Stress, tone, and intonation in Choguita Rarámuri

Gabriela Caballero, Yuan Chai, and Marc Garellek
University of California, San Diego

Choguita Rarámuri is an endangered Uto-Aztecan language from Northern Mexico featuring a complex prosodic system, with both stress and tone in its word prosody, as well as postlexical intonation. Based on the analysis of instrumental data obtained through field research, in this chapter we describe how lexical tones in this language are realized in terms of fundamental frequency (f0); how tonal realization varies by phrasal position in declarative sentences; we also show how correlates of voice quality help increase the acoustic differentiation across tones for some speakers. In addition, we describe how tone and intonation interact and show that the incorporation of f0 over pre- and post-tonic syllables generally improves the statistical discriminability of the three tones in utterance-medial position. Therefore, tonal contrasts in this language are as much about paradigmatic differences across the tone categories during the tone-bearing syllable, as they are about syntagmatic changes over at least three syllables centered around the tone-bearing syllable.

Keywords: stress, tone, intonation, voice quality, Uto-Aztecan

1. Introduction

This chapter addresses the prosody of Choguita Rarámuri, an endangered Uto-Aztecan language from Northern Mexico. The prosodic system of Choguita Rarámuri is typologically relevant for several reasons: (i) it is the only language within the Uto-Aztecan language family documented to date with a three-way lexical tone system in addition to stress; (ii) lexical tonal contrasts involve narrow f0 differences and a high degree of speaker variation; and (iii) fundamental frequency (f0) is deployed in postlexical intonation patterns, resulting in different accommodation strategies.

Based on the analysis of instrumental data obtained through field research, we assess two questions. First, we seek to establish the role of fundamental frequency and phonation in distinguishing the lexical tones of Choguita Rarámuri: cross-linguistically, tone contrasts are usually defined primarily as pitch distinctions (Yip 2002, Hyman 2011). During voicing, pitch is primarily cued by f0. However, it is also clear that other phonetic parameters, such as phonation (Kuang 2013) and duration (Gandour 1977, Gordon 2001, Zhang 2002, Faytak & Yu 2011), are crucially involved in tone realization. Second, we seek to describe how lexical tone and postlexical intonation interact in the language. Here we focus exclusively on declarative sentences, in which certain lexical tones prevent the occurrence of a right-edge boundary tone associated with such sentences. In languages with both lexical tones and intonation, these frequently interact in complex and language-specific ways (see, e.g., discussion in Gussenhoven 2004). Such interactions have ramifications for modeling how tone and intonation relate to and interact with one another (Gibson 2013).

Both tone and intonation are understudied for Uto-Aztecan languages, though there has been some recent work for varieties of Nahuatl (Guion et al. 2010, Patiño 2014, Aguilar 2020) and for Northern Tepehuan (Gil & Carrillo 2019). Our chapter therefore contributes to the description of suprasegmental structure within the Uto-Aztecan language family, as well as to our understanding both of the phonetic multidimensionality of tone contrasts across languages and of cross-linguistic patterns of tone-intonation interactions.

Rarámuri is spoken by approximately 85,000 to 100,000 people in the northern Mexican state of Chihuahua (INEGI 2010, Embriz Osorio & Zamora Alarcón 2012). Choguita Rarámuri (henceforth CR), the focus of this study, belongs to a 'Central' dialect area (ISO: [tar]) within a dialect continuum. The location of Choguita within the municipality (Sp. *municipio*) of Guachochi is shown in Map 1.

Map 1: Location of Choguita (Guachochi municipality, Chihuahua, Mexico).

There were approximately 1,000 speakers of CR over a decade ago (Casaus 2008), but heightened political violence and economic instability have displaced an increasing number of speakers in recent years. There is no data yet available about the current state of intergenerational transmission of the language. The CR data presented in this paper comes from a controlled production study undertaken as part of a language description and documentation project (Caballero 2017) carried out since 2003 by the first author.

CR has a 'hybrid' word-prosodic system, featuring both lexical tone and stress as independent phonological systems, as in Ma'ya (Austronesian; Remijsen 2001, 2002), the Neo-Stovakian dialect of Serbo-Croatian (Slavic; Zec 1999), Fasu (Kutuban, Papua New Guinea; May & Loeweke 1964)), and many Scandinavian languages (Riad 2006), among others; for an overview, see Hyman 2006. Stress in CR is lexically contrastive and may exhibit morphologically conditioned shifts. The examples in (1) show that stress is lexically contrastive in this language:

(1)	a.	ˈsàwa	'smell'	saˈwá	'leaf'
	b.	ˈkôt͡ʃi	'pig'	koˈt͡ʃî	'dog'
	c.	ˈkóri	'visit'	koˈlî	'chile pepper'

Stress is restricted to appear within an initial-three syllable window on the first, second, or third syllable of the word. There is no evidence for secondary stress assignment; i.e., the system is non-iterative (Caballero 2011). There are no constraints on the alignment of stress with respect to word edges nor any constraints against stress clashes across word boundaries.

The CR tone system involves three lexical tones: falling (/HL/, <ô>),[1] low (/L/, <ò>), and high (/H/, <ó>). Lexical tonal contrasts are associated exclusively with surface stressed syllables, i.e., there is only one lexical tone per prosodic word and stressless syllables lack lexical tone. Tone is also obligatory; i.e., there are no toneless words. Tone also encodes grammatical distinctions in CR (e.g., imperative mood may be marked by a L tone), and grammatical tones may override lexical tonal contrasts (Caballero & German to appear). In addition to having a single stress, the prosodic word in CR may be characterized as follows: (i) prosodic words are minimally bimoraic (monomoraic open-class words undergo vowel lengthening when unaffixed); (ii) prosodic words are vowel-final (stress-dependent vowel deletion targets non-final vowels); and (iii) glottal stop is only licensed within the first two syllables of the prosodic word (Caballero 2008).

The data in (2) exemplify the three-way tonal contrast of CR (Caballero & Carroll 2015:465).

(2)	a.	ˈtô	'bury'	ˈtò	'take'
	b.	ˈpá	'throw'	ˈpà	'bring'
	c.	koˈlî	'chile pepper'	koˈlì	'spatial root'

Tone is contrastive; it is not conditioned by voicing or laryngeal setting of the preceding consonant, nor any other phonological features, and stressed syllables may bear any of the three

[1] This falling tone is a phonological unit, a contour tone that is not decomposable into two tonal primitives.

lexical tones. The tone-bearing unit is the mora: falling tones have their high target on the stressed syllable, with the fall starting in the tonic and continuing through a post-tonic syllable, if there is one (Caballero & Carroll 2015). As discussed by Garellek et al. (2015) and in more detail in Section 3 below, H tones may also spread their high f0 to the post-tonic.

The restricted distribution of lexical tonal contrasts to stressed syllables attested in CR is a characteristic shared by all other tonal Uto-Aztecan languages, all of which have hybrid stress and tone systems.[2] Unlike CR, however, these systems are analyzed as involving a binary (privative or equipollent) contrast, which suggests that the three-way split in CR tone may be a recent development. Tone is indeed an innovation in the Uto-Aztecan language family: Proto-Uto-Aztecan is reconstructed to have had only stress (Munro 1977), and the diachronic development of tone has been linked to the loss of laryngeal features in Uto-Aztecan languages such as Hopi (Manaster-Ramer 1986), Northern Tepehuan (Shaul 2000), and some varieties of Balsas Nahuatl (Guion et al. 2010).

While lexical tone is only assigned to stressed syllables, stress and tone in CR are phonologically distinct systems that are also encoded through independent acoustic means. Analysis of f0, intensity, spectral tilt and duration measurements reveals that tone is primarily encoded through f0, while stress is encoded primarily through segmental duration, with both stressed vowels and onsets of stressed syllables exhibiting greater duration than their unstressed counterparts.[3] Additionally, intensity is a reliable acoustic cue of stress for some speakers (Caballero & Carroll 2015). While f0 is a reliable cue to lexical tonal contrasts in CR, it is important to note that the f0 differences reported in our previous work are relatively small compared to other tonal languages that may be equally restricted in terms of the number of tonal contrasts in their inventory: Caballero & Carroll (2015) report that H and L tones in CR differ by approximately 2 semitones, while H and HL tones are primarily differentiated through pitch slope in tonic syllables. Other languages with comparable tone inventories (e.g., Hausa, with the same H, HL, L lexical tones) may have comparable narrow pitch height differences (Maddieson 1978); nevertheless, these results raise questions as to whether other acoustic dimensions are involved in the encoding of lexical tones in CR.

In addition to encoding lexical contrasts, f0 is deployed in the intonational phonology of CR. First, declarative sentences typically end in a high f0 target if the final lexical tone is H or L.

---

[2] These languages include Hopi (ISO code: hop) (Manaster-Ramer 1986), Northern Tepehuan (ISO code: ntp) (Tepiman) (Woo 1970), Balsas Nahuatl (ISO code: ngu) (Aztecan) (Guion et al. 2010), Cora (ISO code:crn) (Corachol) (McMahon 1967), Huichol (ISO code: hch) (Corachol) (Grimes 1959), and Cahitan ((Taracahitan) Yaqui (ISO code: yaq) (Demers et al. 1999]) and Mayo (ISO code: mfy) (Hagberg 1989)). Except for Balsas Nahuatl, these languages are located in the U.S. Southwest and Northern Mexico, an area where tone has also developed in other language families (Caballero & Gordon to appear).

[3] An anonymous reviewer points out that, given the culminative distribution of tone in CR, f0 may be an indication of stress (even if indirect). This would make the CR prosodic system closer to some definitions of pitch accent languages, where f0 would be a correlate of syntagmatic prominence. However, tone in CR is independent from stress phonologically and acoustically, even if tone distribution is dependent on stress: f0 does not play a role in distinguishing stressed vs. unstressed syllables (Caballero & Carroll 2015). Furthermore, we note that some languages that are analyzed as pitch accent in the literature (e.g., Tokyo Japanese) are stressless and have invariant tonal contours on accented syllables.

This is attributed to a H% boundary tone at the edge of the Intonation Phrase (IP) (Garellek et al. 2015). Second, there is also evidence of postlexical tonal targets associated with lexical tones in declarative intonation: all three lexical tones may be preceded by an optional pitch target: a low target before high and falling tones, and a high target before low tones. These tonal targets, which Garellek et al. (2015) labelled “rhythmic lead tones,” are crucially not associated with word boundaries or other prosodic domains, as they appear to be variable in their realization, generally appearing in the pre-tonic syllable. Thus, while the contrast between lexical tones in CR in terms of f0 in tonic syllables is relatively narrow for a low-density tonal system, our prior work on CR intonation suggests that discriminability of lexical tones in terms of f0 involves a window larger than stressed syllables.

In addition to f0, lexical tonal encoding in CR involves non-modal phonation and lengthening in different intonational contexts. Previous work involving analysis of both acoustic and electroglottographic data (reported in Aguilar et al. 2015), reveals preliminary evidence for tone-specific effects at prosodic boundaries: (i) HL tones may be rearticulated phrase-finally (e.g. /î/ → [íʔì], with frequent devoicing after the glottalization) and (ii) L tones are lengthened phrase-finally.

