1. INTRODUCTION

One goal of this paper is to highlight some contributions that continued analysis of prosodic properties (specifically, tone and consonant phonations) of Sino-Tibetan (and Indo-European) languages spoken in Nepal can make towards a number of domains of linguistic inquiry, including: phonological theory, prosodic typology, models of contact-induced language change, approaches to sound change in a more general sense, models of sociolinguistic (socio-phonetic) variation, and also psycholinguistic aspects of tonal production and perception. Another goal here is to discuss the methodological and analytical challenges that fieldwork of this nature in Nepal presents to such contributions. Hopeful outcomes of this account include increasingly flexible and creative approaches to phonetic data collection as well as collaborative interaction on methods of data collection in these types of field settings, such that the findings are meaningful to broader domains of theory, typology and socio-linguistic research.

2. PROSODIC SYSTEMS OF NEPALESE LANGUAGES: ISSUES AND CONTRIBUTIONS

Peter Ladefoged notes that phonetic description begins only once the phonological system of the language has been determined (1994: 13). With respect to the prosodic systems of Nepalese languages, it is in fact the case that the complexity of such systems cannot be fully appreciated until the phonetic details have been better understood, both in terms of category compositions and variation across speech communities and potential paths for change.

A word on the definition of terms like ‘prosody’ and ‘prosodic typology’ in this account is in order first. The term ‘prosody’ is applied variably for different phenomena, including suprasegmental properties (e.g. pitch, loudness, prominence patterns) versus at the segmental level (Crystal 1997: 313-314). In general, prosodic phonology is concerned with the study of phonological features that apply to or extend across ‘larger’ domains beyond the syllable and phonological foot. These features include word-level properties of tone, accent (including stress and pitch-accent) and rhythm, as well as phrasal and clause-level properties of intonation. ‘Prosodic typology’ is concerned with methods of cross-linguistic comparison of prosodic phenomena and the repercussions for phonological theory (see for example Jun 2005; Hyman 2006).

In this account, the focus will mainly be on phenomena surrounding tone and phonation-based systems in some Nepalese languages, although questions of stress, rhythm and intonation in these languages are still important and open ones.
2.1. Tone and the Prosodic Hierarchy

A popular and not totally inaccurate view of the phonological profile of Sino-Tibetan languages is that of ‘monosyllabicity’, with languages from S.E. Asia like (Cantonese) Chinese (Sinitic), Lahu and Lisu (Lolo-Burmese), where the domain of most phonological generalizations (including tone contrasts) is the syllable (Matisoff, J. 1999). Beyond this largely Southeast Asian profile, however, the domain of generalizations about tonal contrasts expand in complex ways.

The Bodish Tibeto-Burman languages from Nepal are a good example of this complexity, with a tone bearing unit (TBU) of either the morpheme or phonological word, regardless of syllable size (Matisoff 1999; Hildebrandt 2007c). A good case of this is found with Manange (Bodish, Tamangic), which has four tones (aligning along parameters of vowel fundamental frequency ($F_0$) and onset consonant aspiration). As is the case with other Bodish languages spoken in Nepal, the TBU in Manange is the phonological word, including the mono- or polysyllabic stem morpheme as well as bound affixes and clitics. So, a single $F_0$ trajectory spreads across morphologically simple and complex words in the same fashion. Figure 1 shows lines for the (mean) starting, mid-point and ending $F_0$ values of a near-minimal set of verbs, all with the nominalizing suffix -$p$ʌ (reproduced from Hildebrandt 2005: 31).

Figure 1
Mean $F_0$ measurements, Verb Stem + Suffix, Manange

![Figure 1](image)

The $F_0$ values spread from root across suffix syllables, indicating that phonologically bound grammatical morphemes pick up the $F_0$ contours of the preceding stem morpheme.

What is particularly interesting to note about languages like Manange is that although tone patterns refer to the phonological word as the domain of application, certain other phonological alternations in these languages fall within a more typically isolating profile of Sino-Tibetan noted above, in that the domain of application is much more restricted in morphological mapping. An example of this in Manange is a process of consonant deletion (a kind of anti-gemination) between the stem and clitics only (see Hildebrandt 2004: 66 for data).
As such, the notion of a prosodic word in languages like Manange necessarily involves an awareness of prosodic mis-alignment. In other words, there are different types of phonological words in these languages, including both ‘exhaustive’ words, where the process targets all available material (e.g. the tone word in Manange, referencing the prefix + stem + suffix + enclitic), and also ‘smaller’ words, where the process targets only some material (e.g. the ‘consonant deletion’ word in Manange, referencing stem + enclitic only). Such prosodic mismatches are attested in other Bodish languages spoken on the Nepal-Tibet border (c.f. Hall and Hildebrandt 2008/Forthcoming for an analysis of mismatches in Kyirong Tibetan), as well as in (non-tonal) Kiranti languages of eastern Nepal (see Hildebrandt 2007a for a description of multiple, mis-aligned words in Limbu).

