Lexical prefixes and Tibeto-Burman laryngeal contrasts
Author(s): David R. Mortensen
*Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society* (2013), pp. 272-286
Editors: Chundra Cathcart, I-Hsuan Chen, Greg Finley, Shinae Kang, Clare S. Sandy, and Elise Stickles

# Lexical Prefixes and Tibeto-Burman Laryngeal Contrasts

DAVID R. MORTENSEN
*University of Pittsburgh*

## Introduction

The correspondences between Tibeto-Burman onsets are complex and often difficult to explain as products of regular, phonetically conditioned sound change. This is particularly true with regard to their laryngeal features. In his ground-breaking reconstruction of Proto-Tibeto-Burman (PTB), Benedict (1972:17–18) provides the following set of reflexes for his reconstructed PTB stops:

| PTB | Tibetan | Jingpho | Burmese | Garo | Mizo |
|---|---|---|---|---|---|
| *p | $p^h$∼p | $p^h$∼p∼b | $p^h$∼p | $p^h$∼p∼b | $p^h$∼b |
| *t | $t^h$∼t | $t^h$∼t∼d | $t^h$∼t | $t^h$∼t∼d | $t^h$∼t∼$ts^h$ |
| *k | $k^h$∼k | $k^h$∼k∼g | $k^h$∼k | $k^h$∼k∼g | $k^h$∼k |
| *b | b | b∼p∼$p^h$ | p | b∼p∼$p^h$ | b |
| *d | d | d∼t∼$t^h$ | t | d∼t∼$t^h$ | d |
| *g | g | g∼k∼$k^h$ | k | g∼k∼$k^h$ | k |

Benedict reconstructs only two series of stops, despite the fact that there are a far larger number of consonant correspondences. Some of these complexities are the result of regular, phonetically conditioned sound changes, the conditioning environment for which can be readily reconstructed. However, even after these changes are factored out, many of the patterns of correspondence cannot be derived mechanically from Benedict's reconstruction. Benedict is aware of this fact, and proposes that laryngeal features, specifically aspiration, were "unstable" in PTB and the unpredictable reflexes of PTB voiceless stops are a reflex of this variability in the proto-language. A similar thread runs through the work of Matisoff, who maintains basically the same two-way laryngeal contrast as Benedict and views deviation from this contrast as the reflex of synchronic variation in the protolanguage and its successors (Matisoff 2003). Matisoff sometimes abstracts over the causes of this variation, which could include dialect borrowing, analogical processes, and (as Matisoff explicitly notes) the effects of morphological operations.

Some scholars who have worked on Tibeto-Burman reconstruction have questioned whether the Benedict-Matisoff system of two laryngeal contrasts is adequate. Peiros and Starostin (1996) reconstruct no less than four series of obstruents for PTB (voiceless, voiceless aspirated, voiced, and voiced aspirated). More recently, Button (2009:74–75) has questioned whether PTB might have at least had a three-way contrast (voiceless, voiceless aspirated, and voiced) among obstruents. He notes that roots, such as **ka* 'bitter', tend to have aspirated reflexes across the Tibeto-Burman family (and even in Chinese), while others such as **kəy* 'barking deer' tend to have unaspirated reflexes throughout the family (reconstructions from (Matisoff 2003)). Button thus entertains the idea of projecting the three-way contrast found in Kuki-Chin languages back to Tibeto-Burman, but remains uncommitted, noting that the only satisfactory answer may be Matisoff's observation that such apparently unconditioned splits could be due to lexical prefixes which are now lost.

The problem with which Button grapples there is an instance of a more general problem: the interaction between lexical affixes and sound change. LEXICAL AFFIXES are defined here as affixes selecting for bases of a particular, potentially arbitrary class and not performing a well-defined grammatical or semantic function. Similar types of morphology are found in other language families, including the Yuman languages of North America (Miller 2001). The existence of quasi-productive affixes of this kind makes possible what I will call the PHANTOM PREFIX MANEUVER: an affix is added to a stem, provides the conditioning environment for a sound change, and is eliminated (perhaps by the aforementioned sound change). This would give rise to apparently unconditioned splits with comparative morphological evidence. The purpose of this paper is to argue that PHANTOM PREFIX EFFECTS should be expected in Tibeto-Burman based on the properties of lexical prefixes in Tibeto-Burman languages, to show that the phantom prefix hypothesis, while it seems excessively powerful, makes specific predictions, and to contend that these predictions are born out for Tibeto-Burman languages of the Burma-India border area (a superset of the Kuki-Chin languages on which Button (2009) concentrated).

