Tonal variability and marginal contrast: Lexical pitch accent in Uspanteko[*]

Ryan Bennett[1], Robert Henderson[2], and Meg Harvey[2]

University of California, Santa Cruz[1], University of Arizona[2]

Abstract

Uspanteko (Mayan) is unique among the languages of Guatemala in having a full-fledged system of lexical tone. Tone is thoroughly integrated into the morphology of the language, and triggers several predictable phonological alternations. The Uspanteko tone system is also relatively simple, involving at most two contrastive categories. However, tone in Uspanteko shows a high degree of phonetic variability, and there is no consensus as to what tones are phonetically or phonologically active in the language. This paper argues that at least some of this unclarity owes to the effect of intonation, which can obscure lexical tone on vowels. A phonetic study attempting to control for intonation suggests that lexical tone in Uspanteko involves a privative [H]∼∅ contrast on short vowels, and either a [H]∼∅ or [H]∼[L] contrast on long vowels. The relationship between the phonetics of tone and its low functional load in Uspanteko is also discussed.

Key terms: Uspanteko, Mayan, pitch accent, intonation, functional load, marginal contrast, phonetics, fieldwork

1 Uspanteko

Uspanteko belongs to the K'ichean branch of the Mayan language family (Bennett et al. 2016, Aissen et al. 2017 and references there). It is spoken in the central highlands of Guatemala, in the city of San Miguel Uspantán and in several surrounding villages, particularly Las Pacayas (Fig. 1; Us Maldonado no date(b)).

Figure 1: To the left: map of Guatemala, with the municipality of Uspantán shaded. Dashed lines indicate inset area to the right. To the right: Inset showing communities with >200 native speakers of Uspanteko, according to Us Maldonado (no date(b)), with major roads shown dashed.

There are currently somewhere between 1500-4000 native speakers of Uspanteko (Richards 2003, Us Maldonado no date(b)). Multilingualism is quite prevalent in Uspantán, and speakers of Uspanteko typically speak Spanish and K'iche' as well, and less frequently Q'eqchi' or Poqomchi'. To our knowledge there are no monolingual speakers of Uspanteko. The language is endangered: many children in the Uspanteko area are now learning Spanish and/or K'iche' as their first languages, as these are the languages most typically used in the public sphere in the Uspantán area. (K'iche', which has over 1 million speakers, is something of a lingua franca in the central and western highlands of Guatemala.)

1.1 Word-level prosody in Uspanteko

Uspanteko is one of the few Mayan languages to have innovated a robust, grammaticized system of lexical tone (Bennett 2016, England & Baird 2017, DiCanio & Bennett 2021).[1] The tone system is intertwined with word-level stress, vowel length, and morphology, as described below.

All descriptions of Uspanteko phonology agree that there are two tonal categories in the language. However, sources differ in how these two tones are characterized. Campbell (1977) reports

---

[1] Along with Uspanteko, the other Mayan languages with lexical tone are Yucatec and Mocho'. These three languages are geographically and genetically distant from each other, suggesting independent innovation of tone rather than shared inheritance. The tonal system of Yucatec is rather different from the tonal systems of Uspanteko and Mocho', consistent with independent innovation. Several Mamean languages spoken along the border between Guatemala and Mexico appear to be undergoing tonogenesis at present, conditioned by pitch pertubations associated with glottal stop [?]. More controversially, tone has been reported for Venustiano Carranza Tsotsil and Southern Lacandon, though the facts here are rather unclear. See Herrera Zendejas (2014), Bennett (2016), England & Baird (2017), Campbell (2017), Sobrino Gómez (2018), DiCanio & Bennett (2021) for details and references.

a contrast between high and low tones, while Grimes (1971, 1972) reports a contrast between falling and rising tones. Can Pixabaj (2007), Bennett & Henderson (2013), Bennett et al. (2019), and Bennett et al. (ms.a) stake out a third view, arguing that tone in Uspanteko involves a privative contrast between the presence vs. absence of a high tone [H] on stressed syllables. This analysis essentially treats tone in Uspanteko as lexical pitch accent system: no more than one lexical tone may occur per word, and toneless words are allowed (e.g. Hyman 2006, 2009). In this section we adopt the privative [H]∼∅ analysis of Bennett & Henderson (2013), and return to a more detailed evaluation of the tone system later on.

The basic stress system of Uspanteko is straightforward. In toneless words, stress falls on the word-final syllable, essentially without exception. There are no secondary stresses. This pattern of stress assignment is exemplified by the toneless words in Fig. 2, and is the same pattern of stress assignment found in other K'ichean-branch Mayan languages, none of which have tone (DiCanio & Bennett 2021, Bennett 2016, England & Baird 2017).[2] As discussed below, stress is phonetically cued by increased duration, greater intensity, vowel quality, and possibly by the anchoring of intonational tones, though this remains to be fully established (see too Baird 2014a,b, 2018 on the closely-related language K'iche', and Adell 2019 on Ixil). Vowel length is phonologically contrastive, but long vowels are restricted to stressed, word-final syllables (Fig. 2). This phonotactic restriction on long vowels gives rise to alternations like *chaak* [ˈt͡ʃaːk] 'work (NOUN)' vs. *tichakuun* [ti.t͡ʃa̬.ˈkuːn] '(s)he works'.
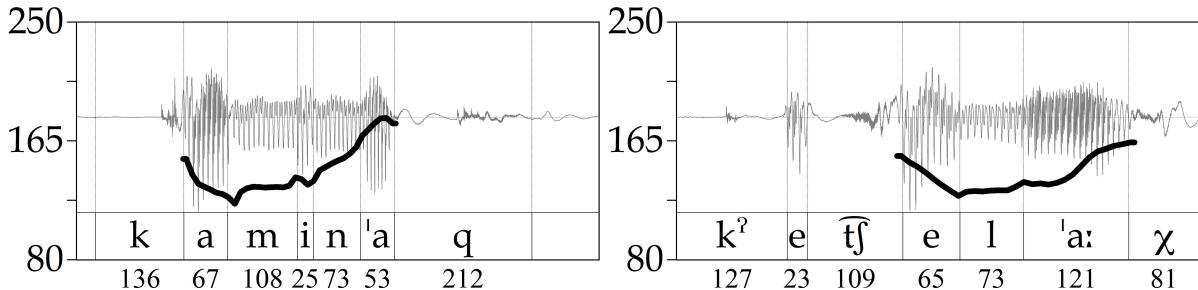
Figure 2: *kaminaq* [ka.mi.ˈnaq] 'deceased (person)' vs. *k'echelaaj* [kʔe.t͡ʃe.ˈlaːχ] 'forest' (speaker 6 JMS 2018). Y-axis marks f0 in Hz, X-axis marks duration in ms.

Recall that Uspanteko has two tonal categories, analyzed as [H] tone vs. the absence of tone in Bennett & Henderson (2013), Bennett et al. (2019). The [H] tone of Uspanteko has a restricted distribution, occurring only in stressed syllables. When the stressed vowel is long, the result is a simple [H]∼∅ contrast in the word-final stressed syllable. Such a contrast is illustrated in Fig. 3, in which tonal [ˈʃáːɓ] 'vomit' (on the right) is realized with a higher f0 than non-tonal [ˈʃaːɓ] 'comb' (on the left).

The distribution of [H] tone is a bit more complex in words containing only short vowels. In words of this type, tone is realized on the penultimate syllable, and stress retracts to the position of tone (Figs. 4-10). Bennett & Henderson (2013) propose that tone in Uspanteko is always realized on the penultimate vocalic mora: this would be the first mora of a final long vowel [. . . V́$_{\mu\mu}$C$_0$#],
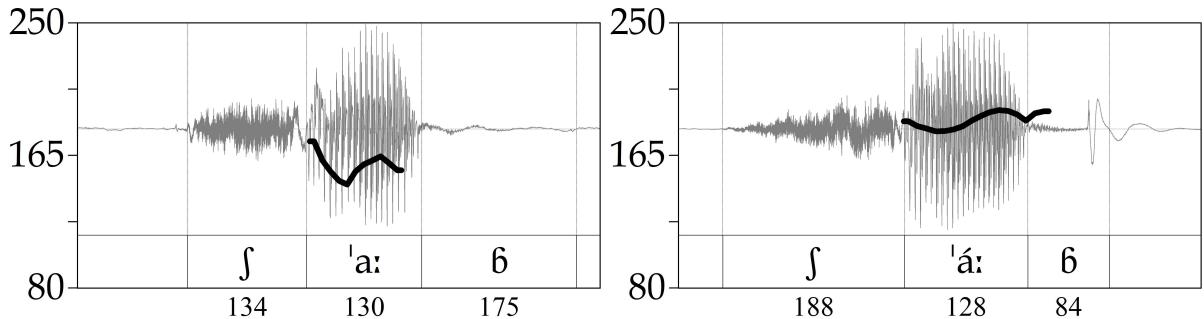
Figure 3: *xaab'* [ˈʃaːɓ] 'comb' (left) vs. *xáab'* [ˈʃáːɓ] 'vomit' (right) (speaker TAML 2019)

or the penultimate vowel when the final vowel is short [... $\acute{V}_\mu C_0 V_\mu C_0 \#$] (see also Campbell 1977). In the latter case, stress retraction to the penult ensures that tone and stress always coincide. (We know that stress retraction is specifically driven by H tone because this is the only context in which penultimate stress is ever observed.)
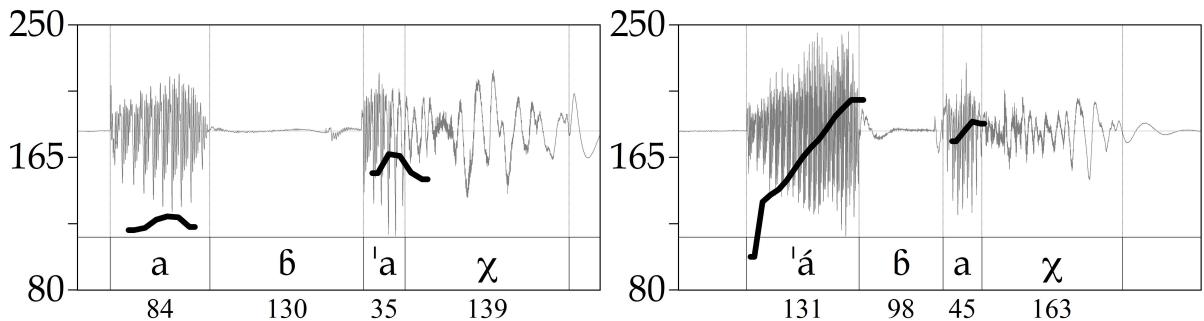
Figure 4: *ab'aj* [a.ˈɓaχ] 'stone' vs. *áb'aj* [ˈá.ɓaχ] 'your name' (speaker 6 JMS 2018)

Tone and stress are phonetically separable in Uspanteko. The primary correlate of [H] tone is raised pitch on the tonal vowel.[3] The primary correlates of stress are duration and intensity, and secondarily vowel quality, which is slightly more peripheral for stressed vowels (Bennett et al. ms.a and 3.6 below).[4] The phonetic differences between stress and tone are plainly visible in Fig. 5. In the non-tonal word [tu.ˈlul] (on the left), stress is cued by vowel duration, and somewhat more weakly by intensity. These same diagnostics indicate that stress has shifted to the penult in the tonal word [ˈín.t͡sʔiʔ] (on the right). Furthermore, tonal [ˈín.t͡sʔiʔ] has a prominent pitch excursion on the stressed penult, which is lacking from the stressed final syllable of the non-tonal word [tu.ˈlul]. Inspection of other comparable diagrams in this article, such as Figs. 4, 7, and 9, make it clear that these phonetic differences between tone and stress are highly systematic in Uspanteko (see also Bennett et al. ms.a). (In section 1.3 we discuss the slight rising pitch seen on some non-tonal stressed vowels in examples like Figs. 2 and 5, which we attribute to intonational boundary tones.)

---

[3]The term 'pitch' refers to a perceptual dimension. We use it here as a slightly inappropriate synonym for the acoustic dimension of f0.

[4]Consonant duration may also play a role in cuing stress in Uspanteko (e.g. Figs. 4, 7, 9), but we have not yet investigated this possibility in any detail.
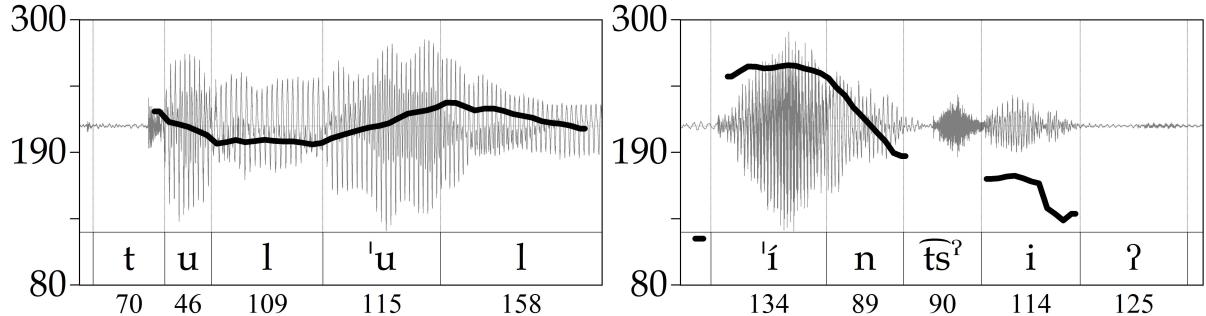
Figure 5: *tulul* [tu.ˈlul] 'zapote (fruit species, *pouteria viridis*)' vs. *íntz'i'* [ˈín.t͡sˀiʔ] 'my dog' (speaker JBAT 2011; Bennett & Henderson 2013)

1.1.1 Tonal morpho-phonology

Tone is closely tied to morphology in Uspanteko. Lexical [H] tones primarily occur on nouns, and to a lesser extent adjectives and adverbs (1).

(1)   a.   *xíkin* [ˈʃí.kin] 'ear'   e.   *xílij* [ˈʃí.liχ] 'inside'
       b.   *íxim* [ˈí.ʃim] 'corn'   f.   *lékej* [ˈlé.keχ] 'up, high'
       c.   *táq'aj* [ˈtá.qˀaχ] 'plains'   g.   *ójor* [ˈó.χor] 'in the past'
       d.   *cháqej* [ˈt͡ʃá.qeχ] 'dry'   h.   *íwir* [ˈí.wir] 'yesterday'

Tone is far less common on verbs and other lexical categories. There seem to be no tonal verb roots, and verbal morphology associated with tone is very limited.[5] The prevalence of tone on nouns owes in part to the fact that some highly productive nominal affixes introduce [H] tone to their stems (Table 1; see also Can Pixabaj 2007:39-49,89-95,110-2, etc.). Figs. 7, 9 illustrate tonal alternations on nouns triggered by the addition of tone-bearing affixes.

While affixation is a frequent and productive source of tone on nouns in Uspanteko, there are numerous morphologically *simple* roots which bear tone, as well as many simple roots (nominal and otherwise) which are lexically toneless (Figs. 3, 5, 6).
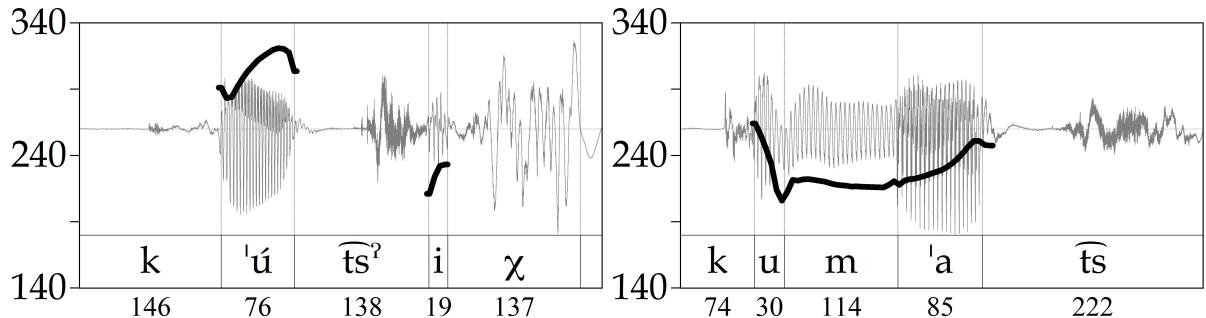
Figure 6: *kútz'ij* [ˈkú.t͡sˀiχ] 'flower' vs. *kumatz* [ku.ˈmat͡s] 'snake' (speaker 9 PA 2018)

---

[5]Can Pixabaj (2007:41,49,88, etc.) reports that the 'status suffix' /-ik/ may introduce tone on intransitive verbs (on status suffixes in general, see Henderson 2012, Coon 2016 and references there). Many of the Uspanteko speakers we have worked with do not make use of this suffix, and some speakers identify /-ik/ as a borrowing from closely-related K'iche', rather than being properly Uspanteko.

Table 1: Some productive affixes associated with tone. The form of the possessive prefixes varies depending on whether the stem begins with a vowel or consonant (Bennett 2016). Capital /V/ indicates a vowel of varying quality

Affix	Gloss	Type	Examples
/in-, [H]/∼/w-, [H]/	1SG.POSS	Inflectional	[ˈín.χal] 'my corncob'
			[in.ˈt͡ʃí.kit͡ʃ] 'my basket'
			[ˈwá.liɓ] 'my daughter-in-law'
/q(a)-, [H]/	1PL.POSS	Inflectional	[ˈqá.laq] 'our plate/dish'
			[qa.ˈpí.ʃiɓ] 'our advice'
/a(w)-, [H]/	2.POSS	Inflectional	[ˈá.piʃ] 'your tomato'
			[a.ˈwá.ɓiχ] 'your milpa'
/-Vɓ, [H]/	Plural	Inflectional (certain animate nouns)	[aχ.ˈt͡ʃá.kiɓ] 'workers'
/-ɓVl, [H]/	Instrumental/ locative	Derivational (noun-forming)	[ˈkʔáj.ɓal] 'marketplace'
/-Vl-ik, [H]/	Adjectival participle	Derivational (adjective-forming)	[sa.ˈt͡ʃá.lik] 'lost'

Tonal alternations triggered by affixes show a certain amount of lexical conditioning in Uspanteko, particularly in cases of nominal possession (Bennett & Henderson 2013). Compare e.g. alternating *pix* [ˈpiʃ] 'tomato' ∼ *ínpix* [ˈím.piʃ] 'my tomato' with non-alternating *pach* [ˈpat͡ʃ] 'friend' ∼ *inpach* [im.ˈpat͡ʃ] 'my friend'.

