THE INTERACTION OF TONE AND INTONATION IN USPANTEKO

Ryan Bennett[a], Robert Henderson[b], and Megan Harvey[b]

University of California, Santa Cruz[a] and University of Arizona[b]
rbennett@ucsc.edu, rhenderson@email.arizona.edu, mharvey3@email.arizona.edu

ABSTRACT

We investigate the effect of intonation and discourse structure on the realization of lexical tone in Uspanteko (Mayan). An H% boundary tone occurs in utterance-final position, and focus produces an upward shift in pitch register. The interaction of these phenomena with lexical tone suggests a privative [H] vs. [∅] tonal contrast for this language.

Keywords: tone, tone-stress interaction, intonation, information structure, Mayan

1. INTRODUCTION

Uspanteko (or *Tz'únun Tziij* 'hummingbird speech') is a K'ichean-branch Mayan language spoken by 1200-4000 people in the region surrounding San Miguel Uspantán, Guatemala [23]. Uspanteko is endangered, with many children in Uspantán learning K'iche' and/or Spanish as their first language.

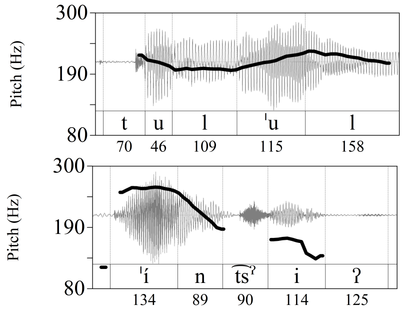
Uspanteko is the only Guatemalan Mayan language which clearly has lexical tone [2, 3, 6, 7, 9, 13]. Tone is quite restricted: there is a binary [H]∼[∅] contrast, limited to stressed syllables. (While this tonal pattern resembles 'lexical pitch accent', we avoid that term because it is poorly defined [16] and easily confused with *intonational* pitch accents.)

Lexical tone in Uspanteko is overlaid on a system of stress and vowel length familiar from related languages [6]. Stress in non-tonal words is final (e.g. *kaminaq* [ka.mi.ˈnaqʰ] 'dead'). Contrastive long vowels only occur in word-final stressed syllables (e.g. *k'echelaaj* [kˀe.t͡ʃe.ˈlaːχ] 'forest'). Tone is restricted to the last two syllables of the word, and interacts with stress. Tone coincides with stress on final long vowels, as in *incháaj* [in.ˈt͡ʃáːχ] 'my ash'. In a word with a final short vowel, like *wíxk'eq* [ˈwíʃ.kˀeqʰ] 'my fingernail', lexical tone appears on the *penult*, and stress retracts to the same position.

The phonetic effects of stress and tone are at least partially independent: stress is associated with increased duration, while tone is associated with pitch raising on the accented syllable (Fig. 1; [3, 4]).

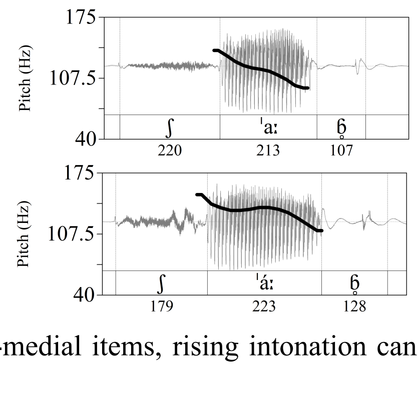
Lexical tone has a low functional load in Uspanteko—there are few true minimal pairs—but it is systematically associated with certain morphological constructions, such as possession (e.g. *ixiim* [i.ˈʃiːm] 'corn' vs. *wíxim* [ˈwíʃ.im] 'my corn').

Figure 1: Non-tonal *tulul* [tu.ˈlul] 'zapote' (top) vs. tonal *íntz'i'* [ˈín.t͡sˀiʔ] 'my dog' (bottom) [3]. X-axis = segment duration (ms); y-axis = f0 (Hz).