These languages will be referred to here as the Borderlands languages. The specific groups that will be examined, following the classification from Burling (2003), are Kuki-Chin (exemplified by Mizo and Tedim), Old Kuki (exemplified by Sorbung), Tangkhulic[1] (exemplified by Ukhrul Tangkhul and Kachai), Zeme (exemplified by Rongmei and Liangmei), and Angami-Pochuri (exemplified by Khonoma Angami). As Burling points out, a genetic relationship between these languages has not yet been established on rigorous grounds and it is possible that each of these groups (except Kuki-Chin and Old Kuki) form a separate top-level branch of Tibeto-Burman. This paper will advance a contrary view, namely that each of

[1] Burling using the term Tangkhul for the group. I use Tangkhulic here to make it clear that the whole group is intended, rather than a particular language.

these groups except Angami-Pochuri form a subgroup within Tibeto-Burman. All of these languages are spoken in a compact area stretching in India from Mizoram through Manipur to southern Nagaland and including contiguous parts of Burma.

Examples of problematic correspondences among Borderlands languages are given in Table 1, which shows cognates having coronal stop onsets. As shown in

| Gloss | Angami | Rongmei | Liangmei | Ukhrul | Sorbung | Mizo | Tiddim |
|---|---|---|---|---|---|---|---|
| 'custom' | de | -dün | — | — | — | — | da:n |
| 'quiet/cool' | — | — | — | — | dài | dai | da:i$^{3}$ |
| 'dew' | zi | — | — | — | ʔə-dái-cɯ̀ɯ | dai | da:i$^{\Lambda}$ |
| 'stand' | — | diŋ | — | — | — | diŋ | diŋ$^{\Lambda}$ |
| 'chop' | du | — | — | k$^{h}$ə-ru | dùu | — | — |
| 'egg' | dzü | roi-dui | — | ha-ru | ʔə-waa-cɯ́ɯ | ar-tui | tu:i$^{\Lambda}$ |
| 'water' | dzü | dui | ta-duiʔ | ta-ru | cɯ́ɯ | tui | tu:i$^{i}$ |
| 'weave' | do | dük | dak | k$^{h}$ə-rak | — | taʔ | — |
| 'short' | ke-dzü | dui-mei | — | — | tòo | toi | — |
| 'large' | ke-di | dai-mei | ka-di-pu | — | -táa- | — | — |
| 'grandchild' | tsu | — | — | ʔa-ru | — | tu | — |
| 'flesh/body' | — | — | pa-teʔ | — | ʔə-táak | tak-sa | — |
| 'ant' | — | n-tieŋ | ma-tiaŋ | — | ciŋ-ʃii-pá | — | — |
| 'black' | ka-ti | — | ka-tik-bu | kə-tsik | — | — | — |
| 'arrow' | tilüsi | — | — | — | t$^{h}$èe | t$^{h}$al | t$^{h}$al$^{i}$ |
| 'deep/thick' | su | t$^{h}$uk | — | kə-t$^{h}$uk | t$^{h}$ùuk | t$^{h}$u:k | t$^{h}$u:k$^{i}$ |

Table 1: Coronal stop correspondences among Angami-Pochuri (Angami Khonoma), Zeme (Rongmei and Liangmei), Tangkhulic (Ukhrul), Old Kuki (Sorbung), and Kuki-Chin (Mizo [=Lushai] and Tiddim) languages. Angami, Rongmei, Liangmei, and Mizo data are from Marrison (1967); Tiddim data are from Bhaskararao (1996); Ukhrul and Sorbung data are from the author's field notes.

the table, there are four different patterns of correspondence among cognates in this group: one where onsets are voiced across the representative languages, one where they are voiced in Angami-Pochuri, Zeme, and Tangkhulic languages but not in Old Kuki or Kuki-Chin languages, one where they are voiceless unaspirated across the languages, and one where they are voiceless aspirated except in Angami. Comparable sets can be assembled for labial and velar stops as well as two series of affricates (coronal and palatal). If one was to reconstruct a laryngeal contrast for each correspondence set, we would indeed have four series of obstruents, as proposed by Peiros and Starostin (1996). Here, I propose instead that the distinction between the two voiced sets and the two voiceless sets arose due to phantom prefix effects.

## 1 Lexical Prefixes in Tibeto-Burman

Tibeto-Burman lexical prefixes display properties not typical of affixes, to the extent that certain scholars have objected to classifying them as affixes at all (Hill 2009:174). While prototypical affixes mark semantic or grammatical features, TB lexical prefixes sometimes do neither. While prototypical affixes sit on a scale of productivity between fully productive affixes and fossilized formatives that cannot, from a synchronic viewpoint, be analyzed from the stem, TB lexical prefixes display a kind of sporadic productivity. This results in extreme variability even across closely related languages. In fact, one of the most compelling reasons for viewing TB lexical prefixes as prefixes, rather than fossilized parts of the stem, is that they are so readily replaced by other prefixes.

