A CASE STUDY OF TONE AND INTONATION IN TWO TIBETIC LANGUAGE VARIETIES

Amos Teo[1], Lauren Gawne[2], Melissa Baese-Berk[1]

[1]University of Oregon, [2]SOAS, University of London
ateo@uoregon.edu, lauren.gawne@gmail.com, mbaesebe@uoregon.edu

ABSTRACT

This paper presents a case study looking at the interaction between lexical tone and post-lexical intonation in two very similar Tibetic language varieties spoken in Nepal: Lamjung Yolmo and Kagate. In these two varieties, we find preliminary evidence that in both monosyllabic and disyllabic words, lexical tone is only specified at the left edge of the word, while the right edge of the word is 'free' to take post-lexical intonation tones. We present evidence of post-lexical intonation on these 'free' right edges both phrase medially and phrase finally. These results suggest that a description of the tone system of these languages without reference to the intonational system is too simplistic, and any future analyses should incorporate descriptions of both lexical tone and post-lexical intonation.

Keywords: fieldwork, intonation, Nepal, Tibetic, tone

1. INTRODUCTION

1.1. Studies on tone and intonation

Much of the research looking at the interaction between tone and intonation has focused on so-called 'syllable' tone languages, such as Mandarin Chinese [16, 20, 21], Cantonese [12], Vietnamese [5] and Thai [15]. A distinction is typically made here between 'lexical tone' which serves to disambiguate lexemes at the word or morpheme-level, and 'post-lexical intonation' which serve more discourse pragmatic functions at the sentence or utterance-level. In syllable tone languages, almost every syllable is specified for tone, while intonation is treated as a separate system that is superimposed over tone sequences, usually without compromising lexical contrasts. Intonation is typically realized phonetically through global changes in pitch range and pitch shape [20]; in addition to changes in duration and phrasing.

Less work has looked at the interaction between tone and intonation in tone languages outside East and Southeast Asia. Much of this work has looked at how intonational tones either replace lexical tones or are added after lexical tones to distinguish questions from statements [11] or to mark prominence [13].

This paper provides an acoustic analysis of the lexical tones in two Tibetic language varieties, and shows evidence that lexical tone is only specified on the left edge of both monosyllabic and disyllabic words, while the right edge is free to take post-lexical tones. The findings are of particular interest because we see a systematic interaction between post-lexical intonation and lexical tone, even on monosyllabic words and in non-phrase final locations, in a way that has not been described in other tone languages of the area, and is not found in more well-studied tone languages of East and Southeast Asia.

1.2. Language background

The two Tibetic language varieties that were examined are Lamjung Yolmo (LY) and Kagate (K). Both are spoken in Nepal by small speaker populations of fewer than 1,000 each. They both share a high degree of mutual intelligibility [3]. Previous work on Yolmo has focused on the variety of Yolmo spoken in the Melamchi area [6, 7]. Hari describes four tones for this variety: low level; low falling, high level; and high falling. However, no minimal sets showing the contrast between falling and level tones have been noted for Melamchi Yolmo, nor are such sets found in LY and K. Furthermore, a preliminary acoustic study of LY showed only evidence for a two-way contrast between low and high tones [4]. Note that in LY and K, word-initial voiced obstruents are always followed by low tone, while word-initial aspirated stops, fricatives and affricates are followed by high tone, suggesting that contrastive phonation and 'redundant' tone exist in some environments.

The first aim of the current study is to provide an acoustic description of tone in these varieties, serving to complement previous descriptions of tone, based only on auditory impressions. It systematically looks at tone minimal and near-minimal pairs in both isolation and in phrase-medial position. The second aim is to examine interactions of tone and intonation in these two language varieties.

2. METHOD

2.1. Speakers

Four participants were recorded in this study: There were two speakers of LY: one female, 48; one male, 20; and two speakers of K: one male, 49, and one female, 18.

2.2. Materials

A list of tone minimal and near-minimal pairs found in LY and K was compiled. A total of 76 separate lexical items were recorded (30 minimal pairs and 8 near-minimal pairs). Three repetitions of each lexical item were recorded in isolation. 3 repetitions of each lexical item were also recorded in the carrier phrase: *Ngà* _____ *làpsing / làpke.* 'I said / say _____.'

