field visit during March to April, 2024. The language is used only by the older generation speakers, the younger generation either shift to Kokborok or Bangla. The language is restricted only in the religious and cultural domain. Even, due to the popularity of Christianity, the newer generation do not know many of their cultural traditions and also do not practice them. There are three major NGOs', that are working to spread the awareness of language among the younger generation such as Uchoi Youth Association, Tripura Uchoi Youth Association (TUYA) and Uchoi Welfare Society (UWS). There is only one comprehensive linguistic work (PhD Thesis) in Uchoi, i.e. 'The Descriptive Grammar of Uchoi' by Tamojoy Brahma, where he failed to mention the degree of endangerment and language vitality regarding this language. So, this paper will try to give a view on language vitality and endangerment status in Uchoi.

Keywords—Uchoi, Tibeto-Burman, UNESCO, language vitality, endangerment.

1. Introduction

Language, then, like everything else, gradually transforms itself over the centuries, there is nothing surprising in this; in a world where humans grow old, tadpoles change into frogs, and milk turns into cheese (Atchison 2001:4). Over the course of the next century, however, the very basic global picture will change drastically, a certain small group of languages (Arabic, Chinese, English, Spanish, to name a few) will be used by an ever-increasing percentage of the world's population, while a great many other languages will completely disappear (Austin, 2007: viii). Language endangerment is typically used to describe the rapidly declining global language population. On the one hand, one might say, language loss has been a reality throughout history; and on the other, the loss of a language is of no great moment either for science or for human intellectual life (Hale 1992:1). However, when a native speaker stops using a language as their first language (L1) or switches to another language because of the impact of the dominant language or perhaps because of political concerns, the language is in danger. Therefore, if native speakers continue to alter their language, the language will eventually go extinct because there won't be any speakers remaining. As Joshua Fishman (1991:2) rightly quoted, 'a language is endangered if it is not being passed on to the younger generation'. When languages become extinct, not only can concepts and methods disappear, but whole conceptions of reality are also erased from the

map (George V. Driem, 2007:11). We can justify the issue of language endangerment into two ways—firstly to preserve all the lesser-known languages and secondly only prioritizing the known languages. But in the continent like Africa where many tribes stay side by side for a long period of time it has become very problematic to implement government policies due to language differences. Tanzania, for example, is a country which is striving for unity, and the spreading of Swahili is regarded as a major force in this endeavour. Tribalism is seen as a threat to the development of the nation, and it would not be acting responsibly to do anything which might seem, at least superficially, to aid in its preservation (Ladefoged 1992: 809).

2. The Objective of the Study

The current study will concentrate on the issues of language endangerment in Northeast India, specifically referring the Uchoi language, based on UNESCO's framework on Language Vitality and Endangerment (2003). Apart from that the study will also try to examine the following objectives-

a. To assess the degree of endangerment.
b. Factors responsible for the endangerment of Uchoi.
c. Also suggest the proposed steps for protection and preservation of the language.
d.

3. Methodology

Both the primary and secondary sources of data were incorporated for this study. Books, journals, articles, documents, the internet etc. are being utilized as the source of secondary data. The primary data was conducted with the assistance of the interview method with the community speakers from South Tripura District and Gomati District, the data collection was mostly based on fieldwork conducted in March—April of 2024. Digital camera (Nikon D5300) and audio recorder (Zoom H4N Pro) were used to capture and digitize the data.

4. Language Endangerment

The study of Endangered languages catches the eyes of the linguists in the recent past, in the late 80s and early 90s. It starts with the International Linguistics Congress in Quebec in 1992, when they attended the meeting and make a statement—

> "As the disappearance of any one language constitutes an irretrievable loss to mankind, it is for UNESCO a task of great urgency to respond to this situation by promoting and, if possible, sponsoring programs of linguistic organizations for the description in the form of grammars, dictionaries and texts, including the recording of oral literatures, of hitherto unstudied or inadequately documented endangered and dying languages." (Crystal, 2000: vii)

To which, UNESCO did reply in the very next year at a Conference in November, 1993. As an outcome they (members of International Linguistics Congress) took two essential steps; firstly, they adopted 'Endangered Languages Project' and secondly, they formed 'Red Book of Endangered Languages' and also submit a progress report after few months. As per the report—

> "Although its exact scope not yet known, it is certain that the extinction of languages is progressing rapidly in many parts of the world, and it is of the highest importance that the linguistic profession realize that it has to step up its descriptive efforts." (Crystal, 2000: vii)

According to Michael Krauss, Language Endangerment can be scaled from 'safe' to 'moribund'. He suggested 4 levels of danger i.e. Safe, Endangered, Extinct and Moribund (a state where the language is not learned any more as a mother tongue) [Krauss 1992:4]. Kincade (1991:160-3) proposed 5 levels, distinguishing from 'safe' and 'not so safe'; *viable, viable but small languages, endangered languages, nearly extinct languages, extinct languages*. According to Stephen Wurm [1992:192] focuses on the weaker languages, and he also mentioned 5 levels as *potentially endangered languages, endangered languages, seriously endangered languages, moribund languages, extinct languages*. Apart from these, Nancy Dorian (1981:27—28) postulated that the type of change in endangered languages was the same as in healthy languages, but the rate of change was greatly accelerated; the defining features of an endangered language were sociolinguistic rather than structural.

On the otherhand, UNESCO also tried to categorize languages on the basis of endangerment as: *Vulnerable*, *Definitely Endangered*, *Severely Endangered*, *Critically Endangered*, and *Extinct*. According to the UNESCO (2003:11), a language is called to be *Vulnerable* when the language is used by some children in all domains or used by all of them in limited domains. *Endangered* language is when the language is restricted with the parental generation only. A language is *Severely Endangered* when the language is only spoken by the older generation, and the successor of the language only understand the language but can't speak among them or pass to their children i.e. the third generation. *Critically Endangered* is when the youngest speakers of the language are at age of grandparents or older, and they don't speak that particular language fluently and frequently. Lastly, an *Extinct* language is when there is no speaker left to speak or pass the language to anyone. In this scale or parameter, it will be tough or nearly impossible for the linguist to revive the language if it passes through the first two stages i.e. Vulnerable and Definitely Endangered. So, how much of the world's linguistic diversity is endangered? We do not have any precise answer. Perhaps the only way to get some idea of the extent of endangerment is to look at the sizes of living languages (Nettle & Romaine 2001:40). The following table shows the percentages of languages according to continent having fewer than indicated number of speakers.

Continent	<150	<1000	<10,000	<100,000	<1,000,000
Africa	1.7	7.5	32.6	72.5	94.2
Asia	5.5	21.4	52.8	81	93.8
Europe	1.9	9.9	30.2	46.9	71.6
North America	22.6	41.6	77.8	96.3	100
Central Am.	6.1	12.1	36.4	89.4	100
South Am.	27.8	51.8	76.5	89.1	94.1
Aus/Pacific	22.9	60.4	92.8	99.5	100
World	11.5	30.1	59.4	83.8	95.2

Table—1 Percentages of languages according to continent having fewer than indicated number of speakers[1]

[1] Source: Daniel Nettle, *Linguistic Diversity*, Oxford University Press, 1999:114

5. Uchoi: An Overview

The name Uchoi, as spelled by Thomas H. Lewin in 1869, refers to a small ethnic community that mainly lives in southern Tripura. Only 2447 people belonged to this tribe as per the 2011 census. The Chittagong Hill Tract is home to some Uchois in the Khagrachori Hill District. However, they consistently identify as Tippera, while the Mrung, the Mranma (Arakanese) word for Reang, is what their neighbouring tribes call them. However, Uchai is more like Bru than Tippera in terms of ethnicity; Debnath (2010, p. 31) describes how their tradition binds Bru and Uchai together as brothers, and they both went through early Tippera rule. The scholars have struggled with Uchai's ethnicity. They are the Kuki origin, according to R. H. Sneyd Hutchinson (1906), although S. Debbarma (1983) believes they are a blend of Tippera and Kuki. According to Debbarma's analysis in 1983, the term Uchai originates from the title 'Ulchoi' bestowed upon them by the Tripura ruler of that era. The Kokborok term for 'Ulchoi' is 'ullo-' which means 'after' and 'čoi' which means 'to follow' (ul-čoi phainai/ 'one who came after following someone). Ulchoi later became Uchoi and is now known as Uchai as time goes on. Once more, according to their stories, Debnath (2010) provides us with an intriguing account of the Uchoi migration: "…a particular monarch of Arakan valley abducted their and put them in the highlands above Maungdaw, where they live under their own Kings in peace for a very long period. Then, one day, a group of people arrived from the west, and their monarch made an effort to get along with the Maungdaw King by inviting him to supper. However, the latter was given poison along with his meal, which caused him to pass away. A messenger was dispatched to Maungdaw in the interim to urge the populace to pledge allegiance to the new monarch; if they did not, they would face the danger of a punitive expedition. Several hundred men, women, and children were brutally killed when they refused to obey. 'Ringsa' and 'Uchya' the only two brothers, were able to get away. However, they separated the survivors into two groups because they were afraid of being pursued. Ringsa, the senior, arrived Tripura was first, before Uchya. The Riangs derived from the clan that came before them, according to the Uchais of today, who claim to be descended from the younger brother, Uchya.

Regardless of the Uchais ethnicity, it is understood that they came from the Chittagong Hill Tracts during the reign of Maharaja Birendra Kishore Manikys in the second decade of the 20th century. Under the guidance of headman Ramananda Uchoi, they made their first foray into modern day Tripura, from where the then King of Tripura rehabilitated them at Srakbai, or modern day Charakbai, under the Belonia sub-division of South Tripura district. Afterwards, they scattered and established themselves in many other regions of Tripura, particularly in South Tripura.

The villages where Uchais currently reside in large numbers, are Kewai Uchoi Para, Dhanachandra Uchoi Para, Purna Uchoi Para, Punniram Uchoi Para, Sogoram Para, Khumbaiha Uchoi Para, Thandachandra Uchoi Para, and Mission Hathai Para under the Karbook sub-division of Gomati district; they also live in somewhat sparsely populated Dasda & Jampui villages in Kanchanpur sub-division of North Tripura District and Ratanpur, Muhuripur, & Khahsingkho villages under Belonia sub-division of South Tripura district.

6. Factors Responsible for Language Endangerment

Various factors are involved in language endangerment, it is not a matter of suddenness, but a gradual decline. Some of the factors that are common in world languages are being described here with the reference to our concerning language Uchoi. Six key criteria are used to evaluate language vitality, as stated in UNESCO's framework on Language Vitality and Endangerment. They are discussed below:

a. Intergenerational Language Transmission.
b. The Absolute number of speakers.
c. Proportion of speakers within the total population.

d. Shifts in Domains of Language Use.
e. Response to new domains and media.
f. Availability of Materials for language education and literacy.

The factors listed above are intended to concentrate on the current investigation in order to determine the extent and degree of endangerment in Uchoi.

6.1 Intergenerational Language Transmission

Fishman (1991:48) rightly pointed out as, *the most commonly used factor in evaluating the vitality of a language is whether or not it is being transmitted from one generation to the next.* Language Endangerment can be categorized on a scale from stability to extinction. Again, under Intergenerational Language Transmission, six degrees of endangerment can be distinguished such as—*safe*, there is an uninterrupted language transmission from one generation to the next and there is no visible threat from any other language; *stable yet threatened*, though the language is spoken in most contexts and un-interruption among the generations but multilingualism or one/two dominant language trying to usurped in some domain; *unsafe*, a language is unsafe when it has some restrictions in its use or restricted to specific domains, as all the members of the community specially the new generation speakers do not consider it as their first language; *definitely endangered*, the language is no longer used as mother tongue or first language in home domain, the most youngest generation of that particular language is the parental generation, at this stage though the parents use their mother tongue with their children but they are not always answer with the native tongue; *severely endangered*, when the language spoken only by the older generation or the generation of the grandparents, though the parental generation understand the language but found it difficult to speak; *critically endangered*, when the language is only used by the great grandparents and again it is transmitted only to the grandparents stage. And the older generation do not speak the language on regular basis but use if when they find someone to speak on that particular language.

They also reside in the somewhat sparsely populated Dasda & Jampui villages in the Kanchanpur sub-division of the North Tripura district, as well as the Ratanpur, Muhuripur, and Khahsingkho villages in the Belonia sub-division of the South Tripura district, as well as Kewai Uchoi Para, Dhanachandra Uchoi Para, Purna Uchoi Para, Punniram Uchoi Para, Sogoram Para, Khumbaiha Uchoi Para, Thandachandra Uchoi Para, and Mission Hathai Para under the Karbook subdivision of the Gomati district. The language is not spoken by the speakers of all generations, and there is very little evidence of the language being passed down to the following generation. Thus, it is determined that the language is critically endangered. The Uchoi language is exclusively used at religious ceremonies and other customary activities; it is not spoken in other circumstances. Nonetheless, a number of non-governmental organizations (NGOs[2]) have been established to encourage the younger generation to acquire and utilize the language. The use of the language among the younger generation is very restricted, not all the children use this language in all domains. As a result, shown in the following Table—2, the language is fragile since not all speakers of the various age groups speak it in all domains especially the young speakers.

[2] The Uchoi Youth Association, an NGO, is an active group that has just lately begun to work for the community's welfare. In this context, the inaugural conference took place in 2022. Other groups, such as the Tripura Uchoi Youth Association (TUYA) and the Uchoi Welfare Society (UWS), exist in addition to this one. Together, they have been promoting both the language and the community's development.

Degree of Endangerment	Grade	Speaker Population
Safe	5	The language is used by all ages, from children up.
Unsafe	4	The language is used by some children in all domains; it is used by all children in limited domains.
Definitively endangered	3	The language is used mostly by the parental generation and up.
Severely endangered	2	The language is used mostly by the grandparental generation and up.
Critically endangered	***1***	***The language is used by very few speakers, mostly of great-grandparental generations.***
Extinct	0	There are no speakers.

Table—2 Intergenerational language transmission

6.2 Absolute Number of Speakers

According to the 2011 Census report, the total number of Uchoi speakers is only 2447. It is always problematic to provide the valid count of absolute numbers. The absolute number of speakers of Uchoi language is unknown. However, we can estimate that there are about 250 Uchoi speakers[3], including priests and older generation people. A good number of speakers has been shifted to the dominant neighboring language Kokborok.

6.3 Proportion of Speakers within the Total Population

The number of speakers in relation to the total population of a group is a significant indicator of language vitality, where 'group' may refer to the ethnic, religious, regional or national group with which the speaker community identifies (Language Vitality and Endangerment, 2003:9).

The districts of Gomati, and South Tripura are home to the majority of Uchoi speakers. The South Tripura district, which includes Belonia, Hrishyamukh, Jolaibari, and others, is home to the majority of the speakers. Only a small percentage of individuals speak the language, mostly members of the priestly class and elder generations, with very few younger generations using it. So, the language can be assessed as a critically endangered language after looking at the proportion of speakers within its overall population in Tripura i.e., approximately 0.0667%. While discussing this point it is quite evident from the Table—3 that the language is critically endangered as very few speak the language.

[3] According to the villagers of Depachhera, during the field investigation in Korbook Subdivision (Gomati District) in 2024, April.

Degree of Endangerment	Grade	Proportion of speakers within the Total Reference Population
Safe	5	All speak the language
Unsafe	4	Nearly all speak the language
Definitively endangered	3	A majority speak the language
Critically endangered	2	A minority speak the language
Critically endangered	***1***	***Very few speak the language***
Extinct	0	None speak the language

Table—3 Proportion of Speakers within the Total Population

6.4 Shifts in Domains of Language Use

The language is considered to be alive, when speakers have the chance to communicate in their native tongue, at least within their own community or in their home domain. However, the lack of prestige causes the less prestigious language speakers to switch to a different language in order to take advantage of socioeconomic opportunities, and as a result of this ongoing process, the speakers of the minority language become more and more integrated into the dominant language and culture. Additionally, speakers of Uchoi switched to Kokborok, the dominant language. Since most speakers have already switched to Kokborok, which is mainly used for chanting prayers or at religious ceremonies, the remaining senior members or speakers are not given the chance to speak even in their home domain. The Uchoi language is exclusively utilized in extremely formal contexts, such as administration and ceremonial. At the community centre, at festivals, and during ceremonial events where senior citizens can interact, the language may still be spoken. Grandparents' and other elderly extended family members' homes, as well as other customary senior gathering spots, may likewise fall under the limited domain. Despite not being able to speak the language, many individuals can understand it. As a result, the language has a very limited domain (Table—4), practiced only in a very restricted number of domains and for very limited functions.

Degree of Endangerment	Grade	Domains and Functions
Universal use	5	The language is used in all domains and for all functions
Multilingual parity	4	Two or more languages may be used in most social domains and for most functions
Dwindling domains	3	The language is used in home domains and for many functions, but the dominant language begins to penetrate even homa domains
Limited or formal domains	2	The language is used in limited social domains and for several functions
Highly limited domains	***1***	***The language is used only in a very restricted number of domains and for very few functions***
Extinct	0	The language is not used in any domain for any function

Table—4 Shifts in Domains of Language Use

6.5 Response to New Domain and Media

The language itself finds no place in social domains such as school, college, any kind of institutions, market or any public places and also it is not found to be listed as the official language of the state. Therefore, the response to new domain is very marginal or more precise to say very restricted and only available on digital or social media platforms such as Facebook, You-tube, Instagram etc. If the communities do not meet the challenges of modernity with their language, it becomes increasingly irrelevant and stigmatized (Language Vitality and Endangerment, 2003:11). Thus, the language falls under the grade point 1 (as per UNESCO's parameters), and the degree of endangerment status is minimal as shown in Table—5.

Degree of Endangerment	**Grade**	**New Domains and media accepted by the Endangered Language**
Dynamic	5	The language is used in all new domains.
Robust	4	The language is used in most new domains.
Receptive	3	The language is used in many new domains.
Coping	2	The language is used in some new domains.
Minimal	***1***	***The language is used only in a few new domains.***
Inactive	0	The language is not used in any new domains.

Table—5 Response to new domains and media

6.6. Availability of materials for Language Education and Literacy

Language vitality is significantly influenced by language education. Some linguistic communities do not want their language to be written down, and others have strong oral traditions. Language literacy is a source of pride in other societies. But generally speaking, literacy, social and economic advancement go hand in hand. Since the Uchoi lack their own writings and scripts, there aren't many resources or texts available in their language, and priests primarily teach it informally. Recently few NGOs (Uchoi Youth Association) are actively working to educate the young learners to learn their own language and culture, but unfortunately majority of the population has already been converted to Christianity so the transmission of culture is not happening among the younger generations. Another significant factor is that in Uchoi literature, written materials exist, but they may only be useful for some members of the community; for others, they may have a symbolic significance. Literacy education in the language is not a part of the school curriculum, as the following Table—6 suggests.

Grade	Availability of Written Materials
5	There is an established orthography and a literacy tradition with grammars, dictionaries, texts, literature and everyday media. Writing in the language is used in administration and education.
4	Written materials exist, and at school, children are developing literacy in the language. writing in the language is not used in administration.
3	Written materials exist and children may be exposed to be written form at school. Literacy is not promoted through print media.
2	***Written materials exist, but they may only be useful for some members of the community; for others, they may have a symbolic significance. Literacy education in the language is not a part of the school curriculum.***
1	A practical orthography is known to the community and some material is being written.
0	No orthography is available to the community.

Table—6 Availability of written materials

7. Amount and Quality of Documentation

It is imperative that languages be documented before they go extinct. A little amount of documentation in the Uchoi language may be discovered in primers, dictionaries, and texts that were initially written in Bangla script. Additionally, there aren't many linguistics studies that focus on the sociocultural or religious characteristics of the language; in addition, descriptive grammar is present in this language, although speakers aren't particularly conscious of it. Therefore, in order for the language to continue, it urgently needs appropriate documentation.

Using the following Table—7, one can ascertain the extent and degree of endangerment to the Uchoi language.

Intergenerational Language Transmission	1	Critically endangered
Absolute number of speakers	250	Critically endangered
The proportion of speakers within the total population	1	Critically endangered
Shifts in Domains of Language Use	1	Highly limited domains
Response to New domains and Media	1	Minimal
Availability of materials for language education and literacy	2	Written materials exist, but they may only be useful for some members of the community; for others, they may have a symbolic significance. Literacy education in the language is not a part of the school curriculum.

Table—7 Documentation

8. Factors responsible for the endangerment of Uchoi

In some sociolinguistic contexts, the causes of language endangerment differ from one language to another. The Uchoi language is in danger because of the following factors:

a. Kokborok, the prevalent and regional language, has a significant influence on Uchoi.
b. Discontinuation of their long-standing traditions and cultural activities, such as sports, proverbs, meditation, hunting, folk music, and dances.
c. Absence of oral traditions and written texts.
d. Insufficient formal instruction in their own tongue.
e. Speakers' negative attitude towards their native tongue.
f. The language is used in a relatively limited domain.
g. Transmission of language between generations is rare.
h. There aren't many speakers left in the language.
i. Due to a lack of status, parents make their children learn Hindi (National language), English (International language), and Kokborok (Dominant regional language), in order to get possibilities.

9. Proposed steps for protection and preservation

a. Children should be encouraged by their parents to learn and use the Uchoi language both within and outside the home.
b. By planning awareness campaigns and academic exercises like reciting folktales and rhymes that have been passed down orally from generation to generation, speakers can cultivate a favourable attitude toward the language.
c. To maintain the language, the Uchoi people should remember their ancient cultural practices and customs and share the knowledge with the younger generation speakers.
d. Folk dance, songs, and other traditional customs should be recorded on audio and video format.
e. Regardless of its potential or population size, the Uchoi language ought to be taught in formal education especially into the villages where Uchoi are in majority.
f. By holding workshops, seminars, and other events, native speakers of the language should be encouraged to contribute to the literary body and enhance their literature.
g. The creation of stories, poetry, proverbs, plays, and novels in their own tongue should be promoted.
h. In addition to using the language in formal schooling, the government should take the appropriate steps to preserve and advance the language among the next generation.

10. Conclusion

Uchoi is a less studied language, when it comes to description and documentation. The Uchoi language is in danger of being extinct due to the shift in the majority language, and Kokborok is the language spoken by the majority. Therefore, before we lose the native speakers of the Uchoi language, it is imperative that thorough research and accurate documentation be done. In this paper, examine the factors that are responsible for endangerment. The study based on the six vital factors as suggested by the UNESCO, regarding language endangerment. After minutely examining the factors, the Uchoi language is critically endangered, as the language is used by very few speakers, mostly of great-grandparental generations; very few speak the language; the language is used only in a very restricted number of domains and for very few functions; and lastly written materials exist, but they may only be useful for some members of the community; for others, they may have a symbolic significance. Literacy education in the language is not a part of the school curriculum. Keeping this in mind, the current study focuses on the Uchoi speakers' fight to preserve their language and culture in Tripura.

References

Austin, P. & Simpson, A. 2007. *Endangered Languages: Languages Loss and Community Response*. Cambridge: Cambridge University Press.
Aitchison, Joan. 2001. *Language change: Progress or decay?* 2nd edn. Cambridge: Cambridge University Press.
Brahma, Tamojoy. 2020. *A Descriptive Study of Uchoi.* PhD thesis. Assam University.
Crystal, David. 2000. *Language Death*. Cambridge: Cambridge University Press.
Dorian, Nancy C. 1981. *Language Death: the life cycle of a Scottish Gaelic dialect*. Philadelphia: University of Pennsylvania Press.
Debnath, R. 2010. *Exploring Highlanders of Tripura and Chittagong Hill Tracts*. New Delhi: Akansha Publishing House.
Fishman, Joshua, A. 1991. *Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to threatened Language*. Clevedon: Multilingual Matters.
Fishman, J.A. 1987. *Language Spread and language policy for endangered languages*. In Proceedings of the Georgetown University Round Table on Language and Linguistics. Washington: Georgetown University Press, 1—15.
Fromkin, Victoria. 2009. *Language: Nature, Psychology and Grammatical Aspects*. Delhi: Cengage Learning.
Haokip, Marykim. 2013. *The Nature of Language Endangerment: A Case study of Thadou* in M. Ganeshan (ed.), Language Endangerment in South Asia Vol. I. Annamalai University Publication No. 10.234—255.
Hale, Ken. 1992a. *On endangered languages and the safeguarding of diversity*. Language 68. 1—3.
Ladefoged, Peter. 1992. *Another view of endangered languages*. Language 68. 809—811.
Lewin, T. H. 1869. *The Hill Tracts of Chittagong and the Dwellers Therein.* Calcutta: Bengal Publishing Company.
Matisoff, James A. 1991. *Endangered languages of Mainland Southeast Asia*. In Robins and Uhlenbeck (eds.), 189—228.
Nettle, Daniel. & Suzanne, Romaine. 2000. *Dying Words: The Extinction of World's Languages*. Oxford University Press.
Nettle, Daniel and Suzanne Romaine. 2001. *Vanishing voices*. Oxford: Oxford University Press.
UNESCO. 2002. *Universal Declaration on Cultural Diversity*. Cultural Diversity Series. Paris: United Nations Educational, Scientific and Cultural Organization.
Wurm, Stephen A. 1991. *Language death and disappearance: causes and circumstances*. In Robins and Uhlenback (eds.), 1—18.

Khortha: Swadesh Word List 100

Ashi Urvi[*] & Mona Parakh[+]
Department of Linguistics,
The Maharaja Sayajirao University of Baroda
ashiurvi7667@gmail.com[*],
mona.parakh-linguistics@msubaroda.ac.in[+]

Abstract

Khortha, an Indo-Aryan language spoken in Jharkhand, Bihar, and parts of West Bengal, is classified under the Eastern Magahi dialect of the Bihari subgroup.

The present paper is based on the data collection of the Swadesh Word List-100 for the Khortha language. The SWL-100 Project run by the Society for Endangered and Lesser-Known Languages (SEL) aims to develop linguistic resources for researchers, and as an initial step, collects SWL-100 words across various Indian languages. This collection is accompanied by audio recordings and IPA transcriptions. The data SWL-100 for Khortha has been collected through structured recording sessions with a native speaker.

Glottolog (2024) serves as a key reference for classification of languages, but Khortha as a lesser-known language does not have an assigned ISO code or Glottocode. Hence, for the present study, the recordings of the Khortha words have been labelled using the code "KRT".

Keywords: Khortha, Swadesh Word List, Lesser-Known Language, Indo-Aryan.

1. About Khortha

Khortha is also known by alternative names such as Khotta, Khontai, Kortha, Khatta, and Khattal. It is spoken primarily in the Indian states of Jharkhand, Bihar, and parts of West Bengal. The name Khortha is believed to come from Kulthi (horse gram), the ancient staple crop of the area, or from Danghal Jamin, meaning "high land" near river banks.