Though the presence or absence of tone is morphologically conditioned, the *position* of tone is determined entirely on the basis of phonological principles. As noted above, the [H] tone is always realized on the penultimate mora of the phonological word, regardless of its morphological origin. As a result, [H] tones introduced by affixes are not necessarily realized on those affixes themselves. For example, when tonal possessive prefixes (Table 1) attach to monosyllabic stems like [ˈɓa] 'head', tone will be realized on those possessive prefixes (Fig. 7; note that the 3.POSS prefix /(i)χ-/ in the lefthand diagram is *not* associated with tone). But when a possessive prefix introduces tone on a polysyllabic stem, it will be the stem rather than the prefix which hosts the [H] tone (Fig. 8).

The same is true for suffixes which introduce tone: if the vowel of the suffix is short, it will assign [H] tone and stress to the preceding vowel, as with the plural suffix /-Vɓ, [H]/ (Fig. 9; note the shortening of the stem vowel /iː/ in non-final position).

Similarly, some possessive prefixes associated with tone consist of just a single consonant. These prefixes cannot themselves host the tone they are associated with, even in principle, and so tone necessarily migrates to the nominal stem (Fig. 10).

Since the position of tone is phonologically determined, tone may drift one or more syllables away from the morpheme which introduces it, in order to be realized on the penultimate mora, e.g.
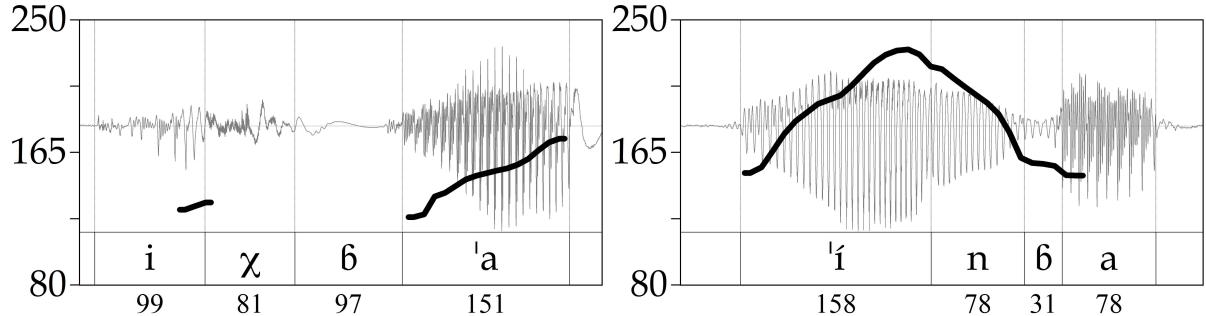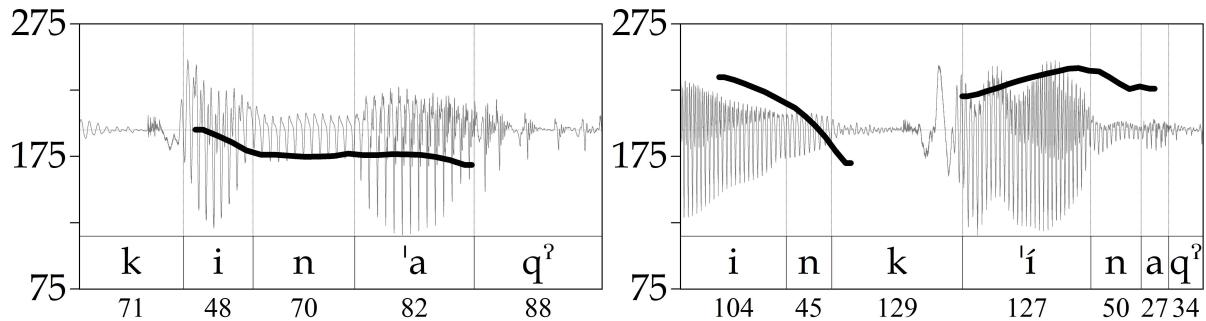
Figure 7: *ijb'a* [iχ.ˈɓa] 'his/her/its head' vs. *ínb'a* [ˈín.ɓa] 'my head' (speaker 6 JMS 2018)

Figure 8: *kinaq'* [ki.ˈnaqˀ] 'bean' vs. *inkínaq'* [in.ˈkí.naqˀ] 'my bean' (speaker 9 PA 2018)

wajtíij /w, [H] - aχ - tiːj/ → [w-aχ-tíːχ] 'my teacher'.[6]

In cases where two morphemes associated with [H] tone occur in the same underlying form, only one [H] tone surfaces phonetically, always in the expected location on the penultimate mora of the word (2):

(2) a. *inkútz'ij* /in, [H] - kut͡sˀiχ, [H]/ → [in.ˈkú.t͡sˀiχ] 'my flower'
  b. *wajtíjib'* /w, [H] - aχ - tiːχ - Vɓ, [H]/ → [waχ.ˈtí.χiɓ] 'my teachers'
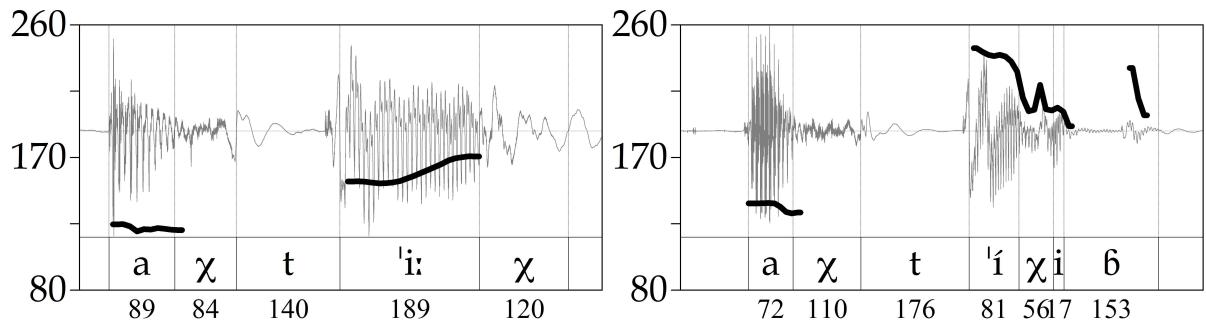
Figure 9: *ajtiij* [aχ.ˈtiːχ] 'teacher' vs. *ajtíjib'* [aχ.ˈtí.χiɓ] 'teachers' (speaker TAML 2019)

---

[6]A reviewer points out that the fact that [H] can shift across unstressed vowels in words like *wajtíij* /w, [H] - aχ - tiːχ/ → [w-aχ-tíːχ] confirms that unstressed vowels are toneless in Uspanteko. Examples in which [H] tone crosses two or more unstressed syllables to arrive at the penultimate mora are difficult to construct, and our consultants find some of the constructed forms that we've tested to be rather artificial. This remains an area for further investigation.
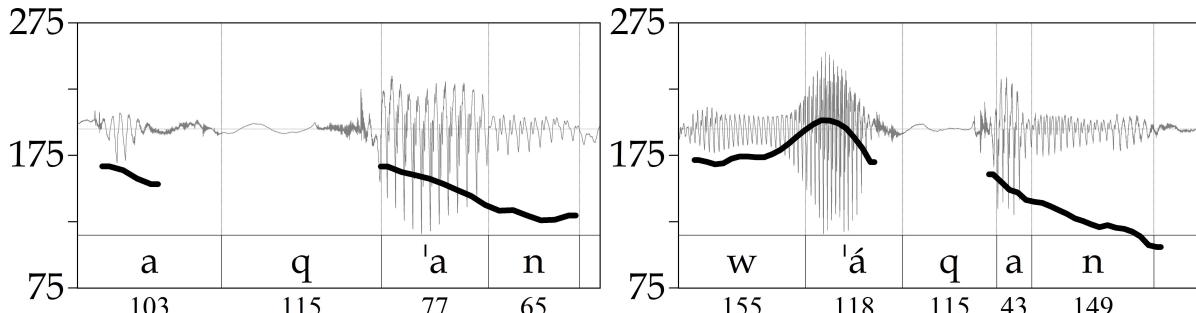
Figure 10: *aqan* [a.ˈqan] 'leg' vs. *wáqan* [ˈwá.qan] 'my leg' (speaker JBAT 2014)

Finally, tone which is specified on roots may also shift under suffixation to satisfy the phonological conditions on tone placement, e.g. *kútz'ij* [ˈkú.t͡sʔiχ] 'flower' (Fig. 6) vs. *jkutz'íjil* [χku.ˈt͡sʔí.χil] 'its flower (as in a flower painted on something)' (Can Pixabaj 2007:76,91-94).

We briefly note here that the deletion of unstressed final vowels can render the placement of tone on penultimate short vowels opaque, e.g. /ʃaχaɓ, [H]/ → [ˈʃáχɓ] 'shoe'. For detailed discussion see Can Pixabaj (2007:27-8, 51-2, 67-8), Bennett & Henderson (2013), Bennett (2016), and Bennett et al. (ms.b).

1.2 The functional load of tonal contrasts in Uspanteko

Tone is deeply ingrained in the morpho-phonology of Uspanteko. Tonal alternations related to affixation are systematic, commonplace, and easy to observe. Tone is associated with particular morphemes and lexical categories, and shows some idiosyncratic (but regular) lexical conditioning. The distribution of [H] tones is also tightly constrained by the phonological grammar. The core generalizations about the patterning of [H] tone—namely, that [H] tone can only occur on penultimate moras, and must always coincide with stress—seem to be exceptionless.[7] These are the characteristics of an extensively grammaticized system of lexical tone.

At the same time, lexical tone plays almost no role in distinguishing words from each other in Uspanteko. Minimal pairs distinguished by tone are almost entirely absent from the language, and so the functional load of tone is essentially nil. We are aware of only three minimal pairs for tone on long vowels (3):

(3) a. [ˈʃaːɓ] 'comb' vs. [ˈʃáːɓ] 'vomit' (Figs. 3, 12, 13)
     b. [ˈsiːp] 'tick' vs. [ˈsíːp] 'gift'
   c. [ˈoːχ] 'us (= 1P.ABS)' vs. [ˈóːχ] 'avocado'

But even these minimal pairs are not in the vocabulary of all speakers. There is a good amount of lexical variation in Uspanteko, no doubt connected to widespread multilingualism in Mayan

---

[7]As noted in the previous section, post-tonic vowel deletion seems to strand [H] tone on the final mora in words like *ínpix* [ˈím.piʃ]~[ˈímpʃ] 'my tomato'. We do not consider such forms to be counterexamples to the claim that [H] tone is strictly limited to penultimate moras, because (i) forms with deletion are typically in free variation with forms that do not delete, and (ii) assuming a phonological derivation in which tone placement occurs before vowel deletion, [H] tone is still limited to the penultimate mora at the point [H] tone is first assigned. Furthermore, Bennett et al. (ms.b) argue that vowel deletion does not in fact eliminate an underlying vowel in Uspanteko, but instead involves phonetic processes of segmental overlap in speech production. If correct, this means that [H] is still phonologically on the penultimate mora even in forms like [ˈímpʃ]. See also Bennett & Henderson (2013).

languages in the area. The notion of 'gift', for example, is lexicalized as [ˈsíːp] (3b), but also as [si.ˈpá.nik], [si.pa.ˈnik], [si.ˈpaːn], [ko.ˈt͡ʃaɓ], and [ˈkoːt͡ʃ(χ)], with a good amount of inter-speaker variation (somewhat akin to English 'gift' vs. 'present'). As a result, there are vanishingly few minimal pairs in Uspanteko in which tone is truly distinctive on long vowels. Compare this with Japanese, where fully 14% of the minimal pairs in the language are tonal minimal pairs (e.g. references in Kawahara 2015). The absence of minimal pairs for tone may be related to the fact that the distribution of tone in Uspanteko is skewed by lexical category: tone is very common on nouns (including both inflected and bare forms), less common on adjectives and adverbs, and rare on verbs (for some speakers, there are *no* verbs with tone; see footnote 5).

For short vowels, there are even fewer minimal pairs (e.g. Fig. 4). In fact, there are arguably *no* true minimal pairs for tone on short vowels. [H] tone on penultimate short vowels always involves stress shift (section 1.1), which means that tonal words of this type are also distinguished from non-tonal words by the *position* of prominence, apart from the type of prominence involved (tonal vs. non-tonal).

Furthermore, tone is frequently introduced by overt morphology in Uspanteko (section 1.1, Table 1). In many cases, this too renders tone functionally redundant for the purpose of distinguishing words (4).

(4) a. [iχ-ˈɓa] 'his/her head' ∼ [ˈín-ɓa] 'my head'
     b. [kʔaχ] 'wheat' ∼ [in-ˈkʔáːχ] 'my wheat'

Additionally, tonal forms are often in free variation with non-tonal forms of the same word: for example, we have recorded numerous narratives in which people switch between productions like [ˈí.ʃim]∼[i.ˈʃiːm] 'corn' or [ˈkú.t͡sʔiχ]∼[ku.ˈt͡sʔiːʔχ] 'flower', with only a few seconds intervening between each rendition. (Such variation may reflect a high degree of bilingualism with K'iche', a language that shares many cognate, non-tonal roots with Uspanteko.)

The low functional load of tone in Uspanteko—particularly the absence of dependable minimal pairs—provides a practical impediment for studying the tone system (Snider 2014). And indeed, there are inconsistencies in how tone has been described by researchers working on the language. Such discrepancies might be attributed, in part, to the lack of minimal pairs which would help establish the relevant tonal contrasts.

In the following section we lay out the varied descriptions of Uspanteko tone that can be found in the existing literature on this language. The divergences across these descriptions help motivate the phonetic study of tone which we present in the remainder of the article.

1.3 Determining the tonal inventory

To this point we have assumed Bennett & Henderson's (2013) privative [H]∼∅ analysis of the Uspanteko tone system. This analysis builds on earlier descriptions of the Uspanteko tone system (particularly Campbell 1977 and Can Pixabaj 2007), but also differs from those descriptions in substantive ways. Grimes (1971, 1972) characterizes tone in Uspanteko as a binary contrast between rising [LH] tone and falling [HL] tone on long vowels. Campbell (1977) also proposes a binary contrast on long vowels, but assumes simple high [H] and low [L] tones, rather than contours. Neither Grimes nor Campbell report a tonal contrast on short vowels, but they do observe stress shift to the penult, which we claim is driven by the presence of an [H] tone in that position. Can Pixabaj (2007) describes a more complex system, with a privative [HL]∼∅ contrast on stressed

long vowels; [H] tone on stressed, penultimate short vowels; and no tone at all on short vowels in final stressed syllables.

The brief, schematic description in Kaufman (1976:19-20,100-2) assumes a three-way [HL]∼[H]∼∅ tone inventory for long vowels, and a privative [H]∼∅ inventory for short vowels, with the proviso that only one tone may occur per word. Kaufman also transcribes some non-final long vowels, contrary to our assertion that long vowels are strictly limited to stressed final syllables (section 1.1).[8] Not all of the tones proposed by Kaufman are directly contrastive: the falling [HL] tone only occurs on long vowels in final syllables, while non-final long vowels are realized with [H] tone instead, unless they are toneless. Toneless words are assigned a default word-final accent. Kaufman treats this word-final default accent as being on a par (at least phonetically) with lexical [H] tone on non-final vowels. Lastly, the earliest descriptions of Uspanteko in Stoll (1884, 1887, 1888, 1896) note some cases of non-final stress, but do not describe tone.

These descriptions essentially agree that Uspanteko possesses two tonal categories on long vowels, and an [H] tone on penultimate stressed vowels. But there is disagreement as to (i) the phonetic values associated with each tonal category, and (ii) whether the tonal system involves a privative contrast between the presence vs. absence of tone, or a binary contrast between two specified tones. Finding empirical evidence to resolve this issue is not straightforward. First, the coarse pitch patterns on long vowels are often compatible with several different analyses. Consider the leftmost diagram in Fig. 11. The pitch contour on this instance of [t͡ʃúːn] is phonetically falling, but this phonetic fact alone underdetermines the corresponding phonological analysis. Is this vowel specified for a falling [HL] contour as such, or does the f0 trace reflect a level [H] specification, with a gradual fall towards neutral f0 following the achievement of the [H] target (Bennett & Henderson 2013)? And if [t͡ʃúːn] is lexically specified with an [HL] tone, should we expect a falling f0 contour to be an invariant property of this word across different utterances? If so, what do we say about the many tokens of this item which *lack* a falling pitch contour, such as the rightmost diagram in Fig. 11? Which properties of these f0 contours reflect the basic phonological specification of lexical tone, and which are due to the influence of contextual factors?

The same issues arise with respect to the question of privative vs. binary contrast. This is essentially a question about tonal underspecification: are all stressed vowels specified for tone, or

---

[8]Short vowels in Uspanteko are often lengthened in stressed, tonal penults, but do not neutralize in duration with true phonological long vowels (Bennett et al. ms.a and section 3.6 below). We suspect that most of the non-final, [H]-toned 'long' vowels transcribed by Kaufman (1976) are in fact phonological short vowels that have been phonetically lengthened under tone. The same can be said for many, if not all of the non-final long vowels reported by Grimes (1971, 1972) and Campbell (1977).

The Uspanteko speakers we have worked with have clear intuitions about vowel length in stressed final syllables, where length is plainly contrastive (e.g. *is* /is/ 'sweet potato' vs. *iis* /iːs/ 'pink (like flesh)'). When asked for judgments about vowel length in non-final syllables, they either report that non-final vowels are short, or have difficulty identifying vowel length in those positions. Furthermore, if non-final long vowels were permitted in Uspanteko — as reported by Grimes (1971, 1972), Kaufman (1976) and Campbell (1977) — the systematic shortening of long vowels in suffixed forms like Fig. 9 (e.g. [aχ.ˈtiːχ]∼[aχ.ˈtí.χiɓ]) is mysterious and calls for explanation (for more examples of this type, see e.g. Can Pixabaj 2007:68,87-8, Us Maldonado 2010:Ch. 1-2).