The sound files were stored and labelled in the EMU Speech Database System (version 2.3.0) [8] and F0 measurements were extracted and analyzed in R (version 2.8.1) [17]. Praat (version 5.3.23) [2] was used to generate pitch traces. For this study, only data from open monosyllabic words and disyllabic words were chosen to demonstrate the interaction between tone and intonation.

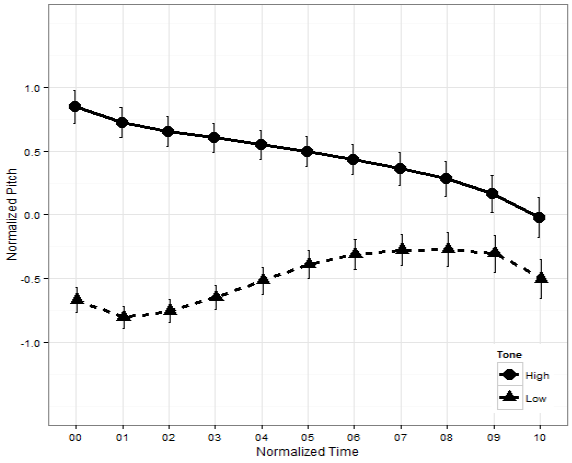
3. RESULTS

3.1. Lexical tone in carrier phrases

First, we examined how the lexical tones are realized in phrase-medial position within a carrier phrase. In this position, words were produced with a generally consistent pitch shape across repetitions.

Figure 2 shows a *z*-score normalized pitch trajectory over the vowel portion of open monosyllabic words across normalized time collapsed across all four speakers. Normalized pitch *z*-scores were plotted at 10% intervals with 95% confidence intervals. At the beginning of the syllable, high tone is produced at a higher pitch than low tone. However, the contrast in pitch begins to neutralize towards the end of the word, as the average pitch of the high tone falls, while the average pitch for the low tone rises, before falling again at the end of the syllable.

Figure 2: Time-normalized pitch trajectory over vowel portion of open monosyllabic words (all speakers)



We analyzed the data using a mixed effects model using tone category, time, gender, and the interaction between tone category and time as fixed factors and speaker and language variety as random effects. The results support the picture shown above: tone category is a significant predictor of pitch, ($\chi^2(1) = 209.78$, $p < .001$). Time alone is not a significant predictor of pitch ($\chi^2(1) = 0.98$, $p = .32$), but there is a significant interaction between tone category and time ($\chi^2(1) = 30.35$, $p < .001$). This reflects the trend we see above where we find the high and low tone starting to converge in pitch towards the end of the syllable. Estimates and *t*-values from the best fitting model are presented in Table 1.

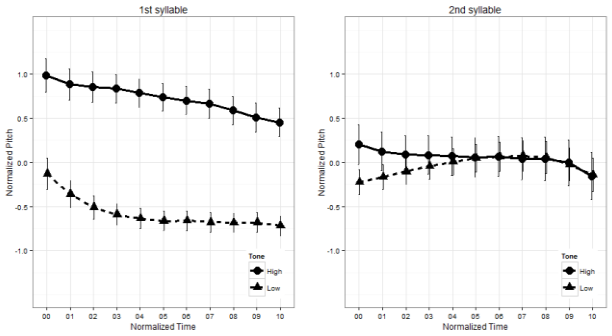
Table 1: Estimates and *t*-values for best fitting model for monosyllabic words.

Fixed effect	Estimate	Standard error	*t* value
Tone category	-26.54	1.75	-15.17
Time	-5.41	1.72	-3.15
Gender	-71.41	8.37	-8.53
Tone-Time Interaction	13.76	2.47	5.57

Figure 3 shows a *z*-score normalized pitch trajectory over the vowel portions of disyllabic words across normalized time collapsed across all speakers. Normalized pitch *z*-scores were plotted at 10% intervals with 95% confidence intervals. We can see that high tone is produced at a higher pitch than low tone only on the first syllable, while no contrast in pitch is produced on the second syllable. For words in both tone categories, the pitch across the second syllable is realized at a pitch height that

lies somewhere between the average pitch of the low and high tones on the first syllable.