Khortha is classified under the Eastern Magahi dialect of the Bihari subgroup within the Indo-Aryan family. According to the 2011 Census, Khortha has approximately 8.04 million native speakers and is written in the Devanagari script. Linguistic perspectives on Khortha vary; while some scholars, such as Prasad and Shastri (1958), consider it a

variant of Magahi, others, including Das (2013), categorize it as a form of Angika spoken in western Jharkhand and Bihar. Grierson (1903), in his *Linguistic Survey of India*, characterized Khortha as a modified form of Eastern Magahi, spoken in Western Malda, with Maithili bordering it to the north and west, and Bengali to the east and south. The language shares significant similarities with Magahi, but it also has influences from neighbouring languages like Maithili, Bengali, and Hindi, particularly due to geographical proximity and social interactions. These influences are evident in vocabulary and certain grammatical structures. Despite its substantial number of speakers, Khortha is facing a decline in use, as Hindi predominates in the region, reducing its presence in formal and public contexts.

2. About the Informant

The Khortha informant from whom the data has been collected is a 40-year-old male, who belongs to the Deoghar district of Jharkhand. His educational background includes a BA, MA, and an M.Phil in Russian. He speaks Khortha, Hindi, Gujarati, Sanskrit, Italian, English, and Russian with varied levels of proficiency.

3. SWL Data

Sl. No.	English	Hindi	Khortha	IPA
1	I	मैं	हाम	hɑ:m
2	you (singular)	तू, तूम	तों	tõ:
3	he	वह	ऊ ई वें यें	u: i: wẽ jẽ
4	we	हम	हमनीह्	həmni:h
5	you (plural)	तूम, आप	तोरनीह्	to:rni:h
6	they	वे, ये	ऊ सब ई सब वें सब यें सब	u: səb i: səb wẽ səb jẽ səb
7	this	यह	ई इटा	i: iʈɑ:
8	that	वह	ऊ उटा	u ʊʈɑ:
9	here	यहाँ	यहाँ इ ठीना	jəhɑ̃: ɪʈʰi:nɑ:

10	there	वहाँ	वहाँ उ ठीना	wəhã: ʊʈʰi:nɑ:
11	Who	कौन	के	ke
12	what	क्या	कि	kɪ
13	where	कहाँ	कहाँ	kɑhã:
14	when	कब	कखनि	kɑkʰnɪ
15	how	कैसा	किरे	kɪre
16	not	नहीं	ने	ne
17	all	सब, सारा	साब	sɑ:b
18	many	बहुत	ढेरी	ɖʰe:ri
19	some	कुछ	थोड़े	tʰo:ɽe
20	few	थोड़ा	कोम एकट्टी तानी	ko:m ɛkʈʈɪ ʈɑ:ni:
21	other	अन्य, दूसरा	दोसर	ɖosər
22	one	एक	एक	ek
23	two	दो	दू	du:
24	three	तीन	तीन	ti:n
25	four	चार	चैर	tʃɛ:r
26	five	पाँच	पाँच	pɑ̃:tʃə
27	big	बड़ा	बोड़	bo:ɽ
28	long	लम्बा	लाम्ब	lɑ:mbə
30	thick	गाढ़ा, मोटा	मोट	mo:ʈə
31	heavy	भारी	भारी	bʰɑ:ri:
32	small	छोटा	चोट	tʃoʈə
33	short	छोटा, नाटा	नाटा	nɑ:ʈɑ:
34	narrow	तंग	साँकर	sã:kər
35	thin	पतला	पातर	pɑ:tar
36	woman	औरत	जेनी	dʒeni:
37	man (adult male)	आदमी	मरद	marad

38	man (human being)	इंसान, व्यक्ति	लोग	lo:g
39	child	बच्चा	बूतरु	butəru
40	wife	पत्नी	कनीये	kəni:je
41	husband	पति	दुल्हा	dʊlha:
42	mother	माता	मैं मेया	mɛ̃: me:jɑ:
43	father	पिता	बाप बाबू	bɑ:p bɑ:bu
44	animal	जानवर	जानवर	dʒɑ:nwar
45	fish	मछली	मच्छ:	mʌʧʰə
46	bird	चिड़िया	चिरियां	ʧirjɑ̃:
47	dog	कुत्ता	कुकुर	kʊkʊr
48	louse	जूँ	ढील	ɖʰil
49	snake	सांप	साँप	sɑ̃:p
50	worm	कीड़ा	पिल्लू कीड़ा	pɪllu ki:ɽɑ:
51	tree	पेड़	गाछ	gɑ:ʧʰ
52	forest	जंगल	बोन	bo:n
53	stick	डण्डा	छेंड़ी	ʧʰẽ:ɽi
54	fruit	फल	फल	pʰəl
55	seed	बीज	बीचा	bi:ʧɑ:
56	leaf	पत्ता	पतै	pətɛ
57	root	जड़	जर	dʒər
58	bark (of a tree)	छाल	बोकला	boklɑ:
59	flower	फूल	फूल	pʰu:l
60	grass	घास	घास	gʰɑ̃:s
61	rope	रस्सी	बरती नेढा	bərʈi ne:ɖʰɑ:
62	skin	त्वचा, चमड़ी	चमड़ी	ʧəməɽi:
63	meat	माँस	मांस	mɑ:ns
64	blood	खून	खून	kʰu:n

65	bone	हड्डी	हाड़	haːɽ
66	fat (noun) grease	चरबी	चरबी	tʃərbi
67	egg	अंडा	डीम्म	ɖiːmm
68	horn	सींग	सिंघा	sɪŋgʰaː
69	tail	पूंछ	नेढ़ी	neːɖʰiː
70	feather	पंख, पर	पैंख	pɛ̃kʰ
71	hair	बाल	केस चूल	kes ʧul
72	head	सिर	मुड़ी माथा	muɖiː maːʈʰaː
73	ear	कान	कना	kənaː
74	eye	आंख	येंख	jɛ̃kʰ
75	nose	नाक	नका	nəkaː
76	mouth	मुख मुँह	मुहाँ	mʊhãː
77	tooth	दांत	दँता	ɖənʈaː
78	tongue (organ)	जीभ	जीहा	dʒihaː
79	fingernail	नाखून	नोंहा	nõːhaː
80	foot	पैर	गोड़	goːɽ
81	leg	टाँग	टेंगा	tɛŋgɑ
82	knee	घुटना	ठेहुना	ʈʰeːhʊnɑ
83	hand	हाथ	हथा	həːtʰaː
84	wing	पंख	डेना	ɖenaː
85	belly	पेट	पेटा भोटी	peʈaː bʰoʈiː
86	guts	अंतड़ी	अतड़ी	əʈəɽi
87	neck	गरदन	घेचा	gʰeːtʃaː
88	back	पीठ	पीठिया	piʈʰɪjaː
89	breast	छाती सीना	छाती	tʃʰaːtiː
90	heart	दिल	करेजा	kəredʒaː
91	liver	यकृत	करजी	kərdʒiː
92	to drink	पीना	पियैल	pijɛl

93	to eat	खाना	खैयल	kʰɛjəl
94	to bite	काटना	काटेल	kɑːʈel
95	to suck	चूसना	चुसेल	tʃʊsel
96	to spit	थूकना	थूकेल	thukel
97	to vomit	उल्टी करना	उल्टी करेल	ulʈiː kərel
98	to blow	फूँक मारना	फुकेल	pʰukel
99	to breathe	साँस लेना	सांस लियेल	sə̃ːs lɪjel
100	to laugh	हँसना	हौसेल	hɔsel

4. Word Forms and Borrowings

Some words in Khortha are similar to Hindi but have unique phonological forms, indicating potential internal evolution or external influence from languages like Maithili, Bengali, and other regional tongues. The words /mʌʧʰə/ **(**fish) and /ʧul/ (hair) in Khortha exhibit Bengali influence, primarily in pronunciation and lexical choice. The words /dʒɑːnwɑr/ (animal), /sə̃ːp/ (snake), and /pʰuːl/ (flower) in Khortha exhibit strong influence from Hindi.

References

Das, P. K. 2013. Ergativity in Khortha: an enigmatic phenomenon, The *Journal of South Asian Studies*, HUFS, South Korea, 18, 3, 1.

Grierson, G. A. 2005. *A Linguistic Survey of India*, vol. I–XI. Calcutta: Government of India, Central Publication Branch.

Kumar, A., and Ray, B. 2023. Analyzing Types of Affixes in Khortha. *Interdisciplinary Journal of Linguistics*, 16, 129-138.

Prasad, N. and Shastri, S. 1958. *Linguistic Survey of the Sadar Subdivision of Manbhum and Dalbhum (Singhbhum)*, Patna: Bihar Rastrabhasha Parishad.

Online References:

Hammarström, H., Forkel, R., Haspelmath, M. and Bank, S. 2024. *Glottolog 5.1.* (Available online at http://glottolog.org, Accessed on 19-03-2025.) https://doi.org/10.5281/zenodo.14006617

Linguistic Survey of India - Jharkhand. 2011. Census, Language Division, Govt. of India. (Available online at https://language.census.gov.inmapdatashow LSI Reports)

वाक् मंथन
ISSN: 2426-2149

Personal Pronouns in Sütsa

Mughalivi
Centre for Naga Tribal Language Studies (CNTLS)
Nagaland University, India
Mughalivimughalivi@gmail.com

Abstract

This paper emphasizes on the personal pronouns in Sütsa (ISO 639-3-nsm), a language spoken by the Sümi tribe of Nagaland, located in North-east India. It belongs to the Tibeto-Burman language family under Angami-Pochury group (Burling, 2003). Sreedhar (1980: 81) references that nouns are incapable of showing any opposition in person and ordinarily show only a two-way opposition in number highlighting the capability of the class pronoun, showing a three -way opposition in both person and number. Based on the study, Sütsa distinguishes three person categories; 1st person, 2nd person, 3rd person: and exhibits three numbers; singular, dual, plural. Sütsa has no gender distinction in the first and the second person pronoun but for third person it distinctly presents feminine gender /li/ along with the gender neutral marker /pa/ 's/he'.

Keywords: Sütsa, Sümi, Naga, Tibeto-Burman, Personal Pronoun**.**

1. Introduction

Sütsa is a Tibeto-Burman language spoken in Nagaland, northeast India by the Sümi/Sema tribe. The GoN on 30th September 1995 announced the official use of 'Sümi' instead of 'Sema', for all official and non-official correspondences for this tribe. Sümis mainly dominates the district of Zunheboto[1] which is located in the heart of Nagaland. They also at large occupy the district of Dimapur, thereafter in other districts of Nagaland viz; Chümoukedima, Niuland, Kohima, Mokokchung, Tuensang, Kiphire. There are also seven Sümi villages in Tinsukia district, Assam. As per 2011 census Sütsa speakers have a population of 140.757 but generally Sümi is estimated to be 300,000.

Burling (2003: 184) grouped Sümi[2] under the Angami-Pochury group along with Rengma, Rengma.N, Chokri, Kheza, and Mao in his genetic classification of TB languages of Northeast India. Sütsa exhibits dialectal variation from village to village, area to area Grierson (1903/1928/1967; Vol. 3, Part 2) mentioned Simi and Zhimomi; Hutton (1921/1968) Zümomi (Zhimomi) and Lazemi (Lazami); Sreedhar (1976) mentioned four dialects of Sümi viz., Western Sema, Eastern Sema, the Chizolimi dialect and the Central dialect. Teo (2014) points out that the dialect spoken at Zunheboto town and Satakha area corresponds to Sreedhar's 'Central' dialect which is used as a standard variety. Sütsa dialects spoken in Aphuyemi area[3], especially the Lazami and the Ighanumi varieties are somewhat unintelligible to the other Sütsa speakers, but its speakers can comprehend the common variety though they may not be fluent in terms of speaking it, Mughalivi (2021: 27). Sümis are bilinguals with fluency in English and

[1] Zunheboto was known as Sümi land/Sümi country before its coinage.
[2] Mentioned as Simi by Burling (2003).
[3] Southern part of Zunheboto district.

Nagamese, the lingua franca of Nagaland. Hindi is also fairly understood and spoken. Sütsa along with Tenyidie and Ao-Chungli are the only three languages that have a daily newspaper [4] among the 17 recognised languages of Nagaland.

2. Methodology

This study is based on the standard variety of Sütsa spoken at Satakha and Zunheboto town area and the data was collected from both primary and secondary sources. Former by means of elicitation, sentence collections, personal interviews and narratives; latter through articles, journals, thesis, and subject related books for contextual relevance.

3. Literature Review

Linguistically not much exhaustive work in Sütsa has been done though some account of it were undertaken in brief viz., Marrison (1967) and Matisoff (1996) on the phonemic inventory of Sütsa, Hutton (1921) a brief glimpse of the language structure in his book "The Sema Naga", Sreedhar (1980) laid down the basic foundation of Sütsa grammar linguistically through his work "Sema Grammar". Ezung and Walling (2021) presented a comparative study on Naga languages thereby segregating them into Group A and B by observing the positions of negation, wh-questions, adverb and tense. The most extensive work on Sütsa relating to topic specificity has been done by Teo (2014-2019) who has worked exhaustively on the sounds and case marking system of Sütsa.

The comparative data of Hutton (1921: 276), Sreedhar (1980: 81-82) and Teo (2019: 60) on Personal Pronouns in Sütsa can be seen in (§Table 1). The data for the present study mostly corresponds to the data provided by Teo.

		SG	DU	PL
1st	H	ni 'I'	ikuzho	niunko 'we, us'
	S	ni	ikuzo	niṅù
	T	ni (or i)	ikujo	Ningu
2nd	H	no 'you'	okuzho	nonko 'you'
	S	No	okuzo	noṅù
	T	No	okujo	Nongu
3rd	H	pa 's/he'	pamho	pananko 'them, they' pananuko
	S	pa 's/he' li 'she' hi 'it' (-human)	pama	panóṅù
	T	Pa li	küma pama paküma	Panongu panoqo ?linongu

Table 1. Personal Pronouns by Hutton, Sreedhar and Teo

[4] Sütsa newspaper called Sümi Zümülhü, Tenyidie called Capi. Ao's have two dailies called Tir yimyim and Ao Milen.

4. Pronouns in Sütsa

According to Payne (1997: 43) "Pronouns are free forms (as opposed to affixes) that function alone to fill the position of a noun phrase in a clause. They normally have all the distributional properties of a noun phrases". Pronouns in Sütsa can be discussed under; personal pronouns, possessive pronouns, reflexive pronouns, interrogative pronouns, demonstrative pronouns, and indefinite pronouns, but this paper solely focuses on the personal pronouns in Sütsa.

4.1. Personal Pronouns in Sütsa

Personal pronouns talk about the three types of speech participants. First and second person are collectively regarded as speech act participants i.e., speaker and hearer, and the third person is generally referred to any non-speech act participant, Payne (1997: 44). Sütsa exhibits three-way oppositions in both person viz; 1st, 2nd, 3rd and number viz; singular, dual, and plural. The presence of gender distinction in Sütsa is present only in the 3rd person category for feminine along with the neuter gender pronoun /-pa/. There is no distinct marker to indicate masculine as opposed to feminine. The first and second person does not identify gender.

"… among the substantives are some pronoun-forms which, by over-differentiation, do not serve as actors: me, us, him, whom" Bloomfield, (1935: 269). Sütsa distinctly recognizes subject and object pronouns having two sets of forms in the 1st and the 2nd pronoun singular (§Table 2).

	SG Subject	**SG Object**	**DU**	**PL**
1st	/nì/, /i/ 'I'	/i/ 'me'	ikudʒo / nìkudʒo	/nīŋu/ 'we, us'
2nd	/nò/ 'you'	/o/ 'you'	okudʒo / nòkudʒo	/nōŋu/ 'you'
3rd	/pa/ 's//he' /li/ 'she'		pama/pakɨma likɨma	/panóŋu/ 'them, they' -

Table 2. Personal Pronouns in Sütsa

There are diverse ways in which person markers are realized. Siewierska (2004) expresses that, with regard to morpho-phonological realization, they may appear as independent words, so called weak forms, clitics, affixes or even only covertly as zero forms. In Sütsa, the object form /-i/ 1st person and /-o/ 2nd person also function as pronominal prefixes on nouns to show possession i.e., [i- pu] '1SG- father = my father', [o- pu] '2SG- father = your father'. They also function for the roles of possessive pronouns [i-u] '1 SG. POSS', [o- u] '2SG.POSS' and reflexive pronouns [i- kutoʃi] '1SG- self = myself' and [o- kutoʃi] '2SG- self = yourself', and not the 1st and 2nd subject singular form /-nì/ and /-nò/.

4.1.1. Singular

The singular pronoun is unmarked in Sütsa and categorized into first, second and third person pronouns

i) First person singular

First person singular			
	Subject		**Object**
	/nì/ 'I'	/i/ 'I'	/i/ 'me'
Marker	/-je/ 'topic'	/-no/ 'agent/focus'	-

Table 3. First person singular in Sütsa.

As mentioned, the first- person singular has two forms i.e., /nì/ and /i/, where /i/ occurs in the position of both subject and object (§1,3). The subject arguments /nì/ and /i/ has different obligations in terms of case marking as shown in Table 3 wherein, the former cannot occur only with the topic marker /-je/ (§2) and the latter occur with the agent/focus marker /-no/ (§1).

First person subject pronoun /nì/ and /i/

1. i no a- le phe a- ni
   1SG AGT NRL- song sing PROG- PRES
   'I am singing a song'

2. nì je a- na ʧu piʧi a- ni
   1SG TOP NRL- food eat lazy PROG- PRES
   'I am feeling lazy to eat food'

First person object pronoun /i/

3. no no i kʰe -a
   2SG AGT 1SG wait -IMPRF
   'You were waiting for me'

ii) Second person singular

Second person singular		
	Subject	**Object**
Marker	/nò/ 'you'	/o/ 'you'
	/-no/ 'agent/focus' /-je/ 'topic'	-

Table 4. Second person singular in Sütsa

The second person singular in Sütsa has two distinct forms exhibiting an exclusive subject and object (§Table 4). /nò/ occurs in a subject position taking the case for an agent and the topic marker and /o/ only takes the position of an object.

Second person subject pronoun /nò/

4. nò je alo
   2SG TOP good
   'You are good'

5. nò no a- kiɣi mutu
   2SG FOC NRL- rope hold
   'You hold the rope'

Second person object pronoun /o/

6. o ʃikimitʰi
   2SG thank
   'Thank you'

iii) Third person singular

Sütsa exhibits gender distinction only in the third person category for feminine /li/[5] (§7, 8) and there is no specific gender marker for masculine. The third person /pa/ occurs as a neutral referent irrespective of gender i.e., s/he (§9, 10). Both the person marker /pa/ and /li/ can take the subject case /-no/ and the marker /-je/. The feminine marker /li/ shows a two - way opposition in number i.e., singular and dual unlike the third person neutral /pa/ that shows a three -way opposition in number viz., singular, dual, plural.

7. li je i- fo
   3SG.F TOP 1SG.POSS- sister
   'She is my sister'

8. nì je li sasü ŋo -a
   1SG TOP 3SG.F COM stay -IMPRF
   'I was with her'

9. pa je amikuʧo timi
   3SG.F TOP honest person
   'S/he is an honest person'

10. nò no pa tsɨ
    2SG FOC 3SG give
    'You give him'

4.1.2. Dual

The first and the second person dual in Sütsa is marked by /-kuʤo/ and the third person is marked by /-ma/, or /-kɨma/ (§Table 2). /-kuʤo/ can be suffixed to both the subject-object forms of first and the second person and use them interchangeably.

11. i- kuʤo je kɨ- pʰi -mi
    2SG- DU TOP NZP- read -person
    'We are students'

12. o- kuʤo a- na ʧu -ve
    2SG- DU NRL- food eat -VM
    'Both of you eat food'

13. pa -kɨma kɨda -ve
    3SG -DU wake -VM
    'Wake them up'

14. li -kɨma ku
    3SG.F -DU call
    'Call them'

[5] During the study it was observed that the usage of /li/ was used more in a formal discourse and written forms. But generally its speakers were more inclined towards the use of the gender neutral /pa/ on most occasions.

4.1.3. Plural

The personal pronoun of plural in Sütsa is marked by /-ŋu/ in the entire person category i.e., first person, second person and third person (§Table 2).

15. nī-ŋu je mɨɣa
    1PL TOP lucky
    'We are lucky.'

16. nō -ŋu je mɨɣa
    2PL TOP lucky
    'You are lucky

17. panó -ŋu je mɨɣa
    3PL TOP lucky
    'They are lucky'

5. Conclusion

This paper has tried to generalize the account of the personal pronouns in Sütsa which showed a three- way contrast in person and number; singular, dual and plural. Wherein, the first and the second person have a distinct two sets of forms segregating the subject and the object/predicate. Gender distinction was present only in the third person pronoun for the feminine category along with the neuter gender /pa/ and no distinction on masculine was found.

Abbreviations:

1SG - first person singular
2SG - second person singular
3SG - third person singular
1DU - first person dual
2DU - second person dual
3DU - third person dual
1PL - first person plural
2 PL - second person plural
3PL - third person plural
AGT - agentive
TOP - topic
VM – verb maker
FUT- future
IMPRF- imperfective
NRL- non-relational
NZP - nominalizing prefix
POSS- possessive
PROG – progressive
PRES – present

References:

Bloomfield, Leonard. 1933. *Language*. Great Britain: The Compton Printing Works.

Ezung, Mimi Kevichüsa &Walling, Pangersenla. 2021. *Typology of Naga Languages*: A Pilot Study

Hutton, John, H. 1921. *The Sema Naga*. Macmillan and Co. Limited, London.

Payne, Thomas, E. 1997. *Describing Morphosyntax*. A Guide for Field Linguists: Cambridge university press

Siewierska, Anna. 2004. *Person*. New York: Cambridge university press.

Sreedhar, Mangadan, V. 1980. *A Sema grammar*. Mysore: Central Institute of Indian Languages.

Teo, Amos. 2019. *Investigating Differential Case Marking in Sümi, A Language of Nagaland, Using Language Documentation and Experimental*. (PhD) University of Oregon.

Linguistic conflicts and identity – How language shapes ethnic identity in Manipur, India

Dr. B. S. Aribam
Assistant Professor,
Department of Anthropology,
Dr. Harisingh Gour Vishwavidyalaya (A Central University),
Sagar, Madhya Pradesh
bsaribam14@gmail.com

Abstract

Manipur, a multi-ethnic and multilingual state in Northeast India, is home to diverse linguistic communities, including the Meitei, Naga, and Kuki-Chin groups. Language serves as a crucial marker of ethnic identity, influencing cultural practices, social structures, and political affiliations. This research explores how linguistic diversity both enriches and complicates inter-ethnic relations in Manipur, contributing to identity formation, political mobilization, and conflict. Historically, colonial policies, post-independence state reorganization, and socio-political movements have shaped the linguistic landscape, reinforcing ethnic boundaries. The study examines the role of language in contemporary ethnic tensions, including disputes over script, medium of education, and recognition of minority languages. It also highlights how government policies, globalization, and modern media impact linguistic identity, with social media and global cultural influences reshaping language use among younger generations. By analyzing recent conflicts, linguistic rights movements, and efforts for reconciliation, this paper argues that inclusive language policies and intercultural dialogue are essential for fostering linguistic harmony in Manipur. The study draws on historical records, linguistic research, media analysis, and policy studies to provide a comprehensive understanding of the intersection between language and ethnic identity in the region.
Keywords: *Linguistic Identity, Ethnic Conflict, Multilingualism, Language Policy, Cultural Identity.*

1. Introduction

Manipur, a northeastern state of India, is home to a diverse array of ethnic communities, each with its own distinct linguistic and cultural identity. The state is predominantly inhabited by the Meitei, Naga, and Kuki-Chin ethnic groups, who speak languages belonging to the Tibeto-Burman family, alongside smaller Indo-Aryan and Austroasiatic linguistic communities (Singh, 2021). While Manipuri (Meitei) serves as the official language of the state, a multitude of indigenous languages thrive, reflecting the deep-rooted ethnic diversity of the region.

Language in Manipur is more than just a means of communication; it is a significant marker of ethnic identity. The linguistic distinctions between the valley-dwelling Meitei majority and the hill-based Naga and Kuki-Chin communities have historically shaped social and political relationships. These linguistic differences reinforce group identities, influencing cultural practices, educational policies, and political affiliations (Bhattacharyya, 2019). Furthermore, the revival of the Meitei Mayek script, alongside the push for English-medium education among tribal communities, illustrates the broader struggle over linguistic and cultural representation.

In recent years, ethnic and linguistic tensions have escalated into violent conflicts, particularly between the Meitei and Kuki-Chin communities. Disputes over land rights, political representation, and

affirmative action policies have been compounded by linguistic divides, with tribal groups resisting what they perceive as cultural and linguistic hegemony imposed by the Meitei-dominated administration (Sharma, 2023). These conflicts underscore the role of language as both a unifying force and a divisive element in Manipur's socio-political landscape.

This paper argues that linguistic differences play a crucial role in shaping and reinforcing ethnic identities in Manipur, contributing to both cultural richness and conflict. While language serves as a means of preserving heritage and fostering community cohesion, it also acts as a source of division, influencing intergroup relations and political dynamics. Understanding the interplay between language and ethnicity is essential to addressing the region's ongoing challenges and fostering inclusive policies that accommodate its rich linguistic diversity.

2. Area and People

Manipur is located in the northeastern region of India, sharing borders with Nagaland to the north, Mizoram to the south, Assam to the west, and Myanmar to the east. Covering an area of approximately 22,327 square kilometers, the state is geographically divided into the central Imphal Valley and the surrounding hilly regions (Government of Manipur, 2022). The valley, which constitutes about 10% of the state's land area, is home to the Meitei majority, while the hill districts are predominantly inhabited by various Naga and Kuki-Chin tribes. As per the 2011 Census of India, Manipur has a population of approximately 2.8 million, with Meitei forming about 53%, Nagas around 24%, and Kuki-Chin communities comprising roughly 16% of the population (Census of India, 2011).

The three major ethnic groups of Manipur are 1. Meitei, the dominant ethnic group residing in the Imphal Valley, speaking Manipuri (Meitei), which is also the official language of the state. Historically, they have played a central role in Manipur's political and cultural development (Singh, 2021). 2. Naga, a collection of tribal groups residing primarily in the hill districts. They speak various Naga languages and maintain strong cultural ties with Nagaland. Prominent subgroups include the Tangkhul, Mao, and Poumai Naga communities (Bhattacharyya, 2019). 3. Kuki-Chin, this group comprises several tribes such as Thadou, Paite, Hmar, and Zou, primarily inhabiting the hill districts. They have distinct linguistic and cultural identities and share historical ties with Kuki-Chin communities in Mizoram and Myanmar (Sharma, 2023).