### 1.1 Variability

An inescapable fact of Tibeto-Burman comparative linguistics is the variability displayed in the distribution of lexical prefixes, even among very closely related languages. The Tangkhulic languages, for example, are a compact family which, despite their mutual unintelligibility, are often considered "dialects" of the same language by their speakers. Nevertheless, they show great variation in the lexical prefixes that appear in cognate lexical items: This variation has two implications:

| Ukhrul | Tusom | Huishu | Kachai | |
|---|---|---|---|---|
| kʰə-**ŋə**-ci | kə-tsɨ | kə-tsik | kʰə-**ŋə**-tsɐ | 'fear' |
| **tʰə**-ruk | **tʰ**-ru-he | **sə**-ruʔ | **ʃə**-ruk | 'six' |
| **ŋə**-luŋ | lũ-kuə | **sə**-luŋ | **kə**-luŋ | 'stone' |
| ha | **ŋə**-ʃi | ʔa-**r**-we | **kə**-fu | 'axe' |

Table 2: Examples of Tangkhulic cognates displaying variability in lexical prefixes.

that lexical prefixes are treated as analyzable morphological constituents, not an unsegmentable portion of the stem, and that changes in the lexical prefix do not necessarily imply changes in the grammatical or semantic properties of stems.

### 1.2 Form

Tibeto-Burman words often have an iambic, sesquisyllabic (syllable and a half) structure. In formal terms, TB prefixes are the light syllables or semisyllables in sesquisyllabic words. In the Borderlands languages, there is often an additional grammatical prefix before this lexical prefix in certain syntactic contexts. Words thus tend to have the formal structure shown in (1):

(1) $(\text{C-})_{\text{pfx}_\text{g}}(\text{C-})_{\text{pfx}_\text{l}}[\text{CV(V)(C)}]_{\text{root}}$

a. $pfx_g$: grammatical prefix (nominalizer, possessor, noun marker).
b. $pfx_l$: lexical prefix.

While there are exceptions, the prefixes tend to lack an underlying vowel and to display predictable vocalism (typically a short mid-central or high vowel).

In general, at most one prefix may occupy each "slot." This means that there are a finite number of "classes" to which each stem (lexical prefix + root) can belong: one for each lexical prefix and one for stems lacking any lexical prefix.

### 1.3 Distribution

The distribution of lexical prefixes in Tibeto-Burman languages is not completely arbitrary. Instead, it is multiply determined. The relevant factors can be classified, in order of decreasing prototypicality, as valency, semantic class, and euphony (phonological typicality).

**Valency-changing prefixes** At least since Wolfenden (1929), it has been noted that there are two widespread prefixes in Tibeto-Burman languages which are associated with certain types of argument structure and event semantics. The prefix **s-* was apparently valency-increasing, associated with "outward directed" action, and **m-* was valency-decreasing, associated with stative, reflexive, and reciprocal events. While neither of these prefixes is productive in most modern Tibeto-Burman languages, there is widespread evidence for their past productivity. For example, in Burmese (and Yi-Burmese languages generally) there are numerous causative-simplex pairs where the causative has an aspirated onset and the non-causative has an unaspirated onset (Cornyn and McDavid 1943; Matisoff 1970). On the basis of comparative evidence, the causatives can often be show to reflect stems containing the **s-* prefix (Matisoff 1970, 2003).

Various subgroups of Tibeto-Burman have innovated additional valency-changing prefixes, some of which are productive. For example, in Sorbung (probably a member of the Old Kuki group) and some of the Kuki-Chin languages, there is a productive causative /m-/ (Hartmann 2001; Mortensen 2010). However, even for the most productive cases, there are usually no neat form-meaning mappings: there are typically formally identical prefixes occurring in semantically incompatible varb stems, nouns, and numerals. For example, the productive Sorbung causative *mə-* overlaps formally with class prefixes found in animal names (*məjúu* 'mouse', *mətìr* 'shrew', *məhít* 'leech') and body parts (*məlúuŋ* 'heart', *mətóo* 'lap', *mətín* 'nail/claw') and reflexes of the PTB **m-* verbal prefix (*mənáam* 'smell', *mənúu* 'laugh/smile', *mət$^h$àk* 'be itchy'), which was characteristically valency *reducing*.

**Semantic class prefixes** Other lexical prefixes in Tibeto-Burman languages are vaguely classificatory in function. These have developed, diachronically, from clas-

sifying compounds where the first constituent is (1) a hypernym of both the complement and of the compound as a whole or (2) an incorporated noun classifying the event to which it refers. For example, some instances of the Sorbung prefix *cə-* are transparently derived from the word *cɯ̀ɯ* 'water' in composition:

| | | |
|---|---|---|
| **Water Related** | cəkòo | 'river' |
| | cəlàm | 'spring' |
| | cəlòk | 'flood' |
| | cəkàaŋ | 'drought' |
| | cəluàŋ | 'flow (v.)' |
| **Other** | cəkàap | 'tongs' |
| | cəhàa | 'leaf-monkey; langur' |
| | cəkùi | 'dance' |

Table 3: Sorbung stems with the prefix cə-, most of which are related to water.

