

1976 07250248
E-4941
C0448911 321001
1177-3914
D002105
E4022111

North East Indian Linguistics

Edited by

Stephen Morey • Mark Post



East
Linguistics Society

FOUNDATION[®]
BOOKS

Contents

Introduction	v
Foreword	ix

Phonology

1	Kurtöp Phonology in the Context of North East India <i>Gwendolyn Hyslop</i>	3
2	Working with Tones in North East India – The Tonal System of Numhpuk Singpho, Assam <i>Stephen Morey</i>	26
3	Tonality and the Analysis of Sub-Minimal Words in Ao <i>T. Temsunungsang</i>	45
4	Change with Continuity in the Boro-Garo Languages <i>Robbins Burling</i>	65

Lexicon

5	Lexicon in a Contact Language: The Case of Bishnupriya <i>Shobha Satyanath and Nazrin B. Laskar</i>	75
6	Shafer's 'parallels' between Khasi and Sino-Tibetan <i>Gérard Diffloth</i>	93

- 7 An Exploratory Study of the Terms of Relationship in
Khamti of Lohit district in Arunachal Pradesh
Bishakha Das 105
- 8 Verbs of Position, Existence, Location and Possession and
their Grammaticalization Pathways in the Tani Languages
Mark Post 127

Morphology, Syntax, and Semantics

- 9 Causative Prefixes in Four Boro-Garo Languages
U.V. Joseph 153
- 10 Similarities in Verbal and Nominal Morphology In Atong
Seino van Breugel 168
- 11 Temporality in Bishnupriya
Nazrin B. Laskar 191
- 12 Explicator Compound Verbs in Assamese and Kashmiri:
A Comparative Analysis
Dipima Buragohain 203
- 13 Explorations in the Nonfinite Verbal System in Asamiya
Runima Chowdhary 221

Language Description and Language Endangerment

- 14 Doing Fieldwork on the Singpho Language of
North Eastern India
Palash Kumar Nath 243
- 15 Quest for a Script
Madhumita Barbora and Mark Post 255

Kurtöp Phonology in the Context of North East India*

Gwendolyn Hyslop
University of Oregon

I. Introduction

Kurtöp is a Tibeto-Burman language of Northeastern Bhutan which has remained only marginally described in the literature (e.g. Michailovksy and Mazaudon 1994; van Driem 1995, 2001). This paper presents a phonemic analysis of Kurtöp and situates the analysis within the context of its neighbors in North East India. By offering a phonological analysis of Kurtöp, this paper contributes to documentary endeavors in Tibeto-Burman linguistics. By situating the phonological description in the context of North East India, this paper further begins the dialogue necessary to ascertain to what extent the languages of North East India have been influenced by each other, or to what extent the area can be considered a linguistic area, and therefore how such influence may impact the historical classification of these languages.

* I am grateful to many people for their assistance with this project. First and foremost I owe my gratitude to Pema Chhophyel and Kezang Wangchuk for patiently sharing their language with me. This work has benefited greatly from conversation with Spike Gildea, Scott DeLancey, Susan Guion, Chris Doty, Racquel Yamada, Jesse Blackburn-Morrow, Michael Ahland, Colleen Ahland, John Busch, Brian Bird and Kun Yue. Travel to the inaugural meeting of the North East Indian Linguistic Society was made possible by a grant from the Center for Asian and Pacific Studies at the university of Oregon and financial assistance from the Aienla Project. Lastly, I am thankful to the participants of NEILS for their comments on this paper.

4 North East Indian Linguistics

As descriptions of previously unstudied languages become available, it is fruitful to maintain theoretical concerns with the new descriptions in mind. Often, an immediate concern for the field of historical linguistics is the historical placement of a newly described language. As many scholars have noted (e.g. Masica 1976, Campbell 2001 [1998], Enfield 2005), areal linguistics is crucial to a thorough understanding of historical linguistics, and therefore the placement of a particular language within its family. That is, in order to confidently situate languages within sub-families, and reconstruct proto-languages, it is imperative to discern borrowings and areal influences from innovations. As Campbell (2001 [1998]:299) states 'in order to reconstruct proto-languages accurately or to determine family relationships, it is necessary to distinguish material which is borrowed from that which is inherited from a common ancestor.'

Examining a certain geographic area in terms of a 'linguistic area' is a practical requisite to discerning borrowings from areal influence. Masica (1976) focuses his study of South Asia as a linguistic area on syntactic features but includes phonological features in the inventory of proposed Indian (South Asian) areal features. One of the defining phonological features he lists is retroflex consonants, which, as he points out, are lacking in Assamese.

That North East India enjoys a thriving mix of linguistic diversity is obvious at first blush. The area, bordered by West Bengal in the west, Bhutan and Tibet in the north, Burma in the east and Bangladesh in the south, is home to approximately¹ 136 languages, coming from Indo-Aryan (16), Tai-Kadai (4), Dravidian (1), Austronesian (7), and Tibeto-Burman stock (106), plus at least two creoles. Clearly, as Burling (2003a:169) notes, North East India is 'even by Indian standards very high in linguistic and ethnic heterogeneity.' In his article justifying the term *Kamarupan*², Matisoff (1999: 179) relays the observation from Dr. Dipankar Moral that 'large scale diffusion has been taking place between Asamiya³ and the TB languages of the area for several centuries, resulting in a Tibeto-Burmanisation of Asamiya and the Indo-Aryanisation of the TB languages of the Brahmaputra valley.' He adds that 'the resultant typological characteristics

¹ Based on an examination of the data for India presented in the Ethnologue (Gordon 2005).

² Matisoff uses the term *Kamarupan* to refer to the Tibeto-Burman languages of North East India and adjacent areas of Western Burma, Bangladesh and Tibet.

³ Asamiya is an alternate name for Assamese.

of the languages of this valley have yet to be studied in detail.' Indeed, there must be a complex and intricate history of contact phenomena which this paper cannot fully explore. However, it is my hope this paper will contribute to the academic discourse necessary to sort out the relationship of these languages to each other, and to the other languages within their respective families.

