

The phonetics and phonology of lexical prosody in San Jerónimo Acazulco Otomi

Rory Turnbull

Department of Linguistics, The Ohio State University

turnbull@ling.osu.edu

San Jerónimo Acazulco Otomi (SJAO) is an underdescribed and endangered Oto-Manguean language spoken in central Mexico. This paper provides an analysis of the phonology of tonal contrasts in SJAO and the phonetics of their realization based on pitch pattern data derived from audio recordings of citation forms of SJAO words. Each SJAO lexical word has one and only one tonal sequence – either /H/ or /HL/. This sequence is underlyingly associated with one syllable in the word. Other syllables are not specified for tone, and their phonetic realization is predictable depending on their position relative to the tonal syllable. A phonetic analysis revealed that underlyingly-tonal syllables are phonetically distinct from non-tonal syllables: those with /H/ are produced with greater vocal effort (measured by spectral tilt), and those with /HL/ are longer, louder, and bear a higher f_0 (fundamental frequency), compared with non-tonal syllables. This analysis differs from previous accounts of lexical prosody in other Otomi varieties, which have either described a three-way system of high, low, and rising tones contrasting on every stem syllable, or a system where one syllable per word is assigned a stress-like ‘accent’. This difference from previous analyses suggests that there is a third possible characterization of lexical prosody for Otomi, which is appropriate for SJAO and potentially other understudied varieties.

1 Introduction

This paper is concerned with the lexical prosody of San Jerónimo Acazulco Otomi (SJAO), an Oto-Manguean language spoken in central Mexico. The starting point for this study is the fact that, like other varieties of Otomi, SJAO features distinct lexical items that are identical in segmental content, but differ in suprasegmentals. The examples in (1) show strings of segments, the lexical identity of which depends on the suprasegmental realization. How these contrasts are represented phonologically and realized phonetically is the central question of this study.

- (1) a. /ʔwini/
 ‘thorn’ or ‘to feed’
 b. /t^heni/
 ‘red thing’ or ‘to grab’
 c. /kihi/
 ‘to bring’ or ‘color, dye’

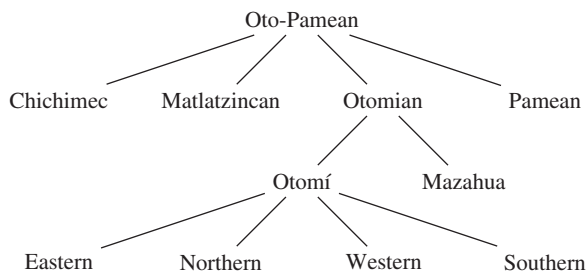


Figure 1 Family relationships among the Oto-Pamean languages.

- d. /k^hini/
'beard' or 'dough'
- e. /tɔʔjo/
'bone' or 'griddle'
- f. /paʃi/
'to travel downriver' or 'grass, trash'
- g. /t'one/
'horn (of animal)' or 'gourd, *jícara*'
- h. /ʔɔhɔ/
'to sleep' or 'yes'

Previous work on other varieties of Otomi has claimed either that these varieties have three tones – high, low, and rising – which can appear freely on almost all syllables (e.g. Sinclair & Pike 1948) or that they have an ‘accent’ system where one syllable per word is accented, and that the realization of f0 follows from regular rules relating to accentedness (e.g. Leon & Swadesh 1949). The goal of the present study is to determine, using phonological and phonetic evidence, whether either of these accounts are appropriate for lexical prosody in SJAQ. This paper proposes an analysis for SJAQ which in fact differs from this previous work on other varieties of Otomi. I argue that in SJAQ each word has one and only one tonal sequence – either high or falling – associated with it. The tonal sequence is underlyingly associated with one syllable in the word and the phonetic realization of the other syllables is fully predictable.

1.1 San Jerónimo Acapulco Otomi

Otomi is one of the two languages in the Otomian branch of the Oto-Pamean subdivision of the Oto-Manguean language family, and is spoken primarily in central Mexico.¹ However, due to widespread dialectal variation and mutual unintelligibility (see Gómez Rendón 2008: Chapter 8, for an overview), Otomi is considered by some (e.g. Palancar 2006, Suárez 1983) to be a diasystem of several distinct languages, or a multidimensional dialect continuum (Lastra 2010). Palancar (2013a) identified four main Otomi dialects or varieties, classified as northern, southern, western, and eastern. These relationships are depicted in Figure 1, with SJAQ being a southern dialect (Hernández-Green 2015). An alternative three-way classification was proposed by Lastra (2001), which groups the southern and eastern varieties together in a single group.

Figure 2 shows a map of central Mexico, with relevant towns and settlements indicated. The eastern dialect is centered approximately in the area around Texcatepec, the northern

¹ The name Otomi is a Nahuatlism (from *otomitl*, possibly related to *totomitl* ‘bird-hunter’ (Jiménez Moreno 1936)), but because of dialectal differences no single autonym has gained widespread use – see Wright Carr (2005) and Lastra (2006). The SJAQ name for the language, and the people, is *yühü* /jühü/.

Table 1 Consonant phoneme inventory of San Jerónimo Acazolco Otomi.

	Labial	Coronal	Palatal	Velar	Glottal
Stop	p ^h p b	t' t ^h t d		k' k ^h k g	ʔ
Affricate		ts' ts ^h ts	tʃ' tʃ		
Fricative		s z	ʃ		h
Nasal	m̥ m	n̥ n			
Tap		r			
Glide	w	l	j		

**Figure 2** Map of central Mexico with state boundaries. Location of San Jerónimo Acazolco is indicated with a cross.

dialect around Querétaro, and the western dialect around Toluca. Varieties from each of these dialect areas have received both descriptive and theoretical attention in the linguistic literature, although the northern varieties are the best described, largely due to the relative vitality of the language in that area. SJAO is exclusively spoken in the village of San Jerónimo Acazolco, in the municipality of Ocoyoacac in Estado de México (indicated with a cross on [Figure 2](#)). The village population is around 4000, although only a few elderly residents are fluent in SJAO, estimated at between 100 and 200 in number. The domain of use of SJAO is limited and shrinking, and most if not all speakers are Spanish-dominant. A related variety of Otomi is spoken in the neighboring village of San Pedro Atlapulco by, at most, 12 individuals (Valle Canales 2008).

Documentation of SJAO is sparse. Lastra's (2001) book contains a short narrative text and some rough grammar details. More recently, a flurry of interest has led to research on aspects of SJAO spatial semantics (Boeg Thomsen & Pharaoh Hansen 2015), linguistic anthropology (Pharaoh Hansen 2012, Pharaoh Hansen et al. 2016), and segmental phonology (Volhardt 2014).² The most comprehensive study to date is Hernández-Green's (2015) Ph.D. dissertation on SJAO verbal morphosyntax (see also Hernández-Green 2011, 2012).

Like other Oto-Manguean languages, SJAO features a rich inventory of consonant phonemes, listed in [Table 1](#) (compare Volhardt 2014, Hernández-Green 2015). Stops and affricates feature a four-way contrast between ejective, aspirated, unaspirated, and voiced, and

² In addition, non-linguistic anthropological studies of the Otomi people of San Jerónimo Acazolco have also been conducted (Figueroa 2012, 2015).

Table 2 Vowel phonemes of San Jerónimo Acazulco Otomi.

	Front	Central	Back
Close	i ĩ	ɨ	u ũ
Mid-close	e	ə	o
Mid-open	ɛ ě		ɔ ɔ̃
Open		a ǣ	

fricatives and nasals feature a two-way contrast between voiced and voiceless. There are gaps in these series, which may simply be due to a rare phoneme not being attested in the present data; therefore, [Table 1](#) is presented as a non-exhaustive list of the consonant phonemes of SJAO. [Table 2](#) lists the nine oral and five nasal vowel phonemes of SJAO. Syllables in native roots are of the form (C)(C)CV(j). Recent Spanish loans often feature codas other than /j/ (e.g. /jor.to.mu/ ← *mayordomo* ‘steward’) or syllables with no onset (e.g. /i.sku.la/ ← *escuela* ‘school’).

1.2 Otomi lexical prosody

All described varieties of Otomi have been observed to feature suprasegmental lexical distinctions, similar to the minimal pairs in (1) above. Given the diversity within the Otomian languages – many varieties are not mutually intelligible – it is not surprising that the suprasegmental distinctions have been described and analyzed differently by different scholars. However, even within the same variety, there are competing analyses. The accounts of Otomi lexical prosody can be broadly categorized into two groups: those advocating tone, and those advocating accent.

1.2.1 Tone in Otomi

In classical non-linear phonology (Leben 1973, Goldsmith 1976), tones are commonly analyzed as autosegments, which are phonological elements not inherently associated with segmental structure (Leben 2011).³ The autosegmental approach to tonal phenomena is assumed in this paper – namely, that words have an underlying form consisting of both segments and autosegments. Through the application of various conditions and rules, the autosegments are associated with the segments, resulting in a surface form with a specific phonetic realization. This approach is used here as a vehicle for description, rather than as a claim about phonological cognition. In fact, the majority of work on Otomi prosody either predates the development of this approach or is not couched in a formal framework.

The tonal description of Otomi is typified by Sinclair & Pike’s (1948) account of Mezquital Otomi, of the northern dialect region. Three phonological tones are posited – high, low, and rising – each of which can appear in any position in a word. Each syllable bears a tone, and there are no restrictions on what tones can follow one another. Sinclair and Pike (1948) reported that all three tones occurred with nearly all vowel phonemes,⁴ and that the tones are not subject to any restrictions with regards their linear ordering within lexemes.

Sinclair & Pike (1948) made passing reference to some phrasal phenomena, such that phrases tend to end with a high tone (suggesting, in modern parlance, a high boundary tone at phrase edges), and some lexical sandhi phenomena, noting that certain affixes can alter the immediately preceding lexical tone. However, their description is brief and does not provide many details.

³ See also Ladd (2014: Chapter 1) for discussion of the theoretical distinction between features and autosegments, and the intellectual history of the concepts.

⁴ The only exception was that the co-occurrence of /i/ with the low tone was not attested; however, Sinclair & Pike (1948) noted that they were working from a relatively small set of data. This fact makes the omission likely to be coincidental rather than of linguistic significance.

Sinclair & Pike (1948) also provided the example sentences in (2), which show ‘all possible sequence combinations in two adjacent syllables’ (p. 93).⁵

- (2) a. rà déhé dícihě dè gá zàbí
 ‘the water which we drink is from a pool’
 b. rà zàbí šà bíxà yăbó
 ‘the pool is very distant’

Indeed, ignoring word boundaries, the two sentences in (2) above show all nine possible combinations of two adjacent tones. However, it is likely to be the case that there are tone sandhi or intonational processes influencing the tonal patterns at the phrasal level. A closer examination of the individual bisyllables in the example sentences and citation forms provided by Sinclair & Pike (1948) reveals a striking pattern in the syllable tonotactics: the vast majority of bisyllabic words (89%) end with a high tone, and very few with low (9%) or rising (2%) tones. This pattern suggests that the tonal melodies observed are not necessarily as free as Sinclair & Pike (1948) suggested.

Sinclair & Pike’s (1948) tonal analysis has become the received view in modern work on other varieties of Otomi, with most researchers positing three tones with limited or no restrictions on placement. Examples include the descriptions of the eastern variety of Highland (Sierra) Otomi (Blight & Pike 1976, Voigtlander & Echegoyen 1985); the southern varieties of Ixtenco Otomi (Lastra 1997), Tilapa Otomi (Palancar 2012), and San Pedro Atlapulco Otomi (Valle Canales 2008); the western varieties of Temoayan Otomi (Andrews 1949, 1993) and San Andrés Cuexcontitlán Otomi (Lastra 1989); and the northern varieties of Mezquital Otomi (Hess 1962, Wallis 1968, Cruz, Torquemada & Crawford 2004), Santiago Mezquititlán Otomi (Hekking & Andrés de Jesús 1984, Hekking 1995), and San Idelfonso Tultepec Otomi (Palancar 2009). Other Oto-Pamean languages have been described in similar ways (see Arellanes et al. 2011 for review). The few studies of SJAO which have mentioned the prosodic system have followed Sinclair and Pike’s (1948) analysis of three tones: low, high, and rising, with no explicit constraints on placement (Hernández-Green 2011, 2012, 2015; Pharaoh Hansen, Turnbull & Boeg Thomsen 2011).

Nevertheless, the Sinclair & Pike (1948) analysis has required some modification to adequately model tonal phenomena in some Otomi varieties. A common modification is stipulating that non-stem morphemes such as prefixes and clitics have restrictions on what tones they can bear. Similarly, it has been claimed that for some varieties, tone is only contrastive on the first syllable of a stem (e.g. Palancar 2009, Hernández-Green 2015). Much of this work, however, has been concerned with morphosyntactic rather than phonological or phonetic questions (e.g. Hess 1962; Andrews 1993; Lastra 1996; Palancar 2004a, b, 2011, 2012, 2013b; Hernández-Green 2011, 2012, 2015) and consequently there has been little critical examination of the combinatorial possibilities of tones within words, or the phonetic properties of said tones.

Similarly, issues of tonal sandhi and phrase or edge tones have been largely untouched by this literature. A notable exception is Wallis’s (1968) study, which identified several sandhi and ‘allotonic’ processes which modify the phonetic realization of tones as a function of speech rate, intonation, and word-level stress. Due to the inclusion of both tone and stress in her account of Otomi lexical prosody, and the blending of word- and phrase-level prosodic units, Wallis’s (1968) comprehensive analysis of Mezquital Otomi is difficult to reformulate in terms of contemporary autosegmental and intonational phonology. Similar phenomena have not been studied in other Otomi varieties.