Heads which occur in a very large number of compounds are vulnerable to reanalysis as part of the prefix and concomitant phonological reduction.

Other examples of this type include body-parts with the prefix **m-*, probably from **mi* 'person' (Benedict 1972) and Tangkhulic nouns and verbs related to the mouth with the prefix **m-*, probably from **mor* 'mouth' (Mortensen 2003).

**Euphonic prefixes** When valency changing prefixes and semantic class prefixes are factored out of Tibeto-Burman lexicons, a substantial residue remains. These are prefixes that serve no discernible function except to make the form of a stem more like that of other stems in the language. These prefixes are subject to, and the product of, various types of local and non-local analogical processes. The next section provides empirical validation for this claim.

## 2 Influences on Prefix Selection

The category to which a stem belongs—whether it includes a prefix and, if so, which—emerges from a competition among formal (phonological), semantic, and grammatical factors. The observed variability in lexical prefixes is best understood as a result of the indeterminacy of this interaction. In the following section, I describe the result of two experiments conducted on lexical collections from Borderlands languages (both from the Tangkhulic branch) which demonstrate the reality of this interaction.

### 2.1 Experiment 1: Kachai lexicon

The first experiment sought to determine whether formal biases in the distribution of lexical prefixes would be detectable in a list of basic vocabulary items from a

Tibeto-Burman language. I chose to use a word list collected from the Tangkhulic language Kachai, originally compiled for general comparative and descriptive purposes.

**Methods** The Kachai word list consisted of 389 unique lexical items, with a bias towards body part terms, kinship terms, and animal names. Words from the list were coded according the formal factors and lexical category (noun, verb, or numeral). Models predicting the occurance of each prefix via phonological factors (onset, nucleus, and coda of the following syllable) and lexical category were explored using the the multiple logistic regression program *Rbrul* for the R statistics environment. Several factors describing the phonological properties of the onset were coded: labial, coronal, dorsal, laryngeal, sonorant, nasal, plosive, aspirated, and fricative.

**Results** One or more of the phonological properties of the onset of the following syllable was a significant predictor for the presence of each of the prefixes but one. For this prefix (*ŋ*-), lexical category was the only significant predictor. It was much more likely to occur in verbs than in nouns and did not occur in numerals. For all others, formal factors were better predictors. The absence of any lexical prefix was also best predicted by a phonological factor (a following non-coronal onset), but was also weakly predicted by lexical category (verbs and nouns were more likely to lack lexical prefixes than numerals). A summary of the best predictors for each of the prefixes is given in Table 4.

| Prefix | Predictors |
|---|---|
| ∅- | ¬Coronal*** > Lexical Category |
| p- | ¬Plosive*** > Dorsal* > ¬Aspirated |
| m- | ¬Labial** > ¬Fricative** |
| k- | Coronal** |
| ŋ- | Lexical Category** |
| c- | Labial** > Dorsal > Laryngeal |

Table 4: Best predictors for Kachai prefixes. *$p < 0.025$, **$p < 0.01$, ***$p < 0.001$.

**Discussion** Effects of form on prefix selection are robust enough to be identified even in relatively small, heterogeneous word lists. In general, it appears that prefixes are preferentially affixed to roots where the following onset is phonologically different from the prefix consonant. Thus, m- is significantly more frequent in stems were the following onset is not labial. The fact that there are no frequent coronal prefixes explains the fact that non-coronal roots are disproportionately likely to lack

a lexical prefix. On the other hand, simple difference cannot explain all of the patterns observed. For example, the specific affinity between k- and coronals (but not labials) suggests a language-specific set of analogical processes.

### 2.2 Experiment 2: Ukhrul verbs

The preliminary experiment on the Kachai word list validated the idea that phonological factors play a role in prefix selection. However, it was not clear whether this effect was robust throughout the data or was confined to the more opaque prefixes in nouns and numerals. The heterogeneous nature of the data set also made it difficult to code systematically for semantic factors. Looking at a larger data set consisting only of verbs would allow these shortcomings to be addressed.

**Methods** The complete set of verbs from a dictionary of Ukhrul Tangkhul (Bhat 1969) was filtered so it only contained one instance of each attested verb stem. This left a set of 1235 verb stems. Verbs were then coded for both phonological and semantic variables (onset, nucleus, coda, stative, causative, reciprocal, reflexive). Semantic categories were coded according to English gloss. For example, words glossed as "cause to X" or "make X" were coded as causative. Models for predicting each prefix based on these formal and semantic factors were explored using the multiple logistic regression program *Rbrul* for the R statistics environment.