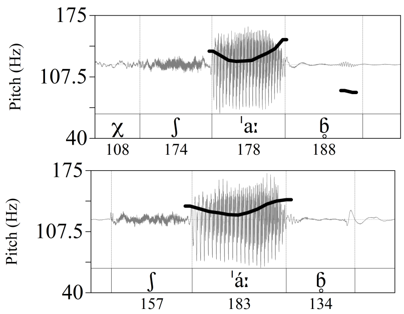
Previous fieldwork [3, 4] suggests that intonation may mask lexical tone distinctions in Uspanteko. Both utterance-final H% rises (§2) and list intonation, for example, affect the realization of tonal contrasts. Consider the minimal pair in Fig. 2. These words were uttered by the same speaker, both as the last item in a list. Lists in Uspanteko are typically produced with rising intonation (LH%) on medial items, but not on final items [15]. In final position, where there is no rising intonational contour, tonal *xáab'* [ˈʃáːɓ̥] 'vomit' is realized with higher pitch and a more pronounced rise-fall pattern than non-tonal *xaab'* [ˈʃaːɓ̥] 'comb'. Both pitch contours may also be influenced by falling intonation (L%) in Fig. 2, common at the ends of lists in Uspanteko.

Figure 2: Non-tonal *xaab'* [ˈʃaːɓ̥] 'comb' (top) vs. tonal *xáab'* [ˈʃáːɓ̥] 'vomit' (bottom), without rising intonation (speaker MH).



On list-medial items, rising intonation can obscure lexical tone distinctions. Fig. 3 shows that the pitch contour on non-tonal *jxaab'* [ˈχʃaːɓ̥] 'her comb' neutralizes with the pitch contour on tonal *xáab'* [ˈʃáːɓ̥] 'vomit' under rising intonation.

Figure 3: Non-tonal *jxaab'* [ˈχʃaːɓ̥] 'her comb' (top) vs. tonal *xáab'* [ˈʃáːɓ̥] 'vomit' (bottom), with rising intonation (speaker MH).



The fact that intonation interacts with tone, possibly masking contrasts, creates both challenges and opportunities. The most immediate challenge is that we must carefully control for intonational factors when studying the phonetic correlates of tone. The primary opportunity lies in the fact that intonation can help diagnose tonal specifications (or the lack thereof) on vowels. Previous phonological analyses of Uspanteko differ mainly in the amount of tonal specification assumed on the last two syllables of the word. Some authors argue for a privative [∅] vs. [H] (or [HL]) contrast, realized only on stressed syllables [3, 7]. Others assume a two-way [HL] vs. [LH] contrast, with tone specified on all (final) long vowels, and on the last two syllables of words with only short vowels [13] (see also [6]). Following [20–22] and others, we presume that intonation may be realized differently on toneless and tonal vowels, and may have distinct effects on vowels bearing different lexical tones (e.g. L vs. H). The particulars of how tone and intonation interact in Uspanteko may thus help decide between competing phonological analyses of the prosodic system. We explored this possibility through a controlled production study centered on tone and intonation.

2. METHODS

Our production study examined word-level prosody under different intonational and discourse conditions in Uspanteko. We focused on tonal contrasts in bisyllabic words, classified into 4 types according to their tone and the length of the final vowel (Table 1).

Target items were elicited in question-answer pairs designed to control for two features likely to significantly impact sentence prosody: (i) phrasal position (medial vs. final), chosen because phrase-

Table 1: Sample target items for the study

	Short	Long
No tone	*k'eten* [kˀe.ˈten] 'hot' $[\sigma\text{ˈ}\sigma_{\mathrm{V}}]$	*qapoop* [qa.ˈpoːpʰ] 'our sleeping mat' $[\sigma\text{ˈ}\sigma_{\mathrm{V:}}]$
Tone	*qálaq* [ˈqálaqʰ] 'our cup' $[\text{ˈ}\acute{\sigma}\sigma_{\mathrm{V}}]$	*qajóoq* [qa.ˈχóːqʰ] 'our corn husk' $[\sigma\text{ˈ}\acute{\sigma}_{\mathrm{V:}}]$

final pitch rises (H%) are common in declarative utterances in Uspanteko, as in other K'ichean-branch Mayan languages [1, 2, 10]; and (ii) information structure (given vs. focused), which is known to affect sentence level prosody both crosslinguistically [14, 18] and in Mayan languages [1, 2, 9, 10].