⁵ Examples are provided exactly using the transcription in the source paper. In their orthography, the high tone is represented with an acute accent (´), the low tone with a grave accent (`), and the rising tone with a caron (ˇ).

An alternative view of Otomi tones was provided by Bernard (1973, 1974), who claimed that only two tones – high and low – are needed to adequately describe (Mezquital) Otomi. This analysis relies on the assumption that aspects of both stress and vowel length operate independently of tone, and that all three have an influence on the final pitch contour of a given word. In this analysis, tone, vowel length, and stress are thus all unpredictable and must be lexically specified; the tonal system is therefore alleged to be simpler than previous accounts, due to there being only two tone types. The system as a whole, however, is potentially more complex due to the requirement that all three of tone, vowel length, and stress must be indicated in the lexicon. In designing a practical orthography for Otomi, Bernard noted that ‘[tone] markings are not needed by the native speakers’ (Bernard & Salinas Pedraza 1976, quoted by Bartholomew 1979: 94), who regarded them as a ‘pesky nuisance’ (Bernard 1980: 136). This fact could suggest that tone is not highly salient to speakers of the language, or that they are not easily able to consciously access tonal information, although speakers having a poor metalinguistic knowledge of prosodic features is not uncommon in tone languages (Bent 2004).⁶ Bernard’s account of Otomi tones has not received a great deal of attention in the literature, and has not been adopted in any descriptive accounts of other Otomi varieties.

1.2.2 Accent in Otomi

The accent hypothesis is represented by Leon & Swadesh’s (1949) study, also of Mezquital Otomi. Tone was dispensed with, and replaced by ‘accent’. It is worth quoting the particulars of their analysis at length:

Our analysis recognizes a partly free pitch-stress accent and also automatic features of tonality. Monosyllables may be either stressed or unstressed, but unstressed ones never stand at the end of the phrase. Disyllables may have either initial or final stress. Trisyllabic and longer words have automatic accent, always on their initial syllable. All accented syllables are high-pitched by comparison with certain unaccented syllables but need not be the highest in the phrase. An unaccented word-final syllable occurring before an accent in the same phrase is low-pitched; thus, *ra* and *n̄te* in *ra pr̄sidente ʔm̄nda*. An unaccented non-final syllable occurring after the accent, as in the second and third syllable of *pr̄sidente* (borrowed from Spanish with shift of accent), is normally at the same pitch as the accented syllable. An unaccented phrase-final syllable is normally as high or higher than the accented syllable; thus, *n̄te* in the simple phrase *ra pr̄sidente* or *nda* in the longer phrase *ra pr̄sidente ʔm̄nda*. (Leon & Swadesh 1949: 100–101)

Leon & Swadesh’s (1949) approach takes advantage of distributional properties of tone that they had noted in Mezquital Otomi. The ‘accent’, which they signify with an acute accent, appears only once in any word. The accented syllable is realized with a higher f₀ than preceding unaccented syllables. The rising contour on a single vowel, Sinclair & Pike’s (1948) rising tone, is described as a ‘geminatē’ vowel, the first portion of which is accented. The rules predict that the second half should be just as high or higher in pitch than the first – thus the percept of rising tonality. This account fits well with a phonetic detail noted in both studies: that the vowels with rising tone have a longer duration than the other vowels. Leon & Swadesh’s (1949) and Sinclair & Pike’s (1948) analyses of four Otomi words are contrasted in Table 3. In this table, and below, Leon & Swadesh’s (1949) accent is marked with the primary stress diacritic, to avoid confusion with Sinclair & Pike’s (1948) high tone.

Similar analyses for Mezquital Otomi were put forward by Leon (1963) and Bernard (1966), with the motivation that the accentual hypothesis is simpler – a speaker need only remember which syllable is accented in a given word, instead of memorizing which tone each

⁶ Indeed, in written English, vowel markings are generally not needed by native speakers (mxst sxntxncxs rxr stll cxmprhxnsxblx xftxr rpxlxcng thx vxwxls), but this fact does not reflect any truths about the representation of vowels in English phonology.

Table 3 Comparison of Sinclair & Pike's (1948; S&P) tonal analysis and Leon & Swadesh's (1949; L&S) accent analysis for four words.

Gloss	S&P	L&S
'horse'	fá.ní	¹ fá.ni
'hole'	ʔò.kí	ʔo. ¹ ki
'wood'	zǎ	¹ za:
'earth'	hà.í	¹ ha.i

syllable bears. Bernard (1966: 15) hoped that the analysis 'will force us to relinquish the notion that Otomi is a tone language at all'.⁷

Unfortunately, like the tonal hypothesis, the accentual hypothesis is scant on phonological details. The treatment of the rising pitch contour is confusing: despite 'earth' having the same pitch contour as 'hole' (both are low-high by the tonal analysis), the accent is initial in the former and final in the latter. It is also not clear whether the word for 'earth' is considered to have one or two syllables; there could be a potential contrast between /ha.i/ and /ha.'i/, but examples like the latter form are unattested and this issue is not addressed by Leon & Swadesh (1949).

The notion of 'accent' was left undefined, and it is unclear to what extent the accent is thought to be a type of tone or a manifestation of an English- or Spanish-like stress system. Like many stress systems, only one main accent is permitted per prosodic word, yet Leon & Swadesh (1949) noted none of the phonetic properties that usually accompany stress: the accented syllable did not sound obviously or consistently longer or louder than the unaccented syllables. 'In Otomi we found no notable length differences and the ratio of loudness seemed not always as great as in other languages we knew' (Leon & Swadesh 1949: 103). Some languages, such as Swedish and Japanese, are considered to have 'pitch accent' systems (see Beckman 1986 and van der Hulst 2011 for overviews), and a sympathetic reading of Leon & Swadesh (1949) yields a formulation of Otomi tone superficially similar to Japanese pitch accent (Uwano 1999, Kubozono 2011). Note, however, that the term 'accentual' is used in this paper to refer specifically to Leon & Swadesh's (1949) proposal for Otomi prosody, and not to any broader conceptions of 'pitch accent' (or other forms of accent) in the phonological literature at large.

Of the two accounts, the tonal hypothesis has gained more traction in the literature, and there are no purely descriptive accounts of Otomi that embrace the accent hypothesis. Nevertheless, neither account has received a great deal of formal phonological or phonetic attention, leaving the descriptions vague and difficult to evaluate. The goal of the present study is to determine which, if either, of these accounts is appropriate for lexical prosody in SJAQ, drawing on both phonological and phonetic evidence.

2 Pitch contours in SJAQ

As Otomi lexical prosody has previously been described in terms of pitch and tone, it makes sense to begin our investigation of SJAQ suprasegmentals by looking at fundamental frequency (f_0). This section outlines the methods used to annotate a corpus of spoken recordings of SJAQ words, and explains why neither of the previous accounts is appropriate for SJAQ. An alternative and novel analysis is then presented, where the attested pitch contours are analyzed

⁷ Bernard later changed his view and claimed that Otomi does have tone; see the last paragraph of Section 1.2.1.

in terms of one of two contrastive pitch target sequences being associated with one syllable in a word.

2.1 Materials and method

In the course of fieldwork on SJAO undertaken in July 2010, a corpus of 584 tokens of 266 words spoken in isolation by an elderly male native speaker was collected. The speaker is regarded in the community as a good speaker of the language, and uses it in his daily life in interactions with others. The productions were collected over the course of several days during elicitation sessions aimed at vocabulary documentation. The method of elicitation was wholly oral, with the consultant being asked in Spanish e.g. *¿Cómo se dice ‘perro’?* ‘How do you say “dog”?’ This format of questioning elicited nouns and verbs in their bare, uninflected form, with no prefixes or other affixes. Minimal pairs were not asked for in direct succession, in order to decrease the possibility of artificial contrast. The recordings were made in a relatively quiet environment, either with a head-mounted microphone (approximately 85% of the recordings) or an omnidirectional microphone.

The vast majority of words in the corpus are native, monomorphemic nouns. Only 15 words were not nouns, namely verbs, numbers, and discourse particles. Like other varieties of Otomi, SJAO does not have adjectives (Palancar 2006). Of the nouns, ten were proper nouns, eight of which were Spanish personal names. Twenty-three of the words in the corpus can be reliably identified as loanwords, 19 from Spanish and four from Nahuatl.⁸ A number of the Spanish loans were clearly borrowed during the colonial era; for example, /ʔaʃo/ ‘garlic’ reflects the 16th-century pronunciation of Spanish *ajo* as /aʃo/ (in contrast with present-day Spanish /axo/). The large number of Spanish loans relative to Nahuatl loans observed here has also been noted for other varieties of Otomi (Lastra 1998, Hekking & Bakker 2009), and is all the more surprising given the extensive history of pre-Columbian Nahuatl–Otomi contact (e.g. Harvey 1972, Wright Carr 2008). Instead of loans, it appears that many Nahuatl words have been calqued into Otomi (Wright Carr 2010), and that this trend has been ongoing since at least the colonial era (Bartholomew 2000). Like other languages, Otomi has been observed to make relatively innovative use of loanwords, often altering their semantic content substantially from that of the source language (Lastra 1998, Hernández-Green, Palancar & Hernández-Gómez 2011). Taken together, these facts suggest that, despite the fact that some of the words in the present corpus have been identified as loanwords, this fact is not a cause for concern with regard to the representativeness of the corpus as a sample of the language.

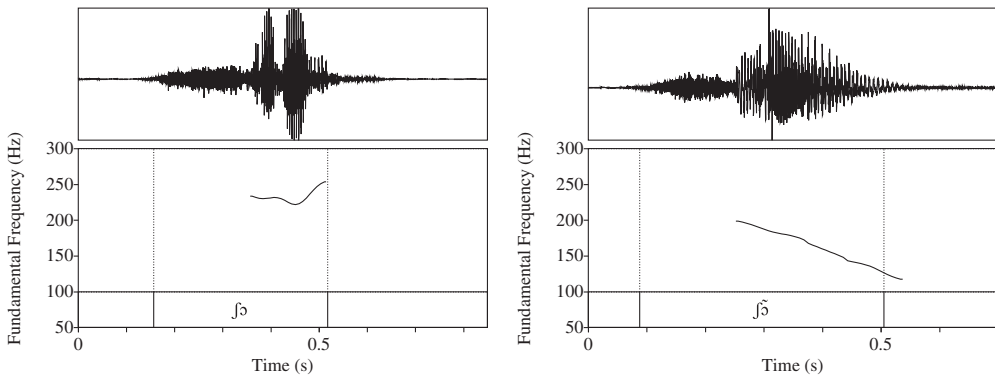
Approximately 45 of the 266 words contained more than one morpheme. This number is approximate due to the occasional difficulty of determining the compositionality of certain words. For example, the word /ʃi-p^hani/ ‘wineskin, *bota*’ is clearly composed of morphemes meaning ‘skin’ and ‘horse’, respectively. However, the similar word /ʃi-ŋgiri/ ‘corn silk’ is more opaque, since /ŋgiri/ is not clearly a morpheme (and not obviously related to /tɛt^hɔ/ ‘corn’). It is not clear to what extent these words are considered semantically compositional by speakers of SJAO (e.g. whether these words are more like *cranberry* and *boysenberry* or more like *perceive* and *receive*).

Eighty-nine of the 584 recordings were excluded from the analysis due to a lack of voicing on every syllable or noise in the recording interfering with interpretation of f₀; a total of 495 recordings were therefore annotated. The f₀ contour of every recording was inspected and noted. This process was carried out by auditory impression and visual inspection of an f₀ track, with each word being labelled for the overall trend of the f₀. Particular attention was paid to local pitch maxima, minima, and excursions, although small and short-term divergences in

⁸ The Nahuatl loans identified are /mbiʃt^hu/ ‘cat’ from Nahuatl *miston*, /ʃidɔmɔ/ ‘(red) tomato’ from Nahuatl *xitomatl*, /kʰaʃti/ ‘yellow’ from Nahuatl *kostik*, and /ʔwe/ ‘sweet’ from Nahuatl *uelik*. The latter two are potentially Otomi loans into Nahuatl, rather than the other way around, but this possibility will not be discussed further here (see Bartholomew 2000).

Table 4 Pitch contours present in the corpus of SJAO words. Cells with > 5% attestation are in boldface.

Pattern	One-syllable		Two-syllable		Three-syllable	
Flat	37	57.8%	37	9.7%	3	5.9%
Rising	1	1.6%	95	25.0%	8	15.7%
Falling	26	40.6%	191	50.3%	8	15.7%
Rising–falling	0	0%	49	12.9%	16	31.4%
Falling–rising	0	0%	5	1.3%	8	15.7%
Rising–falling–rising	0	0%	0	0%	1	2.0%
Falling–rising–falling	0	0%	3	0.8%	7	13.7%
Total	64	100%	380	100%	51	100%

**Figure 3** Waveforms and f0 traces of near-minimal pair. Left panel: flat contour on /ʃɔ/ ‘finger’; right panel: falling contour on /ʃɔ̃/ ‘burrow’.

the overall pattern, likely due to segmental perturbation, were ignored. Although this method is subjective, it was found to be reliable – a representative subset of the data was annotated again, and 83% agreement with the previous annotation was found.