One possibility is that non-final long vowels *were* permitted in Uspanteko, but have been lost in the 50 years separating our fieldwork from the work of Grimes, Kaufman and Campbell. The occurrence of non-final, unstressed long vowels is a point of systematic variation both within and across K'ichean-branch Mayan languages (e.g. Campbell 1977, Dayley 1985, Larsen 1988, Bennett 2016). In fact, the one example we know of that consistently shows a non-final long vowel—*tziijb'al* [t͡siːχ.ˈɓal] 'language'—is described by some speakers as feeling "more like a K'iche' word than an Uspanteko word" (cf. Uspanteko *qayóloj* [qa.ˈjó.loχ] 'our language, speech').
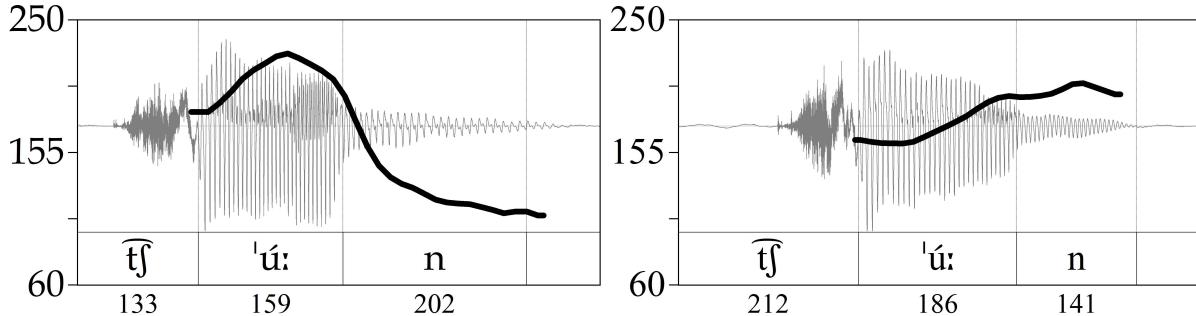
Figure 11: Two renditions of *chúun* [t͡ʃúːn] 'calcium, lime' (speaker TAML 2019)

only those stressed vowels which are associated with raised f0? Consider Fig. 3: the contrast in pitch levels between *xáab'* [ˈʃáːɓ] 'vomit' and *xaab'* [ˈʃaːɓ] 'comb' can of course be interpreted as a binary [H]∼[L] contrast, given the differences in pitch height in these two words. But these f0 differences can just as easily be taken as the realization of a privative [H]∼∅ contrast, on the assumption that tonally unspecified vowels are realized lower in the speaker's pitch range than vowels specified with an [H] tone (e.g. Campbell 2016 on Zenzontepec Chatino).

The upshot is that surface phonetic forms of isolated words are a poor guide to tonal specification. Better evidence for binary contrast would come from grammatical patterns which show that two distinct tones need to be phonologically represented (e.g. Myers 1998 and references there). Were there phonological rules in Uspanteko which clearly referred to two distinct tones, we would have strong evidence for binary specification.

As it stands, we are aware of only two grammatical generalizations involving tone in Uspanteko: the requirement that [H] tone be realized on the penultimate mora, and the requirement that [H] tone coincide with stress. Since both of these generalizations concern just the [H] tone, they cannot be taken as evidence for binary tone specification. Of course, such patterns do not *falsify* the claim that Uspanteko makes use of a binary tone contrast, since a second tone (e.g. [L]) could be phonologically contrastive without also being implicated in phonological rules. The grammatical behavior of tone thus fails to settle the issue of binary vs. privative tonal contrast in Uspanteko.

Making matters worse, even the tonal properties of individual words cannot be taken for granted. Descriptive sources on Uspanteko achieve surprisingly little consensus on which words belong to which tonal categories, a fact already noted by Fox (1978:65). Some examples are shown in Table 2 below.[9]

As Table 2 illustrates, it is not difficult to find numerous discrepancies across these sources. Some of these differences are probably analytical—Grimes' [HL] and [LH] tones, for instance, might reasonably be equated with Campbell's [L] and [H] tones, respectively. But there is a troubling lack of consistency in how words are assigned to tonal classes in these fieldworker descriptions. Grimes' [LH] tone corresponds to both [L] and [H] tones in Campbell (1977); Campbell's [H] tone corresponds to both [HL] and [LH] in Grimes (1972), and to both [H] and [HL] in Kaufman (1976); Can Pixabaj's toneless vowels correspond to both [H] and [L] in Campbell (1977); and so on. The utter lack of systematic correspondence between tonal categories in these sources makes it extremely hard to assess competing proposals about the tonal inventory of the language.

---

[9]Vicente Méndez's (2007) Uspanteko dictionary marks tone, and largely agrees with Can Pixabaj (2007) in how it classifies words into tone categories. Another dictionary, Us Maldonado (no date(a)), does not mark tone.

Table 2: Differing tonal categories across Uspanteko sources (Grimes 1972, Kaufman 1976, Campbell 1977, Can Pixabaj 2007, Bennett & Henderson 2013)

Item	Grimes	Kaufman	Campbell	Can Pixabaj	Our fieldwork
[poːp] 'petate'	[LH]	—	[L]	∅	∅
[aːqˀ] 'tongue'	[LH]	[H]	[L]	∅	∅
[t͡ʃeː] 'tree'	[LH]	[H]	[H]	∅	∅
[kuːkˀ] 'squirrel'	[HL]	—	[L]	[H]	∅ ~ [H] across speakers
[ɓaːq] 'bone'	[HL]	[H]	[H]	∅	∅
[qˀiːχ] 'sun'	—	[HL]	[L]	∅	∅
[t͡ʃiː(ʔ)] 'mouth'	[LH]	[HL]	[H]	Short, ∅ [t͡ʃiʔ]	∅
[t͡ʃaːχ] 'ashes'	—	[HL]	[L]	[H]	[H]

Some of the heterogeneity across these sources likely reflects genuine phonological and lexical variation across speakers of Uspanteko. Inter-speaker linguistic differences are indeed widespread in the Uspanteko community. It may also be that the tone system underwent changes between the time of earlier sources like Grimes (1972), and more contemporary work like Can Pixabaj (2007). But in our view, sociolinguistic and diachronic variation are not sufficient to explain the wide divergences found across firsthand descriptions of the tonal system of Uspanteko.[10]

We suspect that research on Uspanteko phonology has been significantly confounded by the influence of intonation on the phonetic realization of lexical tone (see also Bennett et al. 2019, and Pike 1946 on tone in Yucatec Maya). Fig. 12 provides another example of the tone contrast on long vowels, using the same minimal pair seen earlier in Fig. 3. These items were elicited in final position in a wordlist, and the phonological contrast between tonal categories is phonetically preserved in this context (i.e. f0 is higher for [ˈʃáːɓ] 'vomit' than for [ˈʃaːɓ] 'comb', just as in Fig. 3).

Fig. 13 shows these same items, produced by the same speaker in the same recording session, but now in medial position in a wordlist. In this context the lexical tone contrast is completely masked by a rising intonational contour which is superimposed on both items. Auditorily, the pitch contours in Fig. 13 are almost indistinguishable, suggesting that the lexical tone contrast has been completely neutralized.

These examples illustrate a well-known pitfall of studying tone: isolation forms may not provide reliable information about lexical tone, because the pitch contours on isolated words can be significantly affected by intonational patterns associated with phrases and utterances (e.g. Pike 1948, Bruce 1977, Himmelmann 2006, Himmelmann & Ladd 2008, Hyman 2014, Jun & Fletcher 2014,

---

[10]It is also relevant that different sources on Uspanteko emphasize different levels of description. Bennett & Henderson (2013) are most concerned with the grammatical logic of the tone system, laying out how tone, stress, vowel length, and morphology interact to produce accentual alternations on roots. Other sources describe how tone is phonetically realized on individual words, but are much less concerned with analyzing tonal phonology or describing alternations related to word-level prosody (though Can Pixabaj 2007 provides very good information about phonological alternations in Uspanteko).
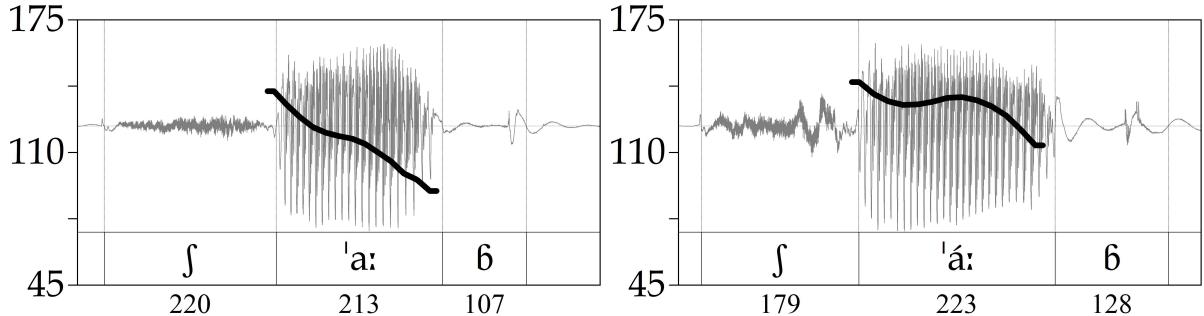
Figure 12: *xaab'* [ˈʃaːɓ] 'comb' (left) vs. *xáab'* [ˈʃáːɓ] 'vomit' (right), produced as final items in wordlist (speaker MH 2016). The lexical tone contrast is phonetically preserved.
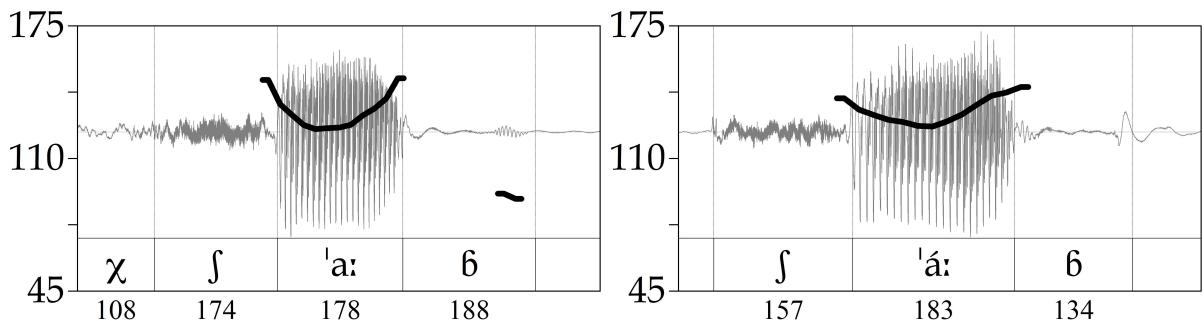
Figure 13: *jxaab'* [ˈχʃaːɓ] 'his/her comb' (left) vs. *xáab'* [ˈʃáːɓ] 'vomit' (right) as medial items in wordlist, with rising intonation (speaker MH 2016). The lexical tone contrast is phonetically neutralized.

Snider 2014, Yu 2014).

Recall that a major disagreement in the description of Uspanteko concerns whether tone should be defined using pitch levels, pitch contours, or some combination of the two. One possibility is that the contours described by Grimes (1971, 1972), Kaufman (1976) and Can Pixabaj (2007) — and the disagreements between them as to what the contours actually *are* (Table 2) — reflect additional pitch variability introduced by higher-level intonation. As in other languages, intonation in Mayan languages is often expressed with large rising or falling pitch contours (Berinstein 1991, Baird 2014a,b, 2018, Bennett 2016, England & Baird 2017, DiCanio & Bennett 2021 and references there). Utterance-final rising intonation is common for both interrogative and declarative sentences, and is illustrated for Uspanteko in Fig. 14.

The same utterance-final rising contour can in fact be observed on some of the single-word examples already presented in this article. Consider Figs. 2, 5 (lefthand side), 7 (lefthand side), 9 (lefthand side), 11 (righthand side), and 13. All of these items show some form of rising contour, though they differ in how sharply the rise is implemented. Some of these items were elicited in isolation or in wordlists. Other items were elicited using frame sentences, which often encourage isolation-type prosody on target items. In all of these cases, it seems plausible that the observed rises are not inherent properties of these words themselves — they are not lexical tones — but rather represent the influence of utterance- or phrase-final intonation on phonetic form (see also Bennett et al. ms.a). This is especially clear in Figs. 12 and 13, where actual alternations in pitch

13
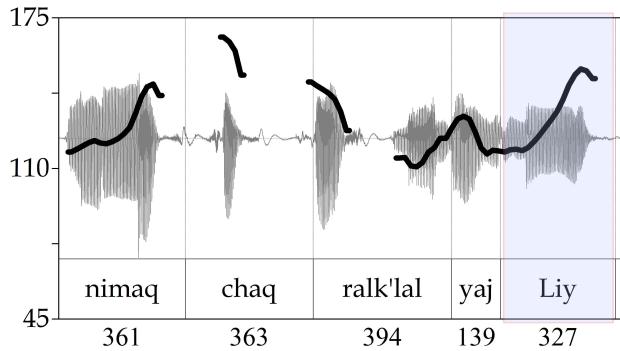
Figure 14: *Nimaq chaq ralk'lal yaj Liy* 'Maria's sons are very large.' (speaker 4 ᴘ, 2016)

shape can be observed depending on the sentential context of the items.

It thus seems worth considering whether the word-level contour tones described in Grimes (1971, 1972) and Kaufman (1976) might be better understood as some combination of lexical and intonational melodies, rather than directly reflecting lexical tone as such. The extensive variation in how tones are transcribed for individual words across sources (Table 2) seems to us to raise exactly the same possibility.

Examples like Figs. 11-13 suggest that intonation may have an extensive influence on how lexical tone contrasts are phonetically realized in Uspanteko. This certainly presents challenges for the study of stress and tone. At the same time, the precise details of how tone and intonation interact in Uspanteko has the potential to shed light on the phonological representation of lexical tone. It has been shown for a range of other languages that intonational patterns distinguish between toneless vowels and vowels which are lexically specified for tone (e.g. Pierrehumbert & Beckman 1988, Myers 1998, Remijsen & van Heuven 2005, Riad 2006, Remijsen et al. 2014). Finding comparable evidence in Uspanteko would provide support for the privative [ʜ]∼∅ analysis of tone developed in Bennett & Henderson (2013), Bennett et al. (2019). In the remainder of the article we discuss a production study designed to probe exactly this question.

2 Production study of tone and intonation

The goal of our study was to examine word-level prosody in Uspanteko under different intonational and discourse conditions. This was done by means of a production study in which native speakers of Uspanteko produced target words in a series of sentential contexts, using a question-answer elicitation method (e.g. Himmelmann 2006, Himmelmann & Ladd 2008, Bennett & Henderson 2013, Clopper & Tonhauser 2013, Baird 2014a, 2018, Hedding to appear).

A quick note on terminology: in discussing our study, we use the term 'accented syllable' to refer to any kind of stressed syllable, regardless of what its tonal specification might be. 'Accent' in this sense includes both tonal prominence (e.g. stress + [ʜ] tone) as well as non-tonal prominence (stress alone).

2.1 Items

Target items were selected to cross two factors: tonal specification of the accented vowel ([ʜ] vs. ∅, following Bennett & Henderson 2013), and length of the accented vowel (short ['V] vs. long ['Vː]; Table 3). There were twelve target items in the study, three in each of the cells in Table 3.

14

A full list of the target items used in our study is provided online at `https://github.com/rbennett24/articles/tree/master/Uspanteko_intonation`.

Table 3: Sample items by condition

	Short [ˈV]	Long [ˈVː]
No tone	/kʔeˈten/ 'hot'	/qaˈpoːp/ 'our sleeping mat'
Tone	/ˈqálaq/ 'our plate'	/qaˈχóːqʔ/ 'our corn husk'

2.1.1 Item design and ecological validity

It is common in studies of tone and intonation to investigate f0 patterns in words containing only sonorant segments (e.g. Jun & Fletcher 2014, Yu 2014). The rationale for doing so is fourfold. First, f0 contours can only occur on voiced segments. Sonorants have particularly robust voicing, making it easier to observe f0 movements. Second, recording words composed entirely of sonorants guarantees that no portion of the f0 contour will be 'hidden' by voiceless intervals. Third, words containing only sonorants may show less microprosodic variation in f0 than words with alternations between sonorants and obstruents, especially voiceless obstruents (e.g. Lehiste 1970, Hanson 2009, Kirby & Ladd 2016). This is even more true for languages like Uspanteko which have glottalized consonants or other laryngeally-complex sounds (e.g. Kingston 2005). Fourth, the shape, size, temporal alignment, and/or inventory of f0 patterns may vary depending on whether sonorants occupy key positions like syllable codas, phrase-final/initial position, and so on (e.g. Zec 1994, Jun 1998, Zhang 2001, Gordon 2006, Morén & Zsiga 2006, Remijsen 2013).

As implied by the sample items in Table 3, we did not follow the convention of studying f0 in words composed only of sonorant segments. This decision was made for both practical and theoretical reasons, mostly connected with the fact that Mayan languages are not rich in sonorants. Uspanteko has 16 phonemic obstruents, /p ɓ t tʔ t͡s t͡sʔ t͡ʃ t͡ʃʔ k kʔ q qʔ ʔ s ʃ χ/. All of these obstruents are voiceless except /ɓ/, and even /ɓ/ is frequently realized as voiceless [ɓ̥] (Bennett 2016). In contrast, there are only 6 phonemic sonorants /m n l r j w/. Furthermore, /r/ is often realized as a voiceless fricative akin to [ʂ] in Uspanteko (Can Pixabaj 2007, Bennett 2016; we've also observed similar devoicing, to a much lesser extent, for /l/).

This phonemic skew has consequences for the distribution of sonorants in the lexicon and in actual speech. Using a corpus of approximately 54,000 words drawn from transcriptions of spontaneous narratives in Uspanteko, we estimate that sonorants make up about 30% of consonants in running speech when counted over individual word tokens (46,003/155,629 consonants), or 27% when counted over unique word types (9535/35,716).[11] These distributions are directly proportional to the representation of sonorants in the phonemic consonant inventory of Uspanteko (6/22 consonants = 27%). The mean length of each unique word type in this corpus is ≈ 6.7 segments (just about what might be expected from the typological survey in Nettle 1995). But the average

---

[11]This corpus comes from the Oxlajuuj Keej Maya' Ajtz'iib' Mayan Languages Collection of transcribed recordings of Uspanteko, as well as our own transcribed recordings of the language. All of these materials are publicly available online at `https://ailla.utexas.org/`. Some word types in this corpus clearly correspond to variant pronunciations of the same word, typically reflecting patterns of vowel deletion and/or vowel length variation (e.g. *woron* [woˈron]∼[woˈroːn]∼[wˈron] 'forest, jungle'). For evidence that small, noisy corpora can provide reliable estimates of some lexical, phonological, and phonetic measures, see Tang & Bennett (2018, 2019), Bennett et al. (2018b), Dockum & Bowern (2019), Strunk et al. (2020) and references there.

number of sonorant consonants per unique word type is only ≈ 1.1, while the average number of obstruents per unique word type is ≈ 3.0.