**Results** For each of the four prefixes tested, Rbrul found a model consisting of significant factors. In each of these models, a formal factor—the onset of the root syllable—was a signficant predictor. For two of the prefixes (*m-* and *p-*) the onset of the following syllable was the *best* predictor. Semantic factors were also significant predictors:

| Prefix | Predictors |
|---|---|
| ∅- | ¬Caus*** > ¬Recip** > Onset** > -Reflex* |
| p- | Onset*** > Reflex** > Caus |
| m- | Onset*** > State*** |
| k- | ¬State*** > Onset** |
| ŋ- | Recip*** > ¬Caus*** > Onset |
| c- | Caus*** > Onset** > ¬State** |

Table 5: The best predictors for Ukhrul Tangkhul prefixes. *$p < 0.025$, **$p < 0.01$, ***$p < 0.001$.

**Discussion** The current distribution of prefixes in Standard Tangkhul (Ukhrul) verbs is best understood as a product of an interaction between formal factors and the morphosyntactic/semantic factors that prototypically govern the distribution of

affixes. As a result, the set of stems contain a particular lexical prefix, or no prefix, is phonologically skewed, as is illustrated in Table 6. The patterns observed are those expected from the Kachai study. Prefixes are dispreferred to the extent that they are similar to the initial onset of the root. The absence of any common coronal prefixes probably accounts for the relatively high frequency at which roots with non-coronal onsets occur without lexical prefixes.

| p- | logodds | m- | logodds | c- | logodds | k- | logodds | ŋ- | logodds | ∅- | logodds |
|---|---|---|---|---|---|---|---|---|---|---|---|
| r | 13.752 | $t^h$ | 5.421 | m | 1.179 | n | 1.741 | w | 1.544 | ŋ | 1.501 |
| j | 13.696 | s | 5.263 | p | 1.026 | ts | 1.634 | $p^h$ | 1.350 | $k^h$ | 0.507 |
| ŋ | 13.441 | c | 5.060 | k | 0.785 | p | 1.423 | $k^h$ | 1.226 | m | 0.338 |
| h | 13.286 | r | 5.056 | h | 0.731 | j | 1.350 | s | 1.077 | $p^h$ | 0.299 |
| n | 12.014 | k | 5.047 | ʃ | 0.527 | t | 1.262 | m | 1.025 | w | 0.158 |
| w | -5.339 | ʃ | 5.036 | n | 0.292 | h | 1.075 | k | 0.914 | t | 0.026 |
| ts | -5.373 | h | 5.033 | ts | 0.289 | $t^h$ | 1.041 | p | 0.912 | $t^h$ | 0.007 |
| m | -5.394 | j | 5.011 | $t^h$ | 0.064 | ŋ | 1.033 | j | 0.813 | c | 0.001 |
| t | -5.402 | t | 4.789 | t | -0.056 | c | 1.003 | n | 0.805 | ʃ | -0.038 |
| ʃ | -5.462 | n | 4.706 | w | -0.119 | r | 0.864 | r | 0.769 | ts | -0.093 |
| k | -5.489 | $k^h$ | 4.458 | $p^h$ | -0.131 | s | 0.726 | c | 0.743 | p | -0.197 |
| $t^h$ | -5.518 | ts | 4.323 | $k^h$ | -0.179 | ʃ | 0.662 | ʃ | 0.659 | s | -0.265 |
| p | -5.541 | $p^h$ | -11.709 | s | -0.249 | k | 0.418 | h | 0.632 | k | -0.277 |
| c | -5.560 | p | -11.764 | ŋ | -0.419 | $p^h$ | 0.396 | ts | 0.583 | j | -0.369 |
| s | -5.562 | w | -11.814 | c | -0.438 | w | 0.071 | t | 0.531 | n | -0.508 |
| $p^h$ | -5.608 | m | -11.949 | r | -0.846 | m | 0.060 | $t^h$ | -0.045 | h | -0.522 |
| $k^h$ | -5.939 | ŋ | -11.965 | j | -2.457 | $k^h$ | -14.759 | ŋ | -13.539 | r | -0.566 |

Table 6: Occurrence of prefixes with onsets.

## 3 Predictions of the "Phantom Prefix" Model

Given that the distribution of lexical prefixes is both multiply determined and diachronically unstable, a scenario in which a prefix is added to a root, then eliminated as part of a sound change that affects the initial consonant of the root, is plausible. Such a hypothesis is problematic, though, in that it seems unconstrained in its explanatory power. At first glance, it appears able to explain any development. Applied indiscriminately, one could even use it to reduce the PTB obstruent inventory to a single series, or a single consonant.