To elicit target sentences, a research assistant who is a native speaker of Uspanteko read from a list of pre-prepared questions. Participants responded by translating a pre-prepared answer from Spanish to Uspanteko. We manipulated information structure and phrasal position by controlling (i) the question type, and (ii) position of the target word in the response. Content questions (e.g. *Nen tiqakoj re ojlejoon?* 'What do we use to make tortillas?') were used to elicit target words in a broad focus context (e.g. *Tiqakoj* ***qaxoot*** *re ojlejoon.* 'We use our **comal** to make tortillas.'). Polar questions (e.g. *Tiqasu'* ***qaxoot****?* 'Do we clean our **comal**?') were used to prompt answers with target words in a discourse-given context (e.g. *Ji'n, tiqasu'* ***qaxoot****.* 'Yes, we clean our **comal**'). Responses were varied so that each target word would be produced in both sentence-medial and sentence-final position, within each discourse-structure condition.

There were 12 target items, 3 in each condition shown in Table 1, produced in 4 contexts each (2 discourse conditions $\times$ 2 sentential positions) for a total of 48 productions per participant. Speakers often produced more than one usable repetition of a target sentence, resulting in 505 total utterances.

Eight native speakers of Uspanteko participated in the study (24-74 years old, median 42, SD 14; 2 male, 6 female). Recordings were made in a quiet room with a headset microphone (Audio-Technica ATM73a) and solid-state portable recorder (Zoom H5), at a 48 kHz sampling rate with 24 bit quantization. The recordings were subsequently downsampled to 16 kHz for forced alignment and acoustic analysis (§3).

3. ANALYSIS

Pitch contours on target words and sentences were analysed in PRAAT [5]. First, transcriptions were semi-automatically time-aligned using forced alignment [11]. Time-normalized pitch contours were then automatically extracted in ERB—a scale which corresponds well to pitch perception [19]—using a script specifying by-speaker pitch ranges [8]. Pitch contours for each speaker were z-score normalized, and outliers > 2.5 SD from the mean removed.

4. RESULTS

We begin with a graphical exploration of our results. Fig. 4 shows time-normalized pitch contours for entire utterances, summarized by means of a loess smoothing algorithm [24]. Dashed gray lines show utterances with tonal target words, and solid black lines show utterances with non-tonal target words. Grey bands around smoothed pitch contours indicate 95% confidence intervals. Columns group utterances by discourse condition, while rows group utterances by the sentential position of the target word. We expect intonational factors to interfere least with the realization of lexical tone in the GIVEN/MEDIAL condition (upper-left corner of Figs. 4-6).

Figure 4: Utterance-level contours.

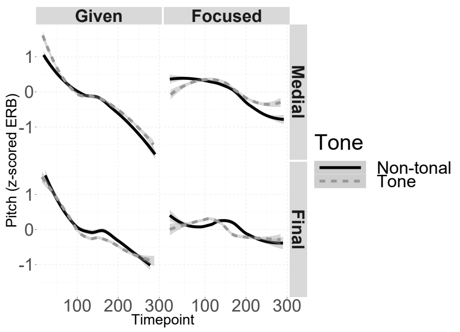
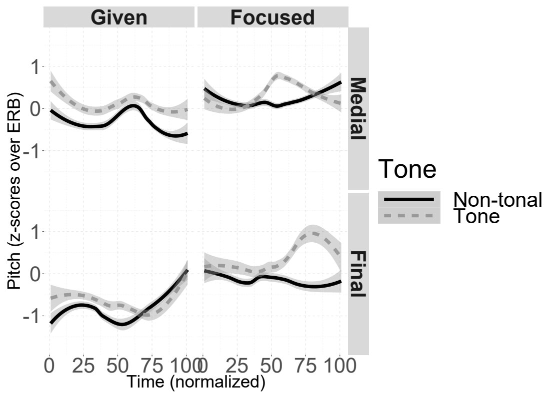