There are very few unique word types in this corpus containing *only* sonorant consonants (278/8667 = 3%). Strikingly, almost all of these words are short function words (e.g. the multifunctional preposition [li]) or borrowings from Spanish (e.g. [lu.ˈwar] 'place' < Spanish *lugar*). Only three of these words have lexical high tone, [íwir] 'yesterday' (often pronounced as [íws̝]), [ínwa] 'my tortilla', and [ánm] 'woman'. These examples constitute 0.03% of the unique words in our corpus; note that there are no examples of tone on a long vowel among those words containing only sonorant segments. Furthermore, our own experience *trying* to find all-sonorant words for each of the four prosodic categories in Table 3 suggests that such words are very sparsely attested in Uspanteko, particularly among common vocabulary items.

While these findings are based on a small and unpolished corpus, it nonetheless seems clear that all-sonorant words are probably quite marginal in the native Uspanteko lexicon and in Uspanteko speech. For that reason, we doubt that f0 patterns in all-sonorant items are actually representative of how f0 is typically used to indicate tonal or intonational distinctions in the language. Words containing obstruents arguably have greater ecological validity for the study of f0 in Uspanteko.

From a theoretical perspective, the preference for all-sonorant items in research on tone and intonation seems to assume, at least implicitly, that there is a 'true' f0 contour for any given phrase or word which may be partially hidden during voiceless intervals. We are skeptical of this assumption. In particular, Barnes et al. (2011, 2012) report that voiceless segments do not contribute to the perception of intonational contours in English. That is, listeners do not seem to reconstruct what f0 'would have been' during voiceless segments, had they been voiced instead. This result is surprising if voiceless consonants mask an underlying f0 contour which listeners must somehow perceptually recover. Rather, listeners seem to judge intonational contours solely on the basis of voiced segments, where they have positive evidence about f0 values, and ignore voiceless segments entirely. These findings suggest that we cannot necessarily take the f0 contours of all-sonorant words to be 'better' exemplars of tonal or intonational patterns than f0 contours on words containing voiceless segments (though to be sure, much more work on these questions is needed, on a range of other languages).

A separate issue in item design arose from the segmental phonology of Uspanteko. Most of our target items were bisyllabic, as we originally intended to normalize the f0 of accented vowels relative to f0 on unstressed vowels in the same word (e.g. *qálaq* [ˈqá.laq] 'our plate'). However, unstressed vowels frequently delete in Uspanteko (e.g. [ˈqá.laq]∼[ˈqálq]), rendering such comparisons impossible to carry out in our data. Phonetically speaking, many of our items were thus monosyllabic, either underlyingly or as the result of vowel deletion, and so our phonetic measures focused on the properties of accented vowels taken in isolation. For some rough, preliminary observations about word-level f0 patterns in Uspanteko, see Bennett et al. (2019); and for the details of vowel deletion see Bennett & Henderson (2013), Bennett (2016), Bennett et al. (ms.b).

2.2 Methodology

As discussed in section 1.3, utterance-final position is often associated with large falling or rising pitch contours in Uspanteko. Rising pitch contours (Fig. 14) are particularly prevalent, in both declarative and interrogative sentences. In order to explore the effect of utterance-final intonation on the phonetic realization of word-level prosodic contrasts, target items were elicited in both sentence-final and sentence-medial position.

In many languages the phonetic realization of stress, tone, and intonation can be affected by the structure of the larger discourse in which a word occurs. Semantic and pragmatic focus are often cued by manipulating prosodic parameters like pitch, duration, and intensity (e.g. Xu 1999, Gussenhoven 2004, Ladd 2008a, Kügler & Genzel 2012, Féry 2013, Féry & Ishihara 2016, Di-Canio et al. 2018, Wagner to appear, Hedding to appear, and references there). This is also true for Mayan languages (see Bennett 2016, England & Baird 2017, DiCanio & Bennett 2021 for details and citations). Baird (2014a, 2018) reports that increased segmental duration, earlier timing of intonational events, and pitch range expansion are all phonetic reflexes of focus in K'iche', particularly for words which are focused *in situ* (focused elements often undergo discourse fronting in Mayan languages; Aissen 1992, 2017). K'iche' is closely-related to Uspanteko, and is spoken by most people in the Uspanteko area. We therefore expect that the information structure of the discourse could have an effect on the phonetic realization of tone and stress in Uspanteko.

We manipulated the discourse context of our target items by means of a question-answer elicitation method. A native speaker of Uspanteko read from a list of pre-prepared question prompts, and participants responded by translating a pre-prepared answer from Spanish to Uspanteko. We used Spanish prompts because literacy in Uspanteko is generally low, and even literate speakers are often more comfortable reading in Spanish.

The question prompts were of two types. To elicit a target word under focus, we used a WH-question as a prompt (5). The target word appeared in the answer as the correlate of the WH-word — that is, the target word could have occurred as a valid answer to the question on its own. In the typology of semantic focus, this corresponds to 'informational narrow focus' (on different types of focus, see Gussenhoven 1983, Féry 2013, Féry & Ishihara 2016). As a shorthand we will simply call this the 'Focus' condition.

(5) Target item focused, in sentence-medial position
    a. **Q**: NEN *kan xawiil li ch'aat?* 'WHAT did you see on the bed?'
    b. **A**: *Xinwiil* SUQ'UK' *li ch'aat.* 'I saw a LOUSE on the bed.'

Baird (2014a, 2018) finds that contrastive focus is more prosodically marked than broad focus in K'iche' ('broad focus' involves focus on the entire expression rather than a sub-part, as in e.g. an answer to the question 'What happened?'). Informational narrow focus — the focus type used in our study — falls between broad and contrastive focus on the 'strength' scale (Féry 2013).

The counterpart of the focus condition was elicited using polar 'yes-no' questions. In this condition the target item was itself mentioned in the question, as in (6). In the corresponding answer the target word is then discourse-given, in the sense of having been previously mentioned in the discourse. The target word is additionally pragmatically backgrounded because it does not by itself provide the information which the yes-no question is seeking.

(6) Target item discourse-given, in sentence-final position
    a. **Q**: *Tiqasu'* **qaxoot***?* 'Do we clean our **comal**?'
    b. **A**: *Ji'n, tiqasu'* **qaxoot**. 'Yes, we clean our **comal**.'

For convenience, we will call this the 'Given' condition, recognizing that other researchers might prefer different terminology (e.g. 'previously mentioned').

There were 48 total question-answer pairs in the study, consisting of twelve target words (vowel

length (2) × tone (2) × 3 items each = 12) produced in sentences reflecting each combination of discourse status and position (2 × 2 = 4 × 12 = 48). Each speaker went through the list of question-answer pairs at least once, often repeating the same question-answer pair in order to guarantee that we had recorded at least one fluent, natural rendition. Overall these speakers produced between 57-113 usable tokens each (mean = 79, median = 69, SD = 21), yielding a total of 945 word tokens for analysis (= 1505 vowel tokens, 945 of which were stressed). 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. A waveform and pitch track for an example response are given below in Fig. 16.

2.2.1 Methodological issues

The question-answer methodology we used in this study is not particularly naturalistic. First, it involves a scripted translation task, which is quite unlike a regular discourse. Second, the answers themselves are unlike the responses typically found in actual Uspanteko conversations. In the Given condition, for example, a natural response would normally involve the omission of discourse-given material (i.e. 'pro-drop'), rather than repetition of the target item (7) (Brody 1984, Du Bois 1987, England 1991, England & Martin 2003).

(7)   a.   **Q**: *Wi' tinik laj qálaq?* 'Is there meat on our plate?'
        b.   **A**: *Ji'n (wi').* 'Yes (there is).'
        c.   **Our study**: *Ji'n, wi' tinik laj qálaq.* 'Yes, there's meat on our plate.'

Similar issues arise for the Focus condition. In Mayan languages focused elements frequently undergo discourse fronting to the left periphery (8) (Aissen 1992, 2017, Can Pixabaj & England 2011).

(8)   a.   **Q**: *Nen rik'il tiqajach qaaw re chwee'q?* 'With what will we serve our food tomorrow?'
        b.   **A**: [*Rik'il li qálaq*] *tiqajach* _____. 'It's with our plate that we'll serve it.'
        c.   **Our study**: *Tiqajach rik'il li qálaq.* 'We'll serve it with our plate.'

In this case the issue is somewhat less severe. Informational narrow focus is a relatively weak type of focus, and need not trigger discourse fronting in Mayan languages (Féry 2013, Aissen 2017). Further, *in situ* focus is attested as a focus strategy in a number of diverse Mayan languages, including K'iche' (Velleman 2014b,a), Tseltal (Polian 2013:773ff), and Yucatec Maya (Kügler & Skopeteas 2006, 2007, Kügler et al. 2007, Gussenhoven & Teeuw 2008, Verhoeven & Skopeteas 2015). More locally, responses to WH-questions which employ *in situ* focus are indeed possible in Uspanteko, and are judged acceptable by our consultants. And since *in situ* focus is more prosodically marked than focus fronting in several Mayan languages (Kügler & Skopeteas 2006, 2007, Kügler et al. 2007, Baird 2014a, 2018, England & Baird 2017), *in situ* focus may be a better testing ground for exploring the interaction of tone and intonation. The somewhat artificial character of these materials should nonetheless be kept in mind when interpreting the results (section 3).

2.3 Participants

Twelve native speakers of Uspanteko participated in this study. Eight of these speakers were recorded in 2017, and four in 2018 using the same elicitation materials. Bennett et al. (2019)

report preliminary results with just the eight speakers recorded in 2017, and analyze different aspects of the data than what is analyzed here. Speaker ages ranged from 24-59 (mean = 40, median = 41, SD = 9.5). Seven speakers were female, and five male. Eight of the speakers were from the town of San Miguel Uspantán, three from the village of Las Pacayas, and one from the village of Chipaj (see map in Fig. 1). One additional participant recorded in 2017 had trouble completing the task, and their data is not reported here.

2.4 Data analysis

The sentences were manually transcribed, and transcriptions were then converted from the Uspanteko orthography to a phonetic transcription using a custom Python script. These phonetic transcriptions were then semi-automatically time-aligned with their associated audio recordings using the Montreal Forced Aligner (McAuliffe et al. 2017). The forced alignment model was trained with the default settings of the Montreal Forced Aligner, and included speaker-specific training of model parameters. Visual inspection of automatic aligments suggested surprisingly high accuracy, but the segment-level alignments were still hand-corrected for the target words.

Pitch contours on target items were then analyzed in Praat (Boersma & Weenink 2016). The recordings were downsampled to 16kHz, then pitch values were automatically extracted with a script specifying by-speaker pitch ranges following the recommendations of De Looze & Rauzy (2009) and Evanini et al. (2011). These pitch tracks were smoothed with a bandwidth of 15Hz. Time-normalized pitch measurements were produced by averaging pitch values in Hz over $1/9$ intervals of the duration of each vowel. The resulting time-normalized contours were manually inspected for obvious errors. While missing values sometimes occurred, especially at the beginning or end of a contour, no other significant errors were observed.

Pitch values for each speaker were then *z*-score normalized (see too DiCanio 2014, Wang et al. 2020).[12] Pitch measurements more than 2.5 *z*-units away from each speaker's mean were treated as outliers and removed from the data. This resulted in the the elimination of 1.1% of the time-normalized pitch data. Along with our time-normalized pitch measurements, we also measured mean f0 for each vowel, and followed the same outlier removal procedure, eliminating 0.8% of the mean pitch data.

All plots in this article were drawn with the GGPLOT2 package (Wickham 2009) in R (R Development Core Team 2020). Scripts used for data preparation and analysis are available online at https://github.com/rbennett24/articles/tree/master/Uspanteko_intonation.

2.5 Expectations

Little is known about intonation in Uspanteko. That fact, along with the unclarity surrounding the phonetic realization of tone, makes it hard to have clearly statable hypotheses about the results of this production study. However, we can state some soft expectations to guide the overall analysis, which is essentially exploratory and descriptive in nature. The following predictions assume that

---

[12]The output of *z*-score normalization was highly correlated with some alternative methods for normalizing pitch values across speakers: semitone transformation relative to each speaker's mean pitch in Hz ($r = 0.97$) (Zhang 2019); Hz divided by the size of each speaker's pitch range as estimated by the De Looze & Rauzy (2009), Evanini et al. (2011) formula ($r = 0.95$); and range normalization using 2% and 98% estimates of floor and ceiling values in Hz ($r = 0.99$) and in semitones ($r = 0.99$) (Bardiaux & Mertens 2014). See also Ladd (2008a:192-202). Converting the raw Hz values to semitones, log(Hz), or ERB prior to *z*-score normalization has effectively no consequences for the results ($r > 0.99$ for *z*-scores in all three cases).

tone in Uspanteko involves a privative [H]∼∅ contrast (Bennett & Henderson 2013), rather than any of the alternative possibilities set out in section 1.3. For the most part, we limit ourselves to the phonetics of stressed (= 'accented') vowels.

(9)　**Guiding expectations for data analysis**:

　　a.　Uspanteko is a tonal language. When intonation is controlled for, vowels with [H] tone should have higher f0 than stressed, non-tonal vowels. (Section 3.1)

　　b.　Vowels with [H] tone will not necessarily show a consistent, fixed f0 shape (as we might expect if they were instead specified with [HL] or [LH] contours). (Section 3.5)

　　c.　Tone contrasts should be clearest in medial position, but may be masked in final position under the influence of intonational boundary tones (particularly (L)H% rises). (Section 3.1)

　　d.　Focus should expand the overall pitch range — as it does in K'iche' (Baird 2014a, 2018) — thereby enhancing the f0 differences between tonal and non-tonal vowels. (Sections 3.1, 3.2; see also the 'focus-to-accent principle' of Gussenhoven 1983)

　　e.　Points (9c) and (9d) imply that tonal contrasts should be most phonetically salient in sentence-medial position under focus. (Sections 3.1, 3.2)

　　f.　Focus may affect the f0 of tonal vowels (which are specified with an [H] target) to a greater extent than the f0 of non-tonal vowels. This assumes that focus has a greater effect on f0 for vowels which are phonologically specified with an abstract tonal or intonational target, when compared to vowels lacking such a target. (See Elordieta 2007, 2008, Elordieta & Hualde 2014 for a pattern of this type in Basque, and Sugiyama 2011 on Japanese.) (Section 3.2)

　　g.　Vowels should be phonetically longer (i) under focus, (ii) in utterance-final contexts, and (iii) when bearing [H] tone. These predictions stem from typological studies of duration (e.g. Myers & Hansen 2007, Kügler & Calhoun 2021) as well as our own prior research on Uspanteko (Bennett et al. ms.a; see also Faytak & Yu 2011).

3 Results

In this section we report the results of three different analyses. First, we plot characteristic f0 trajectories for accented vowels, and statistically model differences in mean f0 as a function of tone, vowel length, discourse status, and sentential position (section 3.1). Second, we analyze durational differences across vowels, taking into account the same factors considered in the analysis of f0 (section 3.4). Third, we use a combination of principal component analysis and *k*-means clustering to assess whether each tonal category might be associated with a particular f0 shape, as well as a particular f0 level (section 3.5). At each step we consider whether and how our phonetic results might help reveal the phonolgical specification of tone in Uspanteko.

3.1 Pitch distinctions by vowel length

Fig. 15 shows loess-smoothed time-normalized f0 contours for tonal and non-tonal long vowels, across all four combinations of discourse status (given/previously mentioned vs. new information focus) and sentential position (sentence-medial vs. sentence-final). Several things are apparent from this plot. First, tonal distinctions on long vowels are preserved in all four conditions. Still, the magnitude of the f0 differences between tonal and non-tonal vowels is not particularly large:

differences in mean f0 across the four cells range from 0.38-0.58 *z*-units. We return to a discussion of the magnitude of the f0 distinctions associated with tone in section 3.6. These results are nonetheless consistent with a tonal contrast such as [H]∼∅ on long vowels. The potential influence of rising (L)H% intonation on the realization of tone in utterance final position is not immediately apparent from these plots.
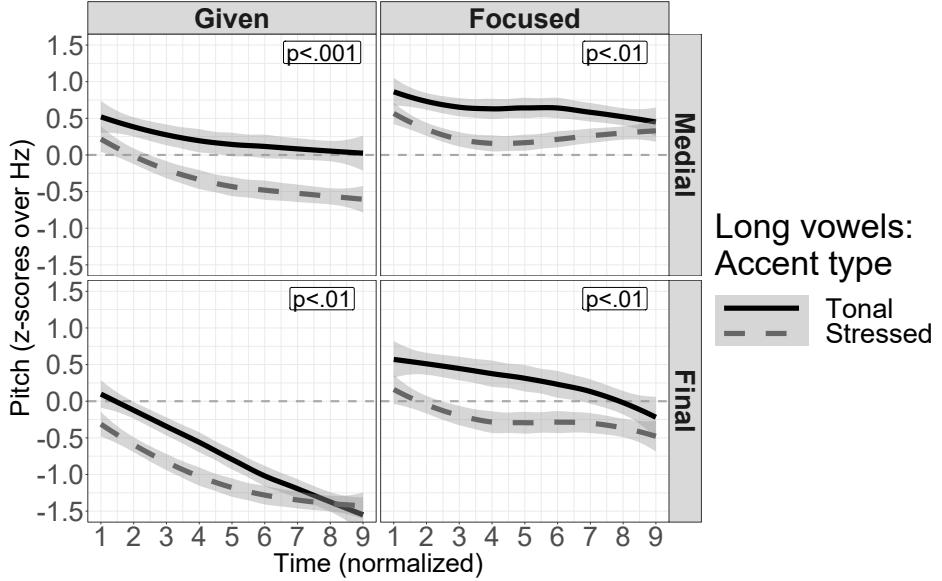

Figure 15: f0 trajectories (in *z*-scores over Hz) for long vowels grouped by discourse condition and sentential position. Grey bands indicate confidence intervals around the estimate of the position of the smoothed loess regression line at each time step. *p*-values refer to a comparison of mean f0 measurements for tonal vs. non-tonal vowels in each cell, using a one-sided Welch *t*-test for unequal variances.