Figure 3: Time-normalized pitch trajectory over vowel portions of disyllabic words (all speakers)



We analyzed the data using a mixed effects model using tone category, syllable, gender, and the interaction between tone category and syllable as fixed factors and speaker and language variety as random effects. Again, the results support the picture shown above: tone category is a significant predictor of pitch, ($\chi^2(1) = 51.18$, $p < .001$). Syllable is also a significant predictor of pitch ($\chi^2(1) = 18.85$, $p < .001$), and there is a significant interaction between tone category and syllable ($\chi^2(1) = 46.78$, $p < .001$). These results reflect the picture we see above where words with high and low tone are produced with significantly different pitch over the first syllable (about 20.1 Hz ± 1.9), but not over the second syllable. Estimates and *t*-values from the best fitting model are presented in Table 2.

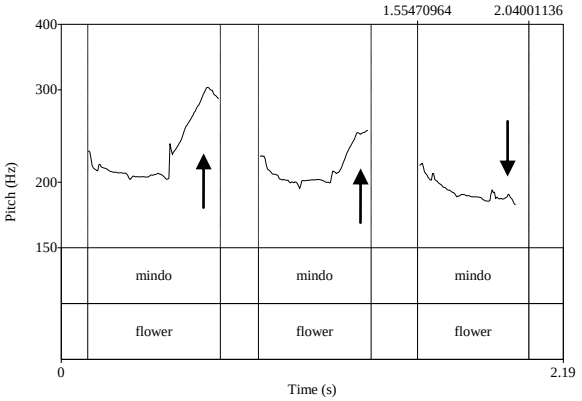
Table 2: Estimates and *t*-values for best fitting model for disyllabic words.

Fixed effect	Estimate	Standard error	*t* value
Tone category	-20.12	1.92	-10.49
Syllable	-4.49	2.03	-2.21
Gender	-67.05	11.46	-5.85
Tone-Syllable Interaction	19.07	2.71	7.04

3.2. Interaction between lexical tone and post-lexical intonation at phrase boundaries

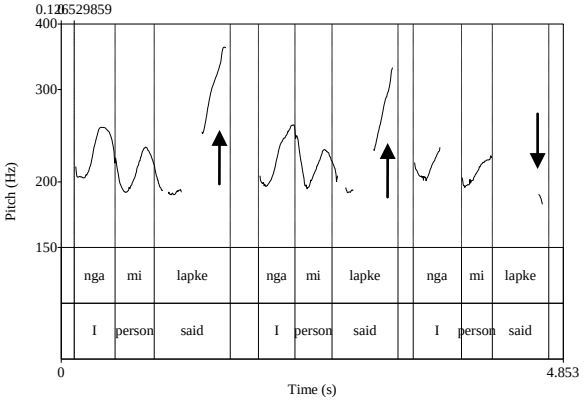
Next, we looked at words repeated in isolation frames. We found that for all speakers' productions of disyllabic words, lexical tone was marked on the first syllable, while non-final repetitions (i.e. the first two repetitions in a set of three) were characterised by a rise over the second syllable. In contrast, the final repetition was accompanied by a low, slightly falling pitch over the second syllable. Figure 4 gives an example of this, with the pitch trace for the word /mìndo/ 'flower' spoken three times by the female K speaker.

Figure 4: Pitch trace over /mìndo/ 'flower' in isolation frame (female Kagate speaker)



This rise in pitch at the end of each non-final repetition was also observed at the end of each repeated carrier phrase. Figure 5 gives the pitch trace of the word /mì/ 'person' produced in a carrier phrase three times by the female K speaker. Here, we see a rise over the second syllable of the word *làpke* in the first two repetitions, but a fall in the final one.

Figure 5: Pitch trace over /mì/ 'person' in carrier phrase (female Kagate speaker)



Even in their production of monosyllabic words in isolation, all speakers produced the same rising pitch over the second half of the word for all non-final repetitions. However, speakers still maintained the contrast in pitch over the first half of the monosyllabic word. For instance, Figures 6 and 7 give pitch traces for the words /mì/ 'person' (low tone) and /mí/ 'eye' (high tone) repeated three times in isolation by the female K speaker. Here we can see that pitch over the first half of /mì/ 'person' is

about 20 Hz lower than the pitch over the first half of /mí/ ‘eye’.