In sum, the goals of this chapter are as follows: (1) to describe how CR lexical tones are realized in terms of f0, and (2) how tonal realization varies by phrasal position in declarative sentences; (3) to show how correlates of voice quality help increase the acoustic differentiation across tones, at least for some speakers; and (4) to describe how tone and intonation interact over the course of a declarative sentence. In the following Section 2, we provide a brief overview of the methods used in the subsequent sections; In Section 3, we focus on the f0 patterns of the lexical tones; in Section 4 we turn to speaker variation in f0 and tone multidimensionality; in Section 5, we turn to sentence-level intonation; we then conclude with an overall discussion of the findings.

2. Method

The data examined in this paper were obtained in the course of fieldwork undertaken in 2014 with four CR native speakers, two females (BFL, age 26; RIC, age 35) and two males (SFH, age 41; CFM, age 36). Three of these speakers have always lived in Choguita (BFL, SFH and CFM). One female speaker (RIC) was born and raised in Norogachi, a community located approximately 20 kilometers east of Choguita, but had lived continuously in Choguita for over 20 years at the time of data recording. All speakers are Rarámuri-Spanish bilinguals, but dominant in Rarámuri, and use CR in their daily interactions with other community members. Three of the four speaker participants (SFH, BFL and RIC) are long-standing collaborators of the Choguita Rarámuri language description and documentation project, and have substantial experience as consultants for linguistic elicitation. All speakers are literate in Spanish and are able to read and write in Rarámuri using a Spanish-based orthography.

The corpus of analyzed tokens includes target words representative of the relevant lexical tone categories (HL, H and L). Targets include monomorphemic nouns and verb roots with a single unstressed suffix, in order to avoid morphologically conditioned stress alternations. Both noun and verb targets are balanced for laryngeal setting of the onset of stressed syllable (h, ʰC, ʔC, C, ʔ). We controlled for phrase position by recording target forms in phrase-final and phrase-medial position in declaratives sentences with intended broad focus.[4] (Phrase-initial contexts were excluded to ensure comparison between targets in equivalent broad focus contexts, avoiding topicalized and fronted elements with narrow focus.) The timing and sequencing of lexical and postlexical tones were assessed by varying the length of words and phrases and stress location (penultimate or final)[5] within prosodic words and with respect to phrasal boundaries. For our analysis of sentence-level intonational patterns (described in Section 5), we also elicited declarative sentences composed of all identical lexical tones (all HL, all H or all L) aligned with stressed syllables. Sample sentences are provided in (3); targets in phrase-final position have penultimate stress (3a-c) or final stress (3d), and are underlined in these examples.

(3)	a.	maˈnuêli	muˈnî	koˈʔá-li			
		Manuel	beans	eat-PST			
		‘Manuel ate beans’					
	b.	maˈnuêli	roˈʔà-li	aʔˈlì	koˈʔá-li		
		Manuel	lay.down-PST	and	eat-PST		
		‘Manuel laid them down (at the altar) and ate’					
	c.	maˈnuêli	paˈt͡ʃî	roˈʔà-li	aʔˈlì	muˈnî	koˈʔá-li
		Manuel	corn	lay.down-PST	and	beans	eat-PST
		‘Manuel laid the corn down (at the altar) and ate beans’					
	d.	maˈnuêli	riˈwá-li	baˈʰt͡ʃî			
		Manuel	see-PST	squash			
		‘Manuel saw a squash’					

Recordings were made in Choguita (Chihuahua, Mexico). Participants read the sentences from a list and repeated each sentence three times. (The sentences were written in Rarámuri using a Spanish-based orthography participants are familiar with. Stress, but not tone, was marked. Spanish translations were also provided.) The data were recorded digitally on a solid state recorder (Marantz PDM660) with a head-mounted microphone (AKG C520) and an omnidirectional lavalier microphone at a sampling rate of 44.1 Hz. The f0 data (in Sections 3 and

[4] As described for other Rarámuri varieties (Burgess 1984, Valdez Jara 2013, Morales Moreno 2016), SOV is the canonical word order in CR, though SVO is also frequently attested in declarative sentences with broad focus.
[5] Noun targets have both final and penultimate stress, while verbs all have penultimate stress in the root final syllable before the unstressed past *-li* suffix.

4) and voice quality measurements (in Section 4) were obtained automatically using VoiceSauce (Shue et al. 2011), using its default algorithms and settings. We excluded 122 tokens because their f0 were outliers 2.5 standard deviations from a speaker's mean. In total, we analyze f0 patterns from 5891 data points; Figures 1-4 represent data from 1181, 3350, 5891, and 1181 data points, respectively.

3. Tonal patterns: F0

In accordance with previous work (Caballero & Carroll 2015), we find relatively small f0 differences between the three tones, at least when measured over the tonic vowel. Figure 1 shows the f0 trajectory over vowel thirds for (penultimate or word-final) tonic vowels in utterance-medial position, averaged over the four speakers (speaker differences will be discussed in Section 4). The difference in f0 between the L and H tones is approximately 3 semitones; the HL tone's f0 falls between that of the H and L tones. Additionally, the HL's falling contour is very small, averaging less than 1 semitone, and only during the latter third of the vowel. The lower f0 onset of the HL tone relative to the H one goes against previous findings (Caballero & Carroll 2015), but this may be due to differences in the sentence types that were collected, and the fact that target tones in that study came from two phrasal positions.

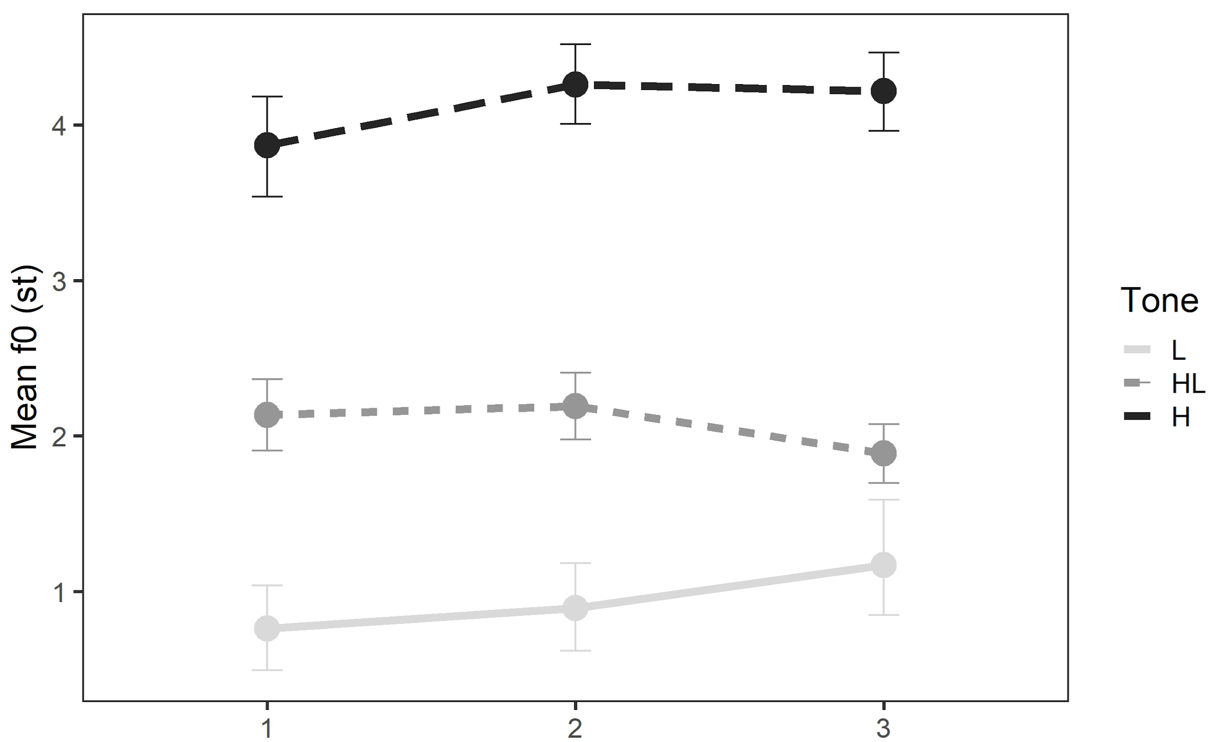


Figure 1: Mean f0 tracks (in semitones with base 100 Hz for male speakers, with base 200 Hz for female speakers) over utterance-medial tonic vowels, averaged over vowel thirds.

In our earlier work on CR tone and intonation (Caballero & Carroll 2015, Garellek et al. 2015), we found that the tones involve changes in f0 over a window larger than just the tonic vowel. Notably, we found that H tones spread their high f0 to the post-tonic, even if it is in the following word; HL tones realize their falling f0 largely over the post-tonic, and pre-tonic positions are often low-pitched before the H and HL tones, and high-pitched before the L tones.

Those results are borne out in the present data set. Figure 2 illustrates the f0 trajectory from the pre-tonic, tonic, and post-tonic vowels (in utterance-medial positions, including only tokens whose pre- and post-tonic syllables were toneless). For H tones, we see that the pre-tonic vowel has a low-pitched target, and that the post-tonic vowel retains the high-f0 target from the tonic-- i.e., there is rightward H-tone spreading. For HL tones, we see that the pre-tonic vowel has a low-pitch target, and that the falling contour is realized over the tonic and post-tonic vowels. For L tones, we see a fall in f0 centered around the onset of the tonic, followed by a rise starting in the second half of the tonic and continuing through the post-tonic. Therefore, the L tone reaches its low target at the onset of the tonic (see also discussion in Garellek et al. 2015). And although it looks as though the f0 targets in pretonic position are comparable across the three tones, in Section 5 we argue that sentence-level f0 patterns reveal this to not be the case; L tones are optionally preceded by a high f0 target, with a subsequent fall into the L at the onset of the tonic syllable. Altogether, these findings show that CR tones, though dependent on and aligning with tonic syllables, affect more than just the tonic.