While some scholars have already contributed to the field by offering a classification of the Tibeto-Burman languages of North East India, (e.g. Burling 2003a) no study to my knowledge has attempted to examine the languages in North East India in terms of areal features. While this paper cannot purport to fully ascertain whether or not north East India can be considered a linguistic area, separate from South Asia in general, it hopes to begin the dialogue necessary to eventually fully appreciate and comprehend the linguistic history of the area. By comparing Kurtöp phonology with that found in the languages of North East India, a number of possible hypotheses could emerge. On the one hand, we may see that Kurtöp shares a great deal in terms of phonology with the languages of North East India, leading us to further pursue the hypothesis that Kurtöp, together with the languages of North East India, could form a linguistic area. Alternatively, we may see that the languages of North East India are more similar to each other, and deviate from the phonology of Kurtöp, encouraging the pursuit of the hypothesis that North East India, as it is defined geographically, is a linguistic area. Finally, a third hypothesis may be that North East India, and the surrounding area, including eastern Bhutan, do not form a linguistic area. Of the three hypotheses this paper will suggest we can continue to pursue the second hypothesis. That is, this paper will support the hypothesis that North East India is a linguistic area and Kurtöp, a language of Bhutan, does not belong within it.

In order to consider potential areal features, in this paper, Kurtöp phonology will be compared⁴ with that of 11 languages⁵, representing four of the five language families spoken in North East India. Of the Ti-

⁴ Unless otherwise noted, the data for each language come from the source mentioned in this section.

⁵ While the both the Tani languages and the Tai languages of Assam are collections of languages, rather than one language, I am treating them as one source of comparison for the purposes of this paper. There is one main reason for this. Rather than arbitrarily pick one language from the Tani or Tai group, I can provide a broader comparison by offering comparison from more than one Tani or Tai language, when relevant.

betto-Burman family, Tshangla (Andvik 1999), Garo (Burling 2004), Ao (Coupe 2006), Meithei (Chelliah 1997), Deuri (Jacquesson 2005), Mizo (Lorraine and Savage 1898, Weidert 1975 and Chhange 1993), and the Tani languages (Sun 1993) are represented in this comparison. One Indic language, Assamese (Goswami 1978), is considered. The Tai languages of Assam (Morey 2005) and the Mon-Khmer language Khasi (Rabel 1961) are also included in this comparison. In addition to being well-described, these languages represent six of the Northeast's seven states: Arunachal Pradesh (Tshangla and Tani), Nagaland (Ao), Manipur (Meithei), Mizoram (Mizo), Meghalaya (Khasi, Garo⁶) and Assam (Assamese, Deuri). A more thorough examination of Northeastern Indian Phonology would include a consideration of the Dravidian languages in the area, such as Kurux, as well as a language from Tripura, such as Darlong.

The paper is organized as follows. Section two describes background information on Kurtöp, including speakers and methodology for the data used in this study. Section Three examines syllable shapes and section four examines consonants. Section five considers vowels, section six discusses tone, and I offer a summary and conclusion in section seven. Each section begins with an analysis of Kurtöp, and then proceeds with a comparison of the phenomenon in the other languages. At the end of most sections I include a chart which summarizes the findings. In the final summary and conclusion I will suggest that North East India may be defined as a linguistic area, in terms of phonology, by the lack of retroflex stops, but more work will be needed to show this definitively.

2. Kurtöp

Kurtöp is a Tibeto-Burman language spoken in Dungkar, Bhutan which lies within the political district of Lhüntse, approximately 50 kilometers west of the border with Arunachal Pradesh in India, and 15 kilometers south of the border with Tibet, shown in Figure One below. Van Driem (1995) estimates there are 10,000 speakers of Kurtöp.

Kurtöp has been the focus of one study (Michailovsky and Mazaudon 1994) and is mentioned in van Driem (2001, 1995). The current findings deviate slightly from Michailovsky and Mazaudon (1994), as will be elucidated throughout the paper.

⁶ The description of Garo in Burling (2004) is primarily representative of the varieties spoken in Bangladesh, rather than Meghalaya.

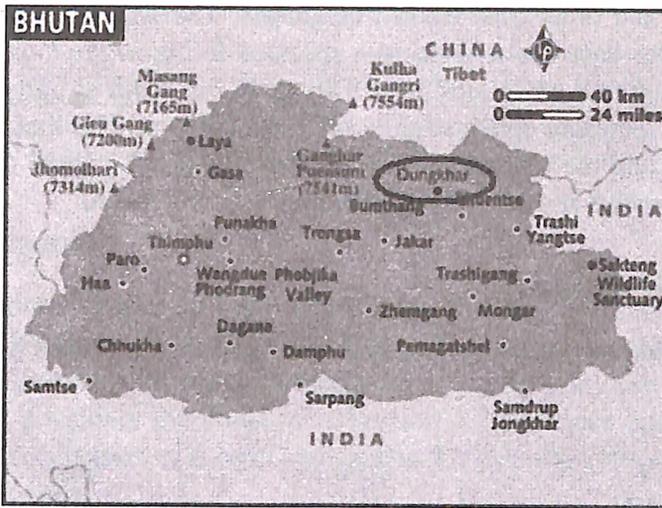


Fig. 1: Dungkar in Bhutan⁷

In terms of genealogy, Kurtöp is considered an East Bodish language. Shafer (1954) appears to be the first to use the term 'East Bodish'. For him the term represented the proto-language from which Dwags, a language spoken southeast of Lhasa, had come.

In his overview on the languages of Bhutan, van Driem (1995) situates eight of Bhutan's 19 languages within East Bodish, including Bumthang, Kheng, Nyenkha, Chali, Dzala, Monkha, Dakpa and Kurtöp. Bradley (1997) proposes that East Bodish is most closely related to Central Bodish (i.e. the Tibetan dialects). In addition to the languages mentioned above, he includes Sherdukpen and the somewhat ambiguous 'Eastern Monpa' in East Bodish. This group joins with its closest genealogical neighbors, the Central Bodish languages. Central and East Bodish together are coordinate with Western Bodish, and the Bodish family then joins with Tshangla and West Himalayan. These three together comprise one side of the Bodic family. A comparison of lexical items in Dwags (Shafer 1954) with the corresponding cognates in Kurtöp indicates the two languages are closely related. Van Driem (1995, 2001) confirms the placement of Kurtöp in the East Bodish branch. The exact position of East Bodish within Tibeto-Burman, however, remains subject to debate. Hyslop (to appear) offers further background on the potential genealogy

⁷ Map downloaded from <http://www.lonelyplanet.com/worldguide/destinations/asia/bhutan>. Reproduced with permission from the Lonely Planet website www.lonelyplanet.com (c) 2006 Lonely Planet Publications.

of Kurtöp and other East Bodish languages. The Kurtöp data discussed in this paper have come from two speakers living in the United States. Both consultants are males from Dungkar, but from slightly different areas, and therefore speak slightly different dialects. The first speaker is from Tabi and the second speaker is from Thuke. The small differences in their speech varieties will be elaborated as is relevant.