Figure 6: Pitch trace over *mi* /mì/ ‘person’ in isolation frame (female Kagate speaker)

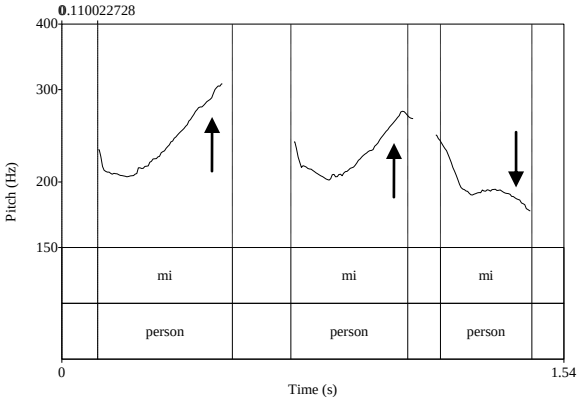
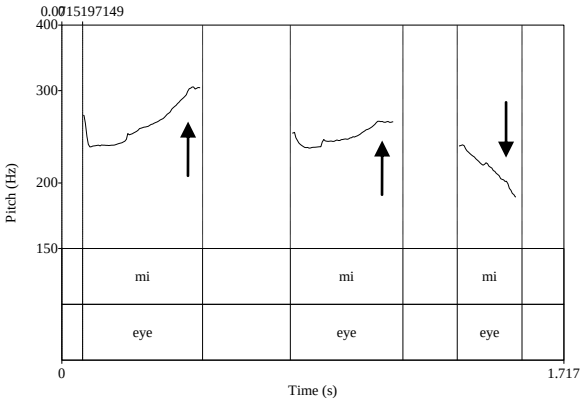


Figure 7: Pitch trace over *mi* /mí/ ‘eye’ in isolation frame (female Kagate speaker)



Due to the preliminary nature of this analysis, we are hesitant to give labels to the intonational patterns here, but it appears that the rise in pitch on non-final repetitions is expressing continuing intonation, while the phrase-final low pitch indicates utterance finality. Nevertheless, what is interesting here is that we find a lexical tonal component on the left-edge and a post-lexical intonational component on the right-edge in the production of these words in isolation frames – this is not the expected situation in the more well-studied tone languages of Asia.

4. DISCUSSION

In this study, we demonstrated that high tone in LY and K is produced with a significantly higher pitch than the low tone. For words in carrier phrases, contrastive pitch is generally found on the left edge of the word. For disyllabic words, the contrast in pitch is produced only on the first syllable, while the second syllable is produced with level pitch at a pitch height about halfway between the pitch heights of the high and low tone of the first syllable. For monosyllabic words, the contrast in pitch is produced mainly across the first portion of the vowel segment.

These findings support previous auditory impressions that tone is only contrastive on the first syllable of disyllabic words in LY and K. The location of the tone contrast on the left edge is not unusual for Tibetic languages, cf. Lhasa Tibetan [18] and Sherpa [19]. The left edge location also supports hypotheses that tones in Tibetic languages arose out of word-initial consonant cluster simplification, along with the loss of an initial voicing contrast [1], although the complete loss of a word-initial voicing contrast has yet to occur in LY and K.

Most work on lexical tone and intonation has found intonation to be realized by global changes to F0, or by the addition of an intonational tone that results in the neutralization of a tonal contrast at a phrase boundary. In contrast, what we are seeing in LY and K is both lexical tone and post-lexical intonation being marked on a single word, including monosyllabic words: the left edge of the word is marked for tone, while the right edge is ‘free’ to express post-lexical intonation information. From a descriptive point of view, this interaction between intonation and tone means that labels like ‘low level’ or ‘high falling’ are not particularly useful to describe the tone ‘contours’ of LY and K, unless we are describing a combination of lexical tone and post-lexical intonation. We suspect such a system may have contributed to Hari’s [6] previous description of four contrastive tones in Melamchi Yolmo.

For researchers more familiar with tone and intonation in languages of East and Southeast Asia, these findings may seem quite unusual. Even where post-lexical tones are found in these languages, as noted in some recent studies of intonation in Cantonese (e.g. [12]), these are often confined to the ends of utterances where they are used to distinguish questions from statements. Crucially, they also often result in the neutralization of tonal contrasts, which is clearly not the case in LY and K.