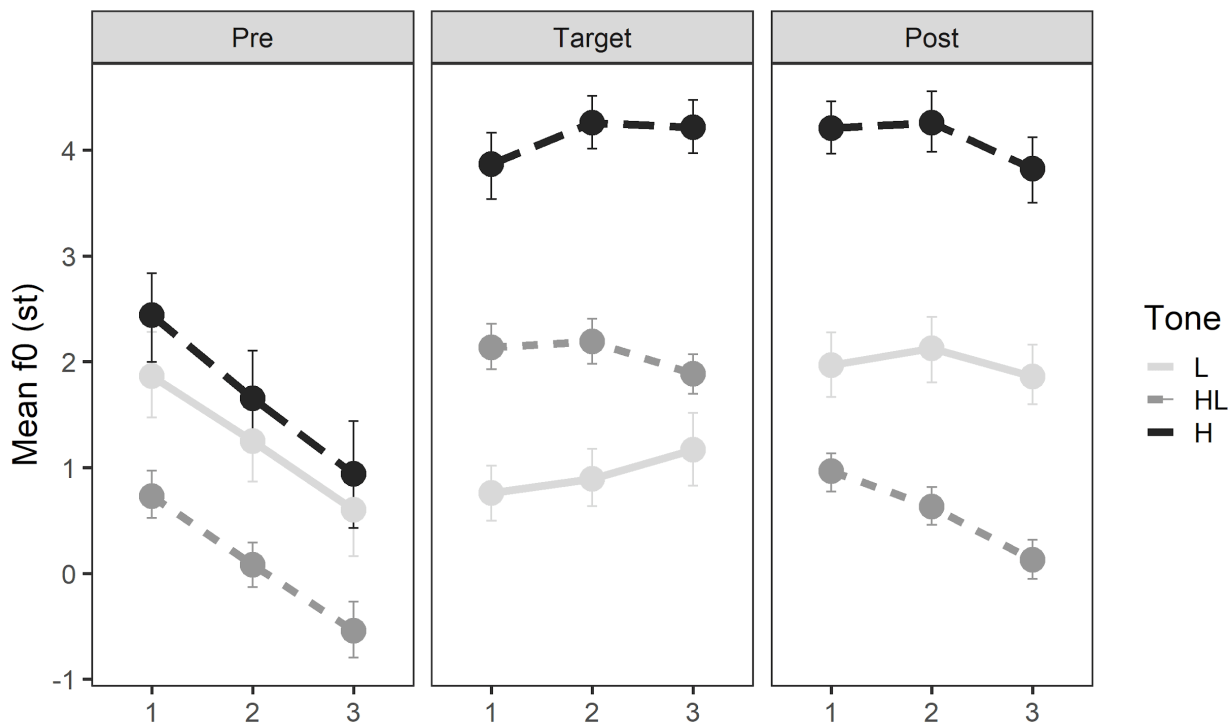


Figure 2: Mean f0 tracks (in semitones with base 100 Hz for male speakers, with base 200 Hz for female speakers) over utterance-medial pre-tonic, tonic, and post-tonic vowels, averaged over thirds for each vowel. Note that the middle panel is identical to Figure 1.

To assess the role of phrasal position on tone realization, Figure 3 illustrates the overall patterns for utterance-medial and utterance-final positions. (The reader should bear in mind that the top panels of Figure 3 are identical to panels of Figure 2; the f0 contours appear more compressed here because the scale of the y-axis has been enlarged.) This brings up an important finding about the pitch scale: in CR, utterance-final positions are associated with pitch range expansion, particularly because the L tone's post-tonic syllable exhibits a dramatic rise in f0. The HL tone in utterance-final position also shows a more pronounced fall, which is driven by the lowered low-pitch target.

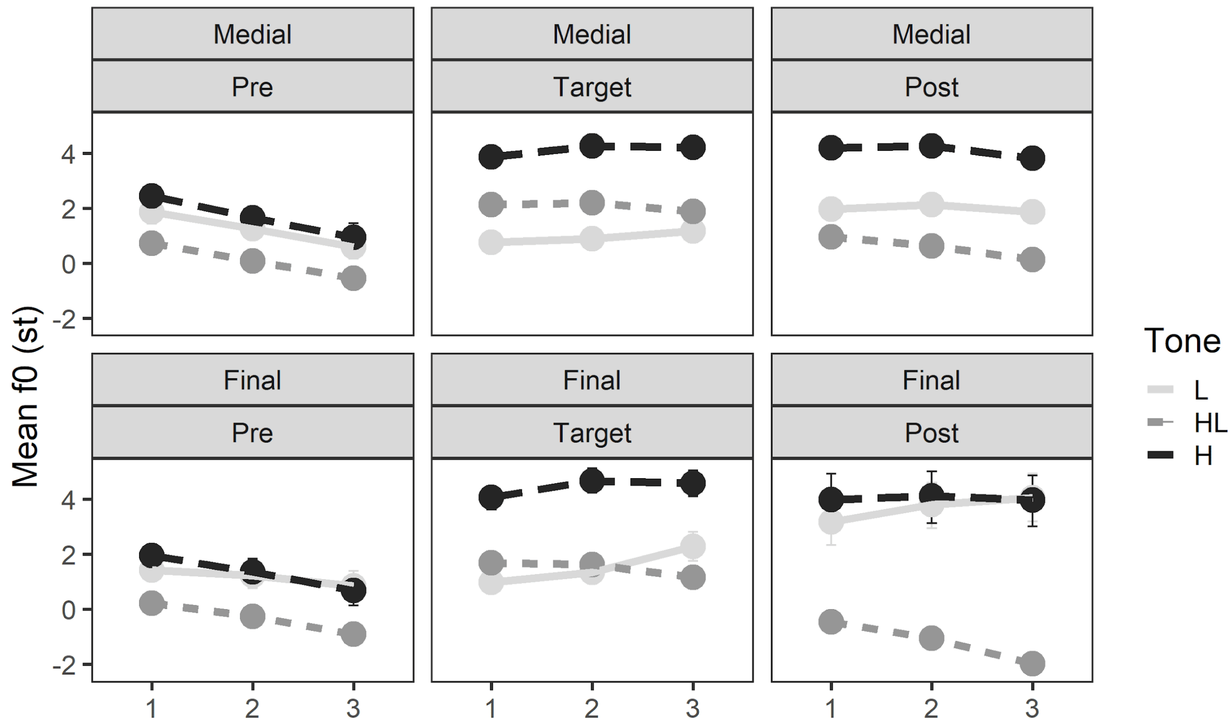


Figure 3: Mean f0 tracks (in semitones with base 100 Hz for male speakers, with base 200 Hz for female speakers) over utterance-medial (top) and utterance-final (bottom) pre-tonic, tonic, and post-tonic vowels, averaged over vowel thirds. Note that the top-middle panel is identical to that of the middle panel in Figure 2, only here it appears with a compressed f0 range.

We attribute the utterance-final pitch range expansion to the presence of a H% boundary tone for declaratives (Garellek et al. 2015). The boundary tone results in a higher f0 target than that of a preceding lexical H tone, and causes a rise from the utterance-final L through to the end of the utterance. In other words, in declarative sentences with a final H or L lexical tone, there is a continuous rise in f0 from the final lexical tone to the end of the utterance. But if the final lexical tone is HL, we find no evidence of a H% boundary tone. We return to these utterance-final patterns, as well as other intonational effects, in Section 5.

4. Speaker variation and tone multidimensionality

In this section, we focus on lexical tones in utterance-medial position, where the mean f0 differences were small (see Figure 1). Specifically, we look at the differences in tonal realization *across* speakers. Figure 4 illustrates how the paradigmatic distinctions across the three tones in tonic position differ by speaker. The distinction between H vs. L is robust for each speaker, but differences in semitones can be small. For instance, the largest mean difference in f0 between H vs. L tones is just over 4 semitones (for RIC); on average, the H and L tones differ by 3.2 semitones. (We note, however, that this is larger than the 2-semitone difference reported in Caballero & Carroll 2015.) Moreover, the difference between the HL tone and the other two tones is not robust over the tonic for most speakers. In the tonic position, HL tone has a similar contour to the H tone for two speakers (CFM and RIC), and a similar contour to the L tone for one speaker (SFH). For the remaining speaker (BFL), there is differentiation across the three tones in tonic position. Therefore, by looking at just the tonic syllable, and just f0, tonal distinctions are small in magnitude, and variable across speakers.

Here we seek to determine whether there are additional acoustic parameters (namely duration and voice quality measures) that help further differentiate the tones in tonic position. Although syntagmatic changes (i.e. changes from the pre-tonic to tonic, and from the tonic to post-tonic syllables) surely help to differentiate the three tones (see Figure 3 in the previous section), it is also possible that additional parameters are involved in strengthening the paradigmatic tonal contrasts in tonic position.

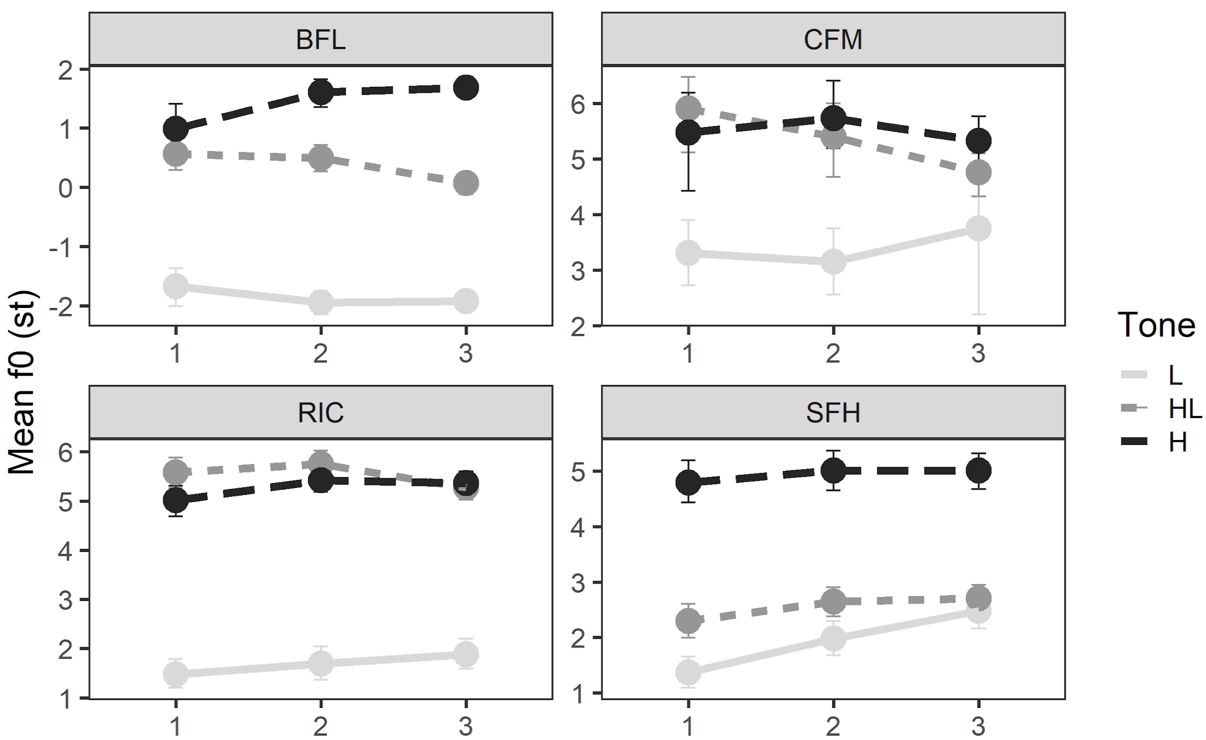


Figure 4: Mean f0 tracks (in semitones with base 100 Hz for male speakers, with base 200 Hz for female speakers) of utterance-medial tonic vowels, averaged over vowel thirds. Each facet represents a different speaker (women on left; men on right).

The acoustic parameters include F0 and voice quality measures. Voice quality is represented using two measures: H1*–H2* and the Harmonic-to-Noise-Ratio (HNR < 500Hz). H1*–H2* is the difference in amplitude between the first and second harmonics (corrected for formant frequencies and bandwidths). A lower H1*–H2* value is correlated with a more constricted laryngeal setting. HNR is a measure of noise, with lower values indicating a noisier, less periodic, signal. Together, H1*–H2* and HNR provide a straightforward way of distinguishing modal voice from both breathy and creaky voice qualities (Garellek, 2019).