However, when the matter is examined more closely, it becomes clear that sound changes triggered by lexical prefixes are likely to have certain properties. The "phantom prefix" hypothesis makes certain predictions which the multiplication of contrasts in the proto-language does not. These predictions are problematic to verify because they tend to be probabilistic rather than categorical, but they are, nevertheless, verifiable. This section identifies several of these predictions and shows that they are consistent with the hypothesis that the three-way laryngeal contrasts found in Borderlands languages is a secondary, prefix-conditioned development.

### 3.1 Localization of deviation

Lexical affixes are most likely to trigger sound changes at the location of attachment. Thus, a language with lexical prefixes would be expected to show more apparently irregular developments in the initial segments of roots than in the final segments. This is difficult to test in a non-circular fashion without a generally agreed-upon reconstruction for the ancestor of the whole language family under discussion and a comprehensive understanding of what developments should be considered "regular." However, as the following sections will make clear, there are various ways in which questions of this type can be investigated. The most important of these are semantically related doublets.

### 3.2 Cross-series patterning of splits

While, as demonstrated, lexical prefixes may have a phonologically biased distribution, they nevertheless tend to occur in a broad range of phonological contexts. This means that changes in laryngeal features triggered by lexical prefixes are likely to occur across whole phonological series, or even multiple phonological series, rather than in a smaller subset of the phonological inventory.

This seems consistent with a development in the Kuki-Chin languages (which is probably shared with at least the Tangkhulic languages (Mortensen and Miller 2009)). In many these languages, there are two series of sonorants: a voiced series and a voiceless series. The Hakha Lai phonological inventory, for example, includes /hm/, /hn/, /hŋ/, /hl/, and /hr/ in addition to /m/, /n/, /ŋ/, /l/, and /r/. (Matisoff 2003) views at least some of the voiceless sonorants in the Kuki-Chin languages as reflecting the fusion of the PTB *s- prefix with the following onset. There are least two reasons for believing this. First, non-Kuki-Chin cognates of some of the etyma displaying this development have prefixes reflecting *s-.

| PTB | Written Tibetan | Lai (Hakha) | Gloss |
|---|---|---|---|
| *s-nap | snabs | hnap | 'snot' |
| *s-ram | sram | -hrem | 'otter' |
| *s-min | smin-pa | hmîn | 'ripe/well-cooked' |

Second, there are numerous causative-simplex pairs which differ only in the voicing of the stem-initial consonant. These appear to reflect stems with and without the historical *s- causative prefix, respectively (VanBik 2009):

| Simplex | | Causative | |
|---|---|---|---|
| maàn | 'crush (v.i.)' | hmàan | 'crush (v.t.)' |
| mit | 'go out (light)' | hmit | 'extinguish (light)' |
| ŋerʔ | 'be entwined' | hŋerʔ | 'entwine' |
| rĭl | 'roll (v.i.)' | hrĭl | 'roll (v.t.)' |
| làaw | 'be alarmed' | hlàaw | 'alarm (v.t.)' |

However, these causative-simplex pairs are not limited to sonorants. There are numerous parallel doublets with voiceless obstruents were the contrast is in aspiration:

| Simplex | | Causative | |
|---|---|---|---|
| pok | 'be open' | p$^{h}$ok | 'open (v.t.)' |
| tolʔ | 'slide (v.i.)' | t$^{h}$olʔ | 'slide (v.t.)' |
| kăaŋ | 'burn (v.i.)' | k$^{h}$ăaŋ | 'burn (v.t.)' |
| kiak | 'break (v.i.)' | k$^{h}$iak | 'break (v.t.)' |
| tsat | 'be severed' | ts$^{h}$at | 'sever' |
| trŭm | 'decrease (v.i.)' | t$^{h}$rum | 'decrease (v.t.)' |

This is especially significant because the aspirated-unaspirated distinction is one of the contrasts that is not directly accounted for by the Benedict-Matisoff reconstruction of PTB. In light of these data, it is likely that at least some of the aspirated onsets and voiceless sonorants in Kuki-Chin languages share a common origin, namely the PTB *s- prefix.

### 3.3 Persistence of semantically related doublets

Unfortunately, not all of the Borderlands language that display unexplained splits present such unambiguous evidence for their origins. However, they do seem to satisfy a more general prediction of the phantom prefix hypothesis, namely that the lexicons of the affected languages should be littered with doublets reflecting pairs of prefixed and unprefixed forms. These should differ either minimally or systematically in semantics and should show the effects of the split conditioned by the prefixes. The causative-simplex pairs from Kuki-Chin languages are a special case of this phenomenon. However, there are many other examples in Borderlands languages.