The results are summarized in Table 4, which shows the number of occurrences of the recorded contour types by number of syllables in the word. As can be seen, some contour patterns appeared on all sizes of word (e.g. falling); some appeared only on two- and three-syllable words (e.g. rising–falling); and some appeared only on three-syllable words (e.g. rising–falling–rising). Some patterns were extremely rare (e.g. the rising pattern on monosyllables), and may be due to talker variability or higher-level prosodic factors. Patterns with an attestation rate greater than 5% are marked in bold. Comparison of loans with native words, of monomorphemic with multimorphemic words, and of nouns with non-nouns did not reveal any patterns or clear differences between the sets. Therefore, these words were not treated any differently from native monomorphemic nouns in the following analysis.

Figure 3 shows waveforms and f0 traces of the near-minimal pairs /ʃɔ/ ‘finger’ and /ʃɔ̃/ ‘burrow’, exhibiting flat and falling contours, respectively.⁹ As the following discussion illustrates, the distinction between the flat contour and the falling contour is central to the SJAO tonal system.

⁹ Note that the flat contour is not always completely flat. However, the scale of the f0 variation in the flat contours is qualitatively smaller than the excursions seen in the non-flat contours.

Table 5 All sixteen theoretically possible tonal contours on bisyllabic words given four tones: high (H), low (L), rising (R), falling (F).

HH	LH	RH	FH
HL	LL	RL	FL
HR	LR	RR	FR
HF	LF	RF	FF

2.2 Analysis

Of the contour types reported in [Table 4](#), setting aside the patterns that occur less than 5% of the time per word size, there are two patterns on monosyllables, four on disyllables, and six on trisyllables. A striking feature of the contour types, for words of all sizes, is the presence of falling contours. These falling contours prove to be fatal for Leon & Swadesh's (1949) accentual account of Otomi prosody, and are difficult to reconcile with Sinclair & Pike's (1948) tonal account. Further considerations of the distributional and combinatorial properties of tones suffice to cast doubt on the adequacy of a purely tonal analysis of Otomi prosody.

2.2.1 The failure of the purely accentual analysis

Under the accent proposal of Leon & Swadesh (1949), pre-accentual syllables are lower in f_0 than the accented syllable, and post-accentual syllables are the same or higher. Therefore, patterns of falling f_0 are impossible for words of any size, and the falling contours in the present data cannot be accounted for. Further, since the accentual account requires one and only one syllable within a word to be accented, the number of possible configurations for a word with n syllables is necessarily n . The current data provides evidence for at least six distinct contour types on words of up to three syllables; a purely accentual account cannot accommodate such diversity of melody.

2.2.2 The failure of the purely tonal analysis

Sinclair & Pike's (1948) tonal inventory is restricted to high, low, and rising tones, making falling contours impossible for monosyllables. It is possible to adjust the tonal proposal by adding a falling tone to the tonal inventory. Indeed, San Gregorio Otomi (Voigtlander & Echegoyen 1985) and San Pedro Atlapulco Otomi (Valle Canales 2008) are reported to have a falling tone in addition to the low, high, and rising tones. This adjustment brings the total number of distinct tones to four, which therefore makes possible sixteen distinct tonal melodies on bisyllables. These melodies are shown in [Table 5](#).

This analysis is therefore able to account for all four of the attested bisyllabic contours observed in the present corpus (that is, flat, rising, falling, and rising–falling). However, this analysis is far from parsimonious, for two reasons. First of all, as noted, sixteen contours are possible, yet only four are actually attested. A number of phonotactic restrictions are needed to ban the unattested contours. Plausible co-occurrence restrictions arising from, for example, the Obligatory Contour Principle, could reasonably exclude the diagonal of [Table 5](#); these restrictions still leave 12 permitted contours, three times the four contours attested on disyllables. These problems quickly multiply when we consider trisyllables, where there are 64 possible contours (4^3), but only six contours are actually attested.

Secondly, this analysis allows for contour alignment distinctions which are not warranted in the present data. For example, a falling–low (FL) tone sequence is theoretically distinct from a high–low (HL) sequence. In the first sequence, the pitch falls quickly in the first syllable and plateaus in the second syllable, while in the second sequence, the fall is between the two syllables. Under a four-tone analysis, these two contour types are phonologically distinct and ought to give rise to different interpretations. However, the present data do not provide any evidence that these distinctions are linguistically meaningful in SJAO. [Figure 4](#) shows

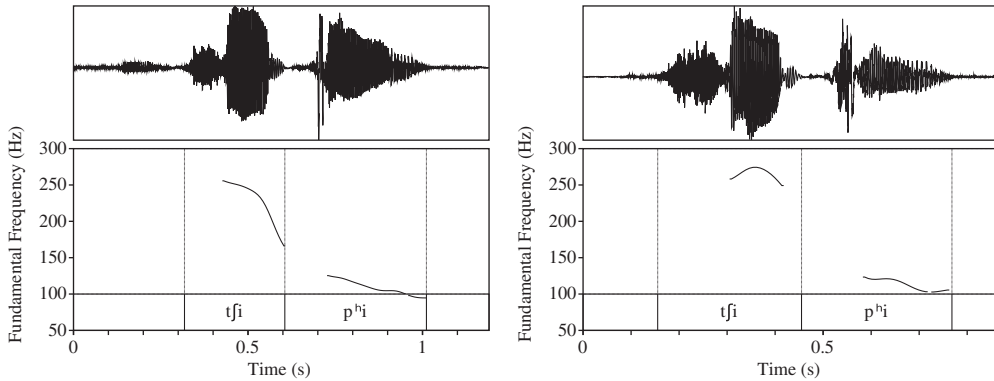


Figure 4 Waveforms and f0 traces of falling contour on two tokens of /tʃipʰi/ 'wasp'. Left panel: falling contour with fall beginning within first syllable; right panel: falling contour with fall between the first and second syllables.

waveforms and f0 traces of two tokens of the word /tʃipʰi/ 'wasp', a bisyllable produced with a falling contour. Here, we see two distinct contour alignments – which would be annotated as FL and HL under the four tone analysis – being used for the same lexical item. The existence of different alignments on the same lexical item provide strong evidence against the four-tone analysis. While the four-tone analysis may be a useful notation of phonetic detail, it fails to shed any light on the phonological system of the language.

For these reasons, the tonal account fails to adequately explain the observed patterns of SJA0 prosody. As we have seen, the accentual account is also unable to explain the patterns. A new analysis is therefore necessary, one which incorporates both accentual and tonal elements.

2.2.3 A novel analysis of Otomi prosody

This novel analysis begins with consideration of the interactions between number of syllables and number of attested contours reported in Table 4 above. There are two patterns on monosyllables, four on disyllables, and six on trisyllables. If a pattern can occur on a word with n syllables, it can also appear on words with more than n syllables. Visual examples of these f0 contours on words of various sizes can be found in the appendix, and several specific examples are discussed in Section 2.2.4.

Following the assumptions of autosegmental phonology, these contours can be treated as being interpolation between high (H) and low (L) pitch targets. For example, the 'falling' contour can be thought of as a sequence of a high pitch target followed by a low pitch target. The sequences HL (falling) and H (flat) can appear on words of all lengths – monosyllabic, disyllabic, and trisyllabic – while the other sequences are restricted to di- or trisyllabic words. This fact suggests that the sequences HL and H may be somehow 'basic' to the SJA0 tone system, and these sequences can be regarded as basic or 'primitive' f0 target sequences. I argue that all of the observed contours can be described in terms of these two basic tone types.

The proposed analysis relies on four basic assumptions (or axioms):

1. There are two tones: /H/ and /HL/.
2. Each lexical word has one and only one tone, which is lexically associated with one and only one syllable (the 'tonic' syllable).¹⁰

¹⁰These assumptions describe how the pitch melody is derived from the location and type of tone in a given SJA0 word. However, it is possible to go in the other direction, and determine the location of the tonic syllable from the pitch contour. In this case, the syllable associated with the pitch peak of the first rise in the word is the tonic syllable. If there is no pitch rise (as in the falling or flat contours), then the tonic syllable is the first one.

Table 6 The tonal patterns and their representation in terms of the primitive tone sequences /H/ and /HL/.

Pattern	Representation	Tone-syllable alignment	Attested on
Flat	/H/	1st	1, 2, 3 syllables
Falling	/HL/	1st	1, 2, 3 syllables
Rising	L+	/H/	2, 3 syllables
Rising-falling	L+	/HL/	2, 3 syllables
Falling-rising	H+L+	/H/	3 syllables
Falling-rising-falling	H+L+	/HL/	3 syllables

Table 7 All possible tonal sequences on words of one, two, and three syllables in length. Tones in /slashes/ are underlying tonal primitives. Rightward spreading is depicted with a dash (-) preceding the spread tone; pre-tonic tones are followed by a plus (+).

Contour	1 σ	2 σ		3 σ		
	σ_1	σ_1	σ_2	σ_1	σ_2	σ_3
Flat	/H/	/H/	-H	/H/	-H	-H
Falling	/HL/	/HL/	-L	/HL/	-L	-L
Rising		L+	/H/	L+	/H/	-H
Rising-falling		L+	/HL/	L+	/HL/	-L
Falling-rising				H+	L+	/H/
Falling-rising-falling				H+	L+	/HL/

3. Post-tonic syllables have the same tone as the tonic syllable (rightward spreading).
4. Syllables immediately preceding the tonic syllable are realized with a low tone; all other pre-tonic syllables are realized with a high tone.

Each of the observed f0 contour patterns follows from these assumptions.

Table 6 shows the representations of all six of these f0 contour patterns in terms of the two tones /H/ and /HL/ and any pre-tonic tones present. For the first two patterns in this table (flat and falling), the primitive tone (/H/ or /HL/, respectively) is aligned with the initial syllable. There are no pre-tonic syllables and rightward spreading takes care of any post-tonic syllables.¹¹ The middle two patterns in the table (rising and rising-falling) occur when the primitive tone is aligned with the second syllable in a word. Since the first syllable immediately precedes the tonic syllable, it is realized with an L tone. Since the tonic syllable is the second one, this pattern is impossible on monosyllables: it is impossible to align the primitive tone to the second syllable in the word if the word has only one syllable. The final two patterns, falling-rising and falling-rising-falling, have the tonal sequence /H/ or /HL/, respectively, aligned with the third syllable, with the initial pre-tonic syllables realized with H and L tones. This account explains why these final two patterns are only attested on trisyllables – because the primitive sequence must align with the third syllable. Table 7 shows all logically possible tonal sequences under this analysis on words of one, two, and three syllables in length.

As can be seen from Tables 6 and 7, some of the contours consist of one of the two primitive tones plus one or two pre-tonic tones. Notably, the number of pre-tonic syllables depends on which syllable the primitive tone is aligned to: if the tone is aligned to the second syllable, there is one pre-tonic syllable; if the tone is aligned to the third syllable, there are two pre-tonic syllables. When the tonic syllable is initial, no syllables are pre-tonic. It can be

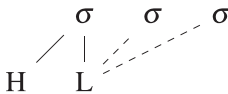
¹¹Such rightward tone spreading is reportedly a relatively common phenomenon in Oto-Pamean languages (e.g. Northern Pame, see Berthiaume 2004).

observed that the tone of the pre-tonic syllable directly adjacent to the tonic syllable is always L; the tone of the pre-tonic syllable adjacent to that L is always H.

This behavior could be explained as simply the default tones for this position. Another possibility is that there exist phrasal tones interacting with the lexical tones, causing these patterns, but since all these data came from words spoken in isolation, it is impossible to determine if this is the case. It is also possible to seek a more principled explanation in terms of the Obligatory Contour Principle (Leben 1973, Goldsmith 1976), whereby adjacent tones of the same type are disallowed, causing the H+L pattern observed prior to the tonic syllable. More abstractly, if the non-underlying H tones are regarded as minor prominences and the underlying H and HL tones as major prominences, then the patterns of pre-tonic syllables can be accounted for as avoidance of stress clash and lapse (Kager 1993, but compare Crowhurst & Olivares 2014). However, the exploration of these explanations is beyond the scope of this paper and will not be discussed further.

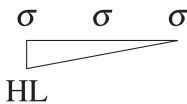
However, the representations provided in Table 7 are not entirely accurate, since they make a prediction which is not borne out in the data. That is, this representation predicts that in cases of rightward spreading of an /HL/ tone, the pitch will fall in the tonic syllable and then stay low for the rest of the word. However, the pitch patterns in the data are much more variable. Sometimes the fall is indeed almost completely contained in the tonic syllable (as in the left panel of Figure 4, or the middle right panel of Figure A2 in the appendix), whereas other times the fall is much more gradual (as in the right panel of Figure 4, or the middle right panel of Figure A1 in the appendix). It is therefore clear that the representation suggested in Table 7 is inadequate. Rather than the L tone spreading to each post-tonic syllable, as in (3), it appears that the /HL/ tone, as a unit, is associated with the entirety of the post-tonic material.

(3) *Rightward spreading of single L tone* (not observed in present data)



This spreading is shown in (4).

(4) *Rightward spreading of entire /HL/ melodic unit* (observed in present data)



In this case, the /HL/ tone is realized over all post-tonic syllables, with the f_0 essentially being interpolated from the peak during the tonic syllable to the nadir in the final syllable. One approach to formalizing this process is via Yip's (1989) notion of a MELODIC UNIT and the use of Span Theory (McCarthy 2004) to align the tones with the appropriate domain (namely, from the tonic syllable to the right edge of the prosodic word). However, detailed discussion of the technical implementation of this kind of spreading is beyond the scope of this paper. The representation of rightward spreading used in (4) will be used for the remainder of this paper.