As covered in Schiering et al (2007), such mis-alignment raises challenges for specific predictions about the relationship of phonological words to other constituents of the PROSODIC HIERARCHY (within the model of PROSODIC PHONOLOGY, the proposed set of hierarchically arranged domains for the application of phonological patterns). The challenge of multiple phonological words in Nepalese languages in particular is to the prediction of PROPER BRACKETING, which stipulates that phonological processes will cluster on a single (aligned) morphological mapping (Nespor and Vogel 1986) (i.e. a single, aligned phonological word contained within super-ordinate units of the hierarchy. The preponderance of both smaller and larger word domains provides evidence that languages can phonologize rules over multiple (and non-aligned) units.

2.2. Prosodic Typology and “Hybrid” Tone Systems
Most of the Bodish languages of Nepal have what can be termed ‘hybrid’ tone systems, in that lexical tone is an amalgam of (vowel) F0 contrasts, as well as different phonation types realized on either consonants and/or vowels. Tamang (Bodish, Tamangic), for example, has four tones that are different in pitch height (high and low) and contour (level and falling), with an additional voice quality difference for the two heights: The vowels in high tone words are modal (i.e. ‘plain’) and vowels in low tone words are breathy (i.e. ‘murmured’ or ‘lax’) (Mazaudon 1973: 82), illustrated in Figure 2.

<table>
<thead>
<tr>
<th>Tone</th>
<th>Correlates</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>/1/</td>
<td>High Falling,</td>
<td>tf³t ‘exclamation of dislike’</td>
</tr>
<tr>
<td></td>
<td>Modal</td>
<td></td>
</tr>
<tr>
<td>/2/</td>
<td>High Level, Modal</td>
<td>tf³hi ‘grass’</td>
</tr>
<tr>
<td>/3/</td>
<td>Low Level,</td>
<td>tf³ji: ‘grip’</td>
</tr>
<tr>
<td></td>
<td>Breathy</td>
<td></td>
</tr>
<tr>
<td>/4/</td>
<td>Low-Rising,</td>
<td>tf³ji: ‘think, remember’</td>
</tr>
<tr>
<td></td>
<td>Breathy</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2
Tamang tones (Adapted from Mazaudon 1973: 94–98)
Beyond the tone-phonation connection, many Nepalese languages without lexical tone still frequently have contrastive segmental phonation types, and this cross-cuts genetic affiliation. For example, Kathmandu Newar (Tibeto-Burman, Himalayish) has a modal and murmured phonation opposition on consonants (e.g. /ji/ ‘I’ and /jhi/ ‘ten’) (Genetti 1994: 37). Nepali (Indo-European, Indic) has a similar opposition (e.g. /ba:t/ ‘talk’ and /bha:t/ ‘cooked rice’), typically referred to as ‘voiced unaspirated and aspirated’ (Acharya 1991: 28). In Kusunda (Isolate), although non-modal phonations are not contrastive, there is a pharyngealization observable on vowels, via the erosion of a voiced uvular plosive [g] (e.g. [a\text{-}\text{Gk}\text{a} ‘up, above’) (Watters 2006: 27).

A number of interesting questions come from these observations, including how to distinguish between ‘tonal’ and ‘phonation prominent’ languages in this area, what sub-phonemic cues speakers attune to for lexical disambiguation, and also how a careful examination of these contrastive and sub-contrastive properties can contribute to our understanding of tonogenesis (and tonexodus). Some observations along these lines are detailed now and in section 2.3.

The question of lexical disambiguation is especially relevant in languages like Tamang and Gurung, where tone is a combination of contrastive pitch, segment phonation and duration properties, and where homophony within tone categories is frequently attested (e.g. /\text{k}\text{w}\text{e}/ ‘bee’, /\text{k}\text{w}\text{e}/ ‘song’ in the Kaski dialect of Gurung).