3. Syllable Shapes

Kurtöp syllables take nine possible shapes, as exemplified by the data in (1).

(1)

V	íp ^h a	'food'
VV	ée	'who'
VC	ím	'hide.IRR'
CV	ba	'target'
CVV	kóo	'hoe'
CVC	gor	'rock'
CCV	ba	'ash'
CCVV	bra	'scratch.IRR'
CCVC	p ^h rúm	'cheese'

Vocalic initials are often realized with a preceding glottal stop, if the following vowel has a high tone, or with a glide if the following vowel has a low tone. In a sense, then, it may be more accurate to state that Kurtöp prefers syllables with an onset consonant.

The Kurtöp syllable structure is remarkably similar to that in languages of North East India. Tani, Meithei and Phake Tai also report CCVC as the maximum syllable shape (excluding notions of tone). Wiedert (1975) also gives examples of CCVC syllables in Mizo. Khasi syllables are maximally CCV[V]C with one example of an onset consisting of three consonants. Tshangla syllables are slightly simpler in that the second member of an onset cluster is limited to /r/. Garo syllables are slightly more complex, allowing complex onsets with three consonants, and Ao syllables are simpler, having a maximum syllable shape of CVGC. Of the languages considered in this paper, Deuri seems to have the simplest syllable structure with approximately 90% of the syllables being open and no consonant clusters found as syllable onsets or codas. Table 1 summarizes the maximal syllable type for Kurtöp and the languages mentioned above.

Table 1: Maximum Syllable Shapes

Kur.	Tsha.	Tani	Deu.	Ao	Mei.	Mizo	Khasi	Garó	P. Tai	Asm.
CCVC	CrVC	CCVC	CVC	CVGC	CCVC	CCVC	CCVVC	CCCVC	CCVC	CCCVCC

4. Consonants

Table 2 shows the Kurtöp consonants using the IPA. Each of the consonants in Table 2 will be considered as both an onset and coda in this section. Beginning with onsets, section 4.1 examines the obstruents and section 4.2 considers the sonorants. Section 4.3 provides examples of possible onset clusters in Kurtöp and section 4.4 discusses the consonants as syllable codas.

Table 2: Kurtöp Consonants

	labial	dental	retroflex	palatal	velar
stops	p, p ^h , b	t, t ^h , d	ʈ, ʈ ^h , ɖ	c, c ^h , ɟ	k, k ^h , g
affricates		ts, ts ^h			
fricatives		s, z		ç	
nasals	m	n		ɲ	ŋ
laterals		l, ɭ			
rhotics		r			
glides	w			j	
aspirates					h

4.1 Obstruents

A three-way contrast in voicing is found for all stops. Voiceless unaspirated, voiceless aspirated and voiced stops are found at the labial, dental, retroflex, palatal and velar places of articulation. High tone is found following the voiceless and aspirated obstruents while low tone is found following the voiced obstruents. There is one exceptions to this generalization; the palatal fricative occurs before both high and low toned vowels.

An example of the contrast amongst the labials is shown in (2) below. The aspirated labial is often realized as [pʰ] or [f].

(2)	p	p ^h	b
	pá	p ^h át	ba
	'slice of meat'	'okay'	'target'

The data in (3) illustrate the contrast at the dental place of articulation. Unlike the labial and velar aspirated stop, the aspirated dental stop does not spirantize.

(3)	t	t ^h	d
	tá	t ^h á	da
	'axe', 'horse'	'weaving pattern'	'fornicate'

The Kurtöp retroflex series is a recent innovation, coming from a series of velar or dental plus rhotic onset clusters. For example, Written Tibetan has སྩ *sgro* 'feather', རྫོང *grong*⁸ 'village' རྩེ *khri* 'throne' and སྨོ *drel* 'mule' where Kurtöp has ɬo, ʈoŋ, t^hi, and ɬe, respectively. The synchronic contrast between the retroflex voiceless, aspirated and voiced retroflex is shown in (4).

(4)	t	t ^h	ɬ
	tá	t ^h áŋ	ɬa
	'flushed'	'climb.vi.'	'good job'

Kurtöp also reports a series of palatal stops, all of which tend to be affricated. The contrast is illustrated in (5) below.

(5)	c	c ^h	ɟ
	cáro	c ^h á	ɟa
	'friend'	'pair'	'tea'

The contrast in velar stops is shown below in (6). Like the labial aspirated stop, the aspirated velar tends to be realized as a fricative [x] or even [h].

⁸ Note that the correspondence between WT and Kurtöp 'village' is not entirely regular; we would expect the Kurtöp form to have maintained the voicing present in WT, rather than become voiceless. The low tone and voiceless initial are likely the result of tonogenesis, though whether the word was borrowed this way, or tonogenesis has happened since the borrowing cannot yet be proved definitively.

(6) k	k ^h	g
ká	k ^h á	ga
'snow'	'language', 'mouth'	'saddle'

Kürtöp utilises two affricates: a voiceless unaspirated and voiceless aspirated. Michailovsky and Mazaudon (1994) report a voiced affricate, though do not offer examples of it. The current study does not find evidence of a voiced dental affricate in Kürtöp. The contrast made by the voiceless unaspirated and voiceless aspirated affricate is demonstrated by the data in (7) below.

(7)	ts	ts ^h
	tsá	ts ^h á
	'nerves'	'salt'

Where Michailovsky and Mazaudon (1994) found four fricatives, the current study finds three. These are demonstrated below in (8).