2.2.4 Examples

To understand the application of these processes more concretely, consider the following example. The word /taʃk^hwa/ 'turkey' has a falling pitch contour. A waveform and f_0 trace of this word is shown in the left panel of Figure 5. Under the present analysis, the falling pitch contour means that there is a primitive HL sequence aligned with the first syllable (as per assumption 2 above), shown in (5a). Following assumption 3, the HL tone of the sequence spreads rightward to the following syllable, shown in (5b), resulting in a surface tonal sequence of HL, as observed in the left panel of Figure 5.

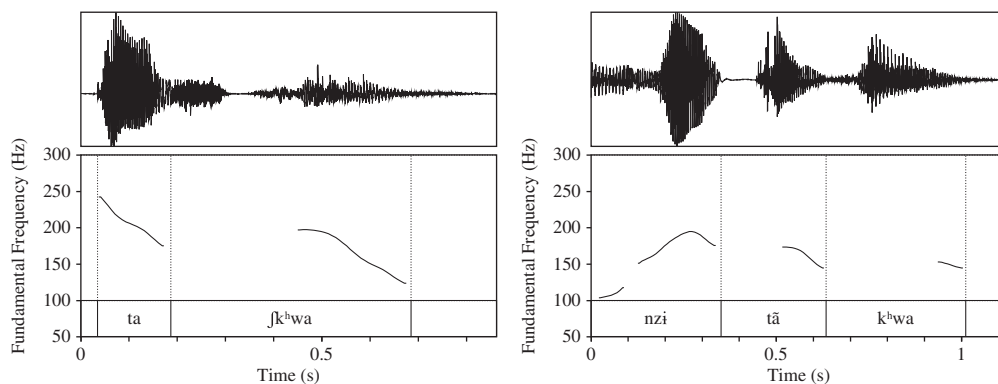


Figure 5 Waveforms and f0 traces of root form and prefixed form of a word, demonstrating default L tone on pre-tonic syllable. Left panel: falling contour on /taʃkʰwa/ ‘turkey’; right panel: rising-falling contour on /nzi-tākʰwa/ FEM-turkey ‘turkey hen’.

(5) a. *Underlying form*

ta ʃkʰwa
|
HL

b. *Rightward spreading*

ta ʃkʰwa
▱
HL


Under assumption 4, pre-tonic syllables are assigned the default tone of L if they are immediately prior to the tonic syllable, and H otherwise. From this assumption a prediction follows that a prefixed form /taʃkʰwa/ will be pronounced with a rising-falling pitch contour, that is, a LHL sequence. This prediction can be tested through the use of the feminine prefix /nzi/ to make /nzi-tākʰwa/ ‘turkey hen’.¹² Given the underlying form shown in (6a), the present analysis predicts that, as with the root form, the HL will spread rightward (shown in (6b)), and that the initial pre-tonic syllable will be assigned a L tone (shown in (6c)). This prediction is borne out in the data: the right panel of Figure 5 shows a waveform and f0 trace for this word.

(6) a. *Underlying form*


nzi tākʰwa
|
HL

¹²The root morpheme /taʃkʰwa/ undergoes changes to become /tākʰwa/; this alternation is irregular and not relevant to this example. Further examples are presented below which do not feature such an alternation.

b. *Rightward spreading*

nzi t̄ā k^hwa

 HL

c. *Pre-tonic syllable assigned default tone*

nzi t̄ā k^hwa

 L HL

Similar analysis can be carried out on shorter words, such as /ʔjo/ ‘dog’, shown in Figure 6, which has a flat pitch contour. Use of the feminine and masculine prefixes leads to the derived forms /nzi-ʔjo/ ‘bitch’ and /ta-ʔjo/ ‘male dog’, respectively. As before, the pre-tonic syllables (the prefixes) are assigned the default tone of L (assumption 4). Again, these predictions are borne out in the data; waveforms and f0 traces of these two words are shown in the left and right panels of Figure 7, respectively.

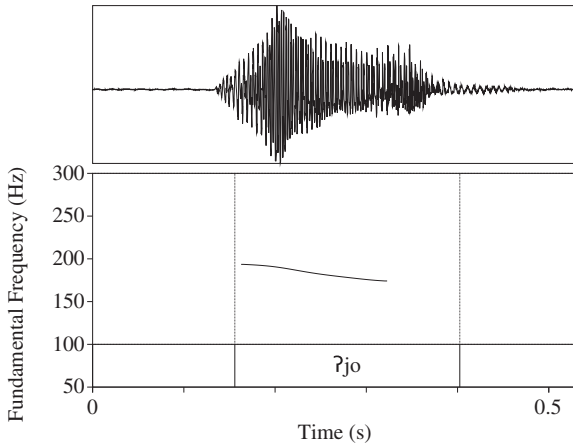


Figure 6 Waveform and f0 trace of flat f0 contour on word /ʔjo/ ‘dog’.

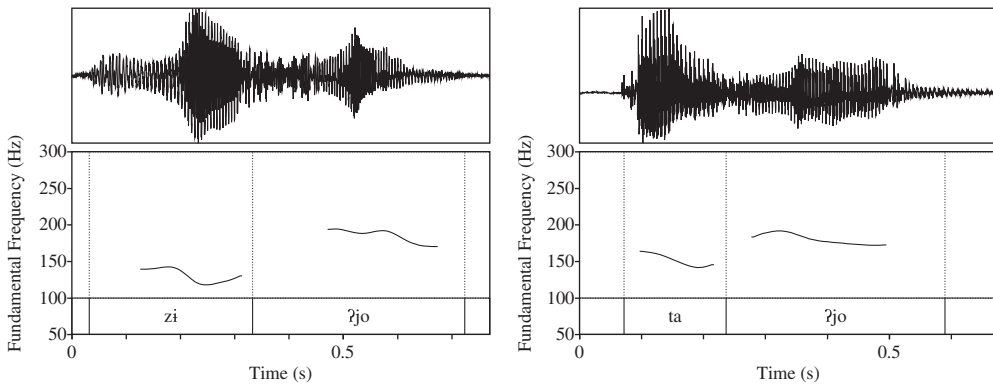


Figure 7 Waveforms and f0 traces of prefixed words with rising contours, demonstrating default L tone of pre-tonic syllable. Left panel: /nzi-ʔjo/ FEM-dog ‘bitch’; right panel: /ta-ʔjo/ MALE-dog ‘dog’.

The same reasoning used to test the behavior of the pre-tonic syllables can be used on the post-tonic syllables. Consider (7a), which shows /tɔʔjo/ ‘bone’, with its underlying HL tone associated with the first syllable. As per assumption 3, the HL tone spreads rightward, as shown in (7b). A waveform and f0 trace of this word is shown in the left panel of Figure 8.

(7) a. *Underlying form*

tɔ ʔjo
|
HL

b. *Rightward spreading*

tɔ ʔjo
└───▶
HL

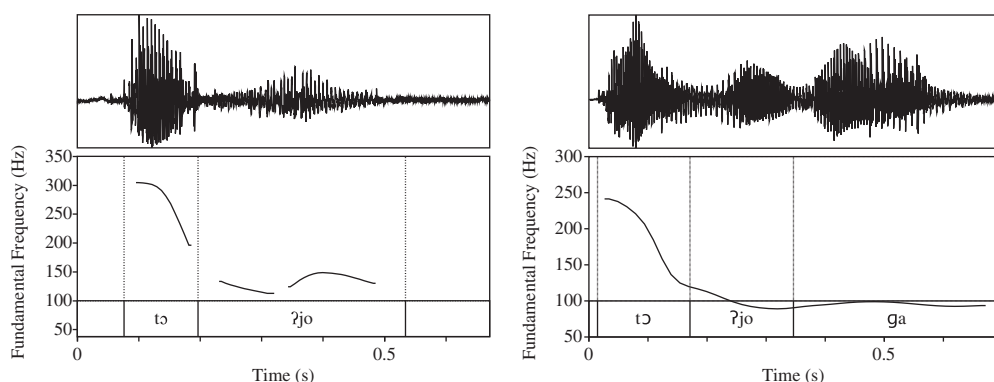


Figure 8 Waveforms and f0 traces of root form and suffixed form of a word, demonstrating rightward spreading. Left panel: falling contour on /tɔʔjo/ ‘bone’; right panel: falling contour on /tɔʔjo-ga/ ‘(my) bone’.

Nouns in SJAO take suffixes to mark possession, as shown in (8). In this example, the suffix /-ga/ marks first person singular possession. Under the present analysis, this suffix should undergo rightward spreading and bear the same tone as the tonic syllable, as shown in (9). As predicted, this spreading is exactly what happens: the right panel of Figure 8 shows a waveform and f0 trace of the word /tɔʔjo-ga/, excised from the phrase in (8).

(8) na-m tɔʔjo-ga
DET.SG-my bone-my
‘my bone’

(9) *Rightward spreading*

tɔ ʔjo ga
└───▶
HL

The same suffix /ga/, can be observed with an H tone when it follows an H-toned tonic syllable, as in /ʃɔ-ga/ ‘(my) finger’. Figure 9 shows a waveform and f0 trace of this word, extracted from the fuller phrase /na-m ʃɔ-ga/ ‘my finger’. A comparison of this figure with the right panel of Figure 8 demonstrates the rightward spreading posited in assumption 3.

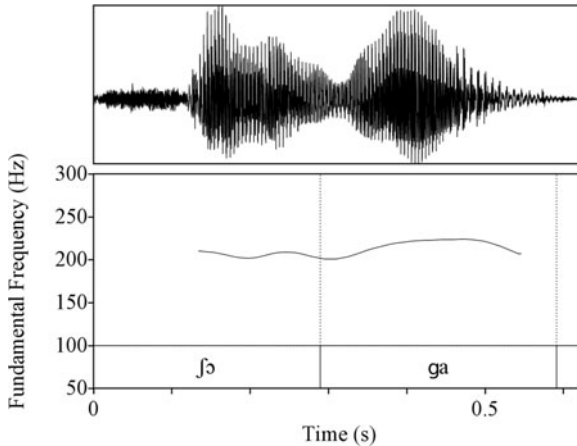


Figure 9 Waveform and f0 trace of flat f0 contour on word /ʃɔ-ga/ 'my finger'.

Under this analysis, there is a lexical specification of one and only one syllable to either a /H/ or /HL/ tone. The pre-tonic syllables are filled in by the default tones mentioned above, and the tone of the tonic syllable spreads to cover all post-tonic syllables. All six of the contour patterns listed in Table 6 and their distribution over word sizes can be accounted for through this analysis. Examples of waveforms and f0 traces of each of these contours are provided in the appendix.

2.3 Minimal pairs

Returning to the minimal pairs identified in (1) above, the present analysis can account for these differences in terms of tone-to-syllable alignment and tone type. Rather than drawing association lines in a multi-tier representation, tone-to-syllable alignment can be depicted through the use of the IPA's tone diacritics, namely an acute accent (´) for an H tone and a circumflex (ˆ) for an HL tone, on the vowel of the tonic syllable. The minimal pairs in (1) are repeated in (10)–(12).

- (10) a. (i) /kʰíhi/ 'to bring'
 (ii) /kʰiḥi/ 'color, dye'
 b. (i) /kʰiḥni/ 'beard'
 (ii) /kʰiḥní/ 'dough'
 c. (i) /tʰɔ́ʃjo/ 'bone'
 (ii) /tʰɔʃjô/ 'griddle'
 d. (i) /páʃi/ 'to travel downriver'
 (ii) /paʃí/ 'grass, trash'
- (11) a. (i) /ʔwîni/ 'thorn'
 (ii) /ʔwíni/ 'to feed'
 b. (i) /tʰêni/ 'red thing'
 (ii) /tʰéni/ 'to grab'

- (12) a. (i) /tʰɔnɛ̂/ ‘horn (of animal)’
 (ii) /tʰɔne/ ‘gourd, *jicara*’
 b. (i) /ʔʰhɔ̃/ ‘to sleep’
 (ii) /ʔʰhɔ̃/ ‘yes’

Note that the pairs in (11) involve contrasts of the type of tone associated with the tonic syllable – HL contrasts with H. The pairs in (10), on the other hand, involve contrasts of the LOCATION of the tonic syllable – a tonic first syllable is contrasted with a tonic second syllable. Finally, the pairs in (12) exemplify contrasts of both tonal type and alignment.

2.4 Caveats

The present analysis successfully accounts for the observed word-prosodic contrasts in SJAQ. Nevertheless, caution is required in the consideration of the generalizability of this analysis, for two principal reasons: variability in the data, and the restrictive nature of the speech corpus used.

In terms of variability, some tokens are not wholly explainable under the current analysis. One such token was shown in the right panel of Figure 5 above. This example was presented as /nzi-tâk^hwa/ ‘turkey hen’, with a /HL/ tone aligned with the second syllable, which the analysis predicts should yield rising-falling pitch contour, with the pitch peak on the second syllable. As can be seen, this contour is observed, but the peak is on the first syllable, rather than the predicted second syllable. There is currently no explanation for this behaviour. A morphological explanation could be pursued: as noted in footnote 12, the root morpheme /taʃkwa/ in the unprefixated form changes to /tâkwa/ in the prefixated form. It is possible that the morpheme’s tonal phonology is also altered, in addition to its segmental phonology, although the present analysis is not able to model such an alteration. Field data of prosodic contrasts are often quite variable, and in this instance the available data do not allow a fuller examination.