Some attempts to examine the acoustic and perceptual properties of tone in two dialects of Gurung spoken in Manang, Nepal were performed by Hünlich (2006) and by Hildebrandt and Hünlich (2007). It was observed that while some acoustic (primarily $F_0$) patterns for words from different tone categories show very little significant distinctive patternings, speakers are still very good at making lexical disambiguation decisions in context-free tasks. In other words, even though the raw pitch properties of words from different tones do not correspond with other published descriptions of Gurung tone, Manang-Gurung speakers still discriminate the meanings of such words with a high degree of accuracy, even when they are heard in semantically and syntactically neutralizing frames or from different speakers. These preliminary results suggest that despite the two-way pitch split characterizing the tones phonologically, speakers attune more towards sub-phonemic factors of segment phonation (and possibly segment duration, which is diachronically young and only marginally contrastive). Such issues form the basis for future research.

2.3. Language Contact and the Role of Sub-phonemic Detail

The amalgam properties of Nepalese languages are also important for investigations of the results of language contact in scenarios of either asymmetrical, bi-/multilingual maintenance or else slow death, both of which are attested in Nepal (Matisoff, J.A. 1991; Kansakar 1996; Noonan 1996; 2005). One model of such change includes two scenarios of structural borrowing outcomes: pattern-borrowing, with the adoption of a particular strategy from one language to another, and matter-borrowing, the borrowing of a structure or form (Matras and
Sakel 2007/Forthcoming). One study of tone change in Manange may be a challenge for both scenarios.

Hildebrandt (2003) documents the phonetic and phonological properties of the tone system in two communities of Manange speakers: Rural Mananges, who were born and raised in the traditional, Bodish-language speaking Manang area of northern-central Nepal where Manange is a regional lingua-franca, and Urban Mananges, who were born and raised in an environment of Nepali bilingualism in urban Kathmandu. Unlike the rural speakers, Urban Mananges show a large-scale phonetic merger of the two contour tones into a two-way high-low opposition. The more conservative phonetic view of the system across three speakers is illustrated in Figure 3, while the merged phonetic view (across three speakers) is shown in Figure 4.

![Figure 3](image1.png)

**Figure 3**
Average F$_0$ Values-by-Tone: Rural Mananges

![Figure 4](image2.png)

**Figure 4**
Average F$_0$ Values-by-Tone: Urban Mananges

It is not obvious that the altered tone system is simply one structural symptom of a larger process of slow language loss (i.e. shift to Nepali). Manange in
Kathmandu appears to be maintained in a situation of diglossia, whereby its place remains in private contexts, while Nepali is the language of necessity in public domains. In addition, both Rural and Urban Mananges can make lexical disambiguation decisions based on a four-way tone contrast. Nevertheless, this phonetic structural result may be a consequence of this maintenance scenario.

It is also not obvious that urban Manange is borrowing anything from Nepali phonology either. Nepali has no tone, and in fact has a four-way obstruent voicing distinction (voiceless plain, voiceless aspirated, voiced plain and voiced aspirated). There is no evidence that urban Mananges are incorporating obstruent onset voicing into their production of Manange.

Another analysis could be structural simplification due to contact and asymmetrical bilingual language use. However, a general ‘simplification’ analysis misses out on both the structural properties behind this uneven diffusion of (surface) tone merger, as well as the usage-based factors responsible for the perceptual retention of a four-way contrast for both populations of speakers. Hildebrandt (2007b) shows that for urban speakers, low-grade phonetic factors, such as intrinsic pitch and voice-onset-timing of onsets, appear to be hastening a phonetic merger in words that are highly frequent, while low-frequency words overall retain the acoustic profile of all words. As such, the role of sub-phonemic detail becomes a crucial part of better understanding the mechanisms of contact-induced language change, as well as predicting the possible path of tone change for this language.

3. PREFERRED METHODOLOGY MEETS HARSH REALITY: THE NEPAL FIELDWORK CONTEXT

The studies in section 2 all have in common the need for exceedingly well planned and careful methodologies for detailed data collection (both at the level of the token and speaker representation, and also at the level of situation and equipment). It is important to acknowledge that even under the best of circumstances, a tight control over all of these factors can be difficult. However, these difficulties can be compounded in fieldwork settings like remotely located villages, or in scenarios of social upheaval, and sometimes, adjustments and compromises are necessary. As such, it is a non-trivial matter to explore what methodological adjustments and compromises are allowable, while still providing for the validity and application of findings to more general debates and discussions. The following sections deal with such complications and issues of compromise.