(8)	s	z	ç
	sá	zɑŋ	çá
	'soil'	'bridge'	'meat'

Michailovsky and Mazaudon (1994), who based their study on data collected in the late 1970s, found Kürtöp to have a voiced palatal fricative in addition to the fricatives displayed above in (8). As will be elaborated in section five, this report, in conjunction with the current findings, suggest that tonogenesis has happened for the palatal fricatives.

In terms of comparison of obstruents, Tshangla is most similar to Kürtöp, reporting the same set of phonemes plus a voiced palatal fricative. Where Tshangla has a voiced palatal fricative, Kürtöp often has a voiceless palatal fricative with a low tone. Compare for example Kürtöp /çor/ to Tshangla /zu/ 'wine' and Kürtöp /çonba/ to Tshangla /zonma/ 'young'⁹.

Deuri, like the Tani languages, exhibits a two-way contrast of voicing in obstruents; no aspiration is reported. The Tani languages considered by Sun (1993) report contrasting stops at four places of articulation, if one considers /ç/ and /ʃ/ to be stops¹⁰. Ao displays a contrast between voiceless aspirated and unaspirated. Garo also reports a two-way contrast in terms of voicing, but with voiced and voiceless aspirated stops. Like

⁹ I am grateful to Erik Andvik for providing these Tshangla data.

¹⁰ Sun (1993:32) reports he uses the symbols /ç/ and /ʃ/ to represent alveolopalatal consonants.

Kurtöp and Tshangla, some Tai languages have a minimal three-way contrast in voicing (Aiton¹¹, Khamyang¹² and apparently Ahom). Both Khasi¹³ and Mizo have a three-way contrast for the labial and dental stops, but a two-way contrast for the velar series. By Chhangte's (1993) account, Mizo contrasts stops at four places of articulation. Khasi reports bilabial 'post-dental' and velar stops, plus a voiced affricate. Meithei and Assamese have the four-way contrast of voice and aspiration that is typical of Indic languages. Kurtöp and Tshangla thus seem unique in contrasting three voicing types at five places of articulation. They are also the only languages to report a series of retroflex stops.¹⁴

Table 3 and 4 below summarize the findings for the stop consonants. Table 3 compares the languages in terms of maximum voicing contrast found on stops, and Table 4 compiles the maximum number of place contrasts in stops made by each language.

Table 3: Maximum Voicing Contrast in Stops

Kur.	Tsha.	Tani	Deu.	Ao	Mei.	Mizo	Khasi	Garó	Asm.	P. Tai
3	3	2	2	2	4	3	3	2	4	3

Table 4: Maximum Place Contrast in Stops

Kur.	Tsha.	Tani	Deu.	Ao	Mei.	Mizo	Khasi	Garó	Asm.	Ai.Tai
5	4 ¹⁵	4	3	3	3	4	3	3	3	4

¹¹ The status of the voiced series in Aiton differs from that found in Kurtöp or Tshangla. Whereas in the latter two the contrast in voicing extends across each of the five places of articulation, the voiced stops in Aiton are found only at the bilabial and dental/alveolar places of articulation. Further, [m] and [n] are variants for /b/ and /d/, respectively. For additional differences and greater detail, see Morey (2005:134-143).

¹² The only voiced stop reported for Tai Khamyang is /b/

¹³ Rabel reports that voiced aspirates are rare; they are found primarily in loan words in initial position. She treats them as consonant clusters for phonological reasons. For more details, see Rabel (1961:6).

¹⁴ Weidert (1975) lists <tr> and <trh> as 'flapped plosive alveolars' in Mizo. Chhangte (1993) describes these segments as palato-alveolar stops, which contrast with the dental stops.

¹⁵ Andvik analyses the palatal series, apparently cognate with the series of palatal stops in Kurtöp, as palatal affricates. If these were included as stops, Tshangla, like Kurtöp, would contrast stops at five places of articulation.

4.2 Sonorants

Sonorants in Kurtöp are found preceding both high and low-toned vowels. The voiceless lateral, found in loans, is an exception to this in that it is found only preceding high tone. Kurtöp has four nasals, one rhotic, two laterals and two glides, which will be shown below.

4.2.1 Nasals

Kurtöp contrasts nasals at the labial, dental, palatal and velar places of articulation, as demonstrated below in (9).

(9)	m	n	ɲ	ŋ
	mar	na	ɲa	ŋáa
	'butter'	'ear'	'fish'	'pillow'

4.2.2 Laterals and Rhotics

Kurtöp contrasts a rhotic and voiced and voiceless alveolar lateral. The rhotic can be realized as a trill, a flap, and more rarely, a retroflex fricative. Bumthang (van Driem 1995), another East Bodish language of Bhutan, reports a more robust system with a four-way contrast in rhotics. Thus, I suspect the wide variation in Kurtöp rhotics is due to a recent merger, though further work is needed to confirm this. In the present sampling of over 1,200 lexical items, the voiceless /l/ is found in only four. Nonetheless, it contrasts with the other liquids, as shown below in (10).

(10)	l	ɭ	r
	la	ɭá	ra
	'mountain'	'deities'	'root'

4.2.3 Glides and h

Kurtöp has two glides, which can precede both high- and low-toned vowels, and an *h* found as a phoneme in a few words. The contrast between each is illustrated in (11) below.

(11)	w	j	h ¹⁶
	wa	jaa	hámpa
	'water tunnel'	'hand'	'shameless person'

¹⁶ Although [h] is sometimes a realization of /k^h/, the contrast between /k^h/ and /h/ is shown by the minimal pair: /k^hámpa/ 'Tibetan people' and /hámpa/ 'shameless person'.