3. Historical Background

Pre-colonial Era*:* The Meitei kingdom, historically known as Kangleipak, was a dominant power in the region, with a rich cultural and linguistic heritage. It maintained trade and political relations with neighboring Ahom, Burmese, and tribal communities (Parratt, 2005). Linguistically, the region was already diverse, with various Naga and Kuki-Chin languages coexisting alongside Manipuri. During this period, the Meitei Mayek script was developed and used for writing the Manipuri language, playing a crucial role in the preservation of Meitei identity and literature (Singh, 2008).

Colonial Influence*:* With the British annexation of Manipur in 1891, colonial policies reshaped the region's linguistic and ethnic landscape. The British administration introduced English and Bengali as administrative languages, sidelining indigenous scripts like Meitei Mayek (Kabui, 1991). The colonial practice of categorizing communities along ethnic lines reinforced divisions between the Meitei, Naga, and Kuki-Chin groups. Missionary-led education also played a role in linguistic shifts, with many hill communities adopting Roman scripts for their languages (Elwin, 1959).

Post-Independence: Manipur merged with India in 1949, a transition that significantly affected its ethnic groups. The subsequent reorganization of states and implementation of affirmative action policies heightened ethnic nationalism. The Meitei community sought greater recognition for their language and cultural identity, leading to the official adoption of Manipuri as a scheduled language in 1992 (Bhattacharyya, 2019). Meanwhile, Naga and Kuki-Chin groups intensified their demands for autonomy, contributing to ongoing ethnic and linguistic tensions (Sharma, 2023).

4. How language shapes one's identity?

One of the closest theoretical frameworks to explain how language shapes linguistic identity in Manipur is Ethnolinguistic Identity Theory (ELIT), developed by Howard Giles (1977). This theory posits that language is a crucial marker of ethnic identity and is used to maintain group distinctiveness. According to ELIT, ethnic groups use language to establish and reinforce in-group solidarity while differentiating themselves from others. For instance: Meitei identity is strongly tied to the Manipuri (Meiteilon) language and Meitei Mayek script. The promotion of Meitei Mayek in schools and its recognition as an official language in 1992 strengthened Meitei linguistic nationalism (Singh, 2021). Naga and Kuki-Chin tribes, speaking distinct Tibeto-Burman languages, resist Meitei linguistic dominance by advocating for their own languages in education and administration (Bhattacharyya, 2019). Social Identity Theory (Tajfel & Turner, 1979) complements ELIT by explaining how language fuels ethnic conflicts. The 2023 ethnic clashes between the Meitei and Kuki-Chin groups were intensified by language-based exclusion and competition over linguistic recognition (Sharma, 2023). Besides, modernization theories suggest that globalization leads to language shift, particularly among younger generations. The rise of English as a medium of education and the popularity of Korean pop culture has influenced youth linguistic preferences, sometimes at the cost of indigenous languages (Sharma, 2023). Besides, the imposition of Hindi as a compulsory paper to learn in schools also add to the linguistic conflict.

In most parts of India, people typically learn one or two languages—usually their native or regional language and Hindi or English for broader communication. For instance, a person in Maharashtra might speak Marathi at home and learn Hindi and English in school, while someone in Tamil Nadu might primarily use Tamil and English. However, in Manipur, linguistic diversity and socio-political realities compel individuals to learn at least four languages to navigate the social structure effectively. A Meitei individual, for example, speaks Manipuri (Meiteilon**)** as their mother tongue but must also learn English for education and administrative purposes, Hindi for communication with non-local communities and central government institutions, and often a local tribal language—especially in workplaces, markets, or inter-ethnic social settings. Similarly, tribal communities like the Tangkhul Naga, Thadou Kuki, or Paite grow up speaking their native tongue but must learn Manipuri to interact with the majority population, English for academic and professional mobility, and Hindi for national communication. This multilingual necessity arises from Manipur's ethnic diversity, historical conflicts, and socio-economic interactions, where different linguistic groups coexist but maintain distinct identities. The linguistic complexity is further compounded by education policies—many tribal schools adopt English or Roman script for their languages, while Meitei-medium schools emphasize Manipuri and Meitei Mayek. This multilingual reality creates both advantages, such as cognitive flexibility and broader communication skills, and challenges, such as identity struggles and language endangerment among smaller communities. Unlike other Indian states where bilingualism is often sufficient, Manipur's multilingual demands reflect its unique ethnic mosaic, historical influences, and socio-political landscape, making language a central force in shaping identity and social mobility.

5. Linguistic Diversity and Ethnic Identity in Manipur

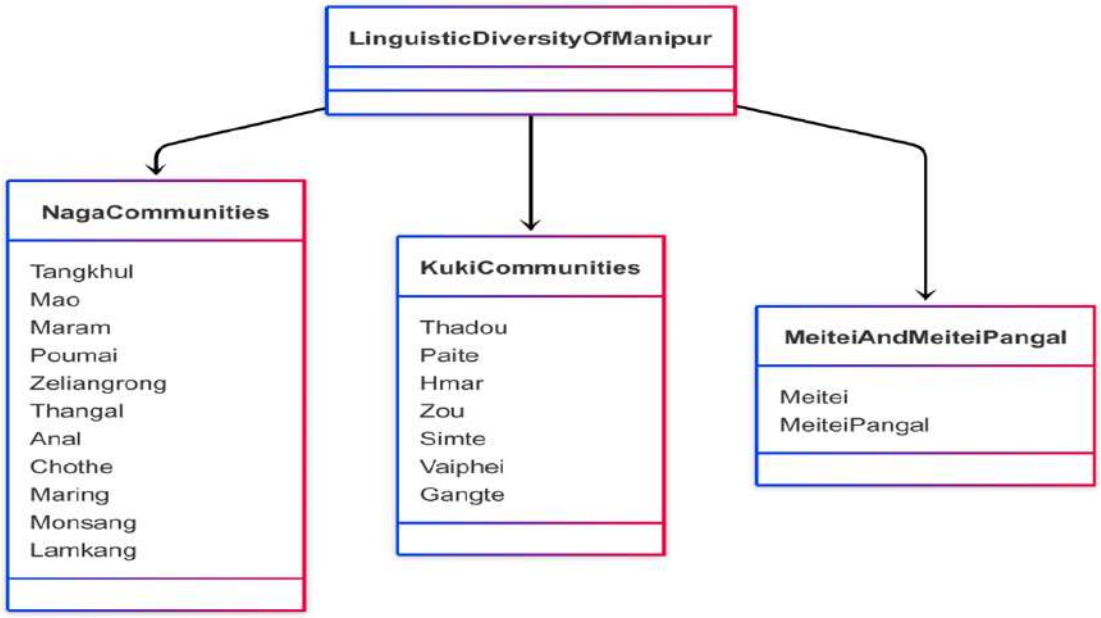


Figure1: Showing Linguistic Diversity of Manipur

6. Meitei Language and Identity

The Meitei language (Manipuri) is a central component of Meitei identity. Recognized as one of India's scheduled languages, it serves as the lingua franca of the state and is integral to the community's cultural heritage. Efforts to preserve and promote the Meitei language have included the revival of the Meitei Mayek script, which was replaced by the Bengali script during colonial rule but has since been reintroduced in educational curricula (Singh, 2008). Additionally, cultural initiatives such as Manipuri literature, dance, and theatre contribute to the reinforcement of Meitei identity through language (Bhattacharyya, 2019).

7. Naga Languages and Identities

The Naga communities in Manipur speak multiple languages, including Tangkhul, Mao, and Poumai, among others. These languages are distinct yet share similarities with Naga languages spoken in neighboring Nagaland. The diversity within the Naga linguistic landscape reinforces subgroup identities, as each Naga tribe takes pride in its unique linguistic heritage (Sharma, 2023). However, this diversity also presents challenges in terms of communication and political unity, leading to efforts to develop a standardized written form and increase the use of Nagamese as a lingua franca (Elwin, 1959).

8. Kuki-Chin Languages and Identities

The Kuki-Chin communities in Manipur speak various dialects of the Kuki-Chin linguistic family, including Thadou, Paite, and Hmar. While these languages share common roots, they have evolved into distinct variations, reinforcing tribal identities (Kabui, 1991). The adoption of the Roman script for most Kuki-Chin languages, largely due to missionary influence, has facilitated literacy and education but has also led to differences in orthographic traditions among different groups. Language plays a significant role in shaping the ethnic consciousness of Kuki-Chin groups, who often use it as a tool for asserting their distinct identity within the multi-ethnic landscape of Manipur (Sharma, 2023).

9. Analysis of Recent Events

Manipur has witnessed a series of ethnic conflicts in recent years, with language playing a crucial role in reinforcing ethnic boundaries. The 2023 violence between the Meitei and Kuki communities was driven by underlying political, economic, and linguistic factors. The demand for Scheduled Tribe (ST) status by the Meitei community exacerbated tensions, as Kuki-Chin and Naga groups perceived this as an encroachment on their constitutional privileges (Sharma, 2023). Language has been weaponized in these conflicts, with both sides using linguistic differences to assert separate identities and justify territorial claims.

The linguistic divide is evident in education and administration. Many tribal groups advocate for English as a medium of instruction, fearing Meitei linguistic dominance. Meanwhile, the Meitei push for greater recognition of the Meitei Mayek script is perceived as an attempt to establish cultural hegemony (Singh, 2021). These dynamic fuels resentment and hinders inter-ethnic integration.

10. Media and Language

Media outlets have played a significant role in shaping perceptions of the conflict. Local newspapers, digital platforms, and social media have often portrayed linguistic differences in a way that reinforces ethnic divisions. Pro-Meitei media emphasize the historical and cultural primacy of the Manipuri language, while tribal media platforms highlight the marginalization of their languages and demand greater representation (Bhattacharyya, 2019).

Misinformation and propaganda on social media have further deepened linguistic and ethnic divides. Fake news regarding language policies, educational reforms, and administrative decisions has contributed to misunderstandings and hostility between communities. The portrayal of linguistic issues in the media reflects broader political alignments, with different factions using language as a tool for ethnic mobilization and resistance (Sharma, 2023).

Addressing these conflicts requires inclusive language policies that recognize and respect Manipur's linguistic diversity. Ensuring equitable representation of all languages in education, administration, and media can help mitigate tensions and promote social cohesion in the state.

11. Conclusion

The linguistic landscape of Manipur plays a crucial role in shaping ethnic identity and fostering both unity and division among its diverse communities. The historical evolution of the Meitei, Naga, and Kuki-Chin languages has reinforced ethnic boundaries, often leading to social and political conflicts. The Meitei language, with its long literary tradition and recent revival of the Meitei Mayek script, has been a symbol of cultural pride for the Meitei people, while the Naga and Kuki-Chin communities continue to advocate for the recognition and preservation of their own languages.

Recent ethnic conflicts in Manipur highlight how linguistic distinctions are deeply intertwined with political and social issues. The demand for Scheduled Tribe status by the Meitei community, language-based educational policies, and competition over linguistic dominance in administration and media have all contributed to rising tensions. The role of media, especially social media, has further exacerbated these divisions by amplifying narratives that reinforce linguistic and ethnic separations.

Moving forward, there is an urgent need for policies that promote linguistic inclusivity and respect for Manipur's diverse languages. Educational reforms that incorporate multilingual curricula, the protection of minority languages, and the encouragement of cultural exchanges between communities can help bridge ethnic divides. Furthermore, responsible media representation and community-driven peace initiatives are essential in fostering mutual understanding and reconciliation. Ultimately,

language should serve as a bridge for cultural connection rather than a tool for division, ensuring a more harmonious future for Manipur.

References

Bhattacharyya, D. (2019). *Ethnic Identity and Linguistic Politics in Northeast India*. Routledge.

Census of India. (2011). "Demographic Profile of Manipur." Government of India.

Elwin, V. (1959). *The Tribal World of Verrier Elwin*. Oxford University Press.

Government of Manipur. (2022). *Statistical Handbook of Manipur*.

Kabui, G. (1991). *History of Manipur: Pre-Colonial and Colonial Periods*. National Publishing House.

Parratt, S. (2005). *The Pleasing of the Gods: Meitei Lai Haraoba*. Cambridge University Press.

Sharma, R. (2023). "Ethnic Conflicts and Linguistic Divides in Manipur: A Contemporary Analysis." *Journal of Northeast Studies*, 12(3), 45-67.

Singh, H. (2021). *The Linguistic Landscape of Manipur: History, Identity, and Politics*. Oxford University Press.

Singh, K.B. (2008). *The Evolution of Meitei Script and Literature*. Imphal Press.

A study on Indefinite Pronouns in Yimkhiung

Payia Maheo
Payiamaheomei@gmail.com
&
Dr. Imlienla Imchen
aien.imchen@nagalanduniversity.ac.in
Centre for Naga Tribal Language Studies
Nagaland University
Meriema

Abstract
This paper investigates the indefinite pronouns in Yimkhiung, a less documented language spoken in Northeast, India from a morphological perspective. Although, indefinite pronouns are classified as subtypes of pronouns, their structure and formations can vary significantly across different languages. This paper provides an overview on the structures, formation and types of indefinite pronouns, providing an insight into the pattern and usage.

Keywords: *Trans-Himalayan, Yimkhiung, Indefinite, Derived, Reduplication.*

1. Introduction

Yimkhiung (ISO 639-3: yim) is a Trans-Himalayan (Tibeto-Burman) spoken in the eastern part Nagaland, Northeast, India. Previously, they were generally known as Yimchunger/Yimchungrü. The Yimkhiung primarily inhabits Shamataor and Kiphire districts of Nagaland with a total population of 74 647[1]. The language consists of several varieties[2], with langa as the standard variety. Burling (2003) categorised Yimkhiung referred to as "Yimchunger" under the Ao group of the Eastern area. Linguistically, the language is largely under documented.

Indefinite pronouns, a subclass of peripheral pronouns[3], express indefinite reference (Haspel math, 1997:11) or an indefinite entity (Payne, 2011: 391)[4]. In Yimkhiung, indefinite pronouns are derived through morphological process. Semantically, they imply both positive and negative indefinites polarity. Based on specificity, they can indicate both specific and non- specific referent. The referent entities are distinguished on the bases of human, non-human and inanimate entities.

Haspel math (1997: 9) note that indefinite pronouns occur in series such as in English where the indefinite consist of *some*-series as in '*somebody*', '*somewhere*', '*something*', and *any*-series, '*anyone*', '*anything*', '*anywhere*'. In Yimkhiung, this phenomenon is marked in almost all the indefinite forms except in specific indefinite place category. There are two series markers *-kʰə*, and *-ʃe*, which are attached to the base indefinite pronouns. The *-kʰə* series is attached to specific indefinites of person and thing categories. In contrast, the *-ʃe* series occurs in both non-specific and negative indefinites.

[1] https://en.m.wikipedia.org/wiki/Yimkhiung_Naga.
[2] Chirr, Makury, Longphur, Phenunger, and Langa.
[3] This term is adopted from Bhattacharyya (2018).
[4] Glossary section

It is worth mentioning that in Yimkhiung the negative indefinite pronouns do not convey negative connotation when it occurs in isolation. It explicitly takes a negative marker to convey the sense of negation in a sentence. In this sense, negation in indefinite pronoun is realized through the negative marker (§ 4.3).

2. Methodology

The data for this study are primarily drawn from primary sources. The analysis is established based on interviews, elicited data, and fieldnotes. Additionally, secondary sources from articles and journals were also referred to for clarity of the study. The data variety is based on 'Langa' speakers of Shamator district.

3. Indefinite Pronouns: Formation types

Indefinite pronouns in Yimkhiung are formed through derivation. Haspel math (1997: 26) identified "two main types of derivational base from which indefinite pronouns are derived in the world's languages: (i) Interrogative Pronouns, and (ii) Generic ontological-category nouns". Similarly, in Yimkhiung, the derivational based for indefinite pronouns consists of interrogative-based indefinites and Generic-noun-based indefinites. Additionally, the base also includes reduplication and one-based indefinites.

3.1 Interrogative-based indefinites

According to Haspel math (2013) interrogative-based indefinites are "based on interrogative pronouns in the sense that they are overtly derived from them by the addition of a bound morpheme". In Yimkhiung, the negative indefinites pronouns of person and place categories are derived from the interrogative-based *ʃepə* 'who' and *kətʃʰiŋ* 'where'. This derivation occurs via suffixing *-ʃe*, resulting *ʃepəʃe* 'no-one' and *kətʃʰiŋʃe* 'nowhere'.

3.2 Generic-noun-based indefinites

An indefinite pronoun is considered generic-noun-based when its base is a noun. In Yimkhiung, generic-noun-based encompass of non-specific, specific indefinite and negative indefinite pronouns. The non-specific indefinite pronoun of the place category is derived from the reduced form of the generic noun *taŋpuŋ* 'place'. This is achieved by combining *kʰəɹə* to form *taŋkʰəɹə* 'somewhere'. Alternatively, specific indefinite and negative indefinite pronouns are derived from the generic noun *təmə* 'thing' by attaching *-kʰəʔ* and *-ʃe*. The attachment of the suffix *-kʰə* to *təmə* yields *təməkʰəʔ* 'something'. Similarly, the suffixation of *-ʃe* with *təmə* results in *təməʃe*, meaning 'nothing'.

3.3 Reduplication

Another derivative strategy for forming indefinite pronouns in Yimkhiung is reduplication[5]. Yimkhiung employs full reduplication, accompanied by the indefinite marker *-ʃe* to form indefinite pronouns.

As shown in **Table 1**, the non-specific indefinite pronoun of the person category *ʃepəʃepəʃe* 'someone' is derived by reduplicating the interrogative pronoun *ʃepə* 'who'. Correspondingly, the place category *kətʃiŋkətʃiŋʃe* 'somewhere' is derived via reduplicating the interrogative *kətʃʰiŋ* 'where', and the non-specific indefinite pronoun of the thing category *təmətəməʃe* 'something' is derived by reduplicating the generic noun *təmə* 'thing'. This reduplication strategy is restricted to non-specific indefinite pronouns.

Source	Reduplication	Gloss
ʃepə 'who'	ʃepəʃepəʃe	'someone'
kətʃiŋ'where'	kətʃiŋkətʃiŋʃe	'somewhere'
təmə'thing'	təmətəməʃe	'something'

Table 1: Indefinite pronouns formed through reduplication

[5] This feature is common cross-linguistically (Haspel math, 1997: 179).

3.4 One-based indefinite

A language exhibit one-based indefinite when an indefinite pronoun is formed using the numeral one with additional affixes. In Yimkhiung, this formation is restricted to specific indefinite *kʰəlaŋkʰə* 'someone' of the person category. This specific indefinite is formed by suffixing the indefinite marker *-kʰə* to the numeral *kʰəlaŋ* 'one', that functions as the base.

4. Types of Indefinite Pronouns

Yimkhiung exhibit three types of indefinite pronouns- specific indefinite, non-specific indefinite and negative indefinite pronouns encompassing person, place and thing categories. **Table 2** presents this classification. The following sub-sections provides the description of each type.

Types	Forms	Gloss	Ontological category
Spc. Ind. pronouns	kʰəlaŋkʰəʔ /kʰəɹəkʰəʔ	'someone/ somebody'	'person'
	taŋkʰəɹə	'somewhere'	'place'
	təməkʰəʔ	'something'	'thing'
Non-spc. pronouns	ʃepəʃepəʃe /ʃenəʃenəʃe	'anyone/anybody'	'person'
	kəʧiŋkəʧiŋʃe	'anywhere'	'place'
	təmətəməʃe /təməpəɹəʃe	'anything'	'thing'
Neg. Ind. pronouns	ʃeʃe/ʃenəʃe	no-one' / 'nobody'	'person'
	kəʧʰiŋʃe	'nowhere'	'place'
	təməʃe	'nothing'	'thing'

Table 2: Types of Indefinite Pronouns in Yimkhiung

4.1 Specific indefinite pronoun

Specific indefinite pronouns are the antonyms of negative indefinite pronouns. Thus, they serve as positive indefinite pronouns. In Yimkhiung, three forms of specific indefinite pronouns are attested:

kʰəlaŋkʰə/kʰəɹəkʰə 'someone'
kʰəɹəʧiŋ/pəpəʧʰiŋ 'somewhere', and
təməkʰə 'something'

The specific indefinite pronoun 'someone' expresses a specific yet an indefinite unknown person. It serves to indicate the existence of a person. This indefinite pronoun has two lexical variants- *kʰəlaŋkʰəʔ* and *kʰəɹəkʰəʔ*.

- The variant *kʰəlaŋkʰəʔ* 'someone' is used to refer to a singular, indefinite person.
- In contrast, the variant, *kʰəɹəkʰəʔ* 'someone' occurs in a neutral form and it is more frequently used.

In a transitive clause, both the variants *kʰəɹəkʰəʔ/kʰəlaŋkʰəʔ* can function as agent. This is illustrated in example (1), where the agent *kʰəɹəkʰəʔ* is marked by the agentive marker *-nə*.

- The specific indefinite pronoun, *taŋkʰərə* 'anywhere' indicates an indefinite location or an unknown place.
- The morpheme *təməkʰəʔ* 'something' expresses specific yet indefinite or unknown thing. It typically indicates the existence of the entity.

Examples of specific indefinite pronouns in sentential form are illustrated in 1-3.

1) kʰəɹəkʰə-nə nə ʃoɹu-pə aɹə-to
someone-AGT 2SG meet-PURP come-PERF
'Someone came to meet you'

2) kiuluŋ jo graund miki taŋkʰəɹə ʧʰiŋ jukinpuŋ aɲikʰə laʔ
village GEN ground near some where LOC shop small
EXIST.DECL
'Somewhere near the village ground there is a small shop.'

3) aɹake i-nə təməkʰə huʔ-pale
come 1SG-AGT something show-IRR
'Come, I will show you something'

4.2 Non-specific pronouns

Non-specific indefinite pronoun expresses the meaning of free choice (Haspel math, 1997:48) to the addressee. In Yimkhiung, non-specific pronouns are formed through full reduplication of the base and by attaching the indefinite marker *-ʃe*. Like specific indefinites, non-specific indefinites also have positive implications. In Yimkhiung, the following non-specific pronouns are exhibited:

ʃepəʃepəʃe 'anyone/anybody',
kəʧiŋkəʧiŋʃe 'anywhere', and
təmətəməʃe/təməpəɹəʃe 'anything'

- The non-specific indefinite pronoun *ʃepəʃepəʃe* 'anyone/anybody' expresses non-specific indefinite human entity. This lexeme is derived from the reduplicated form of the interrogative base *ʃepə* 'who'.

- The lexeme *kəʧiŋkəʧiŋʃe* 'anywhere', is a reduplicated form of the interrogative *kəʧiŋ* 'where'. It is used to express a non-specific indefinite place category.

- The non-specific pronoun *təmətəməʃe* indicate non-specific thing entity. It is formed by reduplicating the generic-noun based *təmə* 'thing' as 'anything'. It can be also expressed by *təməpəɹəʃe*.

Examples (4) to (6) illustrates the occurrence of non-specific pronouns in sentential form.

4) ʃepəʃepəʃe atʰa kʰuk laʔ
anyone speak can DECL
'Anyone can speak'

5) ɣənə kiuluŋ-ahiaŋɹə-po-ʧu kəʧʰinkəʧʰiŋʃe mə ɣə-ʧʰo
there village-guard-PLU-TOP anywhere NEG go-PROG
'Those village guards don't go anywhere'

6) isa-nə aɹə kʰiak təmətəməʃe mə-tʰa kʰuk
1PLU-AGT 3PLU DAT anything NEG-say can
'We can't say anything to them.'

Additionally, the indefinite pronouns can be expressed in its short form depending on the speaker and in informal discourse as *ʃepəʃe*, *kəʧiŋʃe*, and *təməʃe*.

4.3 Negative indefinite pronouns

Yimkhiung has negative indefinite pronouns, which are generally non-specific. As aforementioned, in contrast to English, in Yimkhiung the negative indefinite pronouns itself do not convey negation and

hence it obligatorily takes a negator to provide negative interpretation and to negate the existence of the specified category. There are three indefinite pronouns which functions as negative indefinites- *ʃeʃe*/*ʃenəʃe* 'nobody', *kəʧʰiŋʃe* 'nowhere', and *təməʃe* 'nothing'. These negative indefinite pronouns are interrogative and generic noun based, overtly marked by the indefinite marker *-ʃe* to indicate indefiniteness.

- The negative indefinite pronoun of the person category has two variant *ʃeʃe* and *ʃenəʃe*. *ʃeʃe* 'someone' typically occurs in non-verbal predicates and with intransitive verbs. As shown in example (7), *ʃeʃe* serves as the subject in a non-verbal predicate.

- Alternatively, *ʃenəʃe* commonly occurs in transitive verbs. This is demonstrated in example (8) where *ʃenəʃe* occur with the transitive verb *ʧim*.

- *kəʧʰiŋʃe* 'nowhere' is used to indicate non-existence of a place category.

- *təməʃe* 'nothing' is express for non-existence of a thing category.

In relative to verbal negation, Yimkhiung exhibit 'verbal negation plus negative indefinite'[6] type, in which the negative indefinites and the negator explicitly co-occur to convey negative indefinite meaning.
Structurally, the negative indefinite pronoun precedes the negated clause. Examples (7) to (10) illustrate this pattern.

7) ʃeʃe mə-tək
nobody NEG-perfect
'Nobody is perfect'.

8) ʃenəʃe i kʰiaʔ ʧim mə-ʧiʔ
nobody 1SG DAT food NEG-give
'Nobody gave me food (to eat)'.

9) aŋipuʔ kəʧʰiŋʃe aɣə-puŋ məla
3SG.M nowhere go-place NEG.EXIST
'He has nowhere to go'

10) haʔ-ʧʰiŋ təməʃe məla
basket-LOC nothing NEG.EXIST
'There is nothing in the hah (basket)'

5. Conclusions

The analysis presented in this paper reveals that indefinite pronouns in Yimkhiung are derived form, consisting of four based types. They are interrogative-based, generic-noun based, indefinite numerals and reduplication. The indefinite type consists of specific, non-specific and negative indefinite pronouns, where the specific and non-specific conveys positive indefinites. Yimkhiung indefinites exhibit *-kʰəʔ* and *-ʃe* series, where the former series is attested in specific indefinite and the latter is attested in non-specific and negative indefinite pronouns. The highlight of the negative indefinite pronouns is that, the negative indefinite overtly co-occurs with the negative marker. Thus, this preliminary paper provides morphological overview on Yimkhiung indefinite pronouns.

[6] This term is extracted from Haspel math (1997) classification of negative indefinites.