For each of the acoustic parameters, we measured the mean over the target syllable (F0_mean, H1*–H2*_mean, HNR_mean). Further, to assess changes in the parameters over the tonic, we measured the difference (Δ) between the mean of the last third and the first third of the syllable (ΔF0, ΔH1*–H2*, ΔHNR).

In order to test which acoustic parameters differentiate the utterance-medial tones in tonic position, 1036 tokens were analyzed using linear discriminant analysis (LDA). Given that there are three tone categories, the output of the LDA provides two linear discriminants (LDs) that best discriminate statistically the three tones. These LDs include all the acoustic parameters provided in the analysis, with each parameter having a different coefficient strength. We incrementally increased the number of acoustic parameters in the LDA, to show how individual acoustic parameters affect the classification accuracy. We tested three models in total:

- Model 1: *Tone ~ Tonic F0*
- Model 2: *Tone ~ Tonic F0 + Tonic voice quality (VQ)*
- Model 3: *Tone ~ Tonic F0 + Pre-tonic F0 + Post-tonic F0*

Note that we combined the pre- and post-tonic f0 together, in order to get a sense of their combined role in discriminating the tones. We compared classification accuracy, which is assessed by determining how the LDA classified each token (as belonging to a particular tone) vs. the token's actual tone category. For instance, a 90% classification accuracy for H tones means that 90% of tokens bearing a high tone were classified as such by the LDA. Comparing the classification accuracy of Model 1 with Model 2 and 3, we can tell whether the voice quality of the tonic syllable and the f0 of the pre- and post-tonic syllables improve the differentiation of the tones. Further, by comparing Model 2 with Model 3, we can determine whether the voice quality of the tonic syllable or the f0 of the pre- and post-tonic syllables is more efficient in helping differentiating the tone of the tonic syllable.

In general, the classification accuracy increased when the tonic voice quality (Model 2) and the f0 of pre- and post-tonic syllables (Model 3) were added to the base model (Model 1), as shown in Figures 5 and 7-9. An exception was found for Speaker CFM. Adding the tonic voice quality to the base model did not improve its classification accuracy (cf. Model 2 vs. Model 1).

Overall, the LDA results indicate that the acoustic parameters used in differentiating the tonic tones differ in importance by speaker. For example, for Speaker BFL, Model 2 improved the classification accuracy more than Model 3 (as shown in Table 1). This indicates that voice quality is more important than the f0 of the pre- and post-tonic syllables in distinguishing the tones. Figure 6 shows the mean H1*–H2* and HNR values of the three tones for speaker BFL. H1*–H2* value discriminates HL and L tokens. HL tokens have higher H1*–H2* values than L tokens in general, indicating that HL tokens were breathier than L tokens. Despite the overlap between HL and H tokens, some of the HL tokens have higher H1*–H2* values and lower HNR values than the H tokens, indicating that some HL tokens are also breathier than H tokens. (The breathiness associated with some HL tones is independent of phrasing; cf. both panels of Figure 6.) In summary, for Speaker BFL, voice quality helps distinguish the three tones because HL tokens tend to be breathier than L and (some) H tones.

However, for Speaker SFH, RIC and CFM, Model 3 improved the classification accuracy more than Model 2 (as shown in Table 1). This indicates that the f0 of the pre- and post-tonic syllables are more important than the voice quality of the tonic syllable in distinguishing the tones.

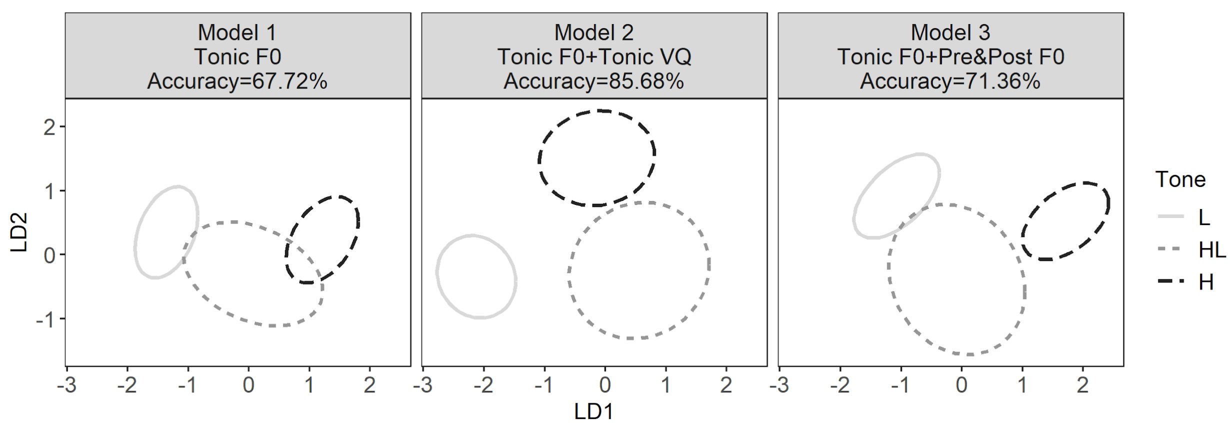


Figure 5: LDA space used for distinguishing the three tones for speaker BFL. The ellipses represent 50% confidence intervals around the mean of each tone group. Model 1 (left) includes only f0 over the tonic; relative to Model 1, Model 2 (middle) also includes voice quality measures over the tonic, and Model 3 (right) also includes the f0 in pre- and post-tonic syllables.

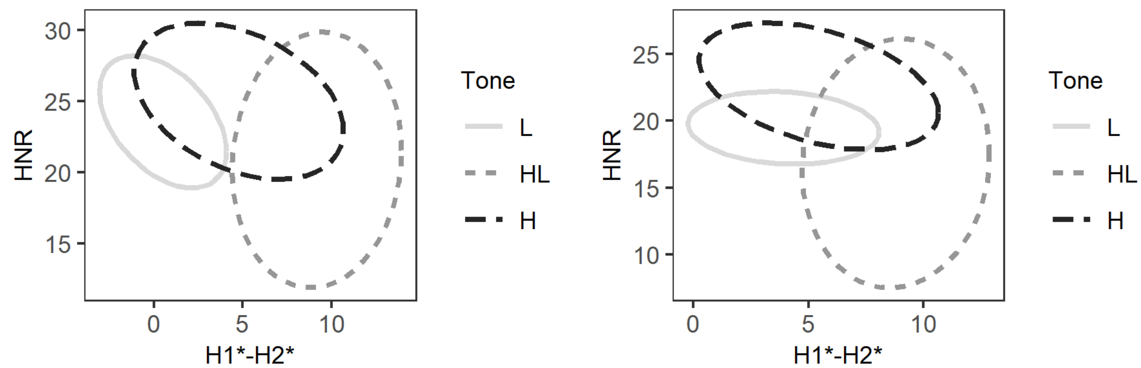


Figure 6: Voice quality space (H1-H2* by HNR) for three tones for speaker BFL, in utterance-medial (left) and -final (right) positions. The ellipses represent 50% confidence intervals around the mean of each tone group. The bottom-right corner of each figure represents breathier voice quality.*

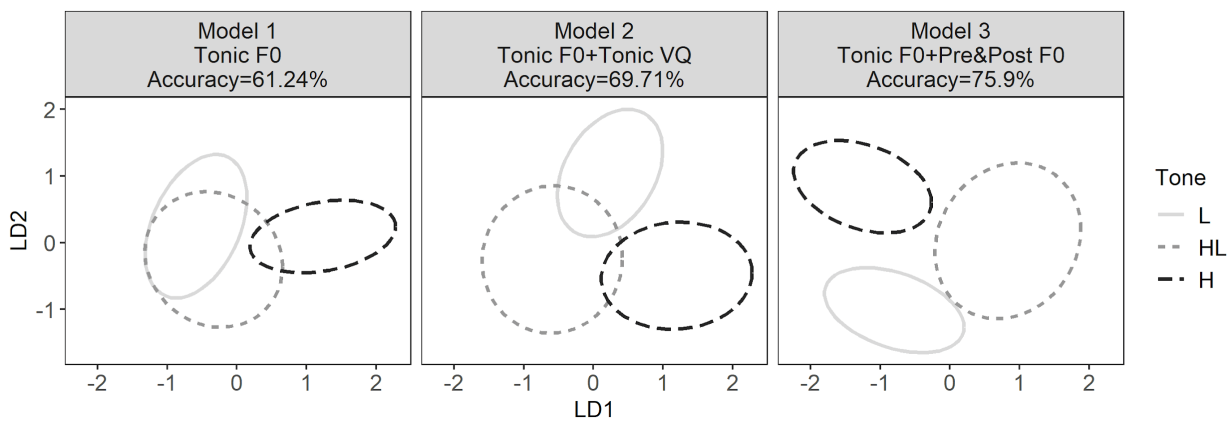


Figure 7: LDA space used for distinguishing the three tones for speaker SFH. The ellipses represent 50% confidence intervals around the mean of each tone group. Model 1 (left) includes only f0 over the tonic; relative to Model 1, Model 2 (middle) also includes voice quality measures over the tonic, and Model 3 (right) also includes the f0 in pre- and post-tonic syllables.

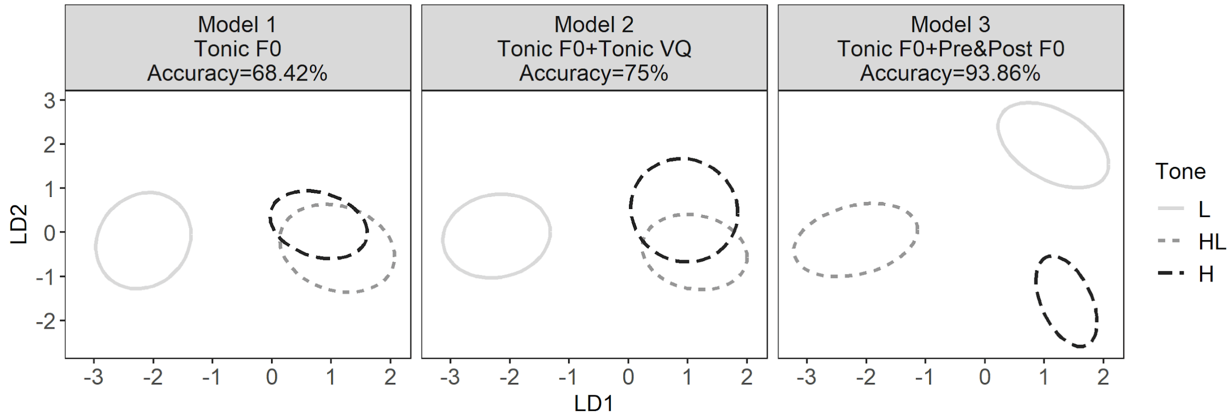


Figure 8: LDA space used for distinguishing the three tones for speaker RIC. The ellipses represent 50% confidence intervals around the mean of each tone group. Model 1 (left) includes only f0 over the tonic; relative to Model 1, Model 2 (middle) also includes voice quality measures over the tonic, and Model 3 (right) also includes the f0 in pre- and post-tonic syllables.