3.1. Fieldwork in Nepal: Some general considerations

Despite its relatively small geographic size, Nepal is ethnically and linguistically diverse, with approximately one-hundred languages from five families (Indo-European, Dravidian, Austro-Asiatic, Sino-Tibetan and an Isolate) (see Yadava 2002 for a report on languages spoken in Nepal). It is also a country with a history of political instability and sudden, violent political shifts. These shifts
have been defined by the often antagonistic relationship between the historically
established monarchy (which has experienced its own episodes of violent internal
upheaval) and the more newly founded but equally fragile parliamentary system,
which in principle strives for a more democratic-oriented manner of rule. (see
Thapa 2003 for a comprehensive historical account of conflicts between the
monarchy and political parties).

On top of this has been an increasingly violent rise of a guerilla insurgency
group, a breakaway faction of the Communist Party of Nepal, (i.e. the Maoists).
The aim of the Maoists is to overthrow what they view as an antiquated and
biased caste-system of absolute rule and exclusion perpetuated by the monarchy,
and replace this with what they call a “people’s democracy” (Josse 2003: 9). In
many Village Development Communities (VDC’s) the result of this insurgency
has been frequent and often violent conflicts between government forces and rebel
groups (who are alternatively viewed with both admiration and hatred by local
populations). Not surprisingly, these conflicts have often lead to more village
population casualties than other types.

Adding to these hardships is the ranking of Nepal as one of the poorest Asian
countries, where the currency is non-negotiable beyond national borders, and
where 40% of the population live below the poverty line (i.e. below a per annum
income of U.S. $77) (Asian Development Bank 2007). Such factors make the
implementation of certain methodological considerations for phonetic linguistic
fieldwork (from both psycholinguistic and sociolinguistic angles) exceedingly
difficult. A few examples of these issues are summarized here, with specific
reference to Bodish-language areas, along with occasional considerations of what
possible alternatives are acceptable before any generalizable findings themselves
are also compromised.

3.2. Population Sampling

Even before the Maoist insurgency, a fact of life for many Nepalese communities
was temporary or permanent migration of some of the population, usually due to
economic or climate-related motivations. With the current insurgency, many
VDC’s overwhelmed by Maoist forces the ensuing armed conflicts have seen the
loss of many younger people (either via forced conscription into the paramilitary,
or else via relocation to comparatively more stable regions in Nepal or India). As
a result, these villages are either rapidly emptying out, or else the sociolinguistic
demographics are skewed towards older speakers.

In Manang, a different type of population movement is currently underway,
namely the dismantling and relocation of several VDC’s due to the building of a
motor road to connect the end of the King Mahendra Highway in Besisahar,
Lamjung to the main Manang village about 50 miles away. The temporary
relocation of Manang peoples because of economic and climate-related factors is
not necessarily a new phenomenon (see Rogers 2004 for an economic history of
Manang). However, the current motor-road project is resulting in fairly rapid
displacement and relocation of some communities, with consequences for the
dialectal tapestry as it currently exists.
As such, the notion of a sociolinguistic sample (e.g. the boundaries of a dialect group, or a balanced sample based on social variables) becomes more complex and the need for such studies takes on a new type of urgency. One way around this issue is to adapt a more flexible definition of what constitutes a language (or dialect) community in Manang, and that a generalization over a smaller proportion of the original community may still provide meaningful findings within larger debates. Another consequence is the need for regular, ongoing studies, even if they are relatively small in scale and population representation or occur in different environments at different times.

3.3. Challenges with Production and Perception Studies

Section 2.3 argues that an understanding of both the (acoustic) production and the perception of prosodic phenomena in Nepalese languages is important to a better appreciation of the role of phonation types (both contrastive and sub-contrastive). Such studies therefore require methodologies that include carefully constructed experimental procedures. Given the largely acoustic issues at hand, these methodologies also require the use of fairly sophisticated equipment, perhaps beyond the level of a microphone-plus-tape-recorder setup. By following advice found in Ladefoged (1993; 1997) and Maddieson (2001), recommended equipment would include a high quality omnidirectional microphone, an easily digitizable medium of recording (e.g. DAT or a solid-state recorder), electroglottographic equipment for examination of non-modal phonations, backup equipment and alternative power supplies.