Abbreviations

1SG- First person singular
2SG- Second person singular
3SG- Third person singular
AGT- Agentive
DAT- Dative
DECL- Declarative
EXIST- Existential
IND- Indefinite
IRR- Irrealis
LOC- Locative
M- Male
NEG- Negation
PERF.ASP- Perfective Aspect
PLU- Plural
SPC-Specific

References

Burling, Robbing. 2003. The Tibeto-Burman languages of Northeastern India. In Graham Thurgood and Randy J. Polla (Ed.), *The Sino-Tibetan Languages*,169-191. London and New York: Routledge.
Haspel math, Martin.1997. Indefinite Pronouns. New York: OUP.
Haspel math, Martin. 2013. Indefinite Pronouns. *In Dryer, Mathew S. & Haspel math, Martin (edn.)*
https://doi.org/10.5281/zenodo.1395059. (18 March, 2025)
Bhattacharya, Renu, Pushpa. 2018. Indefinite Pronouns in Assamese. *Nepalese Linguistics*. 33(1). 8-15.
Koshy, Anish. 2009. Indefinite Pronouns in Pnar. *In Mon-Khmer Studies Journal*. 38. 41-56.
http://sealang.net/sala/mks/htm/38/KOSHYA nish.htm
Payne, E. Thomas. 2011. Understanding English Grammar. Cambridge University Press.

Problems and Prospects in Conceptualising Mother Tongue in the Context of Tiwa Tribe of Assam: with special reference to NEP-2020

Dr. Jyoti Rupa Deka
Department of English
Kampur College affiliated to Gauhati University, Assam
Email: jyotirupadeka@gmail.com

Abstract
National Education Policy 2020 highlights the importance of giving every citizen access to education. For this purpose, NEP outlines the necessity of imparting education in one's mother tongue. One of the major goals of NEP 2020 is to promote multilingualism and multiculturalism by bringing disadvantaged and marginalised groups to the limelight. To achieve this goal, NEP emphasises on Mother Tongue-Based Multilingual Education (MTB-MLE) by advocating for the use of the mother tongue as the medium of instruction, at least until Grade 5. NEP stresses on the need of incorporating the 'three-language formula' into the curriculum, under which a student will learn his/her mother tongue, a regional language, belonging to the 8th Schedule, and English.
Tiwa, one of the ethnic groups of North East India, is settled mainly in central Assam and Ri- bhoi district of Meghalaya. Tiwa language belong to Tibeto Burman language family. Present study is an attempt to contribute to the discourse on mother tongue education. Aim of this study is to discuss the importance of implementing NEP 2020's three language formula in Assam's Tiwa dominated area. Another objective of this work is to highlight the challenges and possibilities in implementing Tiwa as medium of instruction.

Keywords*: Mother tongue education, NEP 2020, Tiwa, Assamese, Tiwamese.*

1. Introduction

Assam is a land of many languages and dialects belonging to Indo- Aryan represented mainly by Assamese and Bangla, Boro -Garo-Koch-Deori, Kuki-Chin and the Tani groups of the Tibeto-Burman family, Siamese-Chinese represented by Tai- kadai language groups like Ahom, Khamti, Turung, Phake, Aiton and Khamyang, and some Dravidian and Munda languages represented by the tea tribes. These languages contribute to make Assam a very heterogeneous linguistic and cultural area with Assamese being the dominant language among non-tribal and tribal languages. Such co-existence of languages belonging to different language families leads to language contact that results language change. In this context, Assamese also serves *almost as a lingua franca* among various speech communities like Tiwa, Mising, Bodo, Garo etc. in the state.

In such a multilingual situation, Assam needs a carefully laid out implementation policy to meet the guidelines of NEP 2020 which will also equally cater to different linguistic groups. In this study, an attempt will be made to discuss the challenges and possibilities that lie in implementing NEP 2020 in Assam's Tiwa dominated area.

2. A brief note on Tiwa Tribe

The Tiwas, also known as Lalung is a tribal community living in Northeast India, mostly in Assam and Meghalaya. The primary source of livelihood of the Tiwa community is agriculture. They are also dependent in other traditional activities like hunting and fishing. The Tiwa tribe celebrates several colourful and vibrant festivals to mark their unique culture and tradition. Some of the significant festivals of this tribe are Sogra Phuja, Langkhon Phuja, Wanshuwa Kham, Yangli Phuja, Sogra Misawa,

and the three Bihu (Pisu) festivals. These festivals involve traditional rites and rituals, feast, and cultural activities.

Tiwa language belongs to the ancient Bodo race of Sino-Tibetan language family. Robbins Burling (1983) classifies the North-Eastern Indian Tibeto-Burman languages into three areas: Central, Eastern, and the Western. According to this classification, Tiwa language is placed in the Central Area and comes under the Bodo-Garo-Koch sub group having close similarities with the Kokborok language.

Tiwa language is spoken in the Morigaon, Nagaon, Karbi- Anglong and Dhemaji districts of Assam and in Ri-bhoi district of Meghalaya. This language falls under UNESCO's list of endangered language. According to the 2011 census report Tiwa tribes has 3,71,000 members whereas only 33,921 of them can speak in that language. *Atlas of the World's Languages in Danger* categories Tiwa as definitely endangered language.

Based on their settlement Tiwa speech community has two types of speech varieties. One is used in the hills and another one is used in the plains, which are known as Hajowali and Datiwali or Tholuwali respectively. In other words,

- The varieties spoken by people settled in the hills are known as Hajowali
- The varieties spoken by people settled near the hills as Datiwali
- The varieties spoken by people settled in plains as Tholuwali.

The variety of Tiwa used in plains has gradually made way for Assamese, which has almost replaced Datiwali and Tholuwali. However, in the hills Tiwa is still being in use along with Assamese. Hajowali Tiwas live mostly in West Karbi Anglong district of Assam and Ri-Bhoi district of Meghalaya (Gohain:1993).

3. Importance of Mother tongue Education

Mother tongue is a child's first language. It is the language that he/she learns from his mother at his home. It is also known as one's native language. Studies show that children who study in their mother tongue learn better and faster than children who study using second languages or non-mother tongues (Akinnaso, 1993; Pong, 1993; Benson, 2005a, b; Lewis, 2006). Those who started education in their mother tongues can read and write better, even in other languages (Lin, 1997; Benson 2000; Hovens, 2002; Malone, 2003; Dutcher, 2004; Baker, 2006). Benson (2000) is of the view that *Teaching beginning literacy in the child's first languages helps him/her make the connection between meaningful speech and written language, rather than struggling to decode a language which she does not command.* Instruction in one's home language also increases self-esteem in at least two ways. First, teachers can get more immediate and comprehensible feedback about what their students know and what they are learning, so they can make more realistic evaluations of their pupils' performances. Secondly, children get an opportunity to express their knowledge and experience perfectly in a language in which they are competent (Benson :2002, Klaus: 2003, Dutcher: 2004). Another advantage of using home languages in schools is that it preserves minority languages and folk traditions (Lin, 1997; Benson, 2002; Dutcher, 2004). In addition to this, it has been argued that mother-tongue instruction increases the active participation of parents in their child's school related activities, thus raising the child's esteem for their parents (Dutcher, 1995; Benson, 2002; Klaus, 2003; Malone, 2003; Dutcher, 2004; Benson, 2005a). Hovens (2002) argues that instruction in home language allows poorer children and girls to overcome their disadvantages and therefore improves the output of education systems.

4. Conceptualization of mother tongue with reference to NEP-2020

The National Education Policy NEP 2020 emphasis on multilingualism and places it at the centre of school education. Titled *Multilingualism and the Power of Language,* the language education section of the NEP 2020 seeks to promote linguistic diversity in school education. This policy recognizes the power of language as a means of effective communication, cognitive development, and cultural understanding among students. It encourages the use of technology to facilitate learning and teaching of languages, and the creation of educational materials in different languages.

NEP acknowledges the importance of mother tongue education as a pedagogical approach as it offers psychological and educational benefits to the students belonging mainly to the lesser- known communities and tribes. Policymakers believe that using mother tongue will help strengthening the bond between learners and their cultural identity, which in turn will help latter understand and appreciate their heritage. With this vision NEP keeps mother tongue as the medium of instruction till Grade 5 while recommending its continuance until the Grade 8 and also in higher education.

NEP 2020 is a boon for all the native Indian languages particularly those who come under endangered language group. Tiwa is one such language of Assam. This language spoken by Tiwa tribe of Assam is used only by the people settled in the hills of West Karbi Anglong district of Assam and Ri-Bhoi district of Meghalaya (Gohain:1993). These Hajowali Tiwas consider Tiwa as their mother tongue. On the other hand, more than 90% of Tiwas living in plains have already adopted Assamese as their mother tongue. Tiwas living in plains and foothills area of Nagaon, Karbi Anglong, Dhemaji, Kamrup and Morigaon districts of Assam are surrounded by Asasamese speaking community and thus Tiwa becomes the linguistically minority language and Assamese becomes the dominant language. The new generation of Tholuwali and Datiwali Tiwas are gradually shifting to Assamese language and culture. They identify themselves with Tiwa ethnicity but consider Assamese as their mother tongue. Moreover, neither their parents encourage and speak in Tiwa nor they get a chance to use Tiwa with their peers at the educational institutions. This results in a mix language called Tiwamese which the plain Tiwas use. Tiwamese is a mix of Assamese and Tiwa language and is greatly differed from native Tiwa language. Due to this difference of mother tongue identity Tiwa tribes has only 33,921 language speakers as per the census report of 2011. This is a major linguistic challenge the Tiwas are facing in present times. NEP 2020 with its focus on mother tongue education brings a ray of hope to the Tiwa community also.

5. Methods of study

While preparing this research article, descriptive method is used. This article is based on both primary and secondary data sources. One primary source was the native speakers of Tiwa who were given unstructured questionnaires for collecting information about the language. Another source was the Policy Document of the NEP 2020, provided by MHRD of the Government of India. For secondary source materials different books, study materials, journals, periodicals online articles have also been consulted.

6. Challenges of implementing NEP 2020

Implementing a language as medium of education depends on several inter related factors like availability of quality text books, reading materials, appropriate curricula, assessment tools and digital platform. However, in Tiwa language there is very limited resources available. To introduce Tiwa language as a medium of instruction at least in the Tiwa dominated areas '*xodou axom tiwa xonmilon*' published two text books. Only a handful of Tiwa writers have produced Tiwa Grammars and Dictionary in the language.

There is lack of trained teachers in Tiwa. Those who are proficient in this language mostly do not have adequate qualification to be employed as language teachers. This makes it difficult for the Government to find qualified teachers for the tribe.

There is also a need for coordination and collaboration among different stakeholders, like central and state governments, community, and educational institutions. But due to various socio- political issues Tiwas are not successful enough in this domain.

Another challenge is lack of awareness among the members of the tribe. Many parents, teachers, students, and policymakers are not aware of the benefits of mother tongue education. New generations of Tiwas have misconceptions or prejudices about certain languages or dialects. Parents also prefer English as the medium of instruction as they believe that English will provide better career and job opportunities for their wards. In this regard, stakeholders such as policymakers, educators, and language communities will have to work together to make sure that the language policy promotes awareness and knowledge about the language.

Curriculum alignment is also one of the challenges policymakers faces while implementing mother tongue education as ensuring that students have access to all-round development while also valuing their linguistic backgrounds is crucial. Therefore, balancing the use of mother tongues or regional languages with national or standardized curricula can be challenging.

7. Possibilities

As starting early education in one's mother tongue provides a strong foundation for learning additional languages, education in Tiwa language will equip the new generation learners with the necessary language skills and competencies to participate in various domains of life, such as education, employment, research, innovation, etc. When students can understand what they are being taught, they are more likely to stay in school and complete their education. This will reduce school dropout rate and increase employability.When a language is recognised as medium of instruction its members also starts respecting the language and feel proud associating with it. This helps saving various minority languages like Tiwas from extinction.

8. Conclusion

National Education Policy 2020 strives to promote regional languages with its three-language formula. This puts emphasis on mother tongue education. It also raises awareness among the member speech communities. Tiwa is an indigenous language of Assam. However, this language is recognised as definitely endangered language. Tiwas living in plains left their original linguistic identity and have started associating with Assamese language and society. They consider Assamese as their mother tongue leading to near extinction of Tiwa language. Only those Tiwa living in hills are retaining their language and culture. In such a scenario NEP, with its focus on mother tongue education, works as boon for the community. However, there are number of challenges like lack of text books, reading materials etc. are to be tackled. It also demands interest and active participation of the community. Then only Tiwa language and tradition can be preserved and its rich and colourful culture will continue flourishing.

References

Acharyya, P & Mahanta, S. (2019). Language vitality assessment of Deori: An endangered language in Language Documentation & Conservation Vol. 13: 514–544.
Akinnaso, F.N. (1993). Policy and Experiment in Mother Tongue Literacy in Nigeria, International Review of Education, Vol. 39 (4): 255- 285.
Baker, Colin (2006). Foundations of Bilingual Education and Bilingualism, 4th edition, Multilingual matters LTD: Clevedon, England.
Benson, C.J. (2000). The Primary Bilingual Education Experiment in Mozambique, 1993 to 1997 in International Journal of Bilingual Education and Bilingualism, Vol. 3 (3): 149-166.
Benson, C.J. (2002). Real and Potential Benefits of Bilingual Programmes in Developing Countries in International Journal of Bilingual Education and Bilingualism, Vol. 5(6): 303- 317.
Benson, C. (2005a), Girls, educational equity and mother tongue, Bangkok: UNESCO.
Benson, C. (2005b), The Importance of mother tongue-based schooling for educational quality in Commissioned study for EFA Global Monitoring Report 2005.
Burling, R. (1983). The Sal languages in Linguistics of the Tibeto-Burman Area, 7(2): 1-31.
Burling, R. (2003). The Tibeto-Burman languages of Northeastern India in Graham Thurgood and Randy J. LaPolla (eds.) The Sino-Tibetan Languages: 169-191. London and New York: Routledge.
Census of India (2011) Office of the Registrar General & Census Commissioner, India. India. [Web Archive].
Debbarma, S. (2021). A comparative phonology of Kokborok dialects: An analytical study of literary Kokborok, Reang and Jamatia dialects, Unpublished PhD Thesis, Tripura, Tripura University.

Dewri, P. and Haokip, D.M. (2021). Language Endangerment: A Case Study of Tiwa, An Endangered Language of North East India in Veda's Journal of English Language and Literature (JOELL) Vol.8(3): 81-93.
Dutcher, N. (2004), Language Policy and Education in Multilingual Societies: Lessons from Three Positive Models. Centre for Applied Linguistics. Washington, DC. USA
Gohain, B. K. (1993). The Hill Lalungs, ABILAC.
Klaus, D. (2003). The Use of Indigenous Languages in Early Basic Education in Papua New Guinea: A Model for Elsewhere, Language and Education, Vol. 17 (2): 105-111.
Lewis, M. and M. Lockheed (2006). Inexcusable Absence: Why 60 million Girls Still Aren't in School and What to do About it, Washington, D.C.: Center for Global Development.
Lin, J. (1997). Policies and Practices of Bilingual Education for the Minorities in China in Journal of Multilingual and Multicultural Development, Vol. 18 (3): 193-05.
Malone, D.L., (2003). Developing Curriculum Materials for Endangered Language Education: Lessons from the Field, International Journal of Bilingual education and Bilingualism, Vol. 6 (5): 332-348.
Nath, Arup (2010). A Lexico-Semantic Study of Tiwa and Deori: Two Endangered Languages of the Tibeto-Burman Family. Jawaharlal Nehru University. New Delhi, India.
Nath, Arup (2011). Revitalizing Tiwa Language: The Phenomena of Language Shift and Maintenance in Language Forum, Vol. 37 (2): 73-84.
Pong, S. (1993). Preferential Policies and Secondary School Attainment in Peninsular Malaysia, Sociology of Education, Vol. 66 (4): 245- 261.

The Structure of Classifiers in Kokborok

Priya Debbarma
priyakhum908@gmail.com
&
S. Indrakumar Singh
iksagol@gmail.com
Department of Linguistics and Tribal Languages
Tripura University

Abstract

This paper attempts to discuss classifiers in Kokborok, a Tibeto-Burman language spoken in north-eastern part of India, Tripura. Kokborok is a Tibeto-Burman language with SOV word order. It is a highly suffixing language. This language shows some cognates features with Bodo, Garo and Dimassa. Classifiers in Kokborok occurs with the noun along with the numerals. The classifiers of this language precede as well as followed by noun. In Kokborok classifications of classifiers are done mainly on the basis of different shapes, sizes or quantities of noun. This work will discuss the structure of classifiers in Kokborok.

Keywords: *Kokborok, Tibeto-Burman, Noun, Classifiers, Structures.*

1. Introduction

Kokborok is one of the major languages spoken mainly in the state of Tripura, north-eastern part of India. Few speakers are also found in Mizoram, Assam and in neighbouring country like Bangladesh. According to the census of 2011 the approximate speakers of Kokborok are 10,11,294. Kokborok is the lingua franca of the majority of the 19 tribal communities of Tripura.

The language is classified in different branches or section by different scholars. G.A Grierson (1903) classified Kokborok under Bodo-Naga group. Robert Shafer (1966-1974) has classified Kokborok which belongs to the western Units of the Barish Section within the Baric sub division of Sino Tibetan. He divides Tibeto-Burman into four main groups; these include Bodic, Burmic, Karenic and Baric. Benedict (1972) also classified Kokborok under Bodo-Garo sub group of the Tibeto-Burman language family. Burling (2003) made a classification of kokborok under the sub group of Bodo languages which is a group of Bodo-Koch part of Bodo-Konyak-Jingphaw group of Tibeto-Burman languages.

Classifiers are generally rich in Kokborok. It occurs along with noun followed by numerals or quantifiers in this language. Classifiers are used for denoting the semantic classification of the referent based on the physical, shapes and sizes of the noun. In Kokborok, classifiers are bound in nature which means it can occur only with noun and numerals. Classifiers in Kokborok can occur before or after the noun. Therefore, this particular study will focus on the structure of classifiers for different kind of objects, nature, shape, sizes and quantity.

2. Methodology

In this study, language data has been collected from two sources: primary and secondary. The primary data includes information gathered through interviews and observations conducted with speakers aged between 28 and 62 years, representing various professions. The data was recorded using a tape recorder and later analysed. In addition, some data was obtained from secondary sources such as books, journals, articles, and theses.

3. Review of Literature

Few authors have mentioned about the classifiers in their work but they failed to mention the structure and functions of classifiers in Kokborok. Among them is Rupak Debnath (2014) '*Kokborok Language Origin & Development'*, briefly discuss about the classifiers for the things denotes to tree/plant is *phaŋ-* , *thai-* for fruit, *koŋ-* for stick, *bar-* for flower, *dek-* for branch, *kaŋ-* for flat objects, *təŋ-* for long and flexible objects, *lep-* for slice objects, *tək-* for pots in terms of measured quantities, *cu-* for food packets. He fails to show the structure of classifier based on different shape, sizes, nature, quantity etc.

François Jacquesson (2008), in his book *'A Kokborok Grammar'* mentions very few classifiers for the objects like *koŋ-* for stick objects, *kaŋ-* for flat objects, *khuŋ-* for round or spherical shape, *dul-* for round or spherical shape and *lep-* for piece. He also did not distinguish classifiers based on the occurrence of different objects.

Pushpa Pai (1976) in her work *'Kokborok Grammar'* mentions 27 numbers of classifier in her work. But she failed to describe the classifiers on the basis of different shape, size, nature, quantity. The structure and functions of classifiers is also not discussed.

Samir Debbarma (2014) in his thesis *'The Structure in Kokborok'* describes classifiers but fail to show the structure of classifiers on the occurrence of different objects and items.

After reviewing some of the works of Classifiers in Kokborok it is still found that many authors did not discuss the structure of classifier based on quantity, used with natural and object things etc.

4. Classifiers in Kokborok

Kokborok have a large set of classifiers. The classifiers are used regularly in this language. In Kokborokclassifiers are used for different kinds of objects including various shape, sizes, nature, quantity etc. It can occur with noun followed by numerals. The nature of classifiers in Kokborok are bound which means it cannot occur alone. Classifiers in Kokborok are mostly monosyllabic structure and few classifiers are di-syllabic. It follows the order of 'Noun CLF+ Numerals'.

4.1 Mono-syllabic Classifiers

Mono-syllabic classifiers refer to those classifiers that have only one syllable and these types of classifiers are also known as basic classifiers. Mono syllabic classifiers which are available in this language are given as follows:

1. *təŋ-* : used to denote any kind of long objects like thread
2. *thop-* : used for denoting to any kinds of liquid things like rain drops, water drops, blood drops, tears, drop of urine
3. *kang-* : used for flat objects like paper, books, mat, and also denotes to clothes
4. *phon-* : used to indicates to piece of meat and fish,
5. *twi-* : used for counting eggs of a birds.
6. *lep-* : used to denotes the object for paise, bread, yeast.
7. *koŋ-* : used to denotes the object for things made of bamboos, wood or pens, body parts like limbs, fingers
8. *ma-* : used to denotes to animals including reptiles, insects and birds
9. *dek-* : denotes to branches of trees and plants,
10. *thai-* : denotes to any kinds of fruits including vegetables fruits, etc.
11. *lai-* : occurs along with the various kinds of leaves including vegetable leaf, flower leaf, banana leaf and also leaf of a tree.
12. *taŋ-* : denotes to indicate long object strings like garlands, banana fruits.
13. *bar-* : occurs with noun to denotes various kinds of flowers.
14. *jor-* : used to indicates pair of any objects like cows, duck, or pair of shoes and sandal
15. *khop-* : used for measuring the quantity of mouthful water and rice.

It is observed that the above classifiers are mono syllabic which are distributed in different domains in Kokborok. The detailed examples are analysed in the following sections.

4.2 Di-syllabic Classifiers

Di-syllabic classifiers are those classifiers which have two syllables. These are very few which are available in Kokborok as given below:

16. *khorok-* : used for counting human beings.

17. *mochom-* : used for the quantity of handful rice, pulses.

18. *buruŋ-* : denote to indicate the bundle of thatch of any kind, stick, wood etc.

These classifiers are disyllabic available in this language. The examples are discussed in the following sections.

4.3 Nominal Classifier

Classifiers in Kokborok occur before or after the noun. Classifiers precede the numerals in Kokborok. The nominal classifiers are those classifiers that denotes to the noun by adding to the numerals as a prefix, as classifiers in Kokborok are bound morpheme. When the classifier is added with the numerals it specifies the quality of noun depending upon its state of object, shape and sizes. The examples are discussed in the following way:

khorok- This classifier is solely used for counting human nouns.

19. *cərai* *khorok-sa*
child CLF-one
'One child'.

20. *borok* *khorok-tham*
person CLF-three
'Three persons.

ma- This classifier is used while counting animals, birds, insects etc. It is used for singular nouns. Also, it has its plural form which is known as ***mak-***. Following instances can be discussed:

21. *pun* *ma-sa*
goat CLF-one
'One goat'.

22. *tok* *mak-tham*
bird CLF-three
'Three birds.'

23. *sui* *mak-bərui*
dog CLF-four
'Four dogs.

It is to be noted that classifier *mak-* is used to indicate plural form as in the e.g. 22 and 23, and for singular classifier *ma-* as in the example 21.

phon- This classifier is used to indicates pieces of meat and fish as can be seen in the following instances:

24. *ah* *phon-sa*
fish CLF-one
'One piece of fish (meat).'

25. *muihan* *phon- bərui*
meat CLF-four
'Four pieces of meat.'

twi- This classifier is used for counting eggs of a birds, insects, and some animals.

26. *toktwi* *twi-tham*
egg CLF-three
'Three eggs.'

27. toktwi twi-ba
egg CLF-five
'Five eggs.'

It is seen from the above examples that all classifiers in Kokborok are nominal i.e. it can occur only with noun followed by numerals.

4.4 Classifier of different shape and sizes

In Kokborok there are many classifiers that are also used for different shape and sizes. For example,

kaŋ- This classifier is used with flat object such as for paper, books, cloth, mat, fan.

1. *bijap* *kaŋ-sa*
book CLF-one
'One book.'

2. *lamthai* *kaŋ-tham*
mat CLF-three
'Three mat.'

3. *kisip* *kaŋ-bərui*
fan CLF-four
'Four fan.'

kol- This classifier is used with round objects like eye, seed, and things made of rice such as rice puff, popcorn, and rice.

4. *mokol* *kol-nui*
eye CLF-two
'Two eyes.'

5. *mairuŋ* *kol-sa*
rice grain CLF-one
'One rice grain.'

6. *bəcului* *kol-bərui*
seed CLF-four
'Four seeds.'

koŋ- This classifier is used with short objects like pen, stick, and some of the body parts of leg, hand, teeth, etc.

7. *suikoŋ* *koŋ-nui*
pen CLF-two
'Two pen.'

8. *latha* *koŋ-sa*
stick CLF-one
'One stick.'

9. *yakuŋ* *koŋ-nui*
leg CLF-two
'Two legs.'

10. *buwa* *koŋ-bərui*
teeth CLF-four
'Four teeth.'

təŋ- This classifier denotes to long and flexible objects like road, wire, thread, hair and bamboo strip.

11. *khənai* *təŋ-tham*
hair CLF-three
'Three hairs'

12. *lama* *təŋ-sa*
road CLF-one
'One road.'

13. *khətəŋ* *təŋ-nui*
thread CLF-two
'Two threads.'

14. *waruk* *təŋ-sa*
strips CLF-one
'One strips.'

taŋ- This classifier also occurs with long string objects for things like banana, garland.

15. *khumtraŋ* *taŋ-ba*
garland CLF-five
'Five garlands.'

16. *thailik* *bətaŋ* *taŋ-nui*
banana string CLF-two
'Two string of banana.'

d*ul-* This classifier is used with round objects like rice ball, ball made of mud

17. *mai dul* *dul-sa*
rice ball CLF-one
'One rice ball.'

18. *ha dul* *dul-nui*
mud bal CLF-two
'Two mud ball/clot.'

phon- This classifier occurs with small pieces for fish, meat.

19. *ah* *phon-sa*
fish CLF-one
'One piece of fish.'

47.*wahan* *phon-tham*
pork CLF-three
'Three pieces of pork.'

20. *muihan* *phon-bərui*
meat CLF-four
'Four pieces of meat.'

lep- This classifier is used with small flat objects like coin, bread, biscuit, yeast.

21. *awaŋ* *lep-sa*
bread CLF-one
'One bread.'

22. *puisa* *lep-nui*
paise CLF-two
'Two paise.'

From the above example, it is observed that Kokborok have classifiers for different shape and sizes such as for long and short objects, flexible objects, flat and round objects.