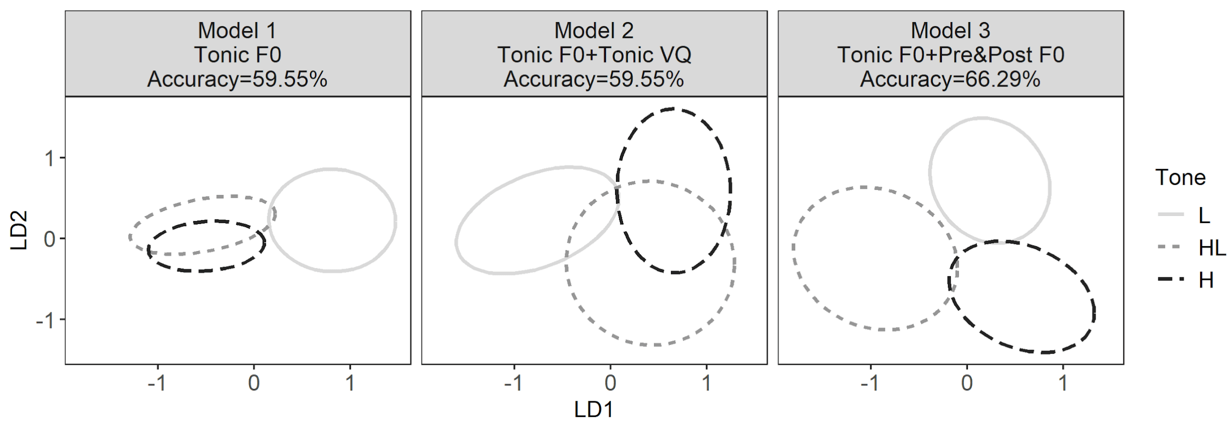


Figure 9: LDA space used for distinguishing the three tones for speaker CFM. The ellipses represent 50% confidence intervals around the mean of each tone group. Model 1 (left) includes only f0 over the tonic; relative to Model 1, Model 2 (middle) also includes voice quality measures over the tonic, and Model 3 (right) also includes the f0 in pre- and post-tonic syllables.

Table 1. Comparison among Models 1-3 in terms of classification accuracy

Speaker	Improvement in accuracy	Model 2 vs. Model 1	Model 3 vs. Model 1
BFL		17.96%	3.64%
SFH		8.47%	14.66%

RIC	6.58%	25.44%
CFM	0%	6.74%

In sum, for three of the four speakers whose speech is analyzed here, the LDA revealed that voice quality also helps differentiate the three tone categories. In particular, the inclusion of measures of voice quality resulted in an improvement in classification accuracy by nearly 18% for one speaker (BFL), because HL tones tended to be breathier than the H and L counterparts. Moreover, the addition of preceding and following f0 leads to modest (speaker BFL) to large (speaker RIC) improvements in tone classification.

5. Sentence-level intonational patterns

In Section 3, we showed how a three-syllable window provides a clearer picture of the lexical tone distinctions in CR than an exclusive look at the tonic. Further, in Section 4 we used a linear discriminant analysis to confirm that the inclusion of pre- and post-tonic f0 helps discriminate the three lexical tones from each other. Therefore, while tone is a property of tonic syllables in CR, its occurrence affects more than just a single syllable. In this section, we will illustrate f0 patterns over the course of the entire CR declarative sentence, showing how the pre- and post-tonic f0 patterns that we have seen thus far contribute to sentence-wide intonational patterns.

5.1 Sentences with only lexical H tones.

Figure 10 shows the typical intonational profile of a declarative sentence composed of lexical H tones aligned with stressed syllables. The first tier represents only lexical tones; the second tier represents all other f0 targets. As we discussed earlier, declarative sentences generally end with a H% boundary tone. Thus, in a sentence like Figure 10 with all H tones, the sentence-final syllable ends on a higher f0 than the preceding lexical H.

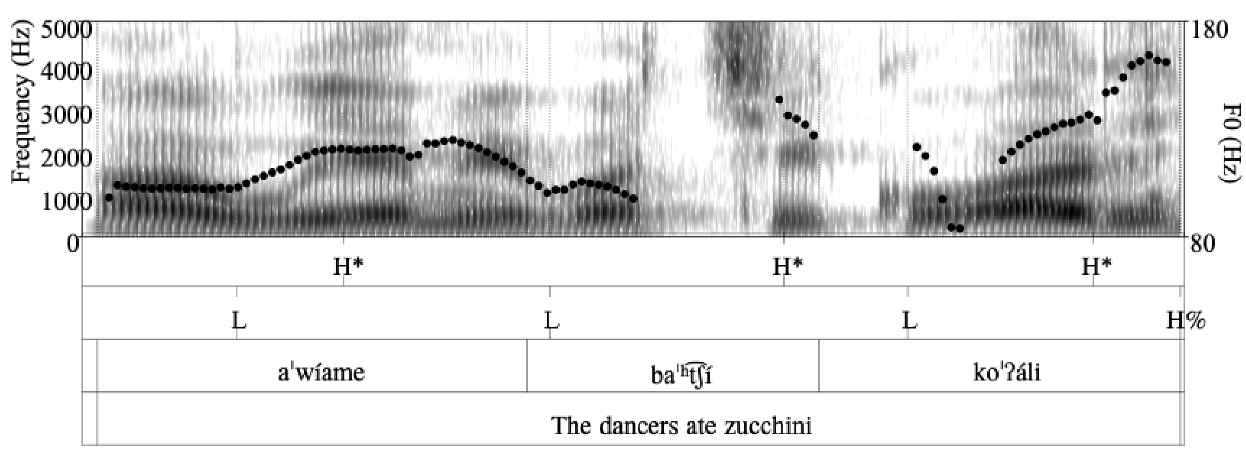


Figure 10: Typical intonation profile of a declarative sentence with H-toned words. Lexical tones are marked as H on the stressed syllables. Non-lexical pitch events, including the H% boundary tone and low "lead" tones, are marked on the second tier. See accompanying text above for more details.*

Further, we see that each lexical H is preceded by a low target, which we have called "lead" tones in earlier work (Garellek et al. 2015). It is generally on account of these lead tones that syllables with lexical H tones are preceded by a lower f0 target, as shown in Section 3 (e.g. in Figure 3). Lead tones are not always found; for example, a syllable with a lexical H tone that is immediately preceded by a word ending in a H tone will not generally have a low-pitched lead tone; see the f0 contours with no lead tones between the H-toned words [riˈʰté ˈpáli] in Figure 11. But their frequent occurrence results in a fairly strong macro-rhythm (Jun 2014), in that a sentence consisting of all H tones has regular occurrences of pitch rises and falls, instead of a relatively flat high pitch.

Sentences with all H tones, as in Figure 10, also reveal additional patterns of interest regarding CR intonation: as discussed in Section 3, a lexical H tone spreads unto the post-tonic; note in Figure 10 the high-pitched plateau during the tonic and post-tonic of [aˈwíame] 'dancers'. However, if the post-tonic occurs on another word, then the H tone doesn't spread consistently; note the presence of a low lead tone between [baˈʰt͡ʃí] and [koˈʔáli] 'zucchini eat-PST' in Figure 10.

Second, sequences of H tones show a progressive rise in the f0 of the lexical H targets, regardless of whether there is a preceding low lead tone. This is seen to some extent in Figure 10, which shows evidence of the lead tones, although the last two H tones have a similar f0. A clearer example of the successive raising in f0 for H tones can be seen in the two sentences shown in Figure 11, which illustrates two sentences that have a sequence of words bearing a H tone, denoted with a H*, and that end with a H% boundary tone between 220-235 Hz (shown on the right-hand side of the spectrograms). In both cases, the f0 rises into the first H tone of [wikaˈná] 'in many places,' and then continues to rise until the H%, even as there are local shallow f0 dips.

In these examples, we also see how the lower-pitched lead targets are often not low in terms of the preceding f0 or the mean f0 for the sentence; instead, they are locally lower-pitched with respect to the *following* lexical H tone. Thus, low-pitched lead pitch targets can be upstepped (represented in the figures as "^L"), meaning that their pitch level is higher than expected for the global f0 pattern. For example, in the sentence shown in the lower panel of Figure 11, there appear to be local lower-pitched targets before all lexical H tones except the utterance-final word, which has a H tone immediately preceded by another H tone in the preceding word. However, these lower-pitched targets are often quite high in f0, given that the utterance as a whole is showing a gradual rise in pitch. (In that figure, one of the upstepped lead low tones is marked with a question mark, because it is unclear if the f0 is lowering due to microprosody from the voiced [j] in [kaˈjámi] 'of all kinds') In Figure 11, the f0 contour is also consistent with terraced upstepping of high tones, where each H is higher in f0 than the preceding one. But note that Figure 10 provides clearer evidence for the presence of lead low tones.