More generally, in light of studies of linguistic variation in communities undergoing language shift from a moribund language (e.g. Dorian 1994), the degree of variability observed in the data is not surprising. Indeed, aspects of lexical and phonological variability in SJAQ itself are documented by Phrao Hansen et al. (2016) and hypothesized to be related to the ongoing language death.

The second reason for caution is the nature of the corpus used. The corpus is relatively small (584 tokens), collected from a single talker, and largely consists of nouns. For the sake of keeping the scope of investigation manageable, morphological alterations were not considered in the construction of the corpus. These and more are left to future research, and it is a possibility that the proposed analysis will need refinement and revision as new facts come to light.

3 Phonetics of tonic syllables

That Otomi is a language which uses f0 (possibly among other cues) to maintain lexical distinctions is uncontroversial. The novel component of the present analysis is the claim that one syllable in a SJAQ word is ‘tonic’ and is lexically associated with an underlying tone. The final pitch contour of the word depends on the location and type of this tone. This analysis is abstract, and based solely on the distributional properties of the observed f0 contours in the language. The goal of this section is to determine whether tonic syllables in SJAQ are physically different from the non-tonic syllables, in addition to being distinct phonologically.

Evidence of such a difference would strengthen the phonological argument in providing an additional basis for claiming that the tonic syllables are ‘special’ or different from the other syllables. There are reasons to expect that such syllables could be physically distinct from non-tonic syllables. For example, Swedish features a lexical prosody that is similar to the present analysis of SJA0, in that there are two tones (assumption 1 in the present analysis) and there is one and only one syllable per word which is associated with a tone (assumption 2) (Shaffler 2005). In Swedish, the ‘tonic’ or ‘stressed’ syllable is routinely realized with greater duration, more peripheral vowels, and more vocal effort than the non-stressed syllables (Engstrand 1988, Fant, Kruckenberg & Nord 1991, Heldner 2003). Note, however, that the absence of evidence would not invalidate the phonological claims made in this paper. In languages such as Japanese, where one syllable per word is ‘accented’ with a tonal specification, the accented syllables are not longer, louder, or produced with more peripheral vowels than the unaccented syllables (Homma 1981, Beckman 1986). The main phonological argument in this paper does not hinge on the evidence provided in this section; but the evidence in this section does lend further empirical support to my claims.

3.1 Materials

Since the goal is to compare tonic syllables to non-tonic syllables, only disyllabic words are considered. This restriction is because the syllables of monosyllabic words are all tonic, and so cannot be compared with non-tonic monosyllables, and comparing syllables in trisyllabic words is statistically intractable due both to collinearity between control factors and the relative paucity of trisyllables in the corpus.

From the corpus of recordings described in Section 2.1 above, 427 disyllabic words were extracted. Each recording was segmented into phones and subjected to acoustic analysis. For each vowel in the word, measures of duration, fundamental frequency, vowel formants, spectral tilt, and intensity were taken.

For each recording, the onset and offset of each vowel and each syllable were identified. The onset (offset) of a vowel was defined as the zero crossing of the waveform at either: the beginning (end) of periodicity; or, in the case of transitions between voiced sounds, when the most significant change in the shape of the waveform occurred. Vocalic segments such as /w/ and /j/ were included with the vowel – e.g. [ja] was annotated as a single segment.

3.2 Measurements

Measurements of duration, intensity, fundamental frequency, vowel centralization, and spectral tilt were obtained. These measurements all reflect common acoustic realizations of tone and accent relations in the world’s languages (Fry 1955, de Jong 1995, Gordon & Ladefoged 2001, Dilley 2005, Gordon 2011). For instance, in some languages, different lexical tones surface with distinct durations (e.g. Mandarin; Blicher, Diehl & Cohen 1990, Whalen & Xu 1992). Duration is also a common correlate of lexical stress, with longer duration making a syllable more prominent (see Fry 1955, Plag, Kunter & Schramm 2011 *inter alia*). The duration of each syllable in a given word was calculated from the demarcation of the onset and offset of the syllable.

For each vowel, the fundamental frequency (f0) was measured at the vowel’s midpoint, and a measure of f0 change throughout the vowel was calculated. The slope of the f0 change throughout the vowel was calculated by linear regression between f0 measurements at three timepoints – one quarter of the way through the vowel; halfway through the vowel; and three-quarters of the way through the vowel. This value of f0 change is referred to as $m(f0)$; in general, measurements of change in unit x are denoted here as $m(x)$.¹³ The f0 values were

¹³The letter m was chosen because it represents the slope value in a linear model formula: $\hat{y} = mx + b$.

calculated with a frame duration of 10 ms, using Boersma's (1993) autocorrelation algorithm. Intensity was measured in dB averaged over the entire vowel.

A measure of vowel centralization provides a measure of how far each token is from the center of the vowel space (Wright 2004). This measure is useful in assessing whether certain prosodic positions undergo segmental reduction – for instance, English commonly reduces unstressed syllables to schwa. For each of the nine vowel phonemes (/a ε e i ɔ o u i ə/), the mean of each of the first three formants was calculated, thus creating a 'mean /i/' vowel point in $F_1 \times F_2 \times F_3$ space, and so on. This averaging was performed because there are not an equal number of tokens of each vowel type. From these nine mean vowel points a grand mean was calculated, which, because the vowel system of SJAO is symmetrical, represents the center of the vowel space. Vowel centralization was then calculated for each token by taking the Euclidean distance, in $F_1 \times F_2 \times F_3$ space, from this center point to the token. Thus, a higher value indicates a more peripheral vowel (more distance from the center to the token), and a lower value is a more centralized vowel (less distance from the center to the token). However, these numbers can only be compared with like vowels because, for example, /ə/ is intrinsically closer to the center of the vowel space than is /a/ or /i/. To correct for this inequality so that the degree of centralization of /ə/ can genuinely be compared with that of /i/, each token's value was scaled by taking the base-10 logarithm of the value divided by the mean centralization value for that phoneme: i.e. $\log_{10}(x / \bar{x})$. Without taking the log transform, the magnitude of the values are skewed, such that tokens with greater-than-average centralization bear less weight than tokens with smaller-than-average centralization. The log transform makes all deviances from the mean equally weighted, regardless of direction.

This measure was calculated from formants estimated at the vowel midpoint; a measure of change in centralization throughout the vowel was calculated using linear regression through three timepoints using the same method as described for f0. The measurement of change of centralization is referred to as $m(\text{centralization})$. Conceptually, a positive $m(\text{centralization})$ value means that the vowel's production becomes more peripheral (and presumably more perceptually clear) toward the end of articulation; a negative value means that the production becomes more centralized and schwa-like. To estimate the formant values, the recordings were first downsampled to 16 kHz before being processed with Praat's 'To formant (burg)' function (time step: 10 ms, maximum formant: 5 kHz; sample window: 25 ms; pre-emphasis from: 50 Hz).

Spectral tilt is a measure of the distribution of energy in a sound's spectrum – a comparison of the component amplitudes at different frequency bins. In this study, two measures of spectral tilt were taken: $H_1 - H_2$, the amplitude of the first harmonic (H_1) relative to that of the second (H_2); and $H_1 - A_3$, the amplitude of the first harmonic (H_1) relative to the amplitude of the third formant (A_3). $H_1 - H_2$ is regarded by many as a more or less direct measurement of the open quotient of the vocal folds (e.g. Klatt & Klatt 1990, but compare Henrich, d'Allesandro & Doval 2001) and is a common measure of voice quality (e.g. Huffman 1987, Ladefoged, Maddieson & Jackson 1988, Wayland & Jongman 2003, Keating & Esposito 2007). Kreiman, Gerratt & Antoñanzas-Barroso (2007) demonstrated that most spectral tilt measures are strongly correlated with $H_1 - H_2$, concluding that it is a useful reflection of both articulatory and perceptual mechanisms, and can be reliably estimated from the acoustic signal. $H_1 - A_3$ has been shown by Hanson (1997) and Kreiman et al. (2012) to be a reliable reflection of the spectral tilt of the glottal source. These findings mean that $H_1 - A_3$ is a good cue for examining non-supralaryngeal effects of stress and vocal effort, such as those examined by Sluijter & van Heuven (1996). For both $H_1 - H_2$ and $H_1 - A_3$, larger values correlate with more breathy or more relaxed voice quality, while smaller or negative values correlate with more tense or more effortful voice quality (Gordon & Ladefoged 2001). To estimate $H_1 - H_2$ and $H_1 - A_3$, the first harmonic must first be identified. Since the first harmonic is the fundamental frequency, the output of the previously-mentioned f0 estimation algorithm was used to extract amplitude measures. The second harmonic was located by finding the

largest spectral peak between $f_0 \times 1.9$ Hz and $f_0 \times 2.1$ Hz – i.e. at twice the fundamental frequency, plus or minus ten percent. The third formant was identified using LPC as described above.

Both of these spectral tilt measures are sensitive to variations in vowel quality, making the comparison of spectral tilt between [i] and [a] nontrivial. To account for this sensitivity, the measures of spectral tilt were corrected to adjust for the influence of the first three formants, using the method outlined by Iseli, Shue & Alwan (2007). Iseli's correction requires an estimate of formant bandwidths. The estimation was carried out using the equations of Hawks & Miller (1995), which are rough approximations based on a fifth-order polynomial regression fitted to recorded data.¹⁴ As with the f_0 and vowel centralization measures, the spectral tilt values were calculated both at the vowel midpoint and as a value of rate of change over the whole vowel, which are referred to as $m(H_1 - H_2)$ and $m(H_1 - A_3)$. These measures capture the change in vocal effort or voice quality over the duration of the vowel.

3.3 Statistical analysis

Linear mixed effect regression models were constructed, with tonicity (/H/-toned, /HL/-toned, or non-tonic), syllable position (initial or final), and the interaction between them as fixed effects. Word identity was entered as a random intercept, with tonicity and syllable position as random slopes. The models were constructed to predict syllable duration, vowel intensity, f_0 , $m(f_0)$, $H_1 - H_2$, $m(H_1 - H_2)$, $H_1 - A_3$, $m(H_1 - A_3)$, vowel centralization, and $m(\text{centralization})$.

3.4 Results

The model predicting intensity did not converge with the interaction between position and tonicity. Once the interaction term was removed from the model specification, the model converged. This result means that the potential interactions between syllable position and syllable tonicity on intensity cannot be assessed. Nevertheless, this more limited model still provided interesting results, discussed here. All of the other models converged with fully specified terms.

Simple effects of position were observed for the analyses of duration, intensity, f_0 , change in f_0 , and change in vowel centralization, meaning that initial syllables differed significantly from final syllables in terms of these acoustic dimensions. Relative to the final syllable, the initial syllable was shorter, had a greater intensity, had a higher midpoint f_0 , a less negative $m(f_0)$ (indicating a less steep decline in f_0 throughout the vowel), and a smaller $m(\text{centralization})$ (indicating that the vowel tended to become more centralized over time). See Table 8 for a statistical summary, including mean values for initial and final position. Note here that 'initial' and 'final' apply both at the word level and the phrase level, since these productions are words produced in isolation. That is, 'initial' means both 'word-initial' and 'phrase-initial'. The phonetic differences reported in Table 8 between syllables in initial and final position are relatively unsurprising, given previous research on phrase-final lengthening, f_0 declination, and positional constraints on prosody (Delattre 1966, Ladd 1984, Cole, Hualde & Iskarous 1999, Turk & Shattuck-Hufnagel 2007, Nakai et al. 2009).

Significant simple effects of tonicity were observed for the analyses of duration, intensity, f_0 , $H_1 - H_2$, and $m(H_1 - A_3)$. Compared to non-tonic vowels, vowels with a tonic /H/ tone had a lower $H_1 - H_2$, indicating a more tense or effortful production (Gordon & Ladefoged

¹⁴A more accurate method is outlined by Hanson (1997, based on House & Stevens 1958); however, its implementation is non-trivial as it requires both a reliable estimate of the length of the speaker's vocal tract and manual demarcation of glottal cycles in the waveform. To my knowledge, there has not yet been a rigorous study comparing the efficacy and accuracy of these and other methods of bandwidth estimation.

Table 8 Summary of means and significant simple effects of position on five acoustic measures.

	Initial	Final	β	t	p -value
Duration (ms)	246	376	-121.58	-12.42	< .001
Intensity (dB)	67.46	62.27	5.02	12.62	< .001
Midpoint f0 (Hz)	193.1	169.2	20.28	5.24	< .001
m (f0)	-4.29	-18.28	16.16	8.86	< .001
m (centralization)	-0.018	0.011	-0.04	-3.51	< .001

Table 9 Summary of means and significant simple effects of /H/ tone on two acoustic measures.

	Non-tonic	/H/-toned	β	t	p -value
$H_1 - H_2$ (dB)	-1.75	-4.28	-3.67	-2.03	< .05
$m(H_1 - A_3)$	1.65	-0.06	-2.41	-1.99	< .05

Table 10 Summary of means and significant simple effects of /HL/ tone on two acoustic measures.

	Non-tonic	/HL/-toned	β	t	p -value
Intensity (dB)	63.7	66.8	1.35	2.98	< .005
f0 (Hz)	171.9	195.8	25.28	4.68	< .001

2001). Additionally, while the non-tonic vowels had positive $m(H_1 - A_3)$ values, the values for the tonic /H/ vowels were around zero. This trend suggests that production of the non-tonic vowels becomes more breathy, or less effortful, as the vowel is articulated. The tonic /H/ vowels, though, had relatively constant $H_1 - A_3$ values throughout production. See Table 9 for a statistical summary of these effects, including mean values for non-tonic and /H/-toned syllables. These results suggest that syllables with a tonic /H/ tone are produced with greater and more consistent glottal effort than non-tonic syllables.