4.5 Classifier of measure/quantity

Kokborok have a classifier for the measurement or the quantity of things based on their physical properties or quantity of noun. Examples are,

mochom- This classifier is used for the quantity of handful rice, pulses, sand.

23. *dal* *mochom-sa*
pulse CLF-one
'One handful of Pulse.'

24. *mai* *mochom-bərui*
rice CLF-four
'Four handful of rice.'

25. *hachiŋ* *mochom-tham*
sand CLF-three
'Three handful of sand.'

khop- This classifier is used for the quantity of mouthful of drinks, water and rice,

26. *mai* *khop-sa*
rice CLF-one
'One mouthful of rice.'

27. *chuwak* *khop-nui*
wine CLF-two
'Two mouthful/sip of wine.'

28. *tui* *khop-sa*
water CLF-one
'One mouthful/sip of water.'

thop- This classifier is used for the quantity of drop of liquid things like blood, water, oil etc.

29. *thəi* *thop-nui*
blood CLF-two
'Two drops of blood.

30. *thok* *thop-sa*
oil CLF-one
'One drop of oil.'

31. *tui* *thop-tham*
water CLF-three
'Three drops of water.'

dol- This classifier is used for the measurement of group of human, birds, animal and insects.

32. *borok* *dol-nui*
people CLF-two
'Two group of people.'

33. *musuk* *dol-tham*
cowCLF-three
'Three groups of cows.'

34. *tok* *dol-sa*
birdCLF-one
'One group of bird.'

chap- This classifier is used for the measurement of bundles like money, clothes, pen, books.

35. *raŋ* *chap-sa*
money CLF-one
'One bundle of money.

36. *re* *chap-bərui*
cloth CLF-four
'Four bundle of clothes.'
37. *raŋ chap-tham*
money CLF-three
'Three bundle of money.'

dek- This classifier is used for the measurement of bunch of branches of tree, plant and flower.

38. *bedek* *dek-tham*
branch CLF-three
'Three bunch of branches.'

39. *khum* *bedek* *dek-sa*
flower branch CLF-one
'One bunch of flower branch.'

40. *buphaŋ* *bedek* *dek-tham*
tree branch CLF-three
'Three bunch of tree branches.'

jor- This classifier is used for pair of animal, bird, shoe, and ornaments like bangle, earing.

41. *yakəlap* *jor-sa*
sandal CLF-one
'One pair of sandals.'

42. *takhum* *jor-tham*
duck CLF-three
'Three pair of duck.'

43. *wakhum* *jor-bərui*
earing CLF-four
'Four pair of earing.'

44. *mathia* *jor-nui*
bangle CLF-two
'Two pair of bangles.'

It is clearly seen from the examples that Kokborok have classifier for the measurement or quantity of group, pair, drop, bunch, bundle etc.

5. Function of Classifier

In Kokborok the classifiers like *thai-, təi, təŋ-, taŋ-, phaŋ-, bar-, thai-, koŋ-, kaŋ-, kol-, phon- ma-, khuŋ-, kai-, lai-, lep-, etc.* are monosyllabic structure and few classifiers like *khorok-, mochom-, buruŋ-* are disyllabic. The classifier of this language follows the order of Noun CLF Numeral as stated in the above examples. The classifiers precede the numeral in Kokborok. The classifiers can occur before or after the noun in Kokborok. The classifiers in Kokborok quantify the objects based on their structures and they are classified based on different kinds of objects. For different shapes and sizes, classifier *kaŋ-, kol-, koŋ-, təŋ-, taŋ-, dul-, phon-, lep-* are used. In Kokborok the classifiers are also classified based on different kind of measurement such as classifier, *mochom-* is used for the quantities of the measurement of handful rice, sand and pulses, *khop-* is used for the quantity of mouthful of water, juice and rice, *thop-* is used for the measurement of the quantity of liquid things like drop of blood, water, oil etc, *dol-* is used for the measurement of groups like human, birds, animals and insects, *chap-* is used for the measurement of bundles like money, clothes, pen, books, papers, *dek-* is used for the measurement of bunch of branches for trees, plants and flowers, *jor-* is used for the counting of pair of animal, bird, shoe, ornaments.

6. Conclusion

Kokborok is a classifier language as discussed in the above examples. The classifiers are regularly used in this language along with the numerals. In Kokborok, classifiers are bound morphemes, it occurs with numerals in the form of prefix. The structure of classifiers is discussed based on quantity, different shape and sizes. Few classifiers are disyllabic and classifiers are mostly monosyllabic in Kokborok. Classifiers in Kokborok have a large set, like other Tibeto-Burman language as Boro, Dimasa, Garo, Rabha.

References

Debnath, Rupak (2014). *Kokborok: Language Origin & Development (KLOD).* Khumulwng: Language Wing.

Debbarma, Samir (2014). *The Structure of Kokborok.* Ph.D. Dissertation, Assam University, Silchar.

Grierson, G. A (1903). *Linguistics Survey of India*. Calcutta: Government Publication, Vol 3, Part 3.

Jacquesson, Francoise, (2008). *A Kokborok Grammar [Agartala Dialect]*. Agartala: Kokborok Tei Hukumu Mission.

Karapurkar, Pushpa P (1979). *Kokborok Grammar*. Mysore: Central Institute of Indian Languages.

Conceptualization of Time in Hindi in Perspective of Cognitive Theory of Metaphor

Adwitee Verma
University of Lucknow
adwiteeverma270796@gmail.com

Abstract

In the past six to ten years, the Theory of Metaphor has become the most impressive and widely used theory. Works and research dealing with this topic exuberate but there is still much to discover. The present paper is an attempt to simplify and comprehend what Cognitive Theory of Metaphor deals. The paper presented is therefore only a small piece that will shed some light on the phenomenon of Cognitive Metaphor, particularly in the field of Conceptualization of Time in Hindi. The differences between this study and the previous study are the object and aims of the study. There have been a number of studies on the conceptualization of TIME in English, though none of them has been in Hindi. In this we will look at and analyse the conceptual metaphors of TIME in Hindi.

The scope of this paper is to find out how we can conceptualize TIME in Hindi in perspective of the Cognitive Theory of Metaphor and the work will also ascertain the Conceptual Metaphor of TIME.

In this study, we will use the descriptive qualitative method. I have collected the data of Hindi expressions from the primary method through observation of daily conversation. To conduct the research, I have, firstly, investigate and gather the conceptual metaphors and grouped the metaphorical expressions of the conceptual metaphors under their general and specific source (PERSON, MONEY, BIRD, etc.) and the target domain (TIME). Then following Lakoff and Johnson's methodology (Conceptual Metaphor Theory), I haveanalysed the whole data.

Keywords*: Metaphor, Cognitive Theory of Metaphor, Source, Domain, Target Domain, Metaphorical Meaning.*

1. Introduction

As the title suggests, this paper deals with the area of the Cognitive Theory of Metaphor. Works and research in this field are abundant, but there is still much to explore. This paper is, therefore, a small contribution that aims to shed light on the phenomenon of cognitive metaphor, particularly in the conceptualization of time in Hindi. In Hindi, time is referred to as *Samaya* or *Kaal*. Time is not a concrete entity; it is an abstract concept. Abstract concepts are often represented through metaphors.

Metaphors are used extensively in everyday communication, politics, education, and science. Moreover, most basic concepts in our world are understood through some form of metaphorical language—for example, concepts such as time, state, and quantity. Consider the conceptual metaphors "GOOD IS UP" and "MORE IS UP" (Lakoff and Johnson, 1980:23), which illustrate the mapping process, where a quality (an abstract concept) is associated with a concrete observation implying vertical movement, as in "prices are high" or "I'm feeling up."

2. Aim and Objective

The scope of this paper is:

- To explore how time is conceptualized in Hindi from the perspective of the Cognitive Theory of Metaphor.
- To identify and analyze the conceptual metaphors related to TIME.

3. Data Collection

This study adopts a descriptive qualitative approach. It is grounded in the claim that human thought is inherently metaphorical, and that conceptual metaphors arise from our bodily experiences (Lakoff and Johnson). Several metaphorical expressions for TIME are used in Hindi. The research corpus was compiled from written literary sources (such as proverbs and poetry) as well as spoken discourse.

Data were collected primarily through the observation of daily conversations in Hindi. Additionally, online resources containing metaphorical expressions in Hindi were used. To conduct the research, I first identified and gathered conceptual metaphors, categorizing the metaphorical expressions according to their general and specific source domains (e.g., PERSON, MONEY, BIRD) and the target domain (TIME). The analysis was carried out using the methodology proposed by Lakoff and Johnson (Conceptual Metaphor Theory). Once the corpus was compiled, the data were analyzed within the framework of the Cognitive Theory of Metaphor.

4. Literature Review

4.1 Metaphor

Traditionally, metaphor was regarded primarily as a linguistic device, with little attention paid to its conceptual basis. For example, Peter Newmark (1988) approached metaphor in the context of translation—a good representation of the traditional or linguistic view. Before the cognitive approach gained traction, metaphor was seen largely as a feature of poetic imagination and rhetorical flourish. It was viewed as a property of language, rather than of thought or action.

From a cognitive perspective, however, metaphor is not just about words—it is about concepts (Lakoff and Johnson, 2003). Metaphors are essential cognitive tools that help us materialize abstract concepts during the process of conceptual development (Lantolf and Thorne, 2006). They are fundamental to how we understand and structure abstract domains. Rather than being limited to literary or rhetorical uses, metaphors are pervasive in daily communication and thinking.

For instance, abstract concepts like LOVE, TIME, THEORIES, and LIFE are structured by conceptual metaphors (CM) such as LOVE IS A JOURNEY, TIME IS MONEY, THEORIES ARE BUILDINGS, and LIFE IS A GAME.

According to the cognitive approach, the three main characteristics of metaphor are:

1. Universality: Metaphor is not limited to a few speakers—it is found throughout both written and spoken discourse and forms a basis for much of our thinking.
2. Conceptual Nature: Metaphor is more than a linguistic phenomenon; it reflects our thought processes and cognitive structures.

3. Systematic Structure: A single metaphorical concept can generate many related linguistic expressions. These metaphorical systems are interconnected and form coherent networks (see Lakoff and Johnson, 2003).

4.2 Metaphorical Meaning

Metaphorical meanings arise through conceptual metaphor mappings, which are grounded in correlations from our experiences. Lakoff and Johnson (2003:247) assert that metaphor is a natural part of both language and thought. To analyze a metaphor's meaning, we identify a source domain and a target domain.

- The Source Domain (SD) is the conceptual domain from which metaphorical expressions are drawn. It is typically concrete, familiar, and explicit.
- The Target Domain (TD) is the abstract concept that is understood metaphorically via the source domain. It is typically abstract, unfamiliar, and implicit.

For example, in the metaphor "ARGUMENT IS WAR," *war* is the source domain, and *argument* is the target domain.

Source Domain	Target Domain
Concrete	Abstract
Familiar	Unfamiliar
Explicit	Implicit

4.3 The Cognitive Theory of Metaphor

The Cognitive Theory of Metaphor (CMT) originated with Michael J. Reddy's (1979) study *The Conduit Metaphor: A Case of Frame Conflict in Our Language*. It was later developed by George Lakoff, Mark Johnson, and their collaborators (Lakoff & Johnson, 1980, 1999; Lakoff, 1987, 1990, 1993; Lakoff & Turner, 1989). Their seminal work *Metaphors We Live By* (1980) marked the beginning of CMT as a formal theory.

Lakoff and Johnson argue that metaphor is not merely a linguistic device used for artistic effect—it is a fundamental aspect of thought. Everyday language provides evidence of how we mentally represent abstract concepts through concrete experiences. For example, metaphors such as "BAD IS DOWN," "GOOD IS UP," "POWER IS UP," and "TIME IS SPACE" show how we project physical experiences onto abstract ideas through conceptual mapping.

Key Features of CMT:

- Metaphors are pervasive in thought and language.
- There is a systematic mapping between two conceptual domains.
- Mapping proceeds from a concrete domain to an abstract one.
- Metaphors primarily originate in thought rather than in language.

5. Previous Studies

Evans' research offers a comprehensive examination of how language and cognitive processes shape our understanding of time. His work emphasizes the importance of metaphors, cognitive models, and embodied experiences in constructing temporal concepts. Cultural context is also highlighted as a key factor in temporal cognition.

One notable example is the metaphor "TIME IS MONEY," which influences how people perceive and manage time. Evans argues that such metaphors act as cognitive shortcuts, linking abstract temporal ideas with concrete, familiar experiences. He further suggests that our bodily experience of spatial dimensions (e.g., movement and direction) contributes to how we conceptualize time as a linear path. Evans also notes that conceptualizations of time vary significantly across cultures. These differences affect social behaviours and temporal organization, as cultural norms shape how time is perceived and managed. Inge Boot's (2010) work on metaphors in abstract thought explores how metaphors structure and influence our understanding of complex concepts. She examines how abstract notions such as time, emotions, and morality are understood through metaphorical mappings. For example, TIME AS A RESOURCE ("spending time") or EMOTIONS AS FORCES ("being overwhelmed") illustrates how metaphors help frame and manage abstract ideas. Boot emphasizes the cognitive implications of metaphorical thinking, showing how metaphors shape our perceptions, reasoning, and communication. Her study highlights the critical role of metaphors in shaping not only language but also our cognitive frameworks.

6. Data Analysis

This section investigates the conceptual metaphors found in metaphorical expressions related to time in Hindi. These expressions, collected from various lexical and cultural sources, provide insight into how time is conceptualized in the Hindi language. The analysis employs the Cognitive Theory of Metaphor, which views language as a reflection of human thought and conceptualization.

6.1. Time is a Person / Agent

Mapping

Source Domain Target Domain

Person Time

In this metaphor, time is personified as an agent capable of actions, intentions, and emotions. Hindi speakers conceptualize time as a human being with influence and control.

Examples:

a. /səməj kɑ: təmɑ:ʧɑ: pədənɑ:/ - *Slap of time*

Time is personified as someone capable of slapping, implying suffering, helplessness, or punishment.

b. /səməj kɑ: sa:θ denɑ:/nə denɑ:/ - *Standby / do not support time*

Just like a person, time can either stand by or abandon someone.

c. /səməj kɑ: səmmɑ:n/ -*To respect time*

Time is seen as a dignified figure—wise, powerful, and worthy of reverence.

d. /samay ke sāth chalnā/ – *Move with time*
Implies keeping up with changes, evolving with society, or working efficiently.

e. /waqt ṭahartā nahī̃ hai/ – *Time does not stop*
Time is seen as a restless being, always moving forward.

f. /samay ke hāth mẽ/ – *In the hands of time*
Time is perceived as a controlling force, like a judge or destiny.

g. /samay kā khel/ – *Game of time*
Time is conceptualized as a player in a game, with outcomes depending on its "moves."

h. /samay kī dawāī / samay kā marham/ – *Medicine / ointment of time*
Time is viewed as a healer, capable of soothing pain.

i. /kāl ke kapāl par likhnā-mitānā/ – *Writing and erasing on the skull of Kaal*
Two metaphors: TIME IS A PERSON and SKULL IS A WRITING BOARD. Suggests that time is not fixed and can be rewritten.

j. /samay karwaṭ le rahā hai/ – *Time is turning*
Time behaves like a person shifting position—indicating change.

6.2. Time is Money

Mapping

Source Domain Target Domain

Money Time

Money is a valuable, limited resource—so is time. In Hindi, time is often framed in economic terms.

Examples:

a. /samay barbaad na karnā/ – *Don't waste time*
Just like money, time is too valuable to waste.

b. /samay kā na honā/ – *Out of time*
Similar to lacking funds; implies urgency or limitation.

c. /samay kā sahī upyog karnā/ – *Use time wisely*
Time must be invested or spent with care.

d. /samay bacha kar rakhnā/ – *Save time*
Just as we save money for future use, time can also be "saved."

6.3. Time is a Thing

Mapping

Source Domain Target Domain

Thing Time

In this metaphor, time is treated as a tangible object that can be changed, possessed, or lost.

Examples:

a. /samay ko badalnā/ – *Change the time*
Suggests time can be altered through effort or perseverance.

b. /tumhārā samay hai/ – *The time is yours*
Time is seen as a possession that grants agency.

c. /samay hāth se nikal gayā/ – *Time slipped from the hands*
Indicates lost opportunity; time as a graspable thing that can be dropped or lost.

6.4. Time is a Material

Mapping

Source Domain Target Domain
Material Time
Time is conceptualized as a substance or resource that can be wasted or preserved.

Example:

a. /samay kā barbaad honā/ – *Waste of time*
Time, like a non-renewable material, must be used judiciously.

6.5. Time is Medicine

Example:

a. /samay har ghaav ko bhar detā hai/ – *Time heals every wound*
Time is viewed as a healing force, soothing emotional or psychological pain.

6.6 Time is a Season

Example:

a. /acchā samay baras rahā hai/ – *A good time is pouring*
Just as rain brings abundance, so does a "good time." Time is likened to a season of prosperity.

6.7. Time is a Bird

Example:

a. /samay ko par lag gae/ – *Time grew wings*
Time is likened to a bird flying away quickly, indicating that time passes swiftly.

6.8. Time is a Water Body

Example:

a. /samay ke bahāv ne uske kalank ko dhō diyā/ – *The flow of time washed away his stigma*
Time is conceptualized as a river or stream with the power to cleanse or erase.

6.9. Time is a Machine

Time is often described in Hindi as a machine that can operate smoothly or malfunction. See the following examples:

a. /acchā samay chal rahā hai/ – *Good time is running*

b. /bahut kharāb samay chal rahā hai/ – *Bad time is running*
Time is understood as a machine that can "run" well or poorly.

7.Conclusion

As discussed above, time is conceptualized in various metaphorical ways in Hindi—such as a person, money, thing, medicine, bird, season, water body, and machine. Each metaphor reflects how humans use concrete experiences to understand the abstract concept of time.

This study reveals that the Conceptual Metaphor Theory (CMT) is highly effective in explaining these metaphorical expressions. It shows how abstract domains are mapped from more concrete, embodied experiences—demonstrating the interaction of culture, cognition, and language.

For instance, when we say Kalachakra (Wheel of Time), the word "chakra" (wheel) evokes an image of something circular and constantly moving. Our understanding of this metaphor is shaped both by cultural knowledge and bodily experience. This highlights the role of embodiment—our bodily experience in the world—and how it intertwines with cultural metaphors to construct meaning.

References

Boot, I. (2010). *Metaphors in abstract thought*. [Master's thesis, Radboud University Nijmegen]. Radboud Repository.
[Include full access link if available]
Evans, V. (2004). *The structure of time: Language, meaning and temporal cognition*. John Benjamins Publishing.
Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. University of Chicago Press.
Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed., pp. 202–251). Cambridge University Press.
Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. Basic Books.
Lakoff, G., & Turner, M. (1989). *More than cool reason: A field guide to poetic metaphor*. University of Chicago Press.
Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.
Newmark, P. (1988). *A textbook of translation*. Prentice Hall.
Reddy, M. J. (1979). The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought* (pp. 284–324). Cambridge University Press.

The Emerging Phonemes of Konyak: Orthographic Influence and the Phonemic Shift in Standard Konyak

Nikay Besa
nikaytajung@gmail.com
&
Dr. Pangersenla Walling
asen.walling@nagalanduniversity.ac.in
Dept. of Linguistics
Nagaland University

Abstract

The Konyak tribe, a major ethnolinguistic group in Nagaland, India, primarily inhabits Mon district, comprising ten areas, two regions, and about 112 villages. Each village has its own mother tongue, with most varieties belonging to the Tibeto-Burman language family and exhibiting varying degrees of mutual intelligibility. While existing literature refers to these varieties collectively as 'Konyak,' this study identifies distinct geographical divisions where villages within the same area speak varieties that are understandable, but it lacks mutual intelligibility when compared with varieties from other areas. This challenges the classification of these varieties as dialects or distinct languages. Due to this linguistic diversity, the Wakching variety was chosen as Standard Konyak for common communication. Consequently, its phonological structure is undergoing notable changes, including the loss of original phonemes and the introduction of new ones. This study examines these shifts, focusing on the emergence of the approximant /ɹ/ and fricative /z/, free variation phonemes which are non-native phonemes influenced by standardization and borrowing, and their implications for phonological change in the language.

Keywords: *Emergence, Approximant, Fricative, Phoneme, Shift.*

1. Introduction

Konyak is a major tribe that inhabits the Mon district of Nagaland. According to the geographical statistics provided by the Government of Nagaland, Mon District is bordered by Arunachal Pradesh to the north, Assam to the west, Myanmar to the east, Longleng district to the southwest, and Tuensang district to the south. Spanning an area of 1,786 square kilometres, it is situated between 94°49′ East longitude and 26°45′ North latitude, with an altitude of 897.64 meters above sea level. In terms of land area, Mon is the third-largest district, covering 10.77% of Nagaland's total area.

The Konyak people speak a range of distinct mother tongues, most of which are unintelligible to one another. Of these languages, the Konyak community has standardized the Wakching variety for general communication and written purposes. Since the various languages and dialects within the community do not have their own script, they have adopted the Roman script for writing. The Konyak language is part of the Naga sub-branch of the Tibeto-Burman language family. Marrison (1967) classified the Naga languages into three groups: A, B, and C. Konyak is part of the Type A.2 group, which also includes the Chang and Phom languages spoken in Tuensang district of Nagaland.

2. Research Methodology

This study is descriptive and comparative in nature. This research is based on fieldwork data, where speeches of both native speakers and non-native were recorded and analysed for the findings. The primary data collection involved elicitation sessions with native and non-native speakers. Comparative analysis with the older linguistic research of Konyak helped contextualize the phonemic shift. For the secondary data, published materials, dictionaries, and books written in the native language were utilized. Fieldwork was carried out in two locations: Mon town and Kohima.

3. Objective of the Research

- To discuss the emerging new phonemes of Standard Konyak and how it affects the language.
- To discuss the phonemic shift of certain sounds in Standard Konyak.

4. Status of the Standard Konyak

The exact number of speakers may fluctuate; however, according to the 2011 Census of India, Standard Konyak had a recorded speaker population of 244,477. This language belongs to the Tibeto-Burman family and remains one of the most understudied and under-documented languages within this linguistic group. In comparison to other Tibeto-Burman languages, Standard Konyak has received minimal scholarly attention, leaving many critical aspects of its phonology, grammar, and lexicon unexplored. While a few preliminary researches have been conducted by both outsider and a handful of local scholars, the overall body of work remains limited. Much of the existing research focuses on foundational linguistic elements, leaving numerous structural, sociolinguistic, and historical dimensions of the language unexamined. There is a pressing need for comprehensive linguistic documentation, particularly in areas such as tonal patterns, morphosyntactic structures, and dialectal variations within Konyak-speaking communities.

Despite the lack of extensive academic research, there have been some notable efforts to sustain and promote the language, primarily through religious, local, and educational initiatives. One of the most significant contributions has been in the realm of translation, particularly of religious texts. The Bible, hymn books, and other Christian literature have been translated into Standard Konyak, making the language a medium for spiritual and communal engagement. In the education sector, the Government of Nagaland has taken steps to institutionalize the language within the school system. Standard Konyak is taught as a subject up to the 8th standard, with textbooks specifically developed to facilitate language learning among younger generations. Reports indicate that educational materials for the 9th and 10th standards are currently in progress and may soon be introduced into the curriculum by the year 2025. The formal inclusion of the language in schools is a promising step toward its preservation, ensuring that the younger generation acquires proficiency in reading and writing in their native tongue.

In addition to government efforts, several local organizations have been playing a crucial role in promoting and revitalizing the language. The Konyak Literature Board (KLB)[1], Konyak Baptist Bumeinok Bangjum (KBBB)[2], and Konyak Union (KU)[3] have all contributed significantly in various capacities. Their efforts include funding language-related projects, developing Sunday School textbooks, encouraging the younger generation to engage with their linguistic heritage, and publishing religious materials and as well as secular literatures in Standard Konyak. Such initiatives highlight the growing recognition of the need to preserve and expand the use of Standard Konyak in both religious and secular contexts. However, for the language to truly thrive, a more interdisciplinary and community-driven approach is needed, one that integrates linguistic research, educational policies, and grassroots movements. By increasing scholarly engagement, digital documentation, and institutional support, Standard Konyak has the potential to be not only preserved but also revitalized in meaningful ways for future generations.

5. Language versus dialect

The Konyak region is broadly classified into two main divisions: the Upper and Lower regions. Each division is further subdivided into ten areas, comprising approximately 112 villages under different areas. According to the Mon District website, Government of Nagaland, 'the Konyaks speak different dialects in different villages. Each village has its own sub-dialect, quite distinct from others.'

[1] Konyak Literature Board is a local body that oversees and regulates all aspects of Konyak language and literature. It plays a central role in language preservation, literary development, orthographic and standardization etc.

[2] Konyak Baptist Bumeinok Bangjum is an organization of the Baptist churches of Konyak-Naga of Nagaland. Its Head Quarter is located at Mon, Nagaland.

[3] Konyak Union is a local apex body of the Konyak.

However, fieldwork observations indicate that villages within the same area generally speak a common variety with minor dialectal variations, maintaining mutual intelligibility. Additionally, neighbouring areas exhibit some degree of mutual comprehension. However, as the geographical distance between areas increases, linguistic intelligibility decreases. For instance, the Tobu, Tizit, and Wakching areas, located at opposite ends of Mon district, exhibit distinct linguistic varieties that are mutually unintelligible which are possibly distinct languages. Also, there are some areas like Tizit area and Wakching area with many villages, where villages despite located within same area exhibits unintelligible varieties.

This variation in mutual intelligibility raises the question of whether Konyak constitutes a single language with multiple dialects or a group of distinct languages. However, this question remains inconclusive in the absence of a comprehensive linguistic classification study. At present, all the speech varieties spoken within Mon district are collectively categorized under the umbrella term Konyak language/ common Konyak. However, the Wakching variety, which serves as the standard language, seems to be generally regarded as the representative of the Konyak language.

6. Adoption of Roman script: advantages and disadvantages

The adoption of the Roman script for writing Konyak has brought several advantages, particularly in terms of accessibility and ease of use. Given its widespread familiarity, the Roman script has facilitated literacy efforts and enabled the documentation of the language in religious, educational, and administrative contexts. However, despite these benefits, the script presents significant challenges in accurately representing the linguistic features of Konyak, particularly its phonemic inventory and tonal system. One of the primary issues lies in the fact that the Roman script is not inherently designed to accommodate all the distinct sounds of Konyak. As a result, certain phonemes lack proper representation, leading to inconsistencies and ambiguities in spelling.