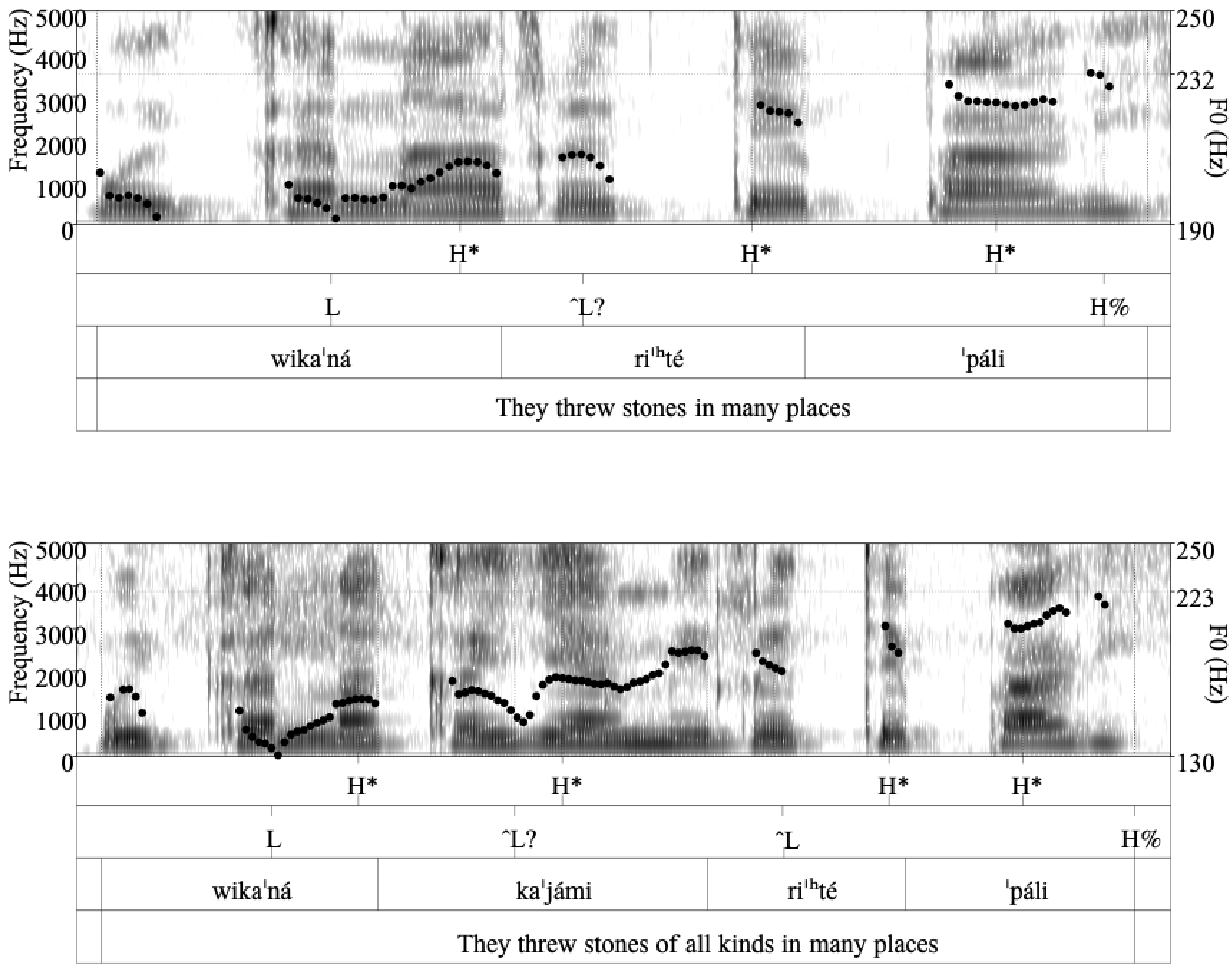


Figure 11: Other examples of declarative sentences with H-toned words. Lexical tones are marked as H on the stressed syllables. Non-lexical pitch events, including the H% boundary tone and low "lead" tones, are marked on the second tier. See accompanying text for more details.*

5.2 Sentences with only lexical L tones.

Figure 12 shows a representative declarative sentence composed of words with lexical L tones. Just as with sentences with only H tones, the intonation is quite rhythmic, only in these sentences there are alternating low-pitched targets for the lexical tones, and higher-pitched targets for the "lead" tones. As with sentences ending with a lexical H-toned word, sentences with a final lexical L tone end in a pitch rise, which we attribute to the H% boundary tone. If a L tone occurs on an utterance-final syllable, then the rise for the H% occurs on the same utterance-final syllable, as seen in Figure 12.

In this Figure, we also see how the lead tones can be realized on the same vowel which bears a lexical tone: the end of [raˈrài] ‘sandals’ shows a rising f0, as both the lexical L and lead H (triggered by the following lexical L tone) co-occur on the final diphthong. It is for these reasons that L tones are often preceded and followed by higher pitch targets (see Figure 12), and why lexical L tones can have a rise in f0 over the tonic (see Figure 3).

While Figure 12 appears to be consistent with an analysis in which the lexical L tone is *followed* by a trailing high target (i.e., it is L*+H), we argue that this is not the correct analysis. First, the high targets that (we argue) precede the lexical L are optional. Figure 13 shows a sentence with all L lexical tones (uttered by the same speaker, SFH, as in Figure 12), but in which no lexical L is preceded by a lead high pitch. Moreover, this sentence shows a clear L% boundary tone, instead of the more typical H% boundary tone. This also implies that when sentences ending in a lexical L tone exhibit a final rise, they do not do so because the lexical tone is actually L*+H. If the sentence ends in a L% boundary tone, then the final lexical L has only a low-pitched target.[6] Therefore, Figure 13 illustrates several important generalizations: sometimes a L% boundary tone is found in sentences ending in a lexical L tone, and the rhythmic fall-rise pattern in sentences with only L tones is best treated as lexical L tones that are *preceded*, rather than followed, by a lead high pitch-- just as lexical H (and HL) tones are preceded by low lead pitch targets.

Figure 12 also shows a gradual lowering in f0 of both high-pitched targets, and both Figures 12 and 13 show gradual lowering of the lexical L tones. Thus, in contrast to sentences with only H tones, sentences with L tones can show declination. Very often though, the f0 rise associated with H%, when it occurs, can go against declination; we saw evidence for this in Figure 3 by the very high mean F0 in utterance-final syllables following a lexical L.

[6] We do not know what conditions such variability in the boundary tone, but the occurrence of H% vs. L% boundary tones does not seem to be determined by differences in intonational meaning or by list intonation effects. Although elicited sentences were sometimes repeated sequentially, this was usually done with a long pause and/or breath in between repetitions, which would disfavor list intonation. Moreover, the H% in sentences with lexical L tones usually occurred in the final repetition of a sentence, suggesting that the H% is not more likely to occur in non-final sentences that are uttered sequentially.

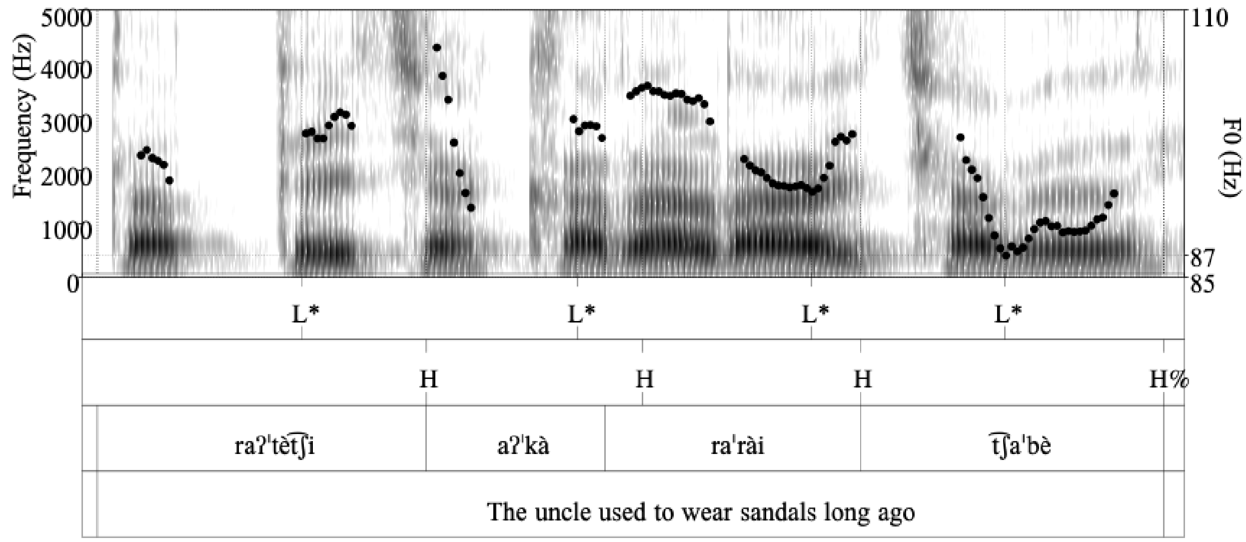


Figure 12: Typical intonation profile of a declarative sentence with L-toned words. Lexical tones are marked as L on the stressed syllables. Non-lexical pitch events, including the H% boundary tone and high "lead" tones, are marked on the second tier. See accompanying text for more details.*

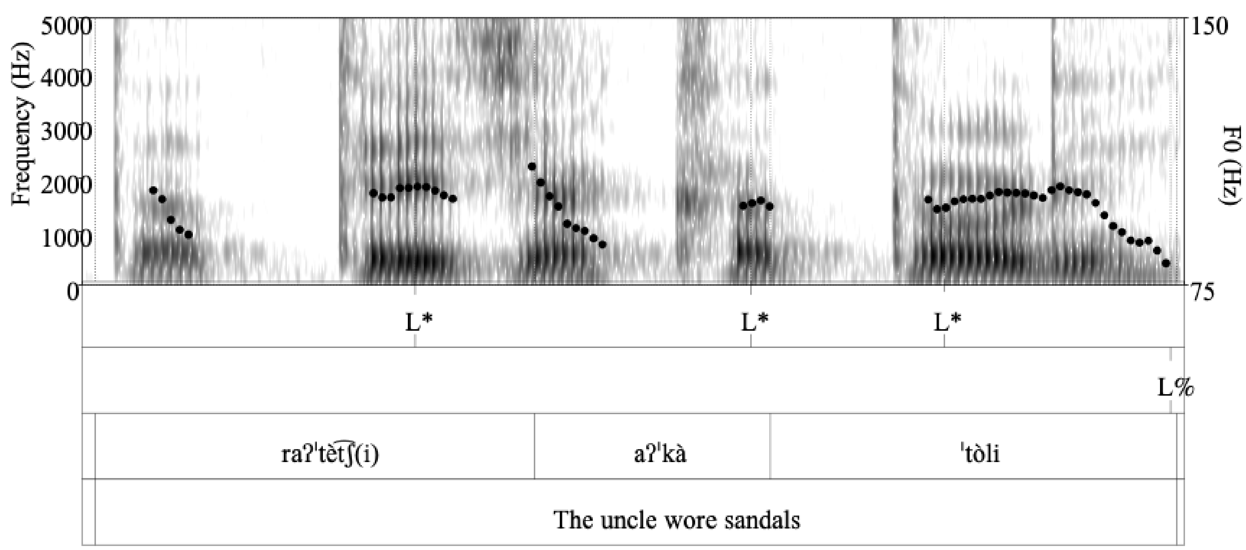


Figure 13: Less typical intonation profile of a declarative sentence with L-toned words. Lexical tones are marked as L on the stressed syllables. This example shows a L% boundary tone and no preceding high "lead" tones. See accompanying text for more details.*

5.3 Sentences with only lexical HL tones.
Figure 14 illustrates a three-word declarative sentence where every word bears a HL tone. (The lexical tones are denoted as H*L in the figure to emphasize that the low-pitch target is a trailing target that occurs on the post-tonic syllable.) There is no H% boundary tone, as expected from what we saw for utterance-final HL tones in Section 3 (e.g. Figure 3). We interpret this as lexical tone preservation-- essentially, the final HL tone overrides a boundary tone if the two were to occur in the last two syllables of a phrase. But we acknowledge that more data is needed to determine whether this is indeed the case. For example, if the final HL of the sentence appears earlier than in the last two syllables of the phrase, does a boundary tone appear? Unfortunately, we did not elicit any such sentences. It is also possible that such sentences take a L% boundary tone, since that is attested in sentences ending in L tones (see previous section). However, since we did not elicit any sentences with multiple syllables following the final HL tone, we cannot determine if a boundary tone is ever found in sentences ending in a HL tone-- and if it is, whether it is a H% or L%. Thus, no boundary tones are marked in Figures 14-15.

Interestingly, the high-pitched target for the final HL is higher in f0 than the previous two high targets, though the high targets for the previous two HL tones show declination. Declination is also found for the low-pitched targets, including the three low targets associated with the lexical HL, as well as the two low-pitched lead tones. Given the declination in low-pitched targets and the higher-pitched target on the final HL, sentences ending in HL tones appear to show pitch range expansion in utterance-final position. But we caution against claiming that pitch range expansion is the norm for utterance-final HL tones: note in Figure 3 that, overall, utterance-final HL tones show a more pronounced fall because their *low* target is lowered-- that is, when averaging across all HL tokens, we do not find evidence for an upstepped high target.