Comparing the left and right columns in Fig. 15, we also see that focus involves f0 raising (see below for a supporting statistical analysis). Alternatively, it may be that focus leads to an overall expansion of the pitch range, as in Mandarin and other languages (Xu 1999, Hartmann 2008). The relative tonal and intonational sparseness of Uspanteko makes it difficult to distinguish between these two interpretations (for related discussion, see e.g. Roettger 2017 on Tashlhiyt Berber, Borise to appear on Caucasian languages, and Baird 2014a, 2018 on K'iche').

Comparing the top and bottom rows of Fig. 15, we see that pitch values are lower in final position, suggesting a pattern of f0 declination across the sentence (Bennett et al. 2019). Declination can indeed be observed on many individual tokens in our study, such as Fig. 16, even for tokens showing final rising intonation (e.g. Fig. 14). Note as well that the tonal vowel in the focused target word *rixóqil* 'his wife' in Fig. 16 is realized with the highest f0 in the sentence.

The overall patterns of f0 contrast are somewhat different for short vowels (Fig. 17). The basic contrast between tonal and non-tonal vowels is preserved in two conditions: the Focused, Medial condition (upper-right; Δ mean *z*-scores = 0.56); and somewhat more marginally in the Given, Final condition (lower-left; Δ mean *z*-scores = 0.31). However, f0 values do not appear to be significantly different between tonal and non-tonal short vowels in the remaining two conditions.

The preservation of tonal contrasts in the Focused, Medial condition is expected: sentence-
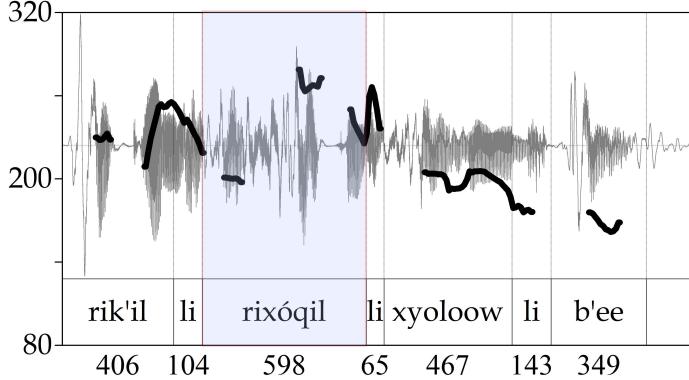
Figure 16: Declination and focus prominence in *Rik'il li* RIXÓQIL *li xyoloow li b'ee* 'I saw him in the street with HIS WIFE' (speaker 3 AP, 2017; Focused, Medial condition). Target word highlighted.

medial position should show the least influence from intonational contours, and focus should exaggerate the contrast between [H] tone and toneless vowels (or possibly [L]-tones) (section 2.5). The reduction and/or neutralization of tonal contrasts in final position may reflect the influence of utterance-final (L)H% rises, a possibility we return to in section 3.2. Otherwise, the same contextual effects observed for long vowels — declination across the sentence, with focus-related pitch raising — are also evident for the short vowel data in Fig. 17.

To confirm the broad trends observed in Figs. 15 and 17 we carried out a linear mixed effects analysis of mean f0 on stressed vowels using the LME4 package (Bates et al. 2020) in R. This model included simple fixed effects for TONE, V LENGTH, SENTENTIAL POSITION, and DISCOURSE STATUS, A V HEIGHT predictor (high vs. non-high) was included to control for the fact that high vowels are often associated with raised pitch (e.g. Sapir 1989; on Uspanteko, see Bennett et al. ms.a). The model also included all two-way and three-way interactions between TONE, DISCOURSE STATUS, SENTENTIAL POSITION and V LENGTH. Lastly, the model included random intercepts for ITEM and SPEAKER, which helps model differences in baseline pitch across participants. By-speaker random slopes for TONE were not included, as the limited amount of data available for some speakers led to convergence issues. Model parameters were estimated using the maximum likelihood procedure.

Step-down model criticism using the log-likelihood test with a threshold of $\alpha = 0.05$ led to the omission of most predictors from the final model. Simple predictors were always retained when participating in a higher-order interaction, and model simplification was biased toward the elmination of higher-order terms. A summary of the final model is given in Table 4; *p*-values were estimated using Satterthwaite's method with the LMERTEST package (Kuznetsova et al. 2020). The final model has fairly low collinearity between predictors ($\kappa = 7.86$; see Baayen 2008:222). (Mean f0 is highly correlated with maximum f0 in our vowel data ($r = 0.99$), and so our results are qualitatively similar if maximum f0 is taken as the dependent measure instead. Fitting the model over log(mean Hz) also produces nearly identical results.)

This statistical model confirms our initial observations: tonal contrasts are present for long and short vowels alike, but the f0 differences associated with tone are modest ($\approx$10Hz). While Figs. 15, 17 might imply that tonal contrasts are more phonetically distinct on long vowels, this inference was not supported by our statistical analysis. The TONE $\times$ V LENGTH interaction was

22
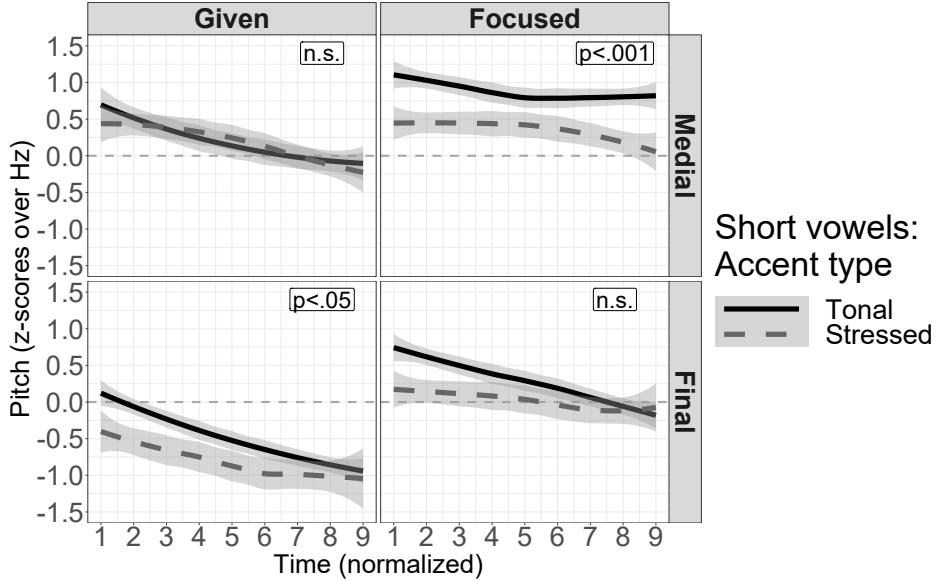
Figure 17: f0 trajectories (in z-scores over Hz) for short vowels grouped by discourse condition and sentential position (see Fig. 15 for other details).

Table 4: Final linear mixed-effects model for mean f0 in Hz

Predictor	β (in Hz)	SE(β)	$|t|$	$p <$
Intercept	181.7	14.2	12.95	.001
SENTENTIAL POSITION (FINAL)	-19.8	1.7	11.41	.001
DISCOURSE STATUS (FOCUSED)	9.8	1.7	5.83	.001
TONE (V́)	9.3	2.7	3.49	.001
SENTENTIAL POSITION (FINAL) × DISCOURSE STATUS (FOCUSED)	7.8	2.4	3.31	.001

eliminated during model comparison, which suggests that the amount of f0 raising associated with [H] tone is roughly comparable on long and short vowels in our data, all else being equal.

As expected, focus is associated with pitch raising, and/or overall expansion of the pitch space, and declination leads to pitch reduction in final position. Additionally, it appears that focus has a greater effect in sentence-final position than in sentence-medial position. This effect differs from the findings of Xu (1999), who reports that in Mandarin pitch range expansion associated with focus is attenuated in utterance-final position. We further explore the interaction of focus with sentence position, and its implications for the analysis of tone, in the following section.

3.2 Focus, phonological representation, and phonetic realization

In section 2.5 we speculated that focus could be a diagnostic for the phonological representation of 'toneless' vowels in Uspanteko. We saw in section 3 that the general effect of focus is to raise pitch and/or expand the overall pitch range. By hypothesis, the effect of focus should be dependent on the phonological specification of tone: vowels which are specified for tone should show more sensitivity to focus than vowels which are not. The presumption here is that focus operates at

a relatively abstract level, manipulating the phonetic realization of specified tones like [H] and [L], and leaving tonally unspecified regions relatively unaffected. (For related discussion, see Liberman & Pierrehumbert 1984, Pierrehumbert & Beckman 1988, Myers 1998, Remijsen & van Heuven 2005, Riad 2006, Ladd 2008a,b, Dilley 2010, Mixdorff et al. 2012, Remijsen et al. 2014, and references there.) On the other hand, if focus affects 'toneless' and tonal vowels equally in Uspanteko, this could provide support for a binary tonal analysis over a privative one (section 1.3).

We begin with the effect of focus prosody on short vowels (Fig. 17), and return to the analysis of long vowels below. Table 5 provides pairwise *t*-tests exploring the effect of focus prosody across all four combinations of tone (toneless ['V] vs. tonal ['V́]) and position (sentence-medial vs. sentence-final). For short vowels, the effect of focus closely tracks the tonal specifications posited in the privative analysis of Uspanteko tone. In medial position, focus has an effect on tonal vowels only—toneless vowels show no focus-related pitch raising whatsoever (i.e. the dashed lines for toneless vowels in the top row of Fig. 17 are about equal in height, regardless of focus). In final position (the bottom row of Fig. 17) both toneless and tonal vowels show pitch raising under focus.

We attribute this difference to the fact that lexically toneless vowels acquire an abstract pitch target—an intonational H% boundary tone—when occurring in sentence final position (e.g. Fig. 14 and Bennett et al. ms.a). Focus prosody may then affect the phonetic realization of the non-lexical, intonational boundary H% in much the same way that it affects the realization of the lexical [H] tone.[13] (See Elordieta 2007, 2008, Elordieta & Hualde 2014 for a parallel phenomenon in Basque.)

Table 5: The effect of focus prosody on short vowels. *p*-values reflect one-sided Welch *t*-test for unequal variances

Comparison	Difference (*z*-scores)	|*t*|	*p* <	Hypothesized tonal specification	Predicted by [H]∼∅ lexical tone contrast?
['V], medial	0.22	1.21	n.s.	∅	✓
['V], final	0.87	4.15	.001	H%	✓
['V́], medial	0.82	4.93	.001	[H]	✓
['V́], final	0.82	5.64	.001	[H] and/or H%	✓

Focus prosody suggests a rather different analysis of tonal contrasts for long vowels (Fig. 15). As Table 6 indicates, focus prosody raises f0 for *all* combinations of lexical tone and sentential position. The key condition involves toneless ['Vː] vowels in sentence-medial position. Under the privative analysis of Uspanteko tone, these vowels should carry neither a lexical tone nor an intonational boundary tone. And yet, focus prosody raises pitch on these vowels, consistent with the assumption that they *are* tonally specified, and not 'toneless' after all.

If focus prosody does indeed provide evidence for the phonological specification of tone, then we are faced with the possibility that tonal specifications in Uspanteko are different on phonemic

---

[13]It would be inaccurate to imply that all of the utterances in our study are realized with the dramatic final rise seen in Fig. 14. Many productions instead show a slight upward contour, and/or raised pitch relative to preceding vowels, in final position (see plots in section 1.1 for some illustrations). We assume that even these slight f0 movements are intonational in nature, rather than e.g. reflexes of stress (see Bennett et al. ms.a for a more detailed phonetic analysis of this claim).

Table 6: The effect of focus prosody on long vowels. *p*-values reflect one-sided Welch *t*-test for unequal variances

Comparison	Difference (*z*-scores)	|*t*|	*p* <	Hypothesized tonal specification	Predicted by [H]∼∅ lexical tone contrast?
['Vː], medial	0.66	5.49	.001	∅	×
['Vː], final	0.97	6.79	.001	H%	✓
['V́ː], medial	0.47	3.06	.01	[H]	✓
['V́ː], final	1.06	6.16	.001	[H] and/or H%	✓

long and short vowels. Short vowels may involve a privative [H]∼∅ contrast, while long vowels host a binary contrast of some sort, perhaps [H]∼[L]. A binary tone specification for long vowels may also help explain why long vowels show a robust tonal contrasts across all conditions in Fig. 15: long vowels simply have more tonal specifications, hence more tonal stability. We comment further on the relative *in*stability of tonal contrasts on short vowels in the next section.

The phonological analysis of Uspanteko tone in Bennett & Henderson (2013) assumes a privative [H]∼∅ contrast on the penultimate mora, for both long and short vowels alike (section 1.3). That analysis could be integrated with the present findings by assuming that long vowels must be specified for tone, and are assigned a default [L] tone in the absence of a lexical [H] tone (see Zec 1999, Zec & Zsiga 2010, Zsiga & Zec 2013 for a somewhat similar approach to Serbo-Croatian pitch accent). This is essentially the analysis of Campbell (1977), but amended to include a tonal component on non-final, accented short vowels. We speculate more on the exact phonetic specification of tonal categories in Uspanteko in section 3.5.

If 'toneless' long vowels are instead specified with an [L] tone, how does this lexical tone interact with the utterance-final H% boundary tone? Recall that the tonal contrast for long vowels is preserved in utterance-final position (Fig. 15). This might mean that the [L] tone resists overwriting by the phrasal H%. Alternatively, it may be that the [L] tone is replaced with a phrasal H% tone, but the phonetic realization of the phrasal H% is lower in f0 than a lexical [H], as reported for e.g. Japanese by Pierrehumbert & Beckman (1988). (See also Myers et al. 2018 on Luganda.)

3.3 The effect of givenness on f0

As evident in Fig. 17, tonal contrasts on short vowels are most robust in the Focused, Medial condition (Fig. 17, upper right). This is as predicted if (i) short vowels host a privative [H]∼∅ contrast, (ii) lexically toneless vowels do not acquire an intonational H% tone in sentence-medial position, and (iii) focus raises pitch on vowels specified with a lexical [H] tone or intonational H% tone (sections 2.5 and 3.2). The lack (or at least marginality) of tonal contrast in final position can be interpreted as the neutralization of the lexical [H]∼∅ contrast to a [H]∼H% contrast under the influence of a final H% boundary tone.

What remains to be explained is the lack of contrast in the Given, Medial condition (Fig. 17, upper left). Phonologically, this should correspond to a [H]∼∅ contrast. And yet the phonetic facts seem to indicate surface-level neutralization of f0 distinctions in this environment.

We suspect that the lack of phonetic contrast in the Given, Medial context may reflect the

relative shortness of vowels in discourse-given words. We provide a full analysis of vowel duration in the following section (3.4), but for the moment make some initial observations about duration which might help clarify the patterns of f0 neutralization seen in Figs. 15 and 17.

Phonemic short vowels are slightly shorter when discourse-given (73ms vs. 81ms under focus, $p < .01$ by one-sided Welch t-test for unequal variances). As emphasized by Xu & Sun (2002), it takes some time to phonetically implement a pitch rise, as the laryngeal articulations involved in raising f0 are relatively slow compared to those which lower f0 (see Myers et al. 2018 for a recent review). It may be that the phonetic shortening of stressed vowels leads to some degree of phonetic undershoot of f0 targets in the Given condition. Such undershoot would plausibly contribute to the phonetic neutralization of an underlying [H]$\sim\varnothing$ contrast in this context. Long vowels are comparatively unaffected by shortening—their mean duration in the Given condition is still 109ms—and so we should not expect to see the same degree of undershoot.

Beyond the effect of duration, it may simply be that f0 is compressed on Given items (Table 4), as in Japanese, German, and many other languages (e.g. Ladd 2008a, Féry & Kügler 2008, Féry & Ishihara 2009, Kügler & Féry 2017 and references there). These effects appear to be weaker in utterance-final position, if the significance of the tonal contrast on short vowels in the Given, Final condition is taken at face value.

3.4 Durational differences

Fig. 18 shows the distribution of vowel duration in our data, grouped by accent type and phonemic vowel length.
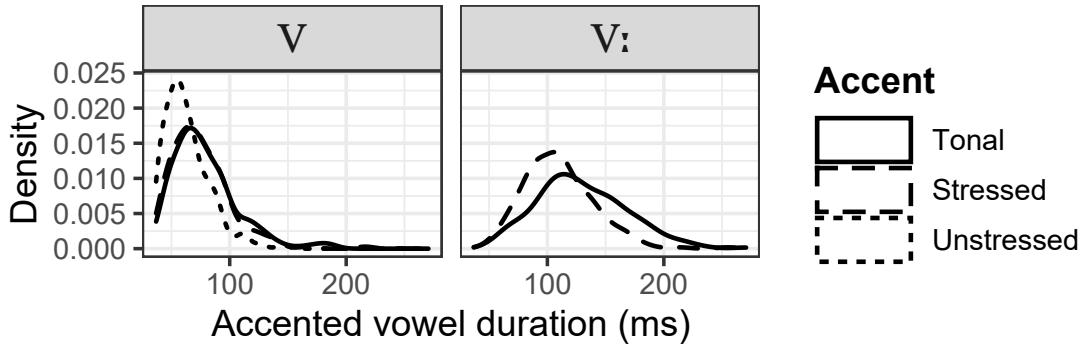

Figure 18: Vowel duration, grouped by phonemic length and accent type.

Despite visual appearances, a statistical analysis of vowel duration suggests that short vowels—but not long vowels—have greater duration when bearing [H] tone, once other factors are controlled for. We fit a linear mixed effects model over vowel duration with the LME4 package in R. This model included simple fixed effects for TONE, V LENGTH, STRESS, SENTENTIAL POSITION, DISCOURSE STATUS, and a control predictor of V HEIGHT (low vs. mid vs. high), since vowel height and duration are often inversely correlated (e.g. House & Fairbanks 1953, Lehiste 1970). We also included two-way interactions for TONE × V LENGTH, STRESS × SENTENTIAL POSITION, and STRESS × DISCOURSE STATUS, as well as random intercepts for SPEAKER and ITEM. Step-down model criticism using the log-likelihood test with a threshold of $\alpha = 0.05$ suggested that all predictors in the initial model should be retained, and so no model simplification was carried out. The

final model is reported in (Table 7), and has moderate collinearity between predictors ($\kappa = 10.76$; p-values were estimated using the LMERTEST package).