Additionally, there has been a tendency within the community to incorporate all the letters of the English alphabet except the letter [x], even when they do not correspond to actual phonetic distinctions in the language. This not only complicates orthographic standardization but also affects pronunciation, as speakers may struggle to reconcile the written form with the actual spoken language. Moreover, Konyak is a tonal language, meaning that pitch variations play a crucial role in differentiating word meanings as well as grammatical distinction. However, the Roman script, in its conventional form, does not account for tonal distinctions, making it difficult to accurately convey meaning in writing. This absence of tonal marking can lead to misinterpretations, especially in contexts where tone alone differentiates words and grammar of the language. While some languages have adapted the Roman script by incorporating diacritics or additional markers to represent tone, such modifications have not yet been systematically

implemented in Konyak orthography. Also, the standardization was made randomly, so inconsistency in the spelling system has created confusion among readers as different authors employ varied spelling systems. For instance, the word 'Ahng'[4] in Konyak which translates to 'King' has varied spellings like, 'Ang', 'Angh' or 'Ahng'.

7. Overview of segmental and suprasegmentals phonology of Standard Konyak

Nagaraja (2010), Ezung (2023), Besa (2017), and Walling & Besa (2025) have all studied the sound system of Konyak. However, their classifications vary, with each author presenting a different number of phonemes in the language. However, this study will overview the sound classification by Walling and Besa (2025), as it is the most recent and comparatively more comprehensive work.

7.1. Segmental: Consonants

Walling and Besa (2025) identifies 17 pure consonantal phonemes, five free variation phonemes (presented in brackets) and two marginal consonantal phonemes (presented in square brackets). Among the 17 pure phonemes, 10 sounds are obstruent sounds and 7 are sonorant sounds occurring at 8 places of articulation and 7 manners of articulation. The consonantal sounds are presented in the table below:

Place → / Manner ↓	Bilabial		Labio-dental		Labial-Velar		Alveolar		Post-Alveolar		Palatal		Velar		Glottal	
Plosive	p	(b)					t	(d)					k	(g)	ʔ	
Plosive aspirate	pʰ												kʰ			
Nasal		m						n				ɲ		ŋ		
Fricative			(f)	(v)			s	[z]	ʃ						h	
Affricate											ʧ					
Approxi mant						w		[ɹ]				j				
Lateral Approxi mant								l								

***Table 1**: Consonantal Phonemes of Standard Konyak.* (Walling & Besa, 2025, p.15)

[4] Ahng refers to traditional Konyak chief or King. The Konyak community has a system where some villages traditionally have a chief or a King, known as 'Ahng'.

7.2. Segmental: Vowels and Diphthongs

Walling and Besa (2025) identify six pure monophthongs and ten diphthongs in Standard Konyak. The vowels are categorized based on tongue height, tongue position, and lip posture. The diphthongs are classified according to the direction of the glides and fall into three types: closing diphthongs, opening diphthongs, and narrow diphthongs. The following sections present the vowels and diphthongs.

	Front	**Central**	**Back**
High	i		u
Mid	e	ə	o
Low		a	

Fig 1: *Vowel chart* (Walling & Besa, 2025, p.38)

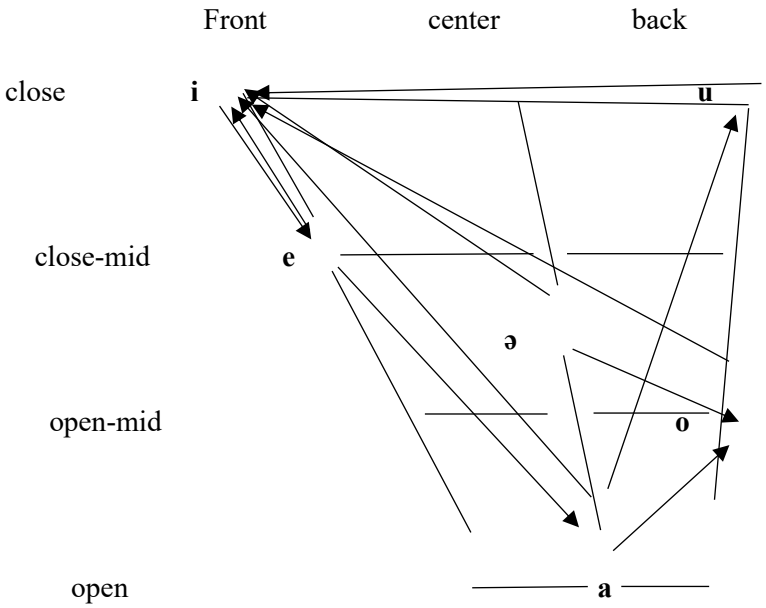


Fig 2: *Diphthong chart of Standard Konyak.* (Walling & Besa, 2025, p.46)

From the figures above, it can be observed that Standard Konyak has vowel sounds such as /i/, /e/, /ə/, /a/, /u/, /o/ and diphthongs such as, /ei/, /ai/, /əi/, /ui/, /oi/, /ie/, /ea/, /əo/, /ao/, and /au/.

7.3. Suprasegmentals: Tone and syllable structure

Walling and Besa (2025) identify three register tones in Standard Konyak: high marked by acute (˴), mid marked by grave (˵), and low marked by macron (¯). Additionally, Standard Konyak exhibits four primary syllable structures—V, CV, CVC, and VC. More complex syllable forms arise through the combination of these primary structures, adhering to the language's phonotactic constraints.

8. The Emerging Phonemic Status of fricative /z/ and approximant /r/ in Konyak

The relationship between orthography and phonology is often bidirectional, with writing systems influencing spoken language and vice versa. In Konyak, the emergence of the approximant /ɹ/ and the fricative /z/ reflects a phonological shift influenced by orthographic standardization. The phonemic status of /z/ and /ɹ/ in Konyak remains a subject of ongoing linguistic transition. Walling and Besa (2025) classified these segments as marginal phonemes. Ezung (2023) classified the approximant /ɹ/ as pure phoneme with limited distribution occurring only at the initial position. However, originally, the Wakching variety which is now standardized, do not have /ɹ/ and /z/ sounds. In order to incorporate the alphabets [r] and [z], the community borrowed three words such as, *'ra'* for servant, *'ree'* for shield and *'zah'* for bee from the neighbouring varieties.

Fieldwork revealed that native speakers, particularly those from Wakching village, do not use the word /zaʔ/ 'bee', as their variety already uses /ɲaʔ/ for the same concept and does not have the word /ɹa/ 'servant' or /ɹi/ 'shield' at all, meaning there is no equivalent word for 'servant' or 'shield' in their original variety. However, written conventions incorporated these borrowed words containing /z/ and /r/ to accommodate dialectal variation, particularly in varieties where these sounds occur. During the research, it was noted that the word '*zah'* 'bee' originates from the Mon area, while /ɹa/ and /ɹi/ 'shield' are used in the eastern Konyak regions, including *Sheanghah, Longwa, Shang, Chuoh, and Chen* among others.

Phonologically, /z/ and /ɹ/ exhibit limited distribution within Konyak, occurring primarily in borrowed forms rather than as native phonemes. While their restricted occurrence currently prevents them from being fully classified as part of the core phonemic inventory, their functional contrast in borrowed word suggests the potential for phonologization over time. Borrowing, particularly when sustained and reinforced through written conventions, has the capacity to introduce new phonemic distinctions into a language. Given these factors, the present study does not claim that /z/ and /ɹ/ have already achieved phonemic status in Konyak. Instead, it argues that their increasing presence through lexical borrowing, combined with their contrastive function, may facilitate their eventual integration into the phonemic system. This process, however, remains an ongoing development rather than a completed shift.

8.1. Sociolinguistic and Dialectal Implications

The integration of /ɹ/ and /z/ into Standard Konyak reflects a broader sociolinguistic effort to unify dialectal variation under a single orthographic system. While these phonemes are not originally robust in Standard Konyak, their prominence in other Konyak varieties justified their inclusion. This decision highlights the role of linguistic ideology in phonological evolution, as the standardization process prioritizes community representation over strict phonological consistency. Moreover, younger generations who acquire literacy in Standard Konyak are more likely to internalize /ɹ/ and /z/ as part of the language sound system. This phenomenon mirrors similar cases in other languages, where orthographic innovations have driven phonological restructuring. For example, French: The historical distinction between /s/ and /z/ in liaison contexts was reinforced by orthographic representations, leading to stable phonemic contrasts (Posner, 1997).

8.2. Consonant Adaptations in borrowed words

Since /ɹ/ and /z/ are not originally part of the Wakching variety, native speakers often adapt borrowed words containing these sounds through re-phonologization, modifying them to align with the native phonological structure. For example, monolingual Wakching speakers, especially elders consistently replace /ɹ/ with /l/ and /z/ with /ʃ/. Borrowed words such as /ɹa/ 'servant,' /ɹi/ 'shield,' and /zaʔ/ 'bee' are often pronounced as /la/, /li/, and /ʃaʔ/, respectively. Also, there are many foreign words from Nagamese, English and Hindi that are constantly used in the communication, and if the borrowed words contain /ɹ/ and /z/ sounds, they often replace it with /l/ and /ʃ/. For example, a native speaker with no influence from other speech varieties would replace the initial sounds in words like /ɹaŋkaŋ/ 'slope/heights' and /zebɹa/ 'zebra' with /l/ and /ʃ/, pronouncing them as /laŋkaŋ/ and /ʃepla/, respectively.

9. The Phonological Shifting of /s/ in Wakching Variety

In the Wakching variety of Konyak, the phoneme /ʃ/ occurs infrequently, though it exhibits phonemic contrast, whereas /s/ is more prevalent. However, Konyak orthographic conventions frequently represent /s/ as *sh*, influencing pronunciation patterns. For instance, the word /səko/ 'women' is written as /ʃeko/ [*sheko*], leading non-native speakers to replace historical /s/ with /ʃ/ over time. Walling and Besa (2025) discuss that while /s/ and /ʃ/ function as distinct phonemes, they also exhibit free variation in several contexts. This phonological shift suggests that the distinction may weaken, with /s/ potentially diminishing in prominence. Despite its foundational role in the language, increasing orthographic influence and non-native pronunciation patterns could contribute to its eventual loss. This underscores the broader impact of writing systems on phonological evolution, particularly in the context of standardization. If this trend persists, the phonemic boundary between /s/ and /ʃ/ may blur, leading to the possible obsolescence of /s/ in future generations.

Such developments highlight the need for further investigation into the long-term effects of orthographic standardization on Konyak phonology.

10. Orthographic Influence and the Phonemic Restructuring of Konyak

The integration of the Roman script into Konyak orthography has introduced phonological ambiguities, particularly concerning the representation of /d/, /g/, /b/, /f/, and /v/. These sounds are not historically part of the Konyak phonemic inventory but are now embedded in the spelling system due to external linguistic influence. As younger speakers are increasingly educated in English-medium schools, they can articulate these sounds with precision, even though they are absent in native phonology. However, inconsistencies in orthographic representation have led to a shift where voiced phonemes appear in positions where their voiceless counterparts should exist.

Over time, this orthographic interference has facilitated a gradual phonological restructuring, with younger speakers replacing traditional sounds with those represented in writing. While Walling and Besa (2025) categorize these sounds as free variants, their increasing stability in pronunciation suggests an ongoing process of phonemicization. This development raises critical questions about the role of literacy in reshaping oral traditions and phonemic boundaries. If left unexamined, the continued reinforcement of orthographic conventions may irreversibly alter the phonological system of Konyak. Thus, further research is imperative to assess whether standardization efforts should accommodate these shifts or seek to preserve the original phonological structure. Examples below are illustrative to support the phenomenon:

/tak/ ‘sick’ is represented as [dak] in orthography; younger generation accurately pronounce it as per orthographic representation.

/wən/ ‘fire’ is represented as [vün] in orthography; younger generation accurately pronounce it as per orthographic representation.

/pan/ ‘morung’ is represented as [baan] is some writings; younger generation accurately pronounce it as per orthographic representation.

/pʰa/ ‘teeth’ is represent as [fa] in dictionary; younger generation accurately pronounce it as per orthographic representations.

/kʰo/ ‘hundred’ is as [gho] in some writings; younger generation accurately pronounce it as per orthographic representations.

11. Orthographic Adaptation and Its Impact on Vowel Perception in Konyak

The representation of the *schwa* sound /ə/ in Konyak has undergone significant transformation due to orthographic simplification and historical pronunciation shifts. Traditionally, this sound was distinct, but its representation using umlauts [ü]

posed practical challenges in writing. As a result, the community adopted /e/ as a simplified alternative in some spellings. Additionally, historical variations in Roman alphabet instruction among earlier generations contributed to this shift. In some contexts, /e/ was pronounced closer to /a/, while in others, it was articulated as a more open /ahh/ sound. Moreover, /e/ was sometimes pronounced as /i/, causing further variation. Over time, these inconsistencies resulted in the widespread replacement of *schwa* (/ə/) with /e/ in writing, while the role of /e/ itself was increasingly taken over by /i/.

As younger generations rely on orthographic cues for pronunciation, this adaptation has become increasingly embedded in spoken Konyak. Given that vowels and diphthongs are fundamental to the phonological identity of Konyak, this shift has had profound implications. The loss of phonetic distinction affects not only pronunciation but also auditory comprehension, particularly when spoken by non-native learners of the standardized variety. In extreme cases, this phonological drift makes speech sound almost like a different language to native listeners. This phenomenon highlights the powerful influence of orthographic conventions on spoken language, reinforcing the need for further investigation into the long-term impact of writing systems on phonological integrity in Konyak. If left unaddressed, these changes could contribute to a divergence between traditional and standardized speech, affecting language continuity and mutual intelligibility. Some examples supporting the discussion are illustrative:

tao-ə		tao-e
1SG.ERG	>	1SG.ERG
'I-case'		'I-case'
ləŋ 'do'	>	leŋ 'do'
təman '2PL'	>	teman '2PL'

10. Conclusion

The interplay between orthographic representation and phonological evolution in Konyak reveals a profound shift driven by literacy, standardization, and historical variations in pronunciation. As demonstrated, the introduction of non-native phonemes through the Roman script, the rephonologization of borrowed sounds, and the adaptation of vowels such as 'schwa' (/ə/) to /e/ illustrate the extent to which writing systems can reshape spoken language. While these developments may facilitate literacy and linguistic unification, they simultaneously erode phonemic distinctions that have historically defined the language. However, in such cases, the decisions and acceptance of the community also becomes more important as language belongs to the community, and language itself is evolutionary in nature; it changes overtime.

The emerging trend, where younger generations rely on orthographic conventions for pronunciation, accelerates phonological restructuring, reinforcing the phonemic status of previously marginal or absent sounds. Furthermore, the perceptual gap between native and non-native speakers highlights the broader consequences of these shifts, where standardized speech may increasingly diverge from traditional phonology. This ongoing transformation raises critical questions about the role of orthographic standardization in language preservation, particularly in tonal and vowel-sensitive languages like Konyak. As linguistic change continues to be influenced by external forces, future research must critically evaluate whether standardization efforts should accommodate these shifts or seek to mitigate them to preserve the phonological integrity of the language. Ultimately, the case of Konyak underscores the intricate relationship between literacy, phonological stability, and language identity, warranting further scholarly engagement in documenting and analysing these evolving patterns.

Abbreviations

KLB: Konyak Literature Board
KBBB: Konyak Baptist Bumeinok Bangjum
KU: Konyak Union
V: Vowel
CV: Consonant Vowel
VC: Vowel consonant
CVC: Consonant Vowel Consonant

References

Besa, N. (2017). *Phonology of Konyak language*. Unpublished M.A. dissertation, Dept. of Linguistics, Nagaland University.
Ezung, M. K. 2022–2023. *Report on a comparative phonological investigation of the Tibeto-Burman languages of Nagaland*. Project sponsored by TRI, Department of Arts & Culture, Government of Nagaland.
Marrison, G. E. 1967. *The classification of the Naga languages of North-East India*. Mouton.
Nagaraja, K. S. (2010). *Konyak grammar*. Central Institute of Indian Languages.
Posner, R. 1997. *Linguistic change in French*. Oxford: Clarendon Press.
Walling, P., Besa, N. A. (2025). *Description of Konyak Grammar.* Woods Publishers.
https://mon.nic.in/
https://en.wikipedia.org/wiki/Mon,_Nagaland

A Study on the Linguistic Attitudes of Deori Speakers, an Endangered Language of Assam

Chandni Khaund
chandnikhaund@gmail.com
The Assam Royal Global University,
Guwahati, Assam

Abstract

The Deoris are one of the indigenous tribes of Northeast India. The language spoken by this ethnic group is also known as Deori or Dari, and it is one of the many endangered languages of the northern and eastern parts of Assam. While the community is divided into four clans or sub-groups based on their geographical distribution—Dibongiya, Tengaponiya, Borgiya, and Patorgiya—only the Dibongiyas have maintained their language. The remaining three are almost non-existent in the current scenario, as these communities have shifted to Assamese. This linguistic shift is due to intensive contact with Assamese, the dominant language in the area. Although Deori and Assamese belong to two different language families, Assamese has a greater influence on Deori. Long-term contact has led to a language shift within three territorial groups of the Deori community, also resulting in the loss of linguistic features in the Deori language.

However, the Dibongiya Deoris have made full efforts to maintain their language despite intensive contact and the threat from Assamese. This study, using a questionnaire (Lahiri & Majumdar, 2018), aims to analyse the outlook, attitudes, and knowledge of the remaining Deori speakers towards their language, attempting to understand their perceptions regarding the significance of maintaining their language amidst the influence of more dominant languages in the region.

The findings of the paper report that younger speakers of the language are equally competent in Assamese, making Deori a language spoken only within the local context and not beyond. However, a positive attitude is observed among the speakers, as they express an interest in maintaining their language and using it in spaces beyond their household and local contexts. Nevertheless, given the extended degree of contact with a widely spoken language in the area, linguistic changes within a minority language are inevitable.

Keywords : *Linguistic Attitude, Endangered, North East, Deori.*

1. Introduction: About the Language

Northeast India has the highest number of endangered languages; out of 170 endangered languages, more than half belong to the Tibeto-Burman family spoken in Northeast India (Moseley, 2007). The highest concentrations are reported in the states of Nagaland and Arunachal Pradesh. Deori also belongs to the Tibeto-Burman family. Categorically, it falls under the Boro-Garo sub-family of languages, spoken in the eastern regions of Assam and in some areas of Arunachal Pradesh. More specifically, the Deori community inhabits the upper Assam districts of Sivasagar, Jorhat, Tinsukia, and Lakhimpur. The 2001 Census reported that the language had a population of 41,161 speakers, but it is currently spoken by only approximately 32,000 people.

Deori, meaning "temple guard," was once the priestly clan appointed to perform sacrificial ceremonies for the Ahom kings (Goswami, 1994; Moseley, 2007). The Deoris refer to their language as Drori, Dari, or Dewri and call themselves 'Jimochaya,' meaning "children of the sun and moon" (Jaquesson, 2005). The dialect spoken in Lakhimpur district has retained the purest form of the language, with a similarity ranging from 77% to 93% to other Deori varieties. The official language of the state of Assam is Assamese, and the literacy rate of the Deoris in their second language (L2), i.e., Assamese, is reported to be 71% (Ethnologue, 2001). The vernacular languages of the Deoris or other minority communities, however, are almost entirely neglected in the school curriculum of the country unless they are scheduled languages.

The Northeastern part of India is known for its rich linguistic diversity. In this region, there is the presence of four major language families: Tibeto-Burman, Austroasiatic, Tai-Kadai, and Indo-European. Compared to its sister states, Assam once had the highest ethnolinguistic diversity. In modern times, however, Assam has a greater population of Indo-Aryan language speakers, with Assamese being one of them (Post & Burling, 2017). This shift has contributed to the dominance of Assamese over other minority languages spoken in the region.

Unfortunately, Deori has been classified as an endangered and "definitely endangered language" by UNESCO (2009) and as a "severely endangered language" in the Encyclopedia of the World's Endangered Languages (Moseley, 2010). Regarding the Deori language, Grierson (1903) remarked: "The Deori language may fairly claim to be the original language of Upper Assam." Today, the number of Deori speakers is dwindling. With only one section of the community continuing to speak Deori, the language is under threat of extinction. Young speakers seldom use their first language, and those who do are equally competent in their second language, Assamese—a dominant language in the area that belongs to the Eastern Indo-Aryan language family. This situation threatens the linguistic identity of many indigenous communities.

Brown (1895) notes that the Deori language is considered "moribund" due to the limited number of speakers compared to other Tibeto-Burman languages. UNESCO (2009) has stated that one of the primary reasons behind Deori's endangered status is the lack of intergenerational language transmission, as the older generation does not pass the language on to the younger generation.

It is worth noting that bilingualism is widespread in this area, and the Deori-speaking community is no exception. Native speakers of Deori are competent bilinguals, proficient in both Deori and Assamese, and they use these languages interchangeably in all social contexts—including at home, in school, and in the workplace.

One of the objectives of this survey is to analyse the linguistic competence of Deori speakers in languages other than their native tongue. It also seeks to investigate whether the Deori language is undergoing linguistic attrition, as many users tend to shift to more dominant regional languages. Bilinguals often experience interference from their second language, particularly if that language becomes a significant part of their daily lives. This interference can result in language attrition, which is more likely among speakers who frequently use a non-native language (Schmid & Köpke, 2014).

The speakers of Deori are multilingual, often switching easily between Assamese and Hindi. The younger generation, especially those educated in English-medium schools, also speak English.

As discussed earlier, the Deori language community can be socially divided into four territorial groups: Dibongia, Tengapania, Bargoyan, and Patorgoyan (Goswami, 1994; Jacquesson, 2005; Deori, 2009; Saikia, 2013). Only the Dibongias have maintained their language, whereas the other three groups have completely merged into Assamese culture and society (Goswami, 1994). They have also adopted Assamese as their first language, leading to a language shift. With only one section of the community speaking their language, the Deori language is under the threat of fading out. Young speakers seldom use their first language, and those who do are equally competent in their second language, Assamese—resulting in the borrowing of several lexical items from Assamese (Brown, 1895). This is one of many consequences of language contact and change.

To investigate the perspectives of native speakers of this threatened language within a society where another language is considered more 'prestigious,' this study aims to:

1. explore how this contact situation has shaped the perception and attitude of Deori speakers,
2. further investigate the factors affecting the ongoing language change,
3. examine how the Dibongia Deoris are striving to keep their language alive amidst the dominance of Assamese and English in the region, and
4. explore the consequences of this contact situation on the cultural and linguistic identities of native speakers.

Due to the intense contact with Assamese, the Deori language has undergone structural changes such as speech stylization, language crossing, borrowing, and code-switching—where speakers deliberately use styles associated with other ethnic groups. These phenomena often involve a stronger sense of social or ethnic boundary transgression, representing one of the most interesting current trends in research on language and ethnicity. However, this paper does not account for the structural changes affecting the Deori language, which will be the subject of future inquiry.

2. Language Contact and Its Effects on Deori

The Deori language has undergone drastic changes: three of the four community groups have shifted to Assamese, the number of speakers has declined, phonetic distinctions have been lost, and a large number of borrowings have occurred. Literature on endangered languages states that a language becomes threatened due to multiple factors. As discussed by Crawford (1996), the most common indicator is the dwindling number of speakers, followed by signs such as fluency in the dominant language among younger generations. Language use declines when it is no longer used at home or in schools. The issue worsens when parents fail to provide adequate language input to their children.

Fishman (1991) comments on the motivations behind language shift in a community. According to him, language shift is driven by both internal factors and external pressures from more dominant communities, which can ultimately weaken bonds within the community. The choices speakers make

about how to use their native language govern its survival. Additionally, Fishman (1991) discusses external factors leading to language shift, such as:

- demographic factors (e.g., in- and out-migration, intermarriage),
- economic factors (e.g., jobs available only to those fluent in dominant languages),
- media influence (e.g., content being accessible only in widely spoken languages), and
- social identifiers, where Indigenous communities try to emulate popular culture particularly evident in younger generations.

The choice to use one's native language in social settings is influenced by factors such as modern attitudes, urbanization (Gal, 1979), and demographic variables including age, education, gender, and place of residence (Huang, 1988). Therefore, the intensity of contact with another language is recognized as a major determinant in the process of language change (Thomason, 2001).

In the case of Deori, the situation appears to be shaped by the factors discussed above. The degree of contact between the Deoris and the Assamese-speaking community has been extensive. This is also true for other tribal minority languages spoken in the region. The Deori language is undergoing attrition due to its limited use in significant domains such as schools, homes, and various social settings.

The co-existence of multiple languages in the same geographic area can trigger language change (Acharyya & Mahanta, 2019). Further, Sallabank (2010), Whaley (1998), and Fishman (1997) define an endangered language as one that is no longer perceived as advantageous due to external or environmental pressures. Consequently, the visibility of an endangered language declines when it does not have official status or is not taught in educational institutions. Deori is not taught in schools, nor is it used in government institutions. This could be a significant reason for the language's endangered status. Studies on endangered languages suggest that the vitality of a language declines when it is no longer transmitted to successive generations.

3. Methodology

A total of 24 young Deori speakers participated in the study. The speakers ranged in age from 19 to 35 years, including 15 females and 9 males. Most participants were from Lakhimpur, as the Dibongiya Deoris inhabit this district of Assam. However, some participants also belonged to the city of Guwahati.
The methodology for this study consisted of a structured questionnaire that assessed language fluency among family members and the use of the language in educational contexts (Lahiri et al., 2018). The questions were designed to gather information on how many languages the participants were fluent in, what languages—apart from Deori—their parents were fluent in, and how the Deori language is represented in popular media such as movies and newspapers. Questions were also included about the existence of Deori literature in the form of books or linguistic resources such as grammar books or dictionaries.

The study, using the questionnaire, focused on determining whether Deori speakers use their native language to communicate with family members such as grandparents, parents, and siblings. The form included questions about participants' opinions on using the language in other spheres of life—such as in schools, businesses, or the workplace—and whether they would want to teach their children the Deori language. Lastly, the questionnaire served an important purpose in understanding the sentiments of

Dibongiya Deori speakers regarding language preservation and the extent to which they wish to carry the language forward. Overall, the questionnaire explored key issues related to language vitality, linguistic attitudes, and the preservation efforts of Deori speakers.

The questionnaire was distributed among members of the Deori community and analyzed to determine their perceptions and attitudes toward their native language, and to specifically understand how important it is for them to preserve it. The questionnaire was shared via Google Forms, and respondents were asked to complete it within a set deadline. Data collection also took into account demographic variables such as age, gender, and location. The participants' responses were examined both qualitatively and quantitatively.

4. Findings

As part of the survey, participants were asked whether they used any alternative names for their language. Most respondents indicated that their language had no other names. However, an interesting finding was that about 29% of the respondents believed Assamese was an alternate name for Deori. This is surprising and may point to two possibilities: either the respondents do not use Deori at all, or they are unaware of any alternate names for their language. Whatever the reason, if Assamese is considered a substitute for Deori, it clearly reflects the dominance of the Assamese language in the region. A few respondents referred to their language as *Jimochaya*, which translates to "children of the sun and the moon." However, this is not the name of the language, but rather a term referring to the people.