Fig. 4 shows a general downtrend across utterances, which is more pronounced when the target word is GIVEN. This suggests that focus is associated with raised pitch on the target word, and/or inhibition of downtrend. From these utterance-level pitch tracks, we also see initial evidence for an interaction between lexical tone and utterance-level intonation, inasmuch as the tonal/non-tonal contrast is somewhat more pronounced in focal contexts.

Turning to word-level pitch contours, we first consider pitch contours in words with final long vowels (Table 1). Fig. 5 shows that in all contexts, with the possible exception of the GIVEN/FINAL condition, words with tonal long vowels tend toward higher pitch, as expected. In the GIVEN condition, tonal and non-tonal words contrast primarily in pitch level, showing roughly the same pitch contours. In the FOCUSED condition, both tonal and non-tonal words show raised baseline pitch relative to their GIVEN counterparts. This difference plausibly reflects pitch range expansion and/or raising associated with focus. Importantly, the pitch contours of tonal and non-tonal words differ clearly under focus: in the FOCUSED/MEDIAL condition, tonal words have a pronounced pitch peak which is absent from non-tonal words; and in the FOCUSED/FINAL condition, tonal words show a more prominent pitch rise. This may indicate that focus prosody sharpens lexical tone contrasts, perhaps as a reflex of the pitch range expansion and/or raising seen under focus (Figs. 4, 5). The flat (rather than rising) contour for non-tonal words in the FOCUSED condition may thus constitute evidence in favor of analyzing Uspanteko tone as a privative [H] vs. [∅] contrast, instead of binary [HL] vs. [LH] [13]. The GIVEN/FINAL condition shows clear evidence of an H% (or LH%) boundary tone in utterance-final position, which swamps lexical tone distinctions on long vowels, and which is itself somewhat obscured in the FOCUSED/FINAL condition by the pitch raising associated with focus.

Figure 5: Target words with final long vowels.

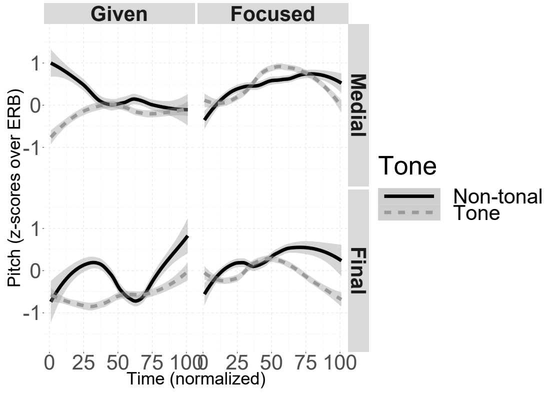


Tonal contours look a bit different for words which contain only short vowels. Recall that in words of this type, tone causes stress to retract from its default final position to the penult, where it coincides with tone (e.g. *wájaj* [ˈwá.χaχ] ‘my sugarcane’; see also Table 1 and Figs. 1, 7). For tonal words with no long vowels, we thus expect a relatively early pitch rise, which should also be absent on non-tonal words (unless intonation interferes).

Setting aside the GIVEN/FINAL condition for a moment, tonal words containing only short vowels generally show an early rise, leading into a fall over roughly the second syllable of the word. This pattern is again clearest in focal contexts, consistent with our earlier finding that focus tends to sharpen tonal contrasts (Fig. 5).