Tshangla again is most similar to Kurtöp in this regard, reporting an identical set of sonorant consonants, including a voiceless lateral found in loans. The Tani languages do not employ a voiceless lateral but otherwise report a set of sonorant consonants identical to those found in Kurtöp. It is also this set that Sun (1993) reconstructs for Proto-Tani. The Tai languages are also remarkably similar to Kurtöp and Tani in this regard. Phake Tai also has four voiced nasals and voiced /r, l, w, j/. Aiton is identical, but Morey does not analyze /w, j/ as phonemes (2005: 143). Khamyang Tai has four nasals plus /l, j/. Ahom likely had the four nasals plus /r, l, j/. Neither Assamese nor Deuri has a palatal nasal or voiceless lateral but are otherwise identical to Kurtöp with regard to sonorant consonants, though for Deuri /l/ and /r/ are in complementary distribution. Garo has more sonorant consonants than Assamese and Deuri in that glottalized versions of the nasals and laterals are also found word-finally. /l/ is found initially only in loan words and, in fact, appears to be in complementary distribution with /r/. Ao and Mizo are set apart by the fact that they have a series of voiceless sonorant consonants in addition to the voiced series. In Ao, the voiced/voiceless distinction is extended further to the glides. Khasi appears to have only a voiced set. Thus, Kurtöp sonorants are very much like those found in North East India, particularly Tshangla and the Tani languages.

Tables 5 and 6 below summarise the comparative findings for the sonorants. Table 5 illustrates the maximum number of place contrast each language maintains for their sonorants, while Table 6 summarises the status of voiceless sonorants in the languages.

Table 5: Maximal Place Contrast for Sonorants

Kur.	Tsha.	Tani	Deu.	Ao	Mei.	Mizo	Khasi	Garo	Asm.	Ai.Tai
4	4	4	3	3	3	3	3	4	3	4

Table 6: Voiceless Sonorants

Kur.	Tsha.	Tani	Deu.	Ao	Mei.	Mizo	Khasi	Garo	Asm.	Tai
marg- inal †	margi- nal †	no	no	full series (plus glides)	no	full series (no glides)	no	no	no	no

4.3 Complex Onsets

In addition to the consonants in Table Two, Kurtöp uses the following clusters as onsets: /pr pc pc^h pr br bʝ bl kw k^hw gw mr mj/. In addition to these, Michailovsky and Mazaudon (1994) found /pl/, which the current study does not find. They also reported palatalized labials which for us are the typologically unusual labial-palatal stop clusters.¹⁷ Each of these clusters is exemplified by the data in (12) below.

(12)	pr	prá	'monkey'
	pc	pcó	'lies'
	pc ^h	pc ^h á	'Bonpa'
	p ^h r	p ^h rúm	'cheese'
	br	braŋ	'fly'
	bl	ble	'four'
	bʝ	bʝa	'ash'
	kw	kwí	'trivet'
	k ^h w	k ^h wí	'dog'
	gw	gwe	'walk.vi'
	mr	mra	'unpolished rice'
	mj	mja	'arrow'

Languages in North East India vary widely with respect to complex onsets. Khasi has what appears to be the richest possibility of complex onsets. Rabel reports 139 possible complex onsets, including numerous clusters for which both members are stops. Garo and Mizo have a number of complex onsets, including a few somewhat unusual for Tibeto-Burman. Complex onsets in Meithei may consist of a voiced unaspirated stop, fricative or voiceless aspirated stop and *l*, *w*, *y* (Chelliah 2003). Phake Tai has clusters mainly in loans and onomatopoeic forms but Aiton Tai has preserved some Proto-Tai clusters, for which /l/, /r/ and /w/ are possible second members. Tshangla has *Cr* as its only consonant cluster although /pç/ is found in a few lexemes. Tani languages vary with respect to possible onset clusters; however, the second member is usually /j/ or

¹⁷ I suspect this is indicative of a sound change in progress: pj, p^hj, bʝ > c, c^h, ʝ, with the labial-palatal stop clusters being an intermediate step. That is, I suspect there was an intermediate stage in this sound change in which the glide had strengthened to a stop but not yet lost the labial initial. For further argumentation of these stop-stop clusters as being an intermediate step in the sound change pj, p^hj, bʝ > c, c^h, ʝ, see Hyslop (to appear).

/r/. Sun also reports clusters as complex as /p^hrj/, /grj/ and /mrj/ in some Tani languages. Bengni reports the largest number of complex onsets in Sun's (1993) 'key' Tani language, which is shown below in Table Seven. Complex onsets are not found at all for Ao or Deuri. Kurtöp, then, appears to fall within the range of what is found within languages of North East India.

Below in Table 7 the findings for complex onsets are summarized. I urge the reader to note that Khasi and Assamese have not been included in the interest of brevity. Those inquiring as to how these 139 possibilities map out are referred to Rabel (1961:21–29) for Khasi and to Goswami and Tamuli (2003) for Assamese.

Table 7: Complex Onsets¹⁸

Initial Place	Kur.	Tsha.	Tani Bengni	Deu.	Ao	Mei.	Mizo	Garó	Ai.Tai
labial	pr, p ^h r, br, bl, pc, pc ^h , bj, mr, mj	pr, p ^h r, br, pç ¹⁹	pj, bj, m, fj			pr, p ^h r, br, pw, (bj, pl)	pr, p ^h r, pl, br, bl, fr, fj, mj	pr, br, mr	pr, p ^h r, pl, (pj)
coronal			rj				sp, st, dr, dj, nj	sp, st, sk, sr, chr, j, spr, skr	(tr)
velar	kw, k ^h w, gw	kr, k ^h r, gr	kj, gj				k ^h r, k ^h l	kr, gr	k ^h r, (kj)

4.4 Codas

Kurtöp utilises the following eight of its consonants as syllable codas, some more robustly than others: /p, t, (k), (s), r, (l), m, n, ŋ, (h)/. Word-finally, /k/ is not usually pronounced, triggering instead a long vowel. Michailovsky and Mazaudon (1994) found a coda /s/ in many environments where we find coda /t/. Nonetheless, a few words in the present sampling have the phoneme as a coda. Coda /l/ is found in one word in the present sampling of over 1,200 words, though occurs frequently in discourse due to apocope of final vowels. For example, /gee-wald/ is usually realised as [geewal]. Note that /h/ has been found as a coda in only three words in the present sampling.

¹⁸ Parentheses around a cluster are indicative of it being found only in loans.

¹⁹ Andvik (1999, 2003) reports this cluster is found only in a few words, such as /pçei/ 'four'.

The possible Kurtöp codas are displayed below in (13).