Syllables with tonic /HL/ tone were also significantly distinct from non-tonic syllables, but in different acoustic dimensions than syllables with tonic /H/ tone. Compared to the non-tonic, the vowels with /HL/ tone had greater intensity and a higher f0 peak. See Table 10 for a statistical summary of these effects, including mean values for non-tonic vs. /HL/-toned syllables. For duration, both a simple effect of tone and an interaction with syllable position were observed. Overall, tonic /HL/ syllables were longer than non-tonic syllables ($\beta = 41.55$, $t = 3.137$, $p < .002$), but this distinction was neutralized in initial position ($\beta = -48.59$, $t = -3.061$, $p < .005$) – initial /HL/ $M = 240$ ms vs. initial non-tonic $M = 252$ ms; compare with final /HL/ $M = 433$ ms vs. final non-tonic $M = 367$ ms.

A potential concern for the interpretation of these results is the presence of nasal vowels in the dataset. This concern is due to the fact that spectral tilt measures involving the first and second harmonics (such as $H_1 - H_2$) have been shown to correlate with nasality (Simpson 2012). If the distribution of nasal vowels across tonic and non-tonic syllables is highly unbalanced, then any observed effects may be due to nasality rather than tonicity. However, this concern is unfounded – of the 63 nasal vowel tokens in the data set, 34 of them are non-tonic and 29 of them are tonic. Even so, the above analyses were repeated on a dataset with the nasal tokens removed, and the same patterns of results were observed, ruling out vowel nasality as a potential confound.

Taken together, these analyses indicate that tonic syllables are acoustically distinct from non-tonic syllables. Compared to non-tonic syllables, syllables with /HL/ tone are longer and have greater intensity and f0 on the vowel; and syllables with /H/ have more consistently tense vowels, produced with greater vocal effort. The generalization here is that the tonic syllables

have greater acoustic prominence, manifested through a variety of acoustic means, than do the non-tonic syllables. In this respect, the tonal system of SJAJO can be said to be more similar to that of Swedish than Japanese. This phonetic evidence supports the hypothesis that the tonic syllables are different from the non-tonic syllables; this difference may be attributed to the underlying tonal sequence associated with these syllables licensing additional prominence.

4 Discussion and conclusion

Sinclair & Pike's (1948) account of Mezquital Otomi prosody argued that each syllable is specified for one of three tones – low, high, or rising. This set of tones cannot account for the present data from SJAJO, since falling f_0 patterns are observed on monosyllables. The data can be accounted for by modifying the tonal inventory to include a falling tone, but such an analysis predicts several unattested forms. My proposed analysis predicts only attested forms, due to the fact that tonal specification is limited to one syllable per word.

Leon & Swadesh's (1949) account of Mezquital Otomi prosody argued that each word has a single accented syllable. Such an account predicts that there are two possible accentual patterns for disyllabic words – initial accent, or final accent. However, the present data from SJAJO exhibit four contrastive f_0 patterns in disyllabic words (see the second column of Table 4); such diversity of patterns cannot be explained by a purely accentual analysis. My proposed analysis retains the requirement that one and only one syllable is specified for prosodic prominence, but this prominence is not a privative distinction but rather one of two tones.

The present analysis differs in several respects from previous descriptions of Otomi prosody. In some ways, the analysis can be regarded as unifying the accounts of Sinclair & Pike (1948) and Leon & Swadesh (1949). SJAJO is argued to feature one prosodically prominent syllable per word (as per Leon & Swadesh 1949); nevertheless, these tonic syllables contrast in underlying tone (as per Sinclair & Pike 1948). In contrast to the high, low, and rising tones of Sinclair & Pike's (1948) tonal hypothesis, however, SJAJO is argued to feature high and falling underlying tonal sequences. This current analysis could be termed a 'pitch accent' system, as opposed to a 'tonal' system. However, the existence of 'pitch accent' as a useful category in prosodic typology has been a subject of extensive debate (see Pulleyblank 1986, van der Hulst & Smith 1988, Odden 1999 and Hyman 2006, 2009, who argue against the term, and Hualde et al. 2002, Donohue 2005 and Hualde 2006, who argue in its favor). Discussion of this debate is beyond the scope of the present paper, and the terms used to describe the typology of SJAJO do not affect the facts of the analysis.

Under one perspective, it is perhaps not surprising that the analyses of Sinclair & Pike (1948) and Leon & Swadesh (1949) are unable to account for the SJAJO data, since their analyses were of Mezquital Otomi, a different language variety. Nevertheless, the Sinclair & Pike (1948) tonal analysis has been applied to every Otomi variety studied, including SJAJO (Hernández-Green 2015) and the related dialects of Ixtenco (Lastra 1997), Tilapa (Palancar 2012), and San Pedro Atlapulco (Valle Canales 2008). Both the data and the analysis presented in this paper are quite different from those reported in Sinclair & Pike (1948) and Leon & Swadesh (1949); this paper is also among the first to present acoustic data (f_0 traces and waveforms) from any Otomi variety (see also Valle Canales 2008, Volhardt 2014, Pharaoh Hansen et al. 2016). It remains to be seen how well the current analysis of SJAJO prosody will account for the prosodic systems of other Otomi varieties.

This question highlights the possibility that being overly hasty in assuming that the analysis of lexical prosody for language X will apply well to the closely-related language Y is to the detriment of thorough linguistic description. Just as other elements of the grammar, such as morphosyntax and segmental phonology, differ, often substantially, between Otomi dialects (Lastra 2001), the lexical prosodic systems also appear to be subject to considerable variation. As such, more phonetically- and phonologically-oriented research is warranted.

The present study was conducted using a relatively small corpus consisting mostly of nouns, spoken in isolation by a single speaker. The data are somewhat variable, and as discussed in Section 2.4, there are a small minority of tokens which do not conform to the present analysis. This analysis also does not account for the minimally attested contours, set in non-boldface in Table 4. Whether these represent infrequent but contrastive contour types or were random quirks or idiosyncrasies of the language consultant is currently unknown. Nevertheless, support for the present analysis is bolstered by the results of the phonetic study in Section 3, which demonstrated that the theorized tonic syllables exhibit physical differences from the non-tonic syllables. It is quite likely that future work on SJA0 will reveal shortcomings of the present analysis, requiring revision and refinement. Accordingly, a promising area for future study is to extend the domain of investigation to include longer phrases and utterances. Tone sandhi processes and phrase edge effects are common in tonal languages (Gandour 1978) and have been documented in several Oto-Manguean languages (e.g. Chiquihuitlán Mazatec, Jamieson (1977); Comaltepec Chinantec, Silverman (1997); Isthmus Zapotec, Mock (1983) *inter alia*) including varieties of Otomi (e.g. Wallis 1968) – it is not yet known whether they exist in SJA0, or – if they do – what properties they exhibit. These and other questions are left open to future research.

The fields of phonetics and phonology have come a long way since Sinclair and Pike (1948) and Leon and Swadesh (1949) first outlined their analyses of Otomi prosody. Our knowledge of tonal and accentual phenomena has grown immensely since then, too, and powerful instrumental and statistical methods can be applied to speech data to help test hypotheses. Given these developments, perhaps the time is ripe for a broader re-evaluation of tone in Otomi, beyond SJA0. Naturally, whether other Otomi dialects have a system similar to that of SJA0 remains an open question. It could well be the case that SJA0 is an outlier among the Otomi dialects, in terms of its lexical prosodic system. Indeed, whether this one speaker's prosodic system can be generalized to other speakers of SJA0 is also an empirical question. Nevertheless, the present data clearly demonstrate a system whereby only one syllable per word is tonally specified, in contrast with previous analyses. This line of inquiry also raises the question of how Otomi prosody developed historically from Proto-Oto-Pamean – a language assumed to be fully tonal (see e.g. Bartholomew 1965), and it further calls into question the supposedly fully tonal status of other Oto-Pamean languages.

Acknowledgements

Many thanks are owed to Cynthia Clopper, Dave Odden, and Judith Tonhauser for their guidance, support, and critique throughout this project. Amalia Falon García Grajeda, Magnus Pharaoh Hansen, Néstor Hernández-Green, Anne Jensen, and Ditte Boeg Thomsen provided me with invaluable assistance and comments both in and out of the field. I would also like to thank OSU's Phonetics and Phonology discussion group, the Prosody and Meaning seminar, Mary Beckman, Bridget Smith, and Kodi Weatherholtz for giving me feedback at various stages of this project. Finally, this work would not have been possible without the assistance, cooperation, and kindness of the people of San Jerónimo Acapulco. All errors remain my own. Portions of this work were supported through financial assistance from the Ohio State University Department of Linguistics.

Appendix. Pitch contour examples

This appendix provides examples of each of the attested pitch contours on words of one, two and three syllables in length.

One-syllable words

Two pitch contours are observed on one-syllable words: flat and falling. Waveforms and f_0 traces of the words /t'ej/ 'hot corn drink, *atole*' and /k'i/ 'blood' are shown in Figure A1.

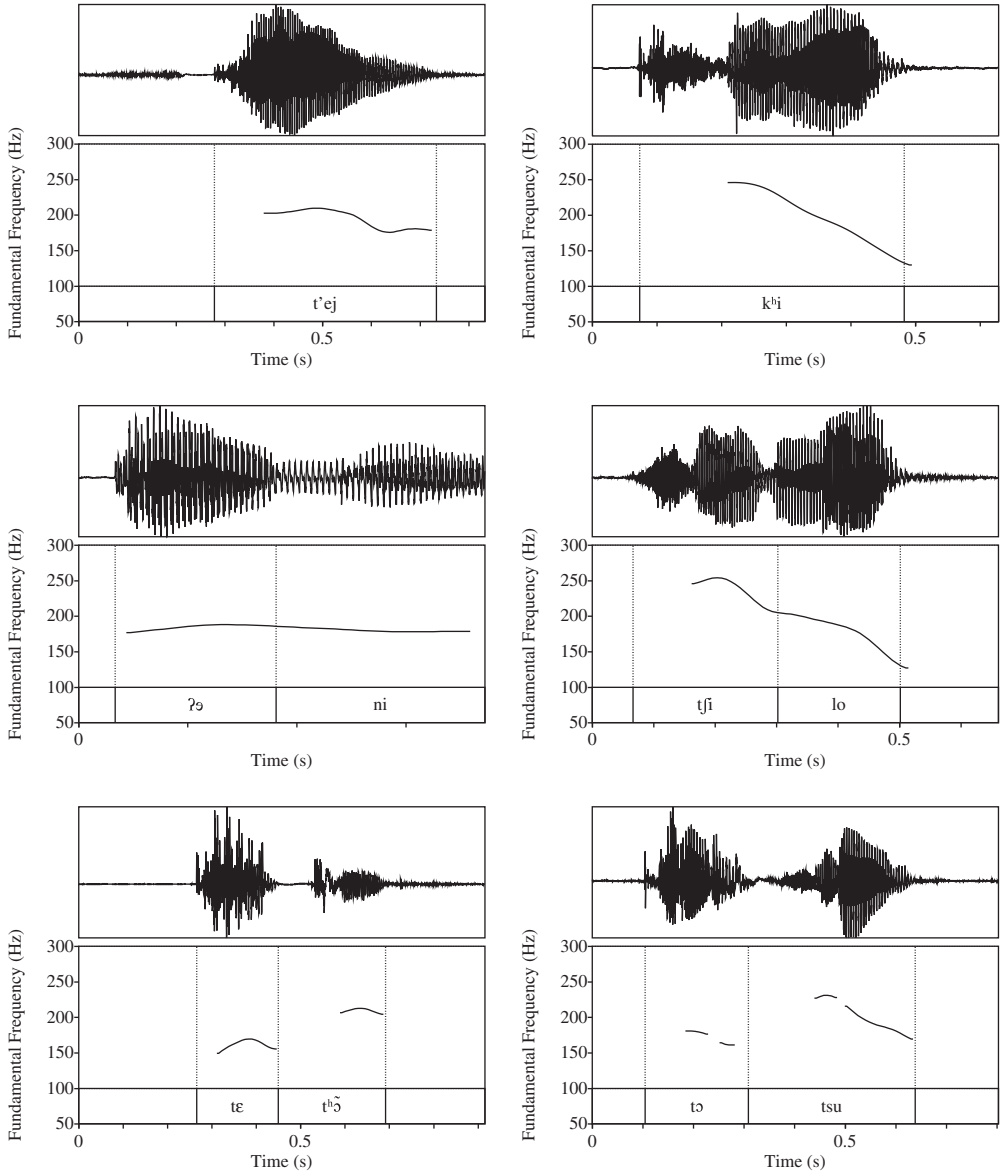


Figure A1 Examples of the two distinct pitch contours attested on one-syllable words and the four distinct pitch contours attested on two-syllable words. Upper left panel: flat pitch contour on /t'ej/ 'hot corn drink, *atole*'; upper right panel: falling pitch contour on /kʰi/ 'blood'; middle left panel: flat pitch contour on /ʔɔni/ 'chicken'; middle right panel: falling pitch contour on /tʃilo/ 'dog'; lower left panel: rising pitch contour on /tɛtʰɔ/ 'corn'; lower right panel: rising-falling pitch contour on /tɔtsu/ 'woman'.