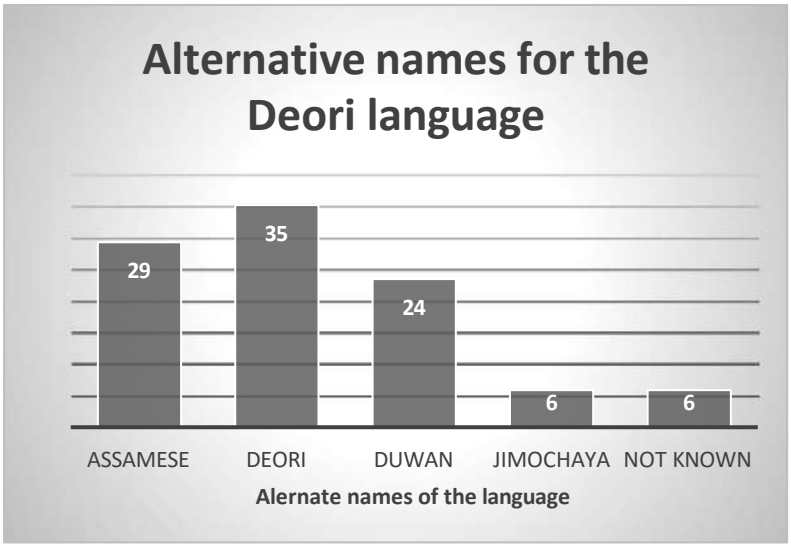


Fig. 2: Knowledge of other names used for the language

In the fluency assessment, it was found that speakers were fluent in more than two languages besides Deori. Thirty-three percent of respondents reported fluency in Assamese, Hindi, and Deori, while sixty-six percent reported fluency in four languages: Deori, English, Hindi, and Assamese. These results highlight that Deori speakers are multilingual, largely due to their education in Assamese or English medium schools, where emphasis is placed on dominant regional or national languages rather than minority languages. Most of the younger Deori speakers in this study had attended English medium schools and were fluent in English.

Although all respondents expressed confidence in their native language, many admitted they lacked strong competence in Deori. More than 60% of respondents were fluent in all four languages, while the remaining participants reported partial fluency. Notably, all participants reported fluency in Assamese.

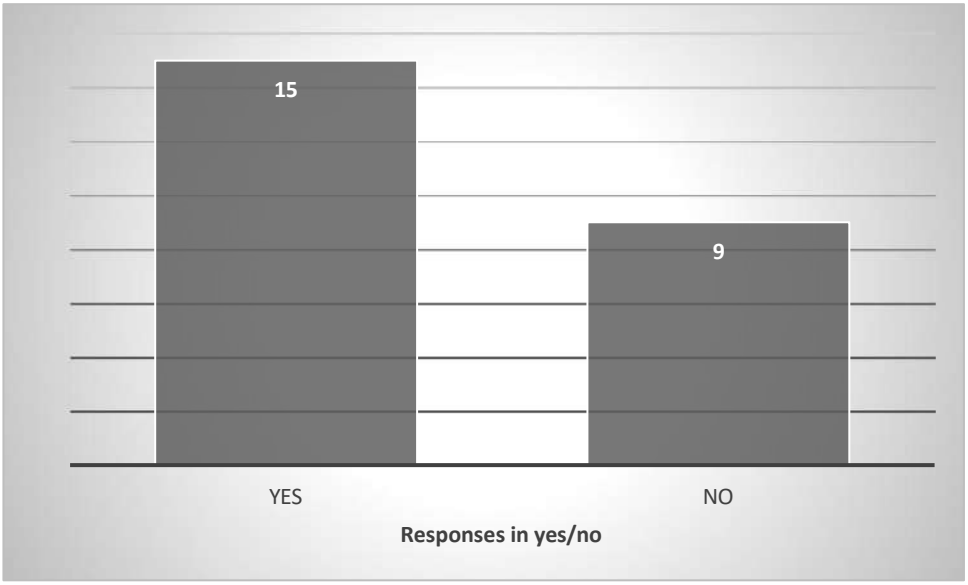


Fig. 3: Fluency in Deori by family members in Deori, English, Hindi, and Assamese

Deori speakers are aware that their language does not appear in popular or digital media. A majority of respondents (58%) confirmed that their language is not represented in print or on television. This was not a concern for them, as they had native competence in Assamese, and mass media in Assam primarily broadcasts in Assamese and English. However, participants observed that minority languages in Assam are generally underrepresented in mass media. They reported that there are no movies in the Deori language and that they watch English or Assamese films for entertainment. They did, however, enjoy listening to Deori folk songs.

When asked about the existence of literary output in their language, many participants responded positively. They mentioned that dictionaries, grammar books, and translated documents have been published in Deori. This indicates a significant effort toward language preservation and highlights the ongoing use of literary traditions for cultural expression.

Participants also explained why Deori is not spoken by all age groups in their community. The main reason, according to them, is the generation gap: the younger generation seldom speaks the language. Migration to urban areas for education or employment was also cited as a major reason. As younger people move to cities, they tend to use the language less frequently.

Inter-caste or inter-clan marriages were another factor mentioned by some respondents. For instance, if a Dibongiya woman marries outside the community and settles in a non-Dibongiya area, their children typically acquire Assamese as their first language. Conversely, if a non-Dibongiya woman marries a Dibongiya man and settles in a Dibongiya village, their children are more likely to learn Deori as their first language (Acharyya & Mahanta, 2019).

Respondents also suggested that members of other Deori clans cannot speak the language, reinforcing the earlier point that only the Dibongiya clan continues to preserve it. Younger respondents reported speaking Deori with their grandparents, both Assamese and Deori with their parents, and mostly Assamese with their siblings. This suggests that the older generations have a stronger affinity for the language, while younger speakers use it less frequently and less consistently.

When giving advice, telling stories, or making suggestions, most Deori speakers said they tend to use both Assamese and Deori. This indicates that Assamese comes naturally to them and that they are as fluent in Assamese as in their native language. Outside the home, Deori is seldom spoken; most communication occurs in Assamese or Hindi, reflecting the multilingual and heterogeneous nature of the communities they live in.

Respondents also shared their thoughts on where they would like to use their language. The majority expressed interest in using Deori in educational institutions, business, and employment. Many emphasized the need to introduce the language in schools as an essential step toward broader use.

Remarkably, every respondent (100%) expressed a strong desire for their children to learn Deori. They are committed to preserving and transmitting their cultural and linguistic heritage to the next generation. Participants noted that the language is not difficult to learn and can be mastered with regular use. Most respondents reported that several organizations and activists are actively promoting Deori through initiatives aimed at raising awareness about their culture and language.

Finally, participants were asked how they would like to use their mother tongue—for example, as a first language (L1), second language (L2), or medium of instruction. The figure below shows that the majority would prefer to use Deori as their L1, which is an encouraging sign for the future of the language.

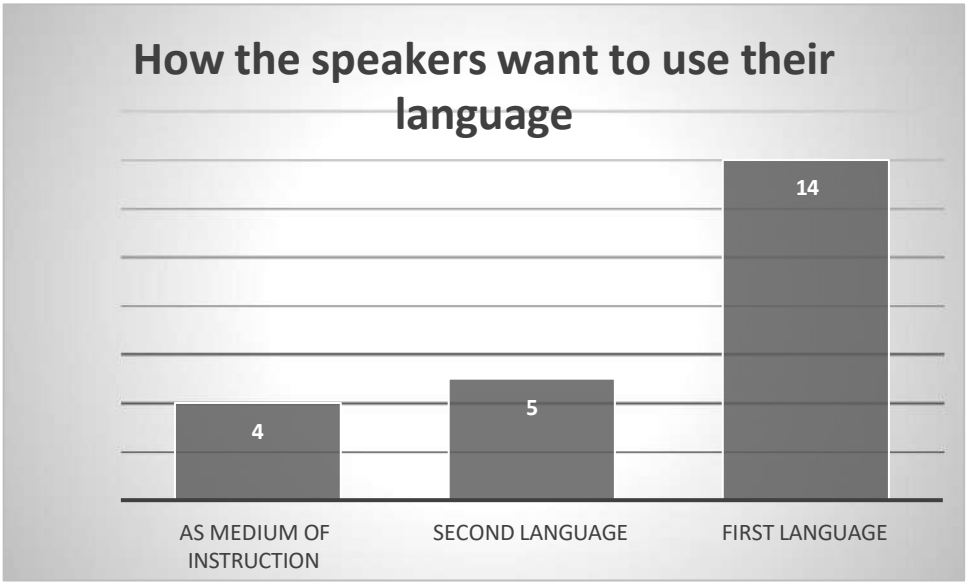


Fig. 4: How the speakers of Deori want to use their language

However, despite these positive attitudes, participants acknowledged that their language cannot be used in all contexts. Since most people in the region speak Assamese, Hindi, or English, Deori is primarily used only within the family and community. Outside of these circles, speakers feel more comfortable using dominant languages.

5. Discussion

The present speakers of Deori are loyal to their language. The number of speakers is therefore seen as growing, indicating stability (Ethnologue). Consequently, it should not be assumed that obsolescence will necessarily result in the extinction of the language. Language loyalty is thus an important element of language maintenance (Singh & Jones, 2005).

Julia Sallabank (2010) points out that the effort to preserve an indigenous identity—whether through language or culture—is especially relevant in the age of globalization. Many Deori speakers have migrated to larger cities within Assam in search of better livelihoods. As a result, many realize they may be losing touch with their indigenous identity due to migration. Consequently, younger speakers have become increasingly conscious and aware of their linguistic identity. Lanza & Svendsen (2007)

suggest that language may become a crucial marker of identity when a group perceives a threat to its identity due to political or social pressures.

In a similar vein, attitudes toward language play a key role in assessing the potential survival of an endangered language. This aspect is captured through the ethnolinguistic vitality framework (Giles et al., 1977) and UNESCO's language vitality scales (2003). A negative attitude—either from the speakers of the endangered language themselves or from those who speak dominant languages—can accelerate the shift toward dominant languages (Austin & Sallabank, 2013).

Deori is classified as a "definitely endangered language," meaning that children rarely learn it at home, and the majority of users are from the older generation. Gal (1979) and Schmidt (1985) observe that, typically, older adults continue to use the obsolescent language, while children switch to the dominant one. However, in the case of the Deori community, as reported by native speakers, both younger and older members use Deori as their first language. The younger speakers, however, are mostly bilingual and switch to Assamese when necessary.

In language contact situations, individuals often face difficult decisions regarding which language to use. In certain settings, they feel secure using their native language; in others, they prefer to use a more widely spoken language.

Thomason & Kaufman (1988) distinguish between language changes that occur due to internal motivations (innate changes within a language) and those that result from external motivations (changes triggered by contact with another language). Internal change is considered a natural tendency of all living languages, occurring regardless of social context (Winford, 2003). In contrast, externally motivated change occurs due to influence from another language (Marie & Singh, 2005).

With equal competence in Assamese—the lingua franca and official language of the state—Deori speakers are bilingual. Assamese has significantly contributed to the Deori vocabulary (Grierson, 1903), which is a clear example of externally motivated change stemming from long-standing contact between the two languages.

The Deori language has undergone considerable transformations and adaptations over time. Despite this, a dedicated section of the community is making concerted efforts to preserve and uphold their language, recognizing its cultural and historical significance.

Winford (2003) identifies three outcomes of language contact:

1. language maintenance,
2. language shift, and
3. the creation of new languages.

Language contact, in its simplest form, involves the simultaneous use of multiple languages within the same geographic area (Thomason, 2001). When the languages in contact are mutually unintelligible, participants may either develop a lingua franca (e.g., a pidgin) or begin learning the dominant language. This is the situation observed in the Deori language context. The community is compelled to learn Assamese, the dominant language, to function in everyday life.For Deori—as with many endangered languages—the contact situation is a key driver of linguistic change. As Thomason (2001) emphasizes, it is unlikely that structural changes in an endangered language occur due to anything other than contact. When multiple languages coexist, numerous foreign elements may enter a language, affecting its structure. Language contact, therefore, plays a significant role in language change. However, this study does not explore the structural changes resulting from contact with Assamese, which will be addressed in future research.

6. Conclusion

From this study, it is evident that the decline in language use among Deori speakers is primarily due to the lack of intergenerational transmission at home—a trend consistent with many other tribal languages in the region. Additionally, these languages receive minimal support from the government, further contributing to their marginalization. One significant challenge is that many of these languages, including Deori, do not have a script of their own, placing them at a disadvantage compared to official, scripted languages.

Nevertheless, the study reveals a hopeful outlook for the Deori language. Speakers across age groups expressed optimism regarding its future. Although there are efforts to preserve the language, young Deori speakers admitted that they are gradually losing touch with it. During interviews, many participants acknowledged uncertainty about numerous lexical items in Deori, indicating a weakening command of their native tongue.

The effects of language change and gradual shift are, to some extent, inevitable in the current sociolinguistic context. The findings also point to a lack of awareness about the importance of preserving and revitalizing indigenous languages in Northeast India. A significant issue is the absence of educational resources; currently, minority languages like Deori are not taught in schools. With increased awareness and promotion of such languages, however, there is potential for positive change.

Furthermore, the use of indigenous languages in popular or digital media is almost nonexistent. No digital platforms or media outlets are known to feature Deori, and the literary output remains limited to just a dictionary and a grammar book. Nonetheless, the Deori community has a rich oral tradition, including folk songs, tales, riddles, and proverbs, similar to other indigenous communities of Northeast India. These oral traditions provide a strong foundation for ongoing preservation and revitalization efforts.

To conclude, the native speakers of the Dibongiya Deori clan remain optimistic about sustaining and preserving their language. However, they need greater visibility, support, and initiatives from public agencies and organizations for full revitalization to take place. A promising development is the National Education Policy (NEP), under which the Assam government has taken steps to introduce indigenous languages as mediums of instruction in primary education. Notably, Deori has been included among seven languages recognized for foundational-level instruction.

In summary, as Austin and Sallabank (2013) argue, language endangerment is often driven by a process of language shift, wherein speakers of minority languages abandon their native tongues in favor of dominant languages supported by governmental and institutional systems. However, as this study shows, the Deori community is actively engaging in revitalization efforts to reverse this shift—especially in light of the language loss already observed in the other three territorial groups.

Acknowledgements

I am immensely grateful to my student, Sneha Deori, for assisting with the data collection process, and to all the respondents who participated in the interviews.

References

Abbi, Anvita. 2001. *A manual of linguistic field work and structures of Indian languages*. Munich: Lincom Europa.

Acharyya, Piyali and Mahanta, Shakuntala. 2019. Language vitality assessment of Deori: An endangered language. *Language Documentation & Conservation* 13. 514–544.

Bowern, Claire. 2010. Fieldwork in contact situations. In Hickey, Raymond (ed.), *The handbook of language contact*, Malden, MA: Wiley-Blackwell.

Brown, William. 1895. *The outline grammar of Deori Chutiya language*. Shillong: Assam Secretariat Printing Office.

Campbell, Lyle. 1998. *Historical linguistics: An introduction*. Cambridge, MA: MIT Press.
Fishman, Joshua A. 1991. *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Clevedon: Multilingual Matters.
Gal, Susan. 1979. *Language shift: Social determinants of linguistic change in bilingual Austria*. New York: Academic Press.
Goswami, Upendranath. 1994. *An introduction to the Deori language*. Guwahati: Anundoram Borooah Institute of Language, Art, and Culture.
Hickey, Raymond (ed.). 2010. *The handbook of language contact*. Malden, MA: Wiley-Blackwell.
Jacquesson, François. 2005. *Le Deuri: Langue Tibéto-Birmane d'Assam*. Leuven: Peeters Publishers.
Jones, Mari C. and Singh, Ishtla. 2005. *Exploring language change*. New York: Routledge.
Kerswill, Paul. 2002. Koineization and accommodation. In Chambers, J. K., Trudgill, Peter and Schilling-Estes, Natalie (eds.), *The handbook of language variation and change*, Malden, MA: Wiley-Blackwell.
Lahiri, B. and Majumder, A. 2018. A questionnaire developed for conducting fieldwork on endangered and indigenous languages. *Jadavpur Journal of Languages and Linguistics* 2(3).
McMahon, April. 1994. *Understanding language death*. Cambridge: Cambridge University Press.
Moseley, Christopher (ed.). 2010. *Atlas of the world's languages in danger*, 3rd edn. Paris: UNESCO Publishing. http://www.unesco.org/culture/en/endangeredlanguages/atlas.
Post, Mark and Burling, Robbins. 2017. *The Tibeto-Burman languages of Northeast India*. https://doi.org/10.4324/9781315399508.
Romaine, Suzanne. 2010. Contact and language death. In Hickey, Raymond (ed.), *The handbook of language contact*, Malden, MA: Wiley-Blackwell.
Sallabank, Julia. 2010. Language endangerment. In Hickey, Raymond (ed.), *The sage handbook of sociolinguistics*, London: Sage.
Sallabank, Julia. 2010. Language endangerment: Problems and solutions. *eSharp, Special Issue: Communicating Change: Representing Self and Community in a Technological World*. 50–87.
Schmid, Monika S. and Köpke, Barbara. 2004. Language attrition: The next phase. In Schmid, Monika S., Köpke, Barbara, Keijzer, Merel and Weilemar, Lina (eds.), *First language attrition: Interdisciplinary perspectives on methodological issues*, Amsterdam: John Benjamins.
Thomason, Sarah G. 2001. *Language contact: An introduction*. Edinburgh: Edinburgh University Press.
UNESCO. 2009. *Atlas of the world's languages in danger*. http://www.unesco.org/languages-atlas/.
Trudgill, Peter. 2002. Linguistic and social typology. In Chambers, J. K., Trudgill, Peter and Schilling-Estes, Natalie (eds.), *The handbook of language variation and change*, Malden, MA: Wiley-Blackwell.
Trudgill, Peter. 2011. *Sociolinguistic typology*. New York: Oxford University Press.

Examining the Vitality of Naga Languages through Intergenerational Language Transmission[1]

Dr. Kethokhrienuo Belho
Assistant Professor
Nagaland University
kethobelho@nagalanduniversity.ac.in

Abstract

This study examines the vitality of Naga languages using UNESCO's (2003) guidelines, focusing on intergenerational language transmission of the languages recognised by the Government of Nagaland. With the dominance of English in education and the rise of Nagamese in other spheres, indigenous languages face significant challenges. Understanding language vitality is crucial for sustainable language development and preservation.

The study examined two age groups to understand intergenerational language transmission among the Nagas. The first group consisted of individuals aged 20-25, while the second group included those aged 8-12. The findings revealed that language transmission is strong, with the indigenous language predominantly used in conversations with grandparents and parents. However, a noticeable and steady decline in indigenous language usage was observed in interactions with siblings and friends. The study also documented a decline in writing and reading skills of the indigenous language among the participants. Additionally, the study highlighted the limited use of the indigenous language on social media platforms, where English overwhelmingly dominates, signalling an alarming trend.

Keywords: *Language transmission, Vitality, indigenous language, English, Nagamese.*

1. Introduction

The global dialogue on language endangerment has predominantly centred on languages with small speaker populations, aligning with Krauss's (1992) assertion that any language spoken by fewer than 100,000 people is at risk. This connection between population size and language vitality is particularly relevant in the context of Naga languages. It is widely acknowledged that at least half of the world's 7,000 languages are endangered and are no longer being passed down to children as first languages. Without intervention, these languages are likely to vanish within the next century as the older generations who currently speak them pass away.Nagaland, a multilingual tribal society, boasts a rich linguistic heritage, with its Naga languages belonging to the Tibeto-Burman family. Despite having a population of around 2 million, the state is home to dozens of indigenous languages, 18 of which are officially recognized. The absence of a unifying language has led to the widespread use of English in education, while Nagamese, a Creole, dominates everyday communication. This linguistic shift poses a challenge to many Naga languages, leaving them increasingly vulnerable to the influence of these more dominant languages.

[1] This paper is an excerpt from the Young Faculty Start-up Project titled *'A study on Naga language endangerment and vitality'* sponsored by Nagaland University.

2. Language Endangerment and Vitality

An endangered or moribund language is one that faces the threat of disappearing, either because its speakers are dying or because they are abandoning it in favor of other languages. A language is considered lost when it no longer has any native speakers, at which point it becomes a "dead language." If such a language has not been properly documented, it cannot be brought back or revived. A language is in danger when its speakers cease to use it, use it in an increasingly reduced number of communicative domains, and cease to pass it on from one generation to the next. That is, there are no new speakers, adults or children.

Language vitality refers to the overall health and strength of a language, demonstrated by its use as a primary means of communication across various social contexts and for diverse purposes. A key indicator of vitality is the language's daily use within the home, as this reflects intergenerational transmission. A language with strong vitality is spoken widely by all generations, both inside and outside the home. Studying language vitality is crucial for assessing the likelihood that a language will continue to thrive in the future. It also helps determine whether efforts to develop and sustain the language are feasible and effective.

Generational language transmission of Naga languages is crucial for preserving cultural identity, passing down traditional knowledge, and maintaining social cohesion. It ensures that indigenous languages remain a living part of daily life, strengthening connections between generations. This process helps prevent language extinction and the loss of unique cultural practices. Moreover, children who learn their ancestral language often benefit cognitively, while communities can resist the dominance of global or regional languages. In essence, transmitting language across generations is vital for cultural continuity and the survival of linguistic diversity. The present study examines the generational language transmissions of Naga languages.

3. Review of related literature

The UNESCO (2003) Preamble on Language Vitality and Endangerment emphasizes the significance of language documentation and preservation. Despite the various internal and external forces driving language extinction and endangerment, the document highlights the necessity of raising awareness and proposes strategies for revitalizing vulnerable languages. It also cautions that assessing language vitality and endangerment is complex due to the diversity of language communities. To address this, UNESCO outlines nine key factors to assist language communities in comprehensively assessing the status of their languages. The first of these key factors is Intergenerational Transmission. The study, referencing the UNESCO scale, attempts to understand the vitality of Naga languages' transmission to the younger generation. Additionally, the study refers to previous research on the vitality and endangerment of Naga languages.

Temsunungsang (2022), focuses on the Chungli-Ao language and its dialects, using UNESCO's framework. While Standard Chungli-Ao is classified as vulnerable, its dialects are critically endangered, mainly due to the influence of English. Despite the widespread use of the standard language, the dialects face a severe risk due to limited intergenerational transmission in some areas.

Daimai (2018), examines the Liangmai (Zeliang Naga) community, analyzing linguistic factors such as code-mixing, code-switching, and the community's attitudes. This study offers a broader view of how language decline occurs across various domains and suggests that societal factors play a significant role in language endangerment.

Belho (2023), investigates the Tikhir Naga language. Although the community is small (under 10,000 speakers), the language remains relatively safe due to the community's cultural and linguistic loyalty. The study highlights that strong unity and commitment can preserve a minority language, but the Tikhir language is still in an early stage of growth in terms of vitality. Collectively, these studies emphasize that while some Naga languages are under threat, strong cultural ties and proactive community attitudes can play a significant role in their preservation. However, the English language and Nagamese growing influence and internal language dynamics present ongoing challenges.

4. Significance of this Study

The significance of this study lies in addressing the limited research on the generational transmission of Naga languages, an area that remains underexplored. While existing studies have focused on language endangerment and vitality, little attention has been given to how effectively these languages are passed down to younger generations. This study will provide crucial insights into current trends in language transmission among Naga youth and children, offering a clearer understanding of the role of external influences, such as English, on indigenous languages. The findings will inform language preservation efforts, shaping strategies for policymakers and community leaders to strengthen cultural identity and safeguard linguistic heritage.

5. Objectives of the Proposed Study

- To assess the current generational transmission of Naga languages among Naga young adults and elementary children living in urban areas.
- To propose ideas and strategies for effective revitalization efforts aimed at promoting the use of Naga languages among Naga youth and children.

6. Major Research Question

- What is the current generational transmission of Naga languages among young adults and elementary children residing in urban areas?
- What ideas and strategies can be proposed to effectively revitalize and promote the use of Naga languages among youth and children?

6.1 Research Design

This study will employ quantitative methods and use a questionnaire to collect data. The participants will be divided into two groups: Group A, consisting of Naga young adults aged 22–25, and Group B, comprising elementary school children aged 8–12. This approach will provide insights from different age groups, offering a broader perspective on the generational transmission of Naga languages.

6.2 Preliminary Observations

Urban children in mixed-tribal settings are adopting English and Nagamese as their first language. Analysing language vitality and assessing the status of language endangerment can provide insights into understanding why this phenomenon is growing.

7. Presentation of Data

The research focuses on two distinct age groups: Group A (ages 20-25) and Group B (ages 8-12), chosen based on their differing experiences with technological advancements and digital media. Both groups have distinct early language environments that shaped their language development.

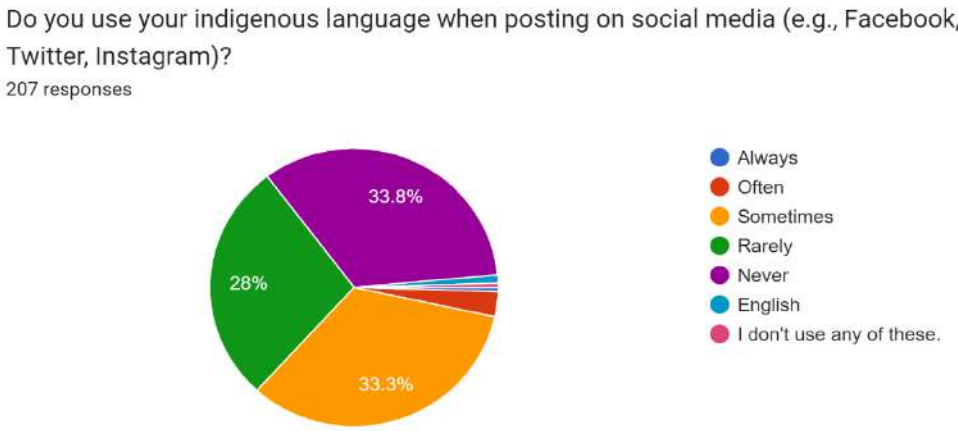
Group A grew up in an early language environment primarily influenced by their local community, schools, and social interactions. Their access to English came mainly through formal education, with media playing a smaller role during their formative years.