- (13)
- | | | |
|---|-------------------|-------------------------------|
| p | náp | ‘phlegm’ |
| t | nat | ‘illness’ |
| k | ŋakpa | ‘cold’ |
| s | dispuŋ | ‘two days after tomorrow’ |
| l | p ^h él | ‘flourishing’ |
| r | p ^h úr | ‘stake’ |
| m | ŋám | ‘many’ |
| n | ŋan | ‘cast a spell to kill.IRR’ |
| ŋ | páj | ‘non-irrigated piece of land’ |
| h | tíh | ‘sieve’ |

Kürtöp codas are typical for the area, if not somewhat richer than those found in North East India. In Khasi, codas can be the voiceless unaspirated stops, sonorants or glides. /l/ occurs only rarely as a coda, as does /s, ʃ/. The Tani languages tend to allow /p, t, k, m, n, ŋ, r/ as possible codas, with Padam Adi also allowing /l/. Mizo also reports /p, t, k, n, m, ŋ, ɣ, l, j, w/ as possible codas (Chhange 1993). Tshangla also has coda *s* but

Table 8: Coda Consonants

Kur.	Tsha.	Tani	Deu.	Ao	Mei.	Mizo	Khasi	Garó	Asm.	Ai.Tai
p	p	p		p	p	p	p	p	p, p ^h , b, b ^h	p
t	t	t		t	t	t	t	t	t, t ^h , d, d ^h	t
k	k	k		k	k	k	k	k	k, k ^h , g, g ^h	k
s	s								s, z	
l		l			l	l		l, lʔ	l	
r		r		ɹ		r	r		r	
m	m	m	m	m	m	m	m	m, mʔ	m	m
n	n	n	n	n	n	n	n	n, nʔ	n	n
							ɲ			
ŋ	ŋ	ŋ	ŋ	ŋ	ŋ	ŋ	ŋ	ŋ, ŋʔ	ŋ	ŋ
h									h	
			j				j			
									x	
							w			

no coda *l* (Andvik 2003). Meithei codas are as in Kurtöp, minus *r*, *s* and *h* (Chelliah 2003). Garo can combine a glottal stop with any sonorant as a coda consonant. Of the Tai languages, Aiton Tai allows /p, t, k, m, n, ŋ/ but I could find no examples of the other consonants. Ao codas are almost identical to Aiton Tai, the difference being can being /ɹ/ is also found. Deuri has less possible coda consonants than any language considered here. Only the nasals /m, n, / and /r, j/ are allowed. All Assamese consonants can serve as syllable codas.

5. Vowels

Kurtöp contrasts seven vowels as shown below in Figure Two. The two diphthongs shown below are often realized as the corresponding front rounded vowels, especially for the speaker from Thuke, but see Hyslop (to appear) for a more detailed history of Kurtöp front-rounded vowels.

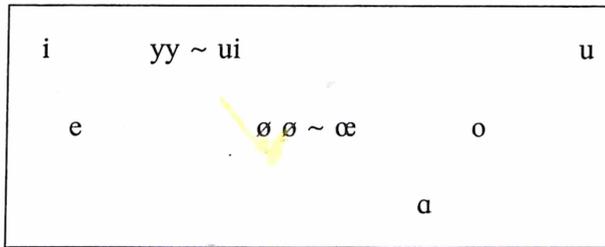


Fig. 2: Kurtöp Vowels

5.1 Monophthongs

The seven Kurtöp monophthongs are shown below in (14).

(14)	p ^h		
	i	p ^h i	'flour'
	y	p ^h ýy	'offer.POL'
	e	p ^h ée	'edge'
	ø	pøø	'incense stick'
	a	p ^h áa	'pig'
	o	p ^h ó	'hole'
	u	p ^h úr	'stake'

5.2 Length

Length is also contrastive on Kurtöp vowels in open syllables. The contrast is not very salient; long vowels are often only 20 ms longer than their short counterparts. Long vowels have been found on the first syllable of bisyllabic words, but no minimal pairs have been found. Examples of the contrast in monosyllabic words follow below in (15).

- (15) kó 'door'
 kóo 'hoe'
 ʔá 'deities'
 ʔáa 'excess'

The contrastive length in Kurtöp vowels seems to have developed from a number of different diachronic sources, the full extent of which has not been explored. One known source, however, is the loss of word-final coda /k/. Where the loss of this segment has conditioned falling pitch and 'checked' tones in other language, it does not appear to be doing so in Kurtöp. While a glottal stop is sometimes an allophone of coda /k/, in natural speech the usual realisation is vowel length with no concomitant glottalization or falling pitch.

None of the languages represented in this paper reports a vowel system identical to that of Kurtöp. Tshangla and Garo both have the five cardinal vowels, as do Khasi, Mizo and Deuri. Khasi and Mizo, however, also contrast long versions of each vowel and Deuri contrasts the nasalized counterparts. Mongsen Ao also has a five vowel system /i, u, a, ə ɤ/, though the varieties spoken in Waromung and Khar villages also

Table 9: Vowels

Kur.	Tsha.	Tani	Deu.	Ao	Mei.	Mizo	Khasi	Garo	Asm.	Kh,Tai
i	i	i	i	i	i	i	i	i	e	i
y										ɯ
e	e	e	e		e	e	e	e	e, ε	e, ε
ø										
ɑ	ɑ	ɑ	ɑ	a, ɤ	ɑ	ɑ	ɑ	ɑ	ɑ	a, aa
o	o	o	o		o	o	u	o	o	o,
u	u	u	u	u	u	u	u	u	u	u,
				ə	ə					ɤ
length		length	nasal.			length	length		nasal.	

have innovated /ɤ/. Meithei has a six vowel system consisting of the five cardinal vowels plus a mid central. The Tani languages typically have a seven vowel system of /a, i, e, o, u, ə, ʊ/ and distinguish length on vowels to varying degrees (Sun 2003). Of the Tai languages in Assam, Aiton Tai has the smallest inventory of /i, ʊ, u, ε, ɔ, a, aa/. Phake and Khamyang Tai both utilize systems of ten vowels, respectively /i, ʊ, u, e, ə, o, ε, ɔ, a, aa/ and /i, ʊ, u, e, ɤ, o, ε, ɔ, a, aa/. Finally, Assamese has an eight vowel system consisting of the cardinal vowels plus /ε, ɔ, ʊ/. Nasalization can be contrastive minimally on the vowels. Thus, Kurtöp is unique when compared to the languages of North East India in that it has a seven vowel system consisting of the five cardinal vowels plus two front-rounded vowels, more reminiscent of Central Bodish. Table Nine summarizes the findings for the vowels. Note that the bottom row of the table includes any additional contrastive features, such as length or nasalization.