Two-syllable words

Four pitch contours are observed on two-syllable words, also exemplified in Figure A1:

- Flat, /ʔɔni/ 'chicken'
- Falling, /tʃilo/ 'dog'

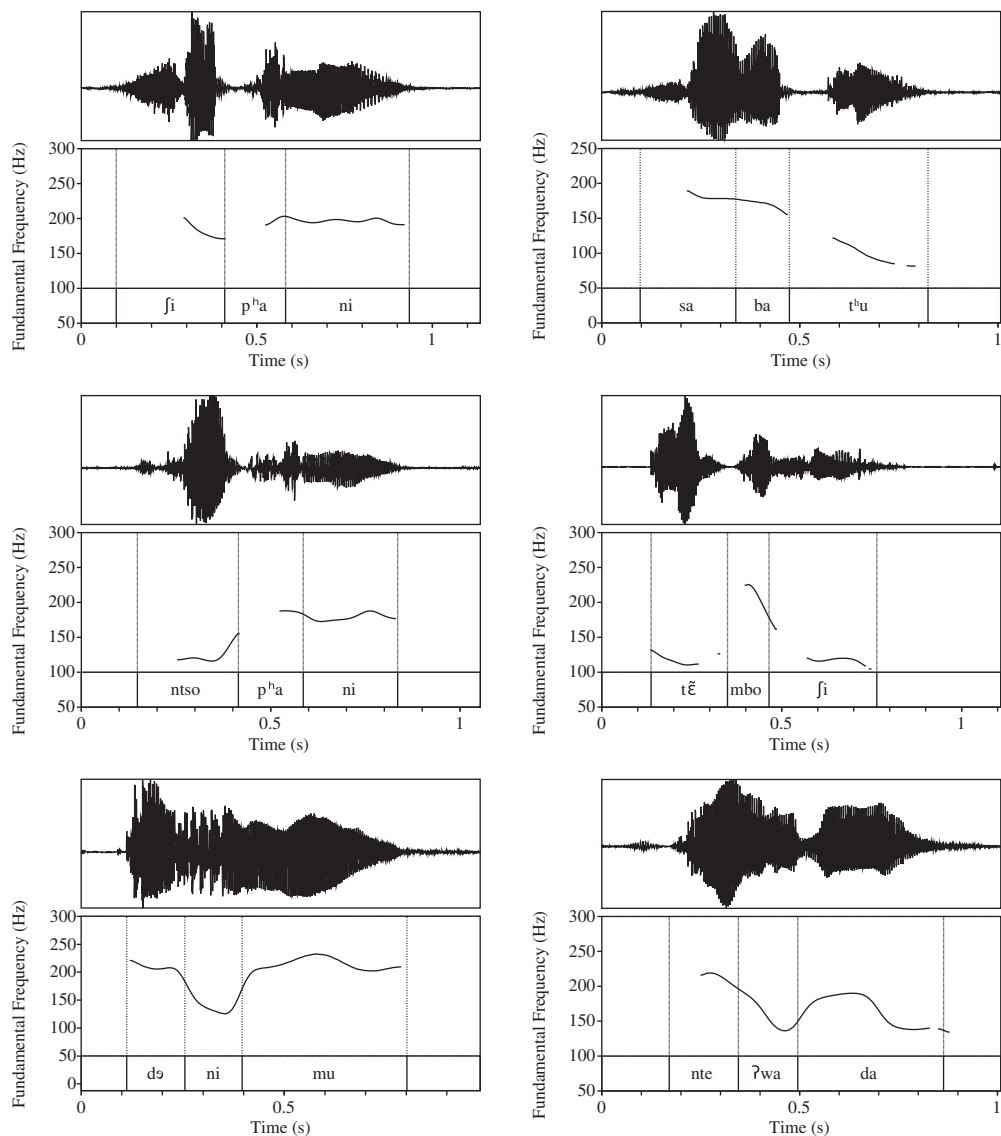


Figure A2 Examples of the six distinct pitch contours attested on three-syllable words. Top left panel: flat pitch contour on /ʃipʰani/ ‘wineskin, *botá*’; top right panel: falling pitch contour on word /sábatʰu/ ‘Sunday’; middle left panel: rising pitch contour on /ntsopʰáni/ ‘La Asunción Tepeyoyuca’ (a local village); middle right panel: rising–falling pitch contour on /tēmbôʃi/ ‘(green) tomato’; lower left panel: falling–rising pitch contour on /dɛnimú/ ‘squash blossom, *flor de calabaza*’; lower right panel: falling–rising–falling pitch contour on /nteʔwadâ/ ‘tool for splitting maguey plants, *partidór*’.

- Rising, /tɛtʰɔ́/ ‘corn’
- Rising falling, /tɔ́tsû/ ‘woman’

A note regarding the word /tʃîlo/ ‘dog’: The word /ʔjó/ was glossed in [Section 2.2.4](#) as ‘dog’, however /tʃîlo/ is a somewhat more common form in SJAQ. The word itself may be

originally derived from what Jakobson (1971: 22) called a ‘nursery form’ – a conventionalized lexical coinage from child speech. In this case it is likely the diminutive prefix /tʃi/ with a child’s production of /ʔjó/ as /lo/. The change in tone attests to its lexicalization as a single morpheme, and most speakers appear to be unaware that /tʃilo/ is at all related to /ʔjó/. The word is not accompanied by any connotations of childishness.

Three-syllable words

Six distinct pitch contours are observed on three-syllable words, exemplified in Figure A2:

- Flat, /ʃip^hani/ ‘wineskin, *bota*’
- Falling, /sâbat^hu/ ‘Sunday’
- Rising, /ntsop^háni/ ‘La Asunción Tepexoyuca’ (a local village)
- Rising falling, /tẽmbôʃi/ ‘(green) tomato’
- Falling rising, /dønimú/ ‘squash blossom, *flor de calabaza*’
- Falling rising falling /nteʔwadâ/ ‘tool for splitting maguey plants, *partidor*’

References

- Andrews, Henrietta. 1949. Phonemes and morphophonemes of Temoayan Otomi. *International Journal of American Linguistics* 15(4), 213–222.
- Andrews, Henrietta. 1993. *The function of verb prefixes in Southwestern Otomí*. Dallas, TX: SIL International.
- Arellanes, Francisco, Leonardo Carranza, Mario E Chávez-Peón, Verónica Fidencio, Alonso Guerrero, Michael Knapp Ring & Antonio Romero. 2011. Hacia una tipología tonal de las lenguas otopames. *Memorias Del V Congreso de Idiomas Indígenas de Latinoamérica*. Austin, TX: Archive of the Indigenous Languages of Latin America.
- Bartholomew, Doris. 1965. *The reconstruction of Otopamean (Mexico)*. Ph.D. dissertation, University of Chicago.
- Bartholomew, Doris. 1979. Review of *Otomí parables, folktales, and jokes*, edited by H. Russell Bernard & Jesús Salinas Pedraza. *The International Journal of American Linguistics* 45(1), 94–97.
- Bartholomew, Doris. 2000. Intercambio lingüístico entre otomí y náhuatl. *Estudios de Cultura Otopame* 2, 189–201.
- Beckman, Mary E. 1986. *Stress and non-stress accent*. Dordrecht: Foris.
- Bent, Tessa. 2004. Mandarin and English listeners’ perception of Mandarin tones in isolation and in context. *The Journal of the Acoustical Society of America* 116(4), 2573.
- Bernard, H. Russell. 1966. Otomi tones. *Anthropological Linguistics* 8(9), 15–19.
- Bernard, H. Russell. 1973. Otomí phonology and orthography. *International Journal of American Linguistics* 39(3), 180–184.
- Bernard, H. Russell. 1974. Otomi tones in discourse. *International Journal of American Linguistics* 40(2), 141–150.
- Bernard, H Russell. 1980. Orthography for whom? *International Journal of American Linguistics* 46(2), 133–136.
- Bernard, H. Russell & Jesús Salinas Pedraza 1976. *Otomí parables, folktales, and jokes* (IJAL Native American Texts Series 2). Chicago, IL: University of Chicago Press.
- Berthiaume, Scott Charles. 2004. *A phonological grammar of Northern Pame*. Ph.D. dissertation, The University of Texas at Arlington.
- Blicher, Deborah L., Randy L. Diehl & Leslie B. Cohen. 1990. Effects of syllable duration on the perception of the Mandarin tone 2/tone 3 distinction: Evidence of auditory enhancement. *Journal of Phonetics* 18(1), 37–49.
- Blight, Richard C. & Eunice V. Pike. 1976. The phonology of Tenango Otomi. *International Journal of American Linguistics* 42(1), 51–57.

- Boeg Thomsen, Ditte & Magnus Pharaos Hansen. 2015. Lenguaje del paisaje: Testimonios lingüísticos del otomí de acapulco. In Silvina Vigliani & Rovertto Junco (eds.), *Bajo el volcán: Vida y ritualidad en torno al nevado de Toluca*, 25–47. Mexico City: Instituto Nacional de Antropología e Historia.
- Boersma, Paul. 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proceedings of the Institute of Phonetic Sciences* 17, 97–110.
- Cole, Jennifer, José Ignacio Hualde & Khalil Iskarous. 1999. Effects of prosodic and segmental context on /g/-lenition in Spanish. In Osamu Fujimura, Brian D. Joseph & Bohumil Palek (eds.), *Proceedings of the Fourth International Linguistics and Phonetics Conference*, 575–589. Prague: The Karolinum Press.
- Crowhurst, Megan J. & Amador Teodocio Olivares. 2014. Beyond the iambic-trochaic law: The joint influence of duration and intensity on the perception of rhythmic speech. *Phonology* 31(1), 51–94.
- Cruz, Luis Hernández, Moisés Victoria Torquemada & Donald Sinclair Crawford (eds.). 2004. *Diccionario del hñähñu (otomí): Valle del Mesquital, Hidalgo* (Vocabularios Indígenas 45). Mexico City: Instituto Lingüístico de Verano.
- de Jong, Kenneth J. 1995. The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *The Journal of the Acoustical Society of America* 97(1), 491–504.
- Delattre, Pierre. 1966. A comparison of syllable length conditioning among languages. *IRAL – International Review of Applied Linguistics in Language Teaching* 4(1–4), 183–198.
- Dilley, Laura C. 2005. *The phonetics and phonology of tonal systems*. Ph.D. dissertation, MIT.
- Donohue, M. 2005. Tone and the Trans New Guinea languages. In Shigeki Kaji (ed.), *Cross-linguistic studies of tonal phenomena: Historical development and the tone/syntax interface*, 33–53. Tokyo: Tokyo University of Foreign Studies Research Institute for Language and Cultures of Asia and Africa.
- Dorian, Nancy C. 1994. Varieties of variation in a very small place: Social homogeneity, prestige norms, and linguistic variation. *Language* 70(4), 631–696.
- Engstrand, Olle. 1988. Articulatory correlates of stress and speaking rate in Swedish VCV utterances. *The Journal of the Acoustical Society of America* 83(5), 1863–1875.
- Fant, Gunnar, Anita Kruckenberg & Lennart Nord. 1991. Durational correlates of stress in Swedish, French, and English. *Journal of Phonetics* 19(3–4), 351–365.
- Figueroa, Sandra. 2012. El pasado otomí de la sierra de las Cruces: Su representación en voces de San Jerónimo Acapulco, Estado de México. Master's thesis, Universidad Nacional Autónoma de México.
- Figueroa, Sandra. 2015. De luna a luna: Quehaceres de las tías para una boda otomí. *Estudios de Cultura Otopame* 8(1), 257–274.
- Fry, Dennis B. 1955. Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America* 27, 765–768.
- Gandour, Jackson T. 1978. The perception of tone. In Victoria A. Fromkin (ed.), *Tone: A linguistic survey*, 41–72. New York: Academic Press.
- Goldsmith, John A. 1976. *Autosegmental phonology*. Ph.D. dissertation, MIT.
- Gómez Rendón, Jorge. 2008. *Typological and social constraints on language contact: Amerindian languages in contact with Spanish*. Utrecht: Landelijke Onderzoekschool Taalwetenschap.
- Gordon, Matthew. 2011. Stress: Phonotactic and phonetic evidence. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. II: *Suprasegmental and prosodic phonology*, 924–948. Oxford: Blackwell.
- Gordon, Matthew & Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29, 383–406.
- Hanson, Helen M. 1997. Glottal characteristics of female speakers: Acoustic correlates. *The Journal of the Acoustical Society of America* 101(1), 466–481.
- Harvey, Herbert R. 1972. The *Relaciones Geográficas*, 1579–1586: Native languages. In Howard F. Cline (ed.), *Handbook of Middle American Indians*, vol. 12, 279–323. Austin, TX: University of Texas Press.
- Hawks, John W. & James D. Miller. 1995. A formant bandwidth estimation procedure for vowel synthesis. *The Journal of the Acoustical Society of America* 97(2), 1343–1345.
- Hekking, Ewald. 1995. *El otomí de Santiago Mexquititlán: Desplazamiento lingüístico, préstamos y cambios gramaticales*. Amsterdam: Institute for Functional Research into Language; Language Use.