For example, in Sorbung (Old Kuki) there are two reflexes of PTB *dzy-, /j/ and /c/. These are reflected in the doublet *cúup* 'breast', *jùup* 'suck' from PTB **dzyoːp* 'breast/suck.' This is paralleled by the doublet *cu̍u* 'water', *ju̍u* 'wet.' In Ukhrul (Tangkhulic) there are two reflexes of PTB *tsy-, /s/ and /ts/. This split is reflected in couplets like **sa* 'be hot', **tsa* 'be ill/be feverish' from PTB **tsya* 'hot/hurt/pain/ill.'

An even more pervasive kind of doublet can be identified only comparatively. In VanBik's (2009) reconstruction of Proto-Kuki-Chin, he notes that there are numerous roots for which both an aspirated and unaspirated form, or both a voiced and voiceless form, must be reconstructed. For example, 'foot/leg' must be reconstructed as **kee* based on evidence from Hakha Lai, Falam Lai, and Mizo but as **k$^{h}$ee* based on evidence from Tedim and Paite. Likewise, Mara, Hakha Lai, Falam Lai, Mizo, Sizang, and Khumi suggest that 'hand' should be reconstructed as **kut* while Mindat Cho, Tedim, Thado, and Paite suggest that it should be reconstructed

as *$k^h$ut. While systematic counts have not yet been made, it appears from VanBik's reconstruction that variation in the laryngeal features of initial consonants, as shown in these examples, is the most common type of variation that must be reconstructed for Proto-Kuki-Chin.

### 3.4 Developments with soft phonological and semantic biases

VanBik (2009) notes with some puzzlement that doublets are not evenly distributed across the Kuki-Chin lexicon. For example, there are few causative-simplex doublets with coronal stops, affricates, or nasals, while there are many with labial and dorsal consonants. There are also relatively few instances of /$t^h$/ from PTB *t- (most instances of /$t^h$/ are regular reflexes of PTB *s-). In light of the results from the study of prefixes in Ukhrul Tangkhul, this is likely to be the result of a selectional bias towards bases that are phonologically unlike the prefix. If the aspirated/voiceless members of these couplets are reflexes of stems containing *s-, and if *s- was more likely to be affixed to roots with non-coronal onsets, the resulting aspiration/devoicing would be less frequent in coronals than in non-coronals.

On the other hand, not all voiceless sonorants or aspirated reflexes of PTB voiceless stops can be derived from causative *s-, nor does the phantom prefix model predict that they should originate from the source, except in a strictly formal sense. For example, neither Hakha Lai *$k^h$ăaw* 'grasshopper' from PTB **ka:w* or Hakha Lai *-hni?* 'two' from PTB **g-ni-s* is causative and neither has cognates with the *s- prefix elsewhere in Tibeto-Burman. What the hypothesis does predict is that there will be probablistic biases in the boundaries of the resulting split in both phonological and semantic domains. This is born out in Kuki-Chin, where stems with initial aspirated obstruents and voiceless sonorants are more likely to be causative in their semantics than roots without these properties.

## 4 Two Series are Sufficient

If lexical prefixes may both expand their domain in unpredictable ways and trigger sound changes that result in their own destruction, it is possible to reconstruct only a two-way laryngeal contrast in PTB and still derive the correspondence sets shown in Table 1. Assume that Zeme, Tangkhulic, Old Kuki, and Kuki-Chin share a common ancestor more recent than PTB and that in this ancestor language a set of prefixes that could induce voicelessness/aspiration (including but perhaps not limited to *s-) proliferated. These prefixes then coalesced with following onsets through a regular sound change, yielding the unaspirated/unaspirated split in voiceless stops and the voiced/voiceless split in sonorants that are seen in various languages within this group.

This hypothesis has the added benefit of explaining why segmental reflexes of the PTB *s- prefix do not seem to occur in the languages belonging to this group. It

is true that they display instances of the PTB animal prefix, which Matisoff (2003) also reconstructs as *s-, but these tend not to be fully reduced and include an underlying vowel. In other words, they were likely to be formally distinct from the *s- causative. Matisoff (2003) also suggests that the Ukhrul Tangkhul causative prefix ʃj\i-/ci- is a reflex of the PTB *s- prefix, but on the basis of internal reconstruction it can be shown the palatal stop allomorph is the historically prior form, making it a poor match for PTB *s- (Mortensen and Miller 2009). While other PTB prefixes are well preserved, *s- is not. This makes sense if segmental reflexes of *s- were destroyed as part of the sound change that gave rise to the aspirated/unaspirated split.