5.3 Vowel Sequences

Kurtöp has two diphthongs shown below in (16).

- (16) /au/ c^háuri ‘eleven’
 /iu/ diu ‘bullet’

North East India varies with regard to the possible diphthongs and/or vowel sequences found in its languages. Kurtöp again appears most similar to Tshangla, which has /au/ and /ai/ as possible vowel sequences (Andvik 2003). Deuri has the richest set of 17 possible vowel sequences, Assamese trails slightly with 11 and Aiton Tai reports nine. Phake Tai, on the other hand, allows only a few phonetic diphthongs. A perusal of Weidert (1975) finds seven possible vowel sequences for Mizo: /ui, ou, a:u, oi, e:u, i:u, ui/, but the extent to which these are representative of the diphthongs is not known. For Garo /ai/, /au/ and /oi/ are diphthongs; /eo/ and /ui/ are marginal diphthongs and other vowel sequences are possible (Burling 2004). In Ao, tautosyllabic phonetic diphthongs can arise by way of morphological processes, if the second vowel is high. Sun argues that there is not strong evidence for diphthongs in the Tani languages and that most, if not all, reported diphthongs are subject to morphological analysis. No diphthongs are reported at all for Meithei. The collection of diphthongs for Kurtöp thus seems unique. However, a controlled comparison, at this point, remains difficult until the extent to which diphthongs are differentiated from vowel sequences is clearer.

6. Tone

In Kurtöp, tone (high or low) is contrastive on the first syllable following sonorant consonants and /ç/, as shown below in (17).

(17)	na	‘ear’	newa	‘relatives’
	ná	‘nose’	néwa	‘hell’
	mi	‘person’	ɲa	‘drum’
	míi	‘aim’	ɲáa	‘pillow’
	waŋ	‘hole’	jaŋ	‘stand.vi’
	wáŋ	‘blessing’	jáŋ	‘prosperous in wealth’
	la	‘mountain’	ra	‘root’
	lá	‘month’	rá	‘hair’
	ça	‘what’		
	chá	‘meat’		

As has been shown by the data in sections 3.1 and 3.2, high tone consistently follows voiceless obstruents and low consistently follows voiced obstruents, with the exception of the palatal fricative, which can contrast high or low on its following vowel. Thus I hypothesize that tonogenesis has happened for the palatal fricatives and is in progress for the remainder of the obstruents in the language. A comparison of lexical items with Tshangla, as described in 4.1, also supports this notion.

Tshangla also has a high/low contrast on some sonorant initials and tonogenesis has recently happened in some environments. In these instances, voiced and voiceless contrasts have collapsed to voiceless with either low (previously voiced initial) or high (previously voiceless initial) tone (Andvik 2003). Deuri contrasts high and low tone in a small selection of words and Meithei also has a two-way system, contrasting high and low tone on roots. Some Tani languages lack tone, some report a two-way contrast (high/low), some report a contrast in voice quality, and in one instance three contrasting tonal melodies are found on words (Sun 2003). Ao has a three-level lexical tone system which contrasts high, mid and low tones and Mizo contrasts four tones. Phake and Khamyang Tai, with six lexical tones apiece, have the richest tone system of the languages considered in this study. Aiton Tai, however, has only three tones. Neither Garo (Burling 2004) nor Assamese has tone. Kurtöp stands out as undergoing tonogenesis and having tone already contrastive following sonorant initials and the palatal fricative. However, this description adequately describes Tshangla as well.

Table 10: Tone

Kur.	Tsha.	Tani	Deu.	Ao	Mei.	Mizo	Khasi	Garó	Asm.	Kh.Tai
2	2	<i>varies</i>	2	3	2	4				6

Table 10 summarizes the number of tones found in Kurtöp and the languages to which it has been compared.

Conclusion

The current findings for Kurtöp deviate slightly from what Michailovsky and Mazaudon (1994) found. Instead of a voiced palatal fricative the voiceless palatal is found preceding both high and low toned vowels, the only obstruent to do so. Coda /s/ was only found in a few environments; where Michailovsky and Mazaudon (1994) reported a coda /s/ we found a /t/. Michailovsky and Mazaudon (1994) also recorded /dz/, for which we found no evidence. Instead of /by/ and /py/ as onsets, the current study finds the labial-palatal stop clusters /bj/ and /pc/. The cluster /pl/, reported in the earlier publication, was not found. In terms of comparison, Kurtöp was unique in reporting five contrastive places of articulation for the stop consonants. Further, Kurtöp and Tshangla were the only languages to report a series of retroflex stops. The latter is particularly interesting in the context of areal linguistics. Masica (1976) lists retroflex stops as one of the features which defines South Asia as a linguistic area, while noting they do not occur in Assamese. In fact, we have noted that, with the exception of Tshangla, retroflex consonants are entirely lacking within the geographic confines of North East India.

Kurtöp displays a seven vowel system consisting of the five cardinal vowels plus two front-rounded vowels. These latter two are in variation with diphthongs (/yy/ ~ /ui/, /øø/ ~ /oe/). For the speaker from Tabi, these vowels were much more likely to be monophthongs while the speaker from Thuke was more likely to produce diphthongs for these segments. While languages in North East India may have seven or more vowels, there was not consistency with regard to what the extra vowels would be. Kurtöp was the only language in this study to report front-rounded vowels.

A comparison of other features, such as length, nasalization and tone, also failed to illuminate any features that would be exclusive to the languages of North East India. For some of the languages vowel length or nasalization was contrastive, while for others it was not. Some of the languages of North East India report up to six contrastive tones, while

others have none. Kurtöp, with contrastive vowel length and two contrastive tones following the sonorants and voiceless palatal fricative, was not remarkable in this regard.