- Hekking, Ewald & Severiano Andrés de Jesús. 1984. *Gramática otomí*. Querétaro: Universidad Autónoma de Querétaro.
- Hekking, Ewald & Dik Bakker. 2009. Loanwords in Otomi, an Otomanguean language of Mexico. In Martin Haspelmath & Uri Tadmor (eds), *Loanwords in the world's languages: A comparative Handbook*, 897–915. Berlin: Mouton de Gruyter.
- Heldner, Mattias. 2003. On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish. *Journal of Phonetics* 31, 39–62.
- Henrich, Nathalie, Christophe d'Allesandro & Boris Doval. 2001. Spectral correlates of voice open quotient and glottal flow asymmetry: Theory, limits, and experimental data. In Paul Dalsgaard, Borge Lindberg & Henrik Benner (eds.), *Proceedings of Eurospeech*, 47–50. Aalborg: Aalborg Universitetsforlag.
- Hernández-Green, Néstor. 2011. Surgimiento de formas no finitas (FNF) en otomí de Acazulco. *Memorias Del V Congreso de Idiomas Indígenas de Latinoamérica*. Austin, TX: Archive of the Indigenous Languages of Latin America.
- Hernández-Green, Néstor. 2012. Tono y flexión verbal en otomí de Acazulco. Handout from talk given at XII Encuentro Internacional de Lingüística en el Noroeste.
- Hernández-Green, Néstor. 2015. *Morfosintaxis verbal del otomí de Acazulco*. Ph.D. dissertation, Centro de Investigaciones y Estudios Superiores en Antropología Social.
- Hernández-Green, Néstor, Enrique L. Palancar & Selene Hernández-Gómez. 2011. The Spanish loanword *lado* in Otomi spatial descriptions. *Language Sciences* 33(6), 961–980.
- Hess, Harold Harwood. 1962. *The syntactic structures of Mezquital Otomi*. Ph.D. dissertation, University of Michigan.
- Homma, Yayoi. 1981. Durational relationships between Japanese stops and vowels. *Journal of Phonetics* 9(3), 273–281.
- House, Arthur S & Kenneth N Stevens. 1958. Estimation of formant band widths from measurements of transient response of the vocal tract. *Journal of Speech and Hearing Research* 1(4), 309–315.
- Hualde, José Ignacio. 2006. Remarks on word-prosodic typology. *Proceedings of the Berkeley Linguistics Society* 32 (BLS 32), 157–174.
- Hualde, José Ignacio, Gorka Elordieta, Iñaki Gaminde & Rajka Smiljanić. 2002. From pitch accent to stress accent in Basque. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology VII*, 547–584. Berlin: Mouton de Gruyter.
- Huffman, Marie K. 1987. Measures of phonation type in Hmong. *The Journal of the Acoustical Society of America* 81(2), 495–504.
- Hyman, Larry M. 2006. Word-prosodic typology. *Phonology* 23, 225–257.
- Hyman, Larry M. 2009. How (not) to do phonological typology: The case of pitch-accent. *Language Sciences* 31, 213–238.
- Iseli, Markus, Yen-Liang Shue & Abeer Alwan. 2007. Age, sex, and vowel dependencies of acoustic measures related to the voice source. *The Journal of the Acoustical Society of America* 121(4), 2283–2295.
- Jakobson, Roman. 1971. *Studies on child language and aphasia*. Berlin: Walter de Gruyter.
- Jamieson, Allan R. 1977. Chiquihuitlan Mazatec tone. In William R Merrifield (ed.), *Studies in Otomanguean phonology*, 108–136. Dallas, TX: Summer Institute of Linguistics & The University of Texas at Arlington.
- Jiménez Moreno, Wigberto. 1936. *Mapa lingüístico del norte y centroamérica*. Mexico City: Instituto Panamericano de Geografía e Historia.
- Kager, René. 1993. Alternatives to the iambic-trochaic law. *Natural Language & Linguistic Theory* 11, 381–432.
- Kaji, Shigeki (ed.). 1999. *Cross-linguistic studies of tonal phenomena: Tonogenesis, typology, and related topics*. Tokyo: ILCAA.
- Keating, Patricia A. & Christina Esposito. 2007. Linguistic voice quality. *UCLA Working Papers in Phonetics* 105, 85–91.
- Klatt, Dennis H. & Laura C. Klatt. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America* 87(2), 820–857.

- Kreiman, Jody, Bruce R Gerratt & Norma Antoñanzas-Barroso. 2007. Measures of the glottal source spectrum. *Journal of Speech, Language, and Hearing Research* 50, 595–610.
- Kreiman, Jody, Yen-Liang Shue, Gang Chen, Markus Iseli, Bruce R. Gerratt, Juergen Neubauer & Abeer Alwan. 2012. Variability in the relationships among voice quality, harmonic amplitudes, open quotient, and glottal area waveform shape in sustained phonation. *The Journal of the Acoustical Society of America* 132(4), 2625–2632.
- Kubozono, Haruo. 2011. Japanese pitch accent. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. V: *Phonology across languages*, 2879–2907. Oxford: Blackwell.
- Ladd, D. Robert. 1984. Declination: A review and some hypotheses. *Phonology Yearbook* 1, 53–74.
- Ladd, D. Robert. 2014. *Simultaneous structure in phonology*. Oxford: Oxford University Press.
- Ladefoged, Peter, Ian Maddieson & Michel Jackson. 1988. Investigating phonation types in different languages. In Osamu Fujimura (ed.), *Vocal fold physiology: Voice production, mechanisms and functions*, 297–317. New York: Raven Press.
- Lastra, Yolanda. 1989. *Otomí de San Andrés Cuexcontitlán, Estado de México*. Mexico City: Archivo de Lenguas Indígenas de México.
- Lastra, Yolanda. 1996. Verbal morphology of Ixtenco Otomi. *Amerindia* 21, 1–8.
- Lastra, Yolanda. 1997. *El otomí de Ixtenco*. Mexico City: Instituto de Investigaciones Antropológicas, UNAM.
- Lastra, Yolanda. 1998. Otomí loans and creations. In Jane H. Hill, P. J. Mistry & Lyle Campbell (eds.), *The life of language: Papers in linguistics in honor of William Bright*, 59–66. Berlin: Mouton de Gruyter.
- Lastra, Yolanda. 2001. *Unidad y diversidad de la lengua. Relatos otomíes*. Mexico City: Instituto de Investigaciones Antropológicas, UNAM.
- Lastra, Yolanda. 2006. *Los otomíes – su lengua y su historia*. Mexico City: Instituto de Investigaciones Antropológicas, UNAM.
- Lastra, Yolanda. 2010. Diversidad lingüística: Variación dialectal actual. In Rebecca Barriga & Pedro Martín Butragueño (eds.), *Historia sociolingüística de México*, vol. 2, 841–880. Mexico City: El Colegio de México.
- Leben, William R. 1973. *Suprasegmental phonology*. Ph.D. dissertation, MIT.
- Leben, William R. 2011. Autosegments. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. I: *General issues and segmental phonology*, 311–339. Oxford: Blackwell.
- Leon, Frances. 1963. Revisión de la fonología del otomí. *Anales Del Instituto Nacional de Antropología e Historia* 15, 315–330.
- Leon, Frances & Morris Swadesh. 1949. Two views of Otomi prosody. *International Journal of American Linguistics* 15(2), 100–105.
- McCarthy, John. 2004. Headed spans and autosegmental spreading. Ms., University of Massachusetts Amherst.
- Mock, Carol C. 1983. Tone sandhi in Isthmus Zapotec: An autosegmental account. *Linguistic Analysis* 12(1), 91–139.
- Nakai, Satsuki, Sari Kunnari, Alice Turk, Kari Suomi & Riika Ylitalo. 2009. Utterance-final lengthening and quantity in northern Finnish. *Journal of Phonetics* 37, 29–45.
- Odden, David. 1999. Typological issues in tone and stress in Bantu. In Kaji (ed.), 187–215.
- Palancar, Enrique L. 2004a. Middle voice in Otomi. *International Journal of American Linguistics* 70(1), 52–85.
- Palancar, Enrique L. 2004b. Verbal morphology and prosody in Otomi. *International Journal of American Linguistics* 70(3), 251–278.
- Palancar, Enrique L. 2006. Property concepts in Otomí: A language with no adjectives. *International Journal of American Linguistics* 72(3), 325–366.
- Palancar, Enrique L. 2009. *Gramática y textos del hñõñhõ: Otomí de San Ildefonso Tultepec, Querétaro*, vol. 1: *Gramática*. Mexico City: Plaza y Valdés.
- Palancar, Enrique L. 2011. The conjugations of colonial Otomi. *Transactions of the Philological Society* 109(3), 246–264.

- Palancar, Enrique L. 2012. The conjugation classes of Tilapa Otomi: An approach from canonical typology. *Linguistics* 50(4), 783–832.
- Palancar, Enrique L. 2013a. The evolution of number in Otomi: The many faces of the dual. *Studies in Language* 37 (1), 94–143.
- Palancar, Enrique L. 2013b. Nominalización y complementación en otomí. In Enrique L. Palancar & Roberto Zavala Maldonado (eds.), *Clases léxicas, posesión y cláusulas complejas en lenguas de Mesoamérica*, 381–412. Mexico City: CIESAS.
- Pharao Hansen, Magnus. 2012. Kinship in the past tense: Language, care and cultural memory in a Mexican community. Master's thesis, Brown University.
- Pharao Hansen, Magnus, Néstor Hernández-Green, Rory Turnbull & Ditte Boeg Thomsen, 2016. Life histories, language attitudes and linguistic variation: navigating the micropolitics of language revitalization in an Otomí community in Mexico. In Gabriela Pérez Báez, Chris Rogers & Jorge Emilio Rosés Labrada (eds.), *Language documentation and revitalization in Latin American contexts*, 215–246. Berlin: Mouton de Gruyter.
- Pharao Hansen, Magnus, Rory Turnbull & Ditte Boeg Thomsen. 2011. From academic salvage linguistics to community-based documentation in only three weeks: Report from a collective and interdisciplinary fieldwork on Acazulco Otomí. Paper presented at the 2nd International Conference on Language Documentation and Conservation.
- Plag, Ingo, Gero Kunter & Mareile Schramm. 2011. Acoustic correlates of primary and secondary stress in north American English. *Journal of Phonetics* 39, 362–374.
- Pulleyblank, Doug. 1986. *Tone in lexical phonology*. Dordrecht: Reidel.
- Shaeffler, Felix. 2005. *Phonological quantity in Swedish dialects*. Reports in Phonetics Umeå University (PHONUM), vol. 10.
- Silverman, Daniel. 1997. Tone sandhi in Comaltepec Chinantec. *Language* 73(3), 473–492.
- Simpson, Adrian P. 2012. The first and second harmonics should not be used to measure breathiness in male and female voices. *Journal of Phonetics* 40(3), 477–490.
- Sinclair, Donald E. & Kenneth L. Pike. 1948. The tonemes of Mesquital Otomi. *International Journal of American Linguistics* 14(2), 91–98.
- Sluijter, Agaath M. C. & Vincent J. van Heuven. 1996. Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America* 100 (4), 2471–2485.
- Suárez, Jorge A. 1983. *The Mesoamerican Indian languages*. Cambridge: Cambridge University Press.
- Turk, Alice E. & Stefanie Shattuck-Hufnagel. 2007. Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics* 35, 445–472.
- Uwano, Zendo. 1999. Classification of Japanese accent systems. In Kaji (ed.), 151–186.
- Valle Canales, Berna Leticia. 2008. Obsolescencia dialectal y evolución de la lengua: la variación fonológica del otomí de San Pedro Atlapulco. Master's thesis, Escuela Nacional de Antropología e Historia.
- van der Hulst, Harry. 2011. Pitch accent systems. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. II: *Suprasegmental and prosodic phonology*, 1003–1025. Oxford: Blackwell.
- van der Hulst, Harry & Norval Smith. 1988. Introduction. In Harry van der Hulst & Norval Smith (eds.), *Autosegmental studies on pitch accent systems*, ix–xxiv. Dordrecht: Foris.
- Voigtlander, Katherine & Artemisa Echevoyen. 1985. *Luces contemporaneas del otomí: gramática del otomí de la sierra*. Mexico City: Instituto Lingüístico de Verano.
- Volhardt, Marc Daniel Skibsted. 2014. Determination of a phoneme set for Acazulco Otomi: Linguistic fieldwork in Ndöngü, San Jerónimo Acazulco. Master's thesis, University of Iceland.
- Wallis, Ethel E. 1968. The word and the phonological hierarchy of Mezquital Otomi. *Language* 44(1), 76–90.
- Wayland, Ratee & Allard Jongman. 2003. Acoustic correlates of breathy and clear vowels: The case of Khmer. *Journal of Phonetics* 31, 181–201.
- Whalen, Doug H. & Yi Xu. 1992. Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica* 49(1), 25–47.

- Wright Carr, David Charles. 2005. Precisiones sobre el término 'otomí'. *Arqueología Mexicana* 13(73), 19.
- Wright Carr, David Charles. 2008. La sociedad prehispánica en las lenguas náhuatl y otomí. *Acta Universitaria* 18(1), 15–23.
- Wright Carr, David Charles. 2010. Préstamos lingüísticos entre el otomí y el náhuatl. In Ana María Salazar & Verónica Kugel (eds.), *Homenaje a Yolanda Lastra. X coloquio internacional sobre otopames*, 211–220. Mexico City: Instituto Nacional de Lenguas Indígenas.
- Wright, Richard. 2004. Factors of lexical competition in vowel articulation. In John Local, Richard Ogden & Rosalind Temple (eds.), *Papers in Laboratory Phonology IV: Phonetic interpretation*, 75–85. Cambridge: Cambridge University Press.
- Yip, Moira. 1989. Contour tones. *Phonology* 6(1), 149–174.