Table 7: Final linear mixed-effects model for vowel duration in ms

Predictor	β (in ms)	SE(β)	$|t|$	$p <$
Intercept	79	4.7	16.81	.001
V LENGTH (Vː)	51	4.0	12.86	.001
STRESS ('V)	12	3.3	3.43	.001
TONE (V́)	13	3.1	4.09	.001
TONE (V́) × V LENGTH (Vː)	-15	4.8	-3.21	.01
V HEIGHT (MID VS. LOW)	-34	3.1	11.04	.001
V HEIGHT (HIGH VS. LOW)	-3	3.1	0.84	n.s.
DISCOURSE STATUS (FOCUSED)	12	1.4	8.50	.001
DISCOURSE STATUS (FOCUSED) × STRESS ('V)	9	2.5	3.67	.001
SENTENTIAL POSITION (FINAL)	1	1.4	0.60	n.s.
SENTENTIAL POSITION (FINAL) × STRESS ('V)	6	2.5	2.27	.05

To summarize the results in Table 7, we can first see that long vowels tend to be about 50ms longer than short vowels. Stress is associated with moderate lengthening (about 12ms) on short vowels (there are no unstressed long vowels, so the effect of stress on long vowels is purely hypothetical). Tone is also associated with moderate lengthening (about 13ms), but only on short vowels: the significant negative TONE (V́) × V LENGTH (Vː) interaction indicates that long vowels do not similarly lengthen under tone. This replicates the results of Bennett et al. (ms.a), who report that short vowels are phonetically longer when bearing [H] tone in Uspanteko in wordlist data (Bennett et al. ms.a did not analyze the effect of tone on the duration of long vowels because of limitations in their data set). Tone-related lengthening does not phonetically neutralize the contrast between long and short vowels—phonemic long vowels are still substantially longer—but phonetic vowel duration does appear to serve as a correlate of [H] tone on short vowels (see also footnote 8).

Focus is associated with lengthening, particularly on stressed vowels. This is consistent with the typological observation that the phonetic effects of focus are most evident in phonologically prominent positions (e.g. DiCanio et al. 2018 and references there). We found a limited effect of utterance-final lengthening: stressed vowels are lengthened very slightly in final position.

Lastly, mid vowels in this data set are phonetically shorter than low vowels, but low and high vowels appear to be equally long. As we made no attempt whatsoever to balance vowel quality in our materials, this result may not be representative of the Uspanteko language as a whole. Further research on this question is needed.

3.5 The shape(s) of lexical tone

The results from focus prosody (section 3.2) suggest a binary tone opposition for long vowels in Uspanteko. But the phonological specification of these two tones remains to be settled—are they level tones, contour tones, or some combination of levels and contours (sections 1.3, 2.5)?

To investigate this issue, we carried out a quantitative analysis of the shape of f0 trajectories on accented vowels, using a combination of principal component analysis and *k*-means clustering. Space restrictions prevent us from laying out this method in full detail, but we try to provide an intuitive explanation of how the procedure works below. For additional details and examples of these techniques, see e.g. Johnson (2008), Mielke (2012), Dillon et al. (2013), Shosted et al. (2015), Gubian et al. (2015), Dockum (2017), Bennett et al. (2018a), Bennett (2019).

The smoothed f0 plots provided in Figs. 15 and 17 imply that the distinction between tonal and non-tonal vowels in Uspanteko is primarily expressed by differences in f0 level, and not f0 shape. But those plots provide only aggregated, summary information. As such, we might worry that they obscure some structured token-to-token variability in the shape of f0, and that such f0 shape differences might also serve to distinguish tonal and non-tonal vowels.

We used principal component analysis (e.g. PCA; Jolliffe 2002) to analyze how f0 shape varies across accented vowels. The output of PCA tells us how f0 contours typically differ from each other in our data set, in terms of level, shape, and any other parameters which might systematically vary across tokens. PCA over *z*-score normalized f0 with the PRCOMP() function in R produces 9 principal components (PCs), which correspond to 9 typical dimensions of variation in f0 shape in our data set.[14] As Fig. 19 shows, the first three PCs account for nearly 100% of the variation in our data. These PCs have straightforward interpretations: PC1, which accounts for most of the variance in the data, corresponds roughly to pitch level; PC2 corresponds roughly to pitch slope; and PC3 corresponds roughly to concavity.
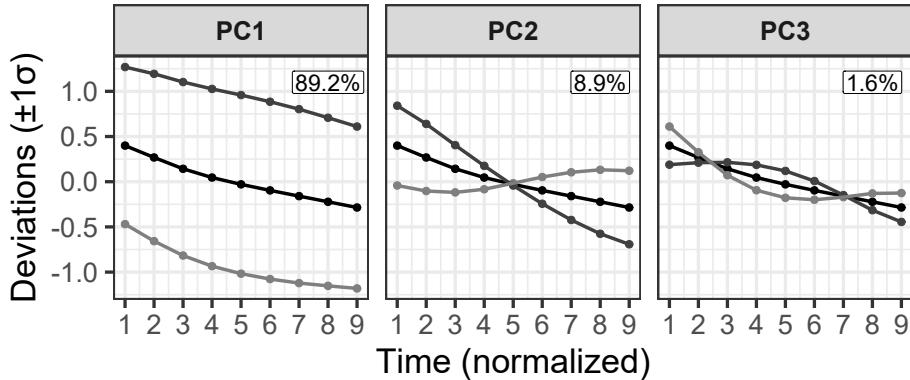

Figure 19: First three principal components. Mean f0 across the entire data set is represented by black midline, and upper/lower lines show changes in each PC value in positive and negative directions.

PC1 is not particularly informative about the tonal specifications of individual vowels, as variation in pitch height in our data is also strongly conditioned by declination and focus-related pitch raising (Table 4). Here, we're most interested in how f0 *shape* might distinguish between tonal and non-tonal vowels. To that end, we used *k*-means clustering to assess how many distinct, characteristic f0 shapes occur on accented vowels in our data. *K*-means clustering is an unsupervised

---

[14]PCA requires fully specified values at each measurement step for each vowel token. To fill in missing values we carried out data imputation using the function NA_INTERPOLATION() from the R package IMPUTETS (Moritz & Gatscha 2020) with Stineman interpolation. Other approaches to dealing with missing values, including simply omitting every incomplete contour from the input to PCA, yielded very similar results.

clustering method, which can be used to classify vowels into discrete categories on the basis of similar f0 shapes. The input to *k*-means clustering was the output of PCA, but with PC1 excluded, so that categorization would be carried out primarily on the basis of f0 shape rather than f0 height.

We ran the *k*-means algorithm multiple times, assuming that between 1 and 15 groups might be present in the data. Comparing models by means of the Bayesian Information Criterion, which is biased against finding too many groups, we found that a *k*-means model with four groups provided the best compromise between data fit and model complexity (i.e. the lowest BIC value). The four groups identified are shown in Fig. 20, represented with their characteristic pitch contours. It appears that accented vowels fall into four broad shape categories: flat f0, rising f0, slightly falling f0, and falling f0.
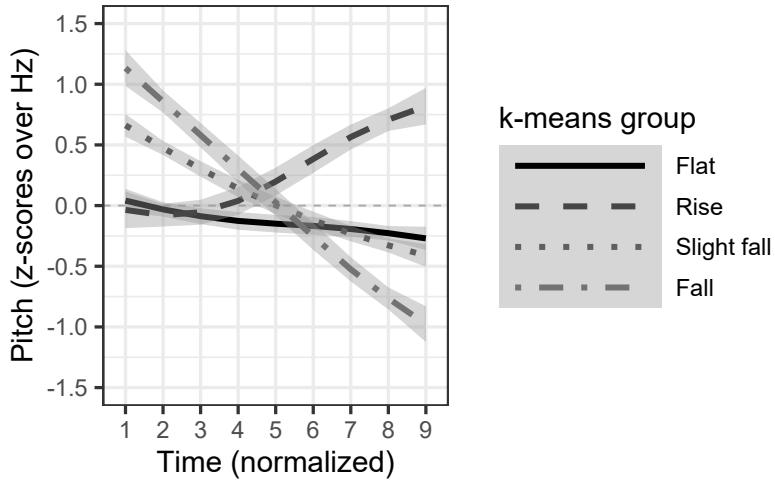

Figure 20: Tone shape categories identified by *k*-means clustering with *k* = 4.

If f0 shape were a distinguishing characteristic of tonal vowels, we should expect some of the shapes in Fig. 20 to be much more common for tonal vowels than for non-tonal vowels. This does not appear to be the case (Fig. 21). All four tone shapes occur on both tonal and non-tonal vowels alike. Still, flat and rising f0 shapes do seem somewhat more common for non-tonal vowels, and falling f0 shapes more common for tonal vowels. These patterns hold equally for short and long vowels, despite the phonetic evidence presented above that 'non-tonal' long vowels might actually be specified for a tonal target.

We assessed these patterns statistically by means of a multinomial mixed-effects analysis using the BRMS package (Bürkner 2018) in R, which fits a Bayesian regression. This model included simple fixed effects for TONE and V LENGTH, and their interaction. It also included random intercepts for SPEAKER and ITEM. The only significant effects in this model are a tendency for tonal vowels to be realized more frequently with f0 falls or slight falls, and less frequently with flat f0 (Table 8; significant effects are those whose 95% confidence interval excludes zero).

We interpret these results in the following way. Tonal vowels are specified for [H], and tone may fall somewhat after the attainment of that [H] target (e.g. Figs. 10, 11 and Bennett & Henderson 2013). All other variability in f0 shape is likely due to the influence of contextual factors which are distributed equally across tonal and non-tonal vowels (e.g. segmental context, phrasal
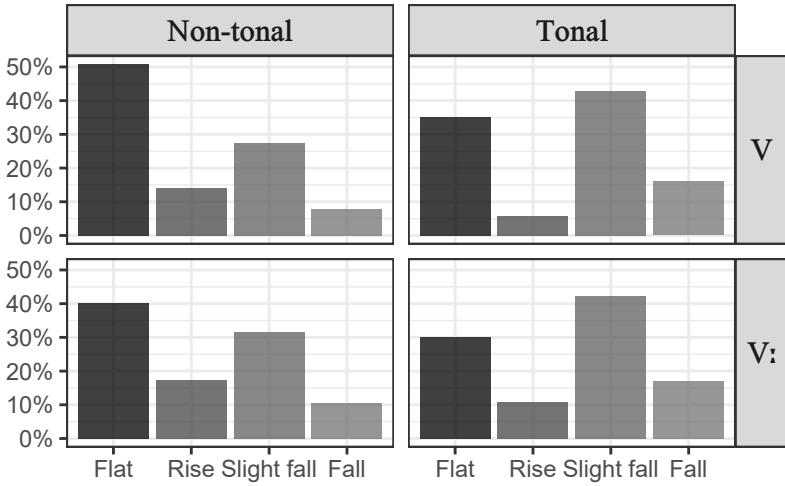
Figure 21: Proportion of different f0 shape categories across for different combinations of lexical tone and vowel length.

Table 8: All significant effects in multinomial mixed-effects model for f0 shape category. Reference level for intercept and TONE predictor is FLAT F0

Predictor	β	SE(β)	95% CI lower bound	95% CI upper bound
Intercept (rise)	-1.55	0.51	-2.65	-0.64
Intercept (fall)	-1.95	0.75	-3.38	-0.42
TONE (V́; likelihood of slight fall)	1.05	0.39	0.33	1.84
TONE (V́; likelihood of fall)	1.29	0.62	0.07	2.57

position, etc.).[15] These patterns seem most consistent with a simple [H] target for tonal vowels, which leaves room for the variability in f0 shape observed in Figs. 20, 21 (see Arvaniti 2016 for a similar approach to analyzing intonational variation, and Xu 1994, 1997, 1998, DiCanio 2014, Wang et al. 2020 for other cases of contextual variability in the shape of level tones).

If 'non-tonal' long vowels are really specified for tone in the surface phonetics (section 3.2), the simplest analysis would be to treat them as bearing a level [L] tone, possibly by default, and allow f0 shape to vary contextually in the same way that it does for the other accentual categories (section 3.2).

3.6 The magnitude of the tonal contrast

In section 3.1 we observed that vowels specified with an [H] tone are realized with higher f0 than toneless (or [L]-toned) vowels. However, the differences involved appear to be rather small, on the order of about ∼10Hz. In this section we consider some explanations for this striking fact.

The first, somewhat banal possibility is that the small effect of [H] tone on f0 reflects a problem

---

[15]The slightly higher likelihood of rising pitch on non-tonal vowels may indicate greater susceptibility to influence from final H% boundary tones in the absence of a lexical tone, though this apparent difference is not supported by the statistical analysis in Table 8.

with the design of our study. Every target item was produced in each of our four experimental contexts ({focused, given} × {medial, final}). Each target item was therefore repeated at least several times by each speaker. As a consequence, even our 'focused' items were in a sense discourse-given—they had often been previously spoken by the participant during the recording session, as part of a different target sentence. The repetition of target items may have conditioned further phonetic reduction of f0 and/or duration across conditions, obscuring the phonetic effect of [ʜ] tone on f0.[16]

Some evidence for this view comes from a comparison with the results of Bennett et al. (ms.a), who present a phonetic analysis of wordlist data in Uspanteko. That study found an effect size of about 50Hz for lexical tone in Uspanteko, after controlling for the effect of utterance-final ʜ% boundary tones on word-final syllables. An effect in the ∼50Hz range is similar to the amount of f0 raising found for [ʜ] tones in lexical pitch accent languages like Japanese, Goizueta Basque, and Serbo-Croatian (e.g. Beckman 1986, Pierrehumbert & Beckman 1988, Hualde et al. 2008, Zec & Zsiga 2010). Importantly, each target word in the wordlist was only produced once by each speaker, so repetition-related phonetic reduction could not have arisen in the same way.

We do not know whether the present study or the wordlist data in Bennett et al. (ms.a) is more representative of the typical phonetic realization of lexical tone in regular Uspanteko speech.[17] But at least in our question-answer materials, it seems fair to say that tonal distinctions in Uspanteko are weakly implemented and phonetically marginal.

There are several other respects in which tone can be considered 'marginal' in Uspanteko. In section 1.2 we observed that tone has an extremely low functional load in the language. There are hardly any minimal pairs for tone—probably fewer than 10 across the entire language—and speakers tend to recognize just one or two of these pairs as being words they would actually use in their own speech. Tone is mostly found on nouns, but tonal nouns often have alternative non-tonal forms which are also in wide use in the Uspanteko community (e.g. *xájab'* [ˈʃá.χaɓ] ∼ *xajab'* [ʃa.ˈχaɓ] 'shoe'). Tone is associated with certain morphological constructions (e.g. possession), but in all such cases tone co-occurs with segmental affixes that independently mark the relevant construction (e.g. *qapátz'an* [qa-ˈpátŝʔan] 'our sugar cane'). By itself, tone plays almost no role in signaling lexical or morphological distinctions in Uspanteko.

On short vowels, tone is also *phonetically* redundant. When an [ʜ] tone occurs on a penultimate short vowel, stress retracts from the final syllable so that [ʜ] tone and stress will coincide (e.g. *wíkij* [ˈwí.kiχ] 'my hatchet'). Hence words like *ab'aj* [a.ˈɓaχ] 'stone' (non-tonal) and *áb'aj* [ˈá.ɓaχ] 'your name' (tonal) are distinguished not only by the presence or absence of tone, but also by the position of stress (Fig. 4). Consequently, the phonological contrast between toneless and tonal short vowels is jointly cued by f0 (the primary correlate of tone) as well as duration, intensity, and vowel quality (the primary correlates of stress). In this context, it doesn't matter much if [ʜ] tone is clearly expressed by raised f0: the phonological difference between words like [a.ˈɓaχ] 'stone' and [ˈá.ɓaχ] 'your name' will be phonetically recoverable from the position of stress alone, even if the pitch raising associated with tone is only weakly implemented (Fig. 22).

---

[16]Another potential issue with our materials is that we made no attempt to control for intonational events occurring at the edges of units smaller than the utterance, such as the INTONATIONAL PHRASE or PHONOLOGICAL PHRASE. We simply do not know enough about the intonational system of Uspanteko to know how to control for such confounds. The lack of such controls may have influenced f0 on some phrase-medial items in ways that we do not yet understand.

[17]We have also recorded a large quantity of more naturalistic speech in Uspanteko, but the transcription and phonetic analysis of those recordings is ongoing.
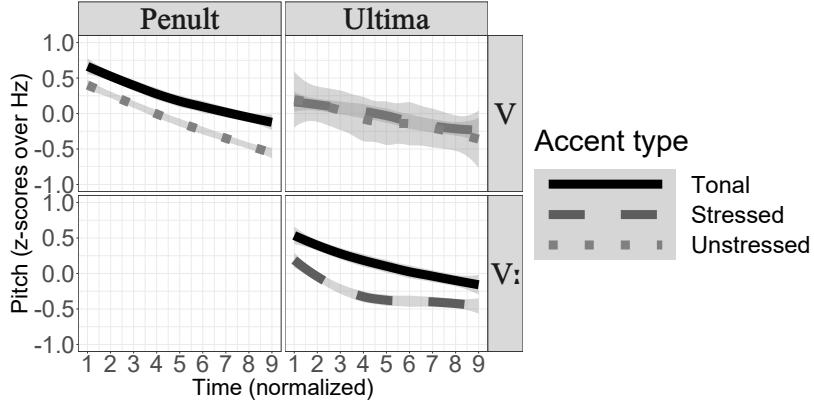
Figure 22: f0 trajectories (in *z*-scores over Hz) for all vowels, grouped by vowel length and position within the word.

These facts are potentially revelant for understanding the relatively small f0 differences associated with contrastive [ʜ] tone in our study. It has been argued that the functional importance of a phonological contrast can affect its phonetic realization: contrasts which are more important for distinguishing words and morphemes tend to be more robustly implemented in speech production (see e.g. Berinstein 1979, Lindblom 1986, Baese-Berk & Goldrick 2009, Nakai et al. 2012, Goldrick et al. 2013, Renwick 2014, Nadeu & Renwick 2016, Nelson & Wedel 2017, Hall et al. 2018, Bennett et al. 2018b and references there). The weak f0 distinctions associated with tone in our results may therefore reflect, in part, the low functional relevance of tonal contrasts in the language. This is especially true for short vowels, which have a particularly low functional load for tone, and which show phonetic neutralization of f0 contrasts in most conditions in our study (Fig. 17).