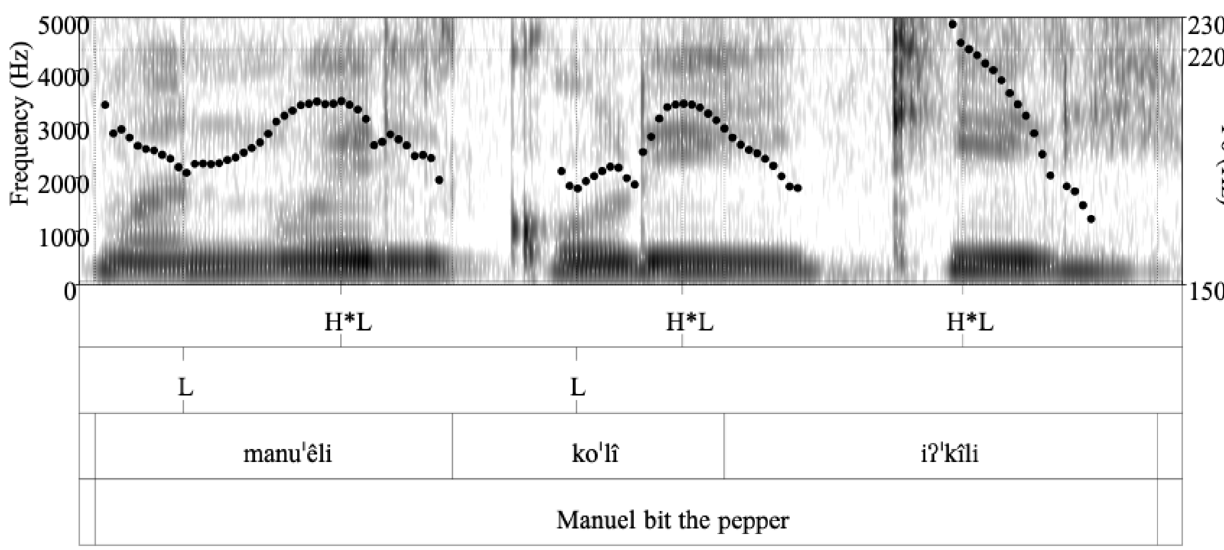
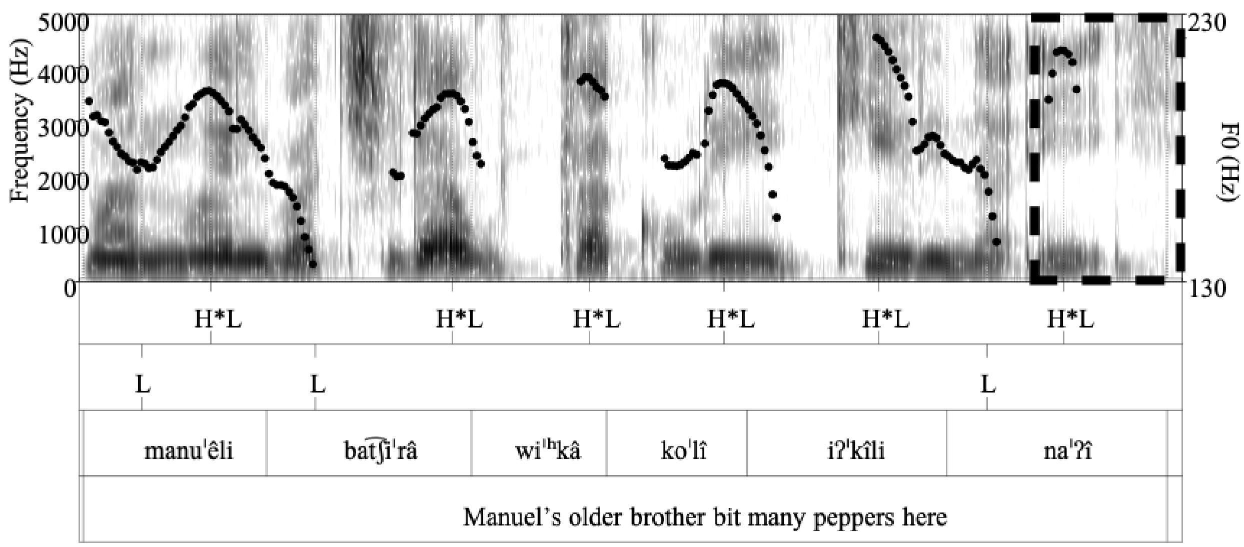


*Figure 14: Typical intonation profile of a declarative sentence with HL-toned words. Lexical tones are marked as H*L on the stressed syllables. Non-lexical pitch events, such as low "lead" tones, are marked on the second tier. See accompanying text for more details.*

In addition, utterance-final HL tones can be optionally rearticulated. That is, a phrase-final /V̂/ can be produced as [V́ʔV̀]. In Figure 15, we show two examples of the same sentence 'Manuel's older brother bit many peppers here.' Both examples were uttered by the same female speaker (BFL). In this sentence each word bears a lexical HL tone; as expected, there is no H% boundary tone. However, the sample sentence in the bottom panel of Figure 15 has no rearticulation of the utterance-final falling-toned vowel of [naˈʔî] 'here,' whereas the one in the top panel does. The final vowel portion is with a dashed border. (Note that the glottalization before the border is due to the phonemic glottal stop of [naˈʔî] 'here,' and not to the rearticulation of the HL tone.) We also note here that the high target of the falling HL tone on word [naˈʔî] 'here' has a lower f0 than the high target of the preceding HL. Determining the reason for this is beyond the scope of this chapter, but may be due to unintended narrow focus on the preceding word [iʔˈkîli]'bite-PST', or to the fact that, after pre-glottalized voiceless stops, vowels (particularly [i], as in [iʔˈkîli] 'bite-PST') tend to have a higher-than-expected f0.



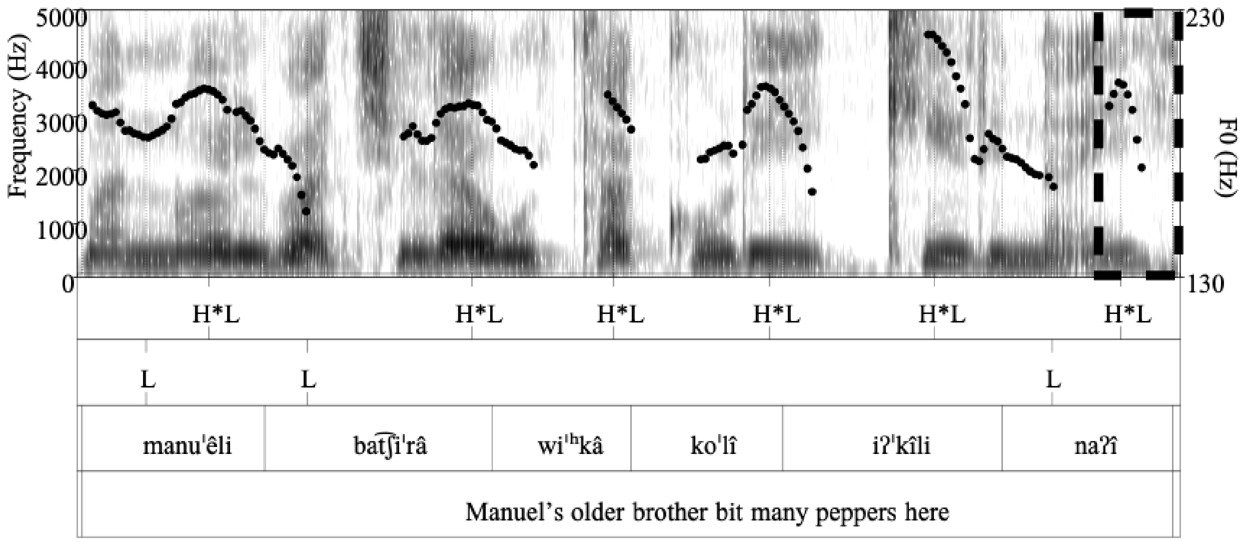


Figure 15: Two versions of a sentence with all HL tones, both uttered by the same speaker. The dashed box illustrates the extent of the final vowel. The top example shows rearticulation of the final vowel bearing a HL tone; the bottom example does not.

6. General discussion and conclusion

The results from the study presented above show that the three lexical tones in Choguita Rarámuri (CR) are realized with predictable f0 contours on tonic syllables: H tones with a high f0, L tones with a low f0, and HL tones with a falling f0. The f0 distinctions are often small (of about 3 semitones; cf. Caballero & Carroll 2015), and some speakers show little distinctions between certain tones (especially H vs. HL) over tonic syllables, though they can be further distinguished by the post-tonic f0 pattern (see Figures 2-3). Still, the tones have predictable f0 realizations in both utterance-medial and utterance-final positions. These stable realizations are to be expected, given the important roles (both lexical and grammatical) that tone has in the language (Caballero & Carroll 2015, Caballero & German to appear).

For three of the four speakers whose speech is analyzed here, a linear discriminant analysis revealed that voice quality also helps differentiate the three tone categories. In particular, the inclusion of measures of voice quality resulted in an improvement in classification accuracy by nearly 18% for one speaker (BFL), because HL tones tended to be breathier than the H and L counterparts, even when they occurred phrase-medially without rearticulation. Thus, CR tones are realized in phonetically multidimensional ways using phonetic parameters beyond f0, such as phonation (Kuang 2013). Still, we note here that the multidimensionality of tones in CR appears to be speaker-dependent. In other contrasts which involve both f0 and voice quality, studies have shown that cue weighting can vary by speaker/listener as well as dialect, such that f0 and voice quality may trade off in importance (Brunelle 2009, Brunelle 2012, Kuang & Cui

2018). For CR then, future research would benefit from perceptual analysis of the (likely) listener-dependent role of voice quality measures as cues to the tonal contrasts.

We note that the multidimensionality of CR tones could be related to the diachronic development of a hybrid prosodic system through the innovation of tone. In the Oapan and Ahueliacán varieties of Balsas Nahuatl (Uto-Aztecan; Mexico), high tone developed on vowels followed by a breathy-voiced coda, where a lower pitch associated with this breathy-voiced coda was reinterpreted as a high pitch in the preceding syllable, with a resulting high-low tonal contour (Guion et al. 2010). Given that stress has been retained, tonal innovation has involved the redeployment of acoustic cues to accommodate both stress and tone in the word prosody, and Guion et al. (2010) speculate that these varieties may eventually transition to being exclusively tonal. They propose hybrid prosodic systems in general may be unstable, raising the question of whether tone multidimensionality is linked to their recent prosodic reconfiguration.

Our study also describes aspects of declarative sentence intonation in CR, and the ways in which lexical tone and intonation interact. Declaratives are characterized by a high boundary tone, in contrast to the general cross-linguistic tendency for low boundary tones in sentences that are not information-seeking (Ohala 1983; see also review of declarative intonation in several languages in Jun 2005, 2014). However, we also find some evidence for a L% boundary tone in declarative sentences, with no concomitant change in intonational meaning. It is also noteworthy that some sentences show no declination, with some utterance-final H and HL tones having higher maximal f0 points than the same tones occurring earlier in the utterance. A lack of declination for sequences of high tones has been similarly observed in other languages, such as Mandarin, Taiwanese, and Yoloxóchitl Mixtec (Xu 1999, Peng 1997, DiCanio et al. 2020/Online First).