However, it is not surprising that the logistics of equipment transport and use are complicated in Nepal. Both politically charged and more neutral issues such as care (and repair) of this equipment in such settings are challenging, and my equipment has been an occasional victim to unfortunate events in the natural world (e.g. sudden rainstorms and slippery paths) as well as to human-caused ones (e.g. leaking water bottles and encounters with self-proclaimed revolutionaries in need of ‘donations’ to the political cause).

Assuming that the equipment makes it to the field site without damage (or backup equipment is arranged), another methodological consideration is the preferred environment of data elicitation (recording) and perception studies. Maddieson (2001: 221) notes that locations with ambient noise and reverberating structures can interfere with the acoustic quality of recordings. Most of my work in Manang is in villages along the shores of the (roaring) Marsyangdi River or in villages on the slopes of the Annapurna mountain range, where windy afternoons are common. The outdoor areas that are sheltered from the wind and water typically contain other ambient noise-producing elements like domestic animals and children. Working indoors brings its own challenges, as houses in upper Manang consist of one or two-room dwellings, built mainly of stone, a naturally reverberating material.

With these environmental challenges, one solution is to apply a more liberal definition of “fieldwork” for phonetic purposes. I personally have found that better results are obtained when language informants perform phonetic elicitation work away from daily existence contexts, but at the same time not entirely
removed from familiar environments (e.g. Gurung language perception experiments undertaken in Kathmandu, and Manange data recorded for acoustic purposes in a wood-built trekking lodge owned and operated by other Mananges).

Another option is to allow for more “homegrown” stimulus methodologies and to view the results as still valid (at least in a preliminary sense). An example of this may be found in the initial study of the perception of tonal contrasts in Gurung with a very limited access to data and informants (Hünnlich 2006), noted in section 2.2. These initial findings, while extremely preliminary, hinted at a pattern that has so far been verified by a more expanded study: $F_0$ does not appear to correlate with perception and discrimination of tone categories in this language. As such, the details of this true ‘amalgam’ system require further examination.

3.4. Challenges for Theoretical Applications

Based on the previous observations, it is not surprising that there remains a lack of detailed published data on which to link these languages and their respective patterns to more general claims about natural language. A number of Nepalese languages do have fairly good descriptions available, some even with high quality, comprehensive reference grammars, analyzed discourse data from various genres, and dictionaries available. However, many of these languages are either under-described (with sketch descriptions, basic-vocabulary wordlists, or else largely non-circulating published information), or else remain undescribed altogether. As discussed in section 2.3, one proposal behind Hildebrandt’s (2003) analysis of the uneven rate of contact-induced tonal merger in one Manange-speaking community is that of lexical frequency, which is grounded in factors of language use. Such an analysis would require access to large amounts of corpora materials from multiple genres in order to get an empirically valid notion of frequency in Manange. The reality of this however is that few published accounts of Nepalese languages currently come with such data.

However, the findings from preliminary and small-scale frequency studies may still be of value within larger contexts. One methodological consideration worth promoting here is an awareness of what statistical tools (e.g. randomization tools) are useful in small or uneven sample types, which provide a greater degree of margin-of-error safety, and which ultimately allow for relatively safe generalizations from the findings at-hand (see Janssen et al. 2006 for randomization tools in typology).

4. SUMMARY

There are other methodological factors and complications not considered in this account, including the personal feelings that informants may have about doing fairly controlled and repetitive types of work required for acoustic and perceptual phonetic research, the appropriateness of this type of work in communities that are undergoing political and economic hardships, the place of spontaneously occurring discourse data in such fine-grained phonetic work, as well as how the
training of informants for this type of work should best go about. These are all issues that are deserving of further discussion.

Ultimately, the goal of this account is not to promote fieldwork in Nepal and the significance of the findings in Nepalese languages as unique or necessarily more challenging (or rewarding) than in any other situation. The goal is also not to promote data of lower quality as necessarily the same thing as ‘acceptable’ quality. Rather, the idea here is to repeat and emphasize many of the themes raised in Ratliff and Newman, eds. (2001), but within the context of phonetic and phonological research in Nepal. These themes include the need for flexibility and a certain amount of fearlessness, especially in the awareness that most findings in these environments will potentially emerge as tentative or based on a less-than-optimal sample. At the same time, given the intriguing nature of prosodic systems in Nepalese languages and the exciting research questions that the observations raise, these types of studies and their findings are in great need. Therefore, while the findings of specific production-perception interfaces and socio-phonetic investigations may lack the strict laboratory-style controls found in similar types of research on other languages in other environments, they are still rewarding in the paths of future investigation they have opened up.

REFERENCES