Two other phonetic observations are worth mentioning. First, the pitch ranges associated with each tonal category may be well-separated for some speakers of Uspanteko, but more overlapped for other speakers. Fig. 23 shows mean f0 values at vowel midpoint for each speaker, grouped by tonal specification and vowel length. This plot is an oversimplification of our data, as it collapses over many other factors which condition f0 variation (e.g. focus and sentential position). For that reason, the *p*-values reported in this plot should not be taken too literally, especially since each comparison is based on a very limited amount of data for each speaker, and we do not correct *p*-values for multiple comparisons. (Most importantly, the *non*-significant *p*-values in Fig. 23 should not be interpreted as indicating much of anything.)

With these caveats in mind, the plots in Fig. 23 might suggest that speakers differ as to how sharply and reliably they use f0 to distinguish the two tonal categories in Uspanteko. If so, the f0 differences associated with tone in our study may have been affected by the specific population of speakers who participated, if they happen to be speakers who produce tonal contrasts in a phonetically attenuated way. (We suspect that speakers who produce weaker tonal contrasts may be more K'iche'-dominant than speakers who produce the tonal contrast more clearly. Can Pixabaj (2007:49) also mentions possible age-grading in the production of tone. We have not yet explored these possibilities in any systematic way.) In future work, we hope to investigate inter-speaker phonetic variation in the Uspanteko tone system using larger data sets better suited for that purpose.

32
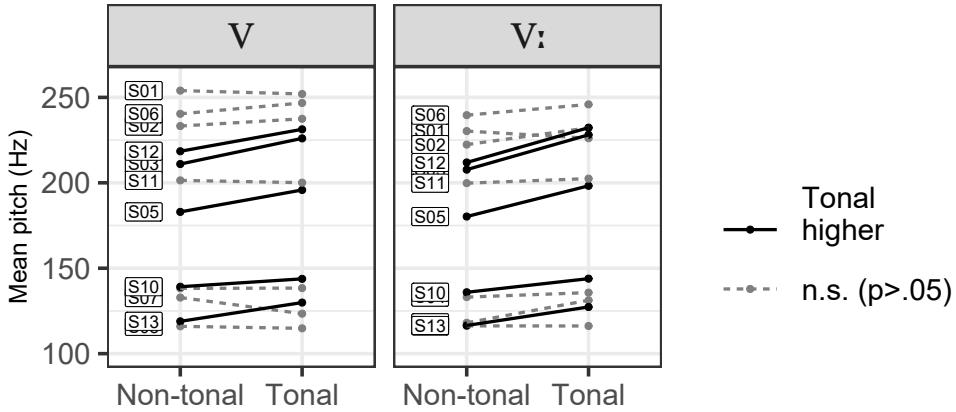
Figure 23: Individual variation in the implementation of f0 for tonal contrasts. Statistical significance assessed with one-sided Welch *t*-test for unequal variances.

A second phonetic observation is that tonal contrasts in Uspanteko may not be exclusively expressed by f0 distinctions. In the current study, we found that phonemic short vowels are phonetically longer when bearing an [H] tone (section 3.4). The use of duration to mark tonal contrasts in Uspanteko provides an interesting counterpoint to lexical pitch accent languages like Goizueta Basque and Japanese: in these languages, f0 robustly distinguishes tonal vowels, but duration is not a cue to lexical tone (e.g. Beckman 1986, Hualde et al. 2008; Fintoft 1965, and Kelly 2015 report similar results for Norwegian tonal accent, though see also Fintoft 1970). The multidimensionality of tonal contrast in Uspanteko may thus help explain why f0 distinctions on short vowels are so marginal in the current study: they are redundantly cued by duration, and possibly other phonetic properties as well. (Impressionistically, tonal short vowels in Uspanteko also differ from non-tonal vowels in voice quality—they seem to be produced with 'pressed' or 'tense' voice. Phonetic investigation of this issue is underway.)

To summarize, the amount of f0 raising associated with [H] tone in our study is smaller than might have been expected on the basis of phonetic patterns in other lexical pitch accent languages, or the wordlist data reported in Bennett et al. (ms.a). We have speculated here on a number of practical and linguistic factors which may have jointly led to the compression of f0 distinctions in our data. We leave further exploration of these issues for future work.

4 Discussion

While there is no doubt that Uspanteko is a language with lexical tone, pinning down the phonological and phonetic properties of the tone system has proven a thorny task. Much of the difficulty involved stems from the extensive variability—phonetic, phonological, lexical, and sociolinguistic—that characterizes the tone system, and its interaction with intonation.

We have argued here that the tone system of Uspanteko is best understood as a privative [H]~∅ contrast on short vowels, and either a privative [H]~∅ contrast or binary [H]~[L] contrast on long vowels. Our evidence has been primarily phonetic, rather than phonological, and our arguments depend on the assumption that structured variability in the phonetic signal can provide a window on abstract phonological specifications. This approach follows a long line of research on intonation in the 'Autosegmental-Metrical' tradition, from which we draw our inspiration (e.g. Pierrehumbert &

Beckman 1988, Ladd 2008a, and other work cited above). At a minimum, we hope that this work has helped clarify the practical issues that arise in the study of restricted tone systems, particularly those in which tone has a low functional load, and in which intonation runs roughshod over tones assigned at the word level.

References

Adell, Eric James. 2019. *The phonetics, phonology, and morphology of Chajul Ixil (Mayan)*. Austin, TX: University of Texas, Austin dissertation.

Aissen, Judith. 1992. Topic and focus in Mayan. *Language* 68(1). 43–80.

Aissen, Judith. 2017. Information structure in Mayan. In Judith Aissen, Nora England & Roberto Zavala Maldonado (eds.), *The Mayan languages*, 293–324. New York: Routledge. doi:10.4324/9781315192345.

Aissen, Judith, Nora England & Roberto Zavala Maldonado (eds.). 2017. *The Mayan languages*. New York: Routledge. doi:10.4324/9781315192345.

Arvaniti, Amalia. 2016. Analytical decisions in intonation research and the role of representations: Lessons from Romani. *Laboratory Phonology* 7(1). 1–43.

Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Baese-Berk, Melissa & Matthew Goldrick. 2009. Mechanisms of interaction in speech production. *Language and cognitive processes* 24(4). 527–554. doi:10.1080/01690960802299378.

Baird, Brandon. 2014a. *An acoustic analysis of contrastive focus marking in Spanish-K'ichee' (Mayan) bilingual intonation*. Austin, TX: University of Texas, Austin dissertation.

Baird, Brandon. 2014b. Acoustic correlates of stress in K'ichee': a preliminary analysis. In Lauren Eby Clemens, Robert Henderson & Pedro Mateo Pedro (eds.), *Proceedings of Formal Approaches to Mayan Linguistics (FAMLi) II*, 21–34. Cambridge, MA: MIT Working Papers in Linguistics.

Baird, Brandon. 2018. Syntactic and prosodic contrastive focus marking in K'ichee'. *International Journal of American Linguistics* 84(3). 295–325.

Bardiaux, Alice & Piet Mertens. 2014. Normalisation des contours intonatifs et étude de la variation régionale en français. *Cahiers de Linguistique Française* 31. 273–284.

Barnes, Jonathan, Alejna Brugos, Nanette Veilleux & Stefanie Shattuck-Hufnagel. 2011. Voiceless intervals and perceptual completion in F_0 contours: Evidence from scaling pperception in American English. In *Proceedings of the XVIIth international congress of phonetic sciences*, 108–111. The International Phonetic Association.

Barnes, Jonathan, Nanette Veilleux, Alejna Brugos & Stefanie Shattuck-Hufnagel. 2012. Tonal Center of Gravity: A global approach to tonal implementation in a level-based intonational phonology. *Laboratory Phonology* 3(2). 337–383.

Bates, Douglas, Martin Maechler & Ben Bolker. 2020. *lme4: Linear mixed-effects models using S4 classes*. Version 1.1-23, retrieved from http://CRAN.R-project.org/package=lme4.

Beckman, Mary. 1986. *Stress and non-stress accent*. Dordrecht: Foris.

Bennett, Ryan. 2016. Mayan phonology. *Language and Linguistics Compass* 10(10). 469–514. doi:10.1111/lnc3.12148.

Bennett, Ryan. 2019. La tensión vocálica en el kaqchikel de Sololá, Guatemala: un estudio preliminar. In Esther Herrera Zendejas & Verónica Reyes Taboada (eds.), *Fonología segmental: pro-*

cesos e interacciones, 13–46. Mexico City: Colégio de México. `http://libros.colmex.mx/ficha/?b=2848`.

Bennett, Ryan, Jessica Coon & Robert Henderson. 2016. Introduction to Mayan linguistics. *Language and Linguistics Compass* 10(10). 1–14. doi:10.1111/lnc3.12159.

Bennett, Ryan & Robert Henderson. 2013. Accent in Uspanteko. *Natural Language & Linguistic Theory* 31(3). 589–645. doi:10.1007/S11049-013-9196-6.

Bennett, Ryan, Robert Henderson & Megan Harvey. 2019. The interaction of tone and intonation in Uspanteko. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*, 452–456. Canberra, Australia: Australasian Speech Science and Technology Association Inc. `https://assta.org/proceedings/ICPhS2019/papers/ICPhS_501.pdf`.

Bennett, Ryan, Robert Henderson & Megan Harvey. ms.a. Uspanteko. *Journal of the International Phonetic Association* .

Bennett, Ryan, Robert Henderson & Megan Harvey. ms.b. Vowel deletion as phonologically-controlled gestural overlap in uspanteko .

Bennett, Ryan, Máire Ní Chiosáin, Jaye Padgett & Grant McGuire. 2018a. An ultrasound study of Connemara Irish palatalization and velarization. *Journal of the International Phonetic Association* 48(3). 261–304. doi:10.1017/S0025100317000494.

Bennett, Ryan, Kevin Tang & Juan Ajsivinac Sian. 2018b. Statistical and acoustic effects on the perception of stop consonants in Kaqchikel (Mayan). *Laboratory Phonology* 9(1). 9. doi:10.5334/labphon.100.

Berinstein, Ava E. 1979. *A cross-linguistic study on the perception and production of stress*. University of California Los Angeles MA thesis. Available online at http://escholarship.org/uc/item/0t0699hc.

Berinstein, Ava E. 1991. The role of intonation in K'ekchi Mayan discourse. In Cynthia McLemore (ed.), *Texas linguistic forum 32*, 1–19. Austin, TX: Department of Linguistics and the Center for Cognitive Science, The University of Texas at Austin.

Boersma, Paul & David Weenink. 2016. Praat: doing phonetics by computer (version 6.0.23). Computer program. Retrieved from http://www.praat.org/.

Borise, Lena. to appear. Tone and intonation in the languages of the Caucasus. In Maria Polinsky (ed.), *Handbook of the languages of the Caucasus*, Oxford, UK: Oxford University Press.

Brody, Jill. 1984. Some problems with the concept of basic word order. *Linguistics* 22(5). 711–736. doi:10.1515/ling.1984.22.5.711.

Bruce, Gösta. 1977. *Swedish word accents in sentence perspective*. Liber Läromedel/Gleerup.

Bürkner, Paul-Christian. 2018. Advanced Bayesian multilevel modeling with the R package brms. *The R Journal* 10(1). 395–411. doi:10.32614/RJ-2018-017.

Campbell, Eric. 2016. Tone and inflection in Zenzontepec Chatino. In Enrique Palancar & Jean Léo Léonard (eds.), *Tone and inflection: New facts and new perspectives*, 141–162. Berlin: De Gruyter.

Campbell, Lyle. 1977. *Quichean linguistic prehistory*, vol. 81 University of California Publications in Linguistics. Berkeley, CA: University of California Press.

Campbell, Lyle. 2017. Mayan history and comparison. In Judith Aissen, Nora England & Roberto Zavala Maldonado (eds.), *The Mayan languages*, 43–61. New York: Routledge. doi:10.4324/9781315192345.

Can Pixabaj, Telma. 2007. *Gramática descriptiva Uspanteka*. Antigua, Guatemala: Oxlajuuj Keej

Maya' Ajtz'iib' (OKMA).

Can Pixabaj, Telma & Nora England. 2011. Nominal topic and focus in K'ichee'. In Rodrigo Gutiérrez-Bravo, Line Mikkelsen & Eric Potsdam (eds.), *Representing language: essays in honor of Judith Aissen*, 15–30. Santa Cruz, CA: Linguistics Research Center. Available online at https://escholarship.org/uc/item/0vf4s9tk#page-15.

Clopper, Cynthia & Judith Tonhauser. 2013. The prosody of focus in Paraguayan Guaraní. *International Journal of American Linguistics* 79(2). 219–251.

Coon, Jessica. 2016. Mayan morphosyntax. *Language and Linguistics Compass* 10(10). 515–550. doi:10.1111/lnc3.12149.

Dayley, Jon. 1985. *Tzutujil grammar*, vol. 107 University of California Publications in Linguistics. Berkeley, CA: University of California Press.

De Looze, Céline & Stéphane Rauzy. 2009. Automatic detection and prediction of topic changes through automatic detection of register variations and pause duration. In *Proceedings of INTERSPEECH 2009*, 2919–2922. ISCA Archive. Available online at http://www.isca-speech.org/archive/interspeech_2009/i09_2919.html.

DiCanio, Christian. 2014. Triqui tonal coarticulation and contrast preservation in tonal phonology. In Ryan Bennett, Rikker Dockum, Emily Gasser, Dolly Goldenberg, Ryan Kasak & Patrick Patterson (eds.), *Proceedings of the workshop on the sound systems of Mexico and Central America*, Yale University. `http://ling.yale.edu/sites/default/files/files/ssmca-proceedings/papers/DiCanio-2014-Triqui-SSMCA.pdf`.

DiCanio, Christian, Joshua Benn & Rey Castillo García. 2018. The phonetics of information structure in Yoloxóchitl Mixtec. *Journal of Phonetics* 68. 50–68.

DiCanio, Christian & Ryan Bennett. 2021. Prosody in Mesoamerican languages. In Carlos Gussenhoven & Yiya Chen (eds.), *The Oxford handbook of language prosody*, chap. 28. Oxford, UK: Oxford University Press.

Dilley, Laura. 2010. Pitch range variation in English tonal contrasts: continuous or categorical? *Phonetica* 67(1-2). 63–81.

Dillon, Brian, Ewan Dunbar & William Idsardi. 2013. A single-stage approach to learning phonological categories: Insights from Inuktitut. *Cognitive Science* 37(2). 344–377.

Dockum, Ricker. 2017. Computational modeling of tone in language documentation: citation tones vs. running speech in Chindwin Khamti. In Julia Nee, Margaret Cychosz, Dmetri Hayes, Tyler Lau & Emily Remirez (eds.), *Berkeley Linguistic Society (BLS) 43*, 43–73. Berkeley, CA: University of California, Berkeley Linguistics Society.

Dockum, Rikker & Claire Bowern. 2019. Swadesh lists are not long enough: Drawing phonological generalizations from limited data. *Language Documentation and Description* 16. 35–54. `https://www.elpublishing.org/itempage/168`.

Du Bois, John W. 1987. The discourse basis of ergativity. *Language* 64(4). 805–855.

Elordieta, Gorka. 2007. A constraint-based analysis of the intonational realization of focus in Northern Bizkaian Basque. In Tomas Riad & Carlos Gussenhoven (eds.), *Tones and tunes: typological studies in word and sentence prosody*, vol. 1, 199–232. Berlin: Mouton De Gruyter.

Elordieta, Gorka. 2008. Constraints on intonational prominence of focalized constituents. In Chungmin Lee, Matthew Gordon & Daniel Büring (eds.), *Topic and focus: Cross-linguistic perspectives on meaning and intonation*, 1–22. Dordrecht: Springer. doi:10.1007/978-1-4020-4796-1_1.

Elordieta, Gorka & José Hualde. 2014. Intonation in Basque. In Sun-Ah Jun (ed.), *Prosodic*

typology II: the phonology of intonation and phrasing, 405–463. Oxford, UK: Oxford University Press.

England, Nora. 1991. Changes in basic word order in Mayan languages. *International Journal of American Linguistics* 57(4). 446–486.

England, Nora & Brandon Baird. 2017. Phonology and phonetics. In Judith Aissen, Nora England & Roberto Zavala Maldonado (eds.), *The Mayan languages*, 175–200. New York: Routledge. doi:10.4324/9781315192345.

England, Nora & Laura Martin. 2003. Issues in the comparative argument structure analysis in Mayan narratives. In John W. Du Bois, Lorraine E. Kumpf & William J. Ashby (eds.), *Preferred argument structure: Grammar as architecture for function*, 131–157. Amsterdam: John Benjamins.

Evanini, Keelan, Catherine Lai & Klaus Zechner. 2011. The importance of optimal parameter setting for pitch extraction. In *Proceedings of meetings on acoustics*, vol. 11 1, Acoustical Society of America. Available online at http://scitation.aip.org/content/asa/journal/poma/11/1/10.1121/1.3609833.

Faytak, Matthew & Alan C.L. Yu. 2011. A typological study of the interaction between level tones and duration. In *Proceedings of the XVIIth international congress of phonetic sciences*, 659–662. Available online at http://www.icphs2011.hk/resources/OnlineProceedings/RegularSession/Faytak/Faytak.pdf.

Féry, Caroline. 2013. Focus as prosodic alignment. *Natural Language & Linguistic Theory* 31(3). 683–734.

Féry, Caroline & Shinichiro Ishihara. 2009. How focus and givenness shape prosody. In Malte Zimmermann & Caroline Féry (eds.), *Information structure: Theoretical, typological, and experimental perspectives*, 36–63. Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780199570959.001.0001.

Féry, Caroline & Shinichiro Ishihara (eds.). 2016. *The Oxford handbook of information structure*. Oxford, UK: Oxford University Press.

Féry, Caroline & Frank Kügler. 2008. Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics* 36(4). 680–703.

Fintoft, Knut. 1965. Some remarks on word accents. *Phonetica* 13(4). 201–226.

Fintoft, Knut. 1970. *Acoustical analysis and perception of tonemes in some Norwegian dialects*. Oslo: Universitetsforlaget.

Fox, James. 1978. *Proto-Mayan accent, morpheme structure conditions, and velar innovation*: University of Chicago dissertation.

Goldrick, Matthew, Charlotte Vaughn & Amanda Murphy. 2013. The effects of lexical neighbors on stop consonant articulation. *Journal of the Acoustical Society of America* 134(2). EL172–EL177. doi:10.1121/1.4812821.

Gordon, Matthew. 2006. *Syllable weight: phonetics, phonology, typology*. New York: Routledge.

Grimes, James. 1971. A reclassification of the Quichean and Kekchian (Mayan) languages. *International Journal of American Linguistics* 37(1). 15–19.