As stated from the outset, it is too early to establish North East India as a linguistic area and we are even further from being able to tease apart the relationship of the over 100 languages spoken in this geographic area. However, by comparing Kurtöp phonology in the context of North East India, at least one potential feature has emerged. The fact that Kurtöp and Tshangla have a series of retroflex stops, while the other languages considered in this study do not, may prove to be meaningful. This gains more interest in the light of Masica's (1976) findings which concluded that the presence of retroflex stops was a defining feature of South Asia. Thus, nearly all the languages of North East India, considered in this study, consistently did not meet this criterion established by Masica (1976). Kurtöp, situated geographically outside of North East India in Bhutan, turned out to be a better match to South Asia in this regard. Perhaps more research in the area of areal and historical linguistics of North East India will support the hypothesis that North East India is a linguistic area, apart from South Asia. Further research may support the observation that Tshangla, a language of western Arunachal Pradesh and Eastern Bhutan, would be just outside this potential linguistic area.

Abbreviations

Kur.	Kurtöp
Tsha.	Tshangla
Deu.	Deuri
Mei.	Meithei
P. Tai	Phake Tai
Asm.	Assamese
Ai. Tai	Aiton Tai
nasal.	nasalization
Kh. Tai	Khamyang Tai
IRR	Irrealis
VI	Intransitive verb

References

- Andvik, Erik. 'Tshangla', In Graham Thurgood and Randy J. LaPolla (eds.) *The Sino-Tibetan Languages*, London/New York: Routledge, 2003, 439–455.

- _____. *Tshangla Grammar*. PhD Dissertation, University of Oregon, 1999.
- Bradley, David. 'Tibeto-Burman Languages and Classification', in David Bradley (ed.) *Papers in Southeast Asian Linguistics No 14: Tibeto-Burman Languages of the Himalayas*, Canberra: Australia National University, 1997.
- Burling, Robbins. *The Language of the Modhupur Mandi (Garo)*. New Delhi: Bibliophile South Asia, 2004.
- _____. 'The Tibeto-Burman Languages of Northeastern India', in Graham Thurgood and Randy J. LaPolla (eds.) *The Sino-Tibetan Languages*. London/New York: Routledge, 2003, 169–191.
- Campbell, Lyle. 2001 [1998]. *Historical Linguistics*. Third Edition. Cambridge, MA: MIT Press.
- Chhange, Lalnunthangi. *Mizo Syntax*. PhD Dissertation. University of Oregon, 1993.
- Chelliah, Shobhana L. 'Meithei', in Graham Thurgood and Randy J. LaPolla (eds.) *The Sino-Tibetan Languages*, New York/London: Routledge, 2003, 427–438.
- _____. *A Grammar of Meithei*. Berlin/New York: Mouton de Gruyter, 1997.
- Coupe, Alexander. *A Grammar of Mongsen Ao*. Manuscript. LaTrobe University, 2006.
- Driem, George van. *Languages of the Himalayas*. Amsterdam: Brill, 2001.
- _____. *Een eerste grammaticale verkenning van het Bumthang, een tal van Midden-Bhutan, met een overzicht van der talen en volkeren van Bhutan*. Leiden: Centrum voor Niet-Westerse Studiën, 1995.
- Enfield, Nick. Areal Linguistics and Mainland Southeast Asia. *Annual Review of Anthropology*. 34, 2005, 181–206.
- Gordon, Raymond G., Jr. (ed.). 2005. *Ethnologue: Languages of the World, Fifteenth edition*. Dallas, Tex.: SIL International. Online: <http://www.ethnologue.com/>.
- Goswami, Upendranath. *An Introduction to Assamese*. Assam: Mani-Manik Prakash, 1978.
- Goswami, G.C. and Jyotiprakash Tamuli. 'Asamiya', in George Cardona and Dhanesh Jain (eds.), *The Indo-Aryan Languages*. London/New York: Routledge (Routledge Language Family Series), 2003: 391–444.
- Hyslop, Gwendolyn. to appear. Kurtöp and the Classification of the Languages of Bhutan. *Proceedings from CLS 42*. Volume 2 - South Asian Linguistics, Case and Voice, and Language Coexistence. Chicago: University of Chicago Press.
- Jacquesson, François. *Le deuri: langue tibéto-birmane d'assam*. Leuven, Belgium: Peeters, 2005.
- Lorraine, J. Herbert and Fred W. Savidge. *A Grammar and Dictionary of the Lushai Language (Dulien Dialect)*. Aizawl: Firma KLM Private Ltd., 1898.
- Masica, Colin P. *Defining a Linguistic Area: South Asia*. Chicago/London: University of Chicago Press, 1976.

- Matisoff, James A. In Defense of *Kamarupan*. *LTBA* 22(2), 1999, 173–182.
- Michailovksy, Boyd and Martine Mazaudon. 'Preliminary Notes on the Languages of the Bumthang Group', in Kvaerne, Per (ed) *Tibetan Studies: Proceedings from the Sixth Seminar of the International Association for Tibetan Studies, the Institute for Comparative Research in Human Culture*. Vol. 2, 1994, 545–557. Also available at <http://lacito.vjf.cnrs.fr/publi/kurtoe.pdf>.
- Morey, Stephen. *The Tai Languages of Assam: A Grammar and Texts*. Canberra: Pacific Linguistics, 2005.
- Rabel, Lili. *Khasi, A Language of Assam*. Baton Rouge: Louisiana State University Press, 1961.
- Shafer, Robert. The Linguistic Position of Dwags. *Oriens, Zeitschrift der Internationalen Gesellschaft für Orientforschung* 7, 1954, 348–356.
- Sun, Jackson T.-S. *A Historical and Comparative Study of the Tani (Mirish) Branch in Tibeto-Burman*. PhD Dissertation. University of California, Berkeley, 1993.
- _____. 'Tani Languages', in Graham Thurgood and Randy J. LaPolla (eds.) *The Sino-Tibetan Languages*, London/New York: Routledge, 2003, 456–466.
- Weidert, Alfons. *Componential Analysis of Lushai Phonology*. Amsterdam Studies in the Theory and History of Linguistic Science. Amsterdam: John Benjamins, 1975.