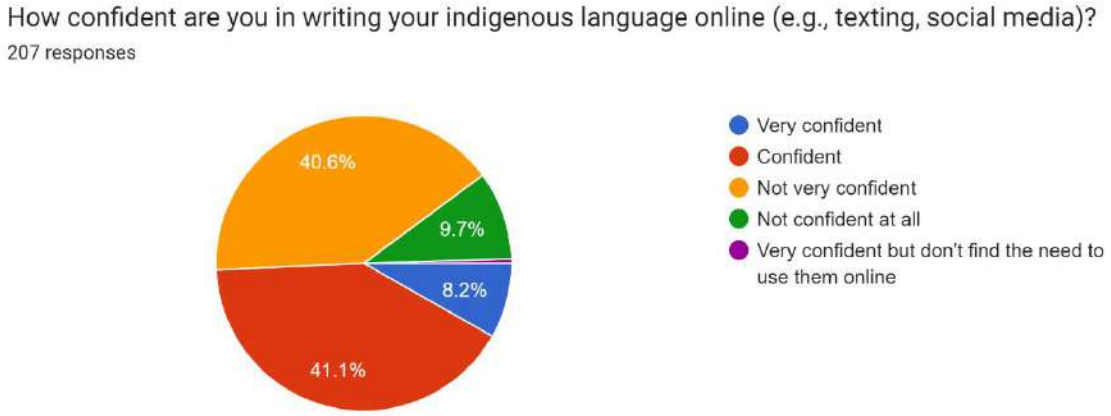
Group B, on the other hand, grew up in an early language environment heavily influenced by digital media and technology. From a young age, they were exposed to English through online platforms, apps, social media, and entertainment. This globalised media environment introduced English as a significant part of their daily communication, alongside their indigenous language.The study compares these two groups, assessing how their distinct early language environments and differing English access levels through digital media have shaped their language preferences, usage, and attitudes.

7.1 Group A findings presentation

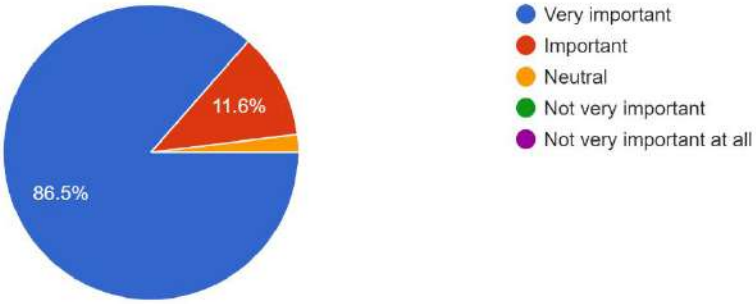
Language	Converse with Grandparents	Converse with Parents	Converse with Siblings	Converse with people belonging to the same community	Converse with Neighbors
Indigenous Language	99%	96.5%	91%	70%	NIL
Nagamese	0.5%	3%	3.8%	13.8%	73.6%
English	0.5%	0.5%	5.2%	5.2%	26.4%

The following charts are excerpt from Google Form Responses.

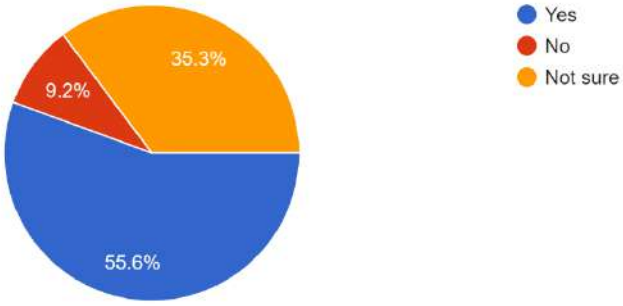




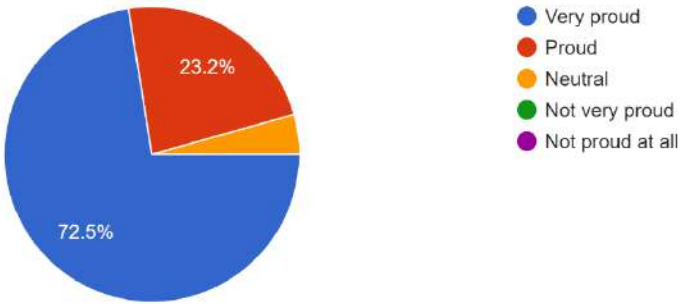
How important is your indigenous language to your cultural identity?
207 responses



Do you think your indigenous language will be as widely spoken in the future as it is now?
207 responses



How proud are you of your indigenous language?
207 responses



7.2 Key findings from Group A (Ages 20-25)

1. The data show that the majority of the participants converse with their grandparents, parents, siblings, and community in their indigenous language. This pattern suggests that Indigenous languages remain a core aspect of familial interactions, representing strong intergenerational

transmission at the family and community level. However, though the usage of Nagamese and English is low, it is alarming to see the penetration of it in this close knitted community.
There is a shift in language use with younger generations although indigenous languages are still used within families; there is a decline in their use when interacting with siblings. Respondents indicated that English or Nagamese is used with their siblings, reflecting a subtle but significant language shift that may indicate a weakening of indigenous language transmission among younger family members. This shift can be attributed to external influences such as schooling, mass media and socialization outside the home, where English and Nagamese are more commonly used.

2. Group A participants demonstrate a growing preference for English over their indigenous languages on social media platforms. While indigenous languages are still maintained at home, English has become the dominant language in both formal settings and casual interactions, largely due to its practicality and global relevance.
3. Although Nagaland is a multilingual community where many indigenous languages are mutually intelligible, Group A participants rarely use them on social media, even when communicating with those who share the same language. This is one of the most striking findings, highlighting the overwhelming dominance of English in digital spaces, including social media platforms, messaging apps, and other online environments.
4. Participants from Group A reported that they seldom, if ever, use their indigenous languages in digital forums. This is significant, as digital media plays a central role in modern communication for young adults. The near absence of indigenous languages in these spaces suggests that they are at a considerable disadvantage in an increasingly digitalized world.
5. Only 41% of participants feel confident in writing their indigenous language, likely due to the absence of formal education in it. Currently, only four recognized languages of the state are introduced in schools, which limits opportunities for structured learning. As a result, many individuals struggle to develop writing proficiency in their native language.
6. Despite this, an overwhelming majority—89%—agree that their indigenous language is an important part of their identity, and 75% take pride in their linguistic heritage. This strong emotional connection is crucial for efforts to revitalize and sustain these languages.
7. However, the findings also reveal a sense of pessimism about the future of indigenous languages, with 55% of respondents expressing doubt that their language will remain widely used in the coming years. This concern highlights the urgent need for intervention to ensure that the vitality of these languages is preserved.

Group B findings presentation

This group consists of children aged 8 to 12 who are elementary students at a diverse urban school with a middle socio-economic background. The school is known for its cultural diversity, with students from various tribes. A total of 110 Naga students from different tribes participated in the interview questionnaire. The participants speak one of the 18 state-recognised language.

Language Background Table

Language	Converse with Grandparents	Converse with Parents	Converse with Siblings	Converse with Extended Family	Converse with Neighbors
Indigenous Language	89%	82%	70%	82.7%	NIL
Nagamese	2.7%	2.7%	12.7%	3.6%	73.6%
English	8.1%	14.5%	17.5%	13.6%	26.3%

(I) Language Proficiency and Comfort

1. Preferred Language for Expression:

Indigenous Language: 85.4% English: 7% Nagamese: 9%

2. Reading Proficiency in Indigenous Language:

Can Read: 80.9% Cannot Read: 19%

3. Writing Proficiency in Indigenous Language:

Can Write: 69.0% Cannot Write: 30.9%

Key findings from Group A (Ages 8-12)

a. Conversations with Grandparents:

Indigenous Language (89%): This is the predominant language used for communication with grandparents. It highlights the strong cultural and familial ties to the indigenous language, reflecting its role in maintaining traditional values and family heritage.

English (2.7%): Minimal use suggests that English is not commonly spoken with grandparents, possibly due to its lesser significance in traditional family settings.

Nagamese (8.1%): A small percentage indicates that Nagamese is occasionally used, likely reflecting some level of bilingualism.

b. Conversations with Parents:

Indigenous Language (82%): The high percentage shows that the indigenous language remains central in the family environment, reinforcing cultural identity and continuity.

Nagamese (14.5%) and English (2.7%): Consistent with the pattern observed with grandparents, indicating limited use of English and Nagamese in family interactions.

c. Conversations with Siblings:

Indigenous Language (70%): While still significant, the indigenous language is slightly less prevalent with siblings compared to other family interactions, possibly due to the influence of peer interactions or educational environments.

Nagamese (17.5%): Increased use of Nagamese with siblings could indicate its role in informal, peer-level communication among children.

English (12.7%): The highest percentage of English usage in this context suggests that children may be integrating more English into their daily conversations with siblings, reflecting its growing influence in their lives.

d. Conversations with Extended Family:

Indigenous Language (82.7%): Similar to interactions with parents and grandparents, indicating the continued importance of the indigenous language in maintaining connections with extended family members.

English (3.6%): The low usage mirrors the pattern seen in interactions with grandparents and parents.

Nagamese (13.6%): A moderate use of English shows that while the indigenous language dominates, there is some level of integration in conversations with extended family.

e. Conversations with Neighbors:

Nagamese (73.6%): The predominant language used for conversations with neighbors, indicating that Nagamese plays a significant role in community interactions.

English (26.3%): Although used less frequently than Nagamese, English remains an important language in community settings, highlighting the high level of multilingualism among the children. The use of both Nagamese and English reflects a balanced linguistic environment in interactions with neighbors.

(II) Language Proficiency and Comfort:

a. Preferred language for expression:

Indigenous Language (85.4%): A strong preference for the indigenous language indicates a high level of comfort and confidence in using it for self-expression. This preference shows the cultural and emotional significance of the indigenous language to the children.

English (7%): A smaller proportion of children prefer English, which might be due to its role as a second language or its use in educational settings.

Nagamese (9%): Similar to English, Nagamese is less favoured for self-expression, possibly reflecting its more utilitarian role in communication rather than as a primary mode of personal expression.

b. Reading Proficiency in Indigenous Language:

A significant percentage of children (80.9%) responded saying they can read in their indigenous language. The present study did not include testing the reading proficiency but asked the participants to respond to grading their proficiency. Therefore, the study could not assess the different levels of reading proficiency. However, from the responses of the participants, the study cautiously adds that the use of Roman script for their indigenous language and the influential role of the church promoting their language at Sunday schools are some facts that cannot be ignored in promoting reading skills. A notable **(19%)** who cannot read in the indigenous language points to potential areas for improvement in literacy programs and support.

c. Writing Proficiency in Indigenous Language:

The majority of the participants **(69.0%)** say they can write in their indigenous language, there is still a significant gap compared to reading proficiency. This suggests that writing skills in the indigenous language may need more targeted development. A considerable proportion **(30.9%)** of children who cannot write in their indigenous language indicates a need for enhanced educational resources and training to bolster writing skills.

The researcher cautions the responses of Group B as they are all under the age of 12 and therefore their response on their ability to read and write in their indigenous language needs further investigation to attest the responses.

8. Conclusion

The study on the intergenerational transmission of Naga languages highlights a troubling trend in the use of Indigenous languages. While these languages are still actively used within family settings, particularly with grandparents and parents, their use is gradually declining in interactions with siblings, friends, and online. For Group A (ages 20-25), there is a noticeable shift towards English and Nagamese in peer interactions, driven by the global utility of English and the informal convenience of Nagamese. This shift reflects broader societal influences, including educational and professional pressures that favour English, raising concerns about the weakening of the vitality of the language.

In Group B (ages 8-10), Indigenous languages remain prevalent in family conversations but are increasingly overshadowed by Nagamese and English in social settings. The growing use of English, even among young children, highlights its rising importance, while the lower proficiency in writing Indigenous languages highlights the need for improved educational resources. Despite strong intergenerational use within families, the broader decline in Indigenous language use suggests that without proactive measures, these languages may face further attrition.

Comparing Group A and Group B, the study finds the slow decrease in usage of the indigenous language within families and the steady increase of English and Nagamese in all domains when compared.

Therefore, despite strong intergenerational use of Indigenous languages within families, particularly with older generations, their prevalence is declining in broader social and digital contexts due to the increasing dominance of English and other languages. This trend suggests that intergenerational transmission may weaken in the future, potentially leading to language attrition if not actively addressed. To address this, it is crucial to implement revitalization efforts such as increasing the presence of Indigenous languages online, integrating them more thoroughly into educational curricula, and fostering community-based initiatives to encourage their use among younger generations beyond family interactions.

References

Belho, K. (2023). *Naga Tikhir language: An outlook on the challenges to develop Tikhir language* (pp. 17–24). Society for Endangered Language (SEL). ISSN: 2426 2149.

Daimai, K. (2018). *Language endangerment: A case study of Liangmai community in Nagaland*. Paper presented at the 6th International Conference on Endangered and Lesser-known Languages, Mysore.

Krauss, M. (1992). The world's languages in crisis. *Language, 68*(1), 4–10. https://doi.org/10.2307/416368

Temsunungsang. (2022). Language vitality: A case study of Ao students in Shillong. In Y. Yanthan & H. L. Khing (Eds.), *Critical conversations in Naga society* (pp. 202–222). Heritage Publishing House.

UNESCO Ad Hoc Expert Group on Endangered Languages. (2003). *Language vitality and endangerment*. UNESCO. Retrieved from https://www.unesco.org

Phonological Variations between Standard Assamese and Palasbaria dialect

Indrani Gogoi
Linguistics and Phonetics
English and Foreign Languages University
Hyderabad
indraniphdlandp19@efluniversity.ac.in

Abstract
This paper explores the phonetic and phonological variations between Standard Assamese (Eastern dialect) and the Palasbaria dialect. It is a sub-dialect of Kamrupi spoken in the Palasbari region of Kamrup district. Based on field data and previous studies, this study identifies key linguistic patterns that differentiate the two dialects. The analysis highlights phonological processes such as paragoge, where an additional vowel increases the syllable count, and syncope, where unstressed vowels are deleted, reducing syllables. The Palasbaria dialect also shows a tendency to avoid complex codas through consonant deletion.

Additionally, the study examines consonantal variations, including the replacement of /l/ with /n/ in word-initial positions, devoicing of glottal consonants, and the absence of nasal vowels, which distinguishes it from Standard Assamese. These findings contribute to the broader understanding of Assamese dialectal diversity. By documenting these phonological differences, this study emphasizes the importance of preserving dialectal variations and encourages further research into the structural aspects of Assamese and its sub-dialects.

Keywords*: Standard Assamese, Kamrupi, Palasbaria, phonetic and phonological variations, syllable.*

1. Introduction

Assamese is an Indo-Aryan language spoken primarily in Assam. Recognized as one of the 22 languages listed in the Eighth Schedule of the Indian Constitution, its native name is Asamiya (/ɔxɔmija/), while "Assamese" is its Anglicized form. It is the state's official language and also serves as its lingua franca with approximately 15 million speakers (according to the Language Census of 2011).

Assamese has many dialectal variations. It has been broadly classified it into two dialects, namely Eastern and Western dialects (Kakati, 1941:16). However, recent studies have identified four different dialects-

a) Eastern dialect spoken in Sivasagar, Golaghat, Jorhat and nearby areas.
b) Central dialect spoken in Sonitpur, Nagaon and Morigaon districts.
c) Kamrupi dialect in Kamrup district and
d) Goalporiya dialect in Goalporia district.

This paper will focus on the phonetic and phonological differences between the Eastern dialect (which is also the standard dialect) and the Kamrupi dialect with a specific focus on the Palasbaria sub-dialect. The following section provides a description of both dialects.

1.1 Standard/Eastern dialect

The standard dialect of the Assamese language originates from the Sibsagar dialect of Eastern Assam. Before British rule (until 1826), the kingdom of Assam was governed by Ahom kings and the then capital was based in the Eastern district of Sibsagar and later in Jorhat. American missionaries set up the first printing press in Sibsagar and in 1846, published a monthly periodical *Arunodoi*, using the variety spoken in and around Sibsagar as the linguistic foundation. Thus, they started using the local Asamiya dialect for writing purposes. This development led to the acceptance of the formal variety spoken in eastern Assam encompassing most of Upper Assam's districts. Despite this, the language spoken in these regions of Assam exhibits some variation from the written form of the 'standard' language.

While the variety spoken in eastern Assam remains relatively homogeneous, significant linguistic variation can be observed in certain districts of western Assam, extending from Kamrup to Goalpara and Dhubri (see also Kakati 1962 and Grierson 1968).

1.2 Kamrupi dialect

Kamrupi dialect, once considered prestigious, is spoken in the Kamrup region. It has three sub-dialects namely Barpetia dialect, Nalbaria dialect and Palasbaria dialect.

During medieval times, Kamrupi was used in the Brahmaputra Valley and its surrounding regions for literary purposes in parallel with Sanskrit, both in prose and poetry. Assamese, arriving from the west through North Bengal, spread into Kamrup or Western Assam. The first Aryan language spoken in Assam was Kamrupi, which was prevalent in Rangpur, Cooch Behar, Goalpara, Kamrup district, and parts of Nowgong and Darrang. Early Assamese literature was primarily written in this Kamrupi language.

The Ahoms, originally Shan invaders who entered Assam from the northeast, established their dominance in eastern Assam, making Sibsagar their capital. As a result, the literary hub shifted from western Assam to the Ahom court in the east. When Assam came under the control of the East India Company, Sibsagar became the focal point of their activities, and its dialect was adopted for literary purposes.

This paper specifically examines the Palasbaria dialect, a sub-dialect of Kamrupi, spoken in the Palasbari region of Kamrup district. The data presented here is primarily gathered from field visits, supplemented by words from previous research. Information was obtained through discussions with native speakers, who provided insights and confirmation regarding the language in question.

2. A comparative analysis between two dialects

These dialects exhibit numerous phonetic, morphological, and phonological differences. Let us first examine the phonetic variations followed by the phonological variations and make a comparative analysis.

2.1 Phonetic differences

A few consonantal changes can be observed in the Palasbaria dialect in comparison to the standard dialect:

2.1.1 /l/ → /n/

The alveolar nasal is seen to be replaced by the alveolar lateral approximant in the word initial position.

Standard Assamese	Palasbaria	Gloss
nemu	*lebu*	lemon
nu:m	*lu:m*	body hair

Table 1: Consonantal changes in Palasbaria dialect

2.1.2 Devoicing

Devoicing, in which a voiced consonant changes to a voiceless one, occurs in the Palasbaria dialect. This process is seen in the case of words with glottal sounds at the word initial position.

Standard Assamese	Palasbaria	Gloss
bʰah	*pʰah*	to float
gʰoh	*kʰoh*	to rub

Table 2: Devoicing in Palasbaria dialect

2.1.3 Absence of nasal vowels

Nasal vowels are seen to be absent in the Palasbaria dialect as opposed to the Standard Assamese dialect. In the initial and medial positions, it tends to denasalize the nasal vowels. However, this needs further investigation.

Standard Assamese	Palasbaria	Gloss
pãs	*pas*	five
bãhi	*bahi*	flute
kãh	*kah*	cough
xõsa	*xosa*	truth
guxãi	*guxai*	god
dʰũwa	*dʰɔha*	smoke

Table 3: Denasalization of vowels in medial position in Palasbaria dialect

In the word final position, the nasal vowel is replaced by the velar nasal /ŋ/.

Standard Assamese	Palasbaria	Gloss
baũ	*baŋ*	left
saũ	*saŋ*	to see
paũ	*paŋ*	to get
gaũ	*gaŋ*	village

Table 4: Denasalisation of vowels at the final position in Palasbaria dialect

2.2 Phonological differences

The phonological differences observed between these two dialects primarily stem from variations in syllable structure. Hence, before exploring these differences, it is essential to first understand the syllable structure of Assamese.

2.2.1 Syllable structure in Assamese

In Assamese, a syllable consists of a peak or nucleus, with a single vowel (V) or a vowel cluster (VV). It can have an onset with one to three consonants (C) and a coda that includes up to two consonants (Goswami, 1966:107). No consonant is found in the nuclear position in an Assamese syllable. Similar to other spoken languages, the CV structure is the least marked syllable type and is widely found in Assamese and its dialectal vocabulary. In Assamese syllable structure; both onsets and codas can be optional. Examples of syllable patterns include CCCV, CCVV, VC, and VVC. Though the system allows codas, many words primarily follow a simple pattern of repeated CV syllables. Onsets are strongly preferred, especially in non-word-initial positions, whereas codas are less required, particularly in non-word-final positions (e.g., */akax/* 'sky' and */urəli/* 'sky full of stars'). This suggests that V syllables commonly appear at the beginning of words, while CVC syllables are more frequent at the end. Assamese allows complex onsets but not complex codas, as seen in structures like CCCV and CCVC, whereas forms such as CCVCC or CVCC do not occur. Additionally, the coda position is marked, with fewer segments and less variety in clusters compared to the onset position.

The following are some syllable-based rules observed in these dialects:

2.2.2 Paragoge (Word-final vowel addition)

Paragoge is a type of epenthesis in which a sound is added at the end of a word, resulting in an increase in syllables.

CVC → CV.CV

In the Palasbaria dialect, certain words have a higher syllable count than in Standard Assamese (Table 5). This occurs by adding the vowel /-a/ at the word-final position, splitting a single syllable into two.

Standard Assamese	Palasbaria	Gloss
gat	*ga.ta*	hole
pʰot	*pʰo.ta*	bindi
xu.kan	*xu.ku.na*	dry
pok	*po.ka*	insect
kʰel	*kʰe.la*	game
rup	*ru.pa*	silver
xun	*xu.na*	gold
tʰup	*tʰu.pa*	bunch

Table 5: Paragoge in Palasbaria dialect

However, this pattern is not observed in all CVC words. Only a few of the words exhibit this pattern as shown in Table 6.

Standard Assamese	Palasbaria	Gloss
kʰɔn	*kʰan*	classifier
gɔs	*gas*	tree
mɔs	*mus*	to wipe
lɔg	*lag*	companionship

Table 6: Exceptions of paragoge in Palasbaria dialect

2.2.3 Syncope

Syncope is the deletion of an unstressed vowel or reduced consonant from the middle of a word, resulting in a reduction in the number of syllables.

CV.CV.CV → CVC.CV

In the Palasbaria dialect, the unstressed vowel in the second syllable is often deleted. As a result, the onset of the second syllable shifts to become the coda of the first (Table 7).

Standard Assamese	Palasbaria	Gloss
xi.di̲.na	*xid.na*	that day
xe.me̲.ka	*xem.ka*	damp
nɔ.kɔ̲.ra	*nɔk.ra*	do.NEG
ha.lo̲.dʰi.ja	*hal.dʰi.ja*	yellow
kɛ.nɛ̲.kɛ	*kɛŋ.ke*	how

Table 7: Syncope in Palasbaria dialect

2.2.4 No complex codas in Palasbaria dialect

The Palasbaria dialect doesn't prefer complex codas unlike Standard Assamese, where complex coda is sometimes preferred. Therefore, to achieve this, they undergo a process of deletion (Table 8).

Standard Assamese	Palasbaria	Gloss
kand	*kan*	to cry
pindʰ	*pinʰ*	to wear
randʰ	*ranʰ*	to cook

Table 8: Deletion in Palasbaria dialect

3. Summary and conclusion

This study highlights the phonetic and phonological differences between the Standard Assamese (Eastern dialect) and the Kamrupi dialect, with a specific focus on its Palasbaria sub-dialect. The comparative analysis reveals distinct patterns in syllable structure, consonantal changes, and phonological processes that differentiate Palasbaria from the standard variety.

One of the key findings is the presence of **paragoge**, where an additional vowel is inserted at the word-final position, increasing the number of syllables in certain words. Conversely, **syncope** is also observed, wherein unstressed vowels are deleted, reducing syllable count. Additionally, Palasbaria does not prefer complex codas, leading to the deletion of coda consonants to maintain simpler syllable structures.

At the consonantal level, several notable changes are observed in the Palasbaria dialect, including the **substitution of /l/ with /n/** in word-initial positions, **devoicing of glottal consonants**, and the **absence of nasal vowels** in various positions. These changes further establish the unique phonological identity of the Palasbaria dialect within the broader Assamese linguistic landscape.

Overall, this study contributes to a deeper understanding of Assamese dialectal variation, emphasizing the need for further research into regional linguistic diversity. Future studies may explore morphological and syntactic differences, as well as sociolinguistic factors influencing dialectal shifts. Such research would not only enrich Assamese linguistics but also aid in the preservation and documentation of dialects that form an integral part of Assam's linguistic heritage.

References

Dutta, H. 2012. Spirantization and the Notion of Phonological Strength Relations in Assamese: An Optimality Theoretic Account. *Journal of Universal Language,* 13, 71-99. 10.22425/jul.2012.13.1.71.

Goswami, G. 1966. *An Introduction to Assamese Phonology*. Deccan College Postgraduate and Research Institute. https://books.google.co.in/books?id=HkkwAAAAMAAJ

Goswami, U. 1970. A Study on Kamrupi: A Dialect of Assamese. Department of Historical and Antiquarian Studies, Assam.

Goswami, G. C. 1982. The Structure of Assamese. Guwahati, Guwahati University, Department of Publication.

Grierson, George A. 1968. Indo-Aryan Family, Mediate Group, Specimens of the Eastern Hindi Language. (Linguistic Survey of India, 6.)

Kakati, B. 1941. Assamese: Its Formation and Development. Guwahati, L. B. S. Publication.

Kakati, B., & Goswami, G. 1962. *Assamese, Its Formation and Development*. Assam, Lawyer's Book Stall. https://books.google.co.in/books?id=u0B-zwEACAAJ

Mahanta, S. 2001. Some Aspects of Prominence in Assamese and Assamese English. M.Phil. Dissertation. Hyderabad, Central Institute for English and Foreign Languages (CIEFL).

McCarthy, J. and Prince, A. 1986. Prosodic Morphology. Amherst, University of Massachusetts.

Medhi, K. 1988. Assamese Grammar and Origin of the Assamese Language. Guwahati, Publication Board, Assam.

Sarma, M. K. 2006a. Phonological Study of Syllables in Assamese and Assamese English. M.Phil. dissertation, Hyderabad, The English and Foreign Languages University.

Sarma, M. K. 2006b. 'Stress Driven Deletion in the Barpetia Dialect of Assamese'. Paper presented at the 26th South Asian Language Analysis Annual Conference (SALA), Mysore.

Sarma, M. K. 2009. Some Aspects of the Phonology of the Barpetia Dialect of Assamese. *North East Indian Linguistics*, *2*, 96-115.

Spencer, A. 1996. Phonology – Theory and Description. Oxford, Blackwell.

SEL

Society for Endangered
and lesser known Languages

www.selindia.org

Society for Endangered and Lesser known Languages
Bharat Ashram, Raja Bazar, Lucknow-226003

Email: contactselindia@gmail.com | Phone: +91 9940138120