Grimes, James. 1972. *The phonological history of the Quichean languages*. Carbondale, IL: University Museum, Southern Illinois University. Available online at http://www.ailla.utexas.org/search/resource.html?r_id=6534.

Gubian, Michele, Francisco Torreira & Lou Boves. 2015. Using functional data analysis for investigating multidimensional dynamic phonetic contrasts. *Journal of Phonetics* 49. 16–40.

Gussenhoven, Carlos. 1983. Focus, mode and the nucleus. *Journal of linguistics* 19(2). 377–417.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge, UK: Cambridge University Press.

Gussenhoven, Carlos & Renske Teeuw. 2008. A moraic and a syllabic H-tone in Yucatec Maya. In Esther Herrera Zendejas & Pedro Martín Butragueño (eds.), *Fonología instrumental: patrones fónicos y variacion*, 49–71. Mexico, D.F.: El Colegio de México.

Hall, Kathleen Currie, Elizabeth Hume, T. Florian Jaeger & Andrew Wedel. 2018. The role of predictability in shaping phonological patterns. *Linguistics Vanguard* 4(s2).

Hanson, Helen. 2009. Effects of obstruent consonants on fundamental frequency at vowel onset in English. *The Journal of the Acoustical Society of America* 125(1). 425–441.

Hartmann, Katharina. 2008. Focus and tone. *Acta Linguistica Hungarica* 55(3-4). 415–426.

Hedding, Andrew. to appear. Asymmetric pitch raising under focus in San Martín Peras Mixtec. In *Proceedings of the Workshop on Structure and Constituency in the Languages of the Americas (WSCLA) 24*, vol. 24 University of British Columbia Working Papers in Linguistics, Vancouver, Canada: University of British Columbia, Department of Linguistics.

Henderson, Robert. 2012. Morphological alternations at the intonational phrase edge. *Natural Language & Linguistic Theory* 30(3). 741–787.

Herrera Zendejas, Esther. 2014. *Mapa fónico de las lenguas mexicanas: formas sonoras 1 y 2*. Mexico, D.F.: El Colégio de México 2nd edn.

Himmelmann, Nikolaus. 2006. Prosody in language documentation. In Jost Gippert, Nikolaus Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation*, 163–181. Berlin: Mouton de Gruyter.

Himmelmann, Nikolaus & D. Robert Ladd. 2008. Prosodic description: An introduction for fieldworkers. *Language documentation and conservation* 2(2). 244–274. Available online at http://nflrc.hawaii.edu/ldc/December2008/.

House, Arthur & Grant Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America* 25(1). 105–113.

Hualde, José Ignacio, Oihana Lujanbio & Francisco Torreira. 2008. Lexical tone and stress in goizueta basque. *Journal of the International Phonetic Association* 38(1). 1–24.

Hyman, Larry. 2006. Word-prosodic typology. *Phonology* 23(2). 225–257.

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

Hyman, Larry. 2014. How to study a tone language, with exemplification from Oku (Grassfields Bantu, Cameroon). *Language Documentation and Conservation* 8. 525–562. Available online at https://scholarspace.manoa.hawaii.edu/handle/10125/24624.

Johnson, Keith. 2008. *Quantitative methods in linguistics*. Malden, MA: Blackwell.

Jolliffe, Ian T. 2002. *Principal component analysis*. New York: Springer.

Jun, Sun-Ah. 1998. The accentual phrase in the Korean prosodic hierarchy. *Phonology* 15(2). 189–226.

Jun, Sun-Ah & Janet Fletcher. 2014. Methodology of studying intonation: From data collection to data analysis. In Sun-Ah Jun (ed.), *Prosodic typology II: the phonology of intonation and phrasing*, 493–519. Oxford, UK: Oxford University Press.

Kaufman, Terrence. 1976. *Proyecto de alfabetos y ortografías para escribir las lenguas mayances*. Antigua, Guatemala: Proyecto Lingüístico Francisco Marroquín.

Kawahara, Shigeto. 2015. The phonology of Japanese accent. In Haruo Kubozono (ed.), *The handbook of Japanese phonetics and phonology*, chap. 11, 4454–492. Berlin: Mouton de Gruyter.

Kelly, Niamh Eileen. 2015. *An experimental approach to the production and perception of Norwegian tonal accent*. Austin, TX: University of Texas, Austin dissertation.

Kingston, John. 2005. The phonetics of Athabaskan tonogenesis. In Sharon Hargus & Keren Rice (eds.), *Athabaskan prosody*, 137–184. Amsterdam: John Benjamins. doi:10.1075/cilt.269.09kin.

Kirby, James & D. Robert Ladd. 2016. Effects of obstruent voicing on vowel F0: Evidence from "true voicing" languages. *The Journal of the Acoustical Society of America* 140(4). 2400–2411.

Kügler, Frank & Sasha Calhoun. 2021. Prosodic encoding of information structure: A typological perspective. In Carlos Gussenhoven & Yiya Chen (eds.), *The Oxford handbook of language prosody*, Oxford, UK: Oxford University Press.

Kügler, Frank & Caroline Féry. 2017. Postfocal downstep in German. *Language and speech* 60(2). 260–288.

Kügler, Frank & Susanne Genzel. 2012. On the prosodic expression of pragmatic prominence: The case of pitch register lowering in Akan. *Language and speech* 55(3). 331–359.

Kügler, Frank & Stavros Skopeteas. 2006. Interaction of lexical tone and information structure in Yucatec Maya. In *Proceedings of the Second International Symposium on Tonal Aspects of Languages (TAL-2)*, 380–388. Available online at http://www.isca-speech.org/archive/tal_2006/tal6_077.html.

Kügler, Frank & Stavros Skopeteas. 2007. On the universality of prosodic reflexes of contrast: the case of Yucatec Maya. In *Proceedings of the XVIth International Congress of Phonetic Sciences*, vol. 2, 1025–1028. Saarbrüken, Germany: Universität des Saarlandes. Available online at http://www.icphs2007.de/conference/Papers/1188/.

Kügler, Frank, Stavros Skopeteas & Elisabeth Verhoeven. 2007. Encoding information structure in Yucatec Maya: On the interplay of prosody and syntax. *Interdisciplinary studies on information structure* 8. 187–208. Available online at http://opus.kobv.de/ubp/volltexte/2008/1946/.

Kuznetsova, Alexandra, Per Bruun Brockhoff & Sofie Pødenphant Jensen Rune Haubo Bojesen Christensen. 2020. *lmerTest: Tests in linear mixed effects models*. Version 3.1-2, retrieved from http://CRAN.R-project.org/package=lmerTest.

Ladd, D. Robert. 2008a. *Intonational phonology*. Cambridge, UK: Cambridge University Press 2nd edn. doi:10.1017/cbo9780511808814.

Ladd, D. Robert. 2008b. Review of Sun Ah-Jun (ed.) (2005). prosodic typology: the phonology of intonation and phrasing. Oxford: Oxford University Press. ix + 462. *Phonology* 25(2). 372–376.

Larsen, Thomas. 1988. *Manifestations of ergativity in Quiché grammar*: University of California, Berkeley dissertation.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.

Liberman, Mark & Janet Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. In Mark Aronoff & Richard Oehrle (eds.), *Language sound structure: studies in phonology presented to Morris Halle by his teacher and students*, 157–233. Cambridge, MA: MIT Press.

Lindblom, Björn. 1986. Phonetic universals in vowel systems. In John Ohala & Jeri Jaeger (eds.), *Experimental phonology*, 13–44. Orlando: Academic Press.

McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner & Morgan Sonderegger. 2017. Montreal Forced Aligner: trainable text-speech alignment using Kaldi. In *Proceedings of the 18th conference of the international speech communication association*, 498–502.

Mielke, Jeff. 2012. A phonetically-based metric of sound similarity. *Lingua* 122. 145–163. doi: 10.1016/j.lingua.2011.04.006.

Mixdorff, Hansjörg, Jennifer Cole & Stefanie Shattuck-Hufnagel. 2012. Prosodic similarity – evidence from an imitation study. In Qiuwu Ma, Hongwei Ding & Daniel Hirst (eds.), *Proceedings of the speech prosody 2012 conference*, 571–574. Shanghai, China: Speech Prosody. Available online at https://www.isca-speech.org/archive/sp2012/.

Morén, Bruce & Elizabeth Zsiga. 2006. The lexical and post-lexical phonology of Thai tones. *Natural Language & Linguistic Theory* 24(1). 113–178.

Moritz, Steffen & Sebastian Gatscha. 2020. *imputeTS: time series missing value imputation in R*. Version 3.1, retrieved from http://CRAN.R-project.org/package=imputeTS.

Myers, Scott. 1998. Surface underspecification of tone in Chichewa. *Phonology* 15(3). 367–391.

Myers, Scott & Benjamin Hansen. 2007. The origin of vowel length neutralization in final position: Evidence from Finnish speakers. *Natural Language & Linguistic Theory* 25(1). 157–193.

Myers, Scott, Elisabeth Selkirk & Yelena Fainleib. 2018. Phonetic implementation of high-tone spans in luganda. *Laboratory Phonology* 9(1). 1–22. doi:https://doi.org/10.5334/labphon.101.

Nadeu, Marianna & Margaret Renwick. 2016. Variation in the lexical distribution and implementation of phonetically similar phonemes in Catalan. *Journal of Phonetics* 58. 22–47.

Nakai, Satsuki, Alice Turk, Kari Suomi, Sonia Granlund, Riikka Ylitalo & Sari Kunnari. 2012. Quantity constraints on the temporal implementation of phrasal prosody in Northern Finnish. *Journal of Phonetics* 40(6). 796–807.

Nelson, Noah Richard & Andrew Wedel. 2017. The phonetic specificity of competition: Contrastive hyperarticulation of voice onset time in conversational English. *Journal of Phonetics* 64. 51–70. doi:10.1016/j.wocn.2017.01.008.

Nettle, Daniel. 1995. Segmental inventory size, word length, and communicative efficiency. *Linguistics* 33(2). 359–367.

Pierrehumbert, Janet & Mary Beckman. 1988. *Japanese tone structure*. Cambridge, MA: MIT Press.

Pike, Kenneth. 1946. Phonemic pitch in Maya. *International Journal of American Linguistics* 12(2). 82–88.

Pike, Kenneth. 1948. *Tone languages: a technique for determining the number and type of pitch contrasts in a language, with studies in tonemic substitution and fusion*. Ann Arbor, MI: University of Michigan Press.

Polian, Gilles. 2013. *Gramática del Tseltal de Oxchuc*. Mexico, D.F.: Centro de Investigaciones y Estudios Superiores en Antropología Social.

R Development Core Team. 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. Version 4.0.2, retrieved from http://www.R-project.org/.

Remijsen, Bert. 2013. Tonal alignment is contrastive in falling contours in Dinka. *Language* 89(2). 297–327.

Remijsen, Bert, Farienne Martis & Ronald Severing. 2014. The marked accentuation pattern of Curaçao Papiamentu. In Sun-Ah Jun (ed.), *Prosodic typology II: the phonology of intonation and phrasing*, 302–323. Oxford, UK: Oxford University Press.

Remijsen, Bert & Vincent van Heuven. 2005. Stress, tone and discourse prominence in the Curaçao dialect of Papiamentu. *Phonology* 22(2). 205–235.

Renwick, Margaret. 2014. *The phonetics and phonology of contrast: The case of the Romanian*

vowel system. Berlin: Walter de Gruyter. doi:10.1515/9783110362770.

Riad, Tomas. 2006. Scandinavian accent typology. *Sprachtypologie und Universalienforschung (STUF)* 59(1). 36–55.

Richards, Michael. 2003. *Atlas lingüístico de Guatemala*. Guatemala City: Instituto de Lingüístico y Educación de la Universidad Rafael Landívar.

Roettger, Timo. 2017. *Tonal placement in Tashlhiyt: How an intonation system accommodates to adverse phonological environments*. Berlin: Language Science Press.

Sapir, Shimon. 1989. The intrinsic pitch of vowels: Theoretical, physiological, and clinical considerations. *Journal of Voice* 3(1). 44–51. doi:10.1016/S0892-1997(89)80121-3.

Shosted, Ryan, Marissa Barlaz & Di Wu. 2015. Modeling Iu Mien tone with eigen-pitch representations. In The Scottish Consortium for ICPhS 2015 (ed.), *Proceedings of the 18th International Congress of the Phonetic Sciences*, 1–5. Glasgow: University of Glasgow. Available online at https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0728.pdf.

Snider, Keith. 2014. On establishing underlying tonal contrast. *Language Documentation & Conservation* 8. 707–737.

Sobrino Gómez, Carlos Martín. 2018. *El sistema fonológico protoyucatecano: reconstrucción y desarrollo*. Ciudad de México: El Colegio de México dissertation.

Stoll, Otto. 1884. *Zur ethnographie der republik Guatemala*. Zürich: Orell Füssli.

Stoll, Otto. 1887. *Die sprache der Ixil-Indianer: Ein Beitrag zur ethnologie und linguistik der Maya-völker. Nebst einem Anhang: Wortverzeichnisse aus dem nordwestlichen Guatemala*. Leipzig: FA Brockhaus.

Stoll, Otto. 1888. *Die Maya-Sprachen der Pokom-Gruppe*. Wien: A. Hölder.

Stoll, Otto. 1896. *Die sprache der K'e'kchi-Indianer. Nebst einem Anhang: Die Uspanteca*. Leipzig: K.F. Köhler's Antiquarium.

Strunk, Jan, Frank Seifart, Swintha Danielsen, Iren Hartmann, Brigitte Pakendorf, Søren Wichmann, Alena Witzlack-Makarevich & Balthasar Bickel. 2020. Determinants of phonetic word duration in ten language documentation corpora: Word frequency, complexity, position, and part of speech. *Language Documentation & Conservation* 14. 423–461.

Sugiyama, Yukiko. 2011. *The production and perception of Japanese pitch accent*. Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Tang, Kevin & Ryan Bennett. 2018. Contextual predictability influences word and morpheme duration in a morphologically complex language (Kaqchikel Mayan). *Journal of the Acoustical Society of America* 144(2). 997–1017. doi:10.1121/1.5046095.

Tang, Kevin & Ryan Bennett. 2019. Unite and conquer: Bootstrapping forced alignment tools for closely-related minority languages (Mayan). In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019*, 1719–1723. Canberra, Australia: Australasian Speech Science and Technology Association Inc. `https://assta.org/proceedings/ICPhS2019/papers/ICPhS_1768.pdf`.

Us Maldonado, Juan Antonio. 2010. *Gramática normativa Uspanteka*. San Miguel Uspantán: Comunidad Lingüística Uspanteka.

Us Maldonado, Juan Antonio. no date(a). *Chololtziij: Diccionario Uspanteko*. San Miguel Uspantán: Comunidad Lingüística Uspanteka.

Us Maldonado, Juan Antonio. no date(b). *Monografía Uspanteka: Aspectos culturales,*

lingüísticos y sociales de la Comunidad Lingüística Uspanteka. San Miguel Uspantán: Comunidad Lingüística Uspanteka.

Velleman, Leah Bridges. 2014a. *Focus and movement in a variety of K'ichee'.* Austin, TX: University of Texas, Austin dissertation.

Velleman, Leah Bridges. 2014b. On optional focus movement in K'ichee'. In Lauren Eby Clemens, Robert Henderson & Pedro Mateo Pedro (eds.), *Proceedings of Formal Approaches to Mayan Linguistics (FAMLi) 2*, 107–118. Cambridge, MA: MIT Working Papers in Linguistics.

Verhoeven, Elisabeth & Stavros Skopeteas. 2015. Licensing focus constructions in Yucatec Maya. *International Journal of American Linguistics* 81(1). 1–40.

Vicente Méndez, Miguel Angel. 2007. *Diccionario bilingüe uspanteko-español.* Antigua, Guatemala: Oxlajuuj Keej Maya' Ajtz'iib' (OKMA).

Wagner, Michael. to appear. Prosodic focus. In Daniel Gutzmann, Lisa Matthewson, Cécile Meier, Hotze Rullmann, Thomas Ede Zimmerman & Dina Voloshina (eds.), *The Wiley Blackwell companion to semantics*, Wiley Blackwell.

Wang, Lei, Carlos Gussenhoven & Jie Liang. 2020. How to pronounce a low tone: A lesson from Kaifeng Mandarin. *Journal of the International Phonetic Association* 50. 199–219.

Wickham, Hadley. 2009. *ggplot2: Elegant graphics for data analysis.* New York: Springer. http://had.co.nz/ggplot2/book.

Xu, Yi. 1994. Production and perception of coarticulated tones. *The Journal of the Acoustical Society of America* 95(4). 2240–2253.

Xu, Yi. 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics* 25(1). 61–83.

Xu, Yi. 1998. Consistency of tone-syllable alignment across different syllable structures and speaking rates. *Phonetica* 55(4). 179–203.

Xu, Yi. 1999. Effects of tone and focus on the formation and alignment of f0 contours. *Journal of phonetics* 27(1). 55–105.

Xu, Yi & Xuejing Sun. 2002. Maximum speed of pitch change and how it may relate to speech. *The Journal of the Acoustical Society of America* 111(3). 1399–1413.

Yu, Kristine. 2014. The experimental state of mind in elicitation: illustrations from tonal fieldwork. *Language Documentation and Conservation* 8. 738–777.

Zec, Draga. 1994. *Sonority constraints on prosodic structure.* New York: Garland.

Zec, Draga. 1999. Footed tones and tonal feet: rhythmic constituency in a pitch-accent language. *Phonology* 16(2). 225–264.

Zec, Draga & Elizabeth Zsiga. 2010. Interaction of tone and stress in Standard Serbian. In Wayles Browne, Adam Cooper, Alison Fisher, Esra Kesici & Nikola Predolac (eds.), *Formal approaches to Slavic linguistics (FASL) 18*, 536–555. Ann Arbor, MI: Michigan Slavic Publications.

Zhang, Jie. 2001. *The effects of duration and sonority on contour tone distribution*: University of California, Los Angeles dissertation.

Zhang, Jingwei. 2019. A comparison of tone normalization methods for language variation research. In *Proceedings of the 32nd Pacific Asia Conference on Language, Information and Computation (PACLIC-32)*, Hong Kong: Association for Computational Linguistics. https://www.aclweb.org/anthology/Y18-1095.

Zsiga, Elizabeth & Draga Zec. 2013. Contextual evidence for the representation of pitch accents in Standard Serbian. *Language and speech* 56(1). 69–104.