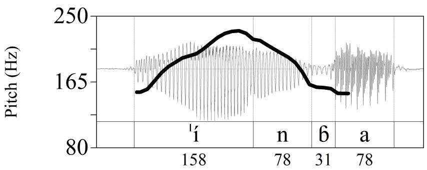
Though the pitch contours in Figs. 5 and 6 are somewhat idealized, being summary representations of pitch tracings from many different words and utterances, they are entirely consistent with pitch tracings taken from individual tokens, and from past work on Uspanteko [3, 4]. Figs. 1 and 7 illustrate

Figure 6: Target words with final short vowels.



the characteristic pattern for tonal words containing only short vowels, which typically have a large pitch rise in the first, tone-bearing, syllable, followed by a rapid fall into the second syllable, as in Fig. 6.

Figure 7: *ínb'a* [ˈín.ɓa] 'my head'.



This characteristic rise-fall pattern is obscured somewhat in the GIVEN/FINAL condition, where an utterance-final H% boundary tone on the final syllable produces a high plateau or slight rise from the lexical [H] tone to the final syllable of the word.

The non-tonal words in Fig. 6 show a more varied set of pitch contours, suggesting greater influence from intonational factors. In the FOCUSED condition we see a relatively smooth upward trajectory, which attribute to pitch raising associated with focal prosody. The GIVEN/MEDIAL condition shows a declining pitch trajectory, which likely reflects the general downtrend seen in Fig. 4. We note that this condition—which should show the least influence from intonational targets—provides evidence against the view that toneless words are instead specified with an [LH] pitch contour [13] (Indeed, we believe [13] mistakenly treats the final H% boundary tone as part of the lexical specification of toneless words—a major risk of investigating tone through lists of words produced in isolation [17], Figs. 2, 3.)

In the GIVEN/FINAL condition, we again see evidence of a final rise triggered by an H% boundary tone. Strangely, we also see raised pitch in the first half of non-tonal words. We aren't entirely sure how to interpret this contour, but speculate that it may reflect anticipatory raising conditioned by the utterance-final H%. In phonological terms, it may be that the H% tone associates with the final two *moras* of the word, i.e. a final long vowel, or the last two syllables in words with only short vowels [3]. (In this case, H% might be better-described as a 'phrase accent' than a boundary tone [12, 20].)

To briefly summarize our main results, we find that focus leads to (i) raised pitch on target words, (ii) inhibition of downtrend, and (iii) sharpening of lexical pitch contrasts. An utterance-final H% boundary tone seems to neutralize the distinction between tonal and non-tonal words with final long vowels (Fig. 5). This follows straightforwardly from the assumption that lexical tone on long vowels reflects a privative [H] vs. [∅] contrast in the final syllable, which is neutralized to [H] vs. H% when toneless long vowels acquire an H% tone in final position. However, this contrast seems to re-emerge under focus, suggesting that lexical tones and boundary tones are enhanced differently on focused words. In particular, focus seems to privilege the expression of lexical tone over the expression of intonational contours.

For words containing only short vowels, lexical tone is realized with an early rise on the penult, falling into the final syllable. This pitch contour is sharpest under focus, and reduced somewhat in discourse-given words. In the GIVEN/FINAL condition, tonal words show a continous, gradual rise, suggesting that the final syllable has acquired an H% boundary tone, yielding an [H]+H% sequence. Again, we believe these patterns follow straightfowardly from the assumption that lexical tone reflects a privative [H] vs. [∅] contrast on the penult in words containing only short vowels.

5. CONCLUSIONS

The interaction of tone and intonation in Uspanteko provides important insight into the phonological analysis of lexical tone. Furthermore, understanding the intonational structure of Uspanteko is crucial to future studies of the phonetics of word- and phrase-level prosody in this language, as intonational factors impinge on the realization of tone and stress in several distinct ways.

6. ACKNOWLEDGEMENTS

This research was supported by NSF grants BCS/DEL-1757473 (to Bennett) and BCS/DEL-1551666 (to Henderson). The authors thank the Comunidad Lingüística Uspanteka for their support with this project, Rosa Lidia Ajpoop for her help with data collection, and the speakers who participated in the study.

REFERENCES

[1] Baird, B. 2014. *An acoustic analysis of contrastive focus marking in Spanish-K'ichee' (Mayan) bilingual intonation*. PhD thesis University of Texas, Austin Austin, TX.

[2] Bennett, R. 2016. Mayan phonology. *Language and Linguistics Compass* 10(10), 469–514.

[3] Bennett, R., Henderson, R. 2013. Accent in Uspanteko. *Natural Language & Linguistic Theory* 31(3), 589–645.

[4] Bennett, R., Henderson, R. 2014. A pilot study on the acoustics of stress and tone in uspanteko. Poster presented at Form and Analysis in Mayan Linguistics (FAMLi) III, Colegio de México, Mexico City, Dec. 2014.

[5] Boersma, P., Weenink, D. 2016. Praat: doing phonetics by computer. Computer program. http://www.praat.org/.

[6] Campbell, L. 1977. *Quichean linguistic prehistory* volume 81 of *University of California Publications in Linguistics*. Berkeley, CA: University of California Press.

[7] Can Pixabaj, T. 2007. *Gramática descriptiva Uspanteka*. Antigua, Guatemala: Oxlajuuj Keej Maya' Ajtz'iib' (OKMA).

[8] De Looze, C., Rauzy, S. 2009. Automatic detection and prediction of topic changes through automatic detection of register variations and pause duration. *Proceedings of INTERSPEECH 2009*. ISCA Archive 2919–2922.

[9] DiCanio, C., Bennett, R. To appear. Prosody in Mesoamerican languages. In: Gussenhoven, C., Chen, Y., (eds), *The Oxford Handbook of Language Prosody*. Oxford, UK: Oxford University Press.

[10] England, N., Baird, B. 2017. Phonology and phonetics. In: Aissen, J., England, N., Zavala Maldonado, R., (eds), *The Mayan Languages*. New York: Routledge 175–200.

[11] Gorman, K., Howell, J., Wagner, M. 2011. Prosodylab-aligner: A tool for forced alignment of laboratory speech. *Canadian Acoustics* 39(3), 192–193.

[12] Grice, M., Ladd, D., Arvaniti, A. 2000. On the place of phrase accents in intonational phonology. *Phonology* 17(2), 143–185.

[13] Grimes, J. 1972. *The phonological history of the Quichean languages*. Carbondale, IL: University Museum, Southern Illinois University.

[14] Gussenhoven, C. 2004. *The phonology of tone and intonation*. Cambridge, UK: Cambridge University Press.

[15] Himmelmann, N., Ladd, D. 2008. Prosodic description: An introduction for fieldworkers. *Language documentation and conservation* 2(2), 244–274.

[16] Hyman, L. 2009. How (not) to do phonological typology: the case of pitch-accent. *Language Sciences* 31(2-3), 213–238.

[17] Hyman, L. 2014. How to study a tone language, with exemplification from Oku (Grassfields Bantu, Cameroon). *Language Documentation and Conservation* 8, 525–562.

[18] Ladd, D. 2008. *Intonational phonology*. Cambridge, UK: Cambridge University Press 2nd edition.

[19] Moore, B., Glasberg, B. 1983. Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *Journal of the Acoustical Society of America* 74(3), 750–753.

[20] Pierrehumbert, J., Beckman, M. 1988. *Japanese tone structure*. Cambridge, MA: MIT Press.

[21] Remijsen, B., Martis, F., Severing, R. 2014. The marked accentuation pattern of Curaçao Papiamentu. In: Jun, S.-A., (ed), *Prosodic Typology II: the Phonology of Intonation and Phrasing*. Oxford, UK: Oxford University Press 302–323.

[22] Remijsen, B., van Heuven, V. 2005. Stress, tone and discourse prominence in the Curaçao dialect of Papiamentu. *Phonology* 22(2), 205–235.

[23] Richards, M. 2003. *Atlas lingüístico de Guatemala*. Instituto de Lingüístico y Educación de la Universidad Rafael Landívar.

[24] Wickham, H. 2009. *ggplot2: elegant graphics for data analysis*. New York: Springer.