We propose that one important difference between these Tibetic language varieties and languages of East and Southeast Asia is that the former are specified for a single pitch target at the left edge of the word, while the latter have at least two targets at the start and at the end of syllables. The presence of a second pitch target constrains the addition of post-lexical intonation. This may reflect the evolution of complex tones in Vietnamese and Sinitic, which are hypothesized to have involved both the loss of a syllable-initial voicing contrast as well as the loss of syllable-final consonants (e.g. [9, 10, 14]).

6. REFERENCES

[1] Bielmeier, R. 1988. On tone in Tibetan. In: Uebach, H., Panglung, J. L. (eds), *Quellen und Studie zur tibetischen Lexiokographie.* Band II. München: Komm. für Zentralasiatische Studien. Bayrische Akademie der Wiss, 43-54.

[2] Boersma, P., Weenink, D. 2013. *Praat: doing phonetics by computer* [Computer program]. Version 5.3.23, retrieved 18 August 2012 from http://www.praat.org/

[3] Gawne, L. 2013. "Notes on the relationship between Yolmo and Kagate." *Himalayan Linguistics* no. 12 (2):1-27.

[4] Gawne, L., Teo, A. 2012. A preliminary analysis of tone in Lamjung Yolmo. Paper presented at the 18th Himalayan Languages Symposium, Banaras Hindu University, Varanasi, Sep 12, 2012.

[5] Ha, K-P., Grice, M.. 2010. Modelling the interaction of intonation and lexical tone in Vietnamese. *Proceedings of Speech Prosody*. Chicago.

[6] Hari, A. M. 2010. *Yohlmo Sketch Grammar*. Kathmandu: Ekta books.

[7] Hari, A. M., and Lama, C. 2004. *Hyolmo-Nepālī-Aṅgrejī śabdakośa = Yohlmo - Nepali - English dictionary*. Kathmandu: Central Dept. of Linguistics, Tribhnvan University.

[8] Harrington, J. 2010. *The Phonetic Analysis of Speech Corpora*. Blackwell.

[9] Haudricourt, A-G. 1954a. Comment reconstruire le chinois archaique. *Word* 10, 351–364.

[10] Haudricourt, A-G. 1954b. De l'origine des tons en vietnamien. *Journal Asiatique* 242, 69-82

[11] Hyman, L. M., Leben, W. R.. 2000. Suprasegmental processes. In Booij, G., Lehmann C., Mugdan J. (eds.), *Morphology: an international handbook on inflection and word-formation*. Berlin: de Gruyter, 587–594.

[12] Ma, J. KY, Ciocca, V., Whitehill, T. L. 2006. Effect of intonation on Cantonese lexical tones. *J. Acoust. Soc. Am.* 120.6, 3978-3987.

[13] Michaud, A. 2006. Tonal reassociation and rising tonal contours in Naxi. *Linguistics of the Tibeto-Burman Area*, 61–94.

[14] Michaud, A., Vaissière, J. 2015. Tone and Intonation: Introductory Notes and Practical Recommendations. *Kieler Arbeiten zur Linguistik und Phonetik* 3, 43-80.

[15] Morén, B., Zsiga, E. 2007. The lexical and post-lexical phonology of Thai tones. *Natural Language and Linguistic Theory* 24, 113-178.

[16] Prom-On, S., Xu, Y., Thipakorn,B.. 2009. Modeling tone and intonation in Mandarin and English as a process of target approximation. *J. Acoust. Soc. Am*. 125.1, 405–424.

[17] R Core Team (2014). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

[18] Tournadre, N., Dorje, S. 2003. *Manual of Standard Tibetan.* Snow Lion.

[19] Tournadre, N., Sherpa, L. N.. Chodrak, G., Oisel, G. 2009. *Sherpa-English and English-Sherpa dictionary with Literary Tibetan and Nepali equivalents.* Kathmandu: Vajra Bookstore.

[20] Xu, Y. 1999. Effects of tone and focus on the formation and alignment of *F*0 contours," *J. Phonetics* 27, 55–105.

[21] Xu, Y. 2009. Timing and coordination in tone and intonation: An articulatory-functional perspective, *Lingua* 119, 906-927.