Utterance-final tones in CR also show patterns of tonal enhancement: the HL tone rearticulates, and no H% boundary tone is realized, allowing the falling contour to be realized in utterance-final position. Moreover, tones (particularly HL) are produced with an expanded f0 range, indicating greater tonal dispersion in utterance-final position. Although this pattern is not unattested across languages, it is more common for languages to show patterns of f0 lowering in utterance-final declaratives: for a recent overview, see DiCanio et al. (2020/Online First).

It is also clear from this study, as well as from previous work on the language (Caballero 2008, Caballero & Carroll 2015, Garellek et al. 2015), that tone extends its influence on f0 to neighboring syllables: H tones are often preceded by a low pitch target on the pretonic, and are followed by a high plateau from the tonic through the post-tonic; L tones are often preceded by a high pitch target (Figure 12; cf. Figure 13); and HL tones are often preceded by a low pitch target, and the falling contour extends to the post-tonic syllable.

Moreover, our analysis presented in Section 4 indicates that the incorporation of f0 over pre- and post-tonic syllables generally improves the statistical discriminability of the three tones in utterance-medial position. In sum then, tonal contrasts in CR are as much about paradigmatic differences across the tone categories during the tone-bearing syllable, as they are about syntagmatic changes over at least three syllables centered *around* the tone-bearing syllable.

Whereas lexical and intonational tones are often described in exclusively paradigmatic terms, reflecting the differences across tonal categories, our study makes clear that their syntagmatic properties are important, at least for understanding how tones are produced; see Dilley & Breen (in press) for a recent review of the topic. Future research should investigate just how important syntagmatic changes in f0 are for CR tone recoverability in context.

References

Aguilar, A.E. (2020). The phonology and phonetics of laryngeal sounds in Chicontepec Nahuatl. PhD thesis, UC San Diego.

Aguilar, A., Caballero, G., Carroll, L., & Garellek, M. (2015). Multi-dimensionality in the tonal realization of Choguita Rarámuri (Tarahumara). Talk presented at the 2015 Meeting of the Society for the Study of the Indigenous Languages of the Americas.

Brunelle, M., (2009). Tone perception in Northern and Southern Vietnamese. *Journal of Phonetics, 37*(1), 79-96.

Brunelle, M. (2012). Dialect experience and perceptual integrality in phonological registers: Fundamental frequency, voice quality and the first formant in Cham. *Journal of the Acoustical Society of America, 131*(4), 3088-3102.

Burgess, D. (1984). Western Tarahumara. (Studies in Uto-Aztecan Grammar, 4.) Dallas: The Summer Institute of Linguistics and The University of Texas at Arlington.

Caballero, G. (2008). Choguita Rarámuri (Tarahumara) phonology and morphology. PhD thesis, University of California, Berkeley.

Caballero, G. (2011). Morphologically conditioned stress assignment in Choguita Rarámuri (Tarahumara). *Journal of Linguistics, 49*, 749–790.

Caballero, G., Carroll, L. (2015). Tone and stress in Choguita Rarámuri (Tarahumara) word prosody. *International Journal of American Linguistics, 81*, 457-493.

Caballero, G. (2017). Choguita Rarámuri (Tarahumara) language description and documentation: a guide to the deposited collection and associated materials. *Language Documentation and Conservation, 11*, 224-255.

Caballero, G. & M. Gordon. (To appear). Prosody in North American Indian Languages. In Carlos Gussenhoven & Ajou Chen (eds.), *The Oxford Handbook of Language Prosody*. Oxford: Oxford University Press.

Casaus, M. (2008). Quantitative ethnobotany and acculturation among the Rarámuri of Choguita, Chihuahua, México. Ms.

Demers, R., Escalante, F., & Jelinek, E. (1999). Prominence in Yaqui words. *International Journal of American Linguistics* 65:40–50.

DiCanio, C., Benn, J., & Castillo García, R. (Online First, 2020). Disentangling the Effects of Position and Utterance-Level Declination on the Production of Complex Tones in Yoloxóchitl Mixtec. *Language and Speech*, 0023830920939132.

Dilley, L. & Breen, M. (in press). An enhanced autosegmental-metrical theory (AM+) facilitates phonetically transparent prosodic annotation: A reply to Jun. In J. Barnes and S. Shattuck-Hufnagel (eds.) *Prosodic Theory and Practice*. Cambridge, MA: MIT Press.

Embriz Osorio, A. & Zamora Alarcón, Ó. (2012). México, lenguas indígenas nacionales en riesgo de desaparición: variantes lingüísticas por grado de riesgo, 2000 . Mexico City: Instituto Nacional de Lenguas Indígenas.

Faytak, M., & Yum A. C. L. (2011). A Typological Study of the Interaction between Level Tones and Duration. In *Proceedings of ICPhS 2011* (pp. 659-662).

Gandour, J. (1977). On the interaction between tone and vowel length: Evidence from Thai dialects. *Phonetica, 34*(1), 54-65.

Garellek, M. (2019). The phonetics of voice. In W. F. Katz & P. F. Assmann (editors), *The Routledge Handbook of Phonetics*, pp. 75–106. Routledge.

Garellek, M., A. Aguilar, L. Carroll, & G. Caballero. (2015). Lexical and post-lexical tone in Choguita Rarámuri. Proceedings of the 18th International Congress of Phonetic Sciences.

Gibson, M. K. (2013). Lexical tone, intonation, and their interaction: A scopal theory of tune association. PhD thesis, Cornell University.

Gil Burgoin, Carlos I., & Carrillo Carrillo, A. (2019). Marcación prosódica de los enunciados interrogativos absolutos en tepehuano del norte [Prosodic marking of polar interogatives in Northern Tepehuan]. In Z. Estrada Fernández, M. Tubino Blanco, & A. Alvarez Gonzalez (editors), *Estudios de lenguas amerindias 4: escenario actual de la investigación sobre lenguas yuto-aztecas. Homenaje a Jane H. Hill*, pp. 105–136. Universidad de Sonora.

Gordon, Matthew. (2001). A typology of contour tone restrictions. *Studies in Language, 25*(3), 423–462.

Grimes, Joseph. (1959). Huichol tone and intonation. *International Journal of American Linguistics* 25:221–232.

Guion, S. G., Amith, J. D., Doty, C. S., & Shport, I. A. (2010). Word-level prosody in Balsas Nahuatl: The origin, development, and acoustic correlates of tone in a stress accent language. *Journal of Phonetics, 38*(2), 137-166.

Gussenhoven, C. (2004). *The phonology of tone and intonation.* Cambridge: Cambridge University Press.

Hagberg, L. (1989). Floating accent in Mayo. Arizona Phonology Conference, vol. 2, ed. S. Lee Fulmer, Masahide Ishihara, and Wendy Wiswall, pp. 32–47. Tucson: University of Arizona.

Hyman, L. M. (2006). Word-prosodic typology. Phonology 23:225–257.

Hyman, L. M. (2011). The representation of tone. In Marcarc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice (eds), *The Blackwell Companion to Phonology* (pp. 1-25). John Wiley & Sons.

INEGI (Instituto Nacional de Estadistica y Geografia). (2010). Censo de población y vivienda 2010: Tabulados del cuestionario básico .

Jun, S.-A. (2005). Prosodic Typology. In Sun-Ah Jun (ed.) *Prosodic Typology: The Phonology of Intonation and Phrasing*. pp. 430-458. Oxford: Oxford University Press.

Jun, S.-A. (2014) Prosodic Typology: By Prominence Type, Word Prosody, and Macro-rhythm. In Sun-Ah Jun (ed.) *Prosodic Typology II: The Phonology of Intonation and Phrasing*. Oxford: Oxford University Press.

Kuang, J. (2013). Phonation in tonal contrasts. PhD thesis, UCLA.

Kuang, J., & Cui, A. (2018). Relative cue weighting in production and perception of an ongoing sound change in Southern Yi. *Journal of Phonetics, 71*, 194-214.

Manaster-Ramer, A. (1986). Genesis of Hopi tones. *International Journal of American Linguistics* 52:154–60.

May, J. & E. Loeweke. 1964. The phonological hierarchy in Fasu. Anthropological Linguistics 7:89–97.

McMahon, A. (1967). Phonemes and phonemic units in Cora (Mexico). *International Journal of American Linguistics* 33:128–34.

Merrill, William & Don Burgess. (2014). Ralámuli kinship terminology: a diachronic perspective on diversity in the Sierra Tarahumara of Northwestern Mexico. *Anthropological Linguistics* 56(3-4): 229-293.

Morales Moreno, J.I. (2016). Estructura de cláusula, tópico y foco en Rarámuri de Rochéachi. Mexico City: Centro de Investigaciones y Estudios Superiores en Antropología Social. MA thesis.

Munro, P. (1977). Towards a reconstruction of Uto-Aztecan stress. Studies in Stress and Accent, ed. Larry Hyman, Southern California Occasional Papers in Linguistics 4, pp. 303–24. Los Angeles: Department of Linguistics, University of Southern California.

Ohala, J. J. (1983). Cross-language use of pitch: an ethological view. *Phonetica, 40*(1), 1-18.

Patiño, E. P. V. 2014. Intonation patterns of Morelos Nahuatl. Proceedings of Speech Prosody 7.

Peng, S. H. (1997). Production and perception of Taiwanese tones in different tonal and prosodic contexts. *Journal of Phonetics, 25*(3), 371-400.

Remijsen, B. (2001). Word-prosodic systems of Raja Ampat languages. Ph.D. dissertation, University of Leiden.

Remijsen, B. (2002). Lexically contrastive stress accent and lexical tone in Maˈya. Laboratory phonology 7, ed. Carlos Gussenhoven and Natasha Warner, pp. 585–614. Berlin and New York: Mouton de Gruyter.

Riad, T. (2006). Scandinavian accent typology. Sprachtypologie und Universalienforschung (STUF) 59:36–51.

Turnbull, R. (2017). The phonetics and phonology of lexical prosody in San Jerónimo Acazulco Otomi. *Journal of the International Phonetic Association*, *47*(3), 251-282.

Shaul, D. L. (2000). Comparative Tepiman: Phonological change and inflectional categories. In E. H. Casad, & T. L. Willett (Eds.), Uto-Aztecan: Structural, temporal and geographic perspectives: Papers in honor of Wick R. Miller by the friends of Uto-Aztecan (pp. 319–356). Hermosillo: Universidad de Sonora.

Valdez Jara, Y. (2013). Predication in Rarómuri (Urique Tarahumara). PhD thesis, University of Oregon.

Woo, N. (1970). Tone in Northern Tepehuan. *International Journal of American Linguistics* 36:18–30.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics, 27*(1), 55-105.

Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.

Zec, D. (1999). "Footed Tones and Tonal Feet: Metrical Constituency in a Pitch Accent Language." *Phonology* 16: 225-264.

Zhang, J. (2002). The effects of duration and sonority on contour tone distribution–typological survey and formal analysis. PhD thesis, UCLA.