Subsequently, after Kuki-Chin and Old Kuki had split from the rest of the group, I propose that there was a second prefix-induced split which resulted in two sets of reflexes for PTB voiced stops. Parallel changes did not occur in Tangkhulic or Zeme, yielding the complicated correspondence seen in Table 1. Unfortunately, little can be said about this prefix currently, or even whether it prevented devoicing, induced voicing, or induced devoicing.

This brings us back to the primary problem of the phantom prefix maneuver: because it relies upon a causal factor that is eliminated by the event that it triggers, it slouches dangerously close to a just-so story. At first glance, it may even seem equivalent to stipulating the existence of one or more additional phonological series. It is a mechanism of massive power that is probably impossible to constrain in strictly categorical terms. In a word, it is an undesirable analytic option. The reason I have argued for it is not that it is inherently desirable, but because it is more desirable than the competing options. Despite its great power, it makes a variety of interesting probabilistic predictions about the lexicons of affected languages. While a great deal of additional work remains to be done, these predictions are generally consistent with my findings for Borderlands languages. Predictions that should be tested, or tested more rigorously, include the following:

(2) a. Doublets that differ in the laryngeal features of initial consonants are significantly more frequent than those that differ in vocalism or the final consonant (excluding independently motivated morphological processes like Kuki-Chin stem alternation).

b. Nouns with initial voiceless sonorants in Kuki-Chin share more formal and semantic properties with nouns having aspirated stops than nouns having unaspirated stops.

c. There are semantic and/or phonological biases distinguishing the two sets of reflexes of PTB voiced stops.

These avenues for future research highlight the greatest advantage of the phantom prefix hypothesis, namely that it invites us to ask more interesting questions than the alternatives.

## References

Benedict, Paul K. 1972. *Sino-Tibetan, a Conspectus*. Cambridge: University Press.

Bhaskararao, Peri. 1996. A computerized lexical database of Tiddim Chin and Lushai. In Tsuyoshi Nara and Kazuhiko Machida, eds., *A Computer-Assisted Study of South-Asian Languages*, 27–143, Tokyo: ILCAA.

Bhat, D. N. Shankara. 1969. *Tankhur Naga vocabulary*, volume 67. Poona: Deccan College Postgraduate and Research Institute.

Burling, Robbins. 2003. The Tibeto-Burman Languages of Northeastern India. In Graham Thurgood and Randy J. LaPolla, eds., *The Sino-Tibetan Languages*, 169–191, London: Routledge.

Button, Christopher Thomas James. 2009. A Reconstruction of Proto Northern Chin in Old Burmese and Old Chinese Perspective. PhD dissertation, SOAS, University of London.

Cornyn, William and Raven McDavid. 1943. Causatives in Burmese. *Studies in Linguistics* 1(18):1–6.

Hartmann, Helga. 2001. Prenasalization and preglottalization in Daai Chin, with parallel examples from Mro and Maru. *Linguistics of the Tibeto-Burman Area* 24(2):123–142.

Hill, Nathan. 2009. Review of Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction. *Languages and Linguistics* 10(1):173–195.

Marrison, Geoffrey E. 1967. The classification of the Naga languages of North-East India. PhD dissertation, University of London, School of Oriental and African Studies.

Matisoff, James A. 1970. Glottal Dissimilation and the Lahu High-Rising Tone: A Tonogenetic Case-Study. *Journal of the American Oriental Society* 90(1):pp. 13–44.

Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. University of California Press.

Miller, Amy. 2001. *A grammar of Jamul Tiipay: Amy Miller*. Hawthorne, N.Y.: Mouton de Gruyter.

Mortensen, David R. 2003. Comparative Tangkhul. Ms. UC Berkeley.

Mortensen, David R. 2010. The Place of Sorbung in Tibeto-Burman. Paper presented at ICSTLL (International Conference on Sino-Tibetan Languages and Linguistics), Lund, Sweden, October 16.

Mortensen, David R. and James A. Miller. 2009. Proto-Tangkhul Onsets in Comparative Perspective. Paper presented at ICSTLL (International Conference on Sino- Tibetan Languages and Linguistics), Chiangmai, November 4.

Peiros, Ilia and Sergej A. Starostin. 1996. *A Comparative Vocabulary of Five Sino-Tibetan Languages (6 vols.)*. Parkville, VIC: Univ. of Melbourne, Dept. of Linguistics and Applied Linguistics.

VanBik, Kenneth. 2009. *Proto-Kuki-Chin: A Reconstructed Ancestor of the Kuki-Chin Languages*, *STEDT Monograph Series*, volume 8. Berkeley: STEDT.

Wolfenden, Stuart Norris. 1929. *Outlines of Tibeto-Burman linguistic morphology*. London: Royal Asiatic society.

David R. Mortensen
University of Pittsburgh
2816 Cathedral of Learning
Pittsburgh, PA 15260

drm31